

Paris, France
April 23-26, 2019



6th International Conference on Control,
Decision and Information Technologies

CoDIT'19

CALL FOR PAPERS
SPECIAL SESSION ON
“Deep Learning for Audio Information Retrieval and Computer Vision”
for CODIT'19
April 23-26, 2019 – Paris, France

Session Co-Chairs :

Fu-Hai Frank Wu, National Tsing Hua University, *Taiwan*

Session description

Recent years, the deep learning(DL) techniques have applied to audio(speech, music, sound etc.) and video(including image). There are differences and commons for these two fields in term of input data type, usually the video is in raw formats (mostly RGB pixel data) and the audio in pre-processed formats, for example spectrogram). Besides, the focus of data augmentation, including synthetic data, strategies are different to be effective in the training phase. In the respective of DL architecture, convolution kernel sizes and pooling strategies are generally distinct. The kernel size of audio DL is rectangular with the long-end in time axis and the other is square due to the import localized characteristic of image in fully convolutional network. The countermeasure beside the rectangular kernel size for the long-term characteristic of audio could be the adoption of recurrent network, for example long-short term memory(LSTM). The DL for audio and video is a broad topic, besides the discriminant problems we mention, it is obvious that tons of issues could be addressed. We also welcome the research of cross-domain inter-activities, although mostly audio IR borrow the DL outcome from the computer vision.

The special session will gather the researchers in the field of DL for audio information retrieval and computer vision to share the research progress, new finding, and state-of-the-art algorithms. We hope to explore and enumerate the common methods could be shared by studying the individual field. We expect to inspire and foster cross-domain improvement and increase the multi-modality research.

The topics of interest include, but are not limited to:

- music lyrics and other textual data, web mining, and natural language processing
- multi-modality
- corpus creation
- musical rhythm, beat, tempo
- optical music recognition
- text in scene
- music synthesis and transformation
- automatic classification
- indexing and querying
- pattern matching and detection
- human-computer interaction
- gaming
- action recognition
- recognition, detection, categorization, indexing
- segmentation, grouping and shape representation

SUBMISSION

Papers must be submitted electronically for peer review through [PaperCept](http://controls.papercept.net/conferences/scripts/start.pl) by **December 05, 2018**: <http://controls.papercept.net/conferences/scripts/start.pl>. In PaperCept, click on the **CoDIT 2019 link "Submit a Contribution to CoDIT'19" and follow the steps.**

All papers must be written in English and should describe original work. The length of the paper is limited to a maximum of 6 pages (in the standard IEEE conference double column format).

DEADLINES

December 19, 2018: deadline for paper submission

February 08, 2019: notification of acceptance/reject

February 28, 2019: deadline for final paper and registration.