Graph Representation Learning Using Graph Neural Networks

Main Papers

- Gated Graph Sequence Neural Networks
  Li Y., Tarlow D., Brockschimdt M., Zemel R., ICLR 2016

- Semi-Supervised Classification with Graph Convolutional Networks
  Kipf T., Welling M., ICLR 2017

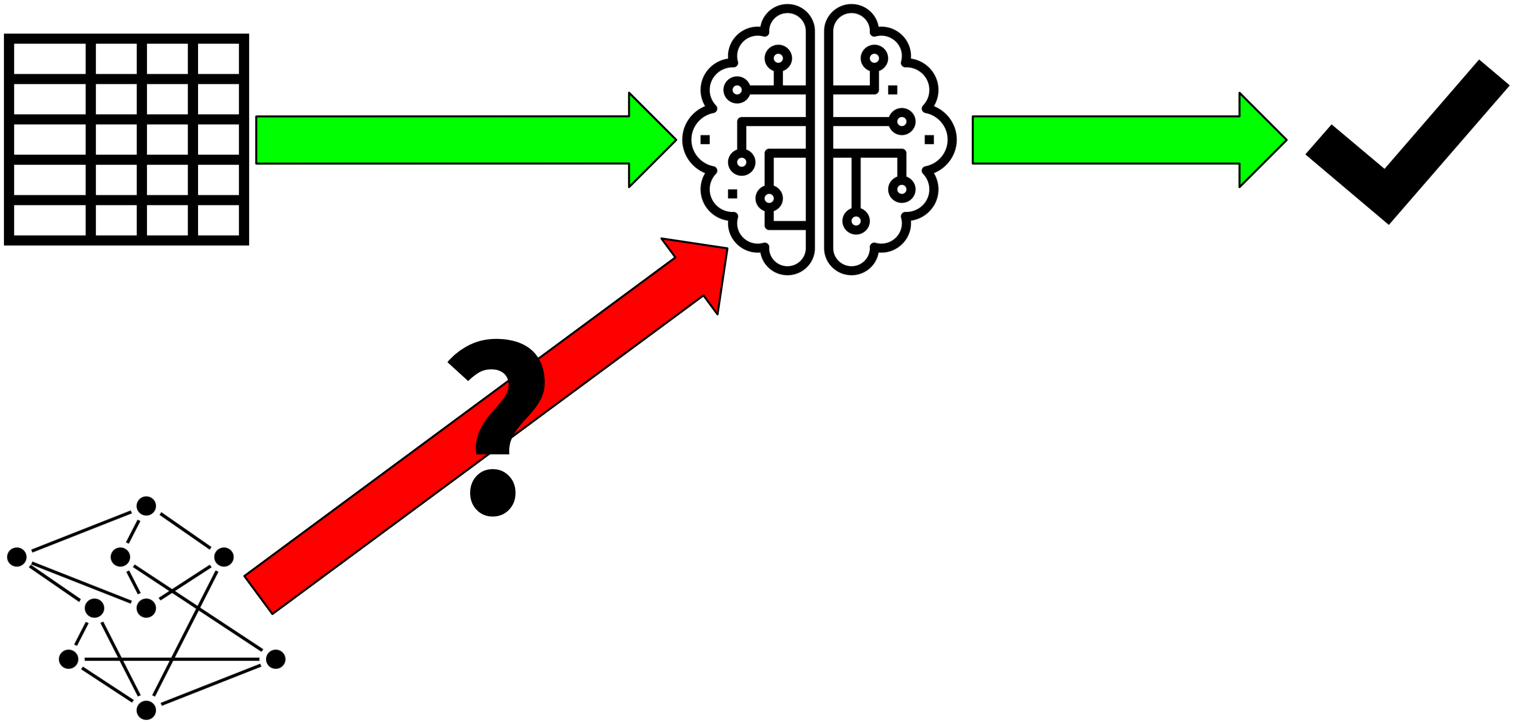
- Inductive Representation Learning on Large Graphs
  Hamilton W., Ying Z., Leskovec J., NIPS 2017

