DeepROCK: Error-controlled interaction detection in deep neural networks

Winston Chen[1], William Stafford Noble*[2,3], Yang Young Lu*[4]

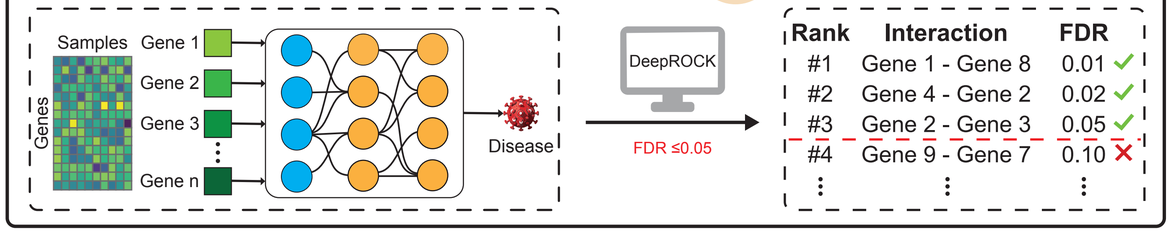
1.Computer Science and Engineering Division, University of Michigan 2.Department of Genome Science, University of Washington
3.Paul G. Allen School of Computer Science and Engineering , University of Washington 4.Cheriton School of Computer Science, University of Waterloo

Paper Code

Motivation

Detecting interactions with false discovery rate (FDR) control is crucial for scientific discovery

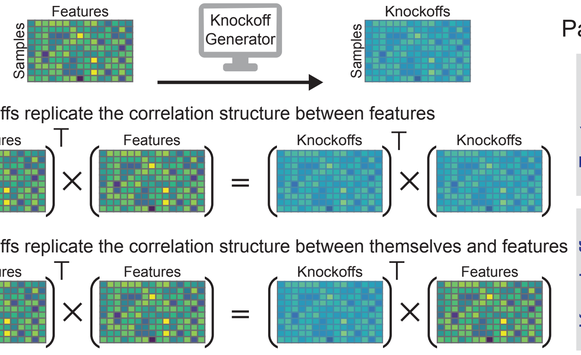
$$\text{FDR} = E\left[\frac{\#\text{ false selected interactions}}{\#\text{ total selected interactions}}\right]$$



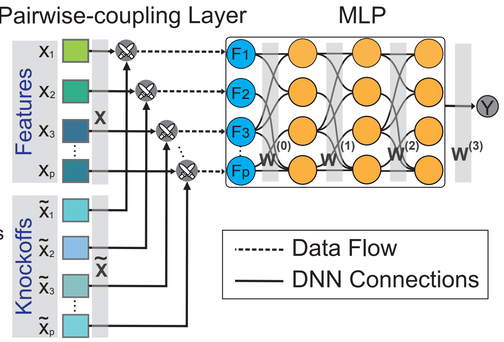
Rank	Interaction	FDR
#1	Gene 1 - Gene 8	0.01 ✓
#2	Gene 4 - Gene 2	0.02 ✓
#3	Gene 2 - Gene 3	0.05 ✓
#4	Gene 9 - Gene 7	0.10 ✗

