Journal of Health informatics journal
Special issue on
Preprocessing for Big biomedical Data

<u>**Overview**</u>

Due to the proliferation of biomedical imaging modalities such as Photoacoustic Tomography, Computed Tomography (CT), Optical Microscopy and Tomography, Single Photon Emission Computed Tomography (SPECT), Magnetic Resonance (MR) Imaging, Ultrasound, and Positron Emission Tomography (PET), Magnetic Particle Imaging, EE/MEG, Electron Tomography and Atomic Force Microscopy, massive amounts of biomedical data are being generated on a daily basis. How can we utilize such big data to build better health profiles and better predictive models so that we can better diagnose and treat diseases and provide a better life for humans? In the past years, many successful learning methods such as deep learning were proposed to answer this crucial question, which has social, economic, as well as legal implications.

A number of significant problems plague the processing of big biomedical data, such as data heterogeneity, data incompleteness, data imbalance, and high dimensionality. What is worse is that many data sets exhibit multiple such problems. A majority of existing learning methods can only deal with homogeneous, complete, class-balanced, and moderate-dimensional data. Therefore, data preprocessing techniques including data representation learning, dimensionality reduction, and missing value imputation should be developed to enhance the applicability of deep learning methods in real-world applications of biomedicine.

This special issue aims to provide a diverse, but complementary set of contributions to demonstrate new developments and applications that covers existing above issues in data processing of big biomedical data. We would also like to accept successful applications of the new methods, including but not limited to data processing, analysis and knowledge discovery of biomedical data.

Suggested topics include but not limited to:

- Feature extraction by deep learning or sparse codes for biomedical data
- Data representation of biomedical data
- Dimensionality reduction techniques (subspace learning, feature selection,
- sparse screening, feature screening, feature merging, etc) for biomedical data
- Information retrieval for biomedical data
- Kernel-based learning for multi-source biomedical data
- Incremental learning or online learning for biomedical data.
- Data fusion for multi-source biomedical data
- Missing data imputation for multi-source biomedical data
- Data management and mining in biomedical data
- Web search and meta-search for biomedical data
- Web information retrieval for biomedical data
- Biomedical data quality assessment
- Transfer learning of biomedical data.

<u>**Schedule:**</u>

Final paper due:	15 Jan 2021
First notification:	28 July 2021
Submission of revised manuscript:	28 Aug 2021

<u>Guest Editors:</u>
Shuihua Wang
University of Leicester, United Kingdom
Email: shuihuawang@ieee.org

Zhengchao Dong

Columbia University

Email: zhengchao.dong@nyspi.columbia.edu

Zheng Zhang
Harbin Institute of Technology
Email: zhengzhang@hit.edu.cn

Yuankai Huo
University of Vanderbilt, United States
Email: yuankai.huo@vanderbilt.edu