北京林业大学
Beijing Forestry University

中国森林生物多样性监测网络（CForBio）讲坛



听见鸟类多样性

基于声景的鸟类多样性评估

汇 报 人	谢将剑 教授
单　　位	北京林业大学 生物多样性智慧监测研究中心
日　　期	2026年6月26日

目录

CONTENTS

1. 基于声景的鸟类多样性监测方法
2. 面向多样性评估的声景数据处理方法
3. 声景评估的缺陷、挑战与深度思考
4. 未来展望

01 基于声景的鸟类多样性监测方法

鸟类调查与监测是多样性保护的核心内容之一



反映生态系统健康状况

鸟类作为环境指示物种，对生境变化敏感。其种群结构、数量和分布变化反映生境质量、污染、资源可用性及气候变化的影响。监测是早期预警。



支持生物多样性保护与管理

监测为划定自然保护区、评估保护效用及制定特定物种管理计划提供科学依据。了解关键栖息地和迁徙通道。



深入理解生态学过程和功能

鸟类通过传粉、种子传播、害虫控制和营养循环等维持生态系统稳定。监测揭示其生态位及对生物群落的作用。

01 基于声景的鸟类多样性监测方法

传统鸟类调查方法存在限制

传统鸟类调查主要依赖样线法、样点法以及专家观测。这些方法在过去几十年中发挥了重要作用，但同时也面临一些限制：

- **调查范围有限**
  受地形、可达性等限制，空间覆盖不足
- **人力成本较高**
  需要大量专业人员投入，成本高昂
- **难以实现全天候连续监测**
  受时间、天气等限制，无法持续获取数据
- **结果受调查人员经验影响较大**
  不同人员识别能力存在差异，标准化难度大

样线法



沿固定路线行进调查，覆盖范围有限，难以全面反映区域鸟类多样性。

样点法



在固定点位定时观察，代表性受点位选择影响，难以反映大尺度变化。

专家观测



依赖专家经验识别与记录，受主观因素影响较大，难以标准化和长期化。

因此，如何实现大尺度、长期化、标准化监测，是当前生物多样性监测面临的重要问题。

- **大尺度**
  更广的空间覆盖
- **长期化**
  更长的时间序列
- **标准化**
  更统一的规范流程

01 基于声景的鸟类多样性监测方法

现实需求与政策助力推动智慧监测广泛应用

顶层战略引领

最高层级

中共中央、国务院 (2023)
《关于全面推进美丽中国建设的意见》

中共中央、国务院 (2023.7)
《习近平总书记在全国生态环境保护大全上的重要指示》

确立数字化、智能化为生态环境治理的基本方向，提供最高层级政治指引

法规制度保障

刚性约束

国务院 (2025.10.17)
《生态环境监测条例》
国务院令第820号，
自2026年1月1日起施行

首次以行政法规形式将“智能化”监测写入法律条款，提供强制性法律依据

核心任务部署

直接任务

生态环境部 (2024.3)
《关于加快建立现代化生态环境监测体系的实施意见》
环监测〔2024〕17号

生态环境部 (2024)
《国家生态环境监测网络数智化转型方案》

生态环境部、中国科学院 (2023)
《全国生态质量监督监测工作方案（2023—2025年）》

数智化转型是监测体系现代化的核心任务，明确技术路线图与时间表

生物多样性专项

领域精准对接

生态环境部等9部门 (2023)
《中国生物多样性保护战略与行动计划（2023—2030年）》

中共中央办公厅、国务院办公厅 (2021)
《关于进一步加强生物多样性保护的意见》

生态环境部等9部门 (2025)
《生物多样性保护重大工程实施方案（2025—2030年）》

智慧监测网络是生物多样性保护的关键工程，纳入重大工程系统推进

地方实践与技术规范

落地示范

浙江省丽水市市场监督管理局 (2025.4.16)
《生物多样性智慧监测体系建设技术规范》
DB3311/T 315—2025

全国首个地方技术规范，提供可操作的技术标准与实践经验



01 基于声景的鸟类多样性监测方法



Devis Tuia et al., NATURE COMMUNICATIONS, 2022(13): 792

01 基于声景的鸟类多样性监测方法

声音监测方法的优势

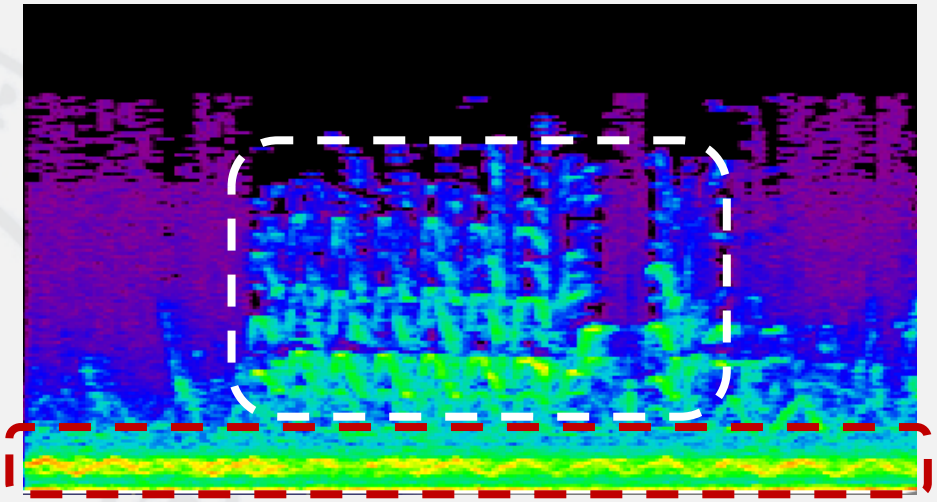


01 基于声景的鸟类多样性监测方法

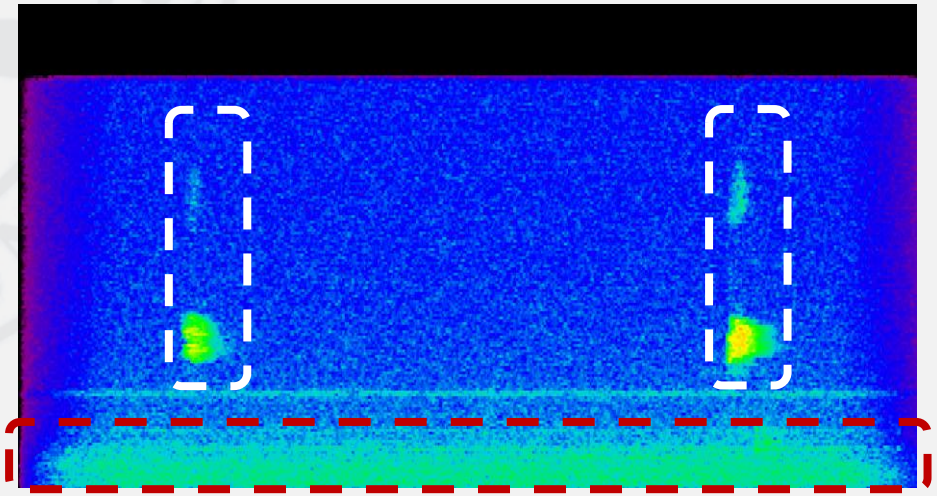
声景：源自景观的生物、地球物理和人为声音的集合



警笛声（暗红色框）
鸟鸣声（白色框）



雨声（暗红色框）
鸟鸣声（白色框）

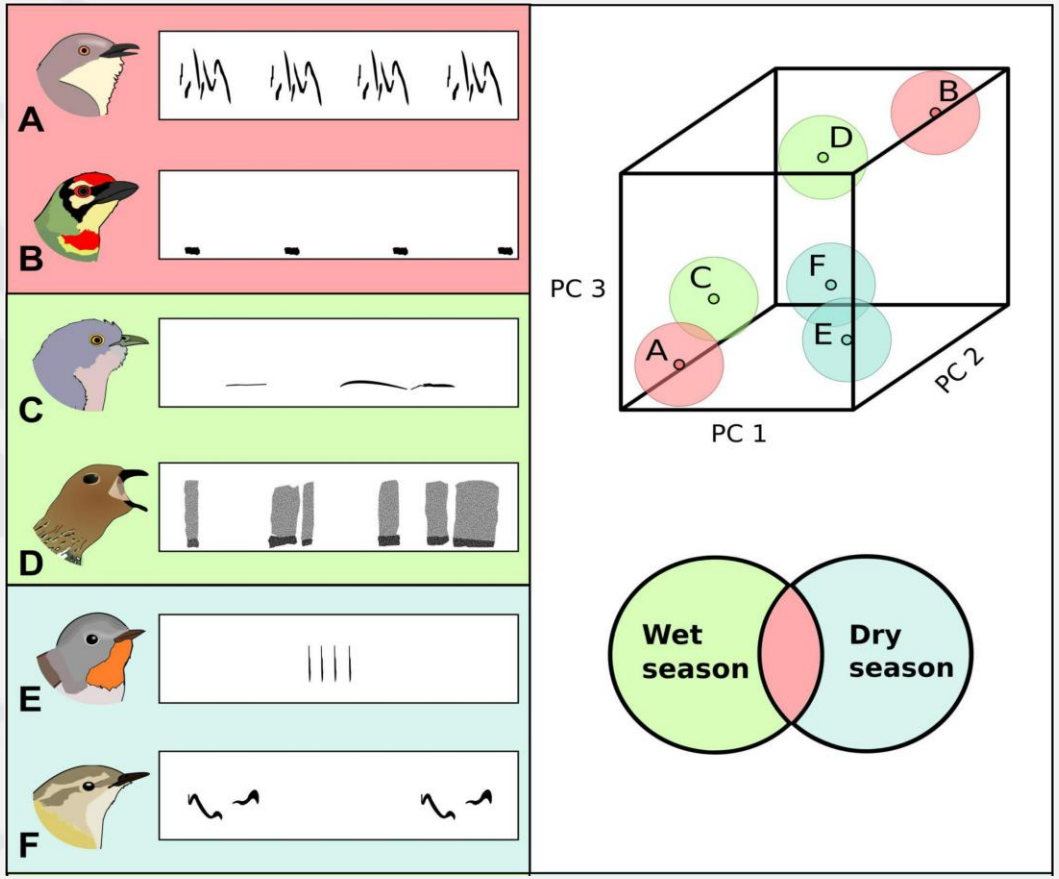


Pijanowski, B. C., et al. BioScience, 2011, 61(3), 203-216

01 基于声景的鸟类多样性监测方法

理论基础：生态声学领域的两个重要理论

- ✓ 声音生态位假说(acoustic niche hypothesis, ANH)：关注动物如何**通过声音的使用来划分生态位**，以减少声音交流中的干扰。
- ✓ 声音适应假说(acoustic adaptation hypothesis, AAH)：强调了**动物声音信号对环境声学特性的适应性**，从而提高声音信号的传播效率。
- ✓ 从理论上预测了**声音整体特征随着生物组成和生境而变化**。



