

A Dataset for Greek Traditional and Folk Music: Lyra



Charilaos Papaioannou¹, Ioannis Valiantzas², Theodoros Giannakopoulos³, Maximos Kaliakatsos-Papakostas⁴, Alexandros Potamianos¹

¹ School of ECE, National Technical University of Athens, Greece; ² Department of Music Studies, National and Kapodistrian University Of Athens, Greece; ³ National Center for Scientific Research - Demokritos, Greece; ⁴ Athena RC, Greece

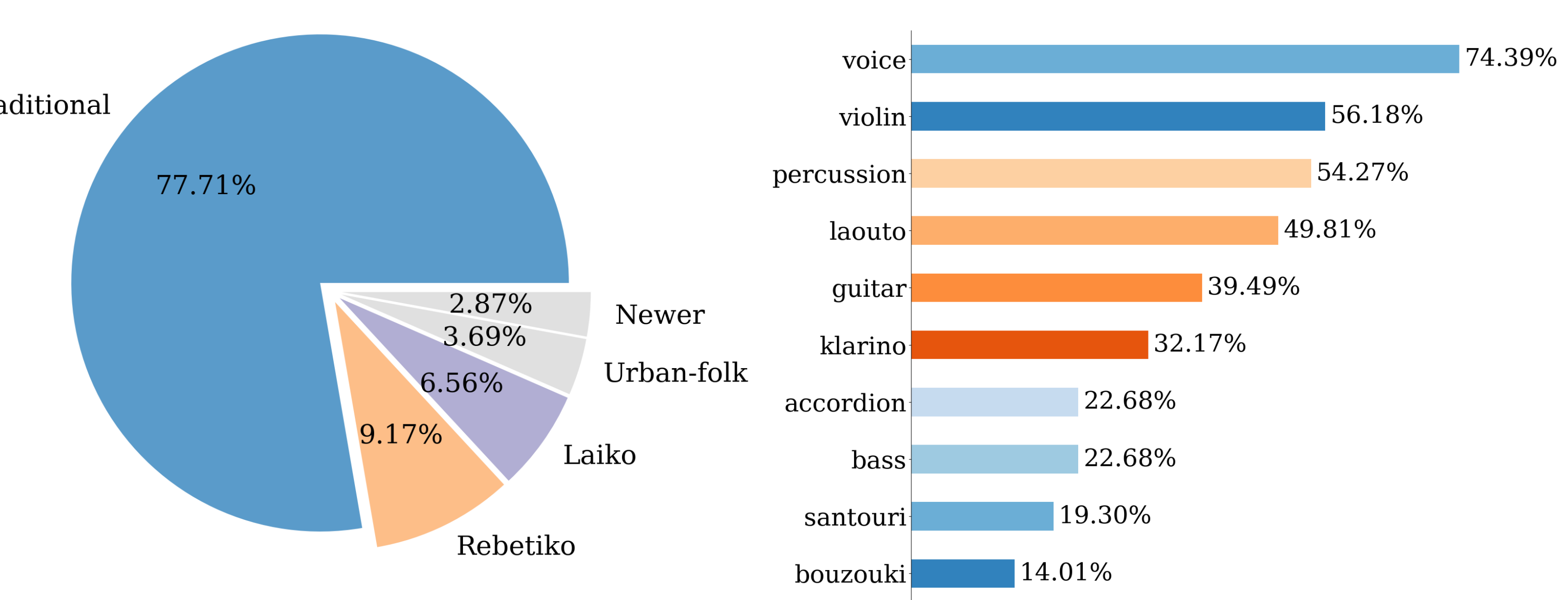
Motivation

- Artificial Neural Networks need lots of data coming from balanced datasets to develop their full potential.
- How can ANNs cope with complex datasets constituting of several imbalanced classes having hierarchical relationships between them?
- How can they help us reveal correlations between musical characteristics such as place of origin and instruments?
- How can they identify similarities/influences between music pieces with different places of origin, instrumentations, genres?