Agenda

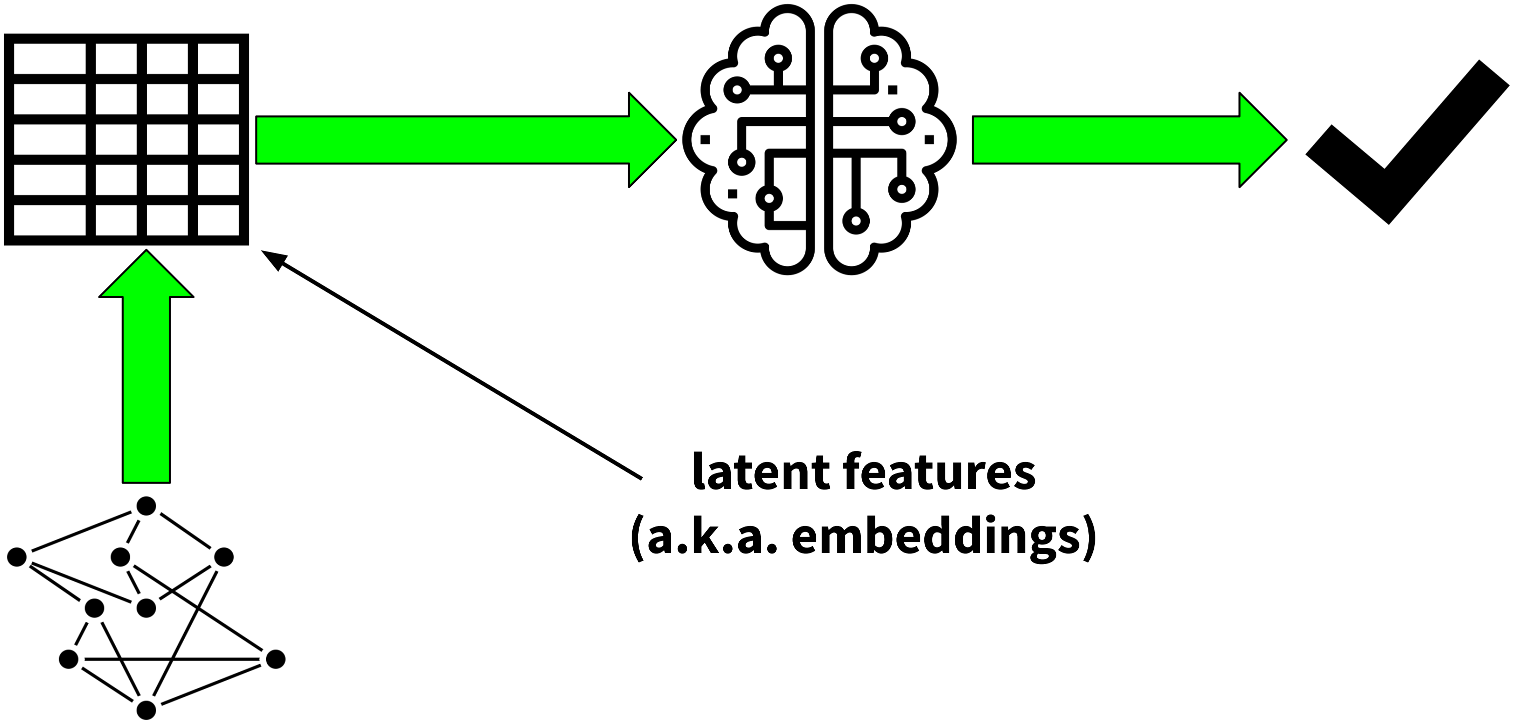
- What is GRL?
- What are GNNs?
- Recurrent GNNs (Li et al.)
- Spectral Conv. GNNs (Kipf et al.)
- Spatial Conv. GNNs (Hamilton et al.)

What is
Graph Representation Learning?

The Problem



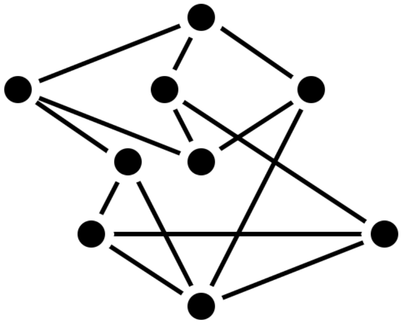
The Solution: GRL!



