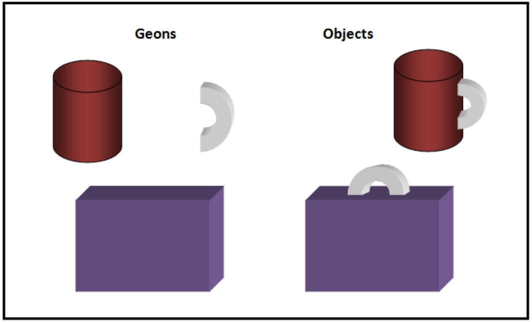
IIT Jodhpur

Biological Vision and Applications
Module 09-01: Graph Neural Networks

Hiranmay Ghosh

Structured representation

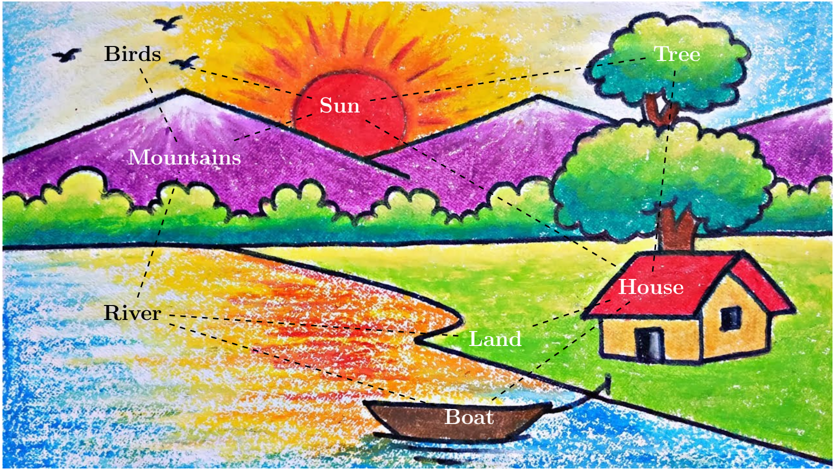
Explicit knowledge



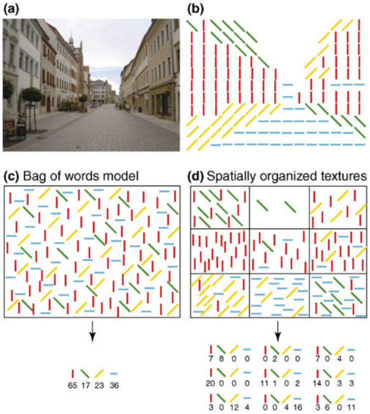
- Inductive Generalization
- Context

Structured representation

Spatial (and temporal) Organization



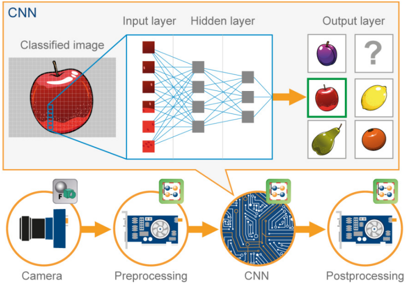
Holistic representation



- **No Inductive Generalization**

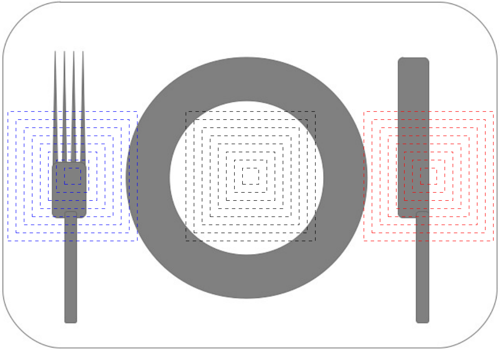
Machine learning

Holistic representation



- Does not use the explicit structured representation
- Flexibility: Knowledge is emergent
  - No dependence on hand-coded knowledge
  - Features and feature weights are machine learned

Does a CNN "see" the structure ?



Can we combine the benefits of the two approaches ?

