Contest title and abstract

PAR 2023 – Pedestrian Attribute Recognition Contest

The Pedestrian Attribute Recognition (PAR) Contest is a competition among methods for pedestrian attributes recognition from images. To allow the participants to train effective models, we provide a dataset, the Mivia PAR Dataset, including images annotated with some of the following labels: color of the clothes (top and bottom), gender (female, male), bag (presence or absence), hat (presence or absence). After the contest, the dataset, also augmented with additional samples and annotations produced by the participants, will be made publicly available for the scientific community and will hopefully become among the biggest dataset of pedestrian attributes with this set of annotations.

The performance of the competing methods will be evaluated in terms of accuracy on a private test set composed by images that are different from the ones available in the training set.

General description of the problem

Pedestrian attributes recognition from images is nowadays a relevant problem in several real applications, such as digital signage, social robotics, business intelligence, people tracking and multi-camera person re-identification.

Multi-task learning

To this concern, there is a great interest for recognizing simultaneously several information regarding the pedestrian, i.e. the colour of the clothes, the gender, the presence or absence of bags or hats and so on. It is worth pointing out that using a single classifier for recognizing each of the abovementioned pedestrian attributes may require prohibitive computational resources (not always available) for obtaining them in real-time; in this scenario, nowadays multi-task learning approaches represent an excellent solution for achieving remarkable recognition accuracy while maintaining the processing time unchanged as the number of pedestrian attributes increases.

Dataset: missing labels and additional samples

For the contest, we propose the use of a novel training set, the MIVIA PAR Dataset, annotated with five pedestrian attributes and we restrict the competition to methods based on multi-task learning. Since not all the training samples are annotated with all the labels, the participants may also propose a learning procedure designed for dealing with missing labels.

Moreover, the participants are encouraged to use additional samples or to produce themselves the missing annotations; this possibility is allowed in the competition only under the constraint that the additional samples and annotations are made publicly available, to give a relevant contribution to the diffusion of public datasets for pedestrian attributes recognition.

Private test set

The effectiveness of the proposed methods will be evaluated by comparing the results on a private challenging test set, so as the design choices in terms of data augmentation, neural network architecture, loss function and learning procedure.

Description of the provided dataset

The MIVIA PAR Dataset consists of more than 80.000 images annotated with the following labels (or part of them):

- Color of the clothes: the considered values are black, blue, brown, gray, green, orange, pink, purple, red, white, and yellow and are represented, in this order, with the labels [1,2,3,4,5,6,7,8,9,10,11].

We provide the color of the upper part of the body and of the bottom part of the body as two different labels.

- Gender: the considered values are male and female, represented in this order with the values [0,1].
- Bag: we consider the absence or presence of a bag, representing it with the values [0,1].
- Hat: we consider the absence or presence of a hat, representing it with the values [0,1].

The unavailability of the specific annotation is indicated with the value -1.

The samples have been collected from existing datasets (e.g. PETA, RAP, Colorful), by manually annotating the missing attributes, and obtained from private images, by manually extracting the image crop of the person and manually annotating the considered pedestrian attributes. Since the images are collected in different conditions, the dataset is very heterogeneous in terms of image size, illumination, pose of the person, distance from the camera. Each image of the dataset contains a single person, already cropped.

We make available to the participants a folder with all the images and a CSV file with the labels of the training samples.

Additional samples and/or annotations

Since the goal of this contest is the development of the research on pedestrian attributes recognition, we encourage participants to use other samples or to add missing labels for training their models, if such additional samples and annotations are made publicly available. The diffusion of samples annotated with pedestrian attributes would make a great contribution to the development of this line of research and to the realization of real applications in this field.

Description of the specific contest tasks

The participants may download the training samples together with the corresponding annotations of the Mivia PAR dataset from the website of the contest.

The participants can use these training samples and annotations, but they can also use additional samples and/or add the missing labels, if they make the additional samples and annotations publicly available.

The participants must provide, for each sample of the test set, the prediction for all the considered pedestrian attributes, by training a single multi-task neural network. They are free to design novel neural network architectures or to define novel training procedure and loss functions for multi-task learning. Particularly welcome are the methods dealing with the missing labels.

Finally, the participants must produce a brief PDF report of the proposed method, by following a template that can be downloaded from the website of the contest.

Operative instructions

The participants must send before the deadline a Python script for executing it on the private test set.

The script must follow these specifications:

1. It must be developed by considering these constraints: Python >= 3.6; Tensorflow >= 2.0; PyTorch >= 1.4.0; OpenCV >= 4.
2. It must be called **test.py** and executed with the following command:

   python test.py –data images.csv

   where images.csv is a file containing an image path for each raw (e.g. 0000000.jpg).
3. It must analyse all the images in images.csv and produce a file of the results. Each raw of the file must include, separated by a comma (according to the CSV standard), the filename of the test sample, exactly as it has been read from images.csv (e.g. 000000.jpg) and the estimated values for each attribute, separated by a comma. Therefore, an example of raw may be 000000.jpg,3,2,1,1,1. The file

of the results is formatted like the file with the training annotations. For the sake of clarity, **the participants will also receive a sample script** which reads the paths of the images from images.csv, load them, predict the attributes, and produce a file with the results.