01 基于声景的鸟类多样性监测方法

声景：源自景观的生物、地球物理和人为声音的集合



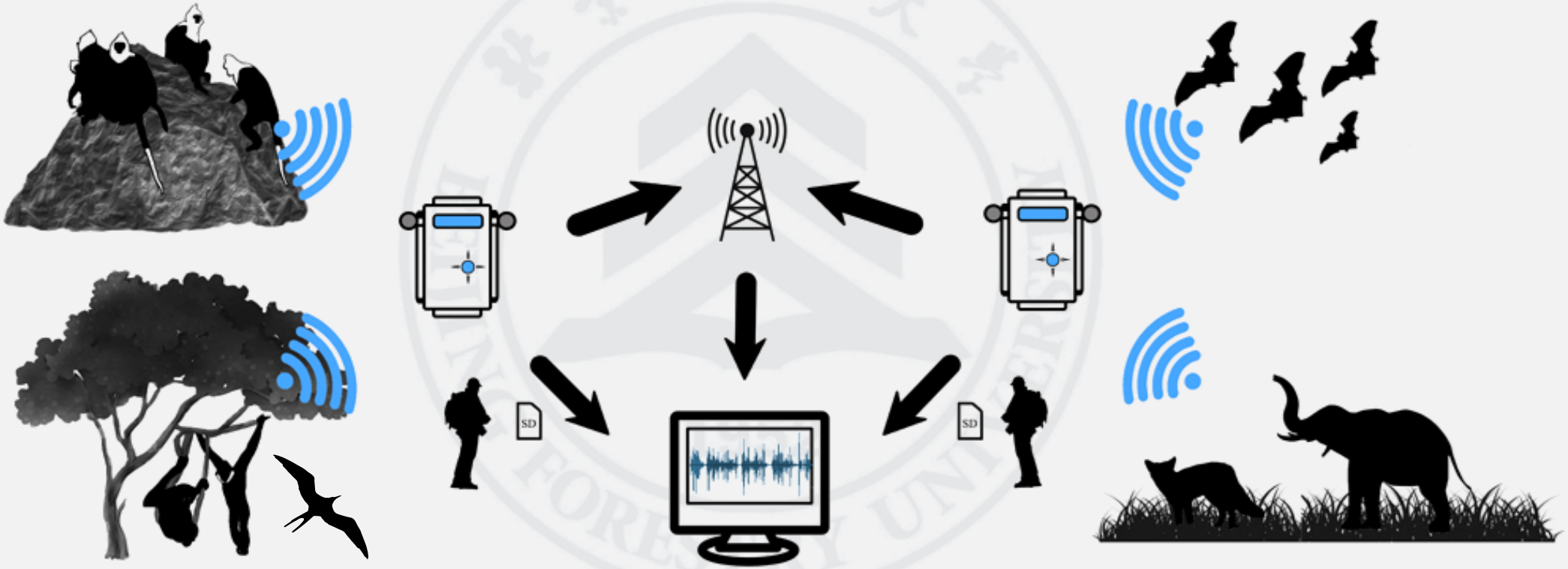
提供多层次生态信息



Sueur, J., & Farina, A. (2015). Biosemiotics, 8(3), 493–502

01 基于声景的鸟类多样性监测方法

自动声学监测有利于实现大范围长时间的生物多样性监测

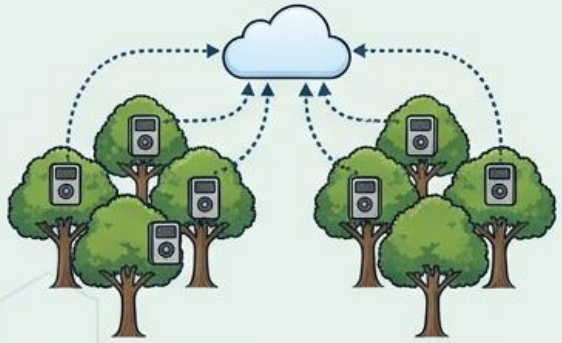


01 基于声景的鸟类多样性监测方法

国外代表性声景监测项目

1. TABMON（欧洲跨国网络）

- **覆盖地区：**挪威、荷兰、法国、西班牙（横跨纬度梯度）。
- **设备数量：**108 台全自动声学传感器（已积25TB+数据）。
- **核心目标：**建立标准化跨国 PAM 网络；利用 AI（如 BirdNet）提取鸟类群落指标； 将声学数据转化为 EBVs，填补欧盟指令报告空白。



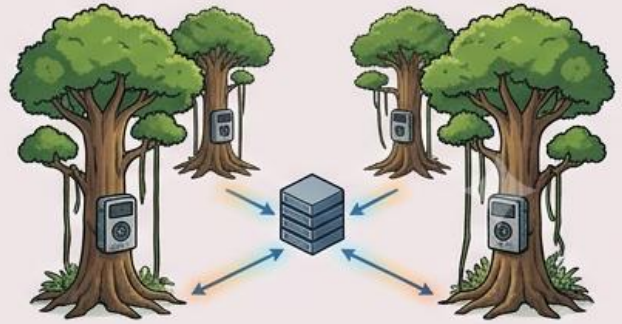
2. Australian Acoustic Observatory (A2O)

- **覆盖地区：**澳大利亚全境（大陆尺度，90 个生态区）。
- **设备数量：**360 个永久监听站（含太阳能传感器，2湿/2干）。
- **核心目标：**构建全球首个大陆尺度陆地声学观测网络；解决火灾/洪水/偏远地区的的长期生态监测问题；打造完全开放的科学平台。



3. Soundscape Baselines Project（声景基线项目）

- **覆盖地区：**全球残存原生完整森林（试点6国：文莱、厄瓜多尔等）。
- **设备数量：**数十台（试点36台以上）高级动声学记录仪（长达1年连续录音）。
- **核心目标：**抢救性建立生态多样性“声音时间胶囊”；解决“基线演变认知（Shifting Baselines）”问题；开发模块化开源协议评估保护成效。



01 基于声景的鸟类多样性监测方法

国内代表性声景监测项目

1 全国范围监测层

声景中国（Soundscape China）

全国尺度的自然声景监测网络



2 公众科学层

聆听生物圈（Listening to the Biosphere）

公众参与的声景采集与传播网络



3 区域级迁飞通道监测层

东亚–澳大利西亚迁飞通道监测网络

多站点声学与影像融合监测候鸟迁徙动态



4 保护区级监测层

典型自然保护区/国家公园声景监测

精细化、长期化的生态声景观测



1 钱江源国家公园

覆盖地区：浙江钱江源国家公园。

建设规模：依托 Sino BON 体系，将被动声学监测（PAM）与红外相机（CT）全网格同位点部署。



2 车八岭自然保护区

覆盖地区：广东车八岭国家级自然保护区。

建设规模：100+ 智能监测节点，构建“物联网 + 无人值守地面终端 + 云端 AI”的近实时网络。



3 东北虎豹国家公园：超大规模天空地一体化子网

覆盖地区：吉林、黑龙江两省交界（横跨上万平方公里温带针阔混交林）。

建设规模：数万个被动声学监测（PAM）与红外相机（CT）。



4 北京林业大学：东亚-澳大利西亚通道候鸟迁徙智能监测网

覆盖地区：纵贯我国东部迁飞关键节点（黑龙江扎龙、山东黄河口、上海崇明东滩、江苏盐城条子泥、江西鄱阳湖）。

建设规模：跨区域多节点部署的“声学监听 + 边缘计算摄像头”联合智能监测网络。



02 面向多样性评估的声景数据处理方法

基于声景的生物多样性评价

评价	用途
物种丰富度 Species richness	被动声学记录中物种数量，用于确定当地群落的多样性
物种丰度 Species abundance	动物群体中发出信号的个体数量，用作种群数量指标
发声丰度 Abundance of sounds	特定物种或特定类型声音的数量，用于确定声学活动的强度
物种均匀度 Species evenness	物种个体数分布的均匀程度，用于反映群落中物种的均衡性
声学活动模式 Acoustic activity patterns	描述物种或群落声学活动的时间和空间模式，用于比较物种的鸣叫物候期
声景组成 Soundscape composition	确定声源（如人声和生物声）对声景的相对贡献，用于描述声景的结构和动态
环境类型 Environmental type	区分生境、生态系统和景观声学的环境特征，用于表示其结构和动力学特征

02 面向多样性评估的声景数据处理方法

有效处理海量数据成为核心挑战

中国生物多样性研究的30个核心问题

张健[1], 孔宏智[2], 黄晓磊[3], 傅声雷[4], 郭良栋[5], 郭庆华[6], 雷富民[7], 吕植[8], 周玉荣[2], 马克平[2*]

1. 华东师范大学生态与环境科学学院, 浙江天童森林生态系统国家野外科学观测研究站/全球变化与复杂生态系统研究中心, 上海 200241; 2. 中国科学院植物研究所, 北京 100093; 3. 福建农林大学植物保护学院闽台作物有害生物生态防控国家重点实验室, 福州 350002; 4. 河南大学地理与环境学院, 河南开封 475004; 5. 中国科学院微生物研究所真菌学国家重点实验室, 北京 100101; 6. 北京大学地球与空间科学学院遥感与地理信息研究所, 北京 100871; 7. 中国科学院动物研究所, 北京 100101; 8. 北京大学生命科学学院自然保护与社会发展研究中心, 北京 100871

(26)如何开发和推广适用于生物多样性调查、监测和研究的新方法与新技术?

(27)如何结合多尺度、多类型的监测数据提高对生物多样性相关属性的预测精度?

张健, 孔宏智, 黄晓磊, 傅声雷, 郭良栋, 郭庆华, ... & 马克平. (2022). 中国生物多样性研究的 30 个核心问题. 生物多样性, 30(10), 22609.