Position paper from DeepMind, Google Brain, MIT, University of Edinburgh (2018)

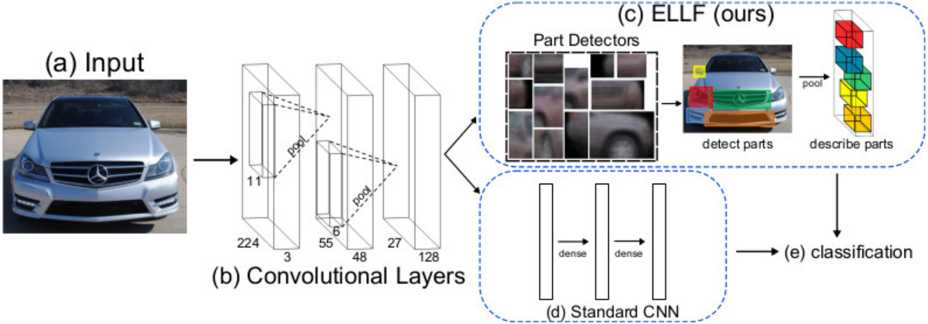
- Key to advancement of AI (cognition): Inductive generalization
- Structure with flexibility (emergent knowledge)
- "Intersection of deep learning and structured approaches"
  - Reason (following DL approach) on structured data (expressed as graph)

- Indeed, this realization is not new!

Battaglia, et al. Relational inductive biases, deep learning, and graph networks (2018)

Fine-grained object recognition

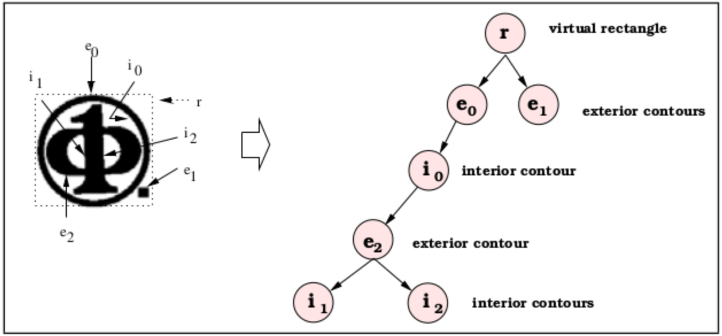
Car model (ICPR'14), Bird species (CVPR'15), ...



- Attention for discriminative parts

Recursive Neural Network (1998)

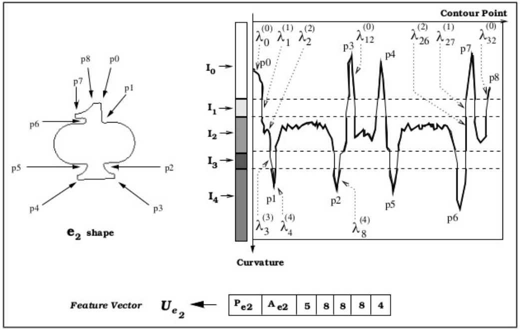
Example: Logo recognition



- Identify the external and internal contours by image processing techniques
  - Edge detection, perceptual grouping
- Create a tree structure

Recursive Neural Network

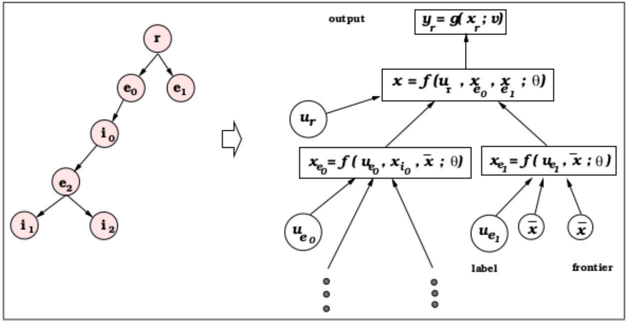
Feature representation



- Shape descriptor for each contour:
  - Perimeter, area
  - Histogram of curvatures

Recursive Neural Network

Processing model



- Functions f() and g() can be realized as deep neural networks
- Identical property descriptions u_x