Method

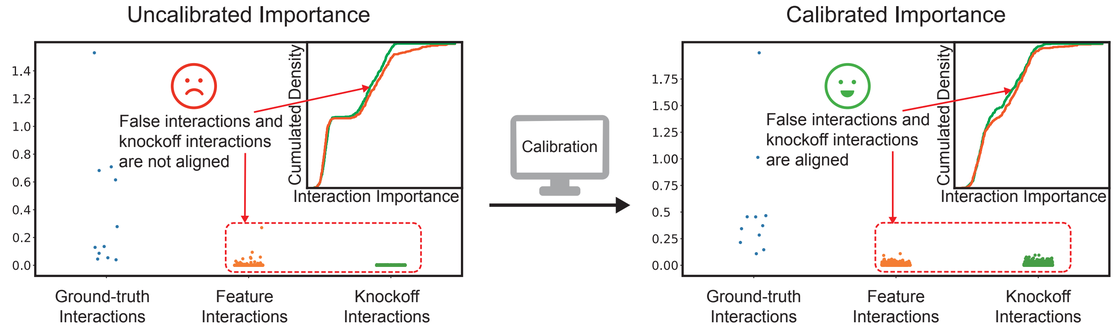
DeepROCK controls FDR using knockoffs



DeepROCK maximizes power using a novel pairwise-coupling layer

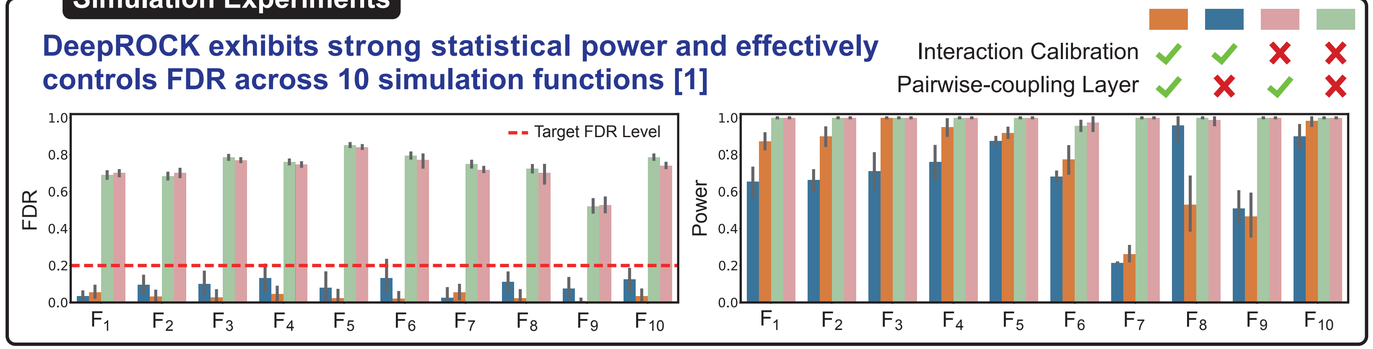


DeepROCK leverages a calibration procedure to correct FDR estimate



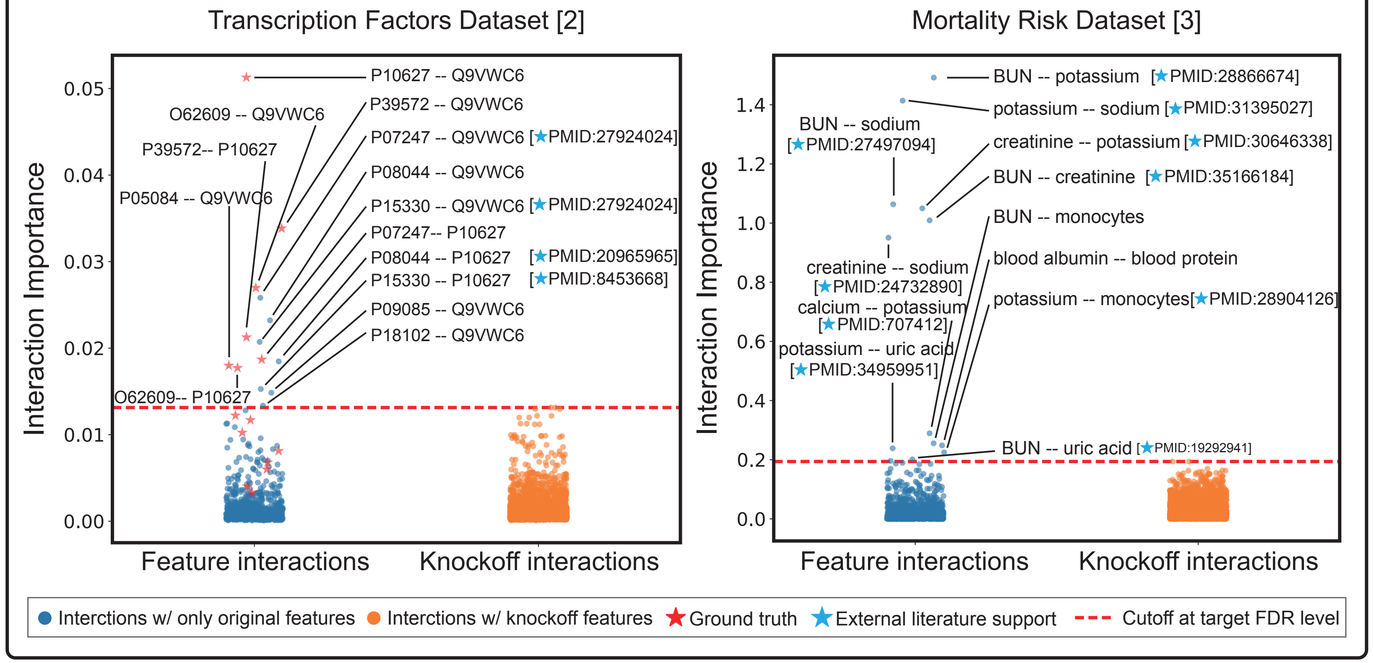
Simulation Experiments

DeepROCK exhibits strong statistical power and effectively controls FDR across 10 simulation functions [1]



Real Data Experiments

DeepROCK discovers experimentally validated and literature-supported real-world interactions



References

[1] M. Tsang, et al. Detecting statistical interactions from neural network weights. International Conference on Learning Representations, 2018.
[2] S. Basu, et al. Iterative random forests to discover predictive and stable high-order interactions. Proceedings of the National Academy of Sciences, 2018
[3] C. S. Cox, et al. Plan and operation of the NHANES I Epidemiologic Follow up Study, 1992. 1997.