MONYC: Music of New York City Dataset

Magdalena Fuentes[1], Danielle Zhao[1], Vincent Lostanlen[2], Mark Cartwright[3], Charlie Mydlarz[1], Juan Pablo Bello[1]

[1]New York University, USA
[2]CNRS, Laboratoire des Sciences du Numérique de Nantes (LS2N), France
[3]Department of Informatics, New Jersey Institute of Technology






https://magdalenafuentes.github.io/
mf3734@nyu.edu

Motivation

- Music is important in human cultures and is an integral part of urban soundscapes
- To make sense of these soundscapes, machine listening models should be able to detect and classify street music
- The **lack of well-curated resources** for training and evaluating models currently hinders their development




Pictures from https://pixabay.com/

MONYC

Gtzan-Genre

hip hop

jazz

pop

MONYC

hip hop

jazz

pop

MONYC recordings have **low Signal-to-Noise Ratio** (SNR), are picked up by the sensors at **far-field distances** (ranging from approximately ten to fifty feet), and have **variant levels of noise from other sources** present in the streets such as cars or people talking.

Data curation

Data curation

Audio acquisition and selection → Urban Sound tagging → Spatiotemporal sampling → Diverse sampling → Annotation

Data curation

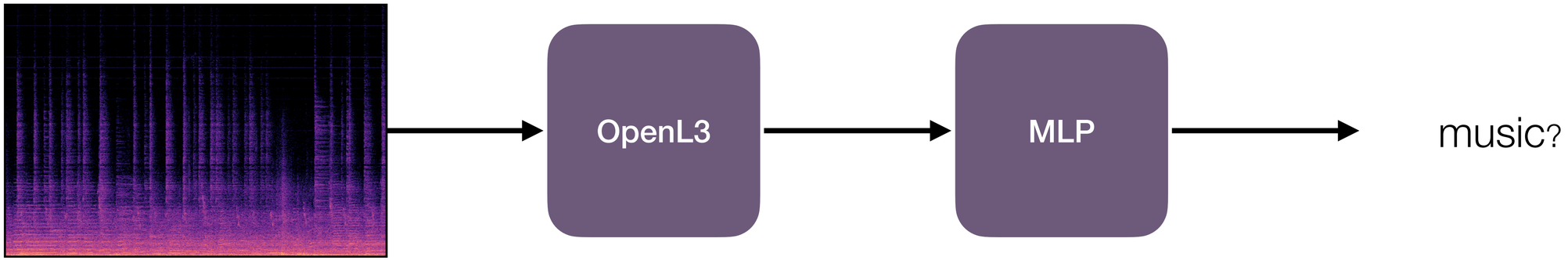
Audio acquisition and selection → Urban Sound tagging → Spatiotemporal sampling → Diverse sampling → Annotation

Data curation

Audio acquisition and selection → Urban Sound tagging → Spatiotemporal sampling → Diverse sampling → Annotation

