**Multichannel Autoencoder for Multimodal DNA Representation**

Ki-wook Leea,b

*a Correspondence to :* [*leek0502@gmail.com*](mailto:leek0502@gmail.com)

*b Department of Digital Health, SAIHST, Sungkyunkwan University, Seoul 06351, South Korea*

# Abstract

Recently several deep learning models have been used for DNA/RNA/Amino acids sequence based classification and regression tasks. In this work, we use a multimodal autoencoder model to learn a latent representation of DNA/RNA/Amino acids sequences in an unsupervised manner. Our model first learns the associations between DNA sequences and other genetic information (e.g. RNA seondary structure, Chromatin accessibility of cell line) . Then the DNA sequences and other genetic information representations are fused through reconstructing their original and associatted embeddings. Using a gating mechanism our model assigns different weights to each modality according to the different concepts. Results on three benchmark concept (e.g. One regression task of Crispr guideRNA efficiency prediction, two classification task of MHC binding affinity prediction, and DNA motif mining) show that the proposed method significantly outperforms than other biologic representation methods.

Keywords: Representation learning, DNA, RNA, Deep learning, Auto encoder

# Introduction

# 

# Related Work

# 

# Methods

## Associative Multichannel Autoencoder

* *Basic Multichannel Autoencoder*
* *Integrating Biologic Associations*

## Architecture of prediction models

* *Crispr guideRNA efficiency task*
* *MHC binding affinity task*
* *DNA motif finding task*.

## Hyperparameter Optimization

# Results

***Comparison of previous representation methods***

*.*

***Crispr guideRNA efficiency task***

*.*

***MHC binding affinity task***

*.*

***DNA motif finding task***

***.***

# Discussion

.

## Author Contributions and Notes

Ki-wook designed research, performed research, wrote software, analyzed data; and wrote the paper.

## Availability of Data and Code

All code is available at https://github.com/LEEKIWOOK/DACO2.

# References

1.