

面向行业应用的机器学习 与智能信号处理研究

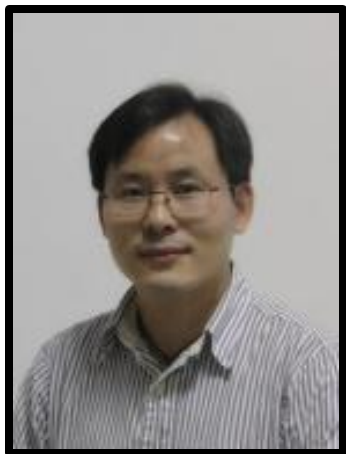
厦门大学信息学院/人工智能研究院

智能数据分析与处理实验室

<https://xmu-smartdsp.github.io>



廈門大學智能分析與處理實驗室



黃悅教授
博士生導師



肖珍龍副教授
碩士生導師



涂曉彤助理教授
碩士生導師

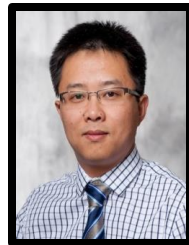
丁興號教授、博導
人工智能研究院雙聘PI
健康醫療大數據雙聘PI
愛思唯爾中國高被引學者
省百千萬人才、省新世紀人才



Columbia
University



Duke
University



UT Austin



華南理工大
學

主要的研究方向：機器學習、機器視覺、智能信息處理、邊緣智能



我们在做什么？

Credit to: IEEE SPS

signalprocessingsociety.org/our-story/signal-processing-101



OUR STORY

PUBLICATIONS & RESOURCES

CONFERENCES & EVENTS

COMMUNITY & INVOLVEMENT

PROFESSIONAL DEVELOPMENT

FOR VOLUNTEERS



- > Our Members
- > Society History
- > State of the Society
- > SPS Branding Materials
- > Contact Us
- > Feedback
- > SPS Staff
- > Sitemap

机器学习与信号处理

Recently, machine learning techniques have been applied to aspects of signal processing, blurring the lines between the sciences, and causing many shared applications between the two.

-- IEEE SIGNAL PROCESSING SOCIETY



Have you shared "what's on your mind" with Facebook friends or searched the Internet today? These are just two of the myriad ways we use machine learning every day. Machine learning brings together computer science and statistics to harness predictive power, and provides the technology behind other applications, including detecting credit card fraud, medical diagnostics, stock market analysis and speech recognition.

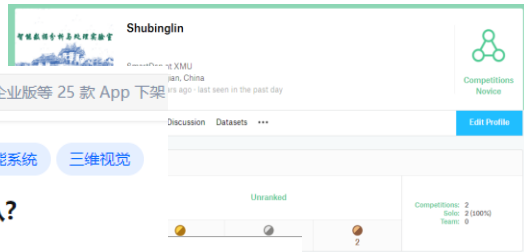
Recently, machine learning techniques have been applied to aspects of signal processing, blurring the lines between the sciences, and causing many shared applications between the two.

廈門大學智能分析與處理實驗室

我們已經拿了



KAGGLE



国内有哪些不错的CV(计算机视觉)团队?

<https://www.zhihu.com/question/332075078/answer/800613546>

知乎

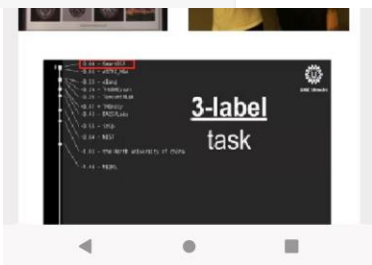
首页 会员 发现 等你来答

滴滴企业版等 25 款 App 下架

廈門

廈門大學：紀榮嶸（媒体分析与计算实验室），丁兴号（智能数据分析与处理实验室）

MICCAI 医学图像分析



CVPR
自动驾驶

Rank	Team	Time	Score	Task
4	Xpeng	2019-05-30 20:48:41 PDT	0.8181 0.3321 0.1717	car+van: 0.6352; bus: 0.4871; truck: 0.9177; person: 0.1120; motorcycle+open-tricycle+closed-tricycle: 0.0861
5	BUPT	2019-05-30 22:43:11 PDT	0.3335 0.32375 0.1068	car+van: 0.6235; bus: 0.4546; truck: 0.5059; person: 0.1068; motorcycle+open-tricycle+closed-tricycle: 0.0851
6	lab	2019-05-31 01:18:41 PDT	0.32001 0.32352 0.1668	car+van: 0.6235; bus: 0.4546; truck: 0.5059; person: 0.1068; motorcycle+open-tricycle+closed-tricycle: 0.0851
7	www.marslab	2019-05-31 09:45:43 EDT	0.32751 0.32291 0.16101	car+van: 0.6235; bus: 0.4812; truck: 0.4926; person: 0.1068; motorcycle+open-tricycle+closed-tricycle: 0.0851





毕业生去哪儿？

01

研究生毕业后去向包括-大厂系列

百度、腾讯、阿里巴巴、字节、寒武纪、海康、云从、华为、平安科技

02

研究生毕业后去向包括-事业单位系列

公务员、国家电网、厦航、专利局、高校、通信运营商、证券和银行等

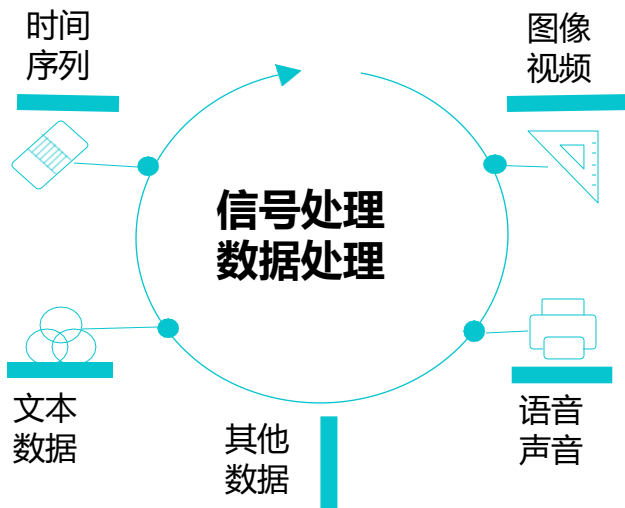
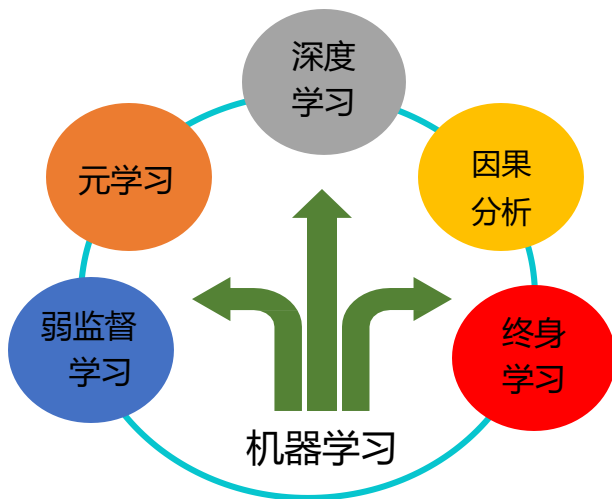
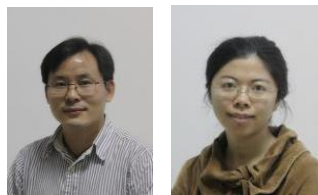
研究生毕业后去向包括-继续深造系列

UT-Austin、新加坡国立大学、香港中文、香港大学、西湖大学、中国科学技术大学教职等

03

廈門大學智能分析與處理實驗室

教授2名，核心科研人員共4名，碩博士60余人。在機器視覺、遙感圖像分析、信號處理等已有多年積淀。已在TPAMI、IJCV、IEEE TIP、IEEE TSP、IEEE TNNLS、Signal Processing、CVPR、ICCV、AAAI 等領域重要期刊和會議上發表學術論文百余篇。負責人論文Google引用1萬+次，h指數:47。主要從模型和數據層面開展深度學習的落地應用研究，具體包括領域/行業大模型研究、邊緣智能、AI for Science及其在語音、圖像處理的應用。





廈門大學智能分析與處理實驗室

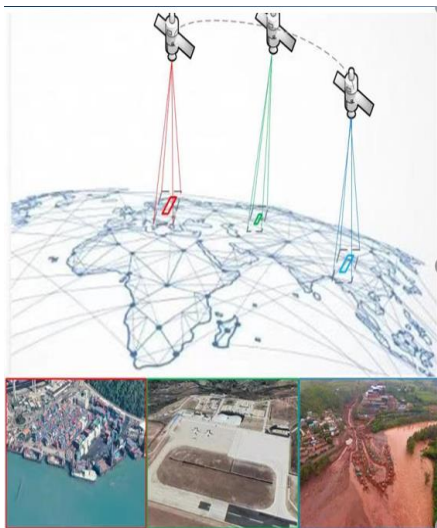
主要围绕**小数据、大任务场景下的低资源机器学习、全链条视角下的探测和声成像设备声像仪开发**。开展机器学习在**工业质检、巡检、卫星对地观测、雷达信号分析**等落地应用相关研究，部分研究已经产业化。

恶劣成像环境下的图像/视频
质量提升、智能分析与理解

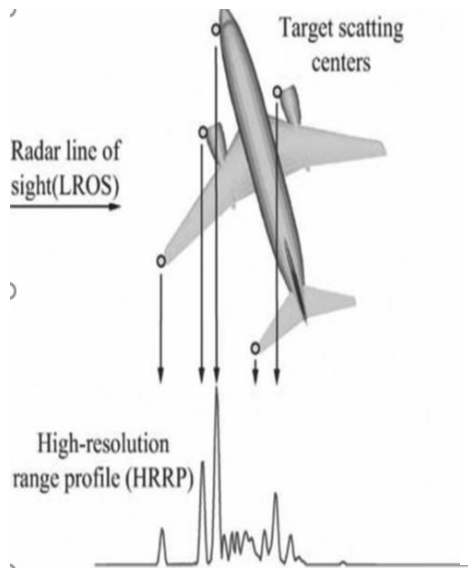
医学影像的智能计算
成像与分析

面向军民融合场景的低资源鲁
棒机器学习方法研究

计算声成像技术



目标识别



雷达信号分析

Stable
Diffusion
(Vanilla)



PACS
(Dataset)



Ours
(Finetune)



可控AIGC



廈門大學智能分析與處理實驗室

主要围绕**小数据、大任务场景下的低资源机器学习、全链条视角下的探测和声成像设备声像仪开发**。开展机器学习在**工业质检、巡检、卫星对地观测、雷达信号分析**等落地应用相关研究，部分研究已经产业化。

恶劣成像环境下的图像/视频
质量提升、智能分析与理解

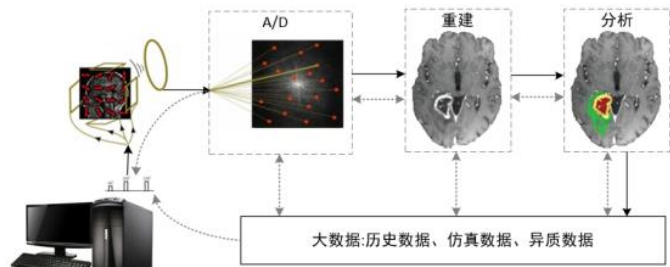
医学影像的智能计算
成像与分析

面向军民融合场景的低资源鲁
棒机器学习方法研究

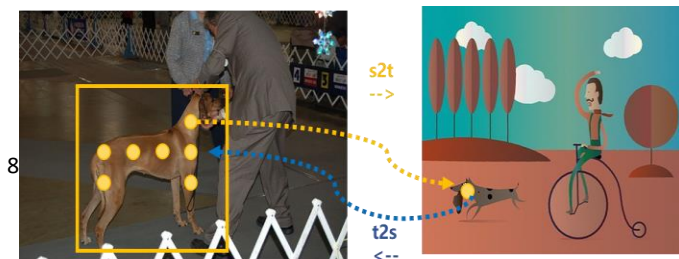
计算声成像技术



雨天图像质量提升



医学计算成像



域适应目标检测算法



廈門大學智能分析與處理實驗室

	項目名稱	來源	合同經費
1	基於***探測目標智能識別研究*	JKW基礎加強基礎重點項目-課題	430萬
2	***信號的通用表征	166項目	190萬
3	面向雷達信號的小樣本目標魯棒識別*	聯合基金/重點項目	102萬
4	基於偽健康圖像生成的知識嵌入深度學習病灶分割方法研究	面上項目	55萬
5	面向交通運輸綜合執法的視覺大數據分析軟件平台	企業委託	200萬
6	AOI Dimple AI判讀系統開發	企業委託	47萬



廈門大學智能分析與處理實驗室

溫馨和睦的大家庭



團隊每年至少組織兩次集體活動

廈門大學智能分析與處理實驗室

充滿激情又安靜的工作氛圍



實驗室設立本科生工作室、硬件實驗室以及配備免費咖啡

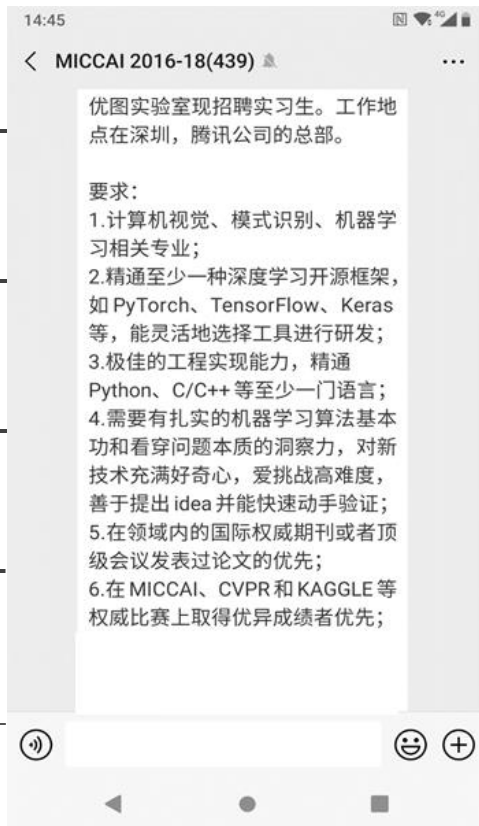
目录

培养目标



廈門大學智能分析與處理實驗室

企業的要求。。





廈門大學智能分析與處理實驗室

企業的要求。。



模式识别、机器学习的浓厚兴趣

精通至少一种深度学习开源框架

极佳的工程实现能力

在领域内的国际权威期刊会议发表过论文的优先

ICML/Neurips/CVPR/ICCV/ECCV/KDD/ACL

在CVPR、KAGGLE和MICCAI等权威比赛上

取得优异成绩



廈門大學智能分析與處理實驗室

我們希望你們擁有



對機器學習理論及等相關應用領域
具有濃厚興趣
較強的分析實際問題的能力

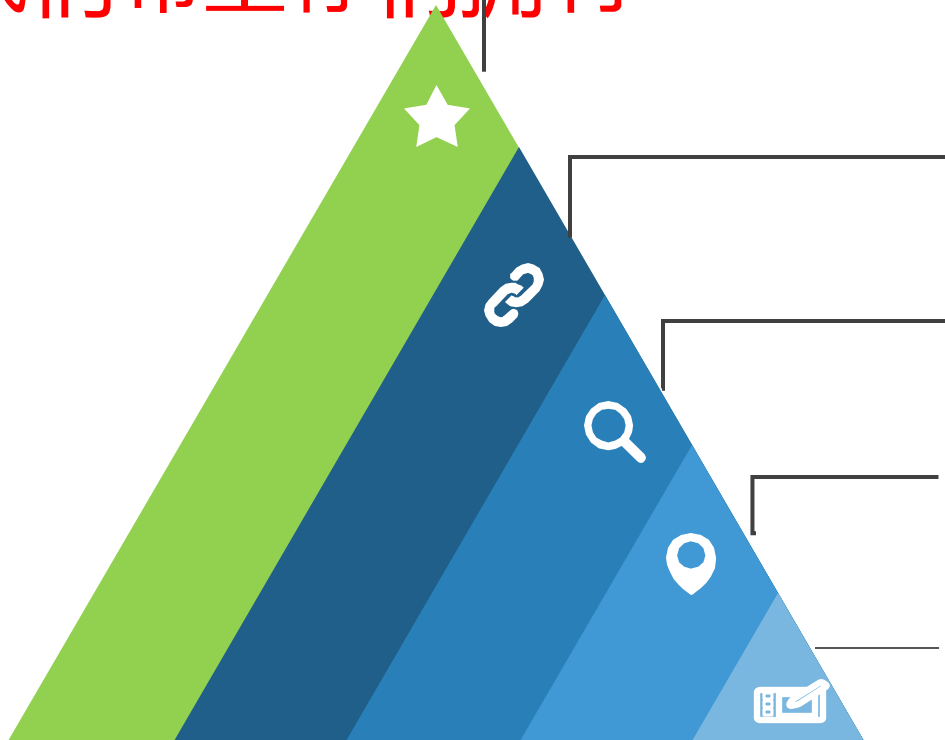
良好的軟硬件綜合能力（我系特色）

良好的數理基礎和專業英語運
用能力



廈門大學智能分析與處理實驗室

我們希望你們擁有



相關領域具

力

通信系特

語運

優先 (企

08:49



← 人工智能是当前最好的计算机专业吗? 写回答 查看全部 93 个回答 >



周博磊

机器学习等话题优秀回答者

已关注

难道大家忘了若干年前的“21 世纪是生物的世纪” doge.

说太多得罪人，简言之，AI本身的内容博大精深，但可惜很少有学校/学院有师资能把这里面的东西融汇贯通作为基础课程教给本科生。所以本科同学不如多花时间把数理基础打好，以及兼顾提升编程能力。目前我还是希望自己招的学生是数学物理自动化电子通讯等本科背景，或者计算机非AI方向。

提早炼丹容易走火入魔，自毁前程，慎重。

另：本科对我影响最大的一门课是信号系统。额，课其实只去上过一次还是两次，不如说是自学Oppenheim's signals and systems课本，基本支撑起了我对这个世界最初的量化认知。以及后来的Gonzalez's digital image processing，现在做CV的一些小朋友可能连这都没听过/读过。。。

编辑于 2019-07-11 · 著作权归作者所有

▲ 赞同 490 ▼

❤ 感谢 39

☆ 收藏

💬 评论 50

Join us



期待您的加入

厦门大学智能数据分析与处理实验室

<https://xmu-smartdsp.github.io>

丁兴号老师



dxh@xmu.edu.cn

6号楼211-5, 18959286868

黄悦老师



yhuang2010@xmu.edu.cn

6号楼211-3, 15959232083

肖珍龙老师



zlxiao@xmu.edu.cn

6号楼211-2, 15359221512

涂晓彤老师



xttu@xmu.edu.cn

6号楼211-1, 18201826883

正在进行的项目



研究背景

什么是大规模语言模型？

大规模语言模型（英文：Large Language Model，缩写LLM），是一种人工智能模型，旨在理解和生成人类语言。它们在大量的文本数据上进行训练。LLM的特点是**规模庞大，包含数十亿的参数**。



- 2020年美国国会将推动**AI模型**进入联邦政府运作，形成《政府AI法案》。
- 2019-2020美国白宫接连以两道总统行政令强调将在未来以**通用大模型**在内的AI突破领域资助力度以实现美国AI领域的领导权。
- GPT&ChatGPT 某种程度对中国禁用，A100&H100等**高端AI芯片**对中国禁售。

中美科技对抗

研究背景

什么是大规模语言模型?

大规模语言模型 (英文: Large Language Model, 缩写LLM), 是一种人工智能模型, 旨在理解和生成人类语言。它们在大量的文本数据上进行训练。LLM的特点是**规模庞大, 包含数十亿的参数**。



2018年

GPT-1

GPT-2

GPT-3

Chat-GPT

GPT-4

2023年

OpenAI

T5

.....

Google

BERT

.....

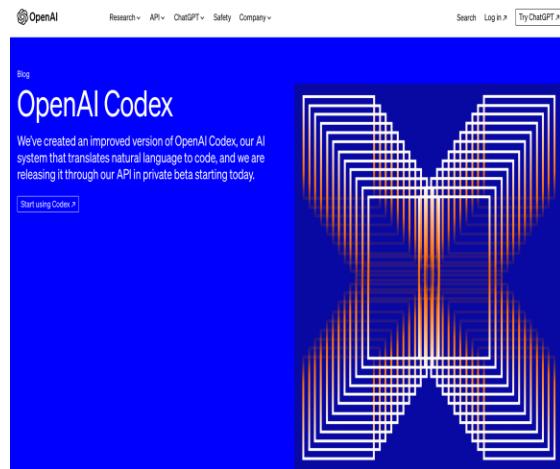
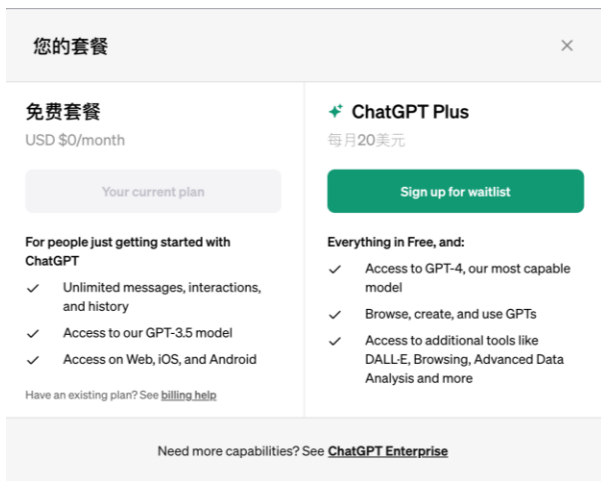
Google

大模型相关领域井喷式发展

研究背景

什么是大规模语言模型?