02 面向多样性评估的声景数据处理方法

三种声景数据处理方法

人工处理

- 动物学家
- 公众科学家
- 观鸟人

自动识别

- 声景类型识别
- 物种识别
- 个体识别

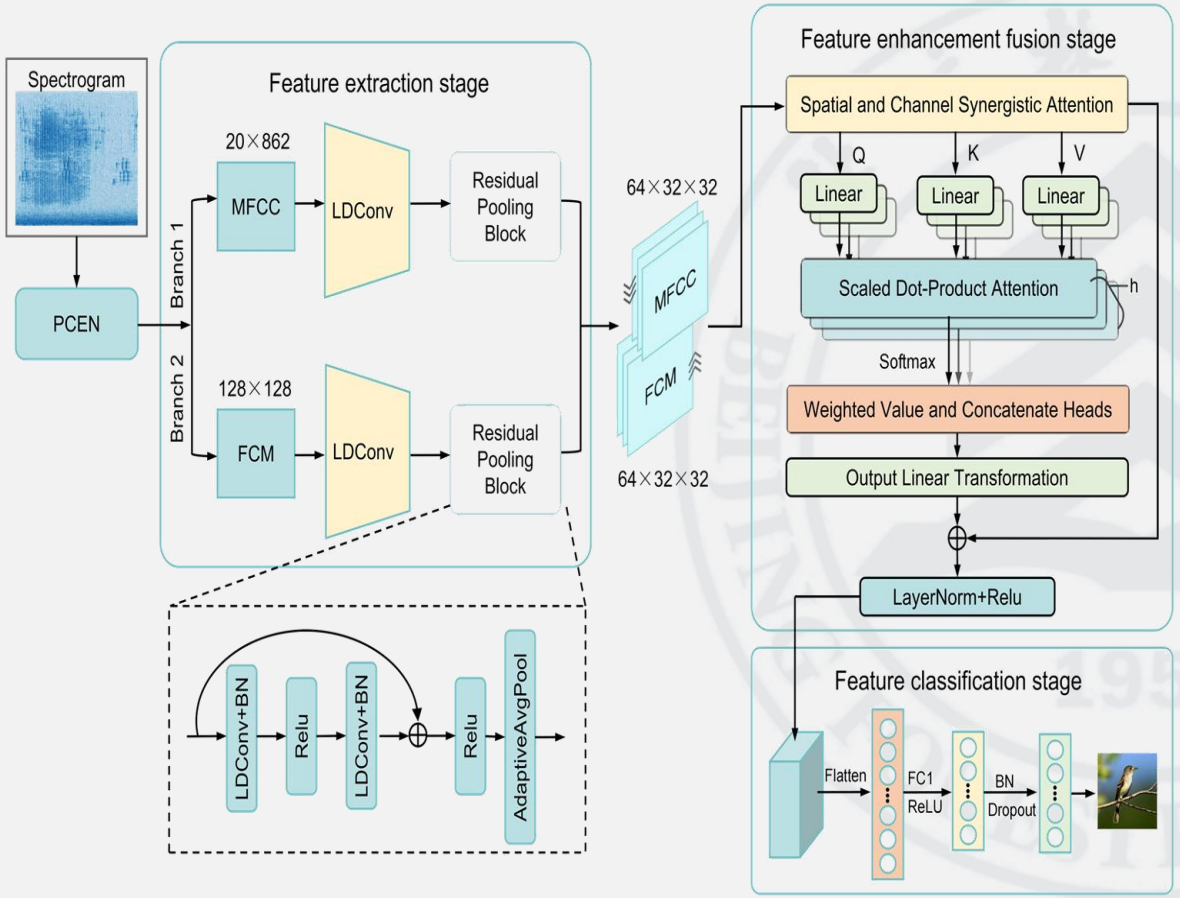
声学指数

- 设计指数
- 寻找指数和多样性的关联

Xie, J., et al. , 2025, Artificial Intelligence Review, 59(1), 10.

02 面向多样性评估的声景数据处理方法：自动识别

融合频率相关性和梅尔特征的边缘鸟类音频检测方法



DCASE2018 数据集上的检测效果

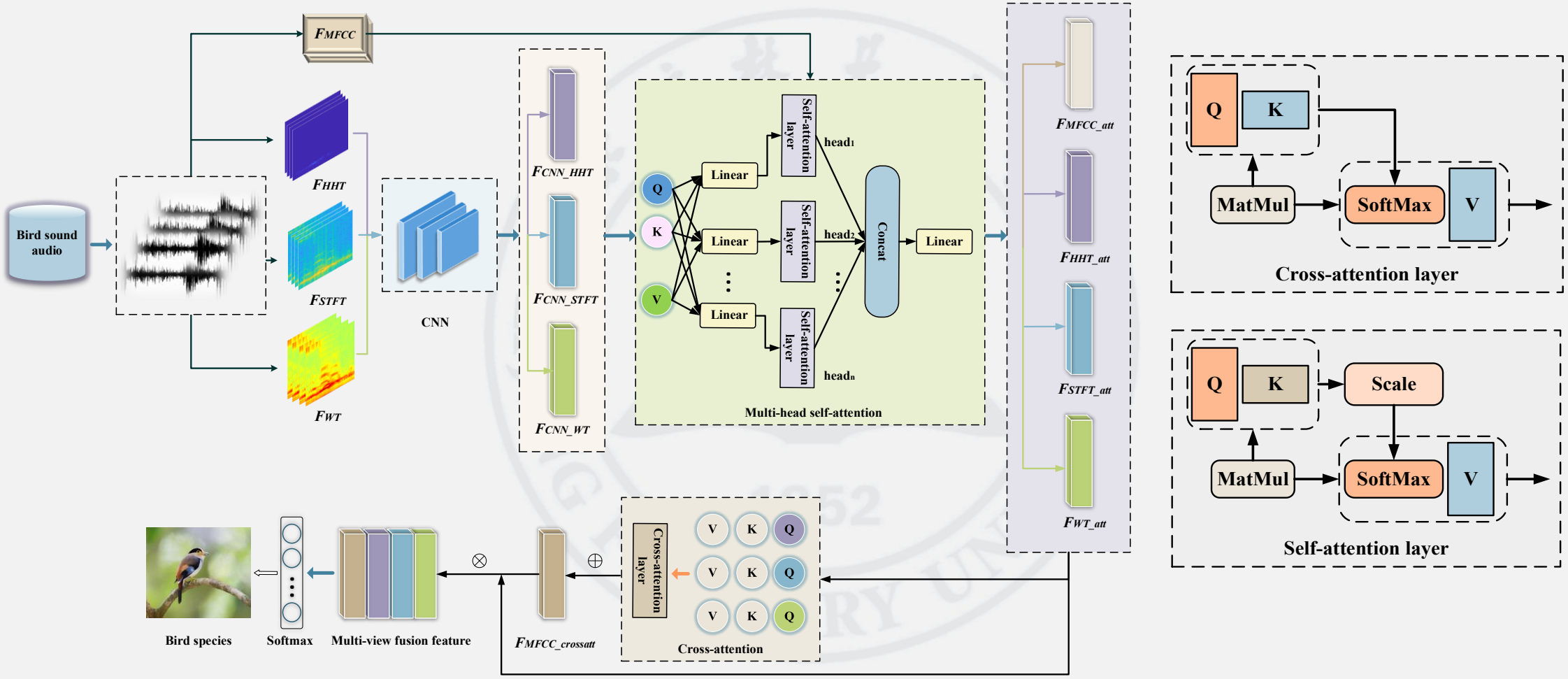
ACC and AUC results for different methods (Value: ACC|AUC).

Methods	BirdVox-DCASE-20k	Freefield1010	Warblrb10k	Total
LDRCM	72.6%|0.725	86.0%|0.755	85.3%|0.765	81.3%|0.748
EfficientNet (Tan and Le, 2019)	80.9%|0.815	86.3%|0.876	84.0%|0.909	83.7%|0.867
ERes2Net (Chen et al., 2023)	79.3%|0.876	86.9%|0.888	85.6%|0.917	83.9%|0.894
CAM++ (Wang et al., 2023)	84.6%|0.913	86.3%|0.892	85.8%|0.927	85.6%|0.911
DFEFM	**89.0%|0.954**	**92.7%|0.962**	**92.5%|0.973**	**91.4%|0.963**

Yingqi Wang, et al. Avian Research, 2025, 16(2): 100232.

02 面向多样性评估的声景数据处理方法：自动识别

基于多视图双注意力融合MDF-Net的鸟鸣声识别方法



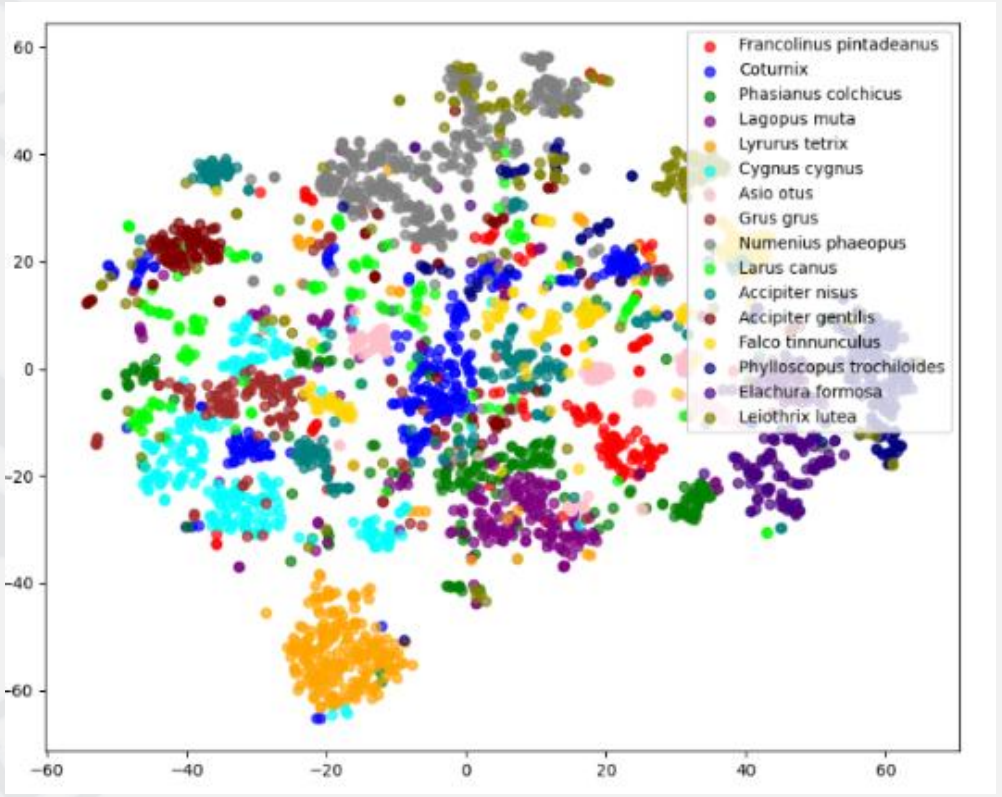
XIE S, et al. Applied Acoustics, 2024, 225: 110138.

02 面向多样性评估的声景数据处理方法：自动识别

基于多视图双注意力融合MDF-Net的鸟鸣声识别方法

不同特征在16种鸟类的测试集上的分类精度

Features	Acc	wAP	wAR	wAF
F_{HHT_att}	84.2	84.23	84.2	84.18
F_{STFT_att}	90.24	90.3	90.24	90.23
F_{WT_att}	90.79	90.83	90.79	90.77
F_{MFCC_att}	85.59	85.53	85.59	85.4
$F_{MFCC_crossatt}$	87.33	87.49	87.33	87.25
$F_{HHT_STFT_WT_MFCC}$	93.50	93.51	93.50	93.49
$F_{HHT_STFT_WT_att}$	93.54	93.57	93.54	93.52
F_{att}	95.7	95.77	95.7	95.69
$\boldsymbol{F_{dual_att}}$	**97.29**	**97.32**	**97.29**	**97.28**