What are
Graph Neural Networks?

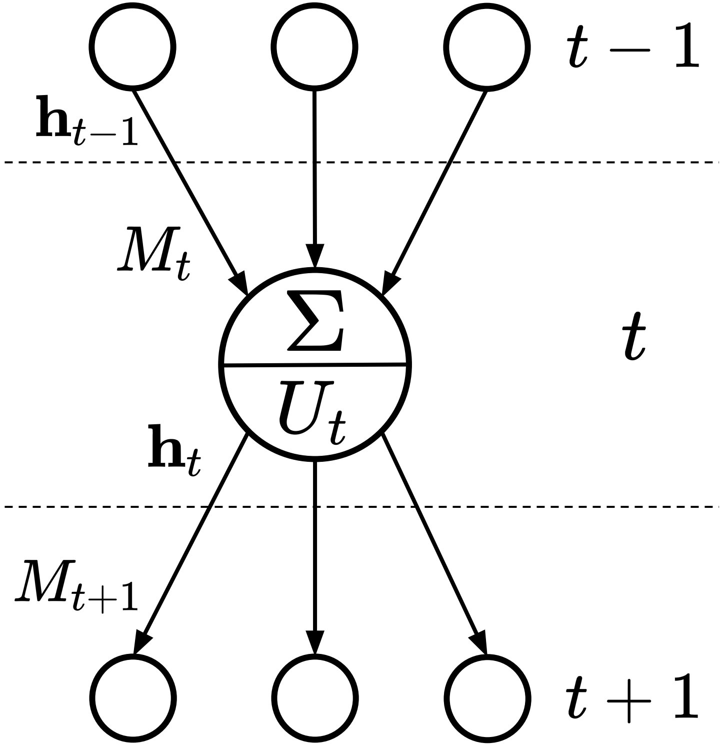
Common GNN Characteristics

- **Inductive** representation learning
- Exploit **node features**
- **Task-specific** training
- **Semi-supervised** training



GNN Architecture

$$\mathbf{h}_v^t = U_t \left(\mathbf{h}_v^{t-1}, \sum_{w \in N(v)} M_t(\mathbf{h}_v^{t-1}, \mathbf{h}_w^{t-1}, \mathbf{e}_{vw}) \right)$$



2 main types:

- Recurrent GNNs
- Convolutional GNNs

Li et al.:
Gated Graph Sequence Neural Networks

Gated Graph Neural Network

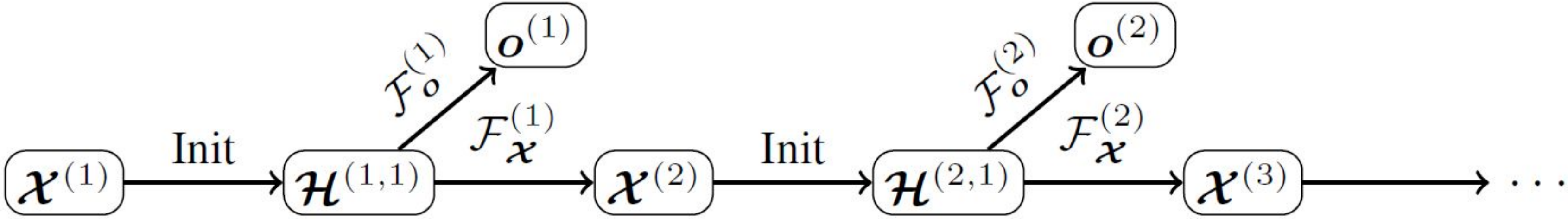
$$\mathbf{h}_v^t = g\left(\mathbf{h}_v^{t-1}, \sum_{w \in N(v)} f(\mathbf{h}_w^{t-1})\right)$$

- ✗ Runs to convergence
- ✓ Memory efficient

$$\mathbf{h}_v^t = GRU\left(\mathbf{h}_v^{t-1}, \sum_{w \in N(v)} \mathbf{W}\mathbf{h}_w^{t-1}\right)$$

