

周报（2023.5.15-2024.5.22） 姓名： 孙瑞阳

每日小结

	周一	周二	周三	周四	周五
早	对比学习代码	Mamba 代码	阅读论文	Mamba 代码套用	Graph 代码框架调整
中	对比学习代码	苏老师那边画图	对比学习代码	GCN 代码 debug	对比学习代码
晚	Mamba 代码套用	苏老师那边画图			

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

<p><b>文献 1</b></p> <p><b>题目：</b>Measuring Chinese mobility behaviour during COVID-19 using geotagged social media data</p> <p><b>作者：</b>Kaixin Zhu, Zhifeng Cheng &amp; Jianghao Wang</p> <p><b>出处：</b>子刊/2024 (Humanities and Social Sciences Communications)</p> <p><b>方法：</b>对标谷歌做了细粒度中国城市动态演化：使用微博社交媒体平台上 1000 万用户的 2.1 亿条地理标记帖子，提供了细粒度的移动数据，描述了 2020 年中国功能性设施的每日访问量与 2019 年基线的百分比变化。</p> <p><b>方法（很简单，都是统计分析，没有标签）：</b></p> <ul style="list-style-type: none"><li>● <b>数据验证：</b>该研究测试了微博的 1000 万用户是否可以在地理分布和人口特征方面代表整个人口。本研究以《中国人口普查年鉴 2020》(<a href="https://www.stats.gov.cn/sj/tjgb/rkpcgb/qgrkpcgb/202302/t20230206_1902003.html">https://www.stats.gov.cn/sj/tjgb/rkpcgb/qgrkpcgb/202302/t20230206_1902003.html</a>)中的常住人口数量为基准，计算各省级区域居民占总人口的比例与各省级区域用户占总用户的比例之间的相关关系。此外，该研究还将微博用户的年龄和性别结构与官方统计数据(<a href="http://www.stats.gov.cn/sj/ndsj/2021/indexch.htm">http://www.stats.gov.cn/sj/ndsj/2021/indexch.htm</a>)中获取的整个人口的年龄和性别结构进行了比较。下一步，本研究将微博流动性数据与百度前情数据进行对比，交叉验证并检验流动性指数是否代表整个人口(Huang et al. 2021;Hunter et al. 2021)。在中国 COVID-19 大流行期间，百度前情数据是使用最广泛的开放移动数据源之一(Hu et al. 2021)。它提供了一个城市运动强度(CMI)指数，显示了 2020 年 1 月至 4 月期间在城市内旅行的人口比例。为了获得与百度千玺 CMI 可比的指标，本研究计算了某一天在城市内访问非住宅场所(包括工作场所、零售娱乐、公园、公交车站和杂货店药房)的用户比例作为微博 CMI 指数。然后，研究将这两个 CMI 指数归一化，并将其从城市层面汇总到省和国家层面，并以《中国人口普查年鉴 2020》中检索的实际城市居民加权。接下来，该研究使用 7 天移动平均法对其进行平滑处理，并通过将其与大流行前(2020 年 1 月 1 日至 1 月 19 日)的中位数进行比较来计算百分比变化。最后，该研究检查了 2020 年 1 月至 4 月的 Pearson 相关性。在相关分析的基础上，通过检验这两个指标是否具有相似的时间趋势来推断微博流动性数据的可靠性。</li><li>● <b>访问行为的时间变化：</b>为了应对大流行的爆发，人们在设施访问、日常活动(Lucchini 等人, 2021 年)以及与娱乐场所的互动方面调整了他们的移动行为。该研究利用从大规模签到数据中提取的人们访问的精确位置，量化了中国 COVID-19 爆发后访问行为的变化。</li><li>● <b>移动距离：</b>行进距离是人类运动的关键描述，是研究移动规律的关键组成部分，也是传染病传播建模的关键参数。然而，目前尚不清楚在封锁政策下，人们是如何改变出行距离到不同的设施的。在这里，该研究计算了在大流行爆发之前和之后个人层面访问六类地方的</li></ul>
---

