02-6

# Interpretable Deep Learning Using Multimodal Graph Convolutional Network for Predicting Compound-Protein Interactions

**Sota Inoue**[1] inoue.sota.65r@st.kyoto-u.ac.jp
**Yoshihisa Tanaka**[2] tanaka.yoshihisa.42n@st.kyoto-u.ac.jp
**Shoichi Ishida**[3] ishida.sho.nm@yokohama-cu.ac.jp
**Ryosuke Kojima**[1] kojima.ryosuke.8e@kyoto-u.ac.jp
**Hiroaki Iwata**[1] iwata.hiroaki.3r@kyoto-u.ac.jp
**Yasushi Okuno**[1] okuno.yasushi.4c@kyoto-u.ac.jp

[1] Graduate School of Medicine, Kyoto Univ., 53 Kawahara-cho, Syogoin, Sakyo-ku, Kyoto, 606-8507, Japan
[2] Graduate School of Pharmaceutical Sciences, Kyoto Univ., 46-29 Yoshida-Shimo-Adachi-cho, Sakyo-ku, Kyoto, 606-8501, Japan
[3] Graduate School of Medical Life Science, Yokohama City Univ., 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama, 230-0045, Japan

**Keywords**: in silico Screening, Compound-Protein Interaction, Graph Convolutional Network

Understanding compound-protein interactions (CPI) is important for identifying hit compounds in the early stages of drug discovery. Recent advances in deep learning technology have enabled rapid and accurate calculations[1], but it remains difficult to interpret the physicochemical and biological aspects of the prediction results. For example, to visualize the prediction results, mapping the weight of Attention mechanism to the 3D structure of the protein or using Grad-CAM has been attempted[2,3]. However, this is still not enough to evaluate the results objectively, and no discussion has been obtained to validate the model. In our study, we propose a multimodal and interpretable CPI prediction model by adapting Graph Convolutional Networks (GCN) for small molecules and Convolutional Neural Networks (CNN) for proteins, and by using Integrated Gradients (IG) to visualize the contribution of each atom in small molecules and each residue in proteins to the prediction results. As a result of visualization with our model, it is suggested that the prediction results are obtained by capturing the features of chemical structures of small molecules and protein sequences. This indicates that when screening hit compounds using Deep Learning-based CPI prediction models, biological reliability of the prediction results needs to be carefully considered.

[1] Hamanaka, M.; Taneishi, K.; Iwata, H.; Ye, J.; Pei, J.; Hou, J.; Okuno, Y.; CGBVS-DNN: Prediction of Compound-protein Interactions Based on Deep Learning, *Mol. Inf.*, **2017**, 36, 1600045.
[2] Tsubaki, M.; Tomii, K.; & Sese, J.; Compound-protein interaction prediction with end-to-end learning of neural networks for graphs and sequences, *Bioinformatics*, **2019**, 35(2), 309–318.
[3] Verma, N.; Qu, X.; Trozzi, F.; Elsaied, M.; Karki, N.; Tao, Y.; Zoltowski, B.; Larson, E.C.; Kraka, E. SSnet: A Deep Learning Approach for Protein-Ligand Interaction Prediction, *Int. J. Mol. Sci.*, **2021**, 22, 1392.