Recursive Neural Network

Training

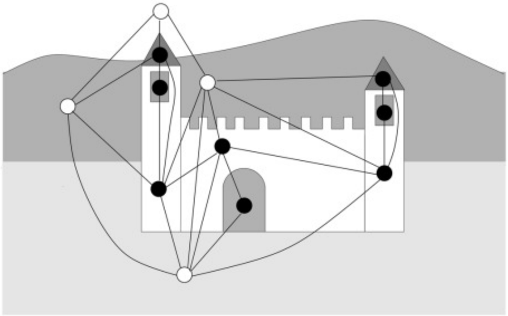
- Trained over a large logo database
  - Each logo generates a different tree
- Goal is to learn parameters for f() and g()
  - Same f() and g() for every node / tree

- Better accuracy than MLP based approach
  - Exploits structure information
  - Parameters (features / weights) are machine learned

- Computations at “Lower” nodes affect that at “higher” nodes, not vice-versa
- Only graph-level (global) inference is drawn

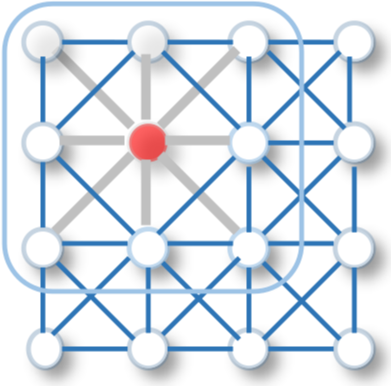
- Frasconi, et al. A General Framework for Adaptive Processing of Data Structures
- Frasesconi, et al. Logo Recognition by Recursive Neural Networks

Graph Neural Network

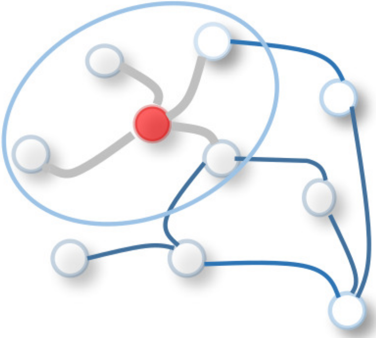


- Find super-pixels in the image
- Create a graph
  - Nodes: Super-pixels
  - Edges: Adjacent nodes
- Inferencing
  - Graph focussed: Castle
  - Node focussed: Tower, Door, Window, ..., Background

2D Convolution vs. Graph Convolution



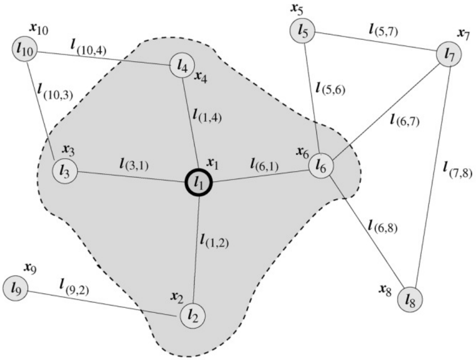
2D Convolution



Graph Convolution

Convolutional Graph Neural Network

Processing model

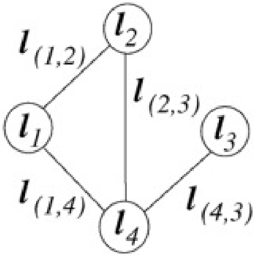


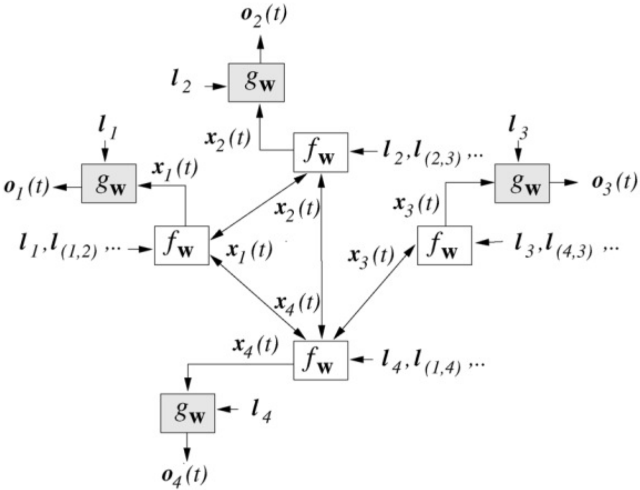
$$x_1 = f_w(l_1, \underbrace{l_{(1,2)}, l_{(3,1)}, l_{(1,4)}, l_{(6,1)}}_{l_{co[1]}}, \underbrace{x_2, x_3, x_4, x_6}_{x_{ne[1]}}, \underbrace{l_2, l_3, l_4, l_6}_{l_{ne[n]}})$$

