



四川大學
SICHUAN UNIVERSITY

基于深度特征融合鸟鸣识别

潘帆

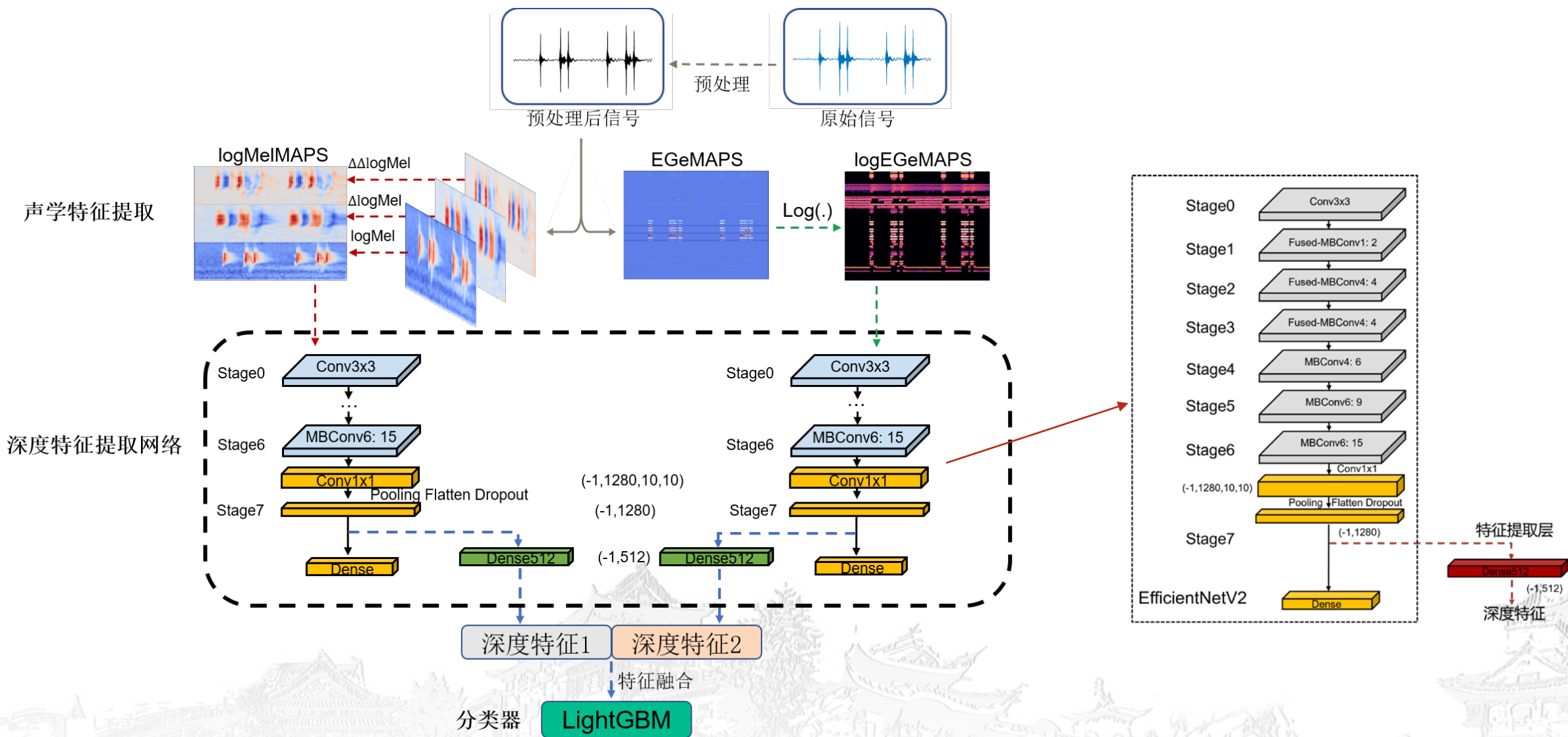
四川大学电子信息学院

- 鸟类监测发现更多的近危或濒危鸟类，科学掌握鸟类的习性、分布和迁徙规律，以及客观反映当地的生态环境状况。能为生态环境保护和管理提供科学依据。
- 相比视觉识别，通过音频识别进行鸟类监测，成本低、监测范围广、环境适应能力强

基于深度特征融合鸟鸣识别方法



四川大学
SICHUAN UNIVERSITY



● 蔡建民, 何培宇, 杨智鹏等. 基于深度特征融合的鸟鸣识别方法及其可解释性分析[J]. 生物多样性, 2023, 31 (07) :142-151.