Lyra Dataset

- 80 hours of audio and video, 1570 pieces
- genre: 32 unique values hierarchically related; 5 genres and 27 sub-genres
- place of origin: 81 unique values, many of them hierarchically related; 20 regions and 61 places
- instruments: 32 unique values, 296 unique groups of instruments
- Other metadata:
 - is-danced: whether a piece is being danced or not during the show
 - coordinates: longitude and latitude of a place of origin

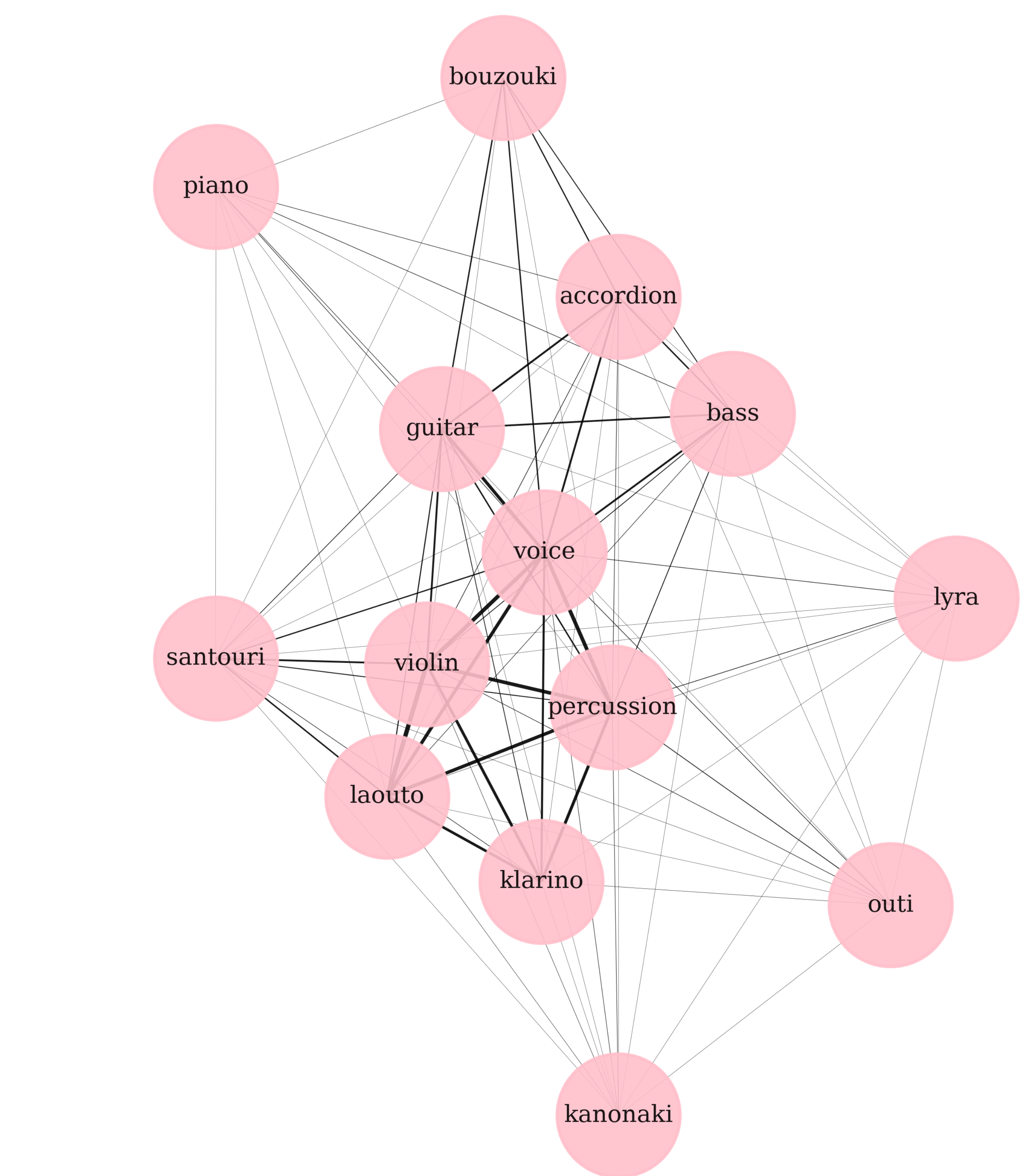
Relative frequencies of the music genres and the most common instruments in the dataset