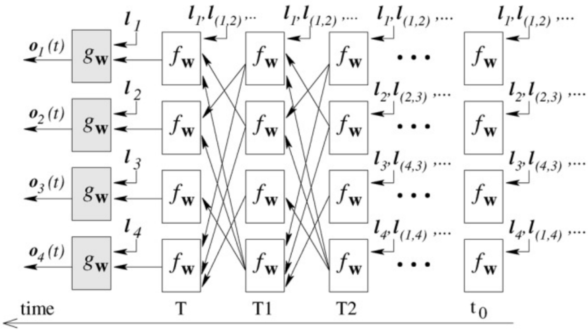
- Nodes have identical property (feature) descriptors
  - e.g. color, texture, shape
- Edges have identical property descriptors
  - e.g. distance between the center of gravities of the nodes

Graph Neural Network

Processing model (contd.)





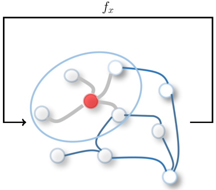


- Take the output after several recursions
  - Is the system guaranteed to go into a steady state after a finite number of iterations?

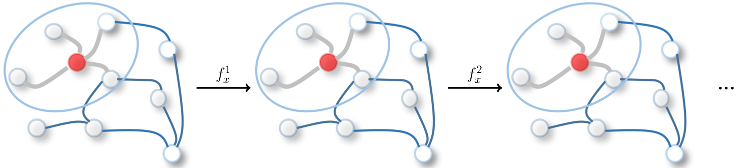
Scarselli, et al. The graph neural network model (2009)

Recursive & Convolutional Graph Neural Network

Rec-GNN & Conv-GNN



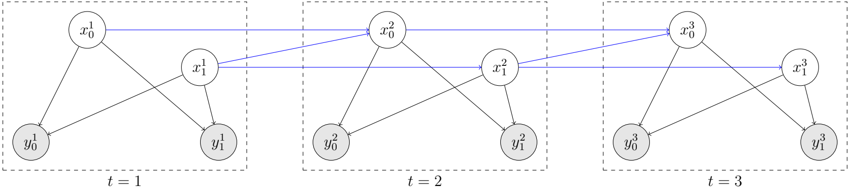
Recursive GNN



Convolutional GNN

Temporal Dependencies

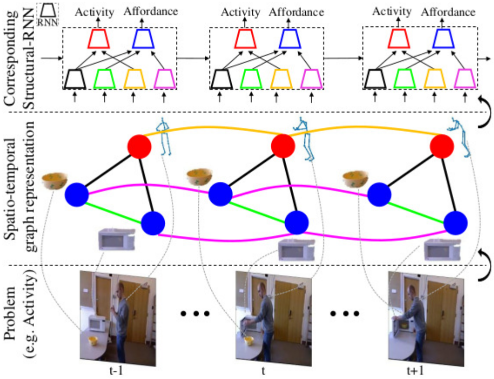
Dynamic Bayesian Network



- Network parameters: $P(X^0), P(Y^t \mid X^t), P(X^t \mid X^{t-1})$
- Prob distribution at $t = T$:
  - $P(X^T) = P(X^0). \prod_{t=1:T} P(Y^t \mid X^t). \prod_{t=1:T} P(X^t \mid X^{t-1})$

Spatio-Temporal Graph Neural Network

Structured RNN (S-RNN)



- CVPR-16 Paper, Presentation Video

Applications

- Rec-GNN / Conv-GNN (still images)
  - Fine-grained classification (e.g. bird species)
  - 3D point cloud processing (LiDAR)
- Spatio-Temporal GNN (video / motion picture)
  - Human action recognition
  - Human/Robot - Object Interaction
  - Human Motion Modeling

Limitations and Future Research

- Model depth
  - Performance (accuracy) drops with depth
- Scalability
  - Graph size needs to be limited
  - Number of nodes / number of edges
- Heterogeneity of graphs
  - Presently graphs are assumed to be homogeneous
- Dynamicity
  - Graph structure changing over time

Quiz

Your feedback for the course please

End of Module 09-01