# 周报 (2024. 2.27-2024.3.5) 姓名: 孙瑞阳

# 每日小结

	周一	周四	周五	周六	周日
早	写 response,	阅读文献	阅读文献	阅读大量 2023	FCN+LSTM代
	FCN+LSTM 代码			年文献	码,写 response
中	阅读文献,写	阅读文献	FCN+LSTM 代	FCN+LSTM代	FCN+LSTM代
	response		码	码,写 response	码,写 response
晚	写 response,		FCN+LSTM代	写 response	代码调参,大数
	FCN+LSTM 调参		码		据集代码

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: Comprehensive Systematic Review of Information Fusion Methods in Smart Cities and Urban Environments

作者: Mohammed A. Fadhel, Ali M. Duhaim, Ahmed Saihood, Ahmed Sewify, Mokhaled N. A. Al-Hamadani, A.S. Albahri, Laith Alzubaidi, Ashish Gupta, Sayedali Mirjalili, Yuantong Gu

出处: Information Fusion 2024

# 方法:

探讨了智慧城市中使用的不同信息融合方法及其优点和挑战。对文献进行了全面的检索并应用了选择性标准。发现了 59 项近期研究,涉及智慧城市应用中的机器学习 (ML) 和深度学习 (DL) 技术。本研究通过补充现有研究,为智慧城市提供更详细的见解。智慧城市论文中机器学习/深度学习和信息融合的词云可视化展示了多样化的景观,涵盖人工智能的技术方面和城市环境中的实际应用。除了技术探索之外,该研究还深入研究了智慧城市中出现的伦理和隐私影响。此外,它还彻底探讨了充分实现这场城市革命潜力所必须解决的挑战



Fig.1: Bird's-Eye View of the Smart City.

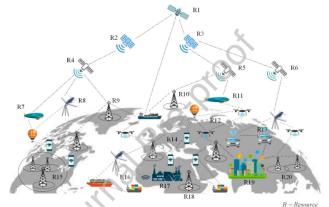


Fig. 13: Integration of 6G Communication Technologies for Smart City and Information Fusion



Fig. 5: Word Cloud.

提出的挑战: 智慧城市在实施信息融合方面面临的挑战,包括各种来源的数据集成、隐私和安 全问题以及质量和可靠性问题:

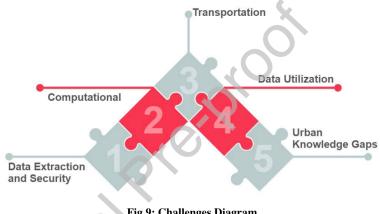


Fig.9: Challenges Diagram.

### 启发:

- 1. 可以增强我自己工作Introduction的深度,在仔细学习这篇综述
- 2. 发现很多有意思的工作,在具体学习,比如:

考虑一个智慧城市正在处理突发交通拥堵事件的场景。多个数据源,包括交通摄像头、车辆的 全球定位系统(GPS)数据以及社交媒体帖子,可以提供碎片信息。信息融合可以结合这些数据 来识别拥堵原因,建议替代路线,并实时触发自适应交通信号调整。

3. 文章整理了比较全面该领域的多源数据集,后面打算详细看看:

Ref.	Name	Description	Size	Link	Input Type and Format	Lawfully	Big Data	Public or Private
[65]	N/A   CRAWDAD RICE/AD_H OC_CITY	Sample subset of user information from the cloud premises   Traces of movements of the city busses in Scattle	Several weeks' worth of data for 1200 Buses spanning 5100 square kilometer region	CRAWDAD rice/ad_hoc_city.\( \) IEEE DataPort (ieee-dataport.org)	.tar.gz	Y	No   Ye s	Priv ate I Pub lic
[48]	Parkinson's disease   NSL-KDD	Parkinson's Disease Prediction   Intrusion Detection	31 patients, 195 instances of speech signalsH1 functions, 150k records	https://archive.ics.nci.edu/dataset/17 4/parkinsons   https://www.unb.ca/cic/dataset/nsl. html	Numeric al: zip (.data and .mames) l .txt and.ARF	Y es	No I Ye s	Pub lic
[66]	AFEW   RECOLA   BAUM-1   EMOEEG   CMU- MOSEI   WESAD	Multi-modal emotion datasets	3301461311811000	Jago, Fern w. Loope, com recording just Wilder-2, 24 Section 1, 24 Secti	Audio, visual I Audio, visual, ECG, EDA I Audio, visual I EEG, ECG, ECG, EDA I Text, visual, audio EDA, Text, visual, ECG, EDA I Text, visual, ECG, EDA I Text, visual, ECG, ECG, ECG, ECG, ECG, ECG, ECG, ECG	Yes	Ye s	Public
[33]	Hetrec2011- lastfm-2k	N/A	1864579 user, artist, label triplet instances	HetRec 2011   GroupLens	Textual and Numeric al: dat	Y es	Ye s	Pub lic
[36]	Winnipeg Transit Open Data Web Service	Winnipeg transport locations and trajectories with trip information (e.g., temperature and road conditions) from multiple users	A year of trip information collected from anonymous users sponning Winnipeg	https://api.winnipegtransit.com/	Textual and Numeric all JSON	Y es	Ye s	Pub lic
[37]	N/A	Government's reports of Dalian city	6000 records fallen into 6 subsets   8000 records fallen into 8 subsets	N/A	N/A	Y es	N/ A	Pub lic
[29]	Simulated Data	Simulated WSN parameters	200 sensor nodes, with each source generating 4000 bits	N/A	Textual and Numeric al	N/ A	N/ A	Priv ate
[41]	UCF-Crime   RWF-2000	Annotated Surveillance Videos   Surveillance Clips	1900 surveillance videos containing 13 instances of crime I 2000 short surveillance YouTube clips involving anomalous behaviours	https://ieeexplore.ieee.org/document /8578576   RWF-2000; an open large-scale video database for violence detection	Visual	N o	Ye s	Pub lic
[23]	Dati Comune Milano   CORINE land cover   Lombardia open data   Telecom	Milano demographics, call phone records, land use classification and points of interest in the region	10s   10Ks   100Ms   1Ks   10 Ks	http://duti.comune.milano.it/   https://www.eco.europs.eu/publicati ens/CORO-handcover   https://www.eco.europs.eu/publicati ens/CORO-handcover   https://www.upenstreetmap.org/#ma pub//224/53/13/3275	Shapefil e and Tabular	Yes	Ye s	Pub