- ✓ Fixed iteration count
- ✗ Large memory requirement

Gated Graph Sequence Neural Networks



Kipf et al.:
Semi-Supervised Classification with Graph Convolutional Networks

Spectral Graph Convolutions

$$h = \mathcal{F}^{-1}\left(w \odot \mathcal{F}(x)\right)$$

$$\mathbf{h} = \mathbf{U}\mathbf{W}\mathbf{U}^T\mathbf{x}$$

$$\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}$$

$$\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

Graph Convolutional Neural Networks

$$\mathbf{h} = \mathbf{U}\mathbf{W}\mathbf{U}^T\mathbf{x}$$

- ✗ O(N) weights
- ✗ O(N^3) time
- ✓ Clear interpretation

$$\mathbf{h} = w\tilde{\mathbf{A}}\mathbf{x}$$

- ✓ Lightweight
- ✗ Less interpretable

GCNN: Full Formulation

$$\mathbf{H}^t = f\left(\tilde{\mathbf{A}}\mathbf{H}^{t-1}\mathbf{W}^t\right)$$

N x K_{OUT}
output channels

N x N
diffusion matrix

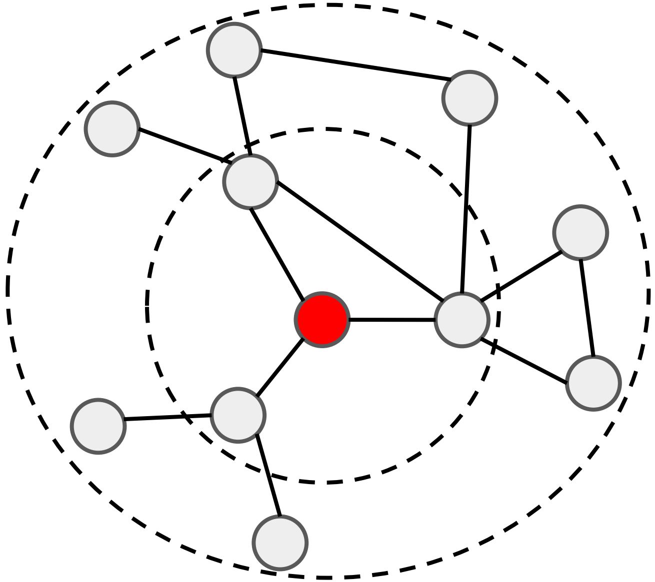
N x K_{IN}
input channels

K_{IN} x K_{OUT}
trainable filter

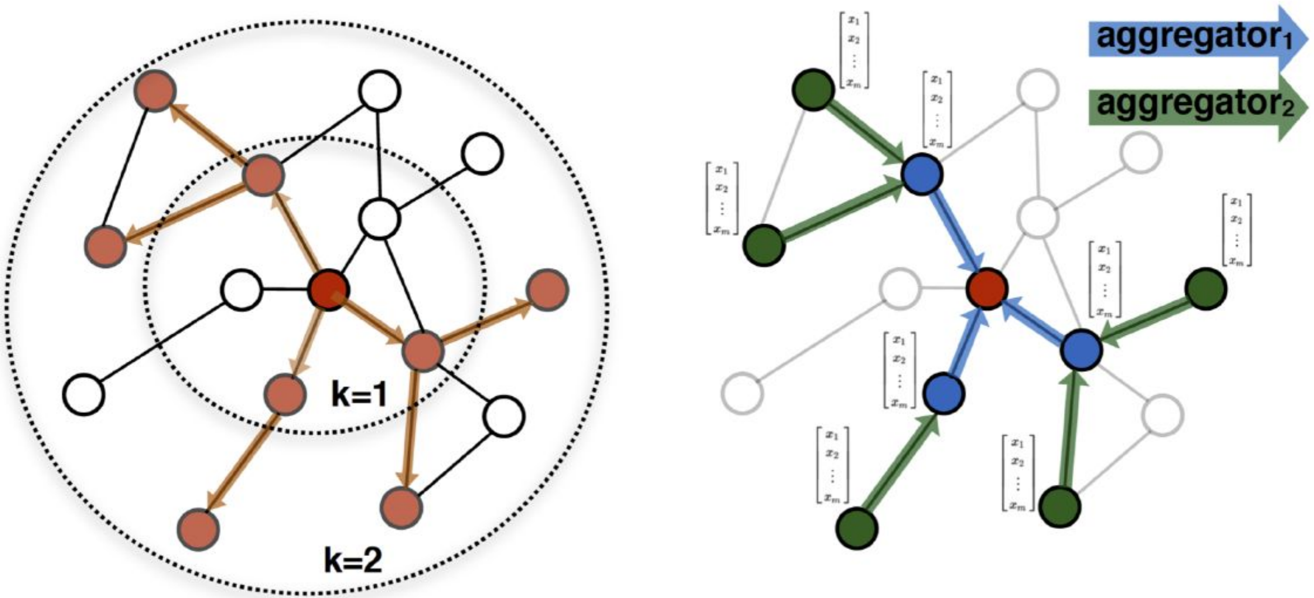
Hamilton et al.:
Inductive Representation Learning on Large Graphs

