

每日小结

	周一	周二	周三	周四	周五
早	学 graph	看代码，学 graph	建立 graph 代码	Graph 代码	Graph 代码
中	Graph 代码	学 graph	处理用户访问数据	Graph 代码	Graph 代码
晚	Graph 代码	组会		Graph 代码	处理用户访问数据

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： Social-STGCNN: A Social Spatio-Temporal Graph Convolutional Neural Network for Human Trajectory Prediction

作者： Abdualлах Mohamed1, Kun Qian, Mohamed Elhoseiny2,3, **, Christian Claudell1, 出处： CVPR 2020

方法：

设计了 Social-STGCNN 来克服上述两个限制。首先，从一开始就将行人的轨迹建模为时空图以替换聚集层。图边缘为行人之间的社交互动建模。我们提出了一个加权邻接矩阵，其中核函数定量地测量了行人之间的影响。之后，为了解决与递归单元相关的问题，模型使用图卷积神经网络（GCN）和时间卷积（TCN）对时空图进行处理。这使模型可以一次预测整个序列。

Social-STGCNN 模型由两个主要部分组成：时空图卷积神经网络（ST-GCNN）和时间外推器卷积神经网络（TXP-CNN）。ST-GCNN 对行人轨迹的图形表示进行时空卷积运算以提取特征。这些特征是观察到的行人轨迹历史的紧凑表示。TXP-CNN 将这些功能用作输入，并预测整个行人的未来轨迹。TXP-CNN 有望通过卷积运算来推断未来的轨迹。图 2 展示了该模型的概述。

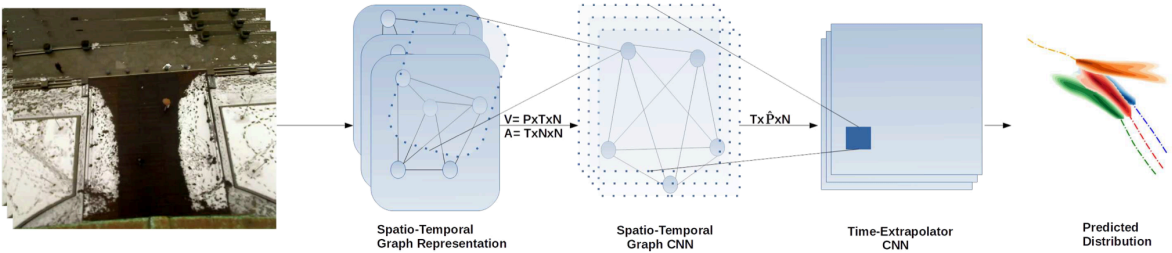


Figure 2. The Social-STGCNN Model. Given T frames, we construct the spatio-temporal graph representing $G = (V, A)$. Then G is forwarded through the Spatio-Temporal Graph Convolutional Neural Networks (ST-GCNNs) creating a spatio-temporal embedding. Following this, the TXP-CNNs predicts future trajectories. P is the dimension of pedestrian position, N is the number of pedestrians, T is the number of time steps and \hat{P} is the dimensions of the embedding coming from ST-GCNN.

启发：

1.在看行人轨迹见时空图的代码，有开源代码

文献2

题目: NodeFormer: A Scalable Graph Structure Learning Transformer for Node Classification

作者: Qitian Wu, Wentao Zhao, Zenan Li, David Wi, Junchi Yan

出处: NeurIPS22 spotlight

方法:

本文首次对大图节点级任务设计了一种 graph Transformer, 称为 NodeFormer, 它对节点数目具有线性复杂度, 可以成功扩展到大规模图上 (例如百万级节点的图) 并取得极具竞争力的性能。和主流 GNN 相比, NodeFormer 在多个不同种类的数据集 (例如同质图, 异质图, 长距离依赖图, 以及没有输入图的图片/文本分类) 都能取得更好的分类精度, 并且显存占用控制在一定范围内。