实验数据集



四川大学
SICHUAN UNIVERSITY

数据集名称	物种数	样本总数	数据集特点
北京百鸟数据集	20	14311	中国 本土常见 鸟种，包含所有 五种鸟鸣频谱类型
CLO-43SD北美林莺	43	5428	北美林莺的迁徙鸣声，具有 持续时间短、频率高等特点
BirdCLEF2022挑战赛	152	130038	涵盖夏威夷群岛链的152个物种，具有 多物种，噪声复杂 的特点。

实验结果 - 北京百鸟数据集 (基准测试)



四川大学
SICHUAN UNIVERSITY

模型	平均准确率	平均F1分数
logMel+mobileNetV3	90.20%	90.11%
深度融合特征 (mobileNetV3) +lightGBM	97.77%	97.76%
logMel+EfficientNetV2	93.08%	93.20%
logMelMAPS+EfficientNetV2	95.69%	95.73%
EGeMAPS+EfficientNetV2	77.41%	77.45%
LogEGeMAPS+EfficientNetV2	91.56%	91.42%
深度logEGeMAPS+LightGBM	97.13%	97.12%
深度logMelMAPS+LightGBM	98.71%	98.69%
深度融合特征+SVM	98.83%	98.82%
深度融合特征+ Random Forest	98.82%	98.81%
深度融合特征+XGBoost	98.64%	98.63%
深度融合特征+lightGBM	98.70%	98.82%

1/2阶差分带
来提升

非线性处理带
来提升

实验结果 - 北京百鸟数据集 (20种)



四川大學
SICHUAN UNIVERSITY

模型	平均准确率	平均F1分数
GWO-KELM(李大鹏, 2022)	91.16%	88.54%
LogMel+CRNN(Adavanne et al, 2017)	92.89%	89.64%
LogMel+CNN(Bold et al, 2019)	91.12%	88.47%
logMel+DSRN+DilatedSAM+BiLSTM(李大鹏, 2022)	96.58%	96.51%
深度融合特征+lightGBM(本文)	98.70%	98.82%

实验结果 - CLO-43SD北美林莺(43种)



四川大學
SICHUAN UNIVERSITY

模型	平均准确率	平均F1分数
深度特征-1-2-3 + 最小最大归一化 +KNN(吉训生 et al, 2022)	93.89%	-
logMel+ SVM(Salamon et al, 2016)	93.96%	-
深度融合特征+lightGBM(本文)	98.32%	98.04%

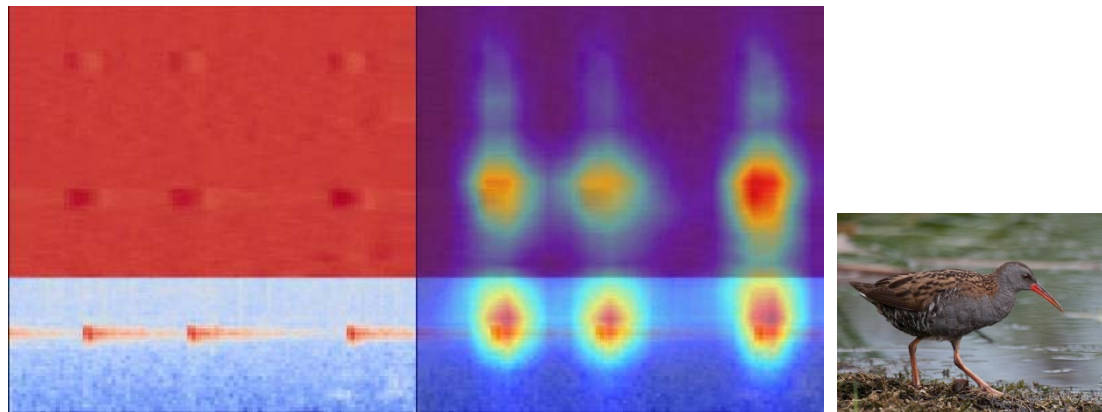
实验结果 - BirdCLEF2022挑战赛(152种)



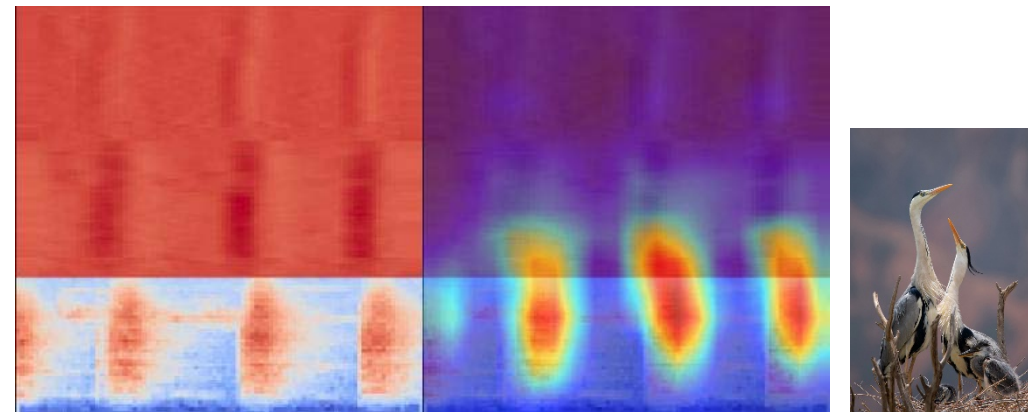
模型	平均准确率	平均F1分数
BirdCLEF 2022比赛public最优	-	91.28%
深度logMelMAPS+ lightGBM (本文)	88.53%	88.32%
深度融合特征+SVM	89.40%	89.22%
深度融合特征+ Random Forest	85.47%	84.77%
深度融合特征+XGBoost	85.71%	85.47%
深度融合特征+lightGBM(本文)	91.12%	91.05%