Table (3): The Smart City Information Fusion Datase

	Italia Big Data Challenge   OpenStreet Map							
[50]	OpenStreet Map   NavInfo Traffic Data platform: Beijing, Shanghai, and Guangzhou	Traffic information of 3 Chinese cities: Road Network topology and speed data	1-month worth of information on 3900 roads across the 3 cities with 8928-time instances	https://www.openstreetmap.org/1 ultrafficindex.com	Shapefil e l Numeric al	N/ A	Ye s	Pub
[67]	GREEND   PLAID   WHITED	Electrical appliance usage patterns and power consumption in domestic buildings	6-month daily recordings of 6 electrical appliances and power footprints of 11 electrical appliances in domestic buildings.	reflub.elsevier.com/S1566- 2535(20)30315-8/sbref0236   reflub.elsevier.com/S1566- 2535(20)30315-8/sbref0237	Numeric	Y es	Yes	Pub
[39]	EnvBodySen s	Nottingham female participants' heart rate, galvanic skin response, SGR, body temperature, movement, and environmental data	40 females	Towards unravelling the relationship between on-body, environmental and emotional data using sensor information fusion approach - Socrace Direct	Numeric al and Visual	Y es	Ye s	Pub
[51]	Citi Bike New York I N/A	Bike trip records (duration, start/stop time, station IDs) 1.5 years of NYC weather data (time, condition, temperature, wind speed, humidity, precipitation)	5.35M records 15 years of hourly records	https://cibbik.rsyc.com/system-data	N/A	Y es IN o	Ye s	Pub
[68]	Traffic Datasets	31 Datasets containing recorded traffic videos and information	31 Datasets	https://www.sciencedirect.com/scie nce/article/pii/S1566253522000859	Visual and Numeric al	Y es	Ye s	Pub
[34]	CIFAR-101 FEMINIST	Colour Images and industrial information	60K color images I 805263 samples	https://www.cx.toronto.edu/~kriz/cif ar.html https://carolinesinders.com/feminist -data-set/	Visual and Numeric al	Y es	Ye s	Pub
[53]	Crime-theft   Ride-hailing   Fire	Urban theft crime information I Urban ride- hailing demand I Urban fire records	5602 time slots   3628   1791 time slots	https://www.sciencedirect.com/scie nce/article/pii/S0925231221007190	Visual and Numeric al	Y es	Ye s	Pub
[49]	BIMCV COVID-19+	COVID-19 RX and CT images	1Ks patients	https://arxiv.org/abs/2006.01174	Visual and Numeric al	es	A/	Pub
[45]	CropScape I Lansat 8	Labellod rerrain data	N/A	https://www.sciencedirect.com/scie nce/article/pii/S0168169912000798	Visual and Numeric al	Y es	Ye s	Pub lic
[56]	BIT-vehicle	Several high-resolution images	N/A	https://link.springer.com/article/10.1 007/s00500-021-05696-3	Visual	N/ A	N/ A	Pub
[42]	MotionSense	Participants walking, running and navigating stairs	24 participants, 4 activities, 15 trials	https://arxiv.org/abs/1802.07802	Visual and Numeric al	Y es	N/ A	Pub
[46]	Set5   Set14   BSD100   Urban100   N/A BSD300	Various images	5114110011001911 200	https://www.sciencedirect.com/scie nce/article/pii/S0165168420303753	Visual	A/	N/ A	Pub
[70]	NYC open data! Barcelona transportatio n data! BCL: Beijing city lab! Chicago data port! N/A	Various types of city data for New York, Beijing, Chicago and London	N/A	https://opendata.eityofnewyork.us/1 https://link.springer.com/article/10.1 007/s13132-012-0884-9   https://link.springer.com/article/10.1 007/s13132-012-0884-9   https://data.eityofchicago.org/1 https://www.data.gov.uk/	Various types of city data: Visual, Numeric al, Textual	Y es	Ye s	Public

# 文献2

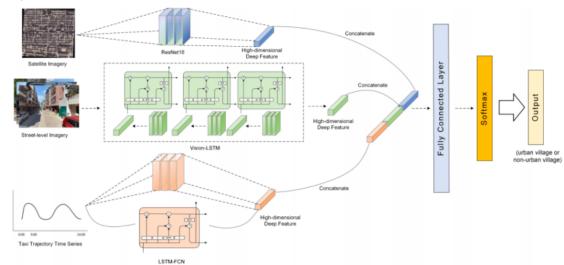
题目: Comprehensive urban space representation with varying numbers of street level images

**作者:** Huang, Y., Zhang, F.\*, Gao, Y., Tu, W., Duarte, F., Ratti, C., Guo, D., & Liu, Y

出处: Computers, Environment and Urban Systems 2023 方法:

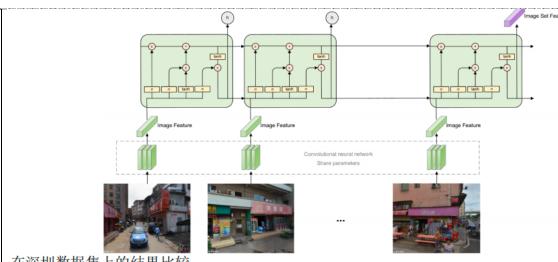
面向城中村分类提出了一种卫星影像、街景影像和出租车轨迹数据融合的视觉长短时记忆(Vision-LSTM)模块,用于从不同数量的街景影像中获取向量表示。该网络使用卫星影像、街景影像和出租车轨迹数据的组合作为输入。这个模型的结构,如下图所示,包括三个分支。

