

Deep Learning for Motif Mining in Biological Sequences

De-Shuang Huang

Prof. & Ph.D, IEEE Fellow, IAPR Fellow & AAIA Fellow

Institute of Machine Learning and Systems Biology, EIT Institute for Advanced Study & Tongji University, China

Email: huangdeshuang@163.com

Personal website: <http://ic-icc.cn/DShuang/DShuangindex.html>

Abstract: Transcription factor (TF) plays a central role in gene regulation. Knowing the binding specificities of TFs is essential for developing models of the regulatory processes in biological systems and for deciphering the mechanism of gene expression. In this talk, we will introduce several novel computational models of TF binding data by combining various types of high-throughput data. Firstly, we will introduce a tensor decomposition model for collaborative prediction of ChIP-seq data, which could overcome its current limitation for integrative analysis. Secondly, we will present a de novo motif learning method based on the area under the receiver-operating characteristic curve (AUC) criterion, which has been widely used in the literature to evaluate the significance of extracted motifs. Finally, based on Fisher Exact Test score (FETS), we propose DirectFS, which is (to our best knowledge) the first FETS-based approach that allows direct learning of the motif parameters in continuous space. Experimental results based on real world high-throughput datasets illustrate that DirectFS outperforms competing methods for refining motifs found by de novo motif elicitation methods, while being one order of magnitude faster. In addition, we also present the possibility for using deep learning technique combining motif mining in biological sequences to address the medical image processing issue.

Bio-Sketch: **De-Shuang Huang** is a Professor in Institute of Machine Learning and Systems Biology, EIT Institute for Advanced Study, China and Tongji University, China. He is currently the Fellow of the IEEE (IEEE Fellow), the Fellow of the International Association of Pattern Recognition (IAPR Fellow), the Fellow of the Asia-Pacific Artificial Intelligence Association (AAIA), and associated editors of IEEE/ACM Transactions on Computational Biology & Bioinformatics and IEEE Transactions on Cognitive and Developmental Systems, etc. He founded the International Conference on Intelligent Computing (ICIC) in 2005. ICIC has since been successfully held annually with him serving as General or Steering Committee Chair. He also served as the 2015 International Joint Conference on Neural Networks (IJCNN2015) General Chair, July12-17, 2015, Killarney, Ireland, the 2014 11th IEEE Computational Intelligence in Bioinformatics and Computational Biology Conference (IEEE-CIBCBC) Program Committee Chair, May 21-24, 2014, Honolulu, USA. He has published over 470 papers in international journals, international conferences proceedings, and book chapters. Particularly, he has published over 240 SCI indexed papers. His Google Scholar citation number is over 20080 times

and H index 76. His main research interest includes neural networks, pattern recognition and bioinformatics. His main research interest includes neural networks, pattern recognition and bioinformatics.

Photo

