CrowdVisor: an Embedded Toolset for Human Activity Monitoring in Critical Environments

Artificial Intelligence for Media and Humanities (AIMH) Laboratory
Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo" (ISTI-CNR) of Pisa
aimh.isti.cnr.it


Giuseppe Amato


Fabrizio Falchi


Claudio Gennaro


Fabio Carrara


Luca Ciampi


Marco Di Benedetto

I-CITIES 2022

Ascoli Piceno --- 14-16 Settembre 2022

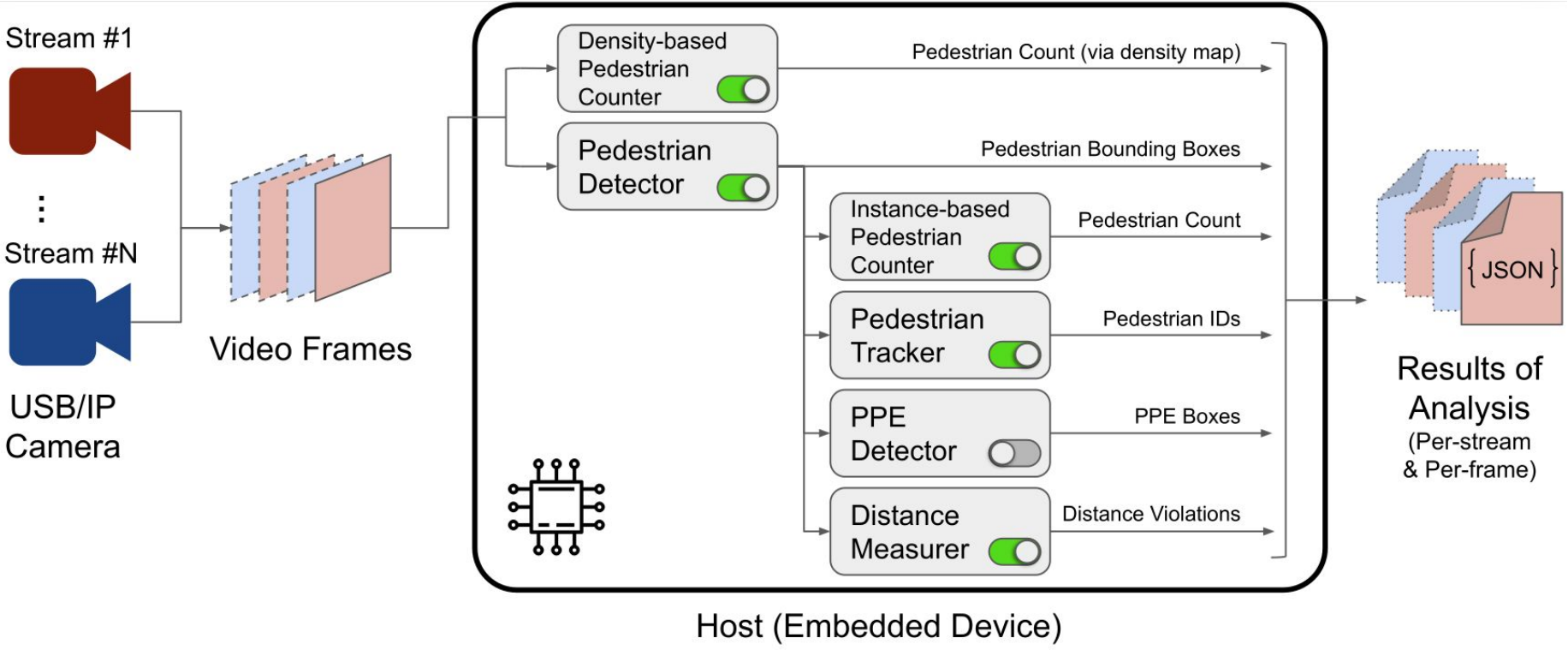
CrowdVisor: An Embedded Toolset for Human Activity Monitoring in Critical Environments

- **Computer Vision at the service of Smart Cities**
  - AI-based applications that **automatically analyse images from visual sources** (city cameras, surveillance cameras, smartphones, …)
- Scenarios where, through image analysis, we want to **monitor compliance with the rules for individual and collective safety**
  - dangerous workplaces, pandemic events like the recent COVID-19 disease
- **Human supervision could not always guarantee this task**
  - especially in crowded scenarios
- **CrowdVisor** → valuable tool to automatically monitor human activities in critical environments through images captured by networked cameras

