

Un riconoscimento internazionale per la comprensione semantica di video

6 Luglio 2022

L'Unità TeV di FBK ha dato un contributo rilevante al team che si è aggiudicato il primo posto alla competizione EPIC-KITCHENS sulla visione egocentrica.

Telecamere mobili e occhiali intelligenti sono dispositivi sempre più diffusi sia come strumenti di studio per la ricerca che come prodotti alla portata di tutti sul mercato. Catturano le interazioni con la realtà di chi li indossa e oggi si inizia a capire come queste tecnologie e le loro applicazioni possano avere un impatto sulla nostra vita.

Si tratta in termini tecnici di dispositivi di **rilevamento "egocentrico"** che potrebbero presto essere in grado di assistere e facilitare chi li indossa, in quanto, con gli opportuni **algoritmi di analisi**, saranno in grado di riconoscere la scena circostante e di comprendere i gesti e le relazioni sociali, migliorando le attività quotidiane come il lavoro, lo sport, l'istruzione e l'intrattenimento.

Per permettere alla comunità scientifica di confrontarsi su questi temi sfidanti, i ricercatori **Giovanni Maria Farinella della Università di Catania e Dima Damen della Bristol University** nel **2018** hanno rilasciato la **prima versione**, poi continuamente estesa, di un database formato da video ripresi attraverso telecamere indossate da persone che, nella propria cucina, compiono azioni quotidiane in modo naturale. Oggi il **database Epic-Kitchens-100**, contiene 100 ore di video ripresi in 45 cucine diverse e annotati con ben 90000 azioni elementari (come "prendere il bicchiere", "aprire il rubinetto", "lavare il bicchiere" ...).

Su questo database si sono cimentati in varie gare i migliori ricercatori dediti allo studio di algoritmi di **Machine Learning** per la realizzazione di sistemi di visione artificiale indossabili. Nel 2022 le gare sono state cinque: **action recognition, action detection, action anticipation, cross-modal retrieval (from captions), unsupervised domain adaptation for action recognition**.

Di questi temi tratta l'annuale workshop [EPIC](#) associato alla conferenza [IEEE / CVF Computer Vision and Pattern Recognition Conference](#) (CVPR), principale evento annuale sulla visione artificiale, durante la quale è avvenuta la nomina dei vincitori.

Il team internazionale a cui ha partecipato FBK, attraverso l'unità di ricerca **TeV** ([Centro Digital Industry](#)) con [Alex Falcon](#), si è classificato al **primo posto nella gara “EPIC_Kitchens-100: Multi-Instance Retrieval”**. Questa gara consiste nell'imparare a ritrovare ed a ordinare per rilevanza i segmenti di video relativi ad una data azione descritta in formato testuale.

Il team vincitore è composto oltre che da **Alex, da Oswald Lanz** (attualmente professore presso la Libera Università di Bolzano), da **Giuseppe Serra** (Università di Udine) e da **Sergio Escalera** (Universitat Autònoma de Barcelona).

Alex è uno studente al terzo anno di un dottorato congiunto tra FBK e l'Università di Udine, supervisionato da Oswald Lanz e da Giuseppe Serra nel ruolo di co-advisor. La sua tesi ha per tema l'analisi congiunta di video e descrizioni in linguaggio naturale tramite tecniche di deep learning.

LINK

<https://magazine.fbk.eu/it/news/un-riconoscimento-internazionale-per-la-comprensione-semantic-di-video/>

TAG

- #industriadigitale
- #intelligenzaartificiale
- #Machine learning
- #realtà aumentata
- #semantica
- #tev

AUTORI

- Silvia Malesardi