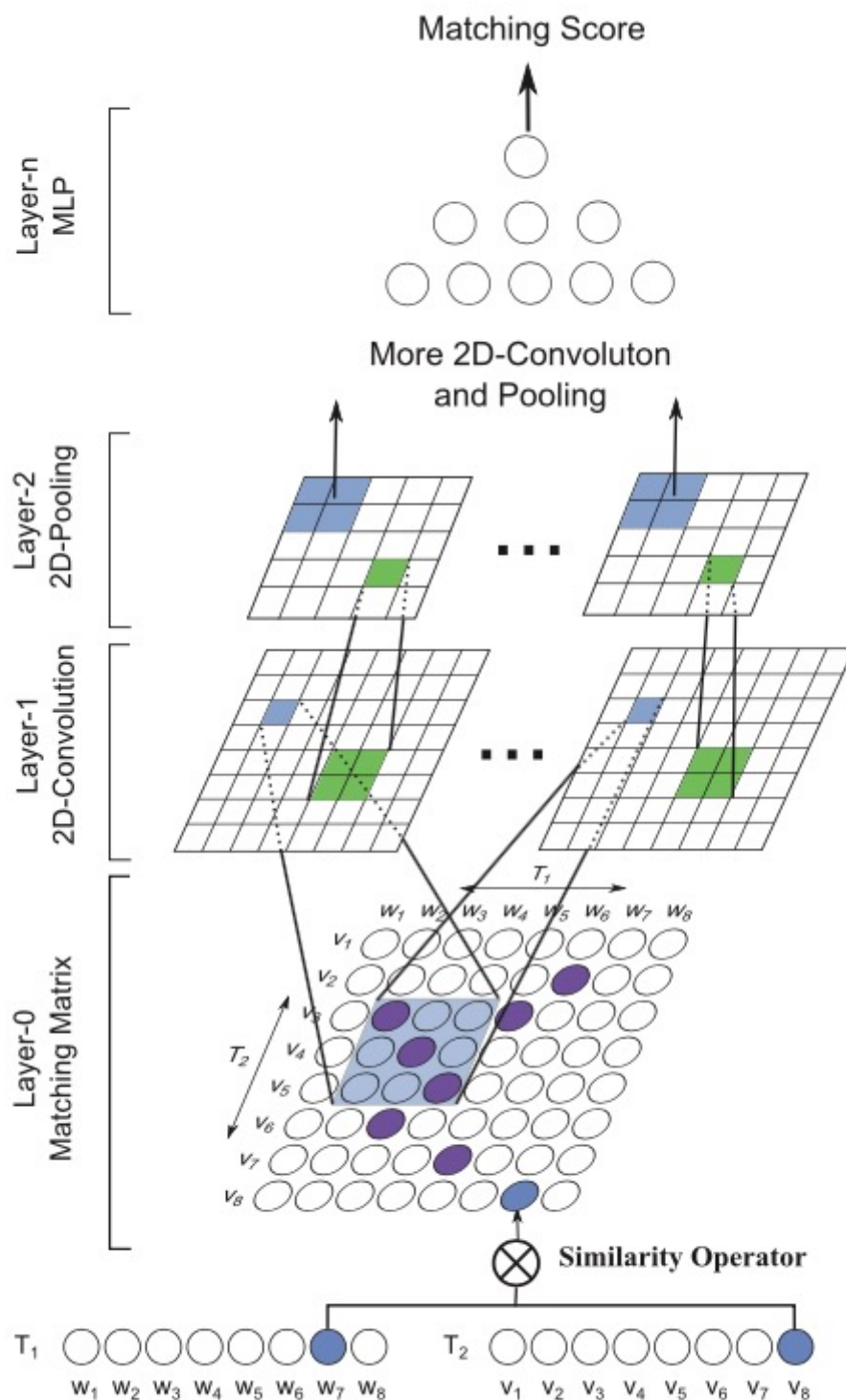


# 2016 MatchPyramid

- 题目：Text Matching as Image Recognition
- 原文：<https://arxiv.org/abs/1602.06359>
- 核心做法：构建矩阵表示文本的相似性，使用卷积神经网络提取相似性矩阵的特征。
- 模型图



## 思想

1.受CNN在图像识别中的启发（可以提取到边、角等特征），作者提出先将文本使用相似度计算构造相似度矩阵，然后卷积来提取特征。把文本匹配处理成图像识别。【想法很特别】

2.根据结果显示，在文本方面使用作者提出的方法可以提取n-gram、n-term特征。作者给了一个匹配的例子：

T1 : Down the ages noodles and dumplings were famous Chinese food.

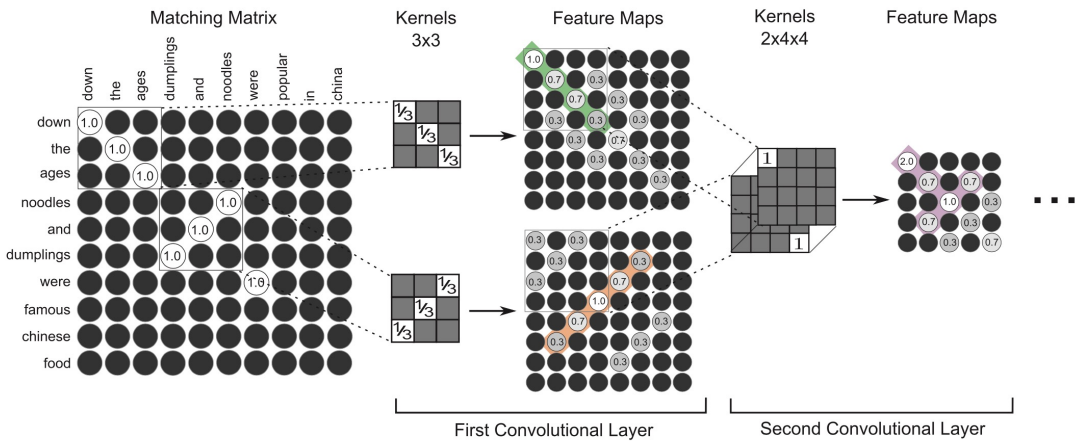
T2 : Down the ages dumplings and noodles were popular in China.

模型可以学习到**Down the ages**（n-gram特征），**noodles and dumplings**与**dumplings and noodles**（打乱顺序的n-term特征）、**were famous Chinese food**和**were popular in China**（相似语义的n-term特征）。

## 实现细节

1. Xi和Yi的相似度计算方式：**完全一样 (Indicator)**，**余弦相似度 (Cosine)**，**点乘 (Dot Product)**。

2. 卷积，ReLU激活，动态pooling（pooling size等于内容大小除以kernel大小）。



3. 通过卷积和池化得到最终的表示向量 $\mathbf{z}$ ，Then，最终结果（匹配或者不匹配）分别表示为 $s_0$ 和 $s_1$ ， $s_0$ 和 $s_1$ 的表示为 $\mathbf{z}$ 经过MLP的结果，这是一个二维向量，分别代表了 $s_0$ 和 $s_1$ 的matching\_score.

$$(s_0, s_1)^T = \mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{z} + \mathbf{b}_1) + \mathbf{b}_2, \quad (10)$$

4. 将上面的结果输入softmax，得到分类的概率，再用交叉熵计算损失。

$$loss = - \sum_{i=1}^N \left[ y^{(i)} \log(p_1^{(i)}) + (1 - y^{(i)}) \log(p_0^{(i)}) \right], \quad (11)$$

$$p_k = \frac{e^{s_k}}{e^{s_0} + e^{s_1}}, \quad k = 0, 1,$$