旅行距离。该研究首先为用户过滤了最常访问的住宅 poi，以确定他们的家的位置，然后计算了每次郊游的 haversine 距离

文章发现了一些有用的规律（验证了常识）：

(i)在疫情爆发后的一个月內，前往工作场所、零售和娱乐场所、公园、中转站、杂货店和药店的次数减少了 16. 8%至 57. 7%；

(二)年轻、本科及以上学历、未婚的外出旅游人数下降幅度较大；

(iii)人们更喜欢去附近的地方，导致在其住所 3 公里半径內的零售和娱乐场所的访问量增加了 4. 3%。

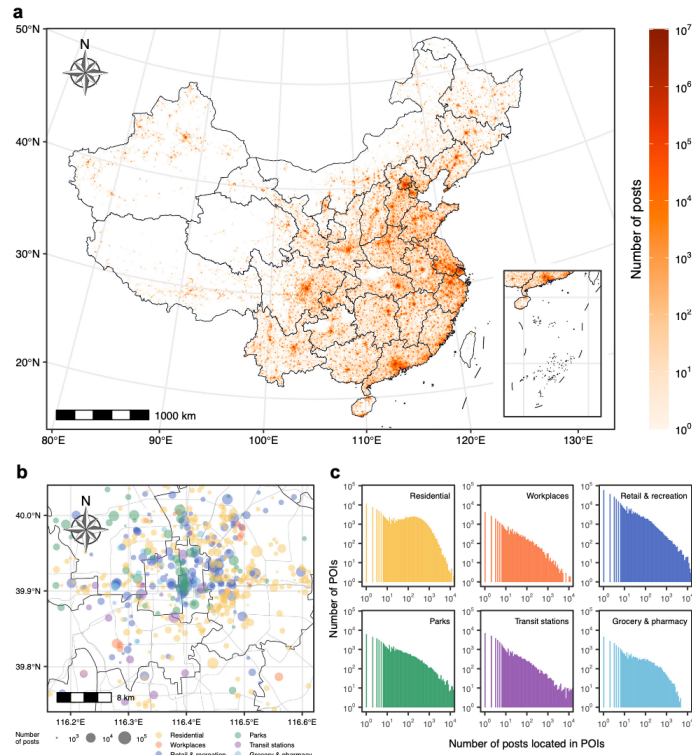
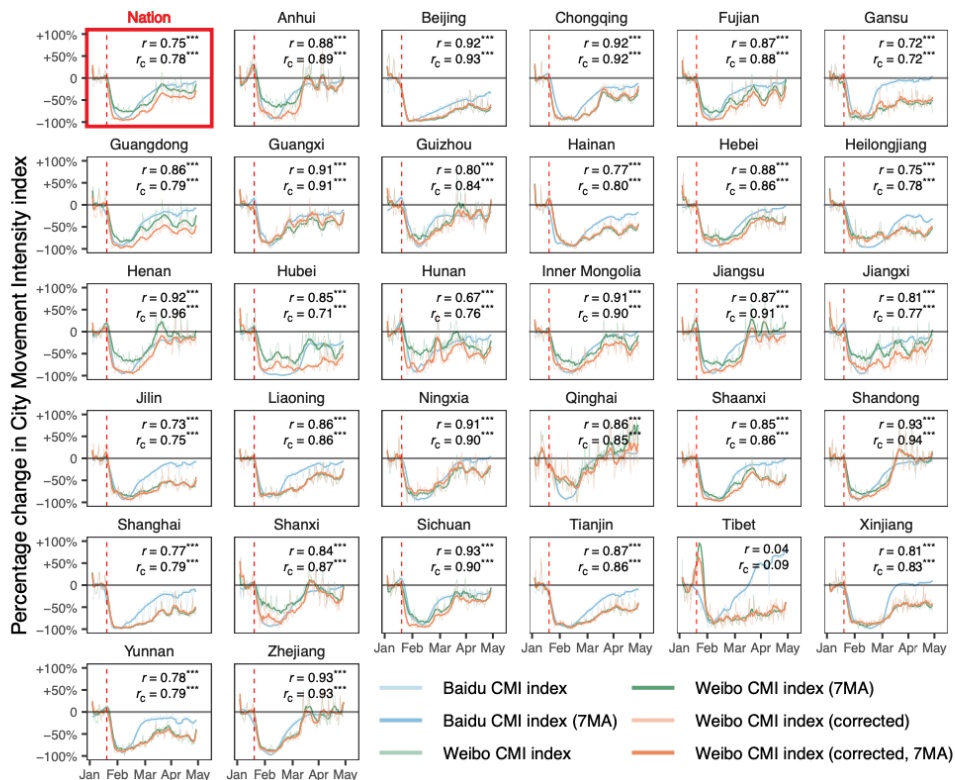
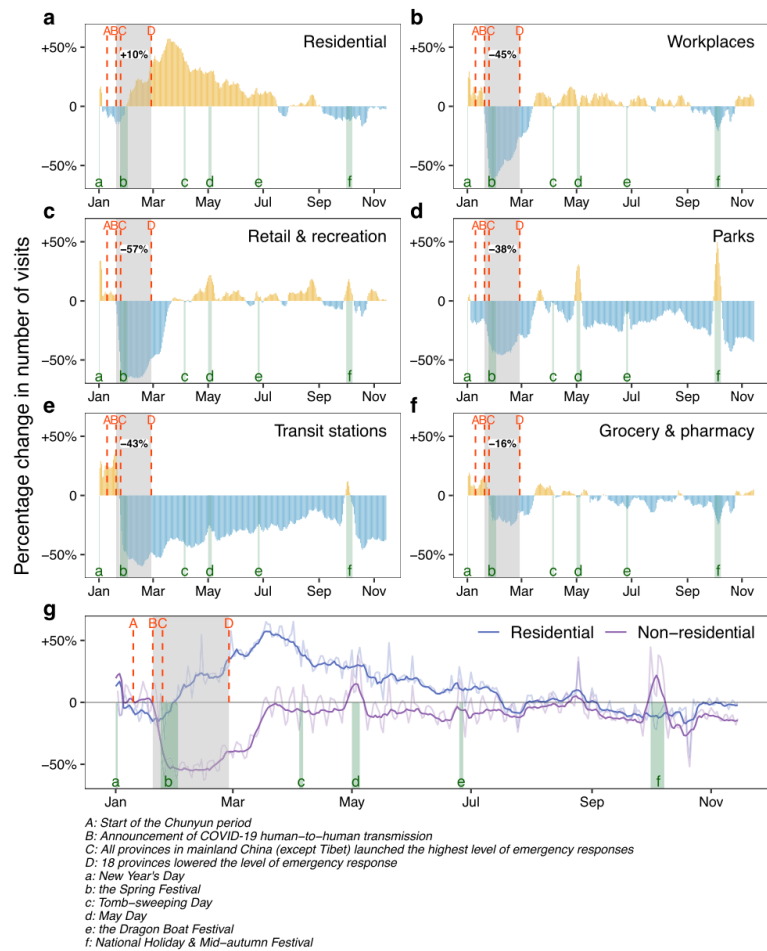