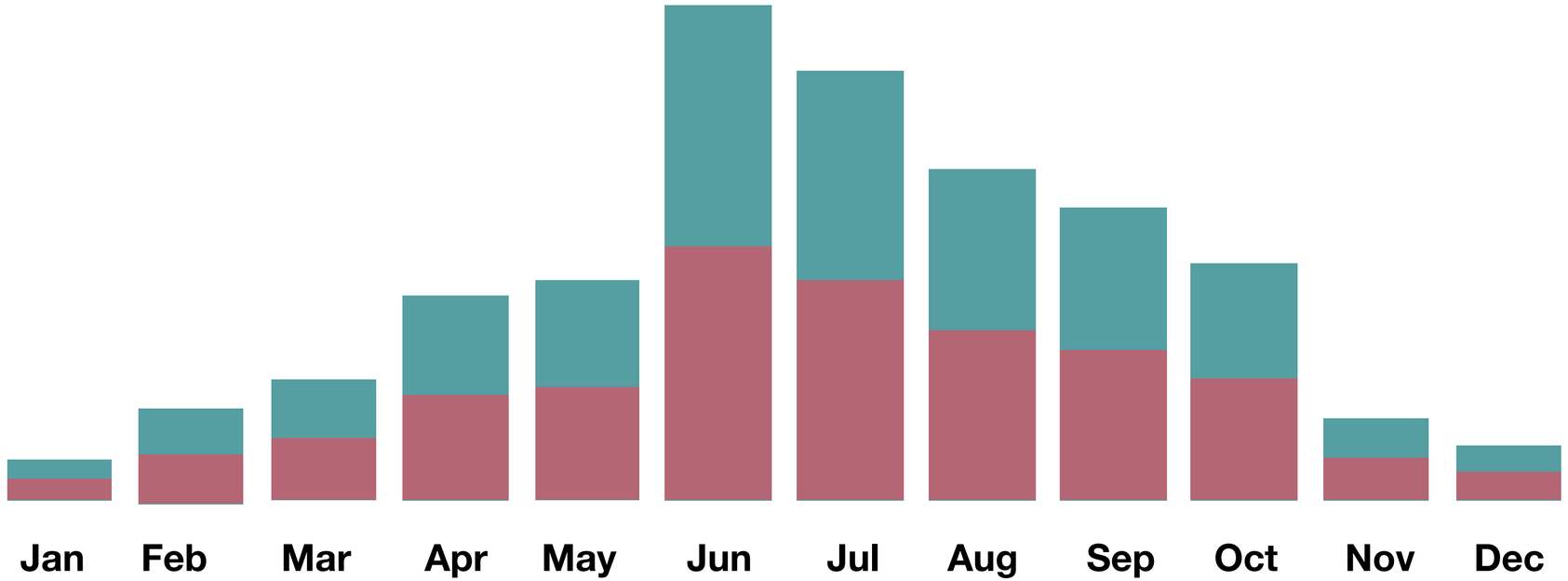
5.8M

Data curation

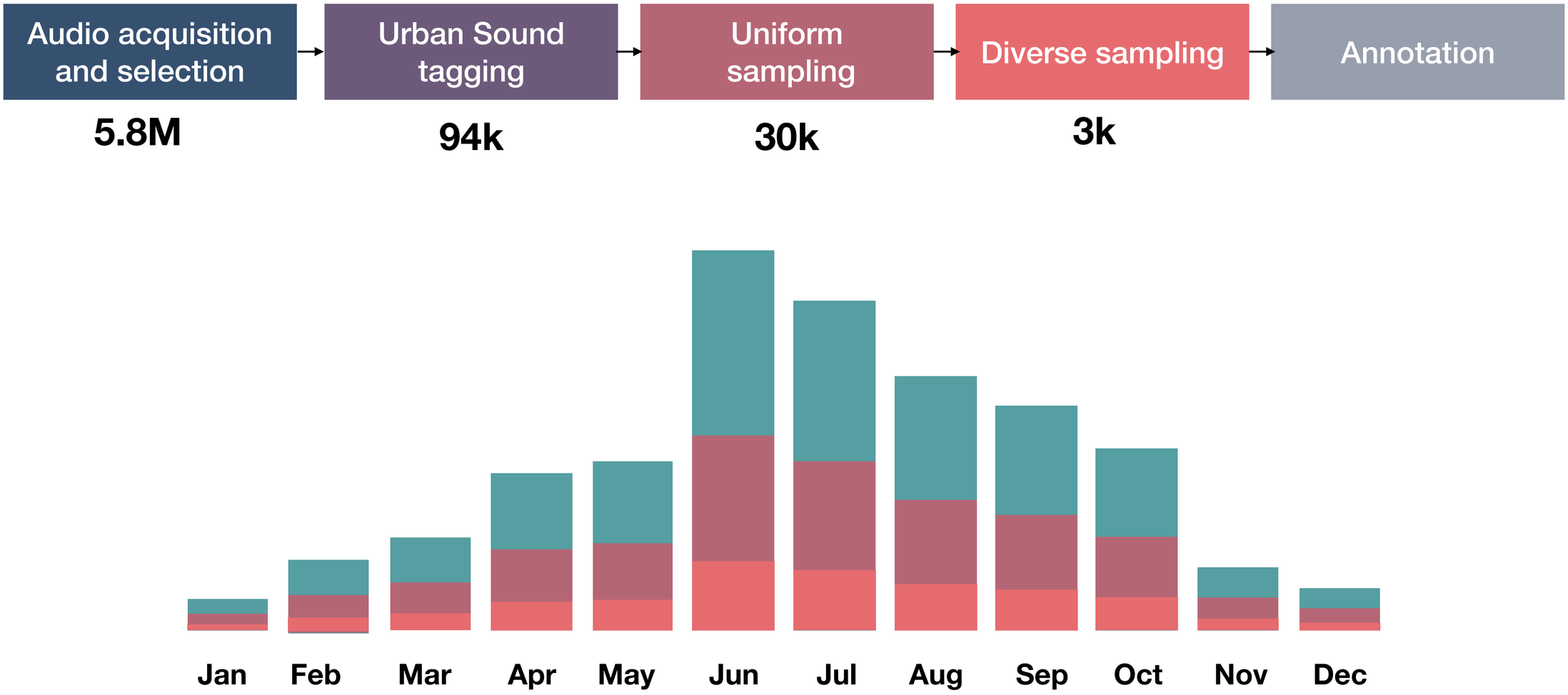


https://github.com/sonyc-project/dcase2020task5-uststc-baseline

Data curation



Data curation



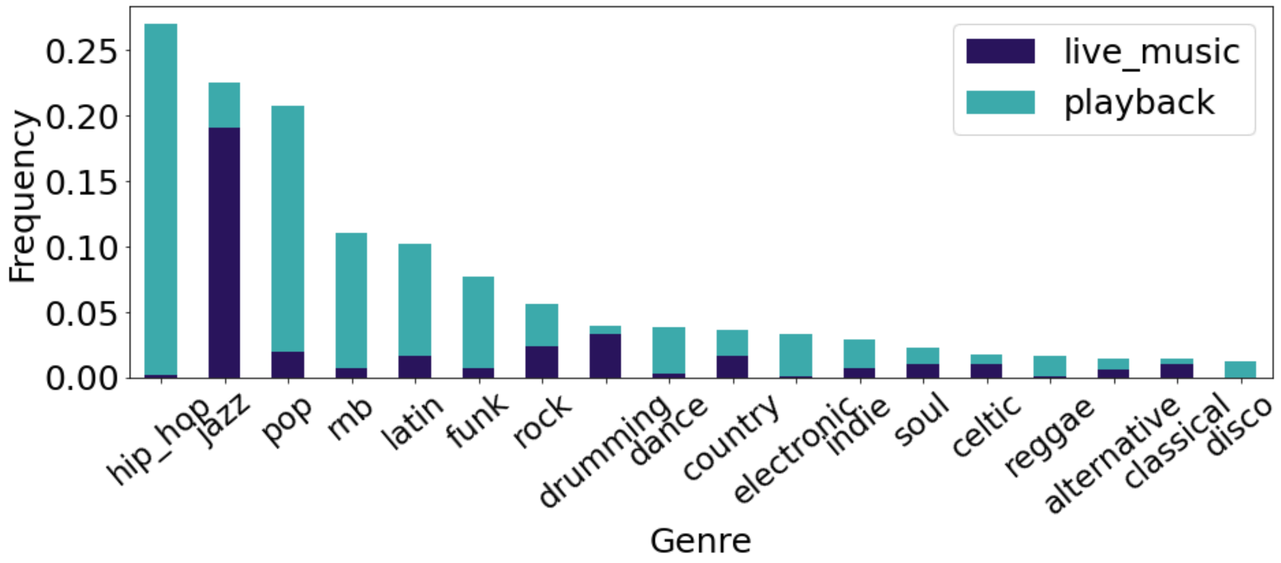
https://github.com/guilgautier/DPPy

Data curation

1. Pre-selection of musical recordings (1.5k)
2. Annotation of **genre**, **live vs. playback music**, **multi- vs. single- instrument**, **loud vs. quiet**
3. Conflict solving

