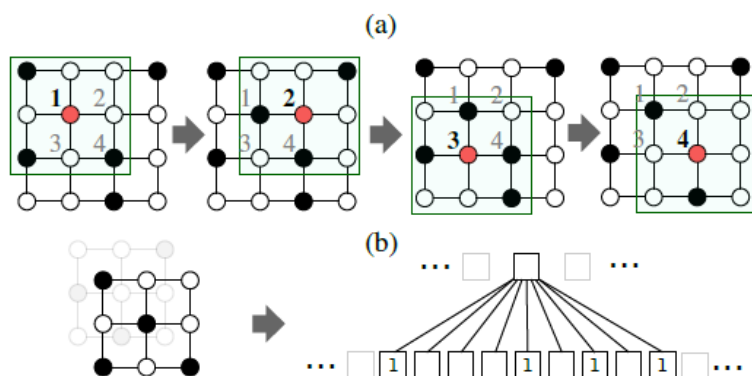


解决的问题

卷积神经网络都是应用在图像数据上，如何将它有效地应用于图类型的数据上。

对于图像数据，应用一个卷积神经网络可以看成将receptive field（图中为 3×3 ）以固定的步长将图像遍历，因为图像中像素点的排列有一定的次序，receptive field的移动顺序总是从上到下，从左到右。这也唯一地决定了receptive field对一个像素点的遍历方式以及它如何被映射到向量空间中。



然而对于图结构数据这种隐式的结构特征很多时候是缺失的，而且当给定不止一张图时，各个图之间的顶点没有必然的联系。因此，在将卷积神经网络应用在图数据上时，需要解决下面两个问题：

1. 决定邻域中顶点的产生次序
2. 计算一个将图映射到向量空间的映射方法

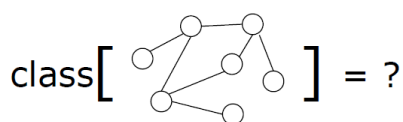
Input: Finite collection of graphs



- Nodes of any two graphs are **not** necessarily in correspondence
- Nodes and edges may have attributes (discrete and continuous)

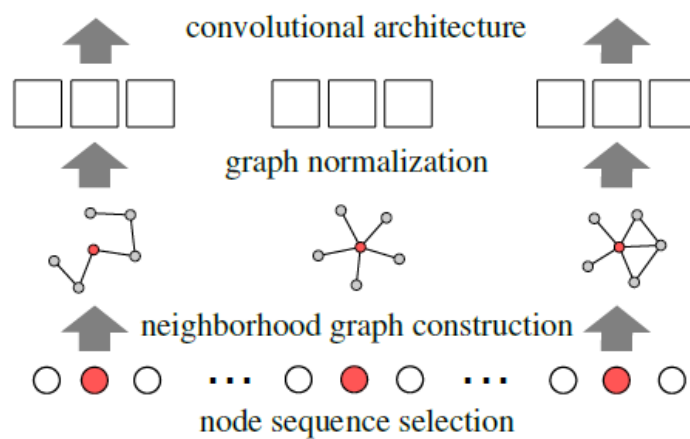
Problem: Learn a representation for classification/regression