What we propose

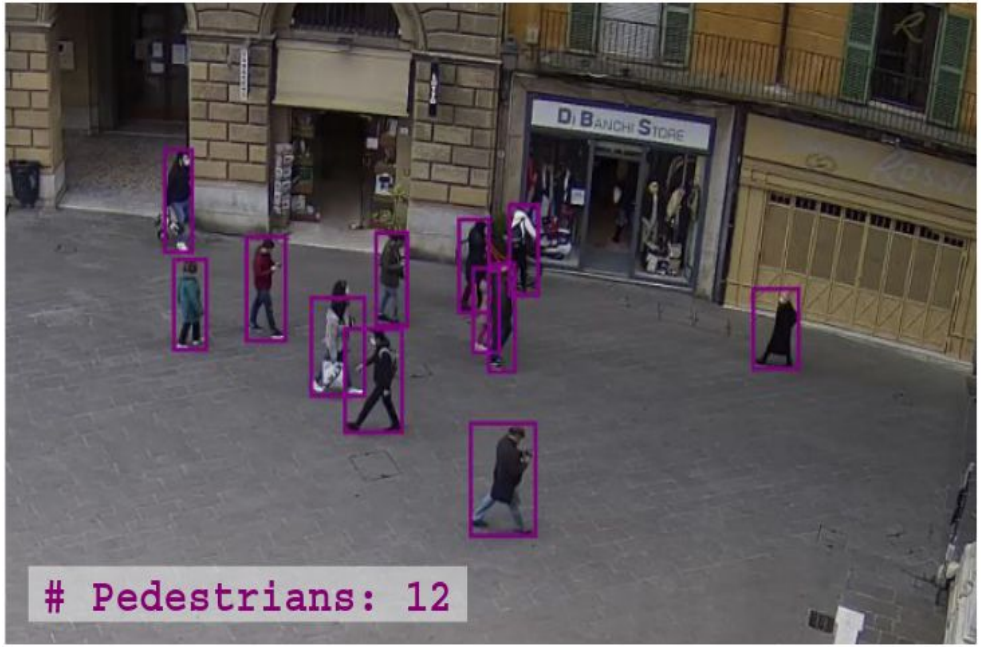
- Embedded modular Computer Vision-based and AI-assisted system that carry out several tasks to help monitor individual and collective human safety rule
- **Real-time but low-cost system**
  - embedded devices, **images** are **captured and processed directly onboard**
  - compute- and storage-limited resources
- **Multiple modules relying on neural network components**
  - each responsible for specific functionalities that the user can easily enable and configure
  - they range from estimation of social distance, estimation of number of people present in a ROI, detection of Personal Protective Equipments, …

Overview of the Modular Architecture



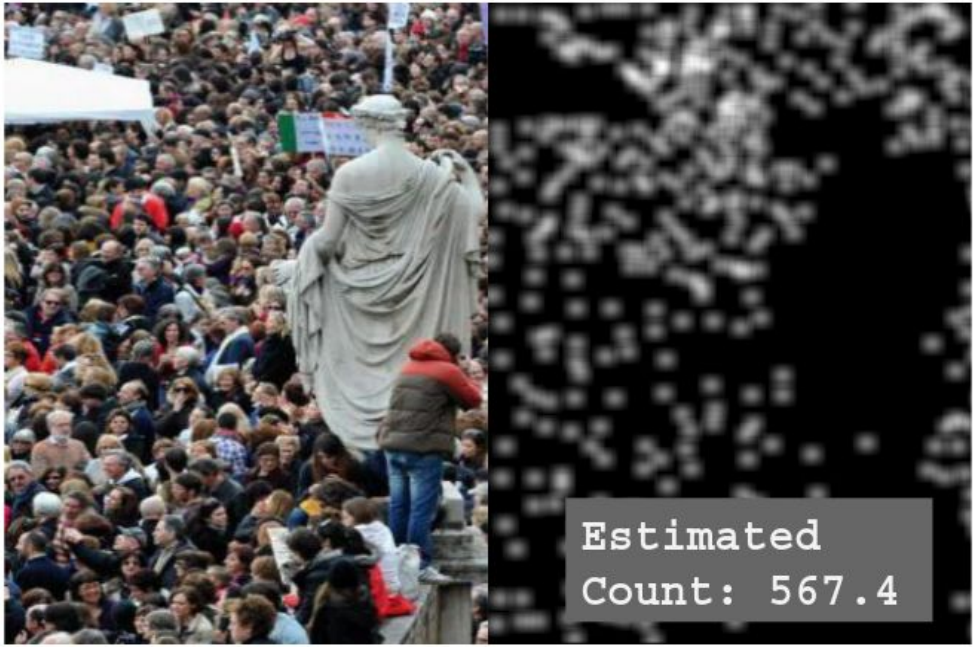
People Detection + Counting by Instances Modules

