



图12. AudioSet标签系统的分类性能。红色、蓝色和黑色曲线是CNN14、MobileNetV1和音频标记系统的AP[20]。蓝色条以对数刻度显示训练片段的数据。

[14] K. Choi, G. Fazekas, M. Sandler和K. Cho, 迁移学习

“音乐分类和回归任务”, 在会议

国际音乐信息检索学会 (ISMIR), 2017, pp. 1411-149.

[15] J. Pons, O. Nieto, M. Prockup, E. Schmidt, A. Ehmann和X. Serra

在Conference on大规模进行音乐音频标签的端到端学习

国际音乐信息检索学会 (ISMIR), 2017年, 第637644页。

[16] 孔、徐、王和普拉布利, “音频设备分类”

《注意力模型: 概率视角》, 发表在IEEE International上声学、语音和信号处理会议 (ICASSP), 2018年,

第316320页。

[17] C. Yu, K. S. Borsim, 孔和徐, 多层次关注

“在检测和声学场景和事件分类 (DCASE), 2018, 第188页 192.

[18] 张、杨, 学会识别

使用注意力监督的瞬态声音事件, ”International, 人工智能联合会议 (IJCAI), 2018, 第3336页 3342.

[19] 王、李、梅译: 五个多实例的比较

学习用于弱标签声音事件检测的池函数, ”