

## Accepted Manuscript

Protein Secondary Structure Prediction by using Deep Learning Method

Yangxu Wang, Hua Mao, Zhang Yi

PII: S0950-7051(16)30471-3  
DOI: [10.1016/j.knosys.2016.11.015](https://doi.org/10.1016/j.knosys.2016.11.015)  
Reference: KNOSYS 3739



To appear in: *Knowledge-Based Systems*

Received date: 9 March 2016  
Revised date: 16 November 2016  
Accepted date: 16 November 2016

Please cite this article as: Yangxu Wang, Hua Mao, Zhang Yi, Protein Secondary Structure Prediction by using Deep Learning Method, *Knowledge-Based Systems* (2016), doi: [10.1016/j.knosys.2016.11.015](https://doi.org/10.1016/j.knosys.2016.11.015)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Protein Secondary Structure Prediction by using Deep Learning Method

Yangxu Wang<sup>a</sup>, Hua Mao<sup>a,\*</sup>, Zhang Yi<sup>a</sup>

<sup>a</sup>*Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu 610065, People's Republic of China*

---

## Abstract

The prediction of protein structures directly from amino acid sequences is one of the biggest challenges in computational biology. It can be divided into several independent sub-problems in which protein secondary structure (SS) prediction is fundamental. Many computational methods have been proposed for SS prediction problem. Few of them can model well both the sequence-structure mapping relationship between input protein features and SS, and the interaction relationship among residues which are both important for SS prediction. In this paper, we proposed a deep recurrent encoder-decoder networks called Secondary Structure Recurrent Encoder-Decoder Networks (SSREDNs) to solve this SS prediction problem. Deep architecture and recurrent structures are employed in the SSREDNs to model both the complex nonlinear mapping relationship between input protein features and SS, and the mutual interaction among continuous residues of the protein chain. A series of techniques are also used in this paper to refine the model's performance. The proposed model is applied to the open dataset CullPDB and CB513. Experimental results demonstrate that our method can improve both Q3 and Q8 accuracy compared with some public available methods. For Q8 prediction problem, it achieves 68.20% and 73.1% accuracy on CB513 and CullPDB dataset in fewer epochs better than

---

\*Hua Mao

*Email addresses:* mellowxu@gmail.com (Yangxu Wang), huamao@scu.edu.cn (Hua Mao), zhangyi@scu.edu.cn (Zhang Yi)

This work was supported by the National Science Foundation of China (Grant No. 61432012 and 61402306).

Download English Version:

<https://daneshyari.com/en/article/4946381>

Download Persian Version:

<https://daneshyari.com/article/4946381>

[Daneshyari.com](https://daneshyari.com)