## RCNN

### Architecture：

图示

描述已自动生成

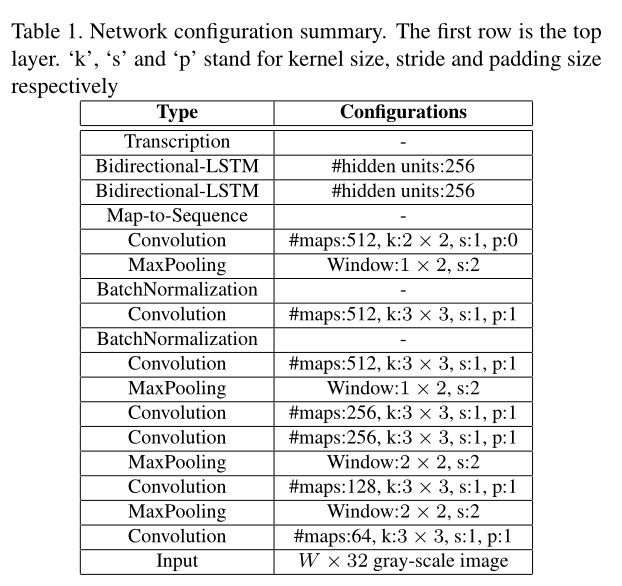
1.forward LSTM + backward LSTM = bidirectional LSTM，充分利用上下文特征

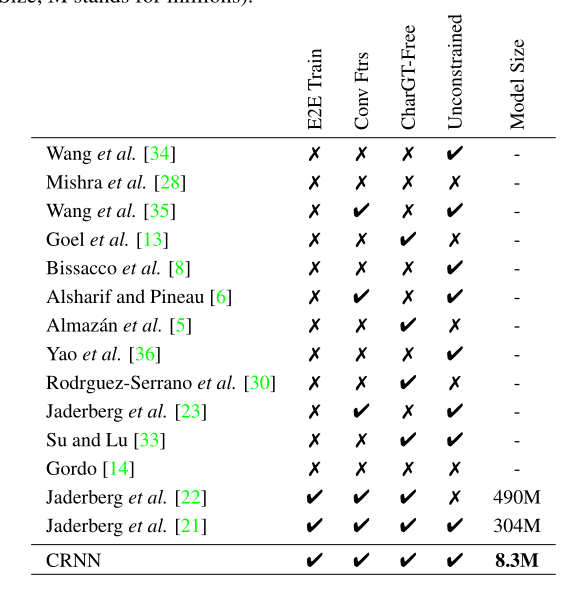
2.由于RNN可以反向传播，因此他们可以unify为一个网络进行训练

3.LSTM相对于RNN，不会有梯度消失的问题，从而可以处理长序列

4.Map-to-Sequence，用于连接CNN和RNN

5.CNN使用VGG-VeryDeep（K. Simonyan and A. Zisserman. Very deep convolu-tional networks for large-scale image recognition. CoRR,abs/1409.1556, 2014.）并为了识别英文做了一定调整，将中间的一个卷积层换成了pooling层，并且是1\*2，目的是为了获取狭长的感受野，更适合识别英文





8.3M为参数量，实际大小应X4

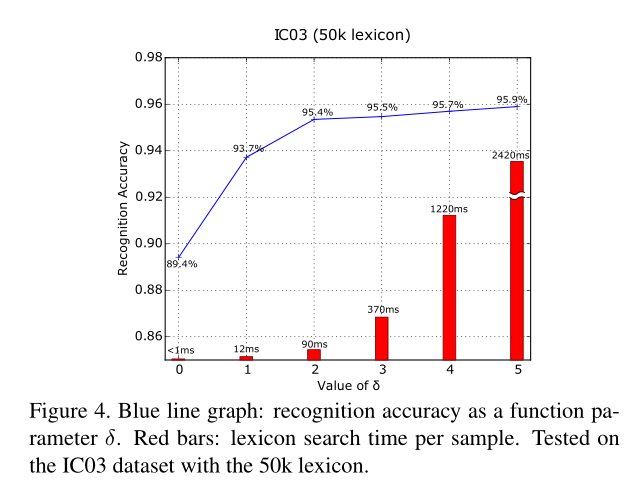
### Train&Evaluation：

1. 训练数据：输入图像以及对应的gt label sequence
2. loss = 图片包含 文本

   描述已自动生成
3. 使用SGD训练，反向传播梯度，转录层使用CTC中提到的方法进行反向传播，在LSTM层，使用BPTT方法
4. 使用ADADELTA（M.D.Zeiler. ADADELTA:anadaptivelearningratemethod.CoRR, abs/）自动计算每层的学习率，相比动量法，ADADELTA不需要手动设置学习率并且学习速度更快

### Skill&Knowledge：

1.CTC使用的N的实验评估，根据这个结果，通常选择2即可达到较好的效果



2.提出观点：减少使用FC层，可以使模型更紧凑和高效

### Questions:

是否可以用RCNN处理图片，进行分类？

map to sequence?

### Improve&Validation：

1.3\*3卷积换成1\*3+3\*1

2.使用VGG，是否可以用前面的方向检测迁移？

## Extension

### Compare:

2022所在公司的CRNN：Resnet+2 \* Bidirectional-LSTM+FC，后处理CTCDecode，模型200M+，远大于论文中的模型

### DCNN：

### LSTM：

S. Hochreiter and J. Schmidhuber. Long short-term memory.

Neural Computation

### CTC：

A. Graves, S. Fernández, F. J. Gomez, and J. Schmidhu-

ber. Connectionist temporal classification: labelling unseg-

mented sequence data with recurrent neural networks. In

ICML, 2006