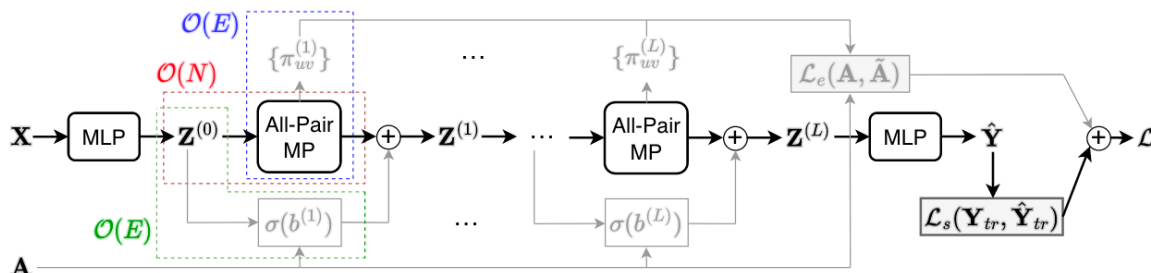
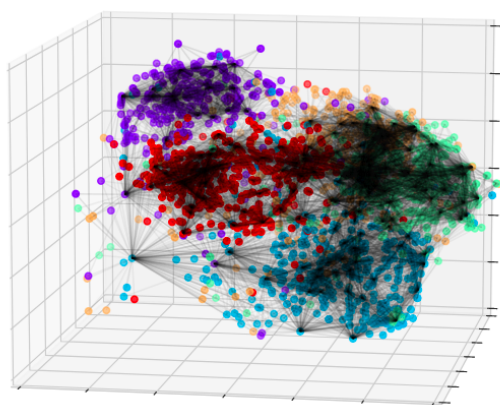
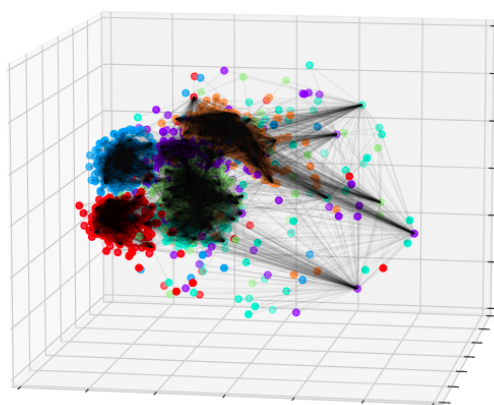


Figure 1: Illustration for the data flow of NODEFORMER which takes node embedding matrix \mathbf{X} and (optional) graph adjacency matrix \mathbf{A} as input. There are three components in NODEFORMER. The first one is the all-pair message passing (MP) module (colored red) which adopts our proposed kernelized Gumbel-Softmax operator to update node embeddings in each layer with $\mathcal{O}(N)$ complexity. The other two components are optional based on the availability of input graphs: 1) relational bias (colored green) that reinforces the propagation weight on observed edges; 2) edge regularization loss (colored blue) that aims to maximize the probability for observed edges. These two components require $\mathcal{O}(E)$ complexity. The final training loss \mathcal{L} is the weighted sum of the standard supervised classification loss and the edge regularization loss.

为了进一步探索模型学到的隐式图结构的特性, 对估计的连边 (筛选出权重大于某个阈值) 和节点 embedding 做了可视化。可以看到, 同类节点之间的连边较为密集, 不同类的节点之间也会存在少许的连边。



(a) 20News-Groups



(b) Mini-ImageNet

对于 graph Transformer, 目前的绝大部分研究都侧重于图分类任务 (每张图是一个样本, 图的规模较小), 例如分子图性质预测。本文首次探索了大规模图上的“两两节点间信息传递”, 实现了面向节点分类大图的 Transformer。

启发:

如果要做节点分类的话, 可以使用这样的 nodeformer, 但是目前节点分类的 transformer 不是特别主流的做法, 效果不一定好。

文献3

题目: Transformer for Graphs: An Overview from Architecture Perspective

作者: Erxue Min^{1*}, Runfa Chen³, Yatao Bian², Tingyang Xu², Kangfei Zhao², Wenbing Huang⁴, Peilin Zhao², Junzhou Huang⁵, Sophia Ananiadou¹, Yu Rong^{2†}