Fig. 1 Description of geotagged Weibo data from 2019 to 2020. **a** The spatial distribution of geotagged posts in China. **b** The spatial distribution of categorized geotagged posts in Beijing. **c** The number of posts located in POIs against the number of POIs.



启发:

1. 这类工作都是通过大数据计算了几个指标来反映城市动态，去验证常识，发现新的规律。微博用户他应该只筛出来 1200 万用户（整个数据集没公开）微博数据（应该是和微博合作的，因为全国范围，且数据量很大，文章开源了微博移动数据量的变化百分比（没什么用，下图 CSV 的第四列）

Code Issues Pull requests Actions Projects Security Insights

Files

main

Go to file

publication

figure1\_data\_description

figure2\_mobility\_index

figure3\_test\_representativeness

figure4\_data\_validation

figure5\_subpopulations\_heterog...

figure6\_mobility\_temporal\_chan...

figure7\_mobility\_distance

geotagged\_data

percentage\_change\_visits\_330...

readme.md

template\_percentage\_change\_...

color\_setting.R

website

gitignore

LICENSE.txt

Socialmedia\_Mobility / publication / geotagged\_data / template\_percentage\_change\_visits\_Beijing\_Shanghai.csv

KaixinZhu update

6a029ec · 2 days ago · History

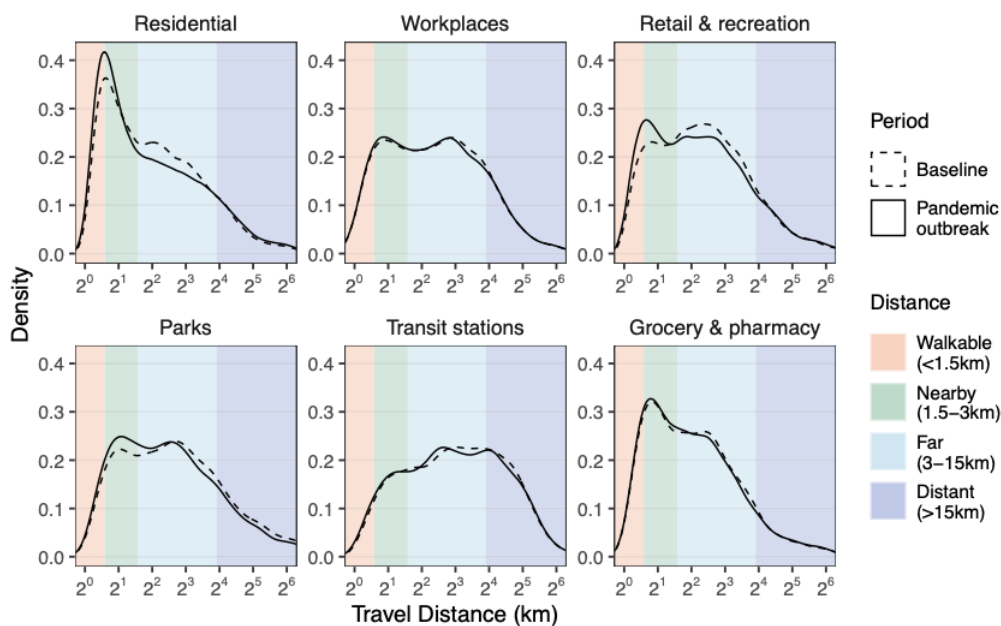
Preview Code Blame 4021 lines (4021 loc) · 397 KB

Search this file

city	category	created_at	visits_7MA_percentage_change	ct_adcode	city_en	city_ch	pr_adcode	pr_ch	pr_en
北京	Grocery & pharmacy	2020-01-05	-1.5703868822209	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-01-12	-2.29949623275379	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-01-19	-8.07627593942794	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-01-26	-16.9377453729669	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-02-02	-42.2321929332586	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-02-09	-53.2249018508132	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-02-16	-50.6449803701627	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-02-23	-55.6926528323051	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-03-01	-36.3432417274257	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-03-08	-46.9994391475042	110100	Beijing	北京市	110000	北京市	Beijing
北京	Grocery & pharmacy	2020-03-15	-48.7380818844644	110100	Beijing	北京市	110000	北京市	Beijing

Data availability : The geotagged social media data used to support the findings of this study is freely available at [https://casgis.github.io/Socialmedia\\_Mobility/](https://casgis.github.io/Socialmedia_Mobility/). The scripts for reproducing the analysis are available at [https://github.com/CASGIS/Socialmedia\\_Mobility](https://github.com/CASGIS/Socialmedia_Mobility). )

- 故事讲的很好，和疫情细节（政策调整）吻合的很好
- 对于个体，六类目的地的旅行距离分布。住宅类别代表用户家中以外的居住场所。实线表示基线期(2019年下半年)，虚线表示大流行爆发期(2020年1月20日之后的一个月)。背景颜色表示不同的距离分区。



**Fig. 7 The distribution of travel distance across six categories of destinations at the individual level.** Specifically, the residential category represents residential places other than the user's home. The solid line indicates the baseline period (the second half of 2019), and the dotted line indicates the pandemic outbreak period (the month following January 20, 2020). The background colours indicate different distance partitions.

