

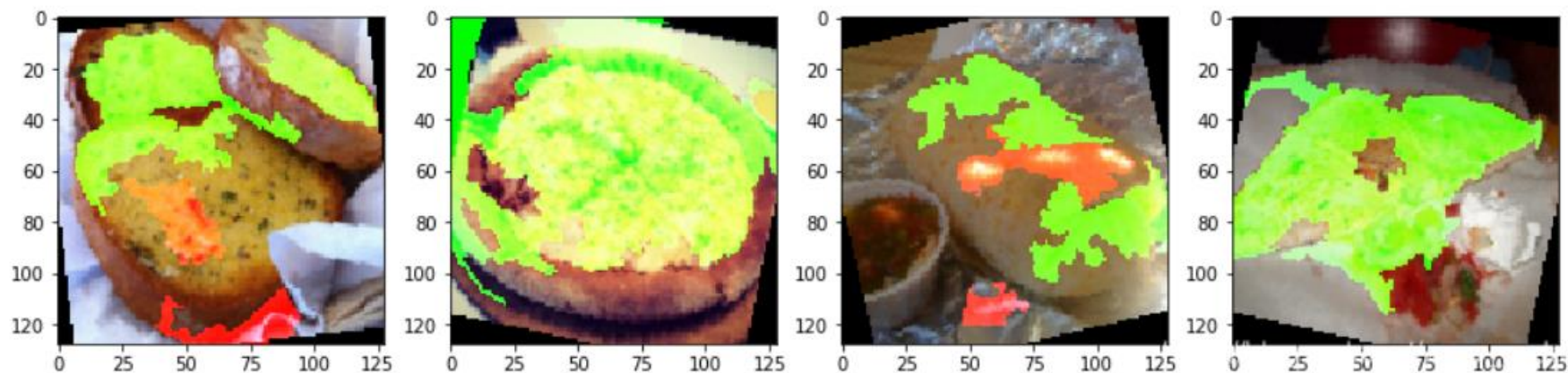
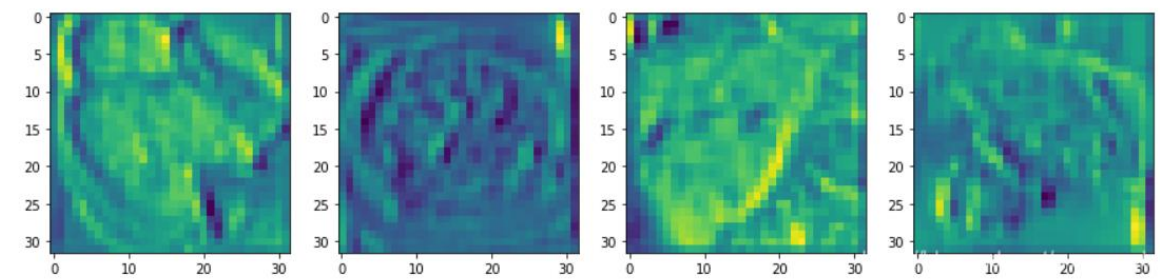
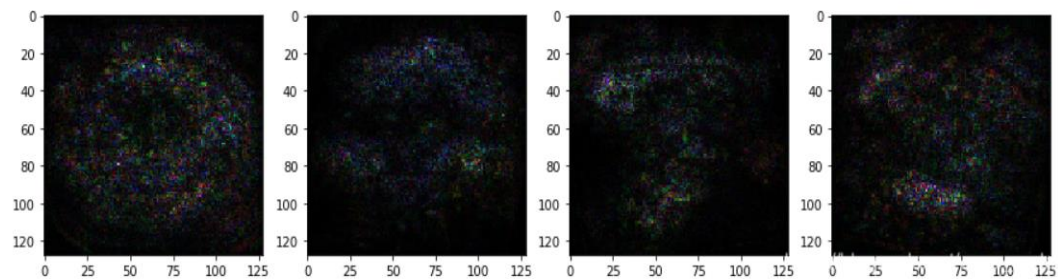
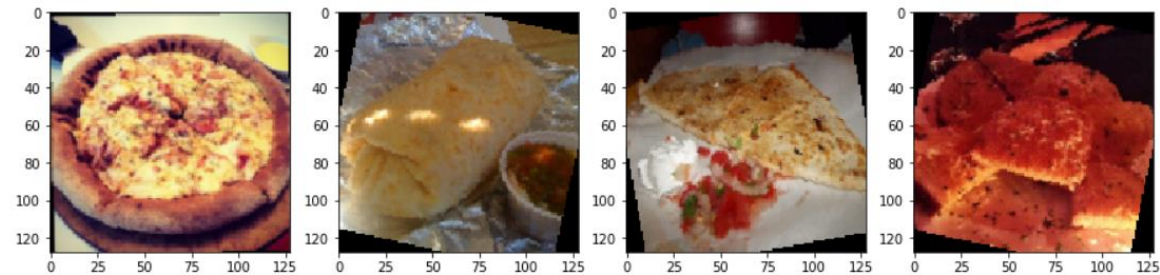
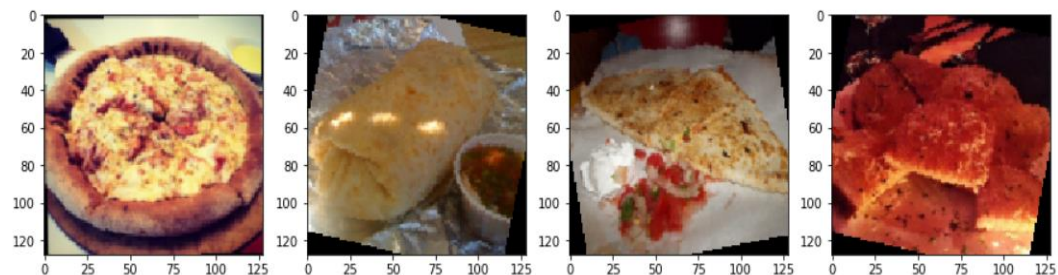
总结

