

## 每日小结

	周一	周二	周三	周四	周五
早	跑 attention, 上课	跑 attention	跑模型, 上课	上课	跑郑州信令
中	跑 attention, 上课	论文阅读	论文阅读	跑代码, 上课	跑郑州信令
晚	上课	上课		信令预处理	跑 attention

注：简单表述当前时间段工作，如看文献1，整理数据等

## 科研详情

## 文献阅读

## 文献1

题目：A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction

作者：Yao Qin1\*, Dongjin Song2, Haifeng Chen2, Wei Cheng2, Guofei Jiang2, Garrison W. Cottrell1

出处：Arxiv 2017

## 方法：

该论文主要通过编码层和解码层都引入双向注意力机制实现时间序列更好的预测。

本文提出了一种新的基于双阶段注意的递归神经网络 (DA-RNN)，它由一个具有输入注意机制的编码器和一个具有时间注意机制的解码器组成。新引入的输入注意机制可以自适应地选择相关的控制序列。时间注意机制可以自然地捕获编码输入的长期时间信息。基于这两种注意机制，LSTM 滤波器不仅可以自适应地选择最相关的输入特征，还可以适当地捕捉时间序列的长期时间相关性。

分为两个阶段：

1. 第一阶段：使用注意力机制从而能够自适应提取每个时刻的特征
2. 第二阶段：使用注意力机制选取与之相关的 encoder hidden states

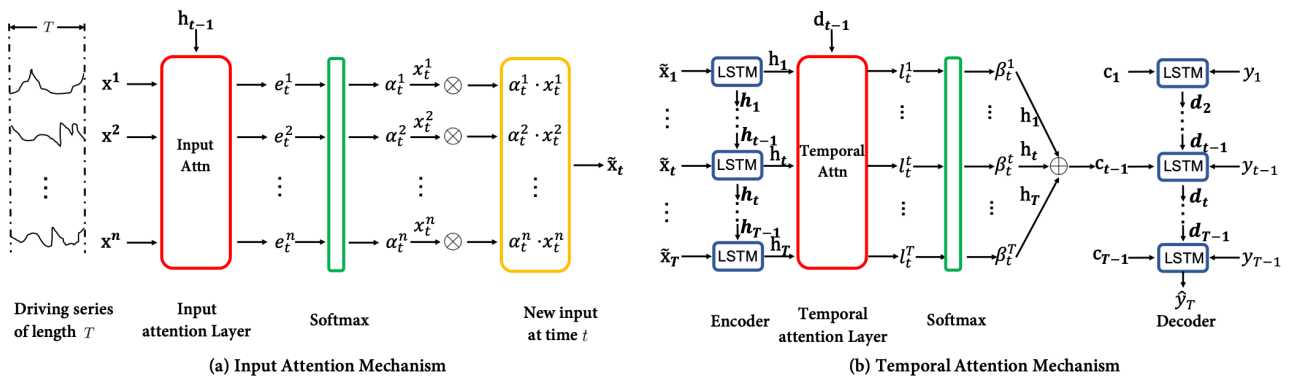
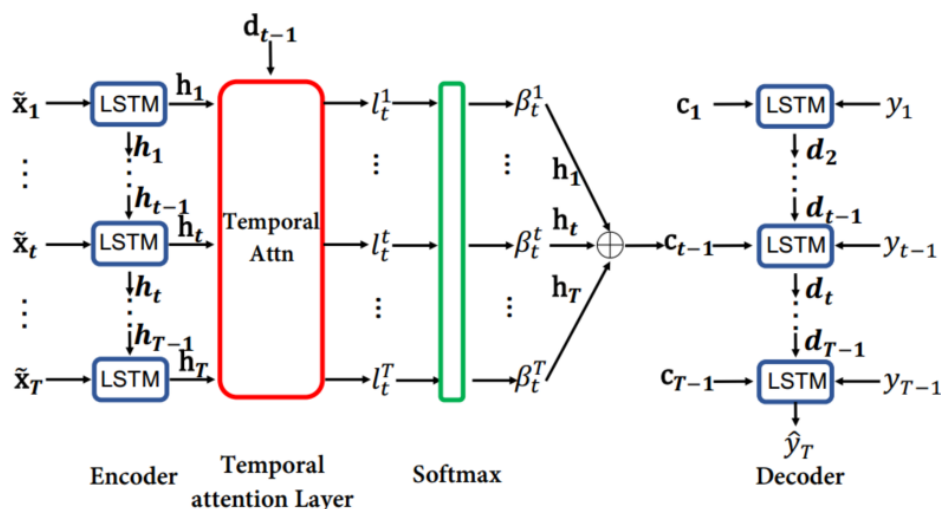