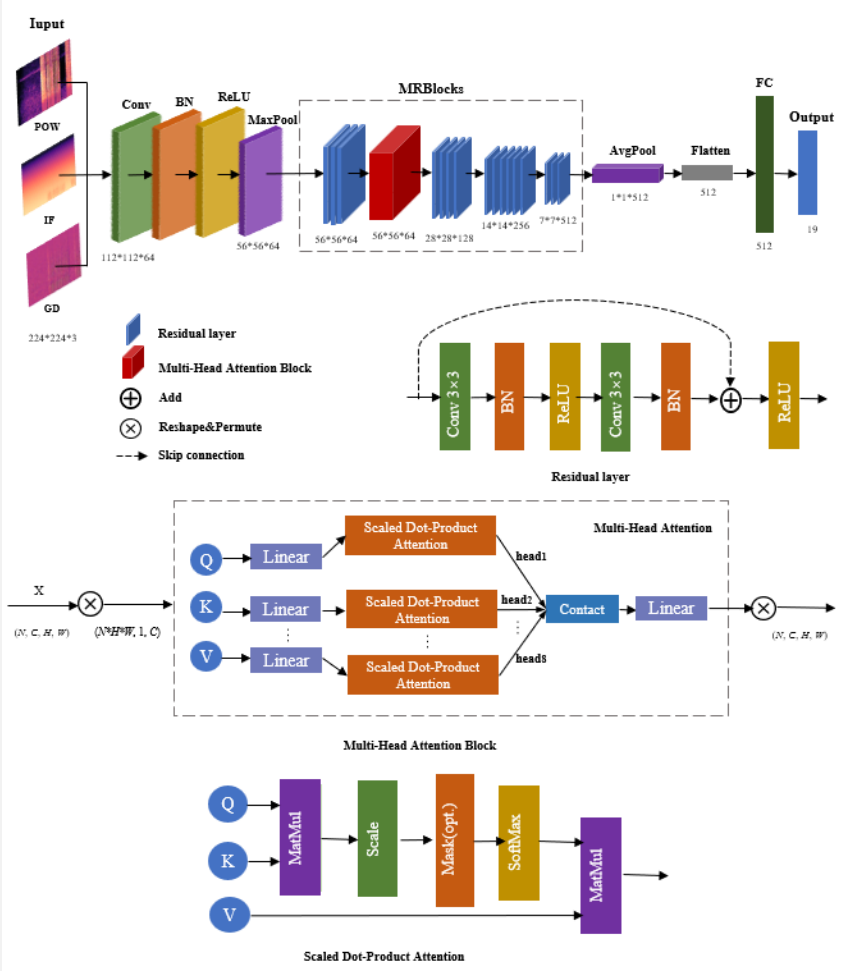


特征的t-SNE可视化

XIE S, et al. Applied Acoustics, 2024, 225: 110138.

02 面向多样性评估的声景数据处理方法：自动识别

综合幅值相位信息的鸟类物种识别方法



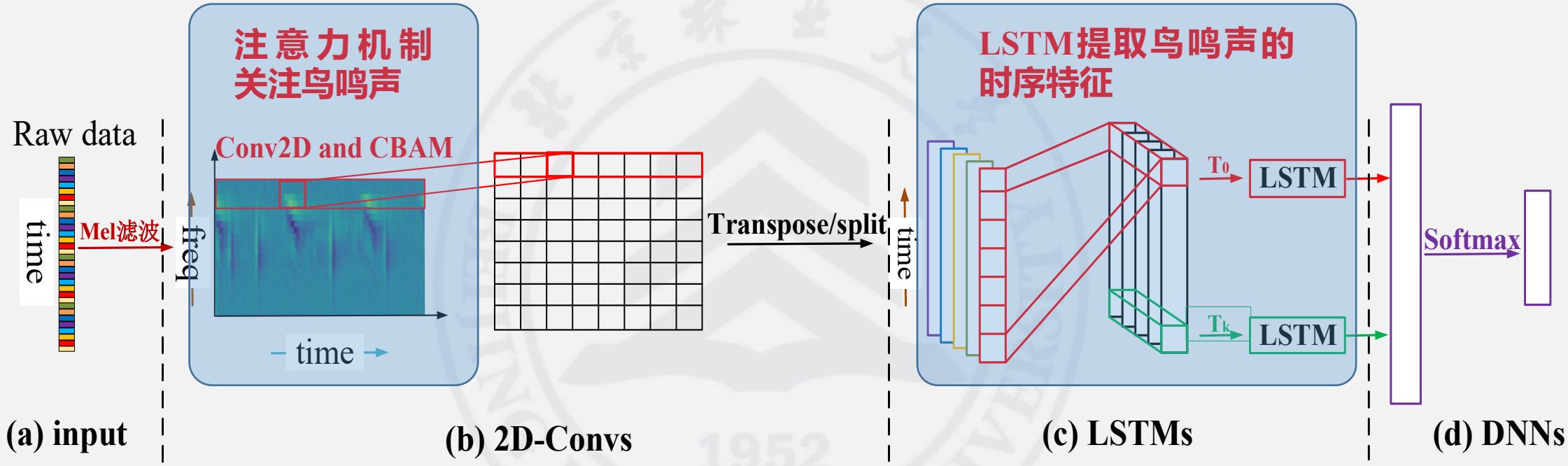
Feature	Accuracy	Recall	Precision	F1-Score
POW	93.1%	92.9%	93.1%	93.1%
IF	85.4%	80.5%	80.6%	80.9%
GD	81.2%	81.1%	81.3%	81.4%
POW+IF	90.2%	90%	90.1%	90.1%
POW+GD	90.9%	90.5%	90.9%	90.9%
IF+GD	84.1%	83.7%	83.8%	83.9%
POW+IF+GD	**94.0%**	**93.9%**	**94.1%**	**93.8%**

Model	Accuracy
GMM	61.1%
HMM	63.2%
ANN	69.3%
ResNet18	88.3%
ResNet34	89.5%
ResNet50	86.6%
ViT	82.8%
BirdNET	82.5%
AMResNet	92.6%
MHAResNet (proposed)	**94.0%**

Xie J, et al. Avian Research, 2025, 16(1): 100229.

02 面向多样性评估的声景数据处理方法：自动识别

基于KD-CLDNN的轻量级鸟类物种识别方法

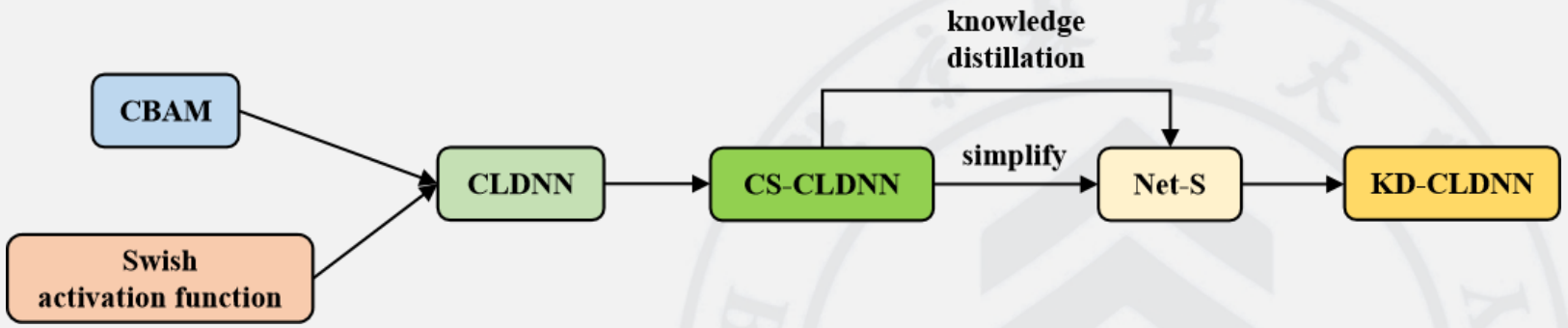


基于 CS-CLDNN(CBAM-Switch-CLDNN)的鸟类物种识别模型

XIE J, et al., Applied Acoustics, 2022, 188: 108550.

02 面向多样性评估的声景数据处理方法：自动识别

基于KD-CLDNN的轻量级鸟类物种识别方法



The performance of three models on the test dataset

Network	Accuracy	Parameter (M)
CS-CLDNN	**0.975**	311.8
Net-S	0.918	97.6
KD-CLDNN	0.946	**97.6**

Average inference time on NVIDIA JETSON TX2

Model	Inference time (s)
CS-CLDNN	0.37
Net-S	0.11
KD-CLDNN	0.11

中文名	音频总数
非洲斑鹑鸽	701
家燕	1887
红嘴鸥	1129
黑啄木鸟	1237
加拿大鹅	1429
小嘴乌鸦	648
煤山雀	1544
乌鹟	1223
苍头燕雀	910
叽咋柳莺	1401
棕柳莺	1792
白腹毛脚燕	1657
红雀	1164
黑水鸡	877
夜莺	2330
雉鸡	837
白腰朱顶雀	807
赤足鹬	898
红尾鸲	1852
芦苇旗	1402

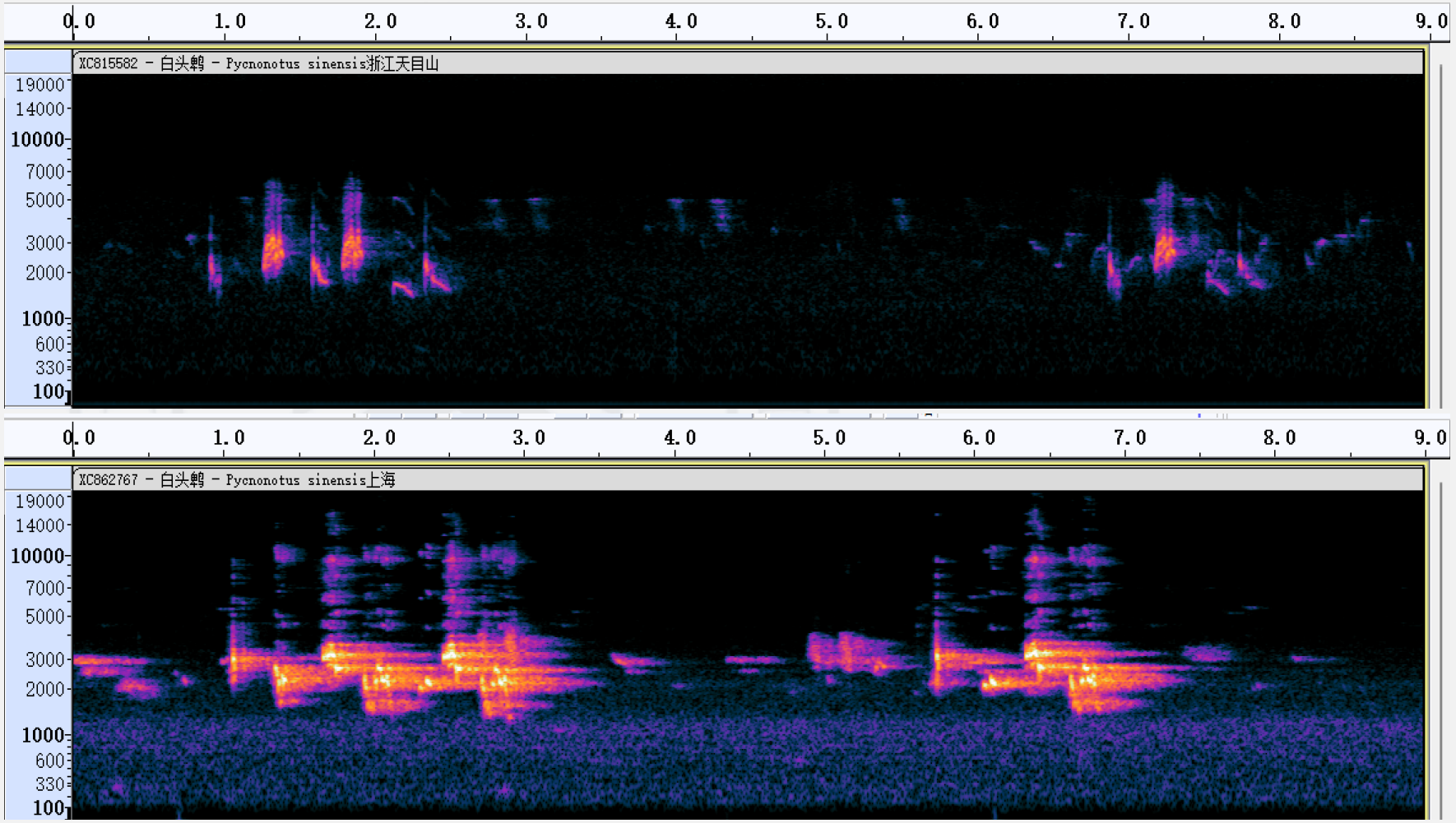
XIE J, et al., Applied Acoustics, 2022, 188: 108550