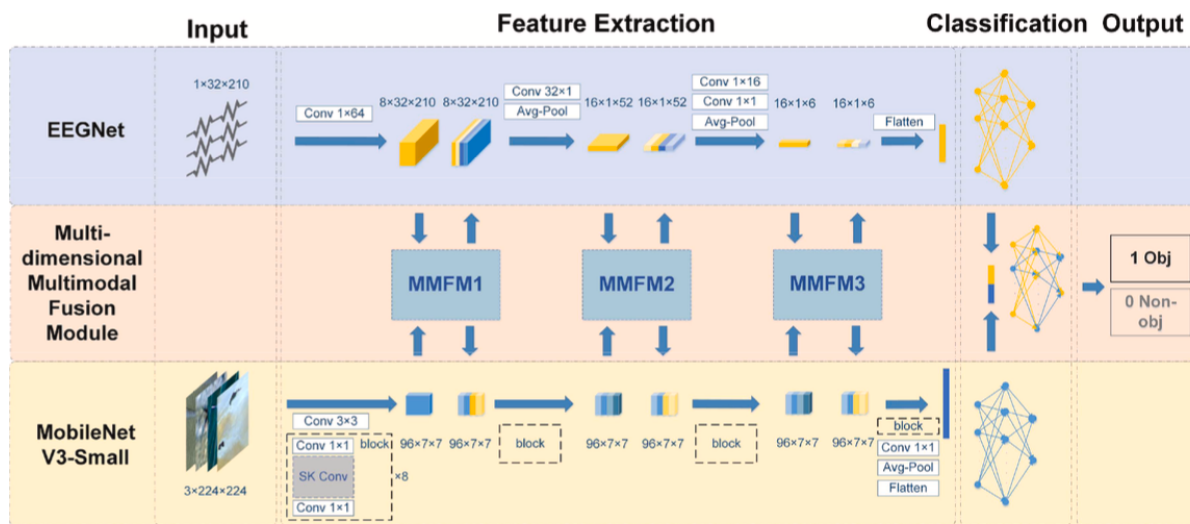
## 文献2

题目: SK-MMFMNet: A Multi-dimensional Fusion Network of Remote Sensing Images and EEG signals for Multi-scale Marine Target Recognition

作者: Jiawen Long, Zhixiang Fang, Lubin Wang

出处: Information Fusion 2024

方法: 在本研究中，提出了一种多式联运海洋目标分类方法框架(SK-MMFMNet)融合遥感图像和脑电信号。



MMFM结构：MMFM的实现分为通道维度融合和空间维度融合。在通道尺寸融合部分，两者都具有A 从EEG信号和图像中提取的B经过平均池化层、拼接层、全连接层，得到激活特征EA和EB。激活特征分别重新分配到A和B，形成初步融合特征A1和B1。空间维度融合组件利用A1和B1，生成最终的融合特征A2和B2。

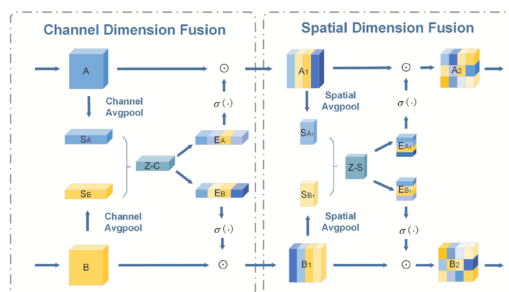


Fig. 3. The implement of MMFM, which is divided into channel dimension fusion and spatial dimension fusion. In channel dimension fusion part, both feature A from EEG signals and B from images undergo average pooling layer, concatenation operation, and fully-connected layer to produce activated features  $E_A$  and  $E_B$ . Activated features are respectively reassigned to A and B, resulting in preliminary fused features  $A_1$  and  $B_1$ . The spatial dimension fusion component utilizes  $A_1$  and  $B_1$ , generating the final fused features  $A_2$  and  $B_2$ .

SK-MMFMNet概述，其中包括EEGNet, MobileNetV3-Small和融合模块MMFM。两种网络分别以脑电信号和遥感图像为输入，在特征提取阶段采用三次MMFM进行特征融合。最后，将两个网络得到的特征进行连接和分类，然后输出最终的预测结果。

