

CALL FOR PAPERS: IEEE Journal of Selected Topics in Signal Processing Special Issue on Self-Supervised Learning for Speech and Audio Processing

There is a trend in the machine learning community to adopt self-supervised approaches to pre-train deep networks. Self-supervised learning utilizes proxy supervised learning tasks, for example, distinguishing parts of the input signal from distractors, or generating masked input segments conditioned on the unmasked ones, to leverage data from unlabeled corpora. These approaches make it possible to use a tremendous amount of unlabeled data available on the web to train large networks to extract high-level, informative, and compact features from raw inputs. Then for downstream applications, a simple task-specific model is added on top of the self-supervised model. The self-supervised model can serve as a feature extractor or be fine-tuned with the task-specific model.

Recently self-supervised approaches for speech and audio processing are gaining attention. This special issue will bring the works on self-supervision for the field of speech and audio processing. Alongside research work on new self-supervised methods, data, applications, and results, this special issue will call for novel work on understanding, analyzing, and comparing different self-supervision approaches for speech and audio processing. We welcome the submission of the work on, but not limited to, the following research directions:

- New self-supervised proxy tasks for speech and audio processing.
- New approaches to use self-supervised models in speech and audio processing tasks.
- Compare the similarities and differences of self-supervised learning approaches.
- Theoretical or empirical studies on understanding why self-supervision methods work.
- Exploring the limits of self-supervised learning approaches for speech and audio processing, for example, universal pre-trained model for multiple downstream tasks, environment conditions, or languages.
- Self-supervised learning approaches involving the interaction of speech and other modalities.
- Comparison or integration of self-supervised learning methods and other semi-supervised and transfer learning methods.

Prospective authors should visit <http://www.signalprocessingsociety.org/publications/periodicals/jstsp/> for information on paper submission. Manuscripts should be submitted using the Manuscript Central system at <http://mc.manuscriptcentral.com/jstsp-ieee>. Manuscripts will be peer-reviewed according to the standard IEEE process.

Important Dates:

Submission deadline: December 31, 2021

First review completed: February 28, 2022

Revised manuscript due: April 30, 2022

Second review completed: June 15, 2022

Final manuscript due: July 31, 2022

Guest Editors

Hung-yi Lee, Lead Guest Editor, National Taiwan University, hungyilee@ntu.edu.tw

Abdelrahman Mohamed, Facebook, abdo@fb.com

Shinji Watanabe, Carnegie Mellon University, shinjiw@ieee.org

Karen Livescu, Toyota Technological Institute at Chicago, klivescu@ttic.edu

Tara Sainath, Google, tsainath@google.com