02 面向多样性评估的声景数据处理方法：自动识别

方言（鸟鸣声地理变异）的影响

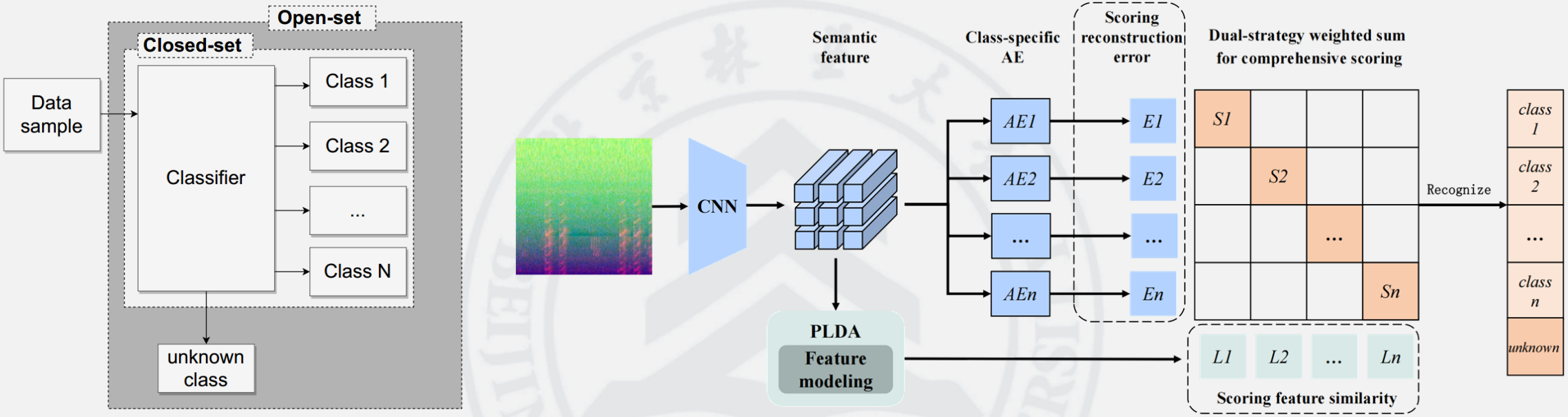


白头鹎
（ Pycnonotus sinensis ）



02 面向多样性评估的声景数据处理方法：自动识别

基于语义特征重构与双策略评分的鸟鸣声开集-跨域识别方法

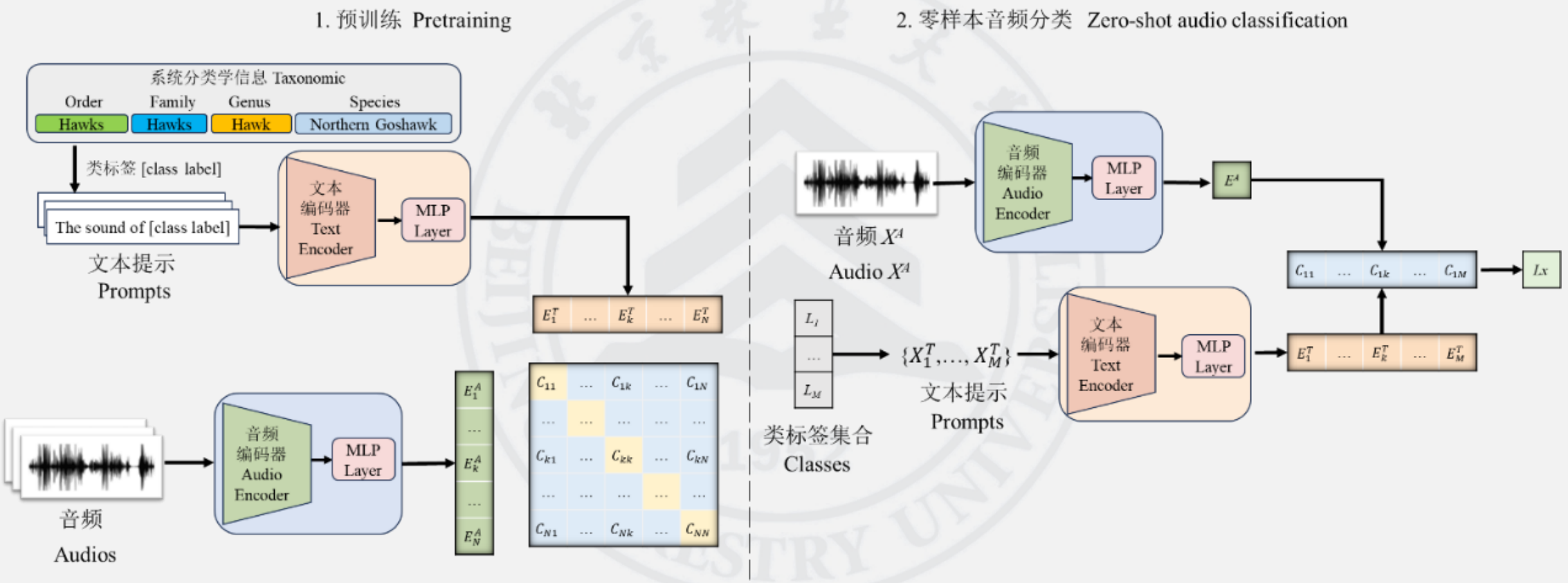


Method	Data	$ACC_k(\%)$	$ACC_u(\%)$	$ACC(\%)$	Data	$ACC_k(\%)$	$ACC_u(\%)$	$ACC(\%)$
Xie *et al.*(2023)	S1S2	50.90	56.67	53.79	S2S1	50.51	75.39	62.95
RES+CEL	S1S2	54.46	67.78	61.12	S2S1	**53.74**	82.12	67.93
RES+CL	S1S2	54.08	83.33	68.71	S2S1	50.58	85.49	68.04
DSCS+CEL	S1S2	44.49	57.41	50.95	S2S1	50.86	80.83	65.85
DSCS+CL	S1S2	**56.97**	**84.81**	**70.89**	S2S1	51.20	**87.09**	**69.13**