Spatial Convolutional GNNs

$$\mathbf{H}^t = f\left(\tilde{\mathbf{A}}\mathbf{H}^{t-1}\mathbf{W}^t\right)$$



GraphSAGE



1. Sample neighborhood

2. Aggregate feature information from neighbors

Comparison

A GRL-wide Issue

- No common application areas
- No common datasets
- Non-comparable approaches
- No overall benchmarks
- An explosion of techniques

Approach	Category
GNN* (2009) [15]	RecGNN
GraphESN (2010) [16]	RecGNN
GGNN (2015) [17]	RecGNN
SSE (2018) [18]	RecGNN
Spectral CNN (2014) [19]	Spectral-based ConvGNN
Henaff et al. (2015) [20]	Spectral-based ConvGNN
ChebNet (2016) [21]	Spectral-based ConvGNN
GCN (2017) [22]	Spectral-based ConvGNN
CayleyNet (2017) [23]	Spectral-based ConvGNN
AGCN (2018) [40]	Spectral-based ConvGNN
DualGCN (2018) [41]	Spectral-based ConvGNN
NN4G (2009) [24]	Spatial-based ConvGNN
DCNN (2016) [25]	Spatial-based ConvGNN
PATCHY-SAN (2016) [26]	Spatial-based ConvGNN
MPNN (2017) [27]	Spatial-based ConvGNN
GraphSage (2017) [42]	Spatial-based ConvGNN
GAT (2017) [43]	Spatial-based ConvGNN
MoNet (2017) [44]	Spatial-based ConvGNN
PGC-DGCNN (2018) [46]	Spatial-based ConvGNN
CGMM (2018) [47]	Spatial-based ConvGNN
LGCN (2018) [45]	Spatial-based ConvGNN
GAAN (2018) [48]	Spatial-based ConvGNN
FastGCN (2018) [49]	Spatial-based ConvGNN
StoGCN (2018) [50]	Spatial-based ConvGNN
Huang et al. (2018) [51]	Spatial-based ConvGNN
DGCNN (2018) [52]	Spatial-based ConvGNN
DiffPool (2018) [54]	Spatial-based ConvGNN
GeniePath (2019) [55]	Spatial-based ConvGNN
DGI (2019) [56]	Spatial-based ConvGNN
GIN (2019) [57]	Spatial-based ConvGNN
ClusterGCN (2019) [58]	Spatial-based ConvGNN

High-Level Comparison

GGNN
(Recurrent GNN)

✓ Low weight count
✓ Long-range dependencies

GCNN
(Spectral Conv. GNN)

✓ Mathematical foundation

GraphSAGE
(Spatial Conv. GNN)

✓ Flexibility
✓ Efficiency

Conclusions

GRL: too hyped, too crowded

Huge effort needed to organize things:

- Small, well-defined set of problems and applications
- Small set of standardized benchmarks
- Ablation studies (or something equivalent)