黄飞虎

2021.1.6

学习情况

看了李宏毅的视频， 完成了实验五



论文情况

A Joint Model of Intent Determination and Slot Filling for Spoken Language Understanding

下载链接: <https://www.ijcai.org/Proceedings/16/Papers/425.pdf>

本文的工作

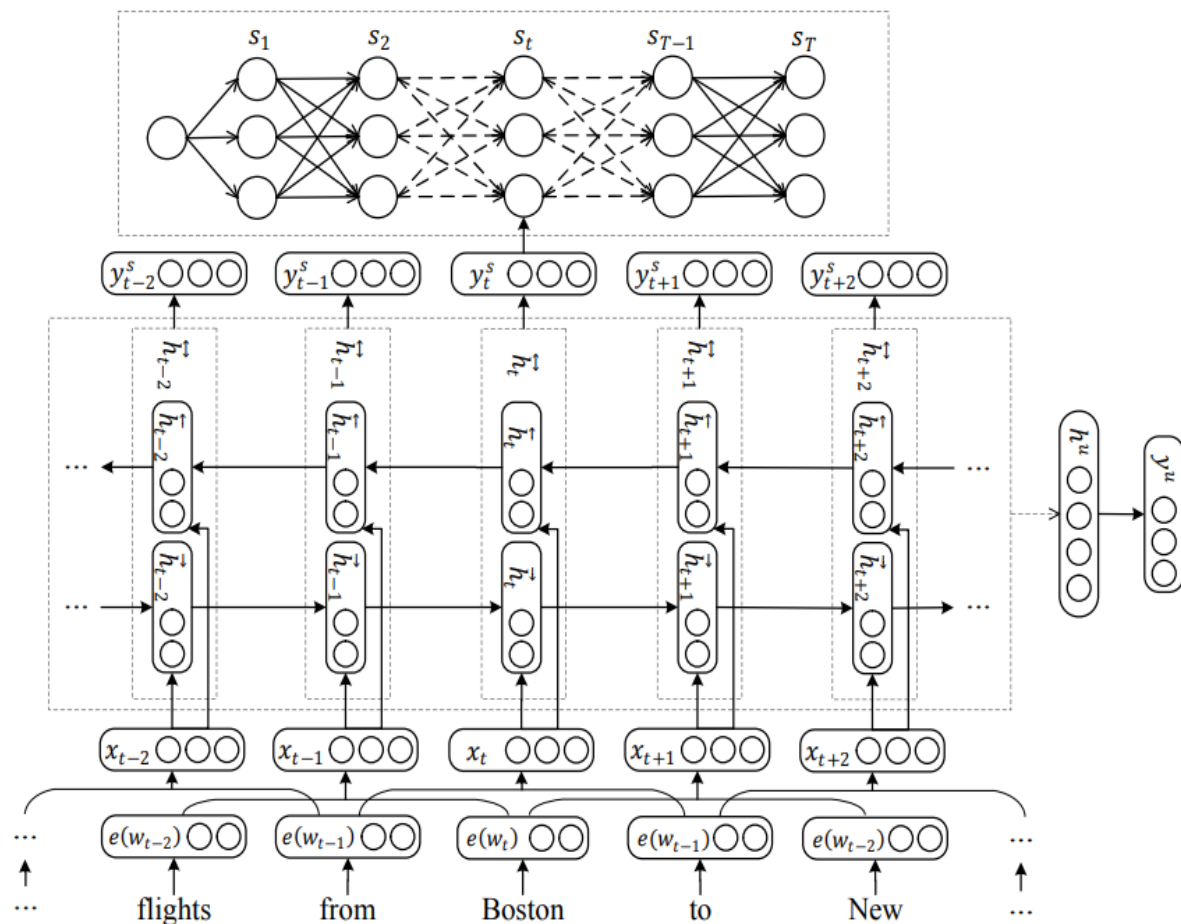
在SLU中的两个主要任务是意图识别 (Intent determination, 即文本分类) 和槽填充 (slot filling, 即文本标记)。本文提出了一个同时解决两个问题的joint model。

