Neural Synthesis of Footsteps Sound Effects with Generative Adversarial Networks

M. Comunità, H. Phan, J.D. Reiss

Queen Mary University of London

Abstract

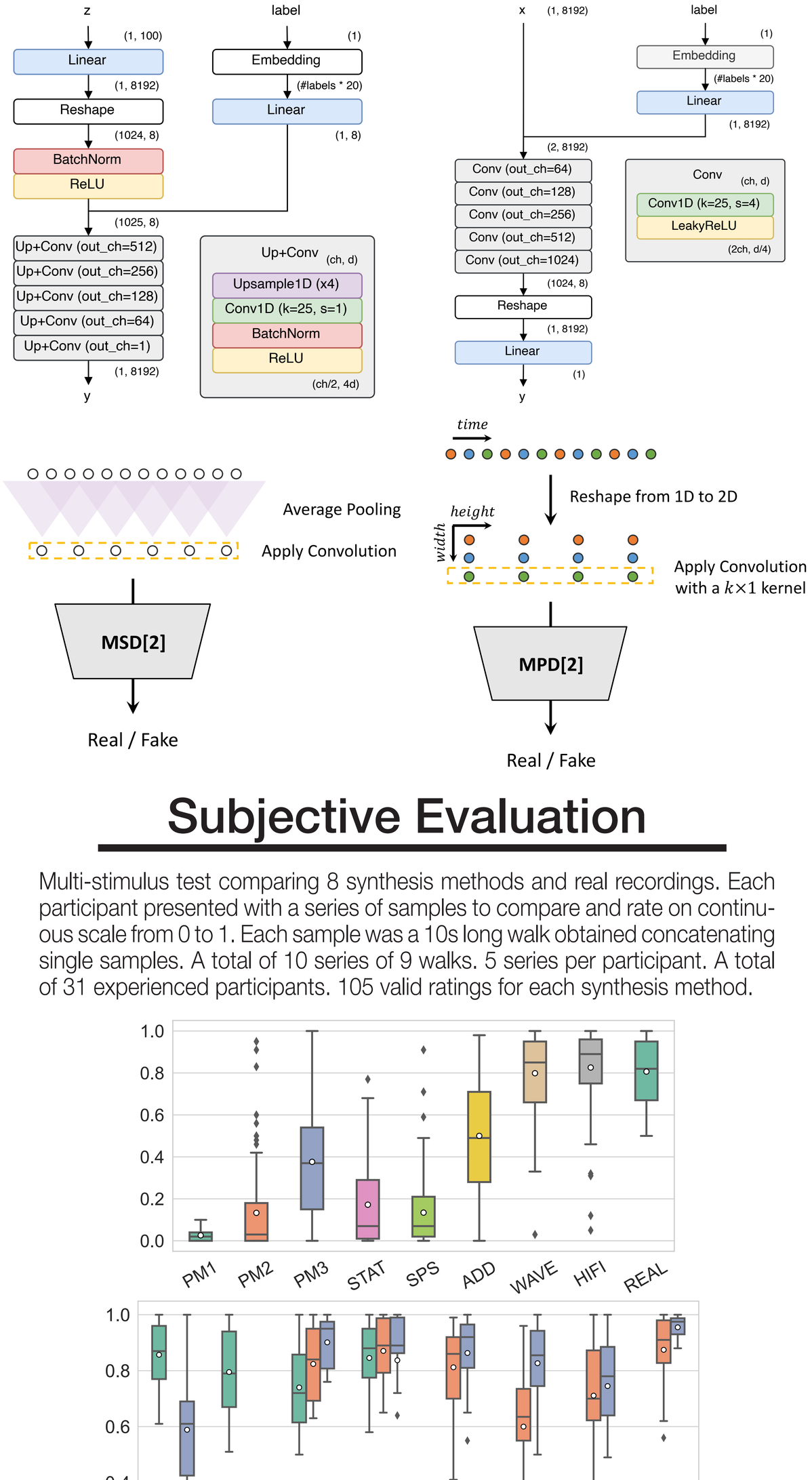
First attempt at neural synthesis of footsteps sound effects

