

# Lecture 17 Scene Graphs and Graph Convolutions

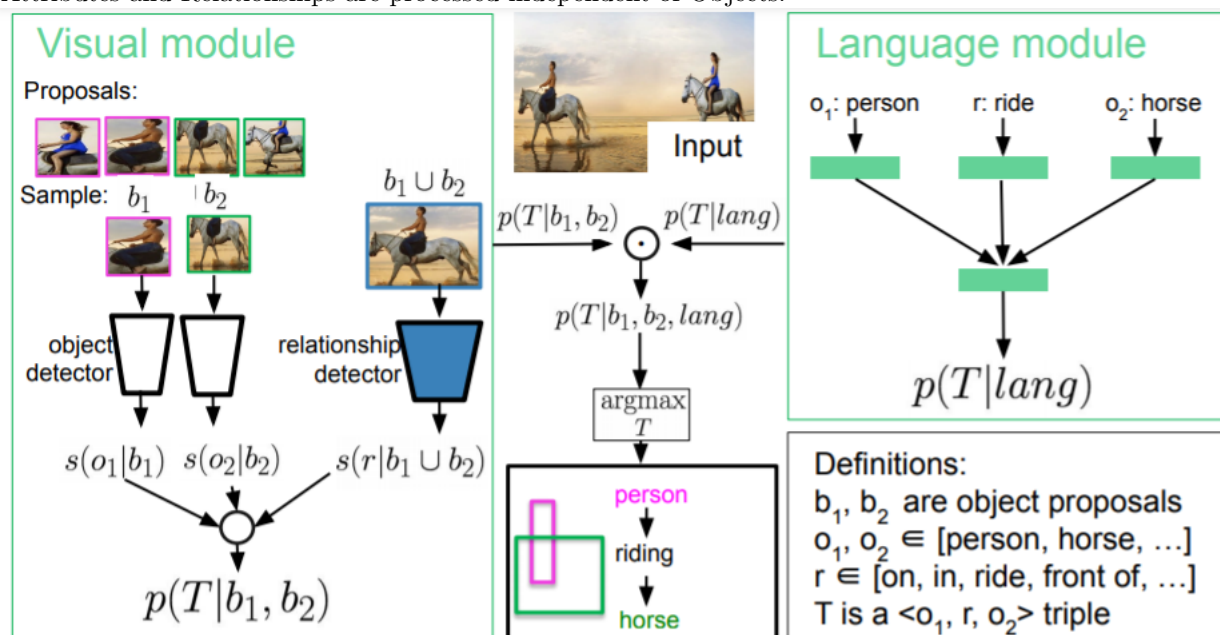
Sinkoo

March 28, 2022

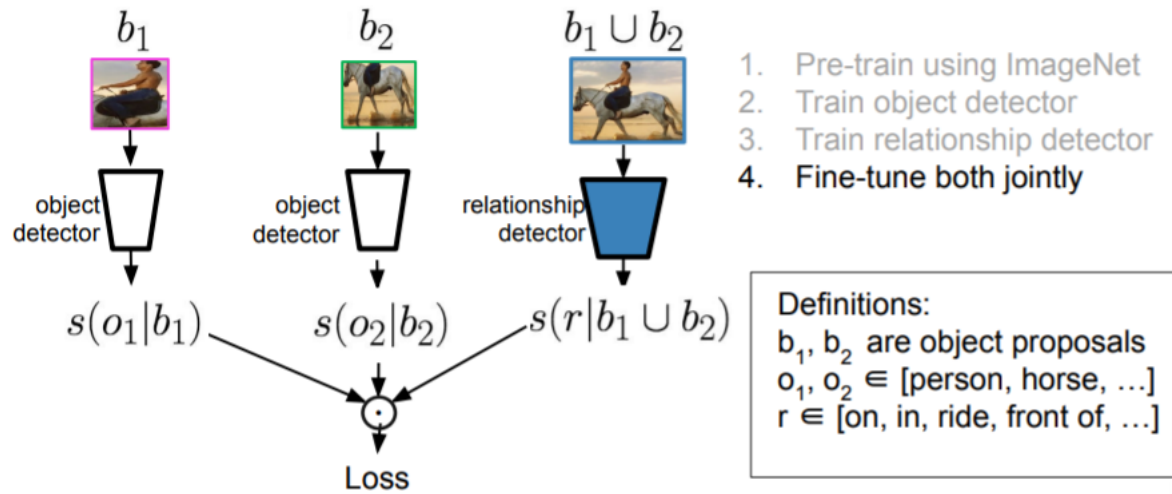
## 1 Scene Graph

Many Vision tasks share a similar underlying structure: Objects, Attributes, Relationships.

Attributes and Relationships are processed independent of Objects.



## Training the visual module

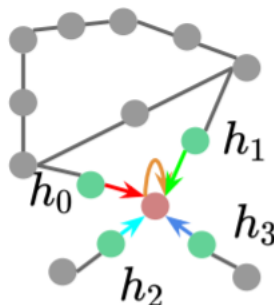


But doesn't consider other relationships when making predictions.

## 2 Generalizing 2D convolutions to Graph Convolutions

Take relationships as Nodes in Scene Graph.

### Generalizing 2D convolutions to Graph Convolutions



- Graph convolutions involve similar **local operations** on nodes.
- Nodes are now object representations and not activations
- The **ordering of neighbors** should not matter.
- The **number of neighbors** should not matter.
- $N(i)$  are the neighbors of node  $i$
- $c_{ij}$  is a normalization constant

$$h_4^{l+1} = W_4 h_4^l + W_0 h_0^l + W_1 h_1^l + W_2 h_2^l + W_3 h_3^l$$

$$h_i^{l+1} = W_{self} h_i^l + \sum_{j \in N(i)} \frac{1}{c_{ij}} W_{other} h_j^l$$

$$h_i^{l+1} = W h_i^l + \sum_{j \in N(i)} \frac{1}{c_{ij}} W h_j^l$$

Increase receptive fields with increasing depth  
Implement Attention

## 3 GCN

提取图的特征:两种层:

1,node level: 不改变图的结构, 只改变图节点的值;

2,graph level:通过某种形式的pooling改变图的结构。

每一个隐藏层可以写成函数:  $H^{(l+1)} = f(H^{(l)}, A)$  论文中采用的是

$$f(H^{(l)}, A) = \sigma(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H^{(l)} W^{(l)})$$

$A$ 是邻接矩阵加单位矩阵, 用于保存自己节点的特征;  $D$ 对 $A$ 做一个对称的归一化。

具体推导见: <https://zhuanlan.zhihu.com/p/107162772>