J. Xie, et al., IEEE Signal Processing Letters, 2025,3549008

02 面向多样性评估的声景数据处理方法：自动识别

基于系统分类信息的鸟类音频零样本分类方法



谢珊珊等, 林业科学, 2025, 61(2): 12-20

02 面向多样性评估的声景数据处理方法：自动识别

基于系统分类信息的鸟类音频零样本分类方法

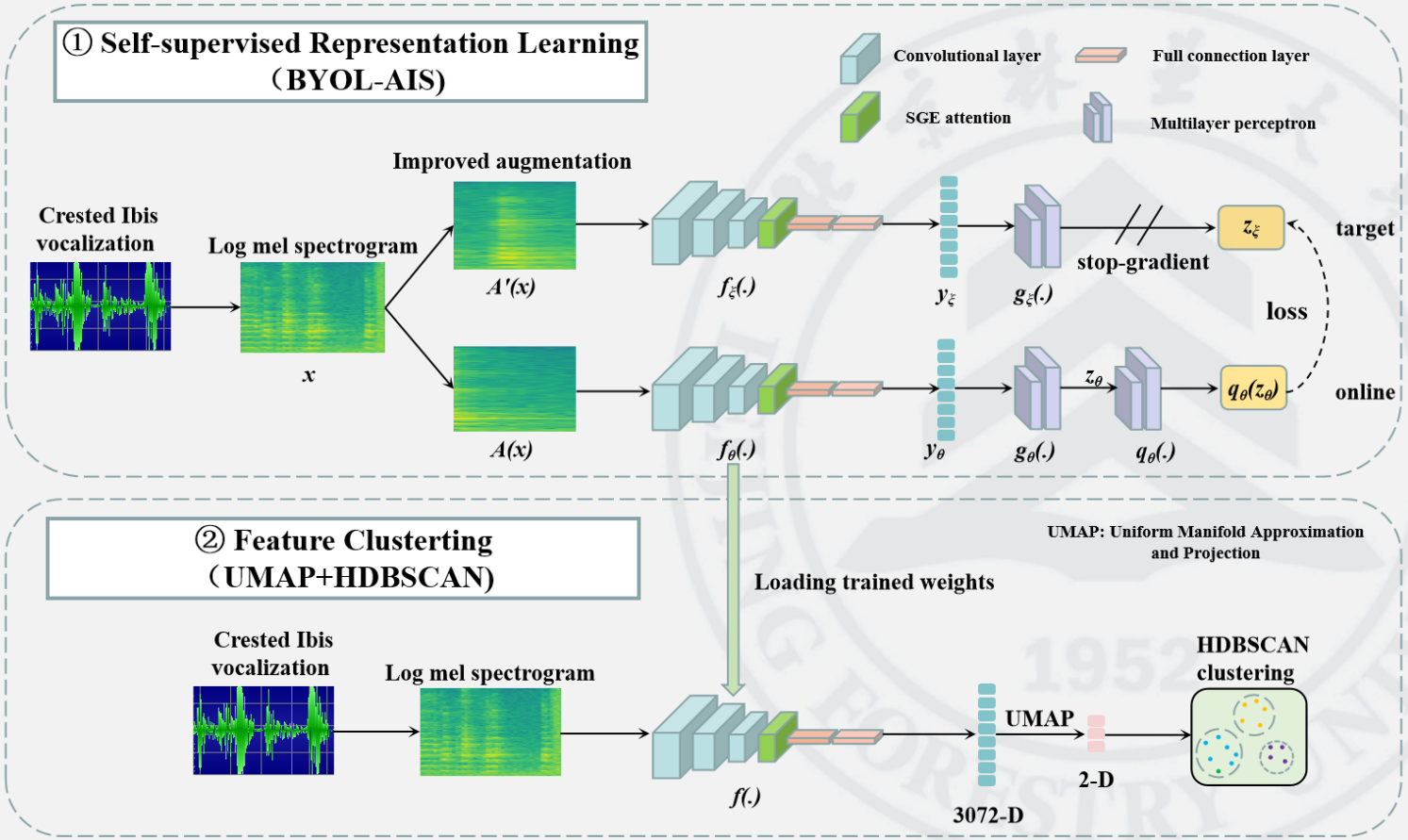
训练集（77类）、开发集（9类）和测试集（9类）

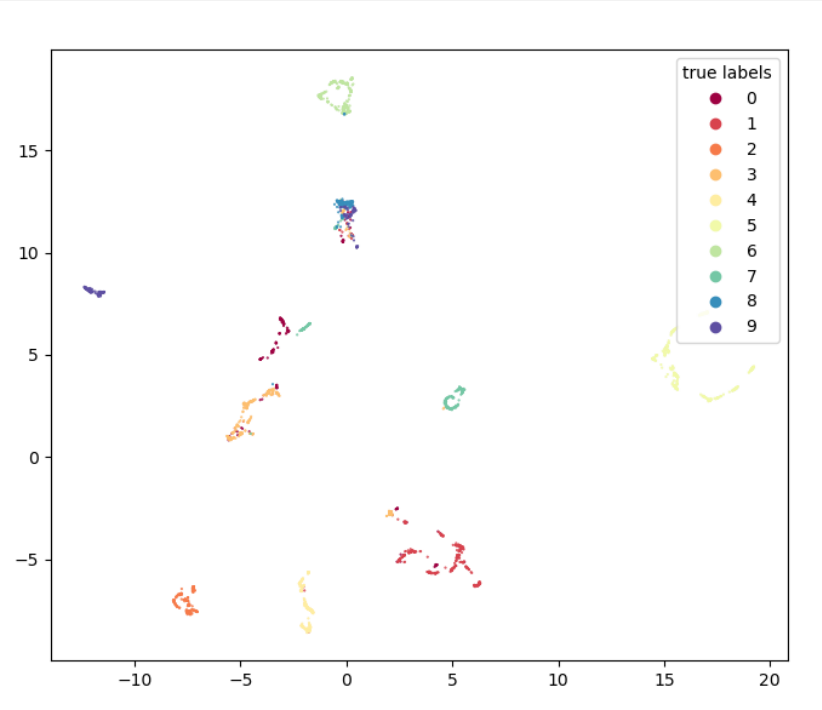
模型 Model	物种类别辅助信息 Side information for species class	开发集 Development 准确率 ACC	开发集 Development 平均召回率 UAR	开发集 Development 平均 F1_score F1_score	测试集 Test 准确率 ACC	测试集 Test 平均召回率 UAR	测试集 Test 平均 F1_score F1_score
CLAP-Laion	学名 Scientific name	0.191	0.134	0.061	0.131	0.111	0.052
CLAP-Laion-LoRA	学名 Scientific name	0.281	0.270	0.245	0.219	0.230	0.196
CLAP-Laion-LoRA	系统分类 Taxonomic	0.372	0.305	0.305	0.269	0.275	0.249
AST-LN	鸟类基础特性和生活史 AVONET+BLH	0.335	0.281	0.244	0.287	0.295	0.233
CLAP-Bird	学名 Scientific name	0.322	0.279	0.266	0.288	0.244	0.244
CLAP-Bird	系统分类 Taxonomic	**0.395**	**0.302**	**0.307**	**0.323**	**0.291**	**0.289**

谢珊珊等, 林业科学, 2025, 61(2): 12-20

02 面向多样性评估的声景数据处理方法：自动识别

基于自监督聚类算法的鸟类个体鸣声识别方法



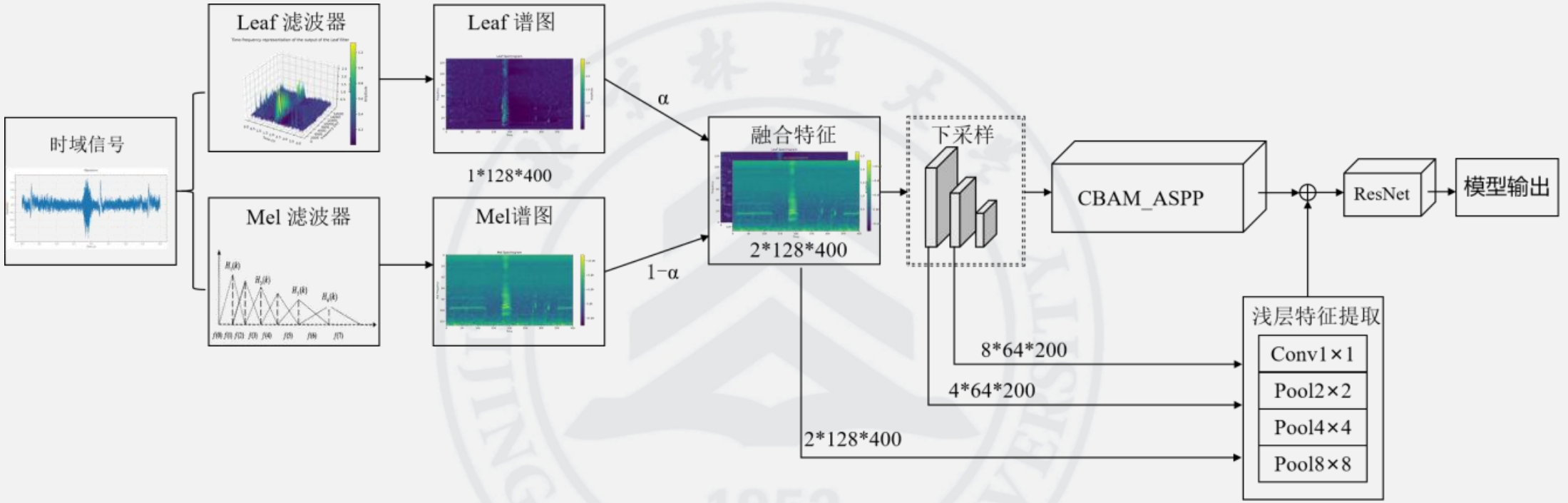


聚类准确率0.8642

XIE J, et al., Ecological Informatics, 2023, 75, 102089

02 面向多样性评估的声景数据处理方法：自动识别

基于多特征多尺度的鸿雁鸣声分类



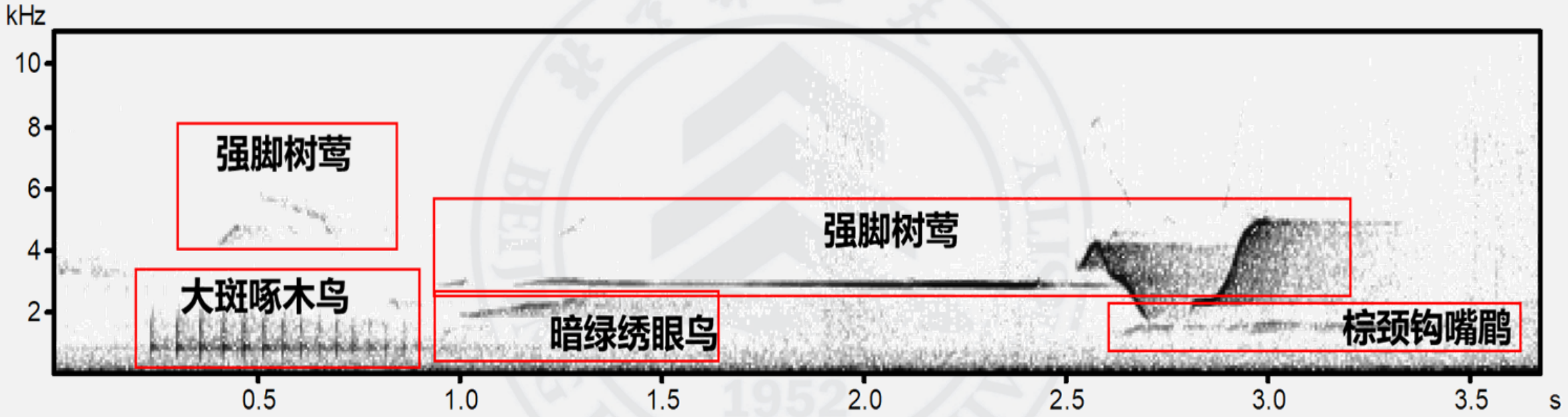
声音类型	英文名	时长/s
联络	Calling	1208
飞行	Flying	640
觅食	Foraging	588

方法	准确率%	F1分数
TCNN	76.65	0.76
KD-CLDNN	83.70	0.84
BirdNet	77.05	0.75
本文方法	**85.48**	**0.85**

石煜炜, 鸿雁鸣声检测与多尺度特征融合分类研究(2025硕士论文)

