

DeepSF: Deep CNN for mapping protein sequences to folds

Zhiyang's presentation about [1]

Zhiyang Ong

Department of Electrical and Computer Engineering
Dwight Look College of Engineering,
Texas A&M University
College Station, TX

March 26, 2020

- ① Problem and Knowledge Gap
- ② Proposed Solution(s)
- ③ Experimental Results
- ④ Discussion and Muddy Points

Table of Contents

- 1 Problem and Knowledge Gap
- 2 Proposed Solution(s)
- 3 Experimental Results
- 4 Discussion and Muddy Points

Context of the Problem

Background Information

Context of the problem.

- ① The structures of most ($> 99\%$) proteins are unknown
- ② Protein fold recognition enables us to associate a protein sequence to a protein fold
 - With this protein fold, determine its protein structure [Tramontano2003]
 - With this protein structure, determine its protein function
 - [Osadchy2011,OConnor2010,Tramontano2003,Hegyi1999]
 - [OConnor2014a, From Contents: Unit 2, How Do Cells Decode Genetic Information into Functional Proteins?: §2.4 The Functions of Proteins Are Determined by Their Three-Dimensional Structures]
 - [OConnor2014, From Contents: Unit 2, How Do Cells Decode Genetic Information into Functional Proteins?: §2.4 The Functions of Proteins Are Determined by Their Three-Dimensional Structures]



Problem Definition

What is the problem that [1] is solving?

Problem description.

- ① Sequence-alignment methods for mapping protein sequences to protein folds are indirect
 - Can't explain relationship between protein sequences & protein folds
- ② Protein fold recognition enables us to associate a protein sequence to a protein fold
 - With this protein fold, we can determine its protein structure
 - With this protein structure, we can determine its protein function
 -



Problem Importance

Why is it important?

Why is this problem important?

- This facilitates protein structure prediction
- Protein structure prediction enables protein function prediction
- Knowing about protein structure and function facilitates:
 - drug/medication design [Nogrady2005, §1.6.4, pp. 54]
[Golan2008, Chapter 1, pp. 4]
 - biotechnology [Walsh2014, Chapter 2]
 - synthetic biology [Zhao2013d, Chapter 2]
 - personalized [Cullis2015, Chapter 2, pp. 26] or precision medicine [Mousa2020, §24.3.2, pp. 778]
 - gene therapy [Wecker2010, Chapter 5, pp. 51]



What are the Knowledge Gaps in [1]?

How would [1] address these knowledge gaps?

List of knowledge gaps:



Table of Contents

- ① Problem and Knowledge Gap
- ② Proposed Solution(s)
- ③ Experimental Results
- ④ Discussion and Muddy Points

Proposed Solution(s) - 1

What do they proposed to address the knowledge gap?

Proposed Solution(s).



Proposed Solution(s) - 2

What do they proposed to address the knowledge gap?

Proposed Solution(s).



Table of Contents

- ① Problem and Knowledge Gap
- ② Proposed Solution(s)
- ③ Experimental Results**
- ④ Discussion and Muddy Points

Experimental Results - 1

How do their proposed solutions compare with existing solutions?

Experimental Results - 1.



Experimental Results - 2

How do their proposed solutions compare with existing solutions?

Experimental Results - 2.



Table of Contents

- 1 Problem and Knowledge Gap
- 2 Proposed Solution(s)
- 3 Experimental Results
- 4 Discussion and Muddy Points**

Discussion of Experimental Results

What do the experimental results tell us?

Discuss the experimental results.



Muddy Points

What do I not understand about [1]?

What do I not understand about [1]?



References



Jie Hou, Badri Adhikari, and Jianlin Cheng.

DeepSF: deep convolutional neural network for mapping protein sequences to folds.

Bioinformatics, 34(8):1295–1303, April 15 2018.