Example: Graph classification problem



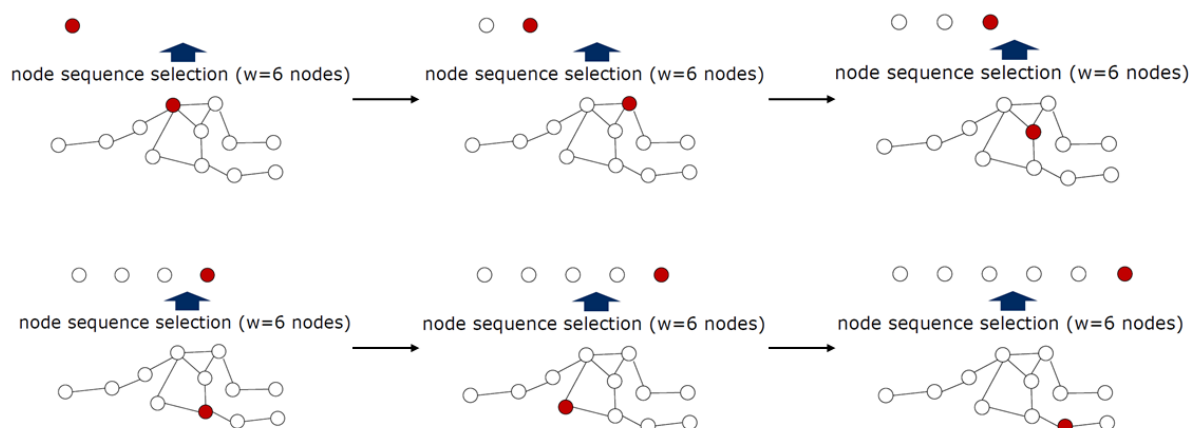
做法及创新

论文提出方法的流程如下：

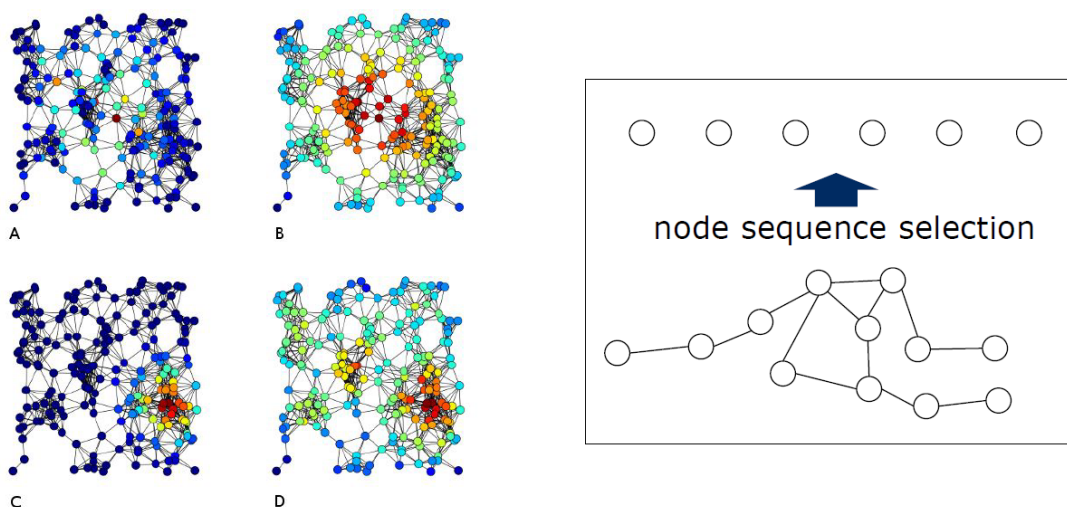


Node Sequence Selection

从图中选取固定数量 w 的顶点，它类比于图像的宽度，而选出的顶点就是卷积操作中小矩形的中心顶点。 w 就是在这个图上所做的卷积操作的个数。如下图所示， $w = 6$ ，代表需要从图中选择6个顶点做卷积操作。论文中选取顶点的方式为DFS，关键点在于图标签函数 l ，这个函数的作用是决定选取顶点的次序，可以选区的函数为between centrality与WL算法等等



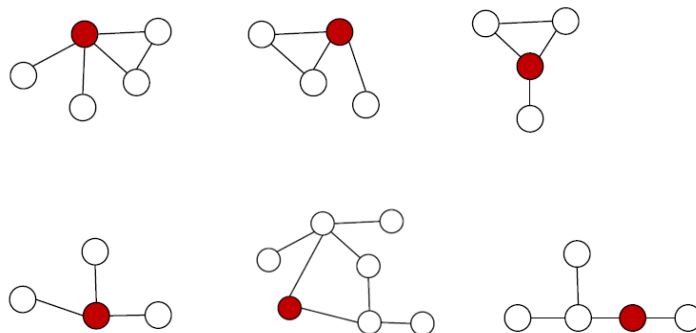
- We use centrality measures to generate the node sequences
- Nodes with similar structural roles are aligned across graphs



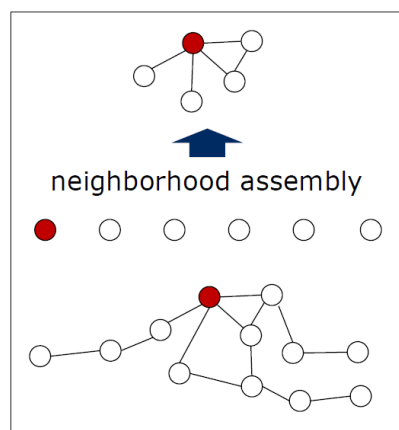
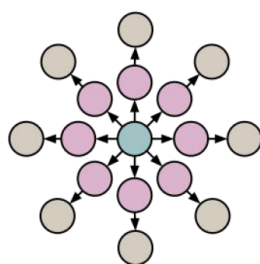
A: Betweenness centrality
B: Closeness centrality
C: Eigenvector centrality
D: Degree centrality

Neighborhood Assembly

选取完顶点后，下一步是为它们构建receptive field，类似于第一张图中的 3×3 矩阵。选取的方式为，以顶点 v 为中心，通过BFS添加领域顶点，直到满足receptive field长度 k ：