Dataset overview

Music genres and live music in MONYC

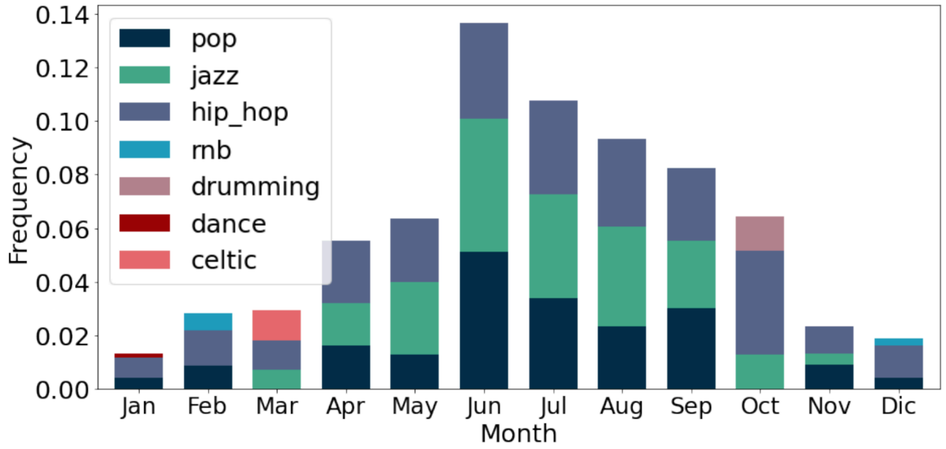


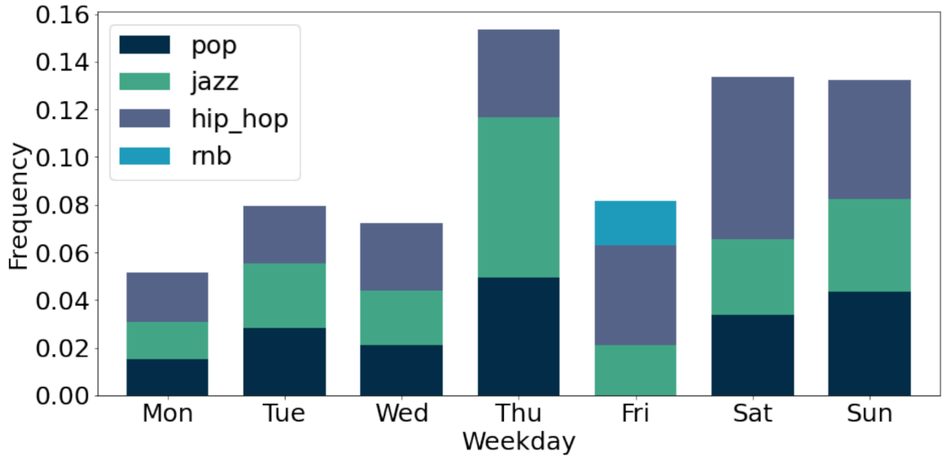
- Top genre in MONYC is **hip hop**, unlike datasets such as AcousticBrainz [1], where the genres with more appearances are rock or pop
- **Most genres are play-backed**, sometimes from cars passing, sometimes from shops, or speakers outside homes. The exception are two genres: jazz and drumming, which are both mostly live

[1] D. Bogdanov, A. Porter, H. Schreiber, J. Urbano, and S. Oramas, “The AcousticBrainz Genre Dataset: Multi-source, Multilevel, Multi-label, and Large-scale,” in Proceedings of the 20th Conference of the International Society for Music Information Retrieval (ISMIR 2019): 2019 Nov 4-8; Delft, The Netherlands.[Canada]: ISMIR; 2019. International Society for Music Information Retrieval (ISMIR), 2019.

