

实验结果分析

下面的混淆矩阵显示了对从 **AudioSet** 语料库中提取的子集的评估结果。该方法对于女性而言表现出色，准确率为 **95%**，而对于男性，准确率为 **76%**。系统的整体精度为 **86%**。

如果我们看一下性能，与训练有素的男性模特相比，训练有素的女性模特似乎可以很好地代表其性别。下一节讨论了可能的原因和需要改进的地方。我们还在展览数据集上评估了

	男	女
男	417	129
女	29	529

同样训练有素的性别模型，该展览数据集由在开放技术展览期间实时环境中收集的 **250** 个音频文件（**130** 个男性和 **120** 个女性）组成。在以下混淆矩阵中显示了对展览数据集的评估结果。

	男	女
男	115	15
女	1 个	119

女性性别模型的准确性为 **99%**，而男性性别模型的准确性为 **88%**。系统的整体精度为 **94%**。