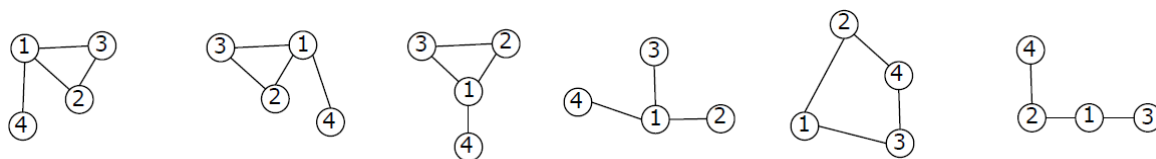


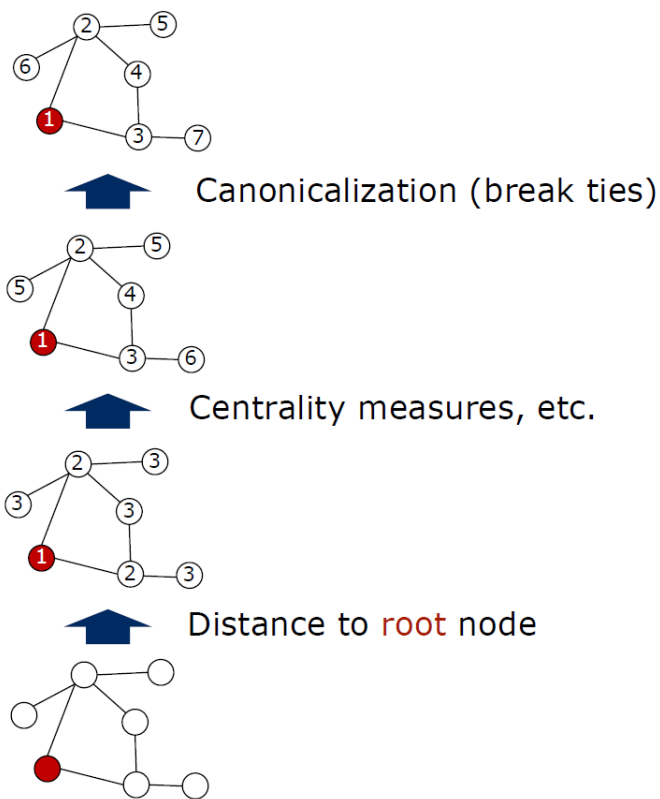
Simple breadth-first expansion until at least k nodes added, or no additional nodes to add



Graph Normalization

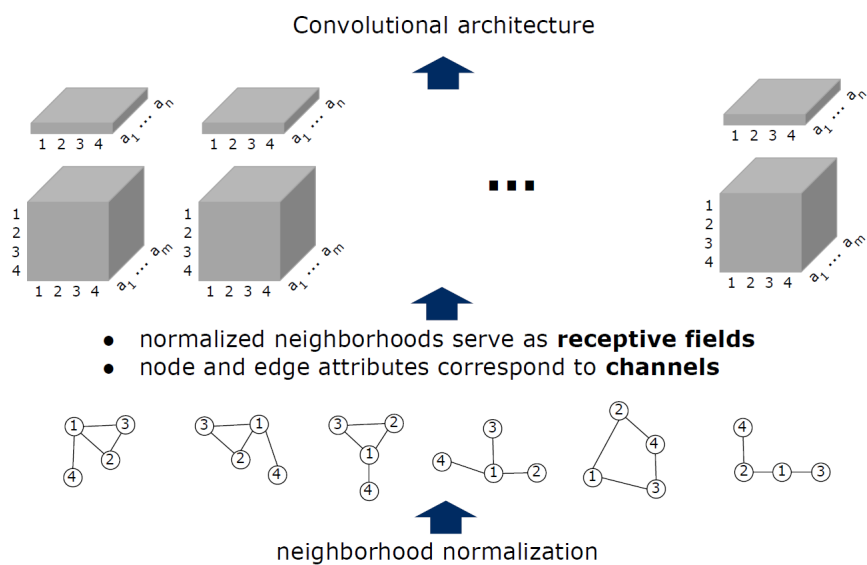
在选取了满足数量的邻域顶点后，下一步是通过图标签函数 l 为这些顶点赋予一个次序，目的在于将无序的领域映射为一个有序的向量：