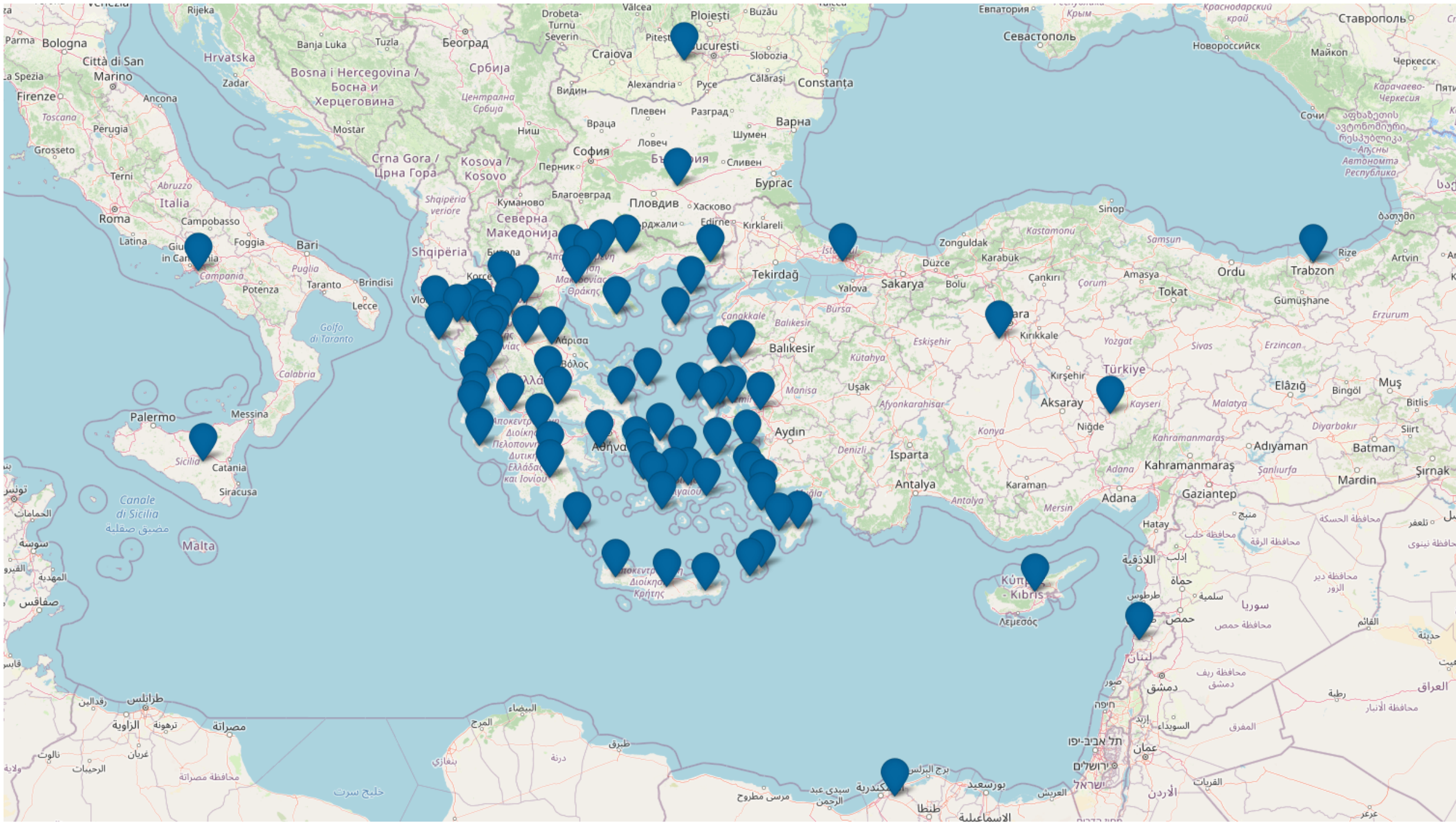
Sample rows of the dataset. Pipe “|” is used to separate values in a field

id	instruments	genres	place	coordinates	is-danced	youtube-id	start-ts	end-ts
alexandra	voice violin percussion laouto klarino	Traditional Epirotic	Epirus Zagori	39.8648 20.9284	0	qrOwc1mLFUk	749	927
choros-tik	percussion lyra	Traditional Pontian	Pontus	40.9883 39.7270	1	Aws0Y3aLaIs	1731	1886
agiothodoritissa	voice violin santouri percussion laouto guitar	Rebetiko	None	None	1	0cj8BNcAhg4	2632	2853
einai-argapoly-arga	voice piano guitar bass bouzouki accordion	Laiko	None	None	0	zkoqg3VRVLA	2365	2614

Graph with instruments co-occurrences



Map of all the places of the dataset



Experiments

- mel-scaled spectrograms per fix-sized segment duration of 10 seconds
- CNN with 4 convolutional followed by 3 fully connected layers to classify each 10-second mel-spectrogram to the tasks: genre, place and instruments
- single-label setup for “genre” and “place” (multiclass) tasks
- multiple binary models for the instrument (multi-label) task
- test set: separation of 20% of the data on an episode level to avoid spreading pieces from the same broadcast across data splits.