大规模语言模型（英文：Large Language Model，缩写LLM），是一种人工智能模型，旨在理解和生成人类语言。它们在大量的文本数据上进行训练。LLM的特点是**规模庞大，包含数十亿的参数**。



成熟商业应用



研究背景

什么是大规模语言模型?

U If you sum up 1 to 100 but exclude all primes, what is the result?



To solve this problem, we should sum all numbers from 1 to 100 and then subtract the sum of all prime numbers within that range.

Firstly, the sum of all integers from 1 to 100 can be calculated using the formula for the sum of an arithmetic series:

$$n*(n + 1)/2,$$

where n is 100 in this case:

$$100*(100 + 1)/2 = 5050.$$

Then we calculate the sum of all prime numbers from 1 to 100. The prime numbers between 1 and 100 are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97. If we sum these up:

$$2 + 3 + 5 + 7 + 11 + 13 + 17 + 19 + 23 + 29 + 31 + 37 + 41 + 43 + 47 + 53 + 59 + 61 + 67 + 71 + 73 + 79 + 83 + 89 + 97 = 1060.$$

Then subtract this from the total sum:

$$5050 - 1060 = 3990.$$

So the sum of all numbers from 1 to 100 excluding all prime numbers is 3990.



研究背景

什么是大规模语言模型？



U What is funny about this image? Describe it panel by panel.



The image shows a package for a "Lightning Cable" adapter with three panels.

Panel 1: A smartphone with a VGA connector (a large, blue, 15-pin connector typically used for computer monitors) plugged into its charging port.

Panel 2: The package for the "Lightning Cable" adapter with a picture of a VGA connector on it.

Panel 3: A close-up of the VGA connector with a small Lightning connector (used for charging iPhones and other Apple devices) at the end.

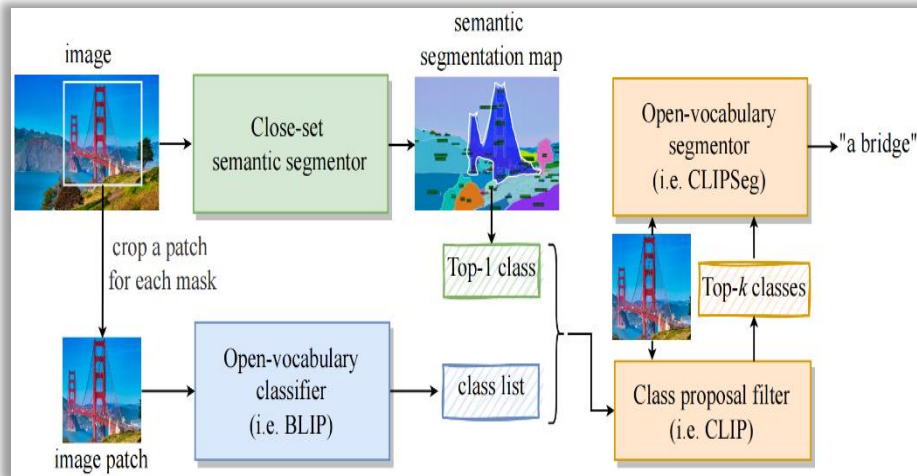
The humor in this image comes from the absurdity of plugging a large, outdated VGA connector into a small, modern smartphone charging port.

一定程度地从感知跃升到认知

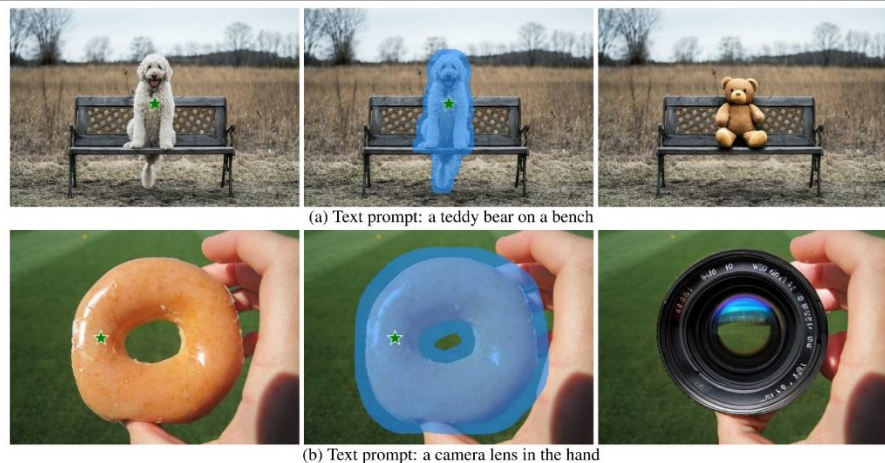
研究背景

什么是大规模语言模型?

语义分割



图像填充



Segment Anything

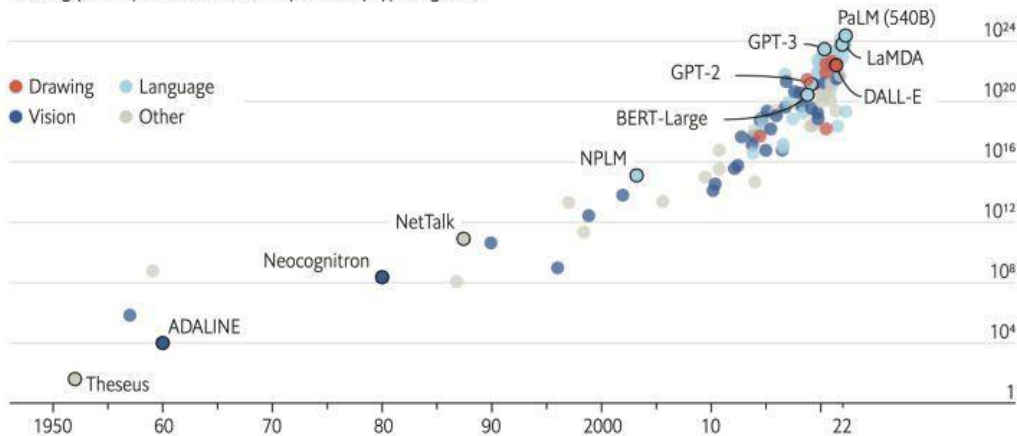


研究背景

什么是大规模语言模型？

The blessings of scale

AI training runs, estimated computing resources used
Floating-point operations, selected systems, by type, log scale



Sources: "Compute trends across three eras of machine learning", by J. Sevilla et al., arXiv, 2022; Our World in Data

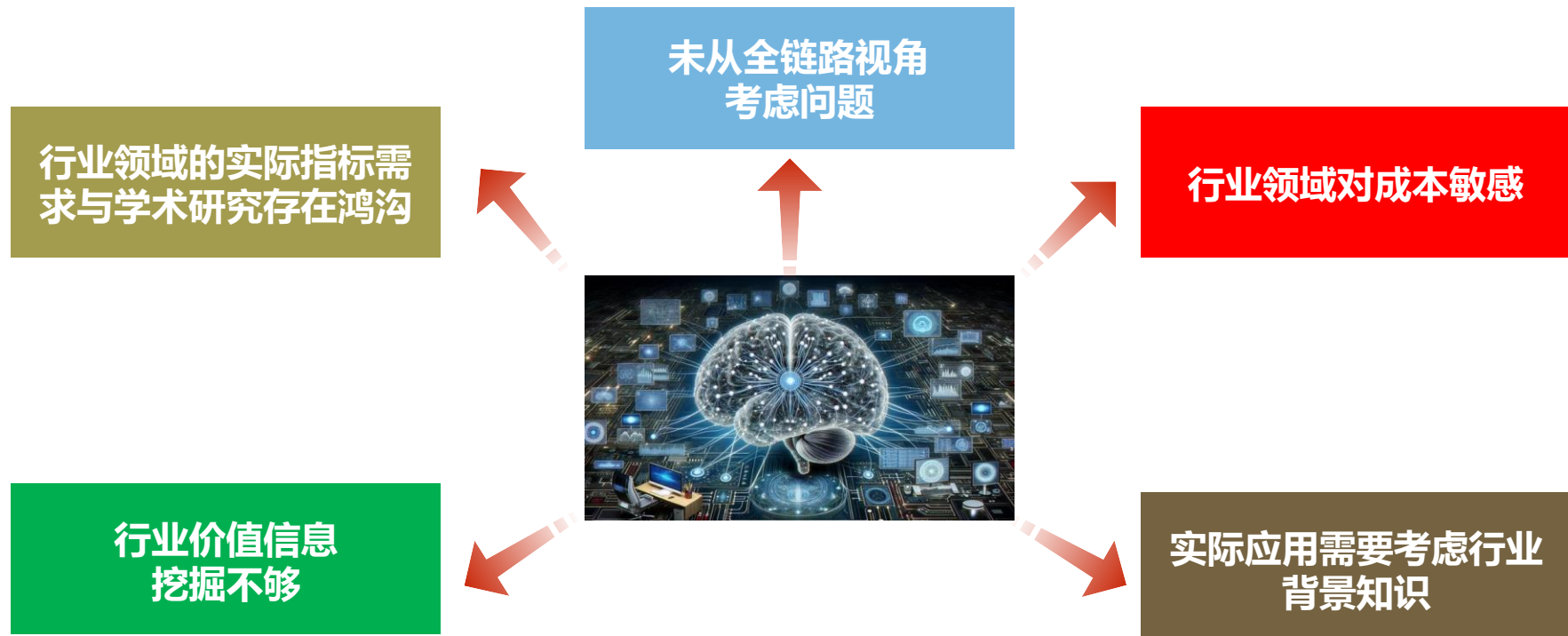


大模型在各个领域都取得了瞩目的成绩，那么大模型会一统天下吗？



研究背景

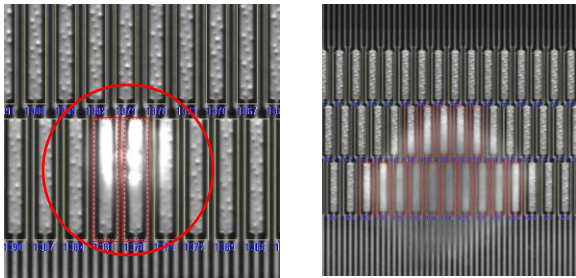
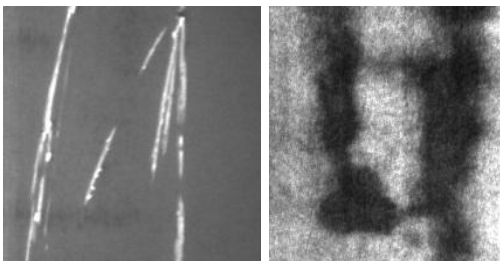
目前AI的挑战?





研究背景

行業領域的實際指標需求與學術研究存在鴻溝



钢板表面缺陷

车载显示屏

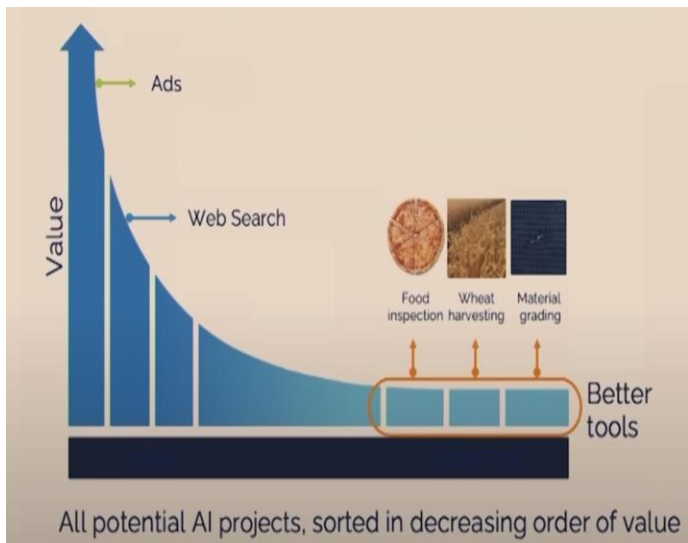
新能源

缺陷样本少，大量的正常样本，要求漏检率接近0



研究背景

行業領域對成本敏感



为什么AI没有广泛应用



声学相机



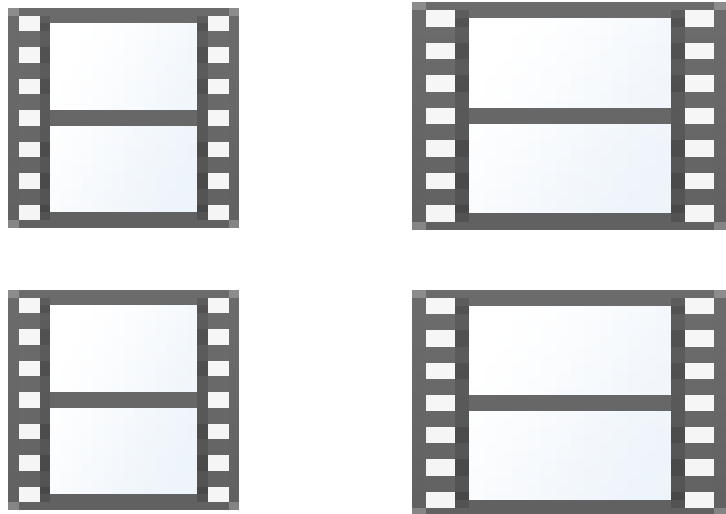
智能会议系统

不同行業對成本的敏感性存在差異



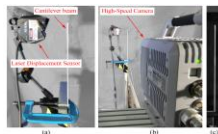
研究背景

行业价值信息挖掘不够



运动放大

原视频



放大后



30.5 Hz



194.5 Hz



535.1 Hz



1056.8 Hz



1747.5 Hz

悬臂梁各阶模态

提取振动，非接触测量

定义新场景实现算法的应用价值



研究背景

实际应用需要考虑行业背景



电子战-感知能力

俄乌战争启示-北约情报支持

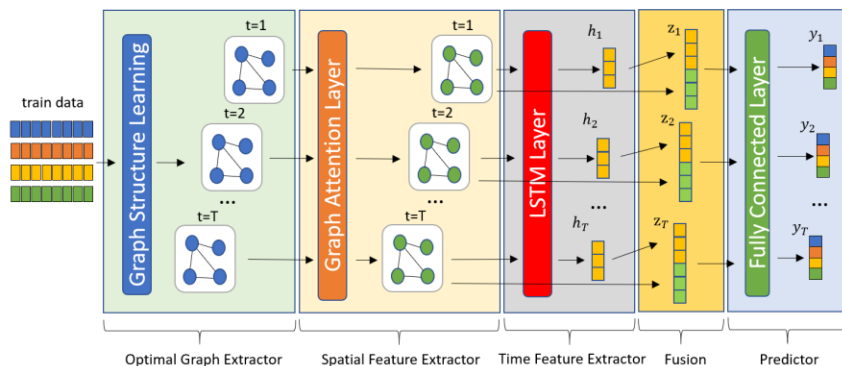
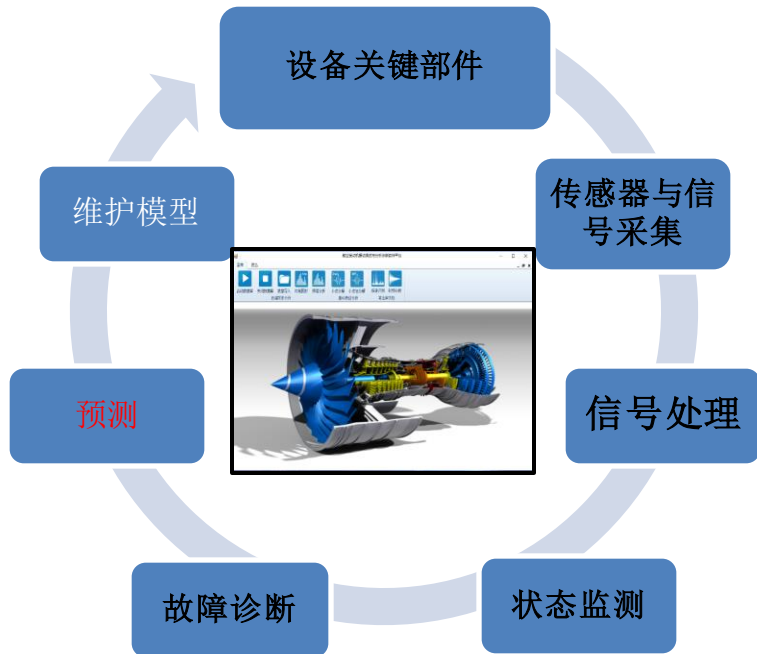
知识和数据驱动



研究背景

未从全链路视角考虑问题

全链路状态感知



从采集 → 感知 → 决策闭环

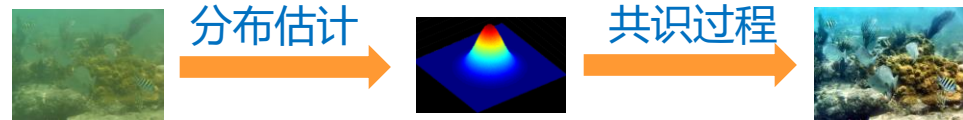
图像增强



廈門大學
XIAMEN UNIVERSITY



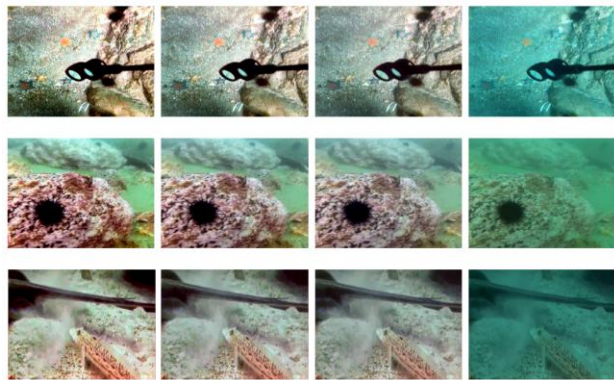
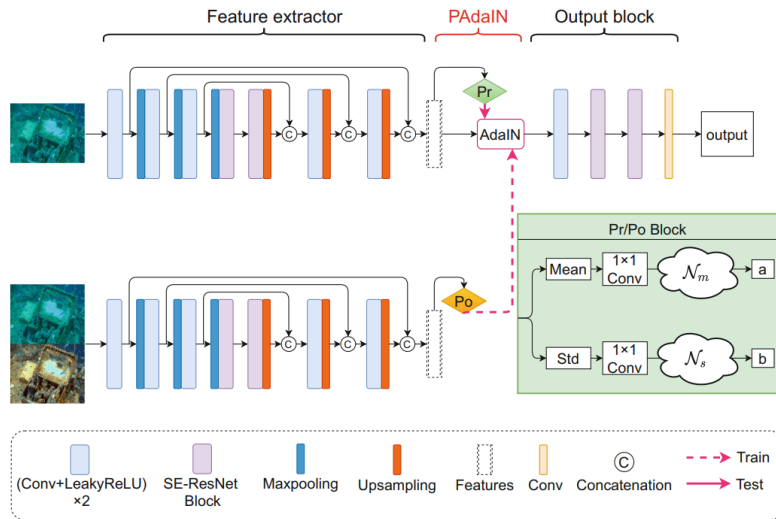
不确定性启发的水下图像增强



PUIE-Net 首个基于概率网络的水下图像增强方法

Zhenqi Fu, Wu Wang, Yue Huang, Xinghao Ding*, Kai-Kuang Ma. Uncertainty Inspired Underwater Image Enhancement[C]//European Conference on Computer Vision (ECCV), 2022: 465-482.

<https://github.com/zhenqifu/PUIE-Net>





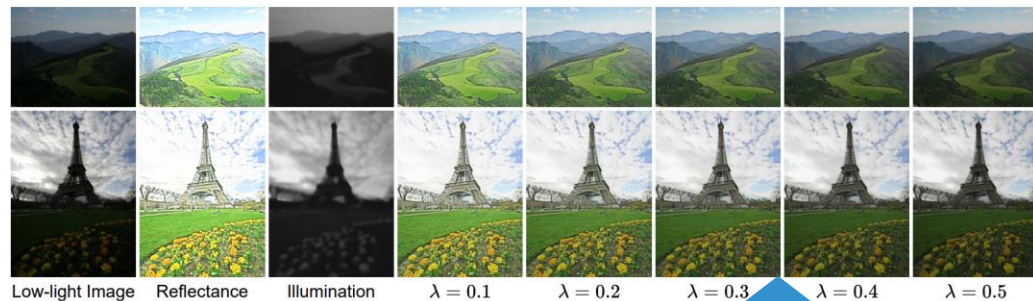
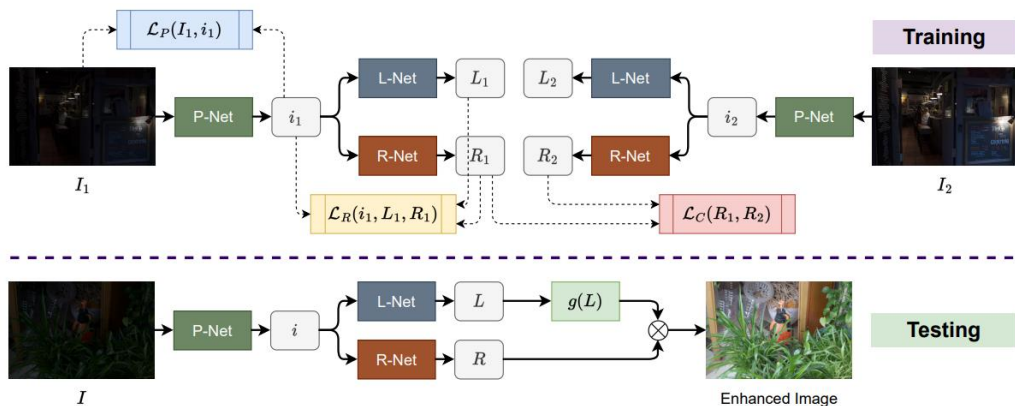
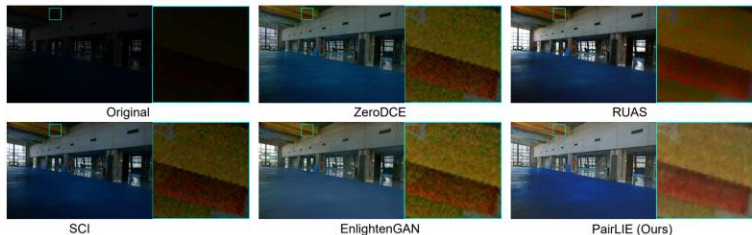
自監督弱光照图像增强

现有方法：低光照-正常光照

提出方法：低光照1-低光照2

Zhenqi Fu, Yan Yang, Xiaoteng Tu*, Yue Huang, Xinghao Ding, Kai-Kuang Ma. Learning a Simple Low-light Image Enhancer from Paired Low-light Instances. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 22252-22261, 2023.

