FONDAZIONE BRUNO KESSLER
MAGAZINE

An international recognition for semantic video understanding

July 6, 2022

The FBK TeV Unit made a significant contribution to the team that came in first in the EPIC-KITCHENS competition on egocentric vision.

Mobile cameras and smart glasses are increasingly popular devices both as research tools and as affordable products on the market. They capture the interactions with the reality of who wears them and we are beginning to understand how these technologies and their applications can have an impact on our lives.

These are in technical terms **“egocentric” detection** devices that may soon be able to assist and facilitate the wearer, since, with the appropriate analysis algorithms, they will be able to recognize the surrounding scene and understand gestures and social relationships, improving daily activities such as work, sport, education and entertainment.

To allow the scientific community to discuss these challenging issues, researchers **Giovanni Maria Farinella** from the **University of Catania** and **Dima Damen** from **Bristol University** in 2018 released the first version, repeatedly extended later on, of a database made up of videos taken through cameras worn by people who, in their own kitchen, perform daily actions in a natural way. Today, the **Epic-Kitchens-100 database** contains 100 hours of videos taken in 45 different kitchens and annotated with as many as 90000 elementary actions (such as “take the glass”, “open the tap”, “wash the glass” …).

Using this database, the best researchers dedicated to the study of Machine Learning algorithms for the creation of **wearable artificial vision systems** took part in various challenges. In 2022, the challenges were five: **action recognition, action detection, action anticipation, cross-modal retrieval (from captions), unsupervised domain adaptation for action recognition**.

These are the topics discussed during the annual EPIC workshop associated with the IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR), the main annual event on artificial vision, during which the winners were announced.

The international team in which FBK participated, through the TeV research unit (Digital Industry Center) with **Alex Falcon**, r**anked first in the "EPIC_KITCHENS-100: Multi-Instance Retrieval" challenge**.This challenge consists of learning to find and sort by relevance the video segments related to a given action described in text format.

Besides Alex, the winning team is composed of **Oswald Lanz** (currently professor at the Free University of Bolzano), **Giuseppe Serra** (University of Udine) and **Sergio Escalera** (Autonomous University of Barcelona).

Alex is a third-year student in an FBK and University of Udine joint doctorate program, supervised by Oswald Lanz and Giuseppe Serra in the role of co-advisor. His thesis work focuses on the joint analysis of videos and natural language descriptions through deep learning techniques.

PERMALINK

https://magazine.fbk.eu/en/news/an-international-recognition-for-semantic-video-understanding/

TAGS

- #artificialintelligence
- #augmented reality
- #digitalindustry
- #Machine learning
- #semantics
- #tev

AUTHORS

- Silvia Malesardi