- **People Detector** → main component on which almost all other plugins rely
- Purpose → localize and classify pedestrian instances from input images
- It is based on *Faster R-CNN*, a popular SOTA CNN-based object detection system
- We specialize it to localize pedestrian instances
- **Counting by Instances Module** → rely on the pedestrian detector, simply count the localized instances



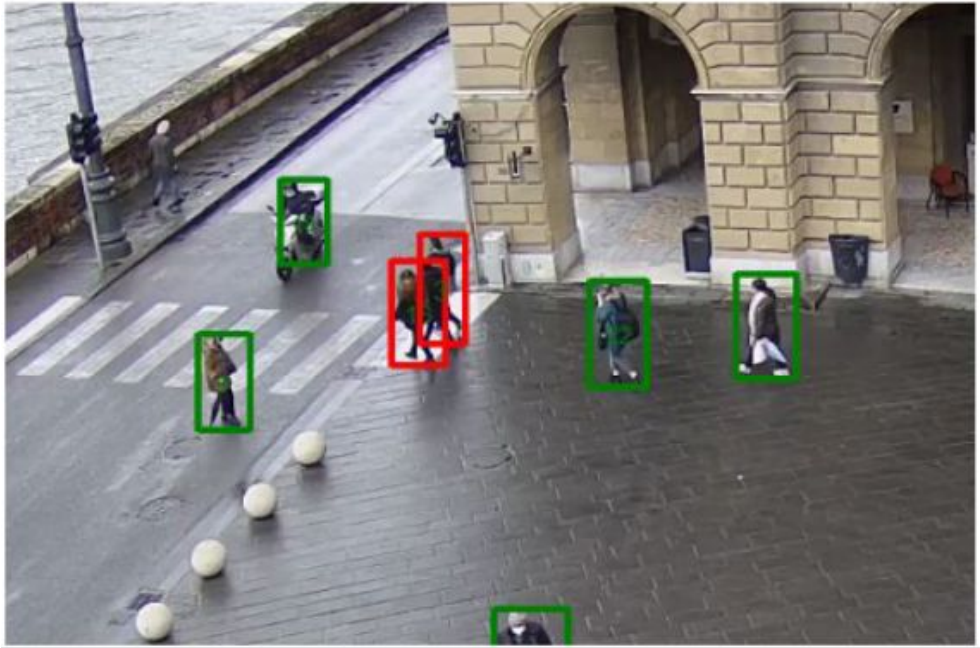
High-Density People Counting Module

- **Estimating the number of people present in a region of interest is crucial to monitor the occupancy area**
- By measuring and limiting the number of people who can visit a location at any one time, it is possible to drastically **reduce the likelihood of setting up people gatherings**
- Dedicated plug-in that can work in **two different modalities** that the user can conveniently pick out, depending on the considered scenario
  - Instance-based counting
  - density-based counting



People Distance Measurer Module

- Although crowd counting is effective in monitoring aggregations, **measuring distances among people becomes critical during pandemic event**
- This module is based on **homography**, i.e., a perspective transformation that projects on two different points of view a set of 3D points lying on the same plane
- Pre-calibrating step to the fixed monitoring camera, using a proper geometrical transformation that places detected items on a common system of reference



People Tracking Module

- **Object tracking** can be an essential tool to increase the robustness to spurious detections and achieve temporal consistency in video analysis
- We apply an object tracker over pedestrian detection to **re-identify people among consecutive video frames**
- Beneficial for assessing temporal rules, such as raising alarms after the same pedestrian occupies a forbidden area for more than a predefined amount of time
- It is based on **DeepSort** tracking algorithm for 2D objects that consider position and area of the bounding box and their speed of variation as well feature vectors that represent their appearance

Personal Protective Equipment Module

- A simple intervention for protecting health and well-being is wearing **Personal Protective Equipment (PPE)**
  - wearing harnesses and helmets on construction sites
  - wearing face masks to prevent infections during pandemic
- Purpose → localize and classify PPE instances from people instances
- It is based on *Faster R-CNN*, a popular SOTA CNN-based object detection system