出处: arxiv 2023

方法:

本文从架构设计的角度对各种图 Transformer 模型进行了全面的回顾。首先对现有模型进行了分解,并总结了三种将图信息纳入 Transformer 的典型方法:1) 图神经网络作为辅助模块,2) 改进的图位置嵌入,以及 3) 改进的图注意力矩阵。

范式(trans (本身是全连接图)是全连接操作,会带来噪声,也可以用 smoothing 解决噪声,可以补全数据,但是和数据相关的,符不符合 inductive bias)

1) 先 GNN 做边 (结构信息,是浅层的裁剪后的 GNN) embedding 更新,后 Transformer 计算 pairwise 注意力 (为啥后用:是因为前面 GNN 有可能学习到与任务无关的节点或者边表示)

2) GNN 将 Trans 的输出进行更新

3) GNN 和 Trans 做 sample 然后并行输出

启发:

看了综述里面几个图 transformer 的代码和论文

文献4

题目: Transformer for Graphs: An Overview from Architecture Perspective

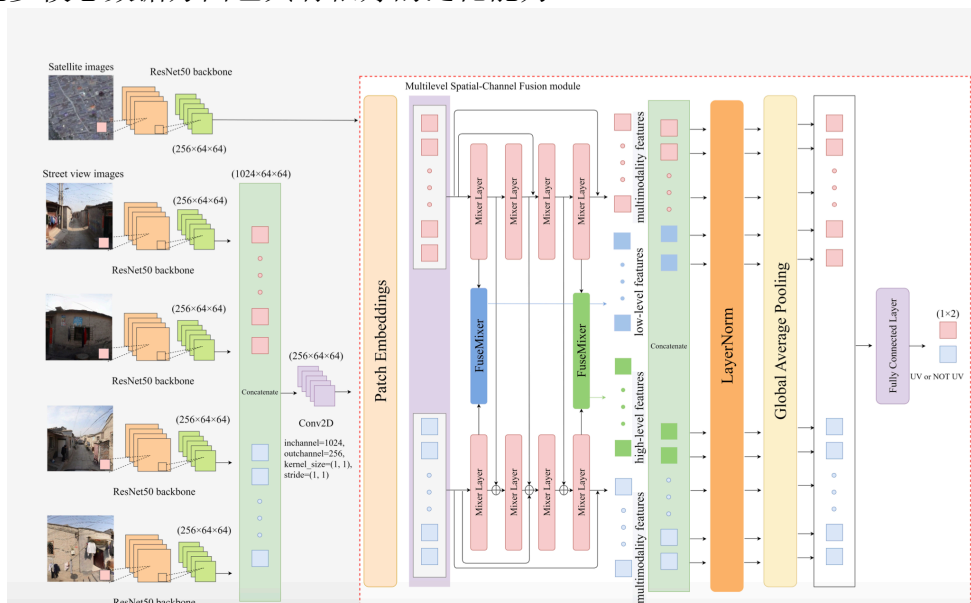
作者: Erxue Min¹, Runfa Chen³, Yatao Bian², Tingyang Xu², Kangfei Zhao², Wenbing Huang⁴, Peilin Zhao², Junzhou Huang⁵, Sophia Ananiadou¹, Yu Rong^{2†}