Figure 1: Graphical illustration of the dual-stage attention-based recurrent neural network. (a) The input attention mechanism computes the attention weights  $\alpha_t^k$  for multiple driving series  $\{x^1, x^2, \dots, x^n\}$  conditioned on the previous hidden state  $h_{t-1}$  in the encoder and then feeds the newly computed  $\tilde{x}_t = (\alpha_t^1 x_t^1, \alpha_t^2 x_t^2, \dots, \alpha_t^n x_t^n)^\top$  into the encoder LSTM unit. (b) The temporal attention system computes the attention weights  $\beta_t^k$  based on the previous decoder hidden state  $d_{t-1}$  and represents the input information as a weighted sum of the encoder hidden states across all the time steps. The generated context vector  $c_t$  is then used as an input to the decoder LSTM unit. The output  $\hat{y}_T$  of the last decoder LSTM unit is the predicted result.

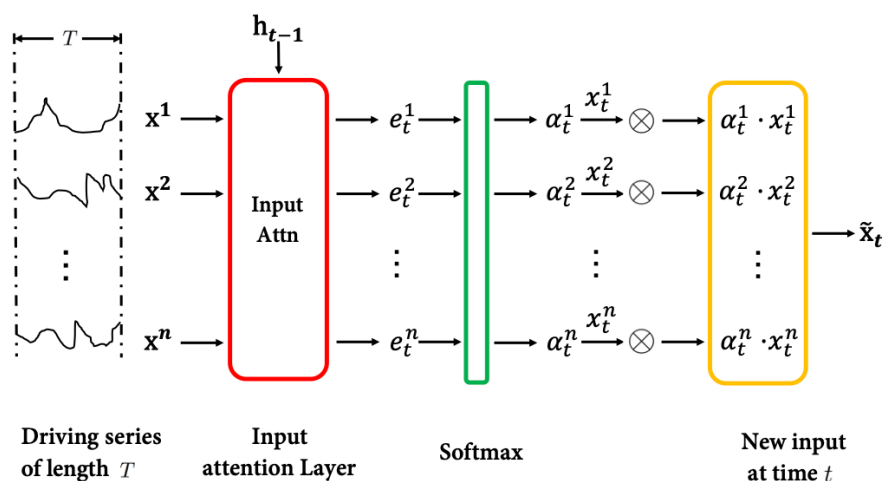
## 启发：

1. 第二阶段temporal Attention机制实现可以借鉴：



(b) Temporal Attention Mechanism

2.第一阶段：使用注意力机制从而能够自适应提取每个时刻的特征, input attention 机制，使得编码器能够关注其中输入特征中重要的特征，而不是对所有特征一视同仁，值得借鉴，但是还需要考虑我所使用的 LSTM、DPN 网络的结构。



(a) Input Attention Mechanism

## 文献2

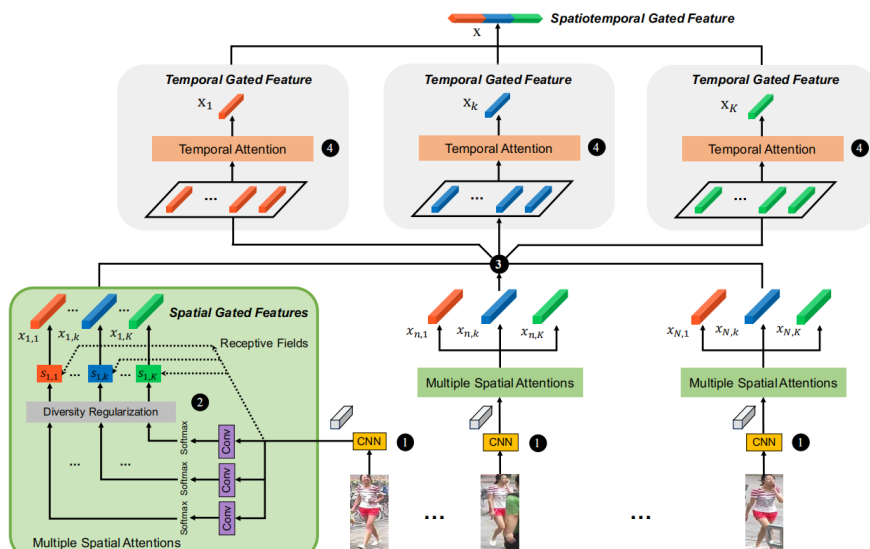
题目：Diversity regularized spatiotemporal attention for video-based person re-identification