All the files necessary for running the experiment must be sent together with the script test.py.

Detailed instructions about the procedure will be provided on the website of the contest.

Evaluation metrics or tool to be used for evaluation

The proposed methods will be evaluated in terms of accuracy over all the considered tasks.

The accuracy A is defined as the ratio between the number of correct classifications (prediction p_i equal to the ground truth g_i) and the total number of samples K:

$$A = \frac{\sum_{i=1}^{K}(p_i == g_i)}{K}$$

The higher the accuracy achieved by a method, the higher its effectiveness on that specific task.

We compute the accuracy for the five pedestrian attributes:

- A_u: accuracy in the recognition of the color of the clothes in the upper part of the body
- A_l: accuracy in the recognition of the color of the clothes in the lower part of the body
- A_g: gender recognition accuracy
- A_b: bag presence recognition accuracy
- A_h: hat presence recognition accuracy

We define the ranking of the contest according to the average accuracy (AA), namely the mean of the abovementioned accuracies:

$$AA = \frac{A_u + A_l + A_g + A_b + A_h}{5}$$

The method which achieves the highest AA will be the winner of the Pedestrian Attribute Recognition Contest since it will demonstrate the highest average accuracy in the various tasks.

Plans for contest web site

The web site of the contest will be available at https://par2023.unisa.it . It will include sections which describe the problem, the dataset and the contest tasks, rules, and deadlines. In addition, the web site will provide information about how to download the training set and how to submit the results. After the deadline, it will also include the results of the contest.

Contact information for the organizer(s)

Antonio Greco, Ph. D., Tenure-Track Assistant Professor
Dept. of Information and Electrical Engineering and Applied Mathematics (DIEM)
Laboratory of Intelligent Machines for the Analysis of Videos, Images and Audio (MIVIA)
University of Salerno, Via Giovanni Paolo II, 132 84084 - Fisciano (SA) Italy
Phone:+39 089 963006
E-mail: agreco@unisa.it
Institutional web profile: http://docenti.unisa.it/antonio.greco
MIVIA web profile: http://mivia.unisa.it/people/greco/
ORCID: https://orcid.org/0000-0002-5495-2432
Scopus: https://www.scopus.com/authid/detail.uri?authorId=55808071886

Scholar: https://scholar.google.it/citations?hl=it&user=g0G9k9oAAAAJ

Bruno Vento, Ph. D Student
Dept. of Electrical Engineering and Information Technology (DIETI)
Pattern Analysis and Intelligent Computation for mUltimedia Systems LAB
University of Naples, Via Claudio, 21 - 80135 Naples - Italy
E-mail: bruno.vento@unina.it
Picus LAB web profile: https://picuslab.dieti.unina.it/index.php/people
Scholar: https://scholar.google.it/citations?user=ffvBB1EAAAAJ&hl=it&oi=ao

Brief CVs of the organizer(s)

Antonio Greco graduated cum laude in Computer Engineering in 2014 at the University of Salerno, Italy. In March 2018 he obtained the title of PhD in Computer Science and Computer Engineering at the same university. In March 2020 he became Assistant Professor (SSD ING-INF / 05 "Information Processing Systems") in the Department of Information and Electrical Engineering and Applied Mathematics (DIEM) at the University of Salerno, Italy, where he is now a Tenure-Track Assistant Professor since November 2022.

His research activities are mainly focused in the areas of Computer Vision and Pattern Recognition and, more specifically, on the design, the implementation and the optimization of algorithms for ambient understanding (gender recognition, age estimation, ethnicity recognition, emotion analysis, fire detection, anomaly detection, people counting, object tracking, audio event recognition) in real time on static (smart cameras, microphones) or moving devices (drones, robots, autonomous vehicles).

He organized in 2021, within the International Conference on Computer Analysis of Images and Patterns (CAIP), the Guess the Age (GTA) Contest.

Bruno Vento obtained his master's degree with honours in Computer Engineering in 2021 at the University of Salerno, with a specialization in Artificial Intelligence.

From 2019 to October 2022, he worked as a Machine Learning Engineer in the R&D division of A.I. Tech, an Italian company that designs and manufactures Artificial Intelligence enterprise solutions in the field of Smart Video Analysis. His activities have been mainly focused on the field of Computer Vision and on the design, implementation, and optimization of algorithms for fire and garbage detection, people counting, crowd estimation, road traffic analysis, parking monitoring, personal protective equipment recognition, license plate recognition, face analysis, intrusion detection, unattended item detection, flooding detection and pedestrian attribute recognition.

Since November 2022, he is a PhD Student in the Department of Electrical Engineering and Information Technology (DIETI) at the University of Naples. His PhD is in the field of "AI & Agriculture (Agrifood) and the Environment". His research project is on neural networks for fire, smoke and garbage detection and re-identification of the perpetrator through intelligent image and video analysis.