<https://github.com/zhenqifu/PairLIE>

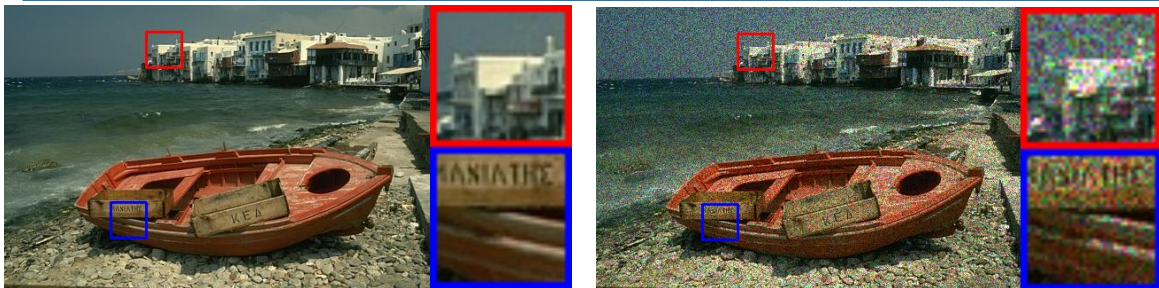


更佳的光照调整和噪声抑制

光照和反射分量可视化

基于自校验深度学习模型的无监督图像去噪方法

针对不同设定条件（数据、噪声）设计相应的高性能的无监督去噪算法，以尽可能大范围地覆盖实际中的去噪问题，以去噪算法为基础，探索更具一般性的无监督图像复原框架。



Ground-truth

Noise



U-Net

IDDP



隐神经表示方法

水下伪装目标分割



廈門大學
XIAMEN UNIVERSITY



基于频域增扩的水下伪装目标分割

- 伪装是自然界广泛存在的现象



显著目标



伪装目标

分割目标更加困难



退化风格多样



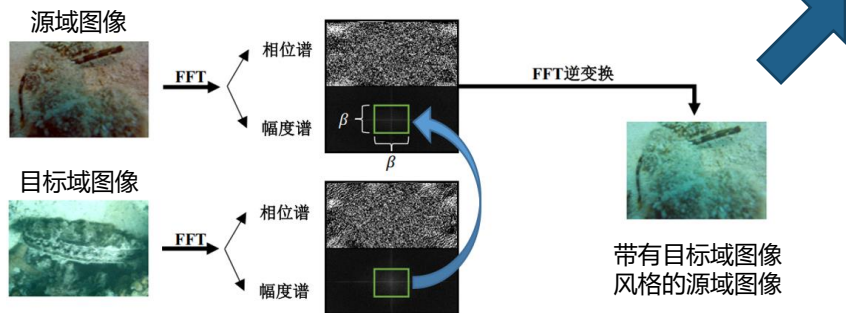
目标伪装特性

陆上图像

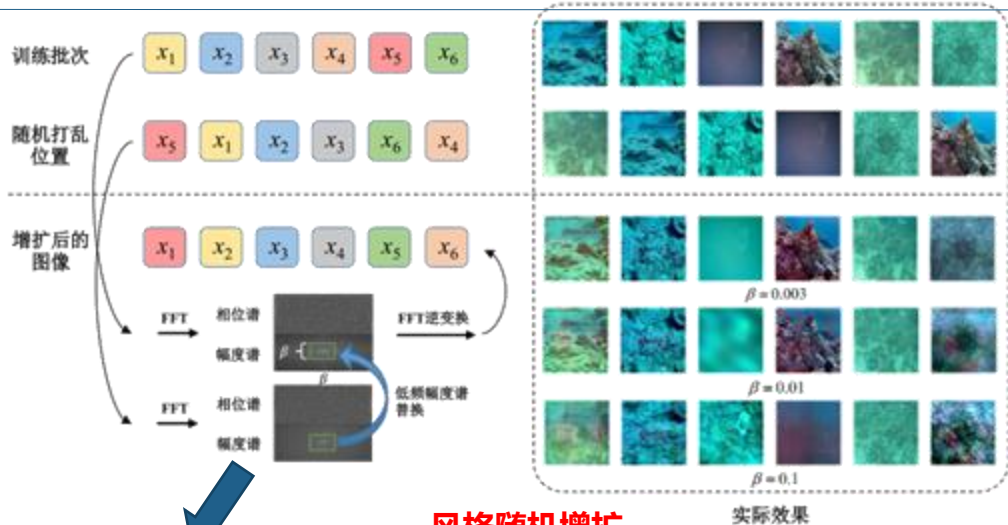
水下图像



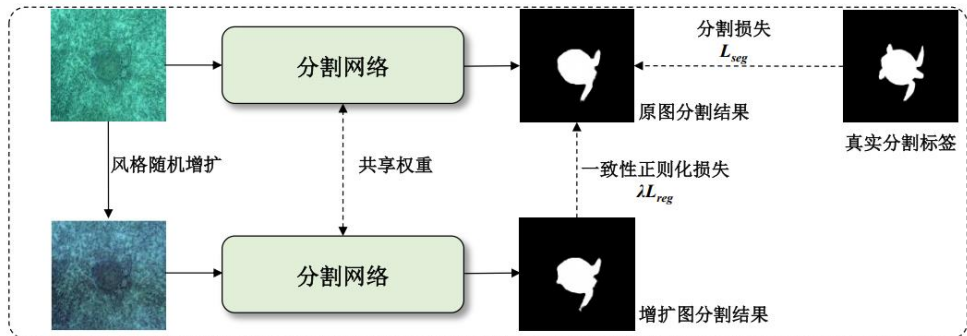
基于频域增扩的水下伪装目标分割



风格增扩



风格随机增扩

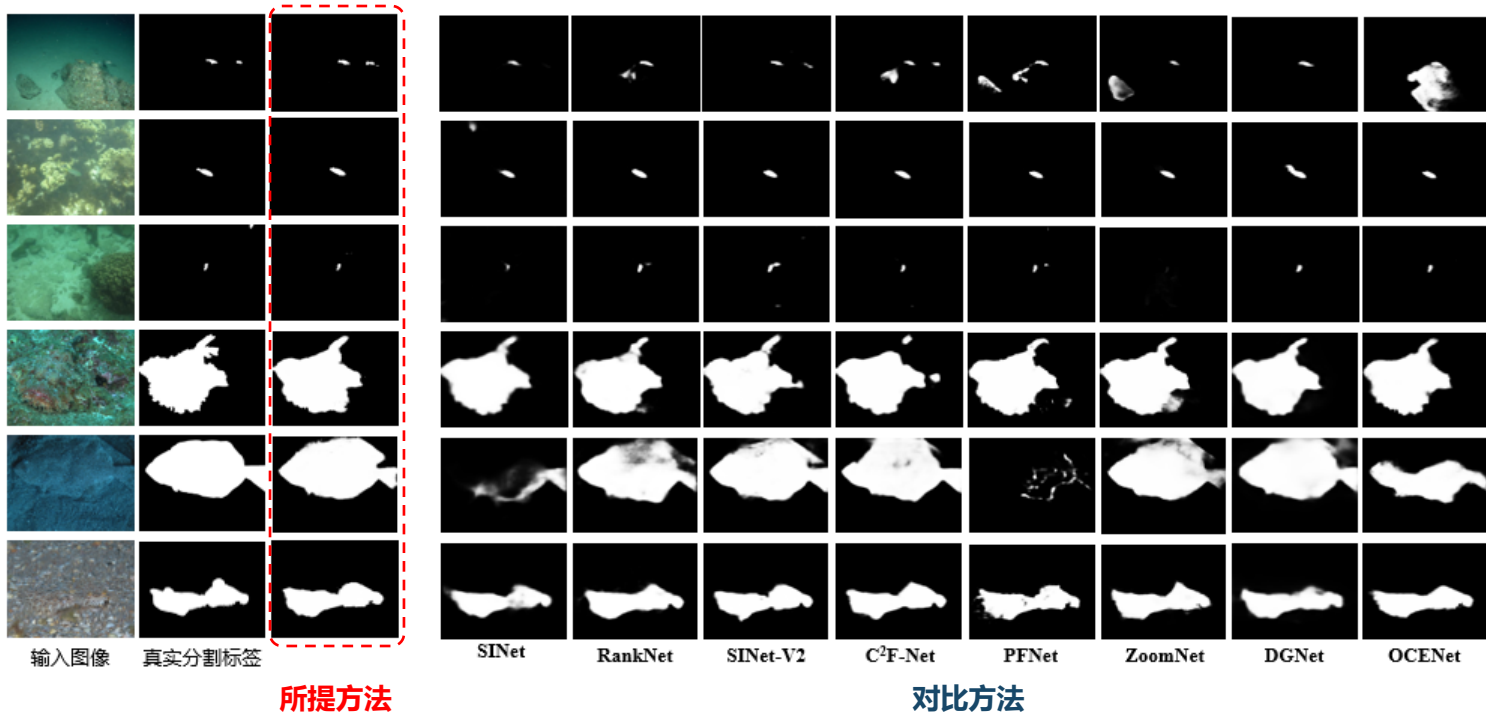


同源一致性约束训



基于频域增扩的水下伪装目标分割

- 在水下图像上的分割效果



域泛化车辆重识别



廈門大學
XIAMEN UNIVERSITY

研究背景

数据 ≠ 信息, 信息 ≠ 知识

Camera Networks in Smart City



More than half of all big data

Candidates for Big Data



■ Big data, fast growth

□ 30 millions camera have been deployed in China

■ Complicated content and rich Information

□ Persons, cars and objects

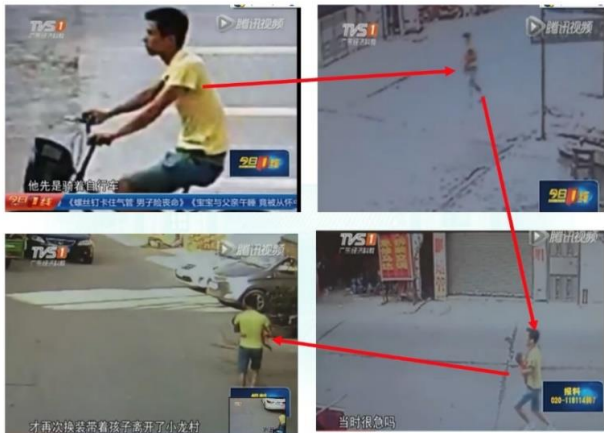


研究背景

什么是重识别?

重识别 (Re-Identification, 简称 **Re-ID**) 技术主要解决跨摄像头或跨场景下行人或车辆的识别与检索。

该技术可以作为人脸识别技术的重要补充, 可以对无法获取清晰拍摄人脸的行人进行跨摄像头连续跟踪, 增强数据的时空连续性, 也可以作为人脸识别技术的补充, 并相互作用, 应用于更多场景。



跨摄像头跟踪黄衣服男子的行进轨迹



识别戴白色帽盔者



面向未知场景下的域泛化重识别

将预训练好的模型直接应用于Re-ID在实际应用场景时，行人/车辆的识别性能会明显下降，且这种性能的损失会受到摄像头的视角差异、光照强度、遮挡等因素的影响。研究域泛化的Re-ID可以显著缓解域偏移所造成的模型精度损失，提高预训练模型在未知场景下的泛化能力。

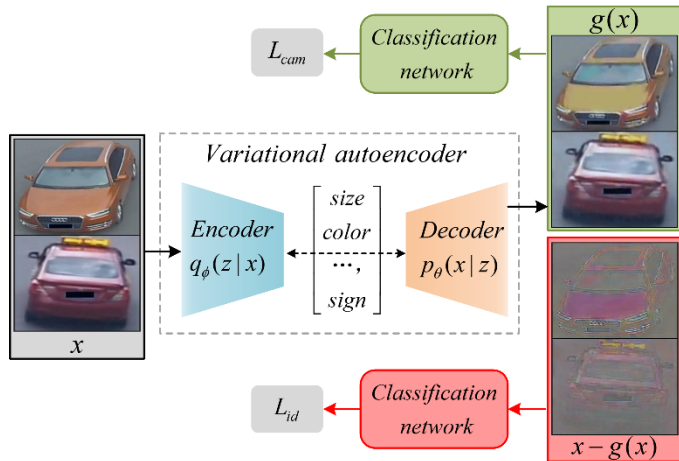
我们构建了细粒度Opri卡车数据集

- 解决现有车辆公共数据集的车辆类型的长尾分布问题



联合图片-特征双重解耦的域泛化重识别方法

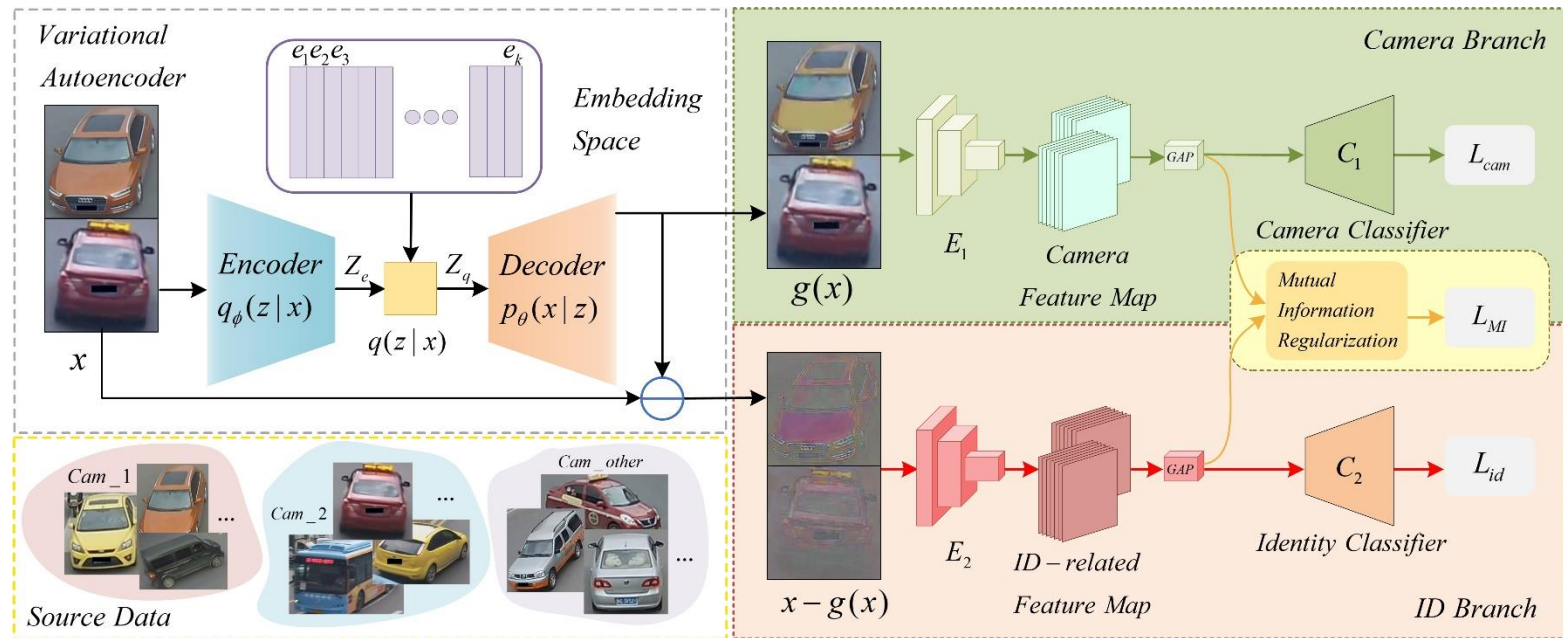
- 以信息瓶颈的视角去除图片冗余信息





▶ 联合图片-特征双重解耦的域泛化重识别模型

网络结构:





实验设置:

数据集: 公共源域训练数据集使用**VeRi-776**与**VehicleX**, 目标数据集为卡车数据集**Opri**。

预处理阶段: 将输入图片全部重新调整为 256×256 的大小。对训练的图片采用随机翻转、裁剪方式进行数据增扩。

网络架构: 以**resnet-50**作为两个神经网络的骨干网络, 分别提取车辆相关和无关的特征。

UDA Methods		VeRi-776→Opri			
		mAP	rank-1	rank-5	rank-10
SpCL[19]	NIPS'2020	40.8	46.8	59.8	65.5
AWB-Pre[97]	TIP'2022	28.3	36.4	47.4	51.9
AWB-Post[97]	TIP'2022	25.2	32.9	43.2	48.1
DG Methods		VeRi-776→Opri			
		mAP	rank-1	rank-5	rank-10
Baseline[1]	TMM'2019	18.3	27.8	35.9	39.7
IS-GAN[25]	TPAMI'2019	26.4	37.8	47.1	51.4
MetaBIN[64]	CVPR'2021	22.0	32.5	41.9	45.9
Mixstyle[82]	ICLR'2021	26.4	36.2	47.1	51.6
OSNet-AIN[94]	TPAMI'2021	27.3	37.9	48.3	52.6
QAConv-GS[69]	CVPR'2022	35.2	44.6	56.4	61.0
TransMatcher[68]	NIPS'2021	34.7	42.8	56.0	61.1
JIFD	This work	25.3	35.0	45.5	49.9
TransMatcher[68]+JIFD		41.2	50.2	63.1	67.4

UDA Methods		VehicleX→Opri			
		mAP	rank-1	rank-5	rank-10
SpCL[19]	NIPS'2020	41.5	47.2	61.0	66.4
AWB-Pre[97]	TIP'2022	29.7	37.6	49.1	54.1
AWB-Post[97]	TIP'2022	26.3	34.6	45.4	50.4
DG Methods		VehicleX→Opri			
		mAP	rank-1	rank-5	rank-10
Baseline[1]	TMM'2019	19.7	29.1	38.9	42.7
MetaBIN[64]	CVPR'2021	21.7	31.0	41.7	45.8
Mixstyle[82]	ICLR'2021	20.6	29.4	40.4	44.8
OSNet-AIN[94]	TPAMI'2021	14.0	21.5	30.2	34.1
TransMatcher[68]	NIPS'2021	28.9	36.7	50.4	55.7
JIFD	This work	23.0	32.8	43.1	47.4
TransMatcher[68]+JIFD		30.9	37.5	52.1	57.9



异常检测

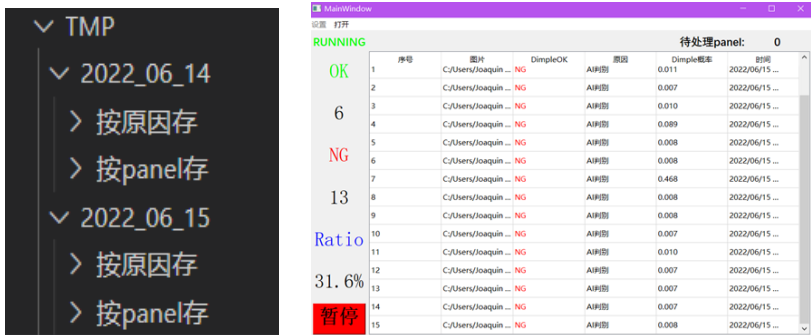
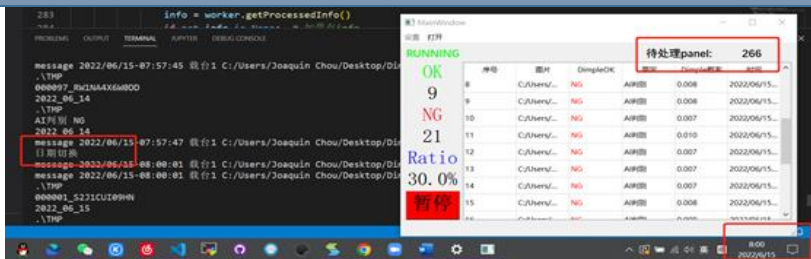


廈門大學
XIAMEN UNIVERSITY

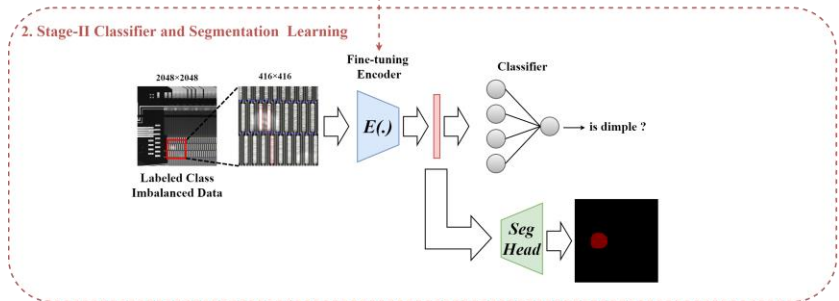
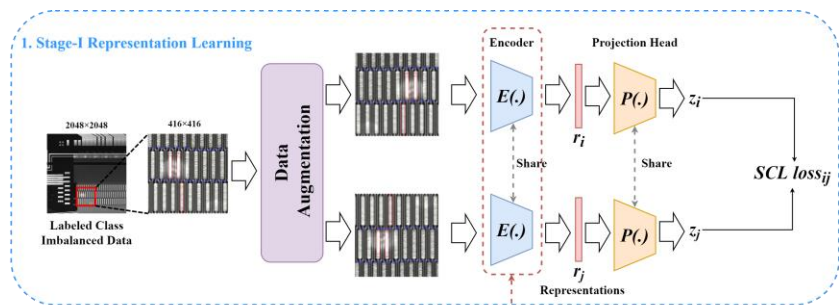
AOI Dimple AI判读系統開發