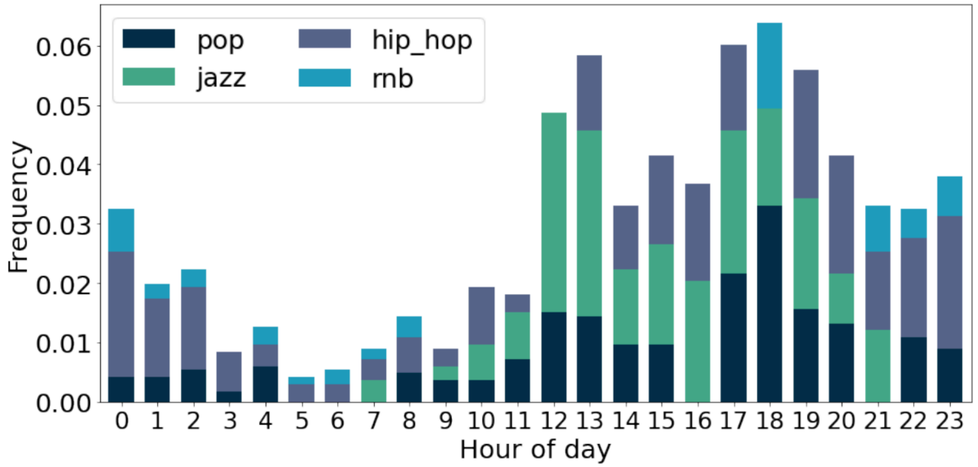
Dataset overview

Spatiotemporal information in MONYC



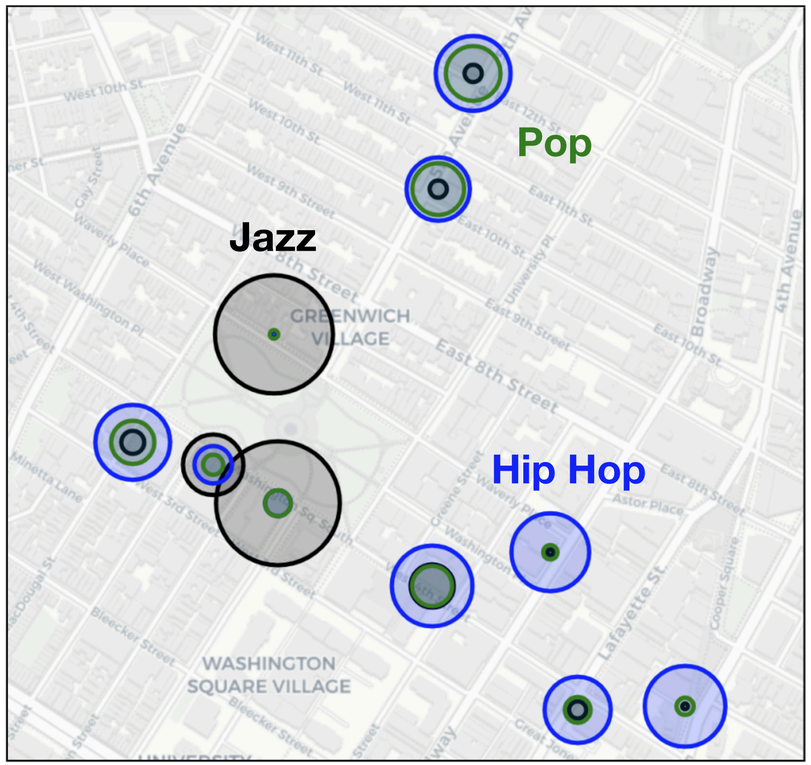


- More music clips towards the Summer months (June, July), considerably less in Winter (November, December and January)
- Genre appearance usually correlated with events, e.g. Celtic music in March due to St. Patrick's day
- Less street music at the beginning of the week, more on the weekend



Dataset overview

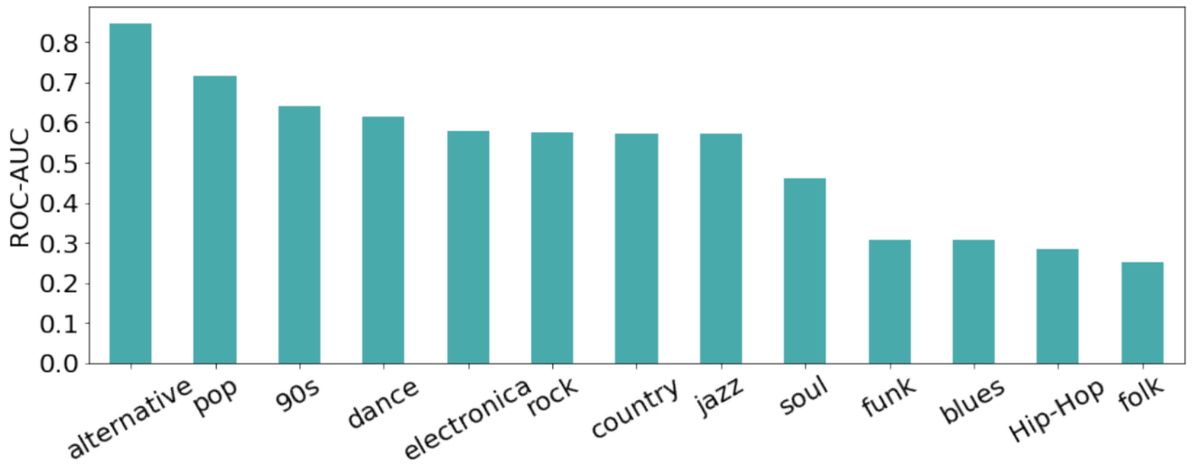
Spatiotemporal information in MONYC



- Spatial distribution of genres, Jazz is mostly live, so is concentrated around Washington Square Park

Dataset overview

Music tagging: a small experiment



- We use an off-the-shelf music genre tagger musicnn [2]
- Overall performance in MONYC is lower than in other datasets [3] (in the range of 90%), with a median ROC-AUC score of 50%
- The model performed 8-12% worse in average in those recordings labeled with high interference of sources

[2] J. Pons and X. Serra, “musicnn: Pre-trained Convolutional Neural Networks for Music Audio Tagging,” arXiv preprint arXiv:1909.06654, 2019.

[3] M. Won, A. Ferraro, D. Bogdanov, and X. Serra, “Evaluation of CNN-based Automatic Music Tagging Models,” arXiv preprint arXiv:2006.00751, 2020

Conclusions and future perspectives

- We present MONYC, the **first-of-its-kind open dataset of music in urban settings**
- MONYC opens the possibility to develop and evaluate machine listening models for the classification of street music
- Such models for the classification of street music offer the opportunity to dig into behavioral patterns related to human activities in urban, such as nightlife, festivals, street celebrations, among others.

- https://magdalenafuentes.github.io/monyc/
- https://github.com/soundata/soundata

Thanks!