- 引入了新特征集eGeMAPS并对其做了对数运算，是对logMelMAPS的补充，融合特征在复杂数据集(BirdCLEF2022)中在平均准确率上取得了**2.59%的提升**

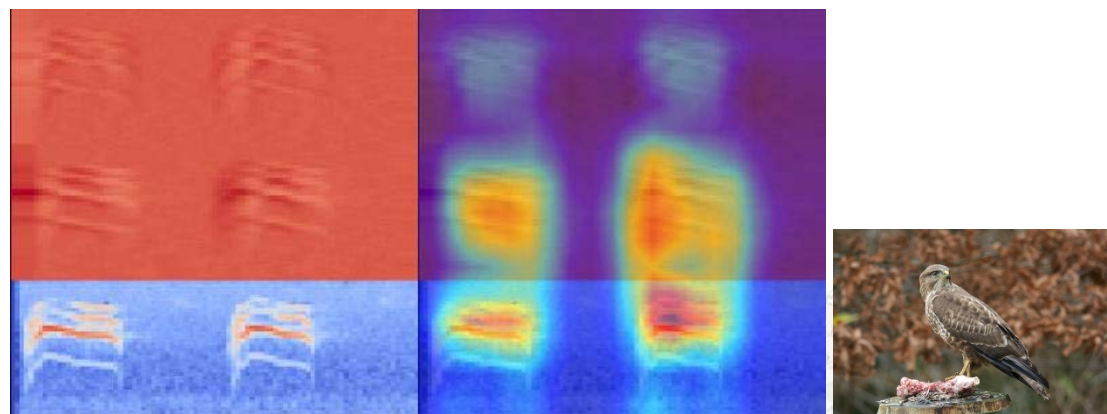
可解释性分析：热力图



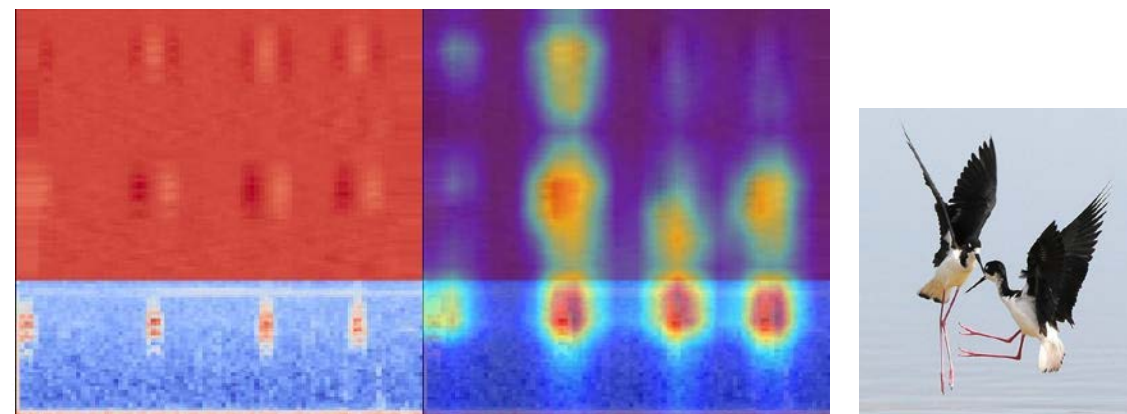
西方秧鸡，单频类型



苍鹭，类噪声类型



欧亚鳶，强谐波类型



黑翅长脚鹬，宽带脉冲类型



左边为logMelmaps，右边为对应的热力图

- 1. 多特征融合。**引入了新特征集egemaps并对其做了对数运算，是对logmelmaps的补充，融合特征在复杂数据集(BirdCLEF2022)中在平均准确率上取得了**2.59%的提升**。深度特征与lightGBM的模型组合相比直接使用神经网络分类准确率上有**3%以上的提升**。
- 2. 可解释性分析。**通过绘制热力图，发现不同类型的鸟鸣网络的关注区域不同。

后续工作



四川大学
SICHUAN UNIVERSITY

- 算法实时性能优化与低计算资源平台部署
- 实现设备端的自动增益控制与鸟鸣事件检测
- 实现设备端+云端的鸟鸣识别



自研生态声学录制/计算设备

谢谢



四川大學
SICHUAN UNIVERSITY