Sub-dataset	Network	Accuracy	Recall	Precision	F1-Score	AUC
Island	SK&MMTMNet	<b>0.98± 0.00</b>	0.91± 0.04	0.89± 0.03	0.90± 0.02	<b>0.99± 0.01</b>
	MMFMNet	<b>0.98± 0.01</b>	<b>0.92± 0.04</b>	0.89± 0.04	0.90± 0.02	<b>0.99± 0.01</b>
	SK-MMFMNet	<b>0.98± 0.01</b>	0.91± 0.04	<b>0.91± 0.05</b>	<b>0.91± 0.03</b>	<b>0.99± 0.01</b>
Wind Turbine	SK&MMTMNet	0.95± 0.00	0.81± 0.06	0.77± 0.04	0.79± 0.02	0.97± 0.01
	MMFMNet	0.95± 0.01	0.80± 0.06	0.77± 0.05	0.78± 0.04	0.96± 0.03
	SK-MMFMNet	<b>0.96± 0.01</b>	<b>0.83± 0.04</b>	<b>0.83± 0.06</b>	<b>0.83± 0.03</b>	<b>0.98± 0.02</b>
Ship	SK&MMTMNet	0.92± 0.01	0.58± 0.07	0.65± 0.05	0.60± 0.04	0.90± 0.04
	MMFMNet	0.91± 0.01	0.49± 0.05	0.60± 0.07	0.53± 0.03	0.86± 0.05
	SK-MMFMNet	<b>0.93± 0.01</b>	<b>0.61± 0.11</b>	<b>0.71± 0.07</b>	<b>0.65± 0.08</b>	<b>0.91± 0.05</b>

启发：

4. 和我们之前的思路是比较相似的，可以进一步尝试
5. 可以学习这部分：T-SNE可视化结果验证，类似于聚类的思想可视化对比两种单模态和多模态特征：

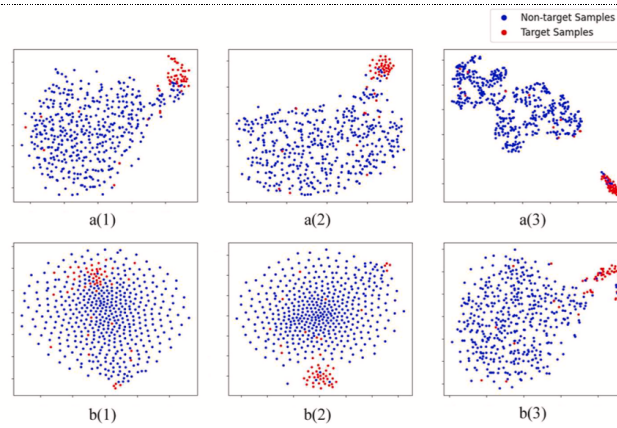
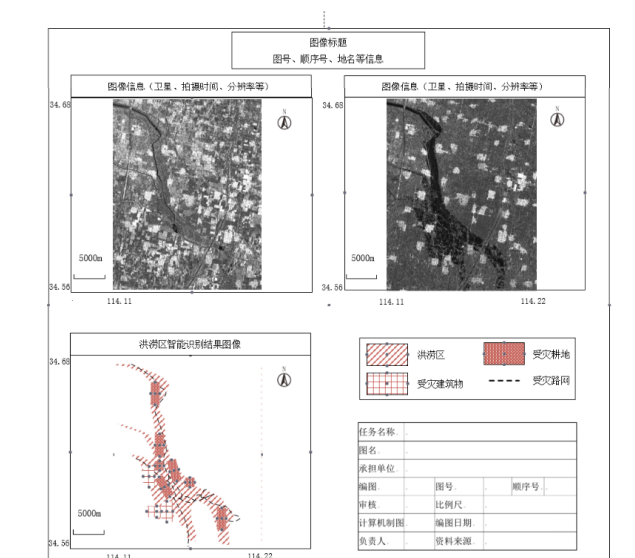


Fig. 13. T-SNE visualization. The a(1) – a(3) is the visualization of the feature map after each fusion module of SK-MMFMNet. The b(1) – b(3) is the feature visualization of MobileNetV3 at the corresponding convolutional block.

## 本周工作

1. 阅读文献；
2. 对比学习代码在增加：有很多节点间流动值为1和0，把节点特征做对比学习（resnet或者VIT的浅层特征如果比较相似的节点，就看作一类地点，同类地点之间流动量为1就和不同类别流动量为1不同，做对比学习）
3. 新方法建图后baseline调参：图建立后Resnet表达结点（resnet和lstm后期还会替换，baseline打算用这个），LSTM表达边信息，用简单的一层GCN尝试，效果跑出来准确度0.61（还需要再调，和以前参数差距很大，应该还能涨点）
4. Mamba的代码加进去，看加在什么位置，要试试
5. 帮苏老师那边画了个图：



## 下周计划

1. 对比学习增加后看看效果