實現了基於深度學習的顯示屏電路導電粒子在線質檢系統，已經成功部署於友達光電廈門廠區，在**26**條生產線應用。極大提升了生產效率。該項目通過提出新的自監督對比學習方案，有效解決**標注樣本稀缺**、**樣本類別分布不均衡**等問題，大幅度提升了診斷精度。目前正與友達光電合作推廣至其它分公司。

軟件界面

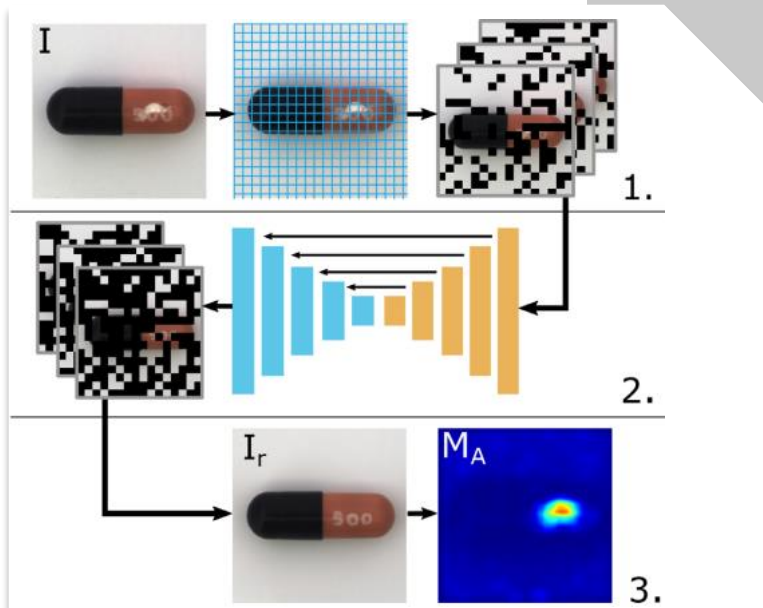


網絡框架

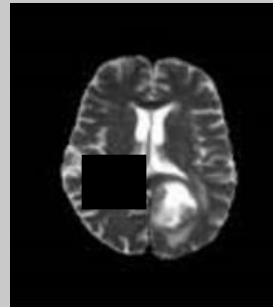
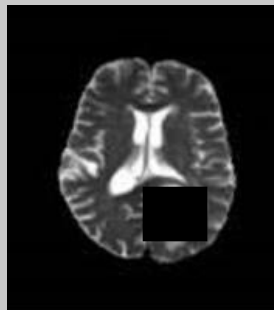




基于自适应傅里叶空间压缩的异常检测



现有基于“减法”的重建方法都是在图像空间域进行不同的掩码操作再编码重建达到伪健康生成的目的。





基于自适应傅里叶空间压缩的异常检测

通过什么方式才能保证一定会丢掉部分的病灶信息呢？





基于自适应傅里叶空间压缩的异常检测



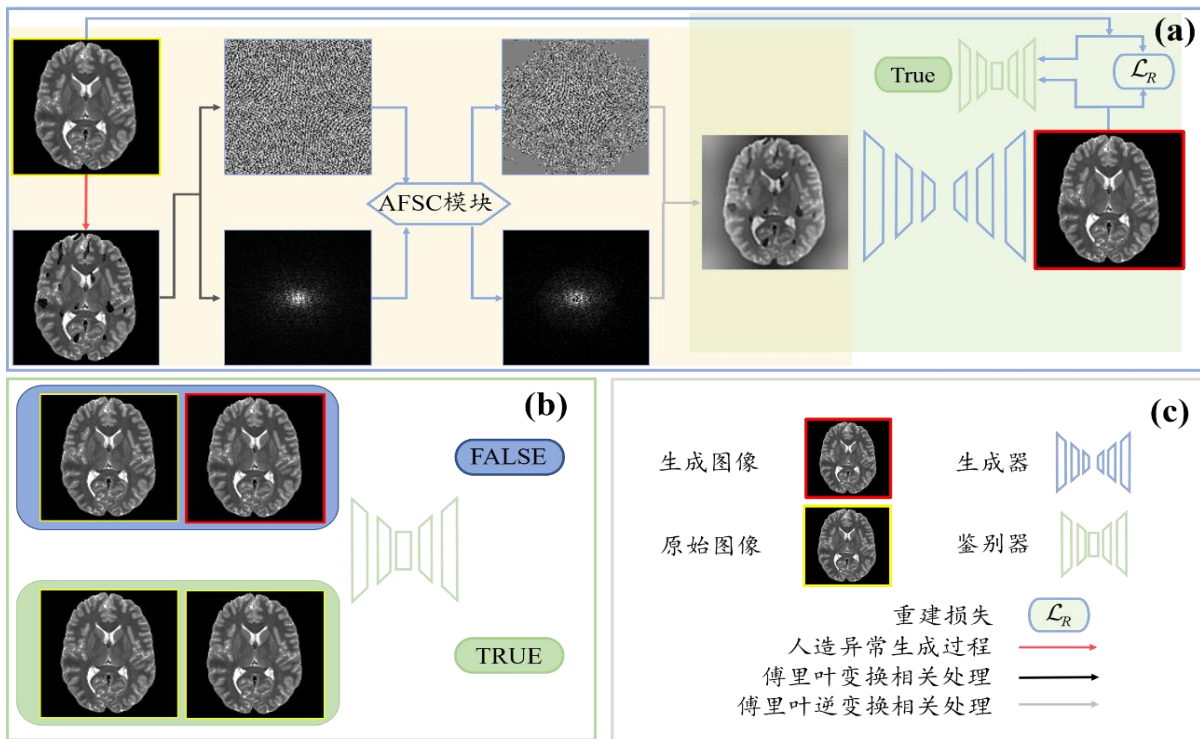
依据帕塞瓦尔定理 (Parseval's theorem) :

傅里叶空间中每个系数均包含着图像的全局空间信息!





基于自适应傅里叶空间压缩的异常检测



基于自适应傅里叶空间压缩 (Adaptive Fourier Space Compression, AFSC) 的异常检测方法架构图



基于自适应傅里叶空间压缩的异常检测

表 2.2 BraTS 数据集各方法对比实验性能表

方法	指标				
	AUROC _{img}	AP _{img}	[DICE]	AUROC _{pix}	AP _{pix}
AnoGAN	0.827±0.0 28	0.775± 0.044	0.314±0.0 065	0.810±0.0 52	0.344± 0.082
AE	0.823±0.0 08	0.791± 0.008	0.353±0.0 005	0.795±0.0 07	0.383± 0.017
AE w/ <u>AFSCm</u>	0.847±0.0 04	0.830± 0.006	0.431±0.0 001	0.862±0.0 01	0.496± 0.003
DRAEM	0.853±0.0 19	0.792± 0.042	0.302±0.0 023	0.819±0.0 08	0.263± 0.027
DRAEM w/ <u>AFSCm</u>	0.841±0.0 02	0.729± 0.017	0.297±0.0 028	0.808±0.0 16	0.210± 0.019
DRAEM_G	0.852±0.0 03	0.835± 0.003	0.479±0.0 005	0.888±0.0 01	0.509± 0.006
DRAEM_G w/ <u>AFSCm</u>	<u>0.868±0.0</u> 07	<u>0.855±</u> <u>0.008</u>	0.500±0.0 009	<u>0.893±0.0</u> 04	0.559± 0.011
AFSC	0.877±0.0 05	0.858± 0.006	<u>0.490±0.0</u> <u>010</u>	0.899±0.0 05	<u>0.520±</u> <u>0.018</u>

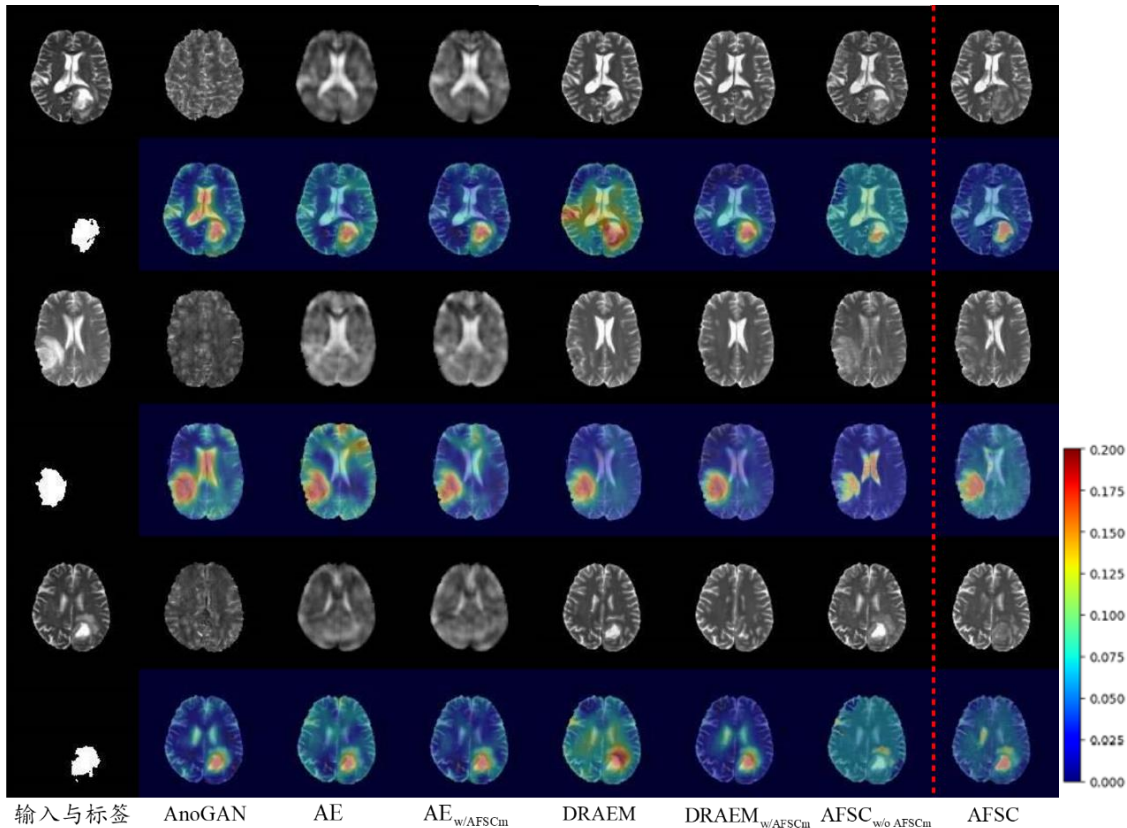
加入AFSC模块后所有方法的各项性能指标都有所提升。



基于自适应傅里叶空间压缩的异常检测

表 2.2 BraTS 数据集各方法对比实验性能表

方法	指标				
	AUROC _{img}	AP _{img}	[DICE]	AUROC _{pix}	AP _{pix}
AnoGAN	0.827±0.0	0.775±	0.314±0.	0.810±0.0	0.344±
	28	0.044	065	52	0.082
AE	0.823±0.0	0.791±	0.353±0.	0.795±0.0	0.383±
	08	0.008	005	07	0.017
AE w/ AFSCm	0.847±0.0	0.830±	0.431±0.	0.862±0.0	0.496±
	04	0.006	001	01	0.003
DRAEM	0.853±0.0	0.792±	0.302±0.	0.819±0.0	0.263±
	19	0.042	023	08	0.027
DRAEM w/AFSCm	0.841±0.0	0.729±	0.297±0.	0.808±0.0	0.210±
	02	0.017	028	16	0.019
DRAEM_G	0.852±0.0	0.835±	0.479±0.	0.888±0.0	0.509±
	03	0.003	005	01	0.006
DRAEM_G w/ AFSCm	<u>0.868±0.0</u>	<u>0.855±</u>	<u>0.500±0.</u>	<u>0.893±0.0</u>	<u>0.559±</u>
	<u>07</u>	<u>0.008</u>	<u>009</u>	<u>04</u>	<u>0.011</u>
AFSC	0.877±0.0	0.858±	0.490±0.	0.899±0.0	0.520±
	05	0.006	010	05	0.018



遥感场景下的旋转目标检测



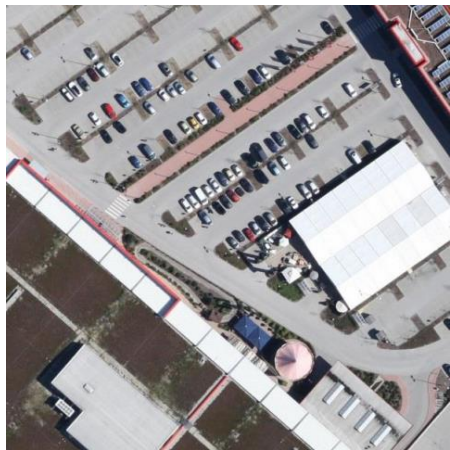
廈門大學
XIAMEN UNIVERSITY



研究背景

什么是旋转目标检测任务?

遥感图像因其特殊的俯视视角以及远距离拍摄，存在物体尺度差异大、小目标排列密集且方向性强等问题，利用定向边框来表示旋转目标，可为目标检测任务提供更完整的几何轮廓与位置信息



航拍车辆形态



水平标注



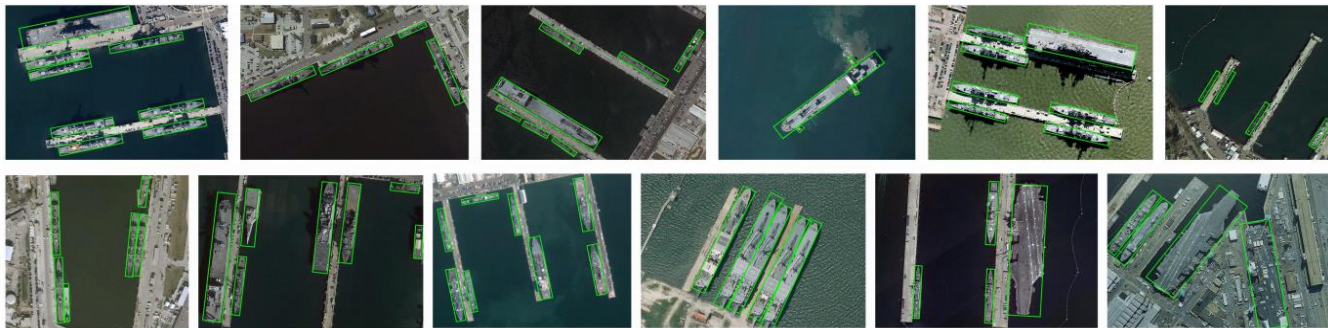
旋转标注



研究背景

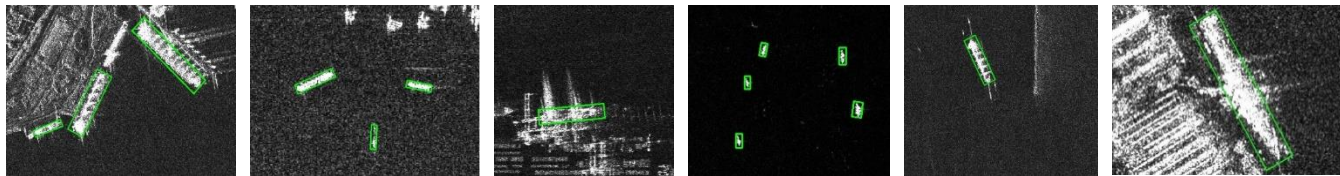
遙感圖像目標檢測難點——標籤的獲取成本高

光學遙感圖像



域適應

SAR圖像



域適應旋轉目標檢測任務具有理論研究意義和現實應用價值!!!



基于多尺度频-空信息融合的域适应旋转目标检测方法

Gabor滤波核与哺乳动物视觉皮层简单细胞2D感受野非常相似，具有优良的空间局部性和方向选择性，可捕获图像局部区域内的多尺度、多方向频率信息和局部结构特征，将其与卷积神经网络所提取的多尺度空间信息融合能够更好地挖掘不同尺度、不同方向的旋转目标信息

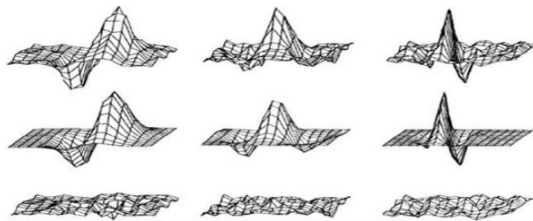
多尺度、多方向目标的感知不确定性



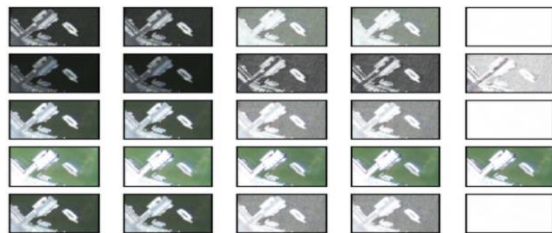
融合多尺度、多方向的Gabor频域信息与多尺度空间像素信息



基于多尺度频-空信息融合的域适应旋转目标检测方法



Gabor滤波器与视觉皮层感受野对比图



5尺度、5方向的Gabor特征

光学遥感卫星影像解译



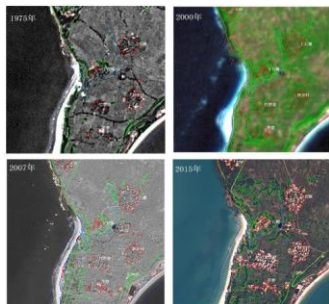
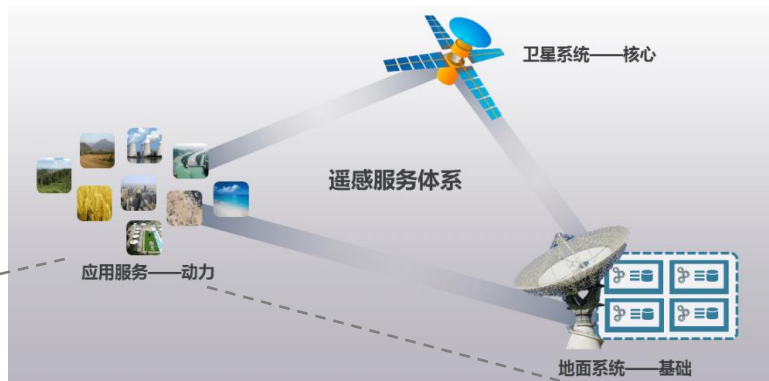
廈門大學
XIAMEN UNIVERSITY



光学遥感卫星影像解译

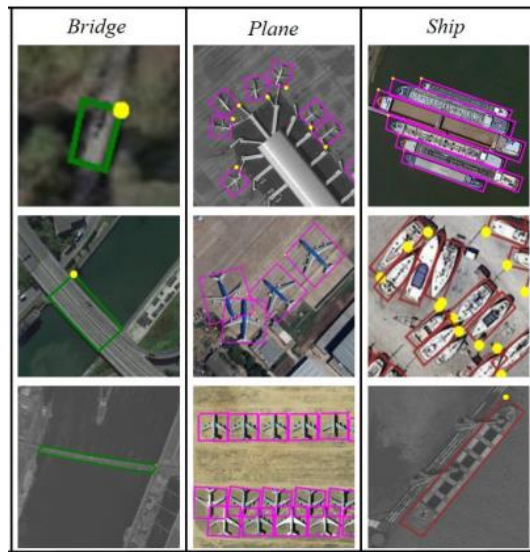
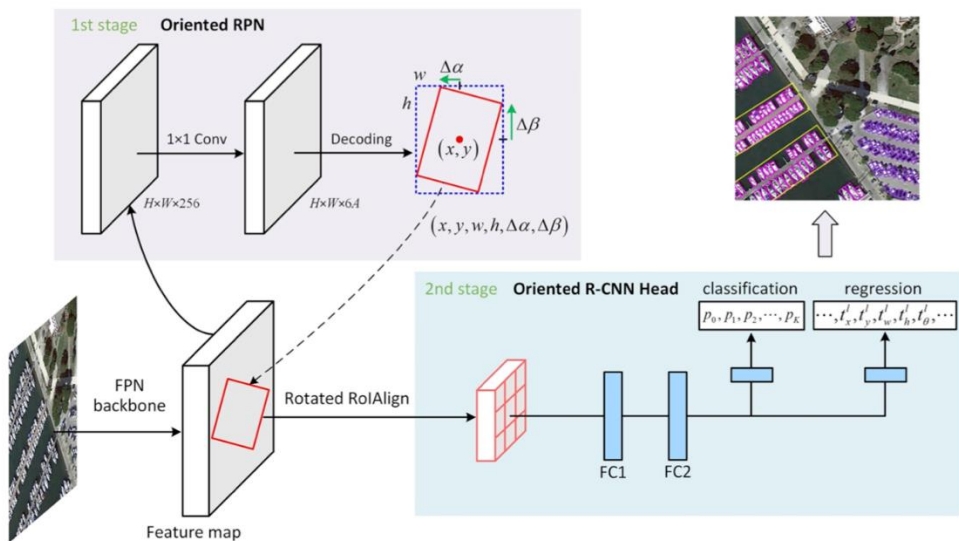
光学遥感卫星影像解译的意义