Hybrid architecture based on WaveGAN and HiFi-GAN

Objective evaluation of quality and diversity of synthesised samples

Subjective evaluation comparing traditional and neural synthesis methods

Architecture

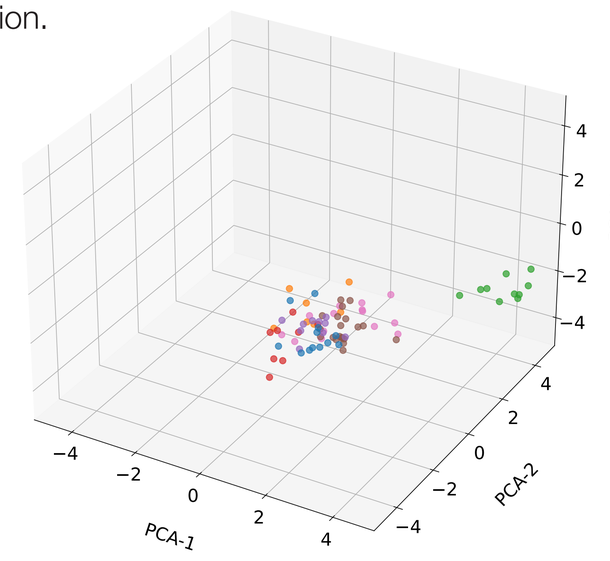


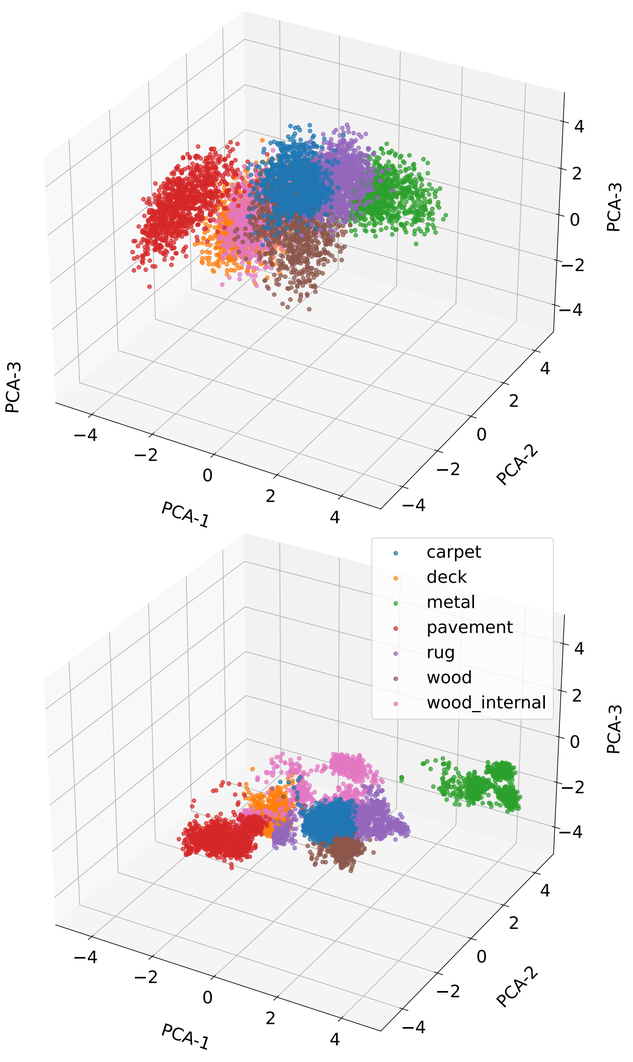
Subjective Evaluation

Multi-stimulus test comparing 8 synthesis methods and real recordings. Each participant presented with a series of samples to compare and rate on continuous scale from 0 to 1. Each sample was a 10s long walk obtained concatenating single samples. A total of 10 series of 9 walks. 5 series per participant. A total of 31 experienced participants. 105 valid ratings for each synthesis method.

