A209 · A Computational Framework for Recovering Molecular Relationships from Biological Pathway Diagrams

Presenter: Xiwen Zhao — Northeastern University

Authors: Xiwen Zhao, Benjamin King, Mingyang Lu

Session: Day 2 · Fri Oct 2, 2026 · 2:15–4:15 PM

Biological pathways contain rich mechanistic knowledge about molecular interactions; however, much of this information exists primarily in static images rather than structured, machine-readable representations. Although some databases provide pathway annotations in structured formats, many regulatory and signaling relationships explicitly depicted in pathway diagrams are not fully captured in these representations. Existing methods for automated parsing of biological pathway diagrams largely focus on local relationship extraction rather than modeling the global structure of pathway diagrams as a whole, and therefore struggle to resolve complex topological patterns such as branching structures and feedback loops.

Here, we propose a computational framework that fine-tunes visual recognition models on manually labeled KEGG pathway diagrams to recognize key visual elements, including line segments, oriented junction points, and biological symbols, and subsequently assembles these elements into a graph representation, thereby recovering molecular relationships directly from pathway image. For visual element recognition, our framework achieves a structural average precision (sAP15) of 81.4% for line segment detection, representing the best performance among the compared methods. It also achieves OKS AP50 scores of 0.95 and 1.00 for junction and symbol recognition, respectively, for which no directly comparable baseline is currently available. This work provides a scalable computational approach for extracting molecular relationships directly from pathway diagrams, with the potential to augment existing pathway resources and support downstream computational analyses.