02 面向多样性评估的声景数据处理方法：声源分离

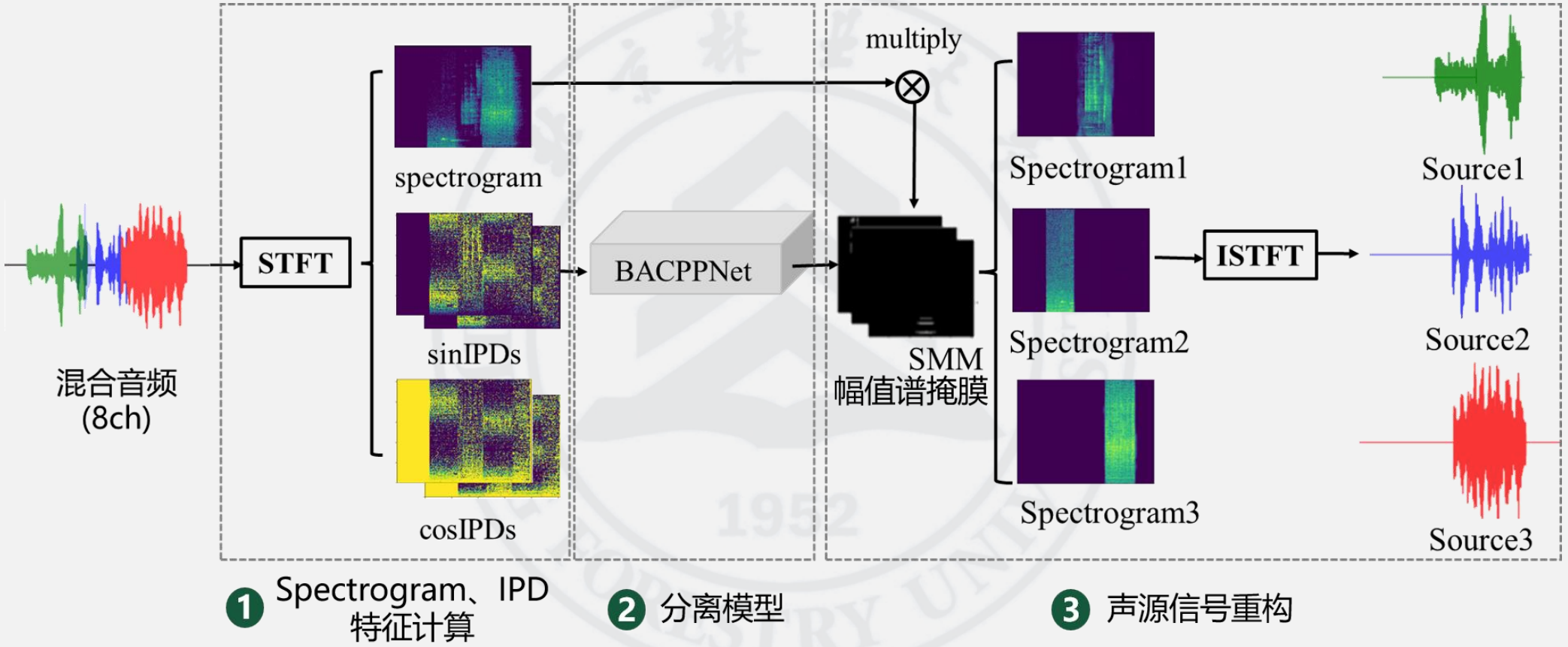
基于BACPPNet的鸟类鸣声自动分离方法



图片来自北京师范大学夏灿伟

02 面向多样性评估的声景数据处理方法：声源分离

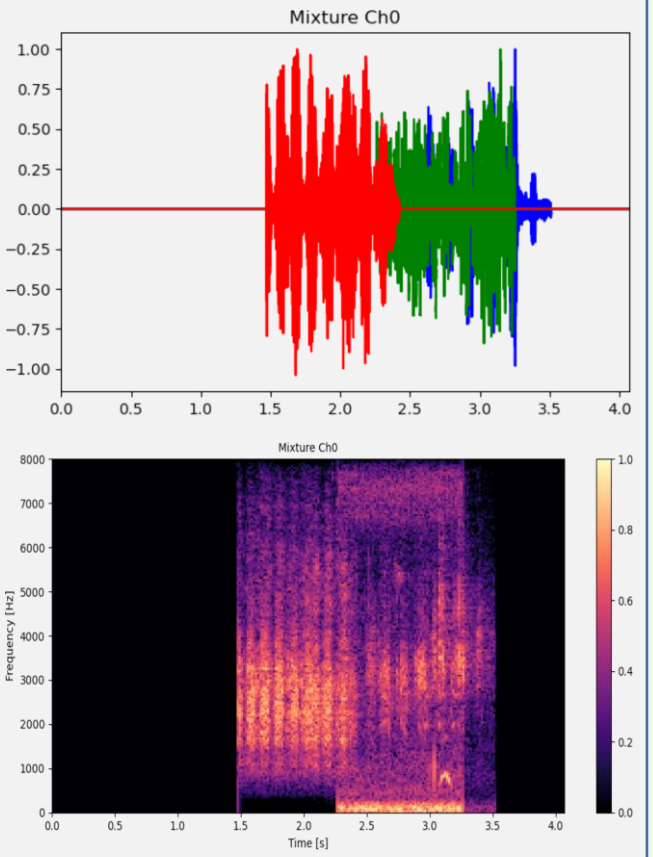
基于BACPPNet的鸟类鸣声自动分离方法



Xie J., IEEE Internet of Things Journal, 2024, 11(9), 16604-16617

02 面向多样性评估的声景数据处理方法：声源分离

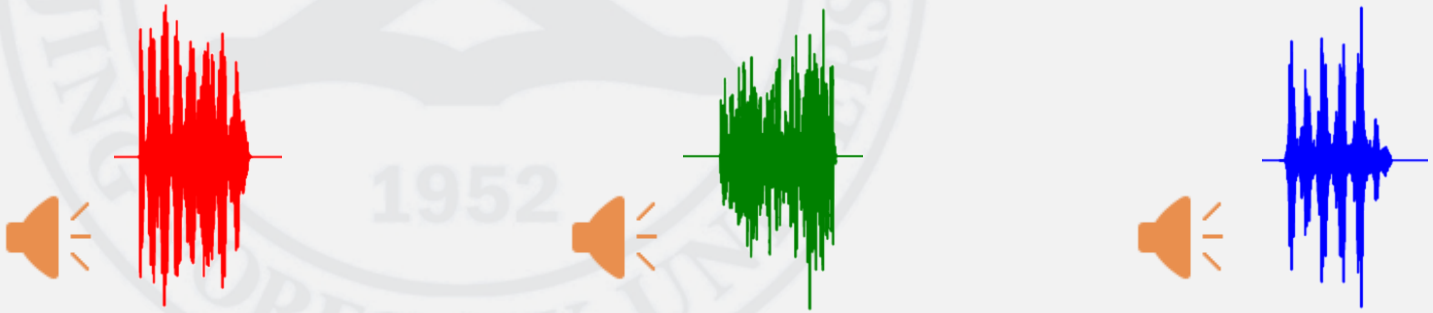
混合音频



混合前声源信号



90-135°-Magpie　　0-45°-Common Cuckoo　　225-270°-Reed Warbler



分离后声源时域信号

02 面向多样性评估的声景数据处理方法：声学指数

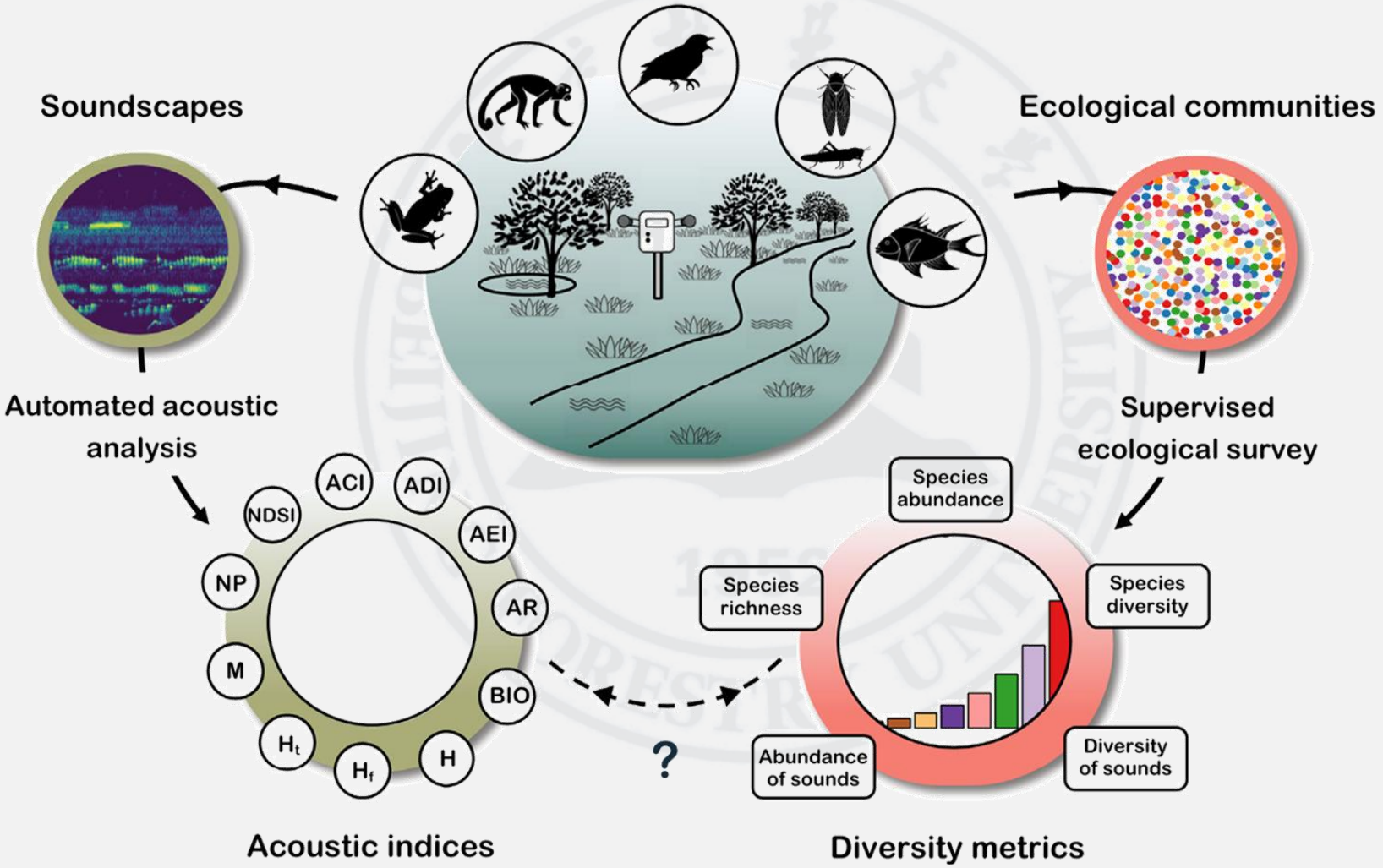
声学指数

- ✓ **alpha声学指数：反映录音内的信息，帮助研究者了解特定环境中的声音结构和生物声学多样性。**
- ✓ **beta声学指数：比较不同录音之间的差异，研究声音景观的时空变化，以及生物群落之间的声学差异。**
- ✓ **alpha声学指数的受关注程度要远高于beta声学指数。**

声学指数 Acoustic index	计算公式 Computing formula	描述 Description	参考文献 Reference
声学复杂度指数 Acoustic complexity index (ACI)		相邻频段窗口音量的变化; 相邻频段音量差的累积(D); 各频段的音量(I_k) The difference in amplitude among samples; summary of intensity difference among adjacent frequency bins (D); intensity in a single frequency bin (I_k)	2011
声学多样性指数 Acoustic diversity index (ADI)	$\Sigma p_i \times log p_i$	基于Shannon指数量化音量(p_i)在不同频段的分布 Reflection of spectral complexity based on Shannon index; relative intensity in each frequency bin (p_i)	2011
声学均匀度指数 Acoustic evenness index (AEI)		基于Gini指数量化音量(I)在不同频段的分布 Gini coef?cient with intensity (I) at each frequency bin	2011
声学丰富度指数 Acoustic richness index (AR)	$[rank(H_t) \times rank(M)]/n^2$	时间熵(H_t)和音量(M)的乘积, 除以频段数量(n) Product of time entropy (H_t) and intensity (M) divided by the number of bins (n)	2012
生物声学指数 Bioacoustic index (BI)		特定频段音量(I_k)的汇总 The sum of amplitude (I_k) in particular frequency band	2007
声学熵指数 Acoustic entropy index (H)	$H_t \times H_f$	时间熵(H_t)和频谱熵(H_f)的乘积 Product of time entropy (H_t) and spectral entropy (H_f)	2008
标准化声景差异指数 Normalized difference soundscape index (NDSI)		生物产生音量(b)与人类产生音量(a)的比率 Ratio of amplitude in biophony (b) and anthropophony (a)	2012
频峰数 Number of peaks (NP)	Σf_i	频谱上频峰(f_i)的数量 The number of frequency peaks (f_i) on spectrum	2013