遥感卫星影像解译是一种利用卫星系统获取地球表面信息并进行分析和解释的技术，在环境保护、灾害监测、智慧农业、军事安全等方面具有重要的实际应用价值。



光学遥感卫星影像解译

双阶段旋转框目标检测算法：首先通过Oriented RPN生成有向候选框，然后Rotated RoIAlign提取固定尺寸大小的特征，最后将提取的特征作为有向检测头的输入，执行分类和精细回归。

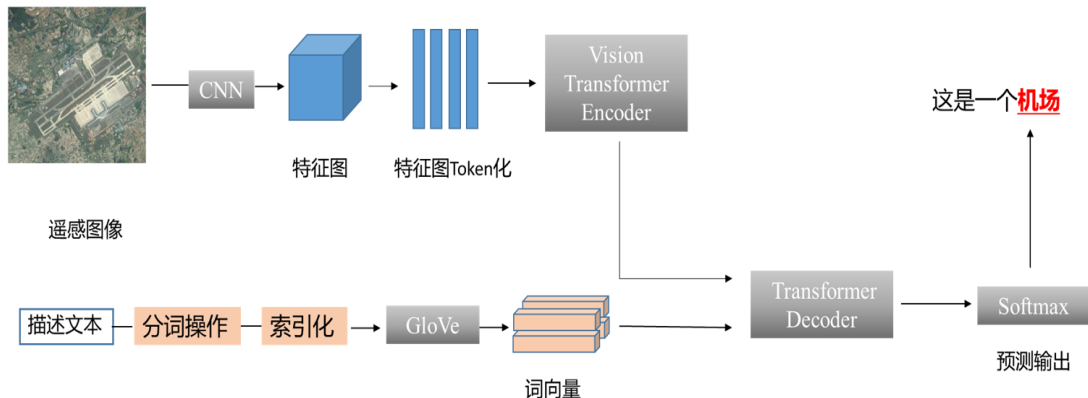




光学遥感卫星影像解译

卫星影像判读文本生成:

- ① 高分辨率遥感图像-长文本描述数据集构建
- ② 基于ViT 的高价值目标描述系统
- ③ 预测结果



机场: 1300张



舰船: 690张



码头: 339张



水坝: 661张



飞机: 1623张

图片	模型预测	参考
	这是一个位于市中心的大型民用机场, 机场有两条黑色沥青跑道。	这是一个民用机场, 周围是城市, 机场上有两条交叉的跑道, 材质为黑色沥青。
	在场景中, 客机分别停靠在三栋建筑物旁边, 跑道在下面。跑道上有五架飞机。底部的飞机很大, 向东。	这是一个大型机场, 机场中有大量的飞机, 其中有 1 架 ARJ21 飞机和 4 架 Boeing737 正在滑行道上滑行。
	该场景的特点是大量的混凝土码头, 所有码头都停满了船只且垂直于岸边, 码头位于一个港口内, 所在港口的南边是另一个港口, 码头靠岸部分是一处大型停车场。	场景中大量水泥材质的突堤式码头, 所有码头都停满了船只且垂直于岸边, 码头位于一个港口内, 所在港口的南边是另一个港口, 码头靠岸部分是一处大型停车场。

多/高光谱图像超分

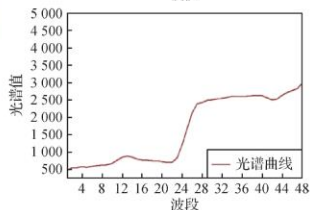
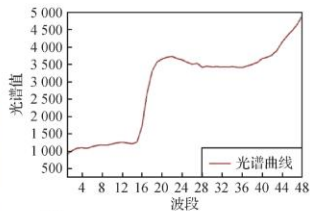
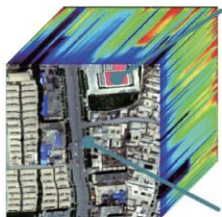


廈門大學
XIAMEN UNIVERSITY

研究背景

什么是高光谱图像?

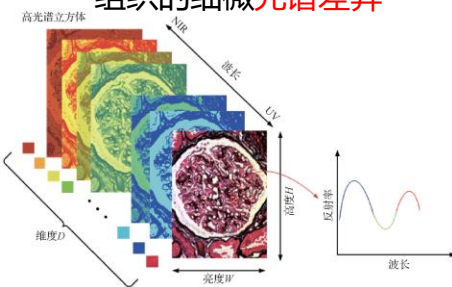
高光谱图像 (Hyperspectral Image, HSI) 是光谱分辨率在 $10^{-2}\lambda$ (小于10nm) 量级范围内的成像的光谱图像, 包含了丰富的**光谱、空间和辐射**等信息, 为**复杂目标分类识别和精细分析**提供了可能。



高光谱图像的应用

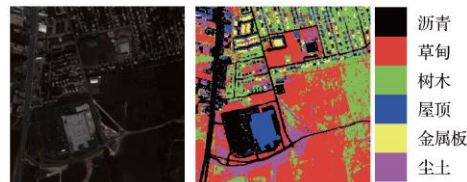
HSI 能够同时提供成像对象的 **2 维空间信息和 1 维光谱信息**, 进而反映其**化学成分信息及物理形态信息**。

捕获不同病理条件下组织的**细微光谱差异**



膜性肾病的HSI

包含丰富的**地物光谱信息**, 用于提升地物分类的**准确度**



(a) 伪彩色图 (b) 地物分布图

多光谱遥感图像中的地物分类情况

HSI 还常用于: **矿物识别、环境监测、植被分类和天体探测**等领域。

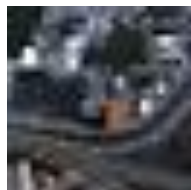


研究背景

由于在物理条件、成本、处理等方面需要综合权衡，实际多/高光谱成像系统常常无法同时呈现高的光谱分辨率和高的空间分辨率。

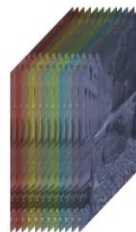


×



✓

较低的空间分辨率



×



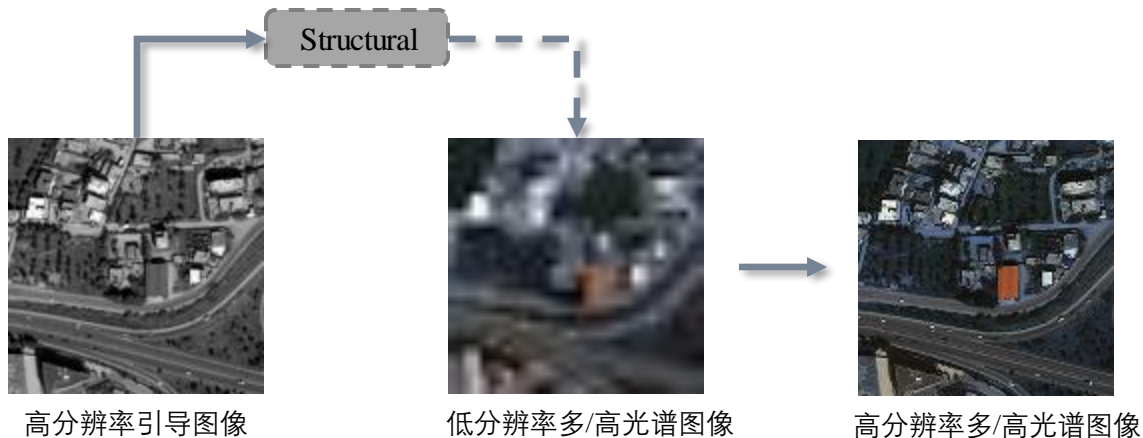
✓

较低的谱分辨率

缺乏高质量的多/高光谱数据!

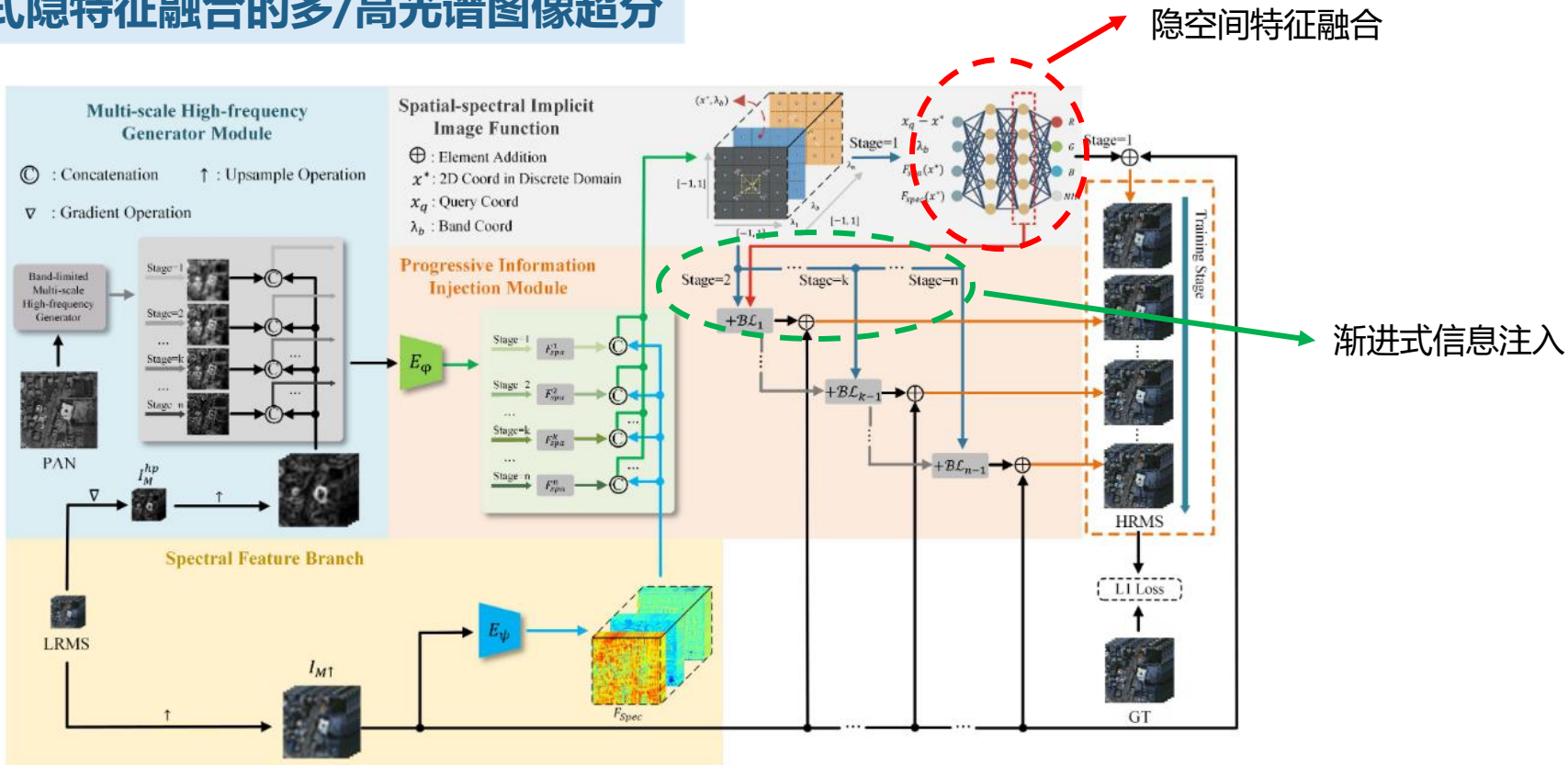


研究背景



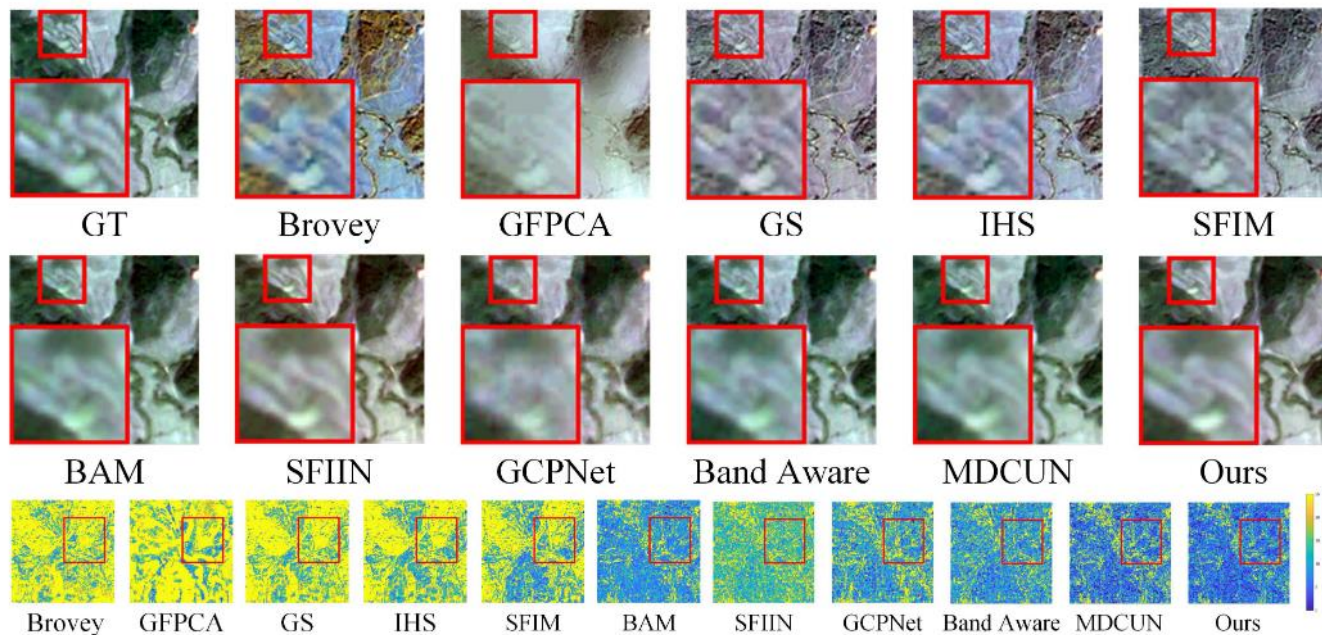
利用高分辨率引导图像的高频信号来增强其对应的多/高光谱图像的空间分辨率

渐进式隐特征融合的多/高光谱图像超分





实验效果:



The background of the slide features a close-up of orange flowers and dark buds on thin branches, set against a blurred background of a building with a red roof and light-colored walls. A dark blue semi-transparent geometric shape is overlaid on the right side of the image, containing the title text.

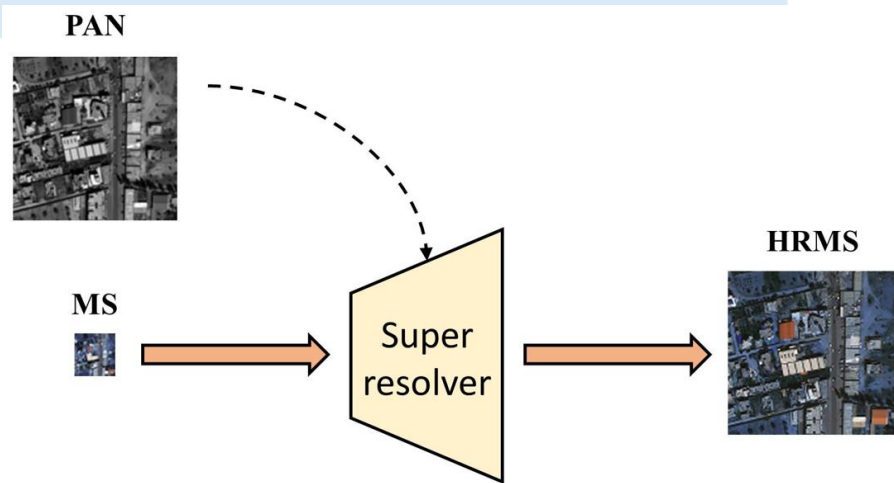
遥感图像全色锐化技术 (Pan-sharpening)



廈門大學
XIAMEN UNIVERSITY



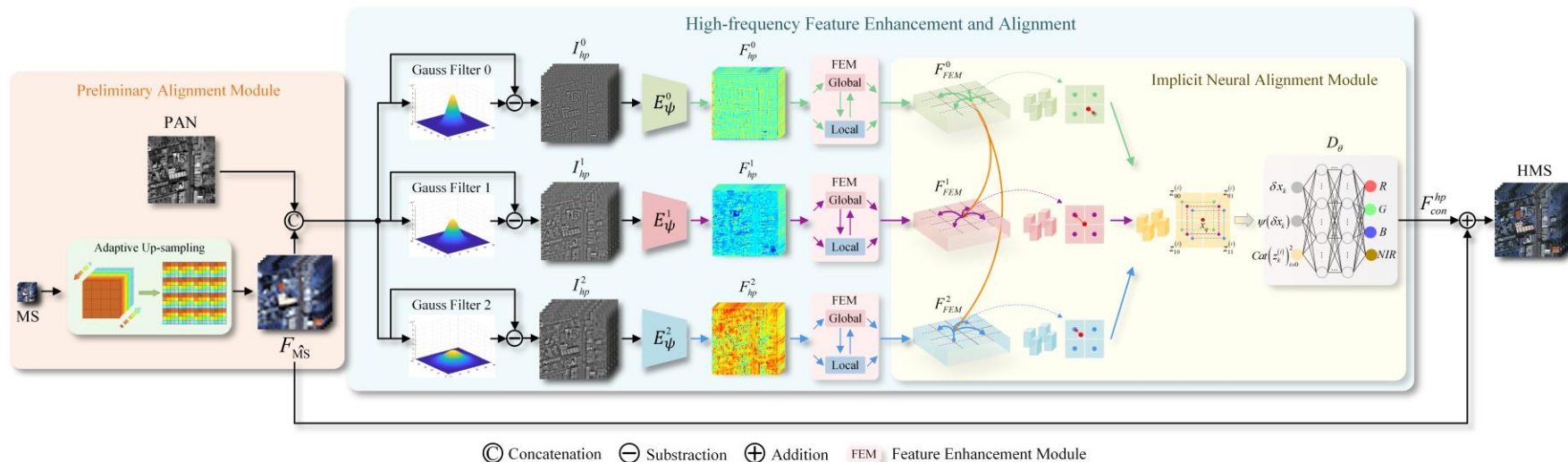
研究背景——遙感圖像全色銳化技術 (Pan-sharpening)



遙感圖像全色銳化技術 (Pan-sharpening) 旨在利用高分辨率全色圖像 (PAN) 來引導其對應的低分辨率多光譜圖像 (MS) 進行超分。其主要挑戰是如何有效提取PAN圖的高頻結構和紋理信息，並將其準確注入MS圖像。

代表性工作

Learning High-frequency Feature Enhancement and Alignment for Pan-sharpening, ACM MM 2023, CCF-A
Yingying Wang, Yunlong Lin, Meng Ge, Zhenqi Fu, Yuhang Dong, Linyu Fan, Hedeng Yu, Xinghao Ding[†], Yue Huang.