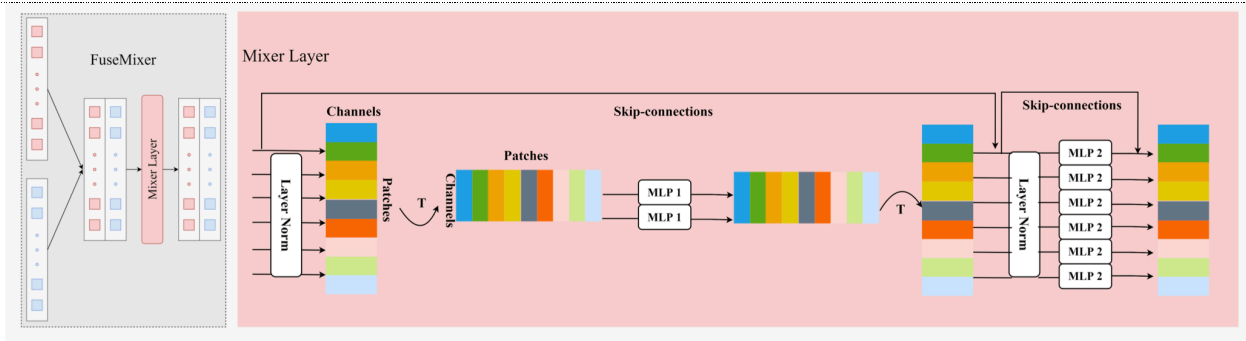
出处: TGRS 2023

方法:

提出了一种结合遥感卫星图像和街景图像对城中村进行提取的新方法。提出了一种新颖的多级空间通道特征融合网络,即 FusionMixer,它集成了基于 CNN 的特征提取模块和多级空间通道特征融合层,对城中村进行提取。实验在深圳市 (人工构建的 RsSt-ShenzhenUV 数据集) 和公共 UV 数据集 (数据集) 进行。与 RsSt-ShenzhenUV 数据集中第二好的融合模型相比,本文提出的 FusionMixer 在验证集和测试集中 OA 分别提高了 8.83%和 8.84%,Kappa 分别提高了 0.1765 和 0.1770。

方法与当前最先进的多模态融合方法 Trans-MDCNN 进行比较,结果表明所提出的 FusionMixer 在融合其他多模态数据方面 also 具有很好的泛化能力。





启发:

论文里对比的多模态融合方法我也可以尝试对比一下，**trans-mdcnn** (B. Chen *et al.*, “Multi-modal fusion of satellite and street-view images for urban village classification based on a dual-branch deep neural network,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 109, May 2022, Art. no. 102794)

RESULTS FOR DIFFERENT MODELS AND THE PROPOSED FUSIONMIXER

	Methods	OA (%)	Kappa
Models with satellite images	VGG-16	87.17	0.7029
	ResNet-50	89.63	0.7662
	DenseNet-121	88.99	0.7464
	ViT	90.80	0.7939
Models with streetview images	VGG-16	83.55	0.6126
	ResNet-50	85.10	0.6404
	DenseNet-121	86.65	0.7028
	ViT	90.03	0.7602
Fused models	Trans-MDCNN [42]	92.61	0.8352
	FusionMixer (ours)	94.43	0.8757

Trans-MDCNN:

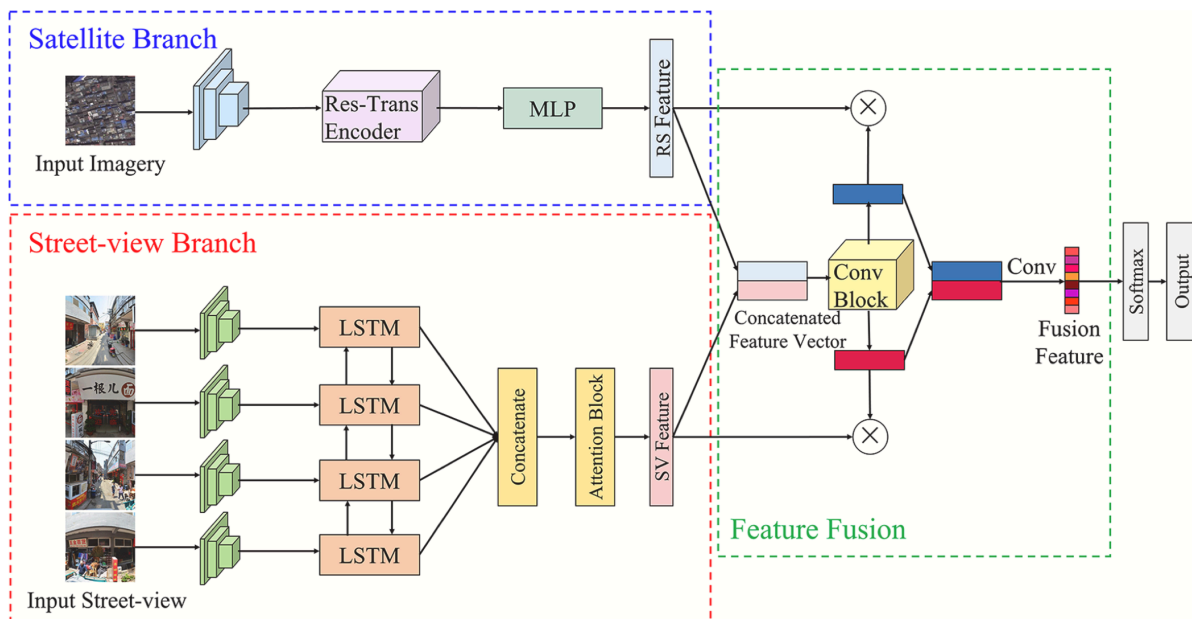


Fig. 6. Structure of the dual-branch deep neural network.

本文用的两个 backbone 都是 VIT（这是在遥感和街景融合里第一次使用双 VIT），可能我那个工作得快点完善了

RESULTS FOR DIFFERENT MODELS AND THE PROPOSED FUSIONMIXER

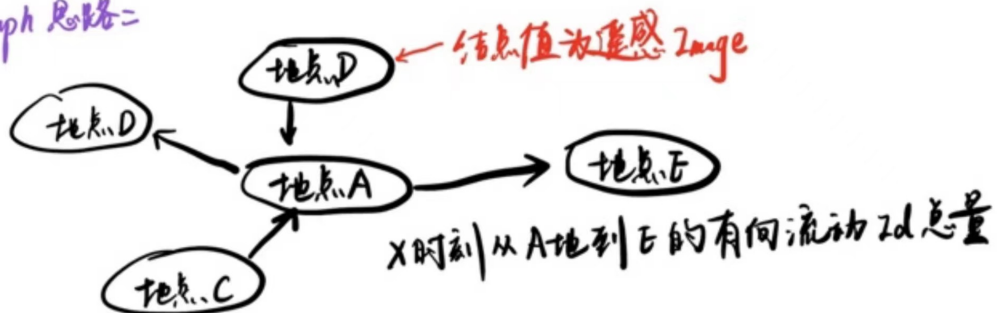
	Methods	OA (%)	Kappa
Models with satellite images	VGG-16	87.17	0.7029
	ResNet-50	89.63	0.7662
	DenseNet-121	88.99	0.7464
	ViT	90.80	0.7939
Models with streetview images	VGG-16	83.55	0.6126
	ResNet-50	85.10	0.6404
	DenseNet-121	86.65	0.7028
	ViT	90.03	0.7602
Fused models	Trans-MDCNN [42]	92.61	0.8352
	FusionMixer (ours)	94.43	0.8757

工作进展

- 1: 阅读文献;
- 2: 还在找长时序 graph, 直接对节点做分类的相关文献。

3/3

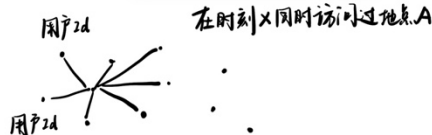
graph 思路二



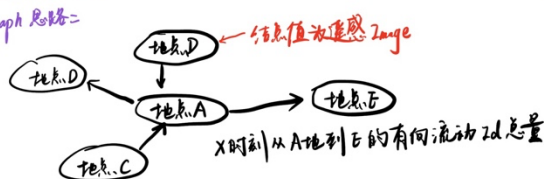
- 3: 重新 python 的 numpy 和 pandas 库清洗用户访问数据, 建立第二种 graph:
在写代码, 代码有点麻烦

16:24 9月12日周二

graph 思路一



graph 思路二



下周计划

1. 找长时序 graph, 直接对节点做分类的相关文献

2. 看 graph 的建立有没有其他思路
3. 补充论文
4. 数据清洗代码和建 graph 代码