作者：Shuang Li, Slawomir Bak, Peter Carr, Xiaogang Wang

出处：Arxiv 2018

## 方法：

对于一个视频序列，首先通过一个有限的采样策略来选择 N 个视频帧。然后将视频帧集合每张图经过预训练的 ResNet50 得到  $N \times 2048 \times 8 \times 4$  的特征图。将特征图送入多区域空间注意力模块得到多个关注不同部位的特征集。然后将不同视频帧同一个部位的特征经过时域注意力模块在时域上进行聚合得到一个部位的一个向量表示。多个部位得到的多个向量连接后经过一个 fc 降维后作为视频序列的特征，训练阶段使用 OIM loss 来训练。



启发:

1. 本文利用一组不同的空间注意模型在多幅图像中一致地提取相似的局部补丁，这种方法自动解决了跨帧对齐对应的图像块（由于身体姿势、相对于相机的方向等改变）并确定身体的特定部分是否被遮挡。
2. 主要学习了一下本文 **spatiotemporal attention** 的使用细节和代码。

## 工作进展

- 1: 阅读文献;
- 2: URFC 实验结果: 加 Spatial Attention

### 加 Spatial & Temporal Attention 在 DPN 前面和后面

Transformer  
DPN26

Temporal-spatial Attention (TSA)

Spatial Temporal Transformer

DPN前加 Spatial Attention :

准确度:	0.592
loss:	1.247
F1指数:	0.498

DPN后加 Spatial Attention:

前几个epoch更差了, 没跑完

不加 Spatial Attention:

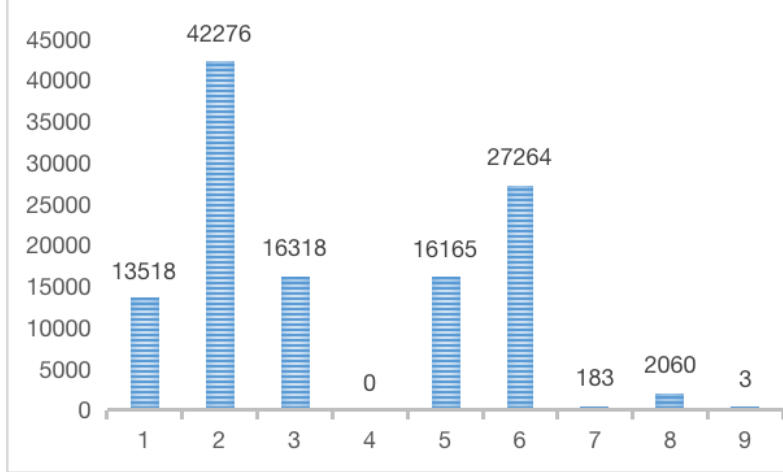
transformer+dpn26(vit b 32)	0.68	0.590
transformer+dpn26(vit b 16)	0.689	0.598

- Temporal Attention 代码维度还没搞懂, 在学
- 在输入去多头注意力的时候增加了一个维度: (batchsize, 26, 7, 24) → (batchsize, 1, 26, 7, 24) → (batchsize\*26, 1, 7, 24)
- (batchsize, t, h, w) → (batchsize, 1, t, h, w) → (batchsize\*t, 1, h, w)

3: 郑州信令数据预处理并输入训练好的Transformer+Bilstm网络得到结果。

预处理: 拼接CSV → 按照区域id转npy文件 → 按照周一到周天的顺序重拼 → 增加一个周三 ✓

郑州数据分类结果直方图



Category label	1	2	3	4	5	6	7	8	9
Function type	Residential	<u>school</u>	industrial	railway station	<u>airport</u>	park	business	Government	Hospital

下周计划

- 1: 觉得 Spatial Attention 代码可能有问题，还在查。
- 2: 寻找 DPN+attention 论文
- 3: Temporal Attention 代码跑通