本文提出一个高频特征增强与对齐网络 HFEAN，网络包括三个模块：基于快速傅里叶卷积的高频特征增强模块 FEM、基于隐式神经表示的高频特征对齐模块 INA、以及预对齐模块 Pre-align。

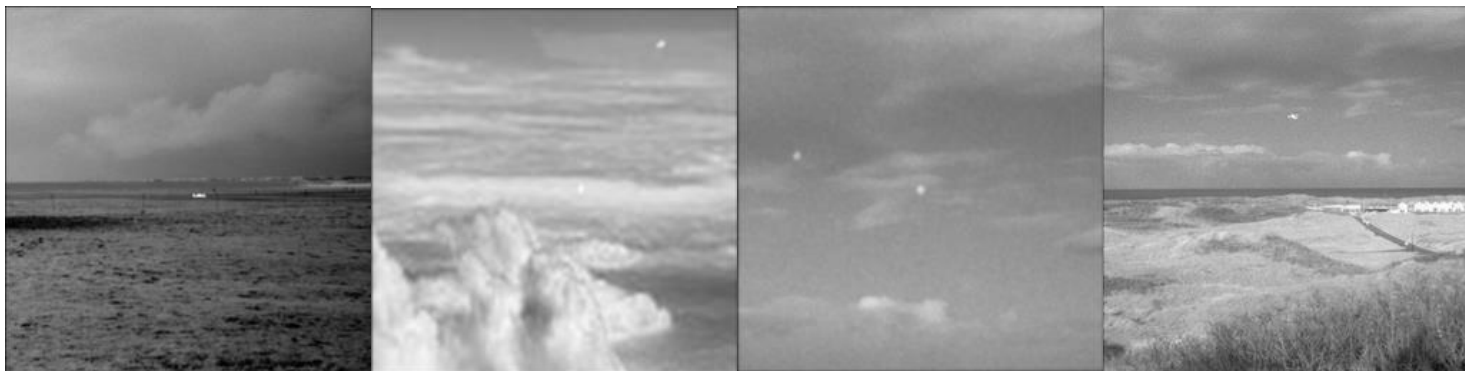
红外弱小目标检测



廈門大學
XIAMEN UNIVERSITY



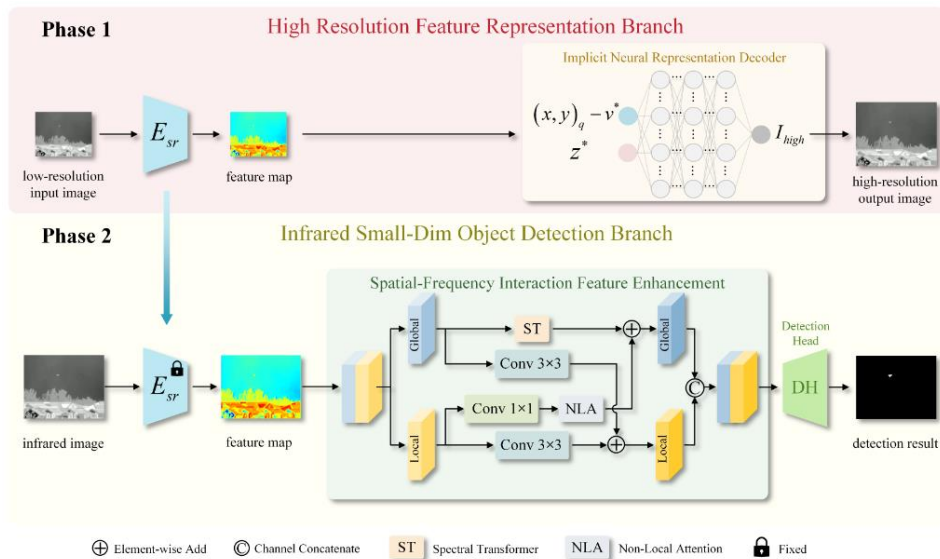
研究背景——红外弱小目标检测



在复杂海空背景下红外小目标具有无明显纹理和结构特征、目标形态模糊、尺寸小、像素少等特点，使得可用于目标辨识的信息较少，导致对目标的检测难度极大。此外，云层和海面杂波的干扰则进一步加剧了检测难度。深入研究复杂海空背景下红外弱小目标的鲁棒检测关键技术具有重要的军事价值和现实意义。

代表性工作

High-Resolution Feature Representation Driven Infrared Small-Dim Object Detection, PRCV 2023, CCF-C
Yuhang Dong*, Yingying Wang*, Linyu Fan, Xinghao Ding[†], and Yue Huang



本文提出了一种高分辨率特征驱动网络用于红外小尺寸目标检测 (HRFRD-Net)，网络包括三个模块：基于隐式神经表示的高分辨率特征表示分支 (HRFR)、红外小尺寸目标检测分支 (ISDOD) 和基于空频交互的红外弱小目标特征增强模块 (SFIFE)。

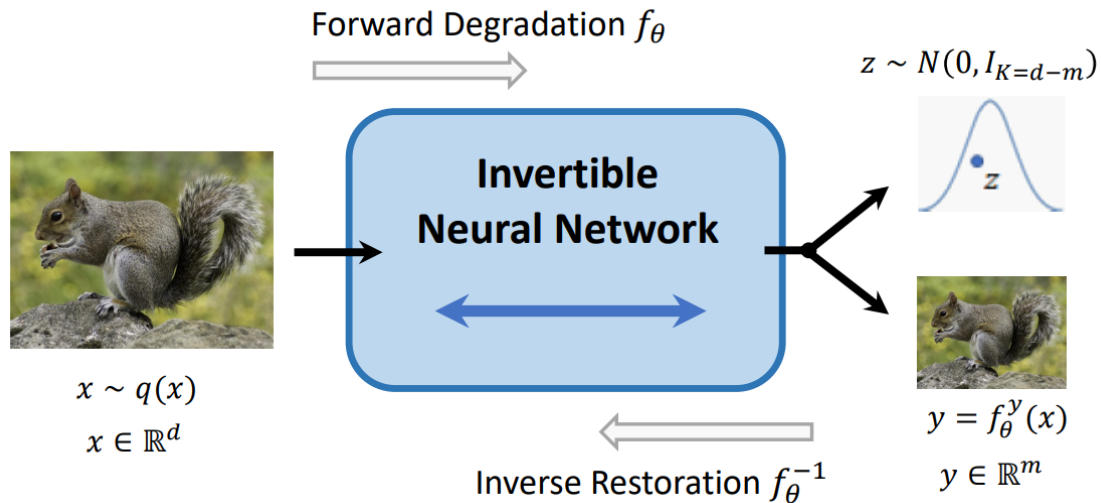
源、边协同图像重缩放



廈門大學
XIAMEN UNIVERSITY



研究背景——源、边协同图像可逆缩放



视频会议应用中，为了有效存储、传输和显示数据，需要对摄像头获得的超高清图像/视频（源头数据采集）进行下采样，而为了能够在更高分辨率显示设备上显示图像/视频，又需要对其进行上采样（边缘端显示）。然而缩小过程中的信息丢失，使得即便是通过超分等方案的放大方法的性能也受限，严重影响超高清视频会议的视觉质量。研究可逆缩放方法，可以在下采样时已考虑上采样时需要恢复的高频信息，从而使得不同分辨率显示器均能够高质量显示源头数据，极大提升视频会议视觉质量。

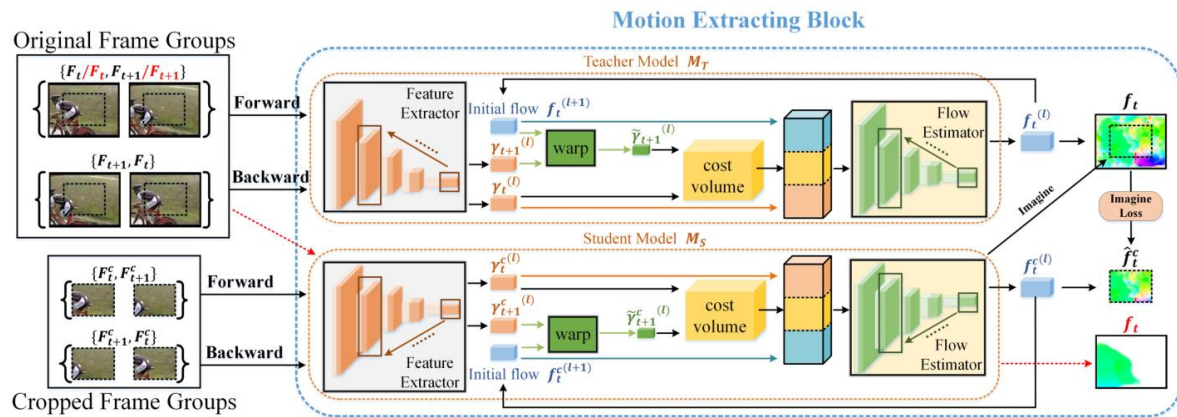
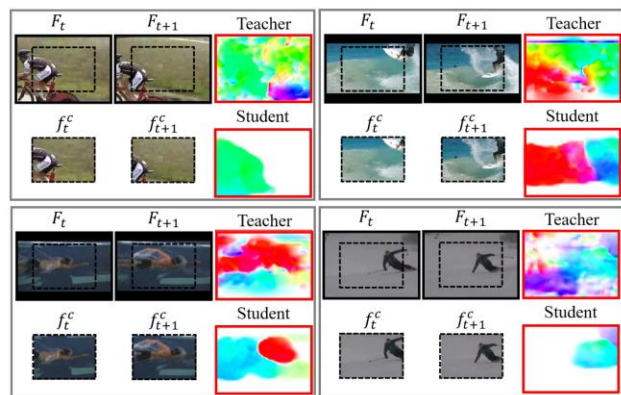
自监督视频表征



廈門大學
XIAMEN UNIVERSITY

自監督視頻表征

自監督視頻表征：現有方法依賴大量計算資源（如大批量，負樣本），在不需要負樣本的前提下通過外觀學習塊來感知視頻中的靜態外觀信息，通過運動提取塊來關注運動物體本身的主要運動信息，綜合學習有效的時空表征。



➤ 不可見運動模擬

對移動物體的想象可以吸引網絡關注其本身信息，通過裁剪來創造運動物體走出視野外的狀況。

➤ 不可見運動定位

基於前後一致性假設，尋找前向流和后向流差值過大的像素區域，即不可見像素區域。

➤ 想象損失

找到由不可見運動模擬導致的人為不可見像素區域，即在裁剪后的幀中不可見但在原始幀中可見的像素區域



網絡被引導在失去部分運動物體信息的狀況下通過想象來學習光流，從而聚焦到運動物體本身

网络结构:

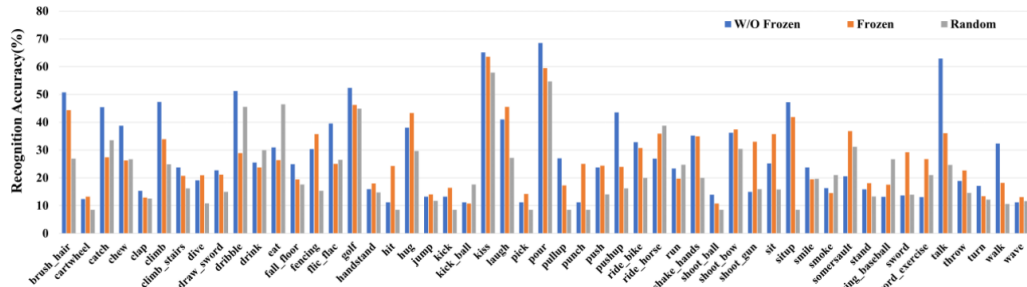
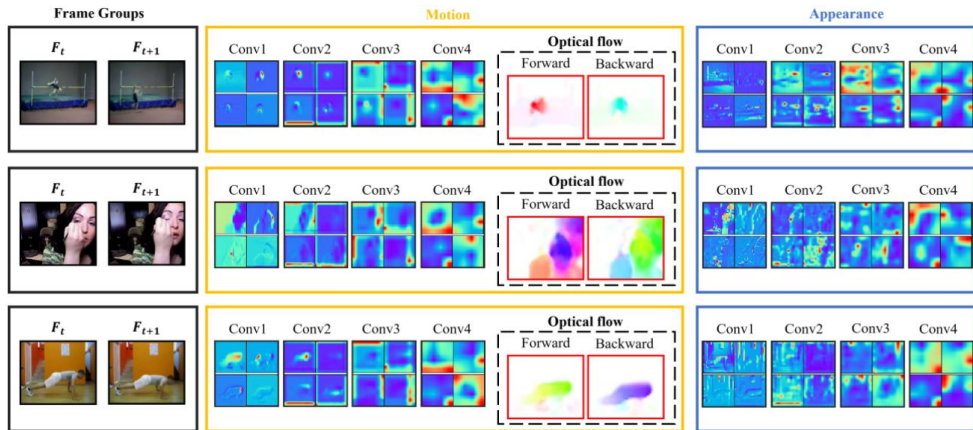
表 4.1 MEB 中特征提取器和流量估计器的网络结构

特征提取器		流量估计器
Conv1: 3×3, 16, stride 2	×2 conv. lReLU	Conv1: 3×3, 128 conv. stride 1, lReLU
Conv2: 3×3, 32, stride 2	×2 conv. lReLU	Conv2: 3×3, 128 conv. stride 1, lReLU
Conv3: 3×3, 64, stride 2	×2 conv. lReLU	Conv3: 3×3, 96 conv. stride 1, lReLU
Conv4: 3×3, 96, stride 2	×2 conv. lReLU	Conv4: 3×3, 64 conv. stride 1, lReLU
Conv5: 3×3, 128, stride 2	×2 conv. lReLU	Conv5: 3×3, 32 conv. stride 1, lReLU
Conv6: 3×3, 192, stride 2	×2 conv. lReLU	Conv6: 3×3, 2 conv. stride 1

表 4.2 ALB 中生成器和判别器的网络结构

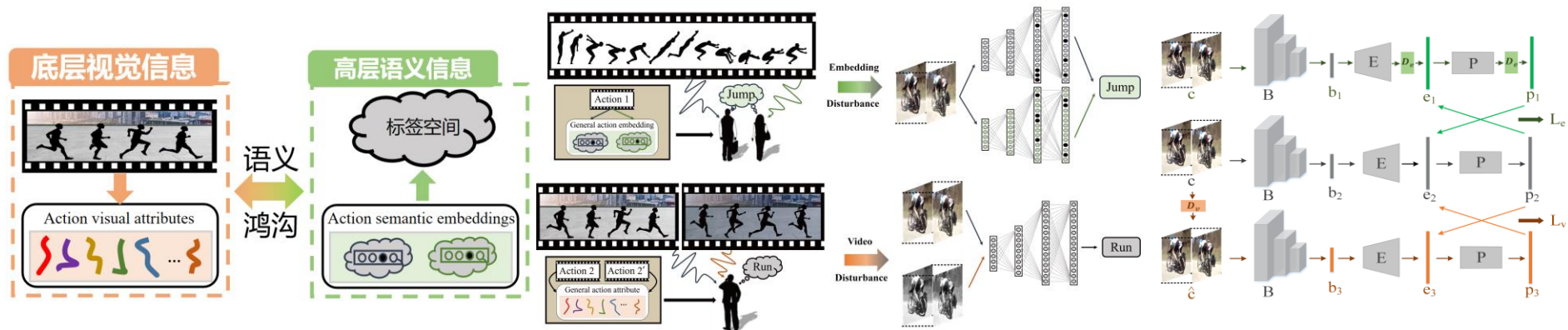
生成器	判别器
DeConv1: 5×5, 256 deconv. stride 2, BNorm, ReLU	Conv1: 5×5, 64 conv. stride 2, lReLU
DeConv2: 5×5, 128 deconv. stride 2, BNorm, ReLU	Conv2: 5×5, 128 conv. stride 2, BNorm, lReLU
DeConv3: 5×5, 64 deconv. stride 2, BNorm, ReLU	Conv3: 5×5, 256 conv. stride 2, BNorm, lReLU
DeConv4: 5×5, 3 deconv. stride 2	Conv4: 5×5, 512 conv. stride 2, BNorm, lReLU

实验效果: 在行为识别数据集上进行测试:



自監督視頻表征

自監督視頻表征：當前基於大數據的深度学习模型通常從視覺信息出發，而視頻特征空間與人為設定的標簽空間之間存在語義鴻溝。有意识地挖掘視頻中的高層語義特性，結合底層視覺和高層語義信息提取更全面的視頻表征。



➤ 视觉片段扰动策略

视觉：同一动作视频的主要信息是一致的。视频中演员或场景的微小变化不会影响对动作的识别。

+

➤ 语义编码扰动策略

语义：一个动作视频的良好编码表示的主要信息也是一致的。不同的人在看同一个动作视频时，他们的认知总是相同的。

=

➤ 目标

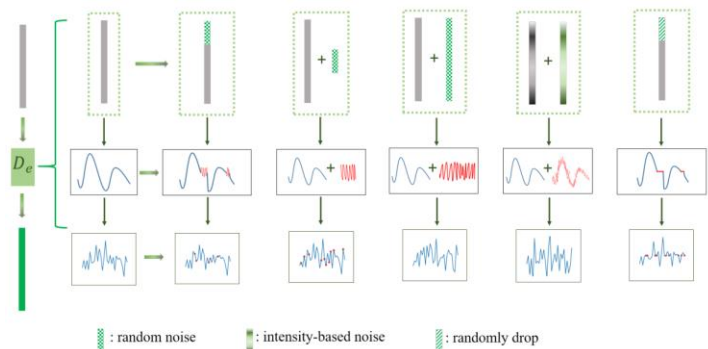
在特征空间中使被扰动的片段/编码和原始片段/编码更接近

➔

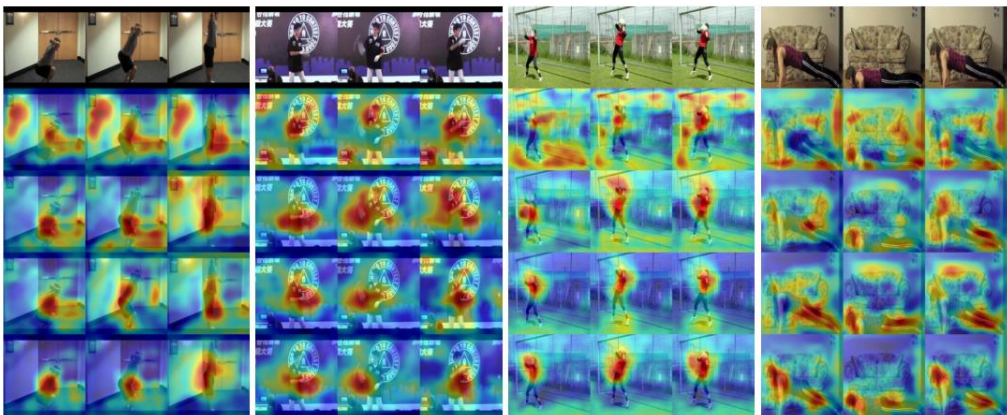
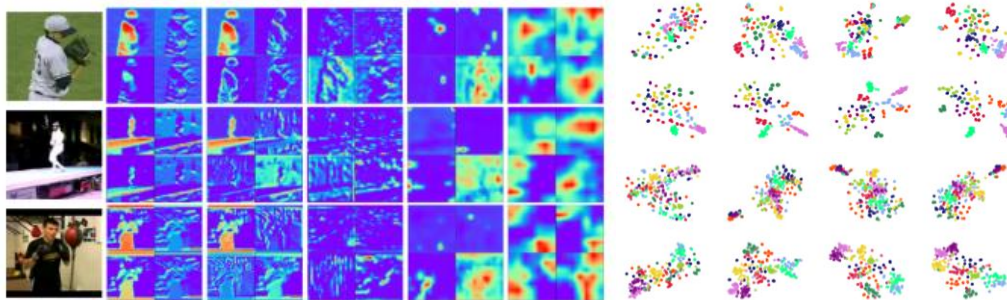
网络在**弱化冗余细节**影响的同时，**更加关注动作的主要信息**



扰动策略:



实验效果: 在行为识别数据集上进行测试:



基于森林的无监督异常检测



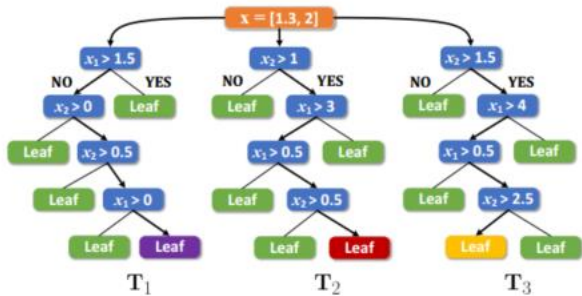
廈門大學
XIAMEN UNIVERSITY

从深度学习视角改进传统森林

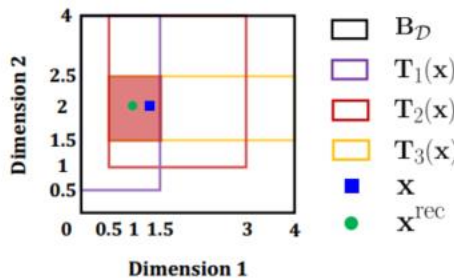
1、结合孤立和重建思想的无监督异常检测算法

完全随机森林

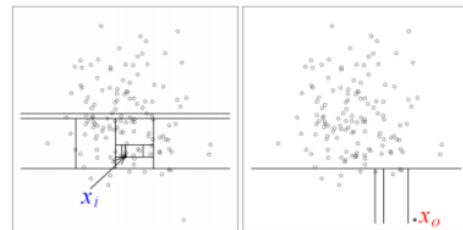
剪枝



(a) Examples on completely random trees

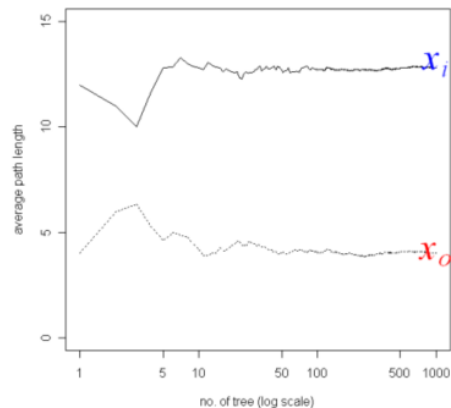


(b) Demonstration on sample reconstruction



(a) Isolating x_i

(b) Isolating x_o



(c) Average path lengths converge

Reconstruction forest

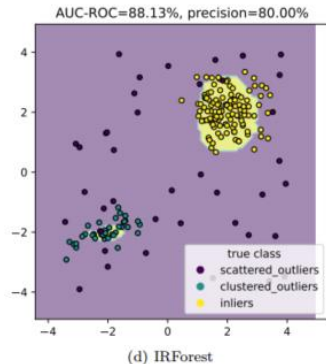
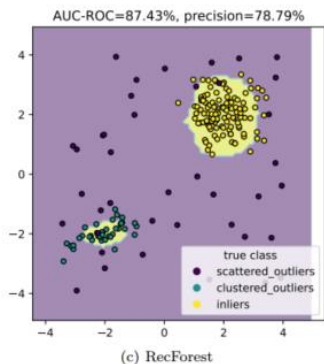
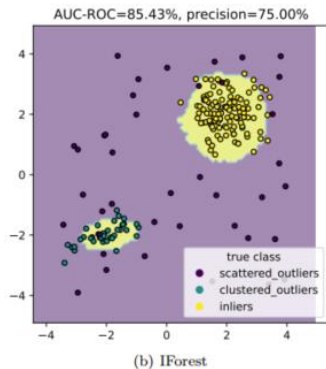
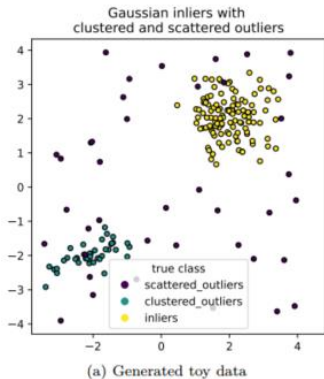
Isolation forest



1、结合孤立和重建思想的无监督异常检测算法

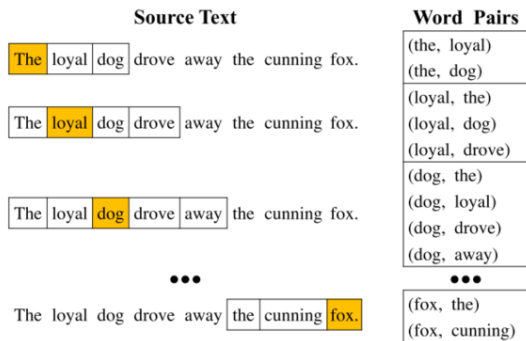
通过较少的计算代价，
获得较大的算法性能提升。

Model	Training (s)	Testing (s)	Total (s)
OCSVM	0.03	0.01	0.04
PIDForest	5.71	2.65	8.36
IForest	0.40	0.11	0.51
RecForest	0.39	0.17	0.56
IRForest	0.42	0.19	0.61