02 面向多样性评估的声景数据处理方法：声学指数

关键问题：声学指数能否与生物组成建立联系



Alcocer, I, et al., Biological Reviews,2022, 97: 2209–2236

02 面向多样性评估的声景数据处理方法：声学指数

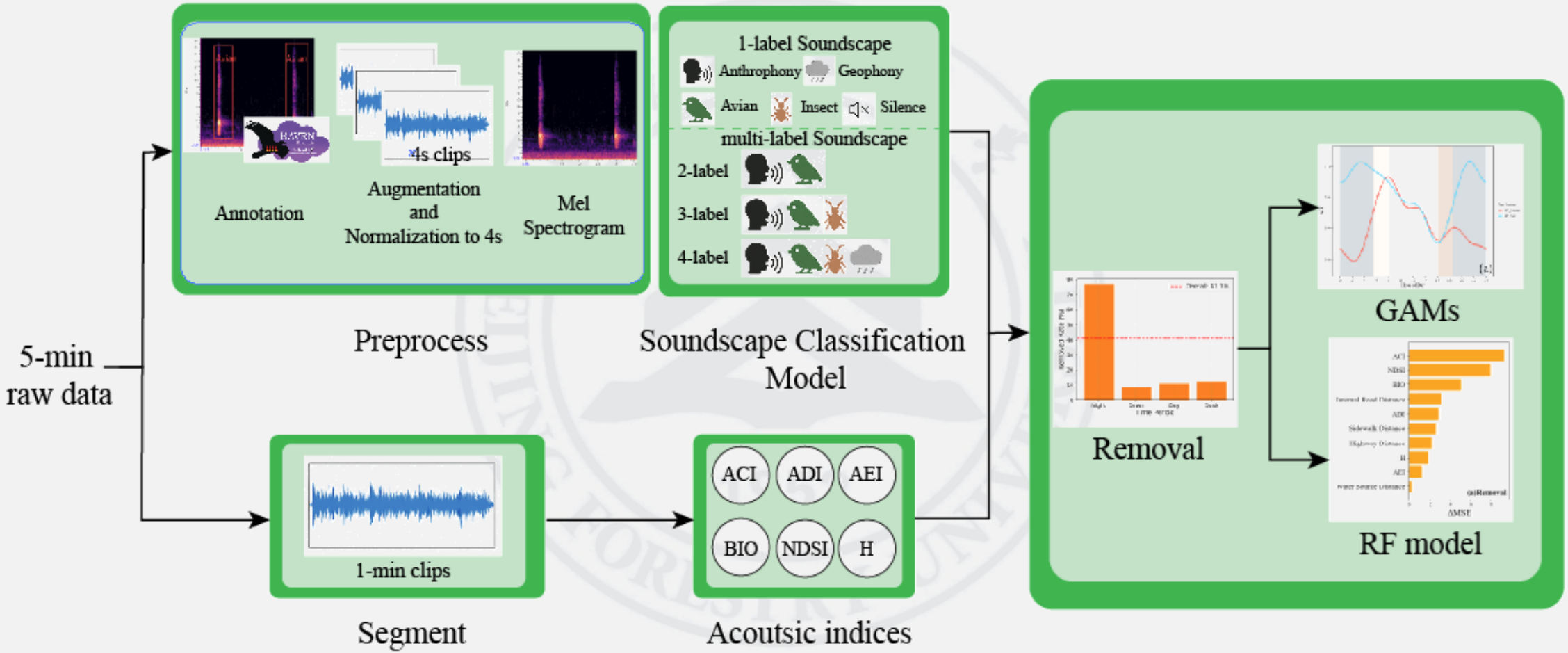
声学指数的困境（以ACI为例）

多样性评价	生态系统	是否关联	文献
鸟类丰富度 Bird richness	原始森林和采伐林 old growth and logged forests	强关联	Mitchell et al., Ecological Indicators, 106717, 2020
鸟类丰富度 Bird richness	草地 grassland in the Northern Great Plains	强关联	Shamon et al., Ecological Indicators, 106928, 2021
鸟类丰富度 Bird richness	棕榈油树林		Mitchell et al., Ecological 17, 2020
物种丰富度 Species richness			al., Ecological 37, 2023
物种丰富度 Species richness	热带岛屿 tropical land-bridge islands in an Amazonian hydroelectric reservoir	弱关联	Bicudo et al., Ecological Informatics, 102241, 2023.
鸟类物种丰富度 Bird species richness	原始森林 Białowieża Forest	弱关联	Budka et al. Ecological Indicators, 110027, 2023
鸟类物种丰富度 Bird species richness	亚热带和热带森林 subtropical and tropical forests in Guangxi Province	弱关联	Galappaththi et al., Global Ecology and Conservation, e02922, 2024

非生物声音对声学指数有负面影响！！！

02 面向多样性评估的声景数据处理方法：声学指数

声学指数的改进（移除非生物声音的影响）



Jiang, Q., et al., 2026, Avian Research, 100361

03 声景评估的缺陷、挑战与深度思考

监测中的“隐形偏差”（生态状态、传播损失、噪声掩蔽）

“完美耳朵”的迷思



录音机不具备“完美听觉”

在一个地点进行录音，并不能保证捕捉到由于固有的采样偏差而存在的所有物种。

可检测性的“隐形偏差”

不完美的检测会导致假阴性，从而扭曲对物种丰度和分布的估计。



扭曲的生物多样性数据

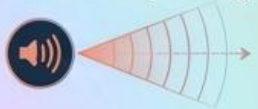
对检测空间的误解会导致生物物候期记录不完整，以及对关关键栖息地的识别不当。



检测空间解剖学

传播损耗因素

Geometric Spreading



Atmospheric Absorption



Habitat Barriers



衰减类型	主要原因	对检测的影响
几何	距声源距离	随声波扩散的标准减少
大气	距声源距离	空气中声波振动的吸收
Atmospheric	温度和湿度	空气中声波振动的吸收
栖息地	植被和地形	物理障碍物引起的波散射和吸收



噪声的掩蔽效应

环境噪声（风、雨、交通）会“隐藏”生物信号，有效地缩小录音机的范围。



可检测性的三种状态

可检测性由生态状态、信号状态和观测过程共同决定。

03 声景评估的缺陷、挑战与深度思考

两种主流声景分析方法的缺陷

为什么现有方法难以准确评估生物多样性？

① 基于声学指数的方法

如 ACI、ADI、BIO 等

简单、快速、可规模化，但生物学含义模糊，
对非生物和环境因素高度敏感。

1 易受非生物声音污染



风声、雨声、昆虫、交通等地球声学
或人为噪声会显著抬高指数值。
高指数 ≠ 高生物活跃度。

2 强烈依赖生境，难以跨区域比较



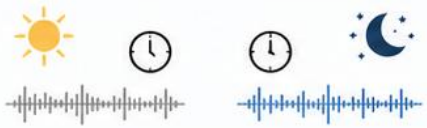
植被结构、地形和传播损耗差异会导致
声能衰减和散射不同。
同一生物群落，不同指数值。
限制跨站点和跨区域比较。

3 生态学意义丧失



指数无法区分物种身份、个体数量与丰度、
行为差异等。
大量信息在压缩计算中丢失。

4 对录音条件敏感



时间（昼夜、季节）、设备（麦克风）、
参数设置等会显著影响结果。
可重复性和稳健性较低。

② 基于 AI/统计推断的方法

如 AI 评分、深度学习、概率推断等

功能强大，但统计意义不确定，依赖数据，
容易产生偏差或不稳定的生态推断。

1 得分不等于概率



模型得分或置信度常被误解为出现概率，
但并非真实概率分布，容易导致
高估或低估。

2 跨域泛化能力差



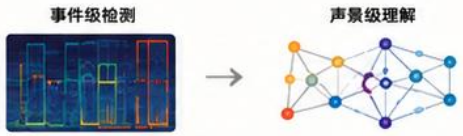
在新地点、新设备、新生境或噪声条件下，
模型性能下降明显。
领域迁移（Domain shift）严重
限制了方法的适用性。

3 数据稀缺与偏差



稀有物种样本不足，标签成本高且不一致，
导致模型学习到数据集偏差，
而非真实生态规律。

4 生态学可解释性有限



多数方法聚焦于“事件/物种”识别，
而非群落结构、生态位重叠和
时空动态等更高层级信息。

核心结论

指数方法：简单但盲目，信息严重压缩。
AI 方法：强大但不确定，偏差与泛化不足。

未来方向

需要融合多源数据、生态学知识与先进 AI 的综合方法，
实现更稳健、可解释、可推广的生物多样性声景评估。

04 未来展望

标准化声景监测体系

第一层：设备与采集标准化
(Standardized Equipment & Sampling)

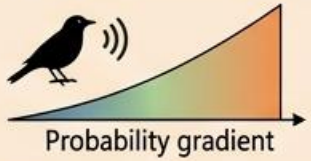
设备校准 (Calibration)
(Microphone Sensitivity, Frequency Response, SNR)

部署标准 (Deployment)
(Height: e.g., 1.5m, Orientation: e.g., North, Spacing: e.g., 500m)

采样标准 (Sampling)
(Sample Rate, Duration, Time Windows: e.g., Daily 05:00-08:00, 1 min recording per 10 min)

第二层：观测过程与模型校正
(Standardized Observation Process & Model Correction)

声学探测模型 (Acoustic Detection Model)



P(Detection)
估计 P(Detection)
物种存在 ≠ 一定被录到

声学传播模型 (Acoustic Propagation Model)

P(Recording | Distance, Habitat)



林分结构，温湿度，地形
对传播损失进行校正

发声行为模型 (Calling Behavior Model)



P(Vocalization)
估计 P(Vocalization)
繁殖状态，季节，时间

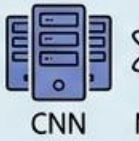
第三层：分析方法标准化
(Standardized Analysis Methods)

声学指数 (Acoustic Indices)



ACI 适用场景
例如：干扰梯度比较 (✓)
不适用于：物种丰富度估计 (✗)

AI 模型 & Benchmarks



跨区域Benchmark
例如建立：Global Bird Sound Benchmark
覆盖：热带，温带，湿地，草原

不确定性报告 (Uncertainty Reporting)



强制报告：
Accuracy = 95%
Detection uncertainty,
Model uncertainty,
Transfer uncertainty

第一阶段：声学指数时代
(过去 (Past)

第二阶段：AI识别时代
未来 (Future)

第三阶段：标准化声景监测体系时代
(标准化采集+探测概率校正+传播模型校正
+多模态融合+生态统计模型)

目标：真实生物多样性推断
(True Biodiversity Inference)
(而不是简单的声音分类)

第一阶段：声学指数时代
(ACI, ADI, BIO)

第二阶段：AI识别时代
(CNN, Transformer, Foundation Model)

声景生物多样性观测体系 (Soundscape Biodiversity Observatory)：
构建一套可重复、可比较、可校正、可长期连续运行的标准化生物性观测网络。

04 未来展望


未来智慧生态监测系统
多模态数据综合监测鸟类多样性
视觉 + 听觉 + 加速度计等多模态融合 | AI智能分析 | 全时空监测
视觉采集
高清相机 / AI识别
物种识别、数量统计
行为分析、栖息地利用
听觉采集
高灵敏度麦克风
鸣声识别、物种分类
活动时间分析
加速度计
数量化传感器
飞行行为识别
活动强度监测
能量消耗估计
视觉 (Vision)
物种识别
数量统计
听觉 (Audio)
鸣声识别
物种分类
加速度计 (Accelerometer)
飞行行为识别
活动强度监测
其他数据
环境数据
(温度 / 湿度 / 气压 / 降雨量)
GPS 轨迹
栖息地信息
(植被 / 水体 / 地形)
...
数据传输
4G/5G/LoRa/卫星
多模态数据融合
AI 智能分析引擎
物种多样性
128 种
物种分布
活动节律
24h
趋势分析
预警与异常
栖息地变化预警
物种减少预警
异常活动预警
鸟类多样性综合监测平台
128 种
28,356 条
56 个
365 天
科学研究
生态保护
生物多样性评估
保护决策支持
栖息地管理
公众科普教育
多模态感知 · 智能融合分析 · 全时空监测
助力鸟类保护与生态可持续发展

05 致谢



国家林业和草原局林草科技创新团队

生物多样性智慧监测研究中心

林木资源高效生产全国重点实验室

重点研发、国自然、北自然

谢

谢

北京林业大学
Beijing Forestry University

敬请批评指正

谢将剑 shyneforce@bjfu.edu.cn