Graph Convolutional Cost Models in TVM: A Drop in the Ocean
Automatic Optimization in TVM

- Expression: $e$
- Schedule Space: $S_e$
- Exploration Module
- Cost Model: $\hat{f}(x)$
- Objective Function
- Code Generator: $x = g(e, s)$
- Hardware Environment
- Experiment Feedback
Main Limitations of Current Cost Models

- Limited generalization between workloads (e.g., convolutions with different shapes)
- Limited generalization between different computation expressions (e.g., conv2d vs. matrix multiply)
Use a Graph Representation Directly

- TVM IR provides an AST: this AST can be lowered to a graph format
Early Results

RMSE of GCN models compared to baselines (ResNet-18 workloads)