FONDAZIONE BRUNO KESSLER
MAGAZINE

A Researcher from Trentino awarded at the most important world conference on voice recognition

March 31, 2017

The reward of $ 1,500 was awarded to Mirco Ravanelli, first author of the paper by FBK and University of Montreal (Canada).

A global competition at the highest levels.

There were over 2,600 works submitted – and of these, 1,317 accepted – at the 42nd ICASSP 2017 international conference on research in the field of speech recognition and “signal processing” which was held March 5 to 9 in New Orleans, USA.

One of the four prizes for best PhD student papers – the “IBM Best Student Paper Award”- was awarded to researcher Mirco Ravanelli, a thirty-one year old researcher from Trentino working under the “SHINE” team of Trento based Fondazione Bruno Kessler’s ICT Center

> “I joined the Foundation to work on my thesis (ed. Telecommunications Engineering) “- he says.
> “I am at my last year of doctorate and this award is an important recognition for me. Last year’s experience has been crucial.
> I spent six months in Montreal, Canada, in the laboratory of Yoshua Bengio, one of the deep learning gurus.
> Bengio has an H-index of 92, and his lab receives lots of applications to work there. Consider this: last year he received about 700 applications, but accepted only 20.
> I was one of these. ”
>
> “During my Mobility stay (Ed. the FBK program that allows researchers and scientists to spend a period of study and work abroad)” – Ravanelli concludes – “I was lucky enough to work in this prestigious lab, where I developed new research methods and work alongside world-renowned scientists.”
> The study by Ravanelli and his colleagues that was awarded in New Orleans, entitled “A network of deep neural networks for distant speech recognition”, although developed in the context of speech recognition, defines a new particularly flexible paradigm, which can be applied in many other areas, including robotics, where different systems must

cooperate at their best to achieve one common goal.
“It’s a top scientific excellence result” – adds Maurizio Omologo,

Head of the SHINE (Speech-Acoustic Scene Analysis and Interpretation) research unit.”The paper awarded at CASSP 2017 describes the work carried out by Mirco last year, during his Mobility stay in Montreal, largely based on ideas and strategies outlined already at the time of the doctoral first year qualifying exam “.

Study Abstract

A network of deep neural networks for distant speech recognition

Authors Mirco Ravanelli, Philemon Brakel, Maurizio Omologo, Yoshua Bengio

Building computers able to recognize speech is a fundamental step for the development of future human-machine interfaces and more generally for the development of artificial intelligence.

For these reasons, in recent years interest in this type of technology has grown a lot, and has led to numerous commercial applications such as Apple’s Siri, Google Voice, and Amazon’s Alexa voice assistants.

However, most of the existing systems provide satisfactory performance only in low noise environments where users speak near a microphone, making interesting the study of more robust technologies capable of operating even at some distance from the talker, in complex environments characterized by various types of noise and reverberation.
The work presented at the ICASSP 2017 (42nd International Conference on Acoustics, Speech, and Signal Processing) International Conference, approaches this problem by using advanced deep learning techniques.

Deep learning is a technology that is dramatically transforming the artificial intelligence world and allows for the development of systems that, starting from the so-called “big data”, are able to solve, with unprecedented precision, many problems such as, the identification of particular objects in images and videos, the translation of texts into another language, or, actually, speech recognition.

More precisely, the paper proposes to make significant changes in the ways in which this technology is used in current distant speech recognition systems.

Even the most advanced systems, in fact, are based on the cascade of different technologies that are typically developed independently, without any guarantee as to their actual compatibility and appropriate level of communication and cooperation.

The current system could therefore be likened to a soccer team which, although made up of good individual players, will never be able to win the championship if these players are not able to play well together.

The network of deep neural networks proposed in the study thus drastically revises the design of current systems, “breaking” this cascade by placing all the technologies involved in speech recognition within a network where there is full communication and cooperation between the various elements.

The research focus was not only on the definition of a new architecture, but above all on the study of innovative techniques to automatically train this network of components.

The work has led to the definition of a new training algorithm (called backpropagation through network) in which all the elements in the network are trained jointly and progressively learn to interact in an automatic way, communicate and cooperate with each other.

The work, though developed in the context of speech recognition, defines a general purpose and flexible paradigm, which can be applied in many other application areas, including robotics, where different systems must work together at their best in order to achieve a common goal.

PERMALINK

https://magazine.fbk.eu/en/news/researcher-trentino-awarded-important-world-conference-voice-recognition/

TAGS

- #best paper award
- #FBK
- #icassp
- #reti neurali
- #ricerca
- #riconoscimento vocale
- #shine
- #speech recognition
- #Technology

RELATED MEDIA

- IBM Best Paper Award: https://magazine.fbk.eu/wp-content/uploads/2017/03/20170322_141308.jpg
- Mirco Ravanelli: https://magazine.fbk.eu/wp-content/uploads/2017/03/Mirco_Ravanelli.jpg

AUTHORS

- Marzia Lucianer