

从时域到频域：基于多分支 CNN 网络的 AI 音频检测模型

NKUMMF2025138

October 6, 2025

摘要

针对 AI 音频的识别问题，本文提出一种基于**多分支卷积神经网络** (Multi-Branch CNN) 的端到端 AI 音频检测与评分模型。

从**时域、频域及声学常见统计量**等多角度提取 11 类特征，分别经五个并行分支建模后融合判别。针对单分支贡献的量化问题，引入**分支探针**机制，基于各分支独立判别准确率确定加权系数，构建可解释的 AI 痕迹综合评分体系。

本文设计多种扰动与对抗性处理（如频谱均衡、高频注入、环境噪声混入等）评估模型鲁棒性，并结合分支贡献分析揭示其在时域包络与共振峰布局上的依赖性。实验结果表明，该模型在验证集上准确率可达 89%-90%，在多数轻中度扰动下保持稳定性能，综合评分在强扰

动下亦具较高稳健性。本文方法具有较低计算开销与良好可扩展性，可推广至语音伪造检测、环境音识别等领域。

关键词： AI 音乐检测；多分支 CNN；音频特征提取；探针机制；AI 痕迹评分