Objective Evaluation

Principal Components Analysis | PCA on OpenL3 embeddings shows whether there is overfitting or mode collapse, and if our models learned the real data distribution.



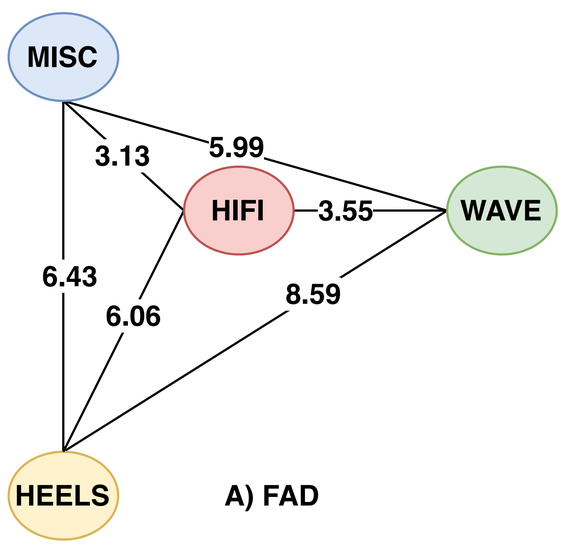


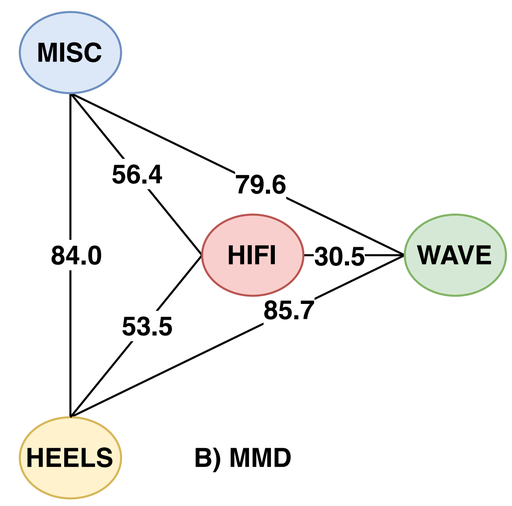
Dataset	IS
Zapsplat Misc	4.139
Zapsplat Heels	3.221
HiFi-WaveGAN	3.411
WaveGAN	3.232

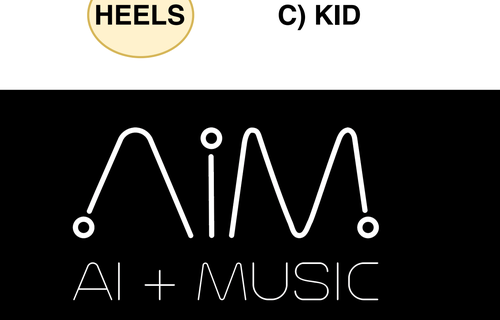
Inception Score | Measures diversity and semantic discriminability. It ranges from 1 to n (with n number of classes) and is maximised for models which can generate samples for all possible classes and that are classified with high confidence

Frechet Audio Distance | A multivariate Gaussian is fitted to the VGG-ish embeddings of real and synthesised data. FAD measures the distance between the two distributions. FAD is robust against noise and consistent with human judgements.

Maximum Mean Discrepancy | MMD between OpenL3 embeddings for real and synthesised samples as a measure of similarity between datasets.







Kernel Inception Distance | Measure of similarity between real and synthesised samples. Based on the squared MMD of embeddings from a pre-trained Inception model.

Paper
Code
Samples

centre for digital music

AI + MUSIC

UK Research and Innovation