本文在两个数据集上进行了实验，在两个数据集上，两个任务都达到了state-of-art的效果。

作者认为联合模式值得一试，原因有二。首先，在SLU系统中，这两个任务通常都是必需的。其次，将一个任务的信息应用到另一个任务中，相互促进，共同预测。

本文采用GRU来学习序列的每个time step的语义表达，一方面这些语义表达会用预测slot的标签，另一方面这些也会进行意图的分类。本文提出的joint model在两个数据集上，在ID和SF任务上都达到了state-of-the-art的效果。

模型结构



1.Embeddedings: 本文采用词窗口作为RNN的输入, 命名实体是SLU的一种重要特性。为了利用这些特性, 原文赋予每个命名实体一个置于[0, 1]的一个embedding, 维度和词向量一样, 并在训练过程中进行微调, 上下文命名实体窗口的定义类似于词窗口。

2.Recurrent Hidden Layers: 本文使用双向GRU来获得文本的表征。

3.Task Special Layers: 双向隐层状态被两个任务所共用。一方面, h 在每个时序捕捉到的特征可以直接用来预测插槽的标签Label, 另一方面, 我们用max-pooling层来得到整个句子的表示。

4. Training:即目标函数的定义

实验

数据集：ATIS dataset（ATIS是SLU领域中使用最广的数据集。）
和CQUD dataset（CQUD是百度知道搜集到的问题，是中文社区最受欢迎的Question Understanding Dataset.）

评价指标：将准确性作为ID任务的评估指标。ATIS中的一些语句有多个意图标签。在之前的研究中，如果预测到任何ground truth标签，就把一个话语算作一个正确的分类。

实验结果

与之前的模型对比

实验结果展示如下，第二列是每个方法使用的特征，其中，W，N和S分别代表单词、命名实体和语义特征，在CQUD数据集中，W代表每一个中文字特征表示。

Model	Features	ATIS		CQUD	
		Intent	Slot	Intent	Slot
SVM [Raymond and Riccardi, 2007]	W	—	89.76	—	81.32
CRF [Mesnil <i>et al.</i> , 2015]	W	—	92.94	—	83.40
CRF [Mesnil <i>et al.</i> , 2015]	W+N	—	95.16	—	—
RNN [Mesnil <i>et al.</i> , 2015]	W	—	95.06	—	85.63
RNN [Mesnil <i>et al.</i> , 2015]	W+N	—	96.24	—	—
R-CRF [Yao <i>et al.</i> , 2014]	W	—	—	—	85.88
R-CRF [Yao <i>et al.</i> , 2014]	W+N	—	96.46	—	—
Boosting [Tur <i>et al.</i> , 2010]	W	95.50	—	93.54	—
Sentence simplification [Tur <i>et al.</i> , 2011]	W+S	96.98	95.00	94.46	—
RecNN [Guo <i>et al.</i> , 2014]	W+S	95.40	93.22	—	—
RecNN+Viterbi [Guo <i>et al.</i> , 2014]	W+S	95.40	93.96	—	—
Our model	W	98.10	95.49	96.05	87.12
Our model	W+N	98.32	96.89	—	—

总结

联合模型与分开模型比较

介绍了用于联合意图确定和空位填充的递归神经网络，这是语音理解中的两个主要任务。双向GRU用于学习两个任务共享的序列表示。通过共享表示的最大池化来获取全局表示，以预测意图标签。时隙的标签由共享表示预测，并在序列级别进一步推断。通过统一的损失函数和共享表示，学习了两个任务的相关性，从而促进了彼此的相互促进。

Model	ATIS		CQUD	
	Intent	Slot	Intent	Slot
ID only	97.53	—	95.34	—
SF only	—	95.14	—	85.78
Pipeline	97.53	95.41	95.34	86.96
Joint (equal)	98.10	95.49	96.05	87.12
Joint (ID oriented)	98.10	95.49	96.35	86.63
Joint (SF oriented)	97.87	95.61	95.93	87.23

Thank you