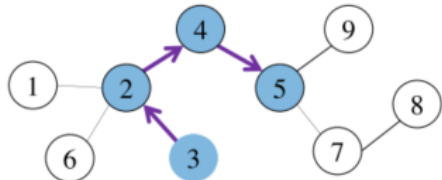




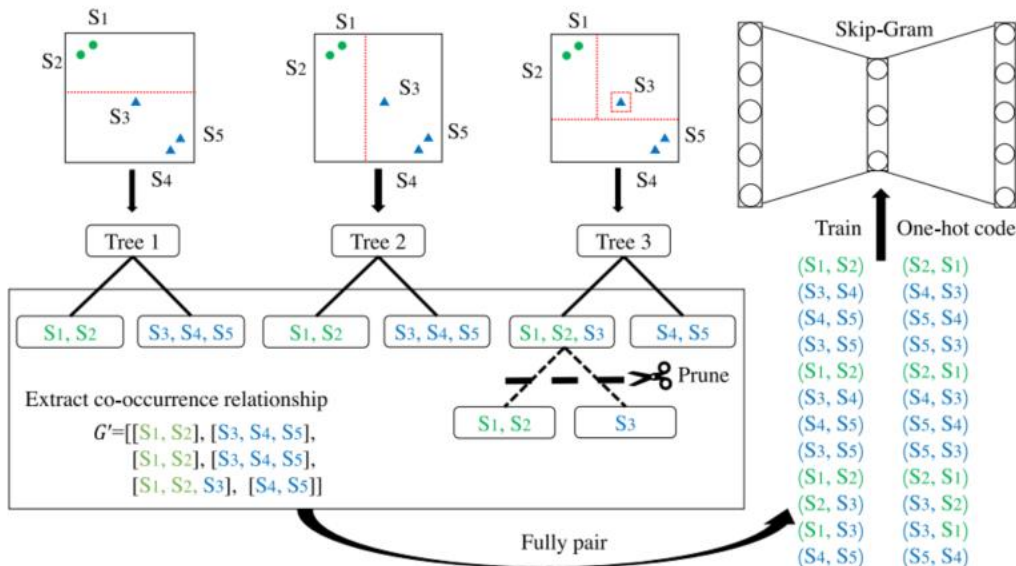
从深度学习视角改进传统森林 2、传统森林缺乏分布式表征能力



word2vec



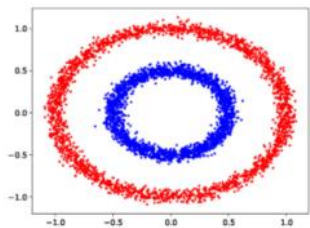
deepwalk



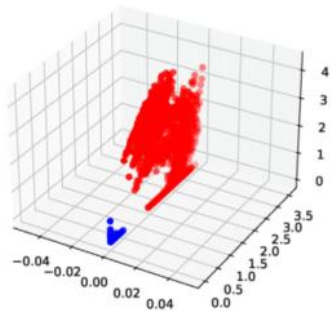
DRL-TF



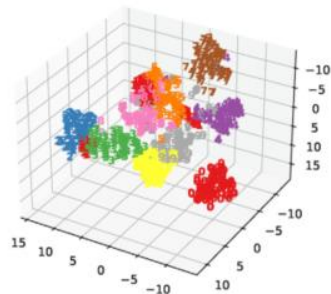
2、传统森林缺乏分布式表征能力



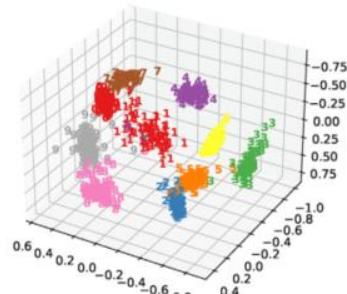
(a) input



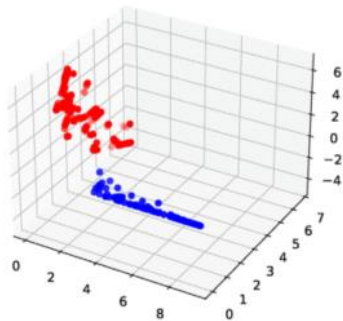
(b) MLP



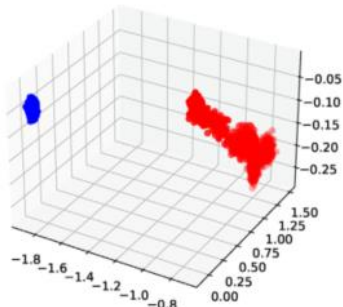
(a) t-SNE



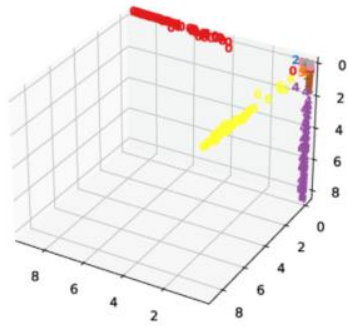
(b) DRL-TF



(c) hashing



(d) DRL-TF



(c) hashing

模型压缩与轻量化

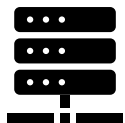
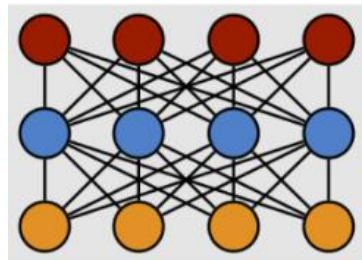


廈門大學
XIAMEN UNIVERSITY



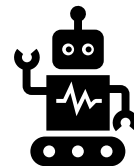
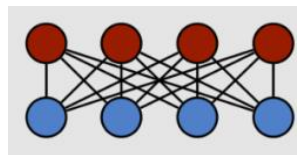
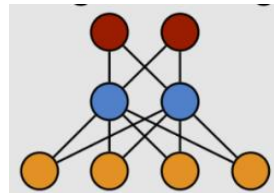
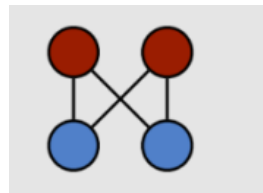
基于教师微调的量化蒸馏

- 当前深度学习范式：大数据 + 大算力（模型）



模型压缩关键技术

- 模型剪枝
- 参数量化
- 轻量化设计
- 知识蒸馏**





基于教师微调的量化蒸馏

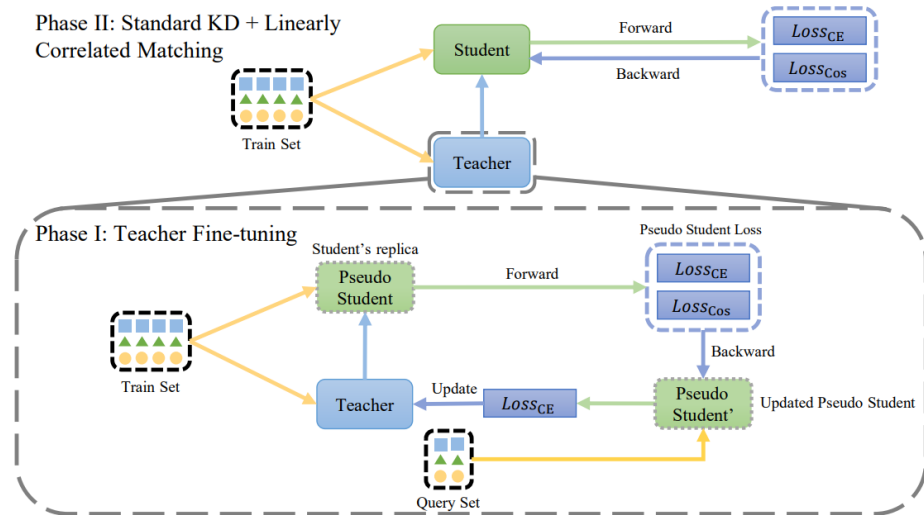
量化场景下，量化后模型精度会出现明显下降，且这种损失会随着量化比特数的减小而增大。知识蒸馏可以显著缓解量化所造成的模型精度损失，但如何将知识蒸馏与量化更好的结合有待挖掘。

基于教师微调的量化蒸馏框架

- 对教师进行针对性微调
- 允许教师在量化蒸馏过程中根据低精度学生学习情况进行针对性更新

余弦相似度代替KL散度

- 以线性相关匹配方式进行量化蒸馏
- 较为宽松的约束方式提升低精度学生模型的蒸馏效果





动态知识线索蒸馏

当前知识蒸馏范式中，不同知识线索之间的比例系数依然是依据经验设定的超参数，但不同线索对于不同样本在整个蒸馏过程中对于学生模型的重要性是不一致的，这会抑制学生模型的能力上限。

➤ 知识线索权重估计网络

- 对不同知识线索的最佳分配方式进行**动态估计**，借鉴元学习训练思路生成的有关知识线索的权重估计

➤ 内外循环框架

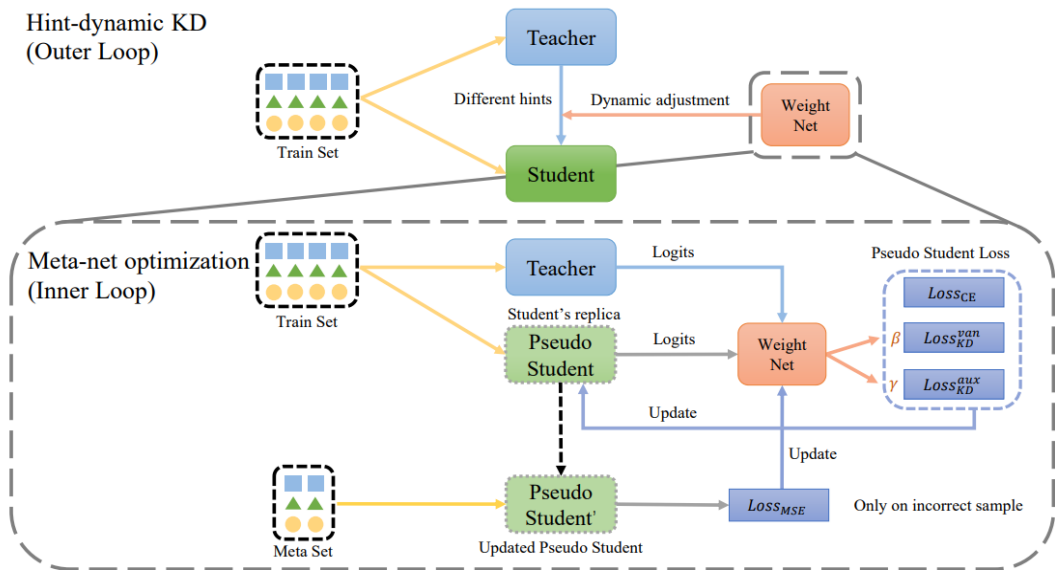
- 对权重估计网络进行有效训练

➤ 伪学生

- 辅助权重估计网络训练

➤ 元数据集

- 检验权重估计网络训练效果



基于价值感知和知识摘要的自适应蒸馏

传统知识蒸馏范式中，存在教师向学生传递的知识中出现冗余的情况，从而导致学生模型从蒸馏训练中收益降低，性能变差。如何定义知识价值和抽取有价值知识以实现高效蒸馏有待探索。

知识价值编码

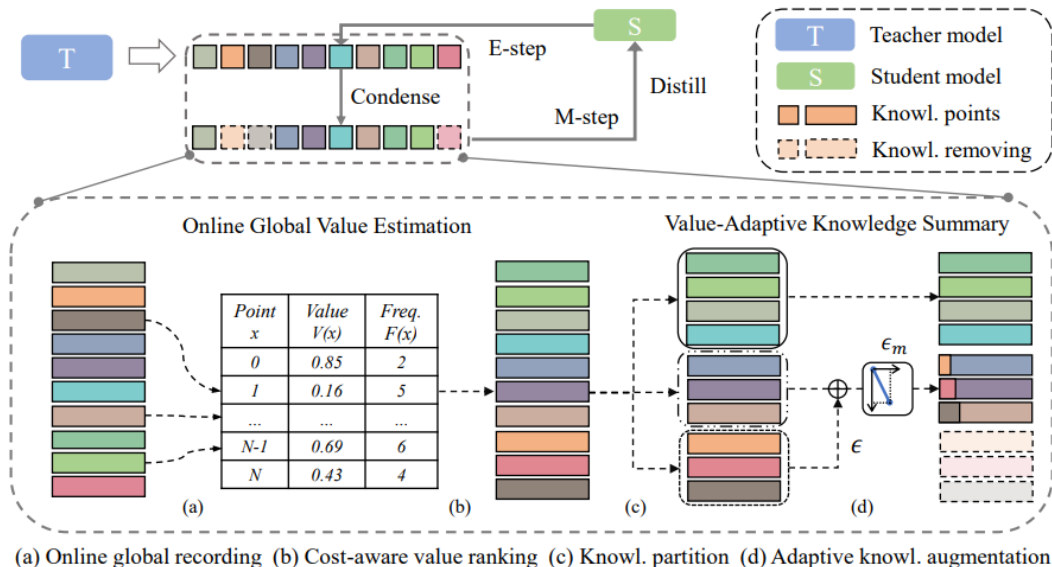
- 知识浓缩和模型蒸馏交替进行
- E步对知识点排序估计知识价值
- S步根据估计的知识价值来进行知识蒸馏

全局知识价值评估

- 在线价值记录和价值感知排序
- 对训练集 X 进行重新加权，评估价值

价值自适应知识摘要

- 在蒸馏前进行价值自适应知识增强，进行知识浓缩



语音增强与DOA估计



廈門大學
XIAMEN UNIVERSITY



麦克风阵列声源定位

将麦克风设计排列成特定的阵列结构，对具有空间特性的声场采样，通过波束形成算法进行优化和估计声源位置

- ✓ 信号采集
- ✓ 噪声抑制
- ✓ 定位追踪

在军事，工业，交通，民生等各个领域都有着广泛的应用，能够为智能设备提供目标定位的功能



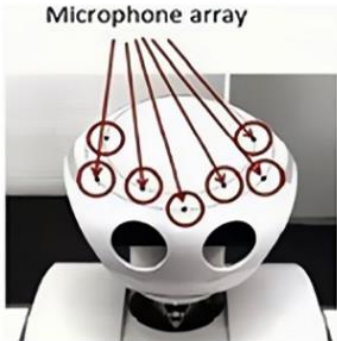
(a)



(b)



(c)



(d)

a. 反直升机雷达

b. 声学成像仪

c. 鸣笛抓拍系统

d. 智能机器人

麦克风阵列定位技术的一些应用

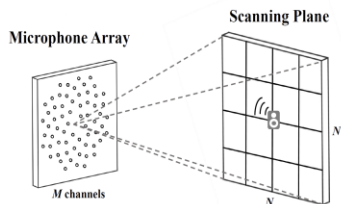
基于麦克风阵列的计算声成像算法研究

目标：通过麦克风阵列，让声音可视化，实现声源高分辨定位。

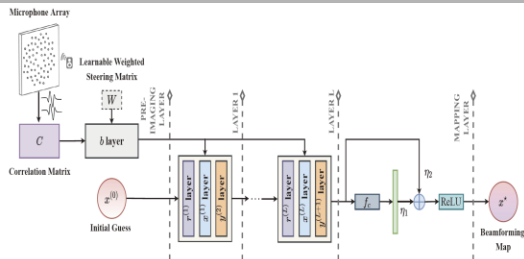
意义：可服务于交通、电力、能源化工领域的异常检测，涉及语音增强以及降噪等技术。

特色：通过自监督降噪技术，改变了传统定位算法的范式，解决了传统成像算法分辨率低，抗干扰能力弱、**运算效率低的问题**，提升了泛化性，产品实现了国产化替代，达到国外Fluke产品性能。

麦克风阵列



基于领域知识的AI成像算法



局部放电



气体泄漏

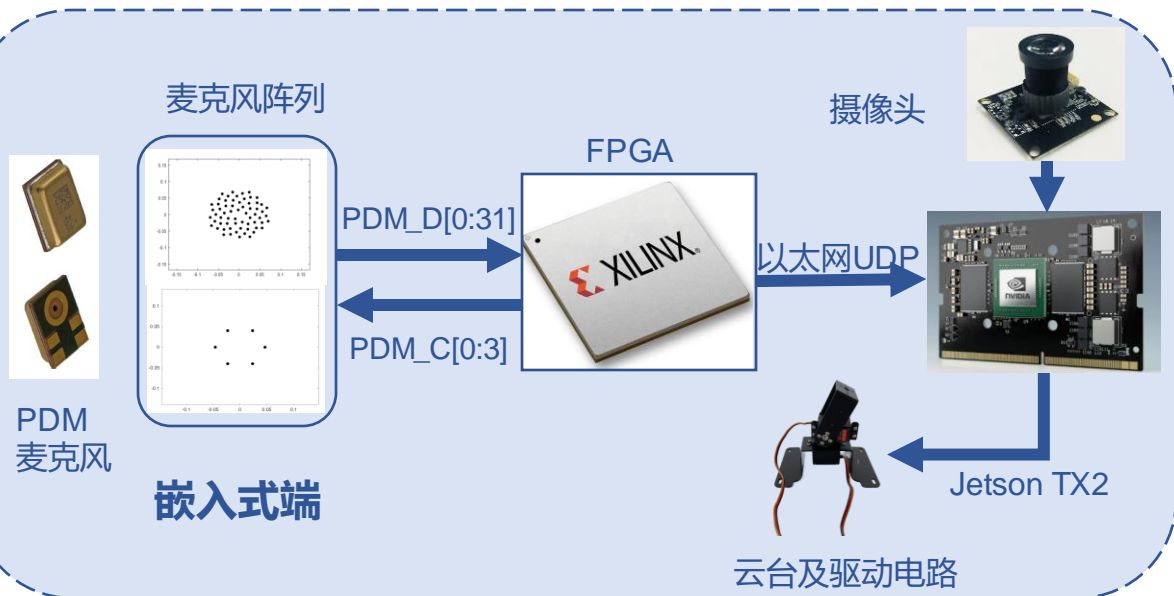


基于麦克风阵列的计算声成像算法研究

目标：通过麦克风阵列，让声音可视化，实现声源高分辨定位。

意义：可服务于交通、电力、能源化工领域的异常检测，涉及语音增强以及降噪等技术。

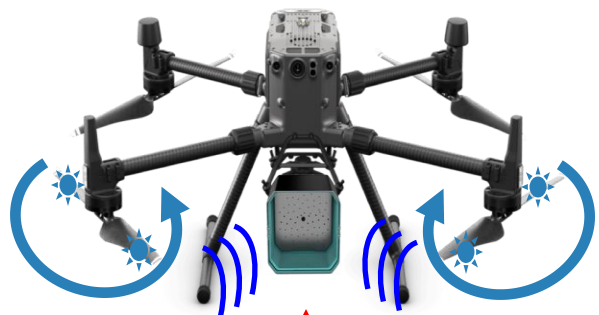
特色：通过自监督降噪技术，改变了传统定位算法的范式，解决了传统成像算法分辨率低，抗干扰能力弱、**运算效率低的问题**，提升了泛化性，产品实现了国产化替代，达到国外Fluke产品性能。





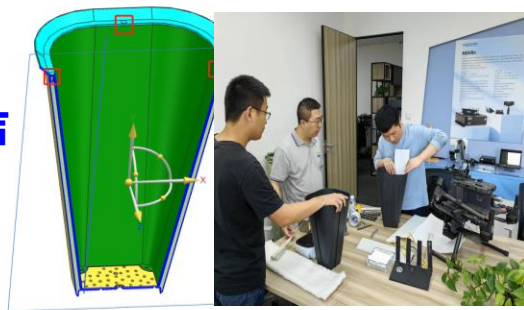
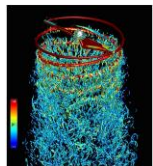
基于麦克风阵列的计算声成像算法研究

通过物理降噪和高效成像算法，实现极低信噪比条件下的稳定成像和定位，该产品可广泛服务于城市噪声防治，智慧电网巡检以及复杂环境下的搜救和语音增强处理。



目标声

旋翼直达声



物理+算法降噪



声源定位

声源定位网络DAMAS-ALISTA: 将传统基于模型的算法DAMAS通过深度网络展开, 兼顾基于模型和基于神经网络的声源定位算法的优点, 极大地提高了算法的泛化性能

基于模型的算法

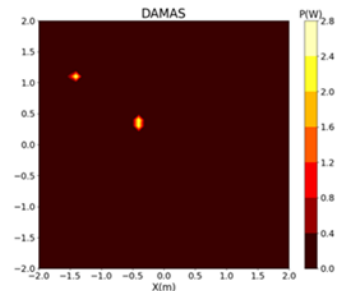
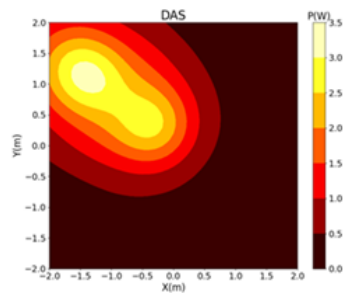
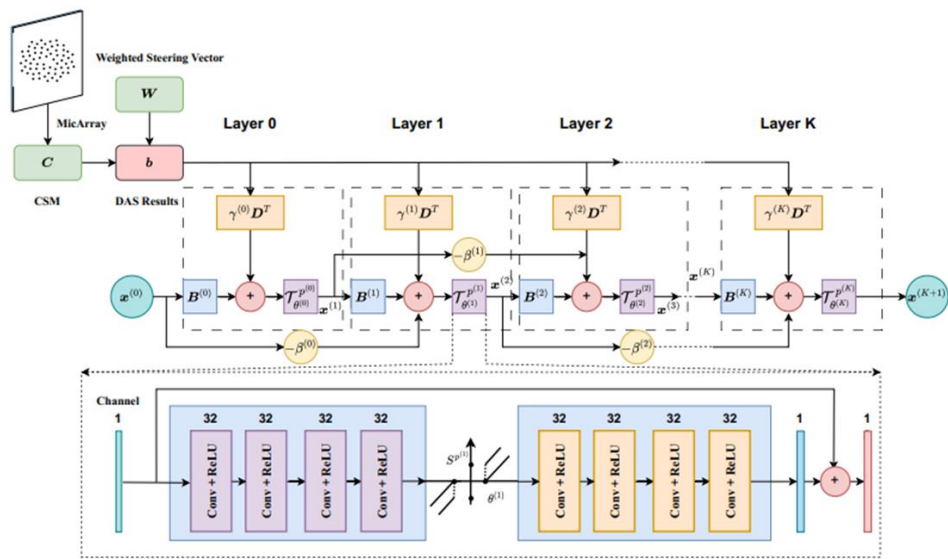
- 需要大量迭代, 影响实时性
- 超参数选择困难, 对结果有一定的影响