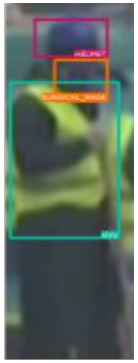
The CrowdVisorPisa Dataset

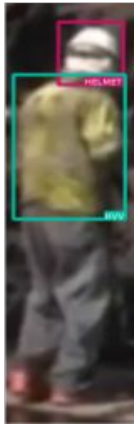





The CrowdVisorPPE Dataset

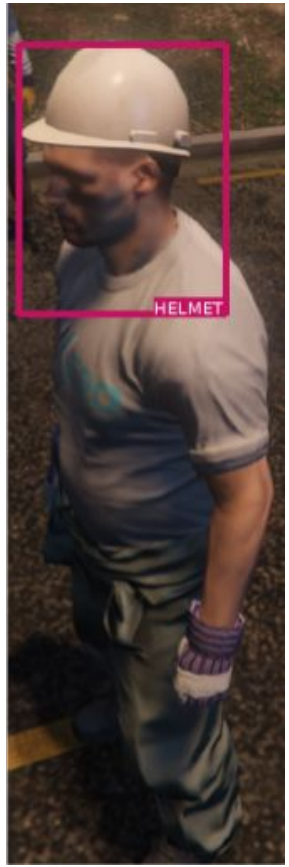




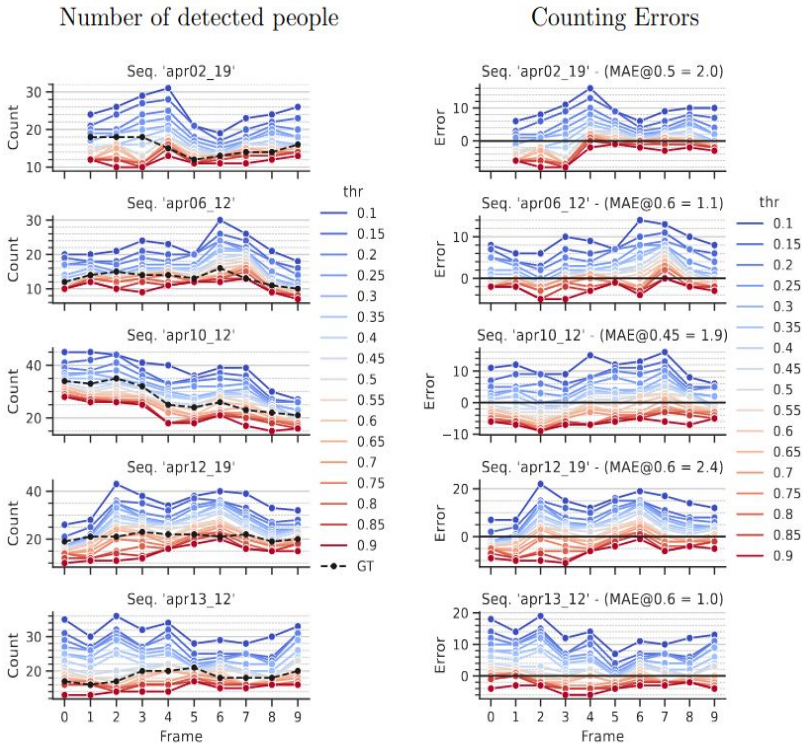
		# PPE instances		
Train Split	# img	Helmet	HVV	Mask
GTA V (**V**)	28,078	9,575	21,374	0
Web (**R**)	21,820	10,673	10,686	1,630
Test Split				
Web (**R**)	4,119	2,163	2,017	271



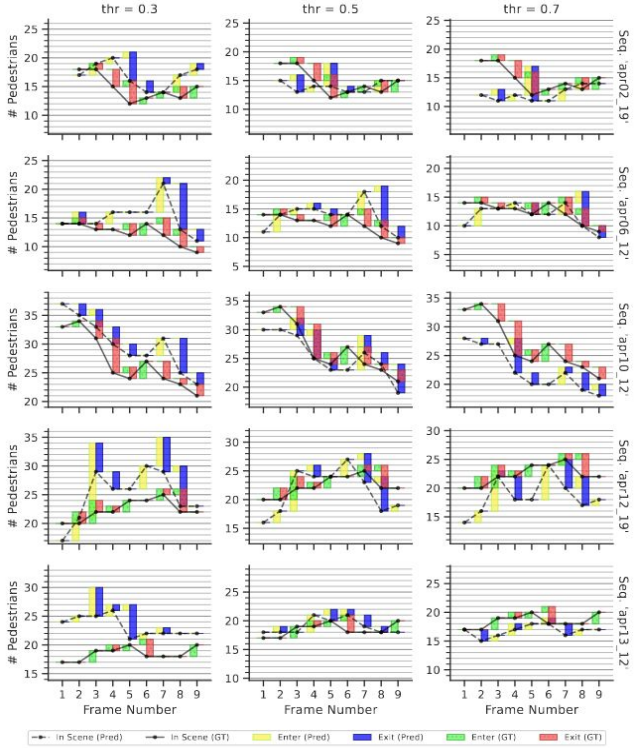


(Some) Results

Evaluation of *Counting by Instances* functionality considering the single still frames of the *CrowdVisorPisa* dataset.