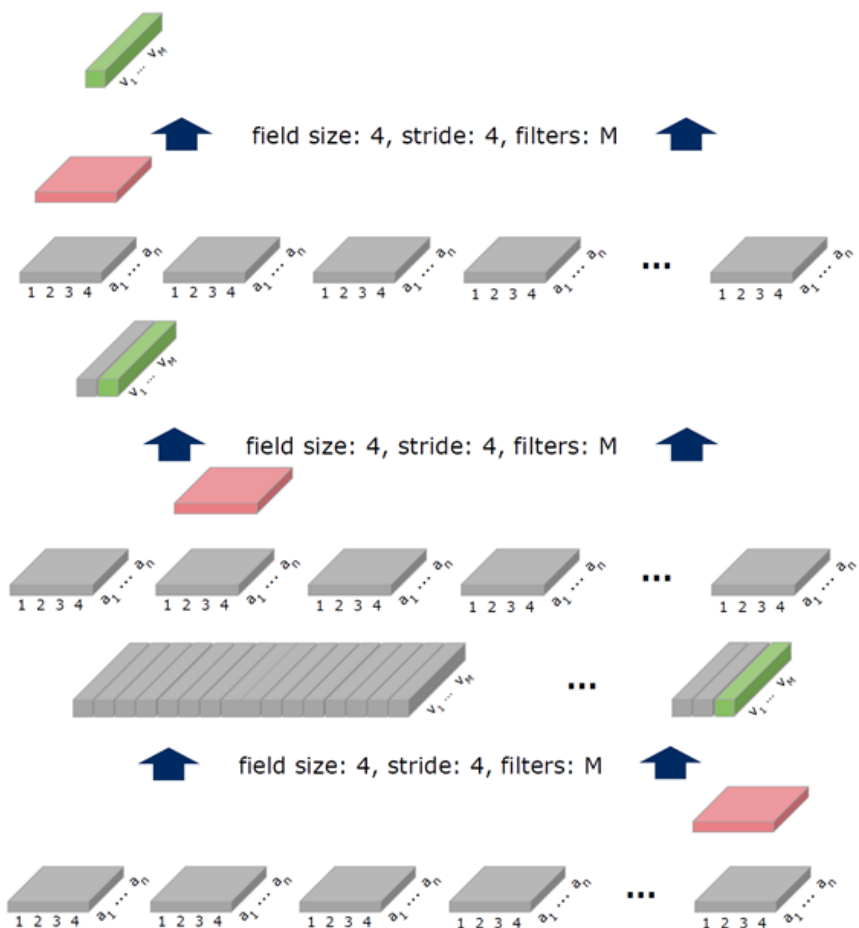


Convolutional Architecture

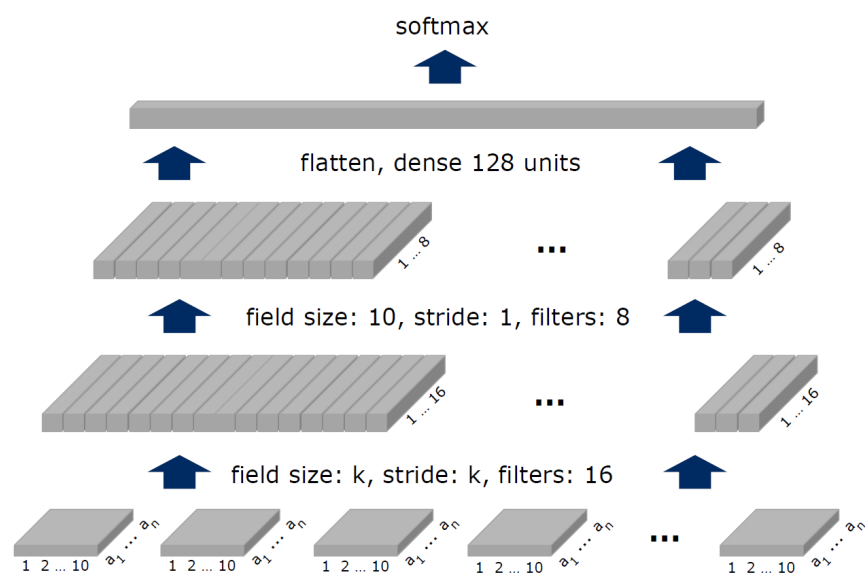
最后一步就是应用卷积层提取特征，顶点和边的属性对应于传统图像CNN中的channel:



假设顶点特征的数目为 a_v ，边的特征个数为 a_e ， w 为选取的顶点个数， k 为receptive field中的顶点个数，则对于输入的一系列图中的每一个，可以得到两个张量维度分别为 (w, k, a_v) 、 (w, k, a_e) ，可以变换为 (wk, a_v) 、 (wk^2, a_e) ，其中 a_v 与 a_e 可以看成是传统图像卷积中channel的个数，对它们做一维的卷积操作，第一个的receptive field的大小为 k ，第二个的receptive field的大小为 k^2 。



整体卷积结构:



数据集

MUTAG、PTC、NCI、D&D