基于网络

- 需要大量的数据训练, 真实数据没有标签 (label)
- 黑盒模型, 泛化性能难以保障

模型驱动神经网络

- 提出传统算法解决DAMAS中的线性逆问题
- 通过Deep-Unfolding, 得到新的成像算法DAMAS-ALISTA





语音去混响

从被噪声环境损坏的语音信号中恢复干净的语音



语音识别、移动通信、工业检测、声源定位等诸多应用。



深度学习 取得成功 的前提

较为完备的、配对声源独立的、数据量充足的训练数据。

训练数据与测试数据噪声分布已知或可被建模。



矛盾

解决方案



自监督学习

现实应用 中存在的 局限性

高质量的配对声源数据收集需要花费大量的时间与精力。

现实世界噪声分布的复杂性。

语音去混响

假如 $x(t)$ 是接收到的混合信号:

$$x(t) = s(t) * h(t) + z(t)$$

STFT(short Fourier transform)转换到时频域

$$X(k, f) = S(k, f) \cdot H(k, f) + z(k, f)$$

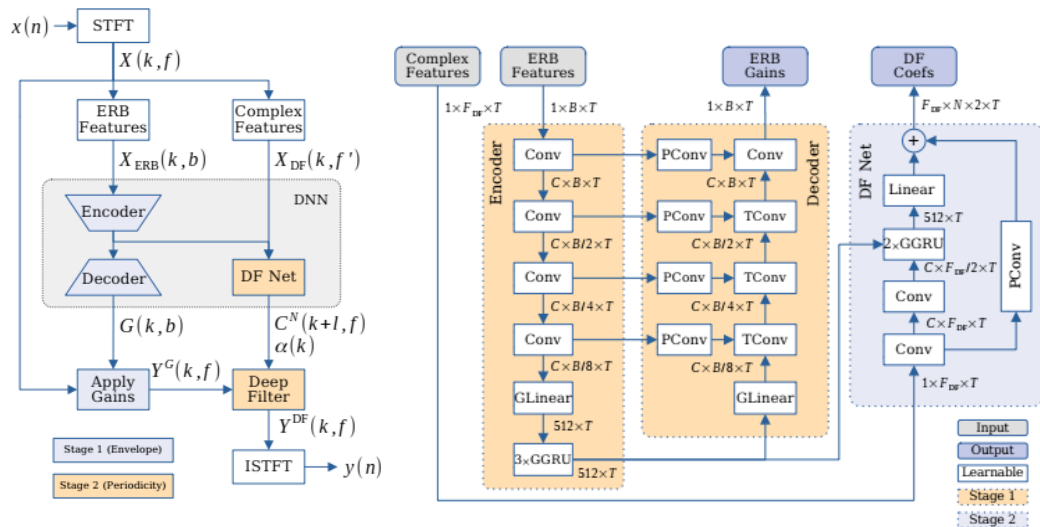
定义时频域复数滤波器Deep filtering:

$$Y(k, f) = \sum_{i=0}^N C(k, i, f) \cdot X(k - i + l, f)$$

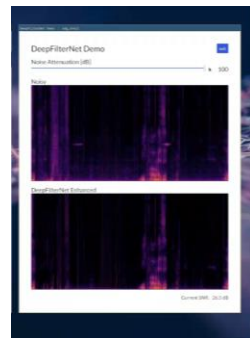
实值增益和深度滤波结合, 网络输出复数时频谱为:

$$Y^{DF}(k, f) = \alpha(k) \cdot Y^{DF'}(k, f) + (1 - \alpha(k)) \cdot Y^G(k, f)$$

- 第一阶段: 使用ERB(equivalent rectangular bandwidth, ERB)比例增益来**增强语音包络**, 其对人类频率感知进行建模
- 第二阶段: 采用深度滤波**增强语音中的周期分量**
- 除了利用语音的感知特性外, 还通过可分离卷积和线性递归层中的广泛分组来增强网络的稀疏性, 以设计低复杂度的网络架构



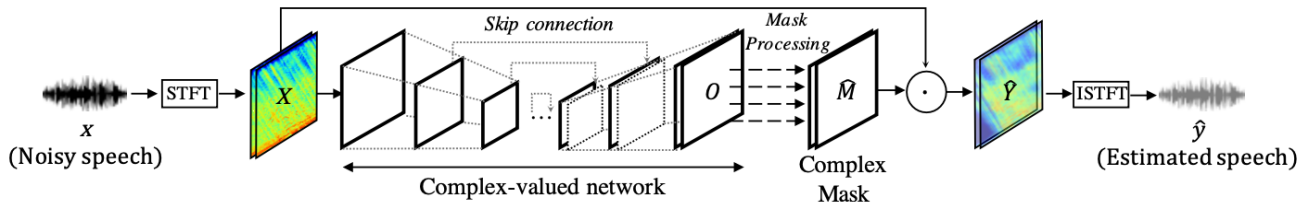
实测演示





语音增强基础模型: DEEP COMPLEX U-NET

网络结构:



损失函数:

在语音降噪任务中，将估计语音和clean语音做loss是常规操作，在此基础上，将噪声也作为一种估计与noise做loss,构成多任务降噪的一个实例

$$loss_{wSDR}(x, y, \hat{y}) := \alpha loss_{SDR}(y, \hat{y}) + (1 - \alpha) loss_{SDR}(z, \hat{z})$$

其中，x是mix的带噪信号，y是mix前的干净语音信号，z是mix前的噪声信号； \hat{y} 是估计的语音信号， $\hat{z} = x - \hat{y}$ 是估计的噪声信号





实验设置:

数据集: 训练数据集使用**VoiceBank+Demand**, 语音数据来自VCTK。

预处理阶段: 最初的原始波形首先从**48kHz**下采样到**16kHz**。对于实际的模型输入, 通过具有**64ms**大小的Hann窗口和**16ms**跳长的STFT从下采样波形获得复值谱图。

网络架构: 将实现10层网络架构:

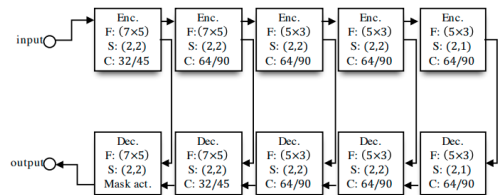
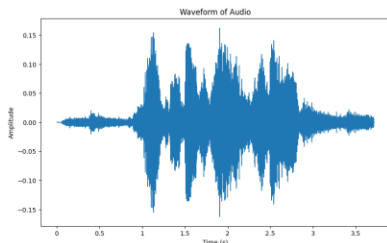
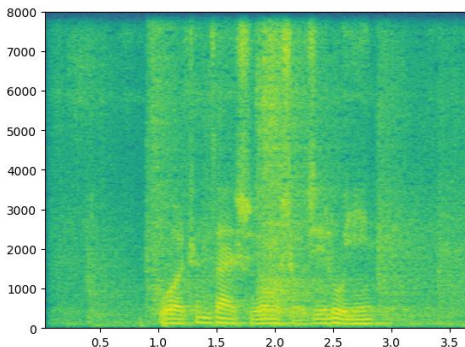
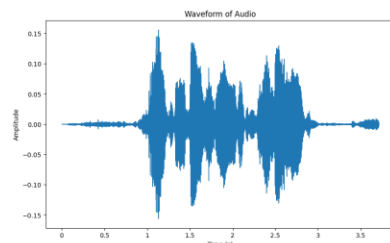
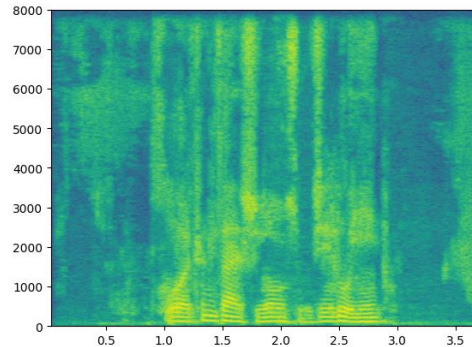


Figure 9: DCUnet-10 (10-layer): a model with 10 convolutional layers.

实验效果: 实际采集单通道数据进行测试



含噪语音



增强后语音



自監督語音增強

現有的無監督方法應用受限，嚴重依賴大數據和配對的數據類型，數據及網絡本身的信息有待挖掘利用。

❑ 自監督學習算法+孿生網絡

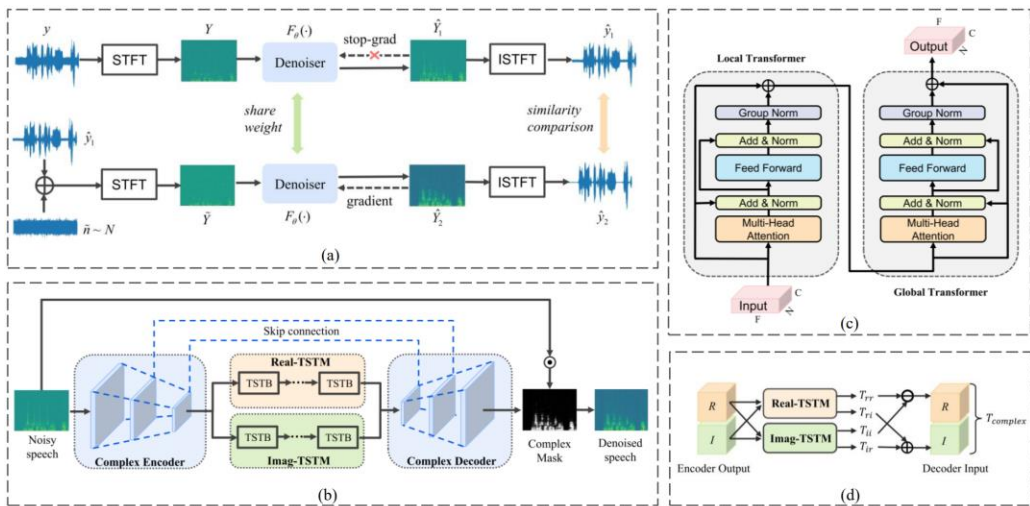
❑ 自校驗深度學習模型

➤ 網絡在自我校驗的過程中逐漸提升去噪性能

❑ 基於上下文感知的復值網絡

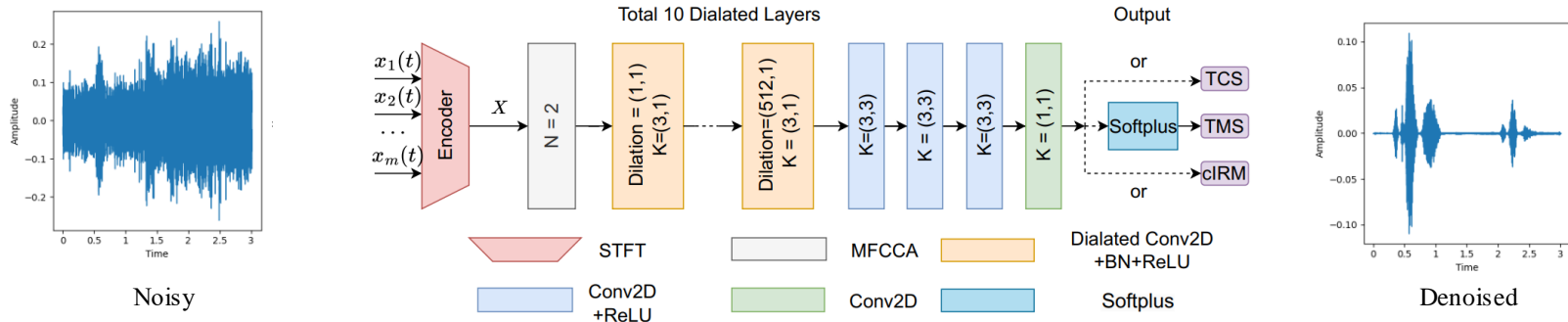
❑ 挖掘有限數據和網絡結構中的多維度信息

➤ 基於時頻譜的特徵提取網絡，融入注意力機制深度感知上下文信息

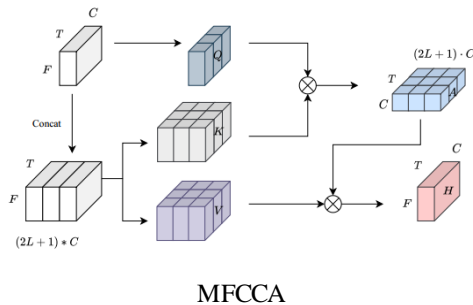


多通道语音增强

多通道语音增强：为了在较低信噪比条件下实现语音增强，通过多帧跨通道注意力机制，探究帧级和通道级之间的相关性，达到比单通道更好的去噪效果，甚至能够进一步适配下游任务（如定位）



- 空洞卷积能够在得到更大的感受野的同时减少计算量，使网络更加轻便
- 通过多通道音频输入，使用多帧跨通道注意力机制 (MFCCA)，在通道间和帧间进行建模，对音频信号实现降噪



$$Q_i^{mf} = X W_i^{mf,q} + (b_i^{mf,q})^T \in \mathbb{R}^{T \times C \times F}$$

$$K_i^{mf} = X_{cc} W_i^{mf,k} + (b_i^{mf,k})^T \in \mathbb{R}^{T \times (2N+1) \times C \times F}$$

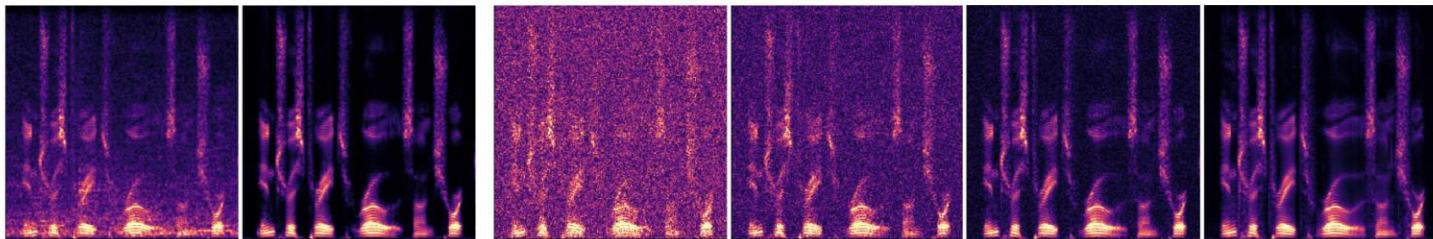
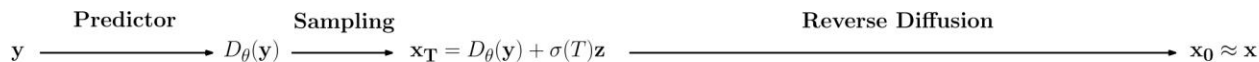
$$V_i^{mf} = X_{cc} W_i^{mf,v} + (b_i^{mf,v})^T \in \mathbb{R}^{T \times (2N+1) \times C \times F}$$

$$H_i^{mf} = \text{Softmax} \left(\frac{Q_i^{mf} (K_i^{mf})^T}{\sqrt{D}} \right) V_i^{mf} \in \mathbb{R}^{T \times C \times F}$$



基于扩散模型的语音增强

传统的有监督语音增强模型，几乎不可能在训练数据中覆盖所有可能的声学条件以保证泛化。此外，一些鉴别性方法已被证明会导致令人不快的语音失真。



□ 扩散模型

□ 正向过程将数据慢慢转化为易处理的先验，并训练神经网络解决反向过程，以从该先验生成干净的数据

➤ 高计算负担，在不利条件下易产生发声和呼吸伪影

□ 随机再生方案

□ 结合预测和生成模型来产生高质量的样本

➤ 减少扩散模型的计算负担及其产生的伪像



含噪语音



增强后语音



干净语音

Data-centric Artificial Intelligence



廈門大學
XIAMEN UNIVERSITY

基于通道高效处理的数据集蒸馏

传统的数据集蒸馏蒸馏范式中，直接匹配压缩数据 S 和原始训练数据 T 。这里我们提出了一个中间过程，在增加极少量存储负担的情况下创建更多的合成数据。

➤ 合成数据量的价值

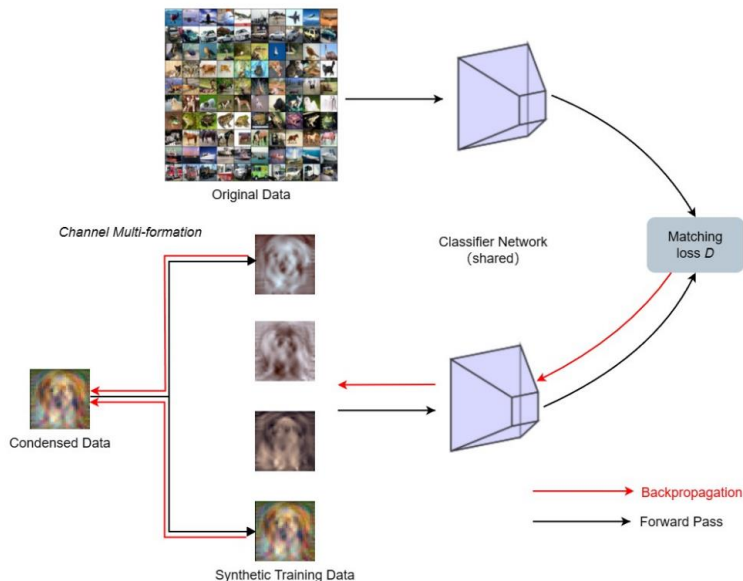
- 合成数据集的数量直接影响训练的精度
- 使用增扩方法增加数据量能有效提升数据质量

➤ 通道数据扩展模块

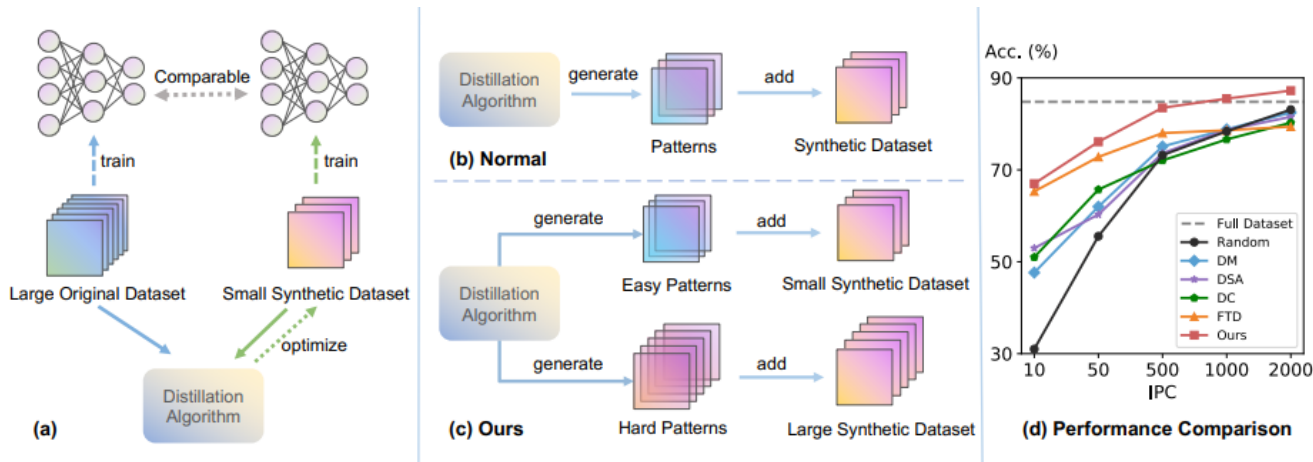
- 使用伪彩处理对每一通道进行扩展
- 伪彩处理的映射器作为元参数在训练中进行学习

➤ 即插即用

- 所提出模块适用于不同数据集蒸馏范式



基于难例轨迹对齐的无损数据集蒸馏



- 指出轨迹匹配中，早期轨迹比后期轨迹更容易学习，可通过限制轨迹匹配范围来控制对齐难度。
- 提出DIFFICULTY-ALIGNED TRAJECTORY MATCHING，依次学习容易特征和hard特征；使优化过程足够问题，以便蒸馏过程中学习软标签，从而带来性能的显著改善。实现了在CIFAR100提取到原始大小的1/5，Tiny ImageNet提取到初始大小的1/10，并且在ConvNet上训练不造成任何性能损失。

Under review as a conference paper at ICLR 2024, 初审得分2*Accept +2*Week Accept

Guo Z, Wang K, Cazenavette G, et al. Towards Lossless Dataset Distillation via Difficulty-Aligned Trajectory Matching[J]. arXiv preprint arXiv:2310.05773, 2023.

A decorative graphic in the top right corner featuring several stylized birds in flight above a series of wavy, horizontal lines that resemble a horizon or a path.

谢谢!

Smart DSP

地址：信息学院 6号楼211、201

A decorative graphic in the bottom left corner consisting of several wavy, horizontal lines that resemble a horizon or a path, mirroring the design in the top right.