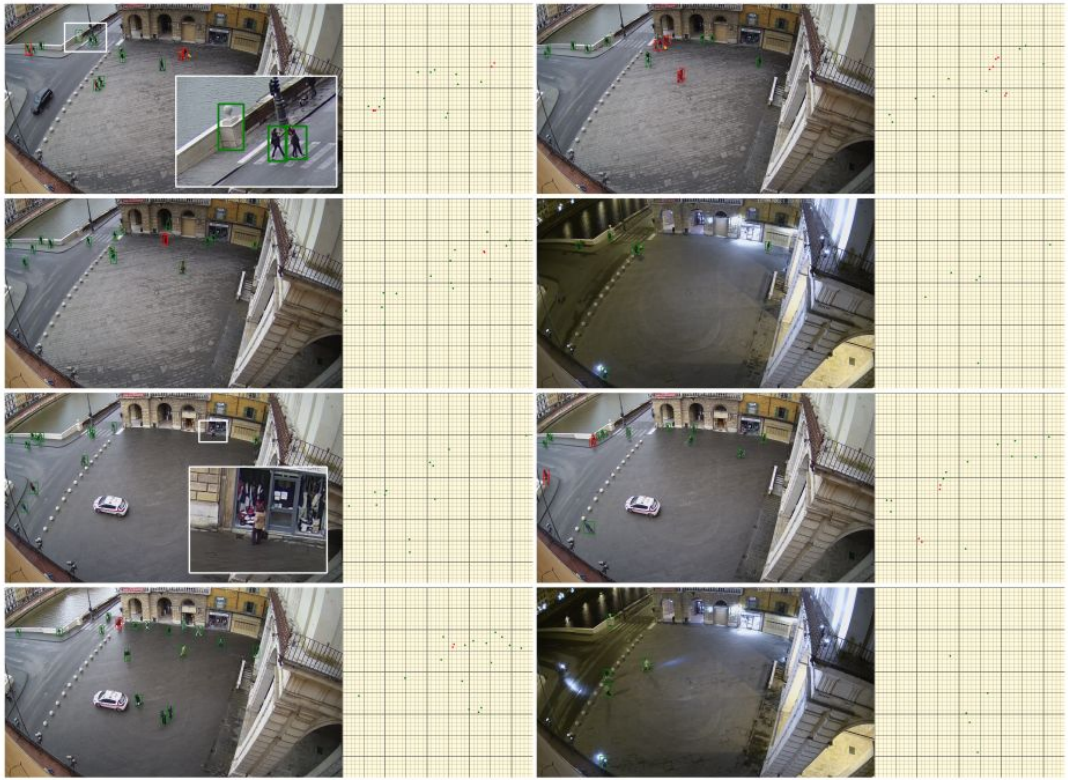
Evaluation of *Counting by Instances* functionality taking into account also the tracker module.



System and GPU Memory Usage

PD	DC	PPE	SysRAM	GpuRAM
Light	✗	✗	2.36	0.55
	✓	✗	2.44	0.86
	✗	✓	2.35	2.10
	✓	✓	2.44	2.20
Full	✗	✗	2.51	0.62
	✓	✗	2.52	0.94
	✗	✓	2.51	2.20
	✓	✓	2.51	2.30

(Some) Results



Conclusion

- We presented a modular framework based on Computer Vision and AI technologies aimed at **monitoring human activities in critical conditions**
- The system has the peculiarity to be **expandable** in the future, simply adding new modules in charge of performing new functionalities that the user can easily enable (or disable) according to their needs
- We implemented a set of visual-based modules for *pedestrian detection*, *tracking*, aggregation *counting based on instances and density maps*, *social distancing calculations*, and *personal protection environment detection*
- To test the effectiveness of our solution, we monitored a known place in Italy proving satisfactory accuracy in terms of detection, counting, and physical distance measurements

CrowdVisor: an Embedded Toolset for Human Activity Monitoring in Critical Environments

Artificial Intelligence for Media and Humanities (AIMH) Laboratory
Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo" (ISTI-CNR) of Pisa
aimh.isti.cnr.it

Thanks! Questions?

luca.ciampi@isti.cnr.it (Speaker)

giuseppe.amato@isti.cnr.it (Lab Head)

Visual Engines

AI4media
ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

Extension