- 在卫星影像分支中,采用 ResNet18 骨干模型来提取卫星图像正射视图的特征。
- 同时,所提出的 Vision-LSTM 模块用于在街景影像分支中提取不同数量的街景影像的视觉特征。
- 在出租车轨迹分支中,采用 LSTM 全卷积网络(LSTM-FCN)骨干模型,根据出租车轨迹学习出行量时间序列数据的移动性特征。
- 最后,从所有分支提取的特征被连接起来,并通过一个 softmax 层进行城市和非城市村落的分类。



#### Vision-LSTM

提出的 Vision-LSTM 模块,如图 2 所示,包括一个具有共享权重的卷积神经网络(CNN)和一个递归神经网络(RNN)。不同的街景影像数量在空间单元中可以被视为不同长度的图像集。为此,使用零填充方法,这是一种在自然语言处理中常用的技术,来处理这种可变长度的街景影像集。随后,图像集中的每张图像都单独输入到一个具有共享权重的 CNN 模型中,以提取单个图像的语义特征,并将输入图像集处理为图像特征集。在零填充过程中添加的任何空白图像都会被忽略,确保它们在训练过程中不会影响模型参数。由于 LSTM 不考虑二维空间信息,在训练过程中,这些图像特征集以随机顺序作为时间步骤输入到 LSTM 中。得到的深度特征反映了从行人的视角观察到的物理环境,并根据街景影像集的不同长度,代表了空间单元的整体特征。



在深圳数据集上的结果比较:

Data	Model	OA (%)	Kappa	F1
SI	ResNet18	81.8	0.541	0.650
SLI	Vision-LSTM (proposed in this study)	82.8	0.540	0.647
Taxi trajectory	LSTM-FCN	71.9	0.311	0.477
SI + SLI	ResNet18 + Vision-LSTM	89.0	0.647	0.715
SI + Taxi trajectory	ResNet18 + LSTM-FCN	87.8	0.645	0.721
SLI + Taxi trajectory	Vision-LSTM+LSTM-FCN	87.2	0.588	0.668
SI + SLI + Taxi trajectory	ResNet18 + Vision- LSTM+LSTM-FCN	91.6	0.720	0.773

#### 不同融合方法比较:

Resolution	Method	OA(%)	Kappa	F1
250 m	No fusion (random image)	79.7	0.382	0.484
	Average Pooling	77.0	0.359	0.471
	Maximum Pooling	72.5	0.311	0.436
	Element-wise Sum	69.1	0.273	0.411
	Vision-LSTM (proposed in this study)	80.5	0.407	0.491
500 m	No fusion (random image)	88.1	0.634	0.708
	Average Pooling	89.1	0.656	0.727
	Maximum Pooling	79.3	0.461	0.588
	Element-wise Sum	77.4	0.432	0.566
	Vision-LSTM (proposed in this study)	91.6	0.720	0.773

#### (融合方法这里依然没有什么创新, concat三个分支)

# 启发:

- 1. 最大的创新就是在街景那里用到的 VisionLSTM, RNN 的输入是共享权重的 CNN
- 2. 在社交数据,使用的时序出租车出行量,采用 LSTM 全卷积网络(LSTM-FCN)骨干模型,还 是之前姚尧老师在遥感-电力数据那里用的模型(下面文献3介绍的网络)
- 3. 代码跑通,精度不如我们的ConvTransformer

#### 文献3

题目: Multivariate LSTM-FCNs for Time Series Classification

作者: Fazle Karim, Somshubra Majumdar, Houshang Darabi, Samuel Harford

出处: IEEE Access, 2019

# 方法:

提出将 FCN 与 LSTM 结合构建模型,对时序数据进行分类。通过使用长短期循环神经网络(LSTM RNN) 子模块(称为 LSTM-FCN) 或具有注意力的 LSTM RNN(称为 ALSTM-FCN) 来增强 FCN 模块来提高 FCN 的性能。与 FCN 类似,两种提出的模型都可以用来可视化卷积层的类激活图(CAM),以检测 对类标签有贡献的区域。此外,注意力 LSTM 还可以通过注意力 LSTM 单元的上下文向量来检测

输入序列中对类标签有贡献的区域。LSTM-FCN 和 ALSTM-FCN 模型的一个主要优点是它不需要大量的预处理或特征工程。结果表明,新提出的模型 LSTM-FCN 和 ALSTMFCN 在加州大学河滨分校 (University of California Riverside, UCR)基准数据集上显著提高了性能。在大多数 UCR 基准数据集上,LSTM-FCN 和 ALSTM-FCN 比几种最先进的集成算法产生更好的结果。

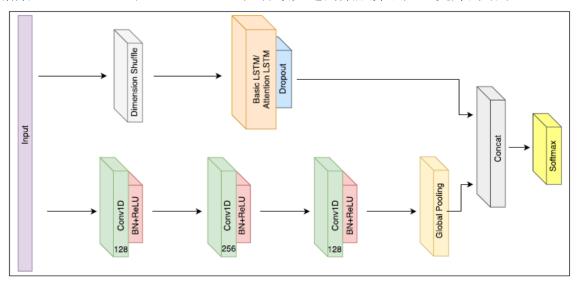


Fig. 1: The LSTM-FCN architecture. LSTM cells can be replaced by Attention LSTM cells to construct the ALSTM-FCN architecture.

在所提出的模型中,全卷积块由 LSTM 块增强,然后是 dropout,如图所示全卷积块由三个堆叠的时间卷积块组成,滤波器大小分别为 128、256 和 128。每个块由一个时间卷积层组成,伴随着批归一化 13,然后是 ReLU 激活函数。在最后的卷积块之后应用全局平均池化。同时,时间序列输入被传送到一个维度。然后将维度变换后的时间序列传递到 LSTM 块中。LSTM 块由一般 LSTM 层或注意 LSTM 层组成。全局池化层和 LSTM 块的输出被连接并传递到 softmax 分类层。

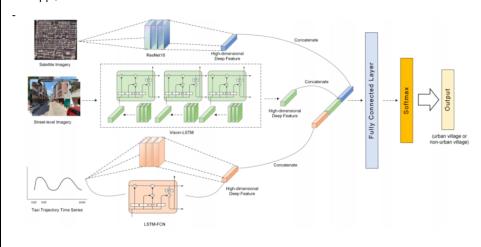
如图示,上面一条路线是通过 LSTM,下面路线是通过 FCN,然后将结构 concat,再经 softmax 进行分类。

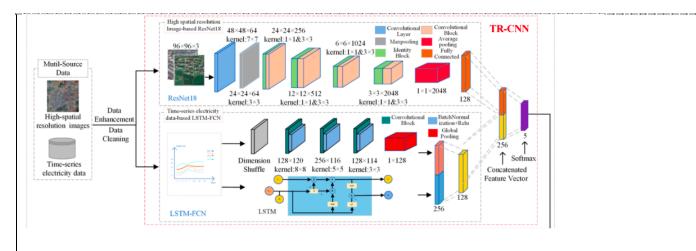
## 启发:

1. Network Input

这里 LSTM 的输入格式跟 FCN 的输入格式是不一样的。例如 以一个多维的时序数据为例子,输入到 LSTM 的时候,网络输入采用单个时间步长的多元时间序列接受。 输入到 FCN 的时候,网络采用多个时间步长的单变量时间序列接受。

2. 23年那个工作中就使用了这个网络作为时序特征提取模型。(和姚尧老师20年的工作一样)

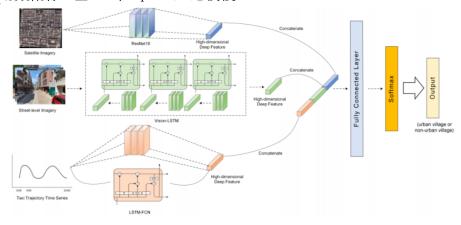




3. 代码跑通,精度不如我们的ConvTransformer

- 1: 阅读文献;
- 2:代码跑通: 23 年工作对比 Resnet 18 和 FCN-LSTM 模型
- 小数据集已跑完: **准确度 0.652,目前还在调参,应该还能涨点** 我们的方法 0.695

大数据集还差 20 个 epoch, 比较慢



3.写 Response 邮件 8页

下周计划

- 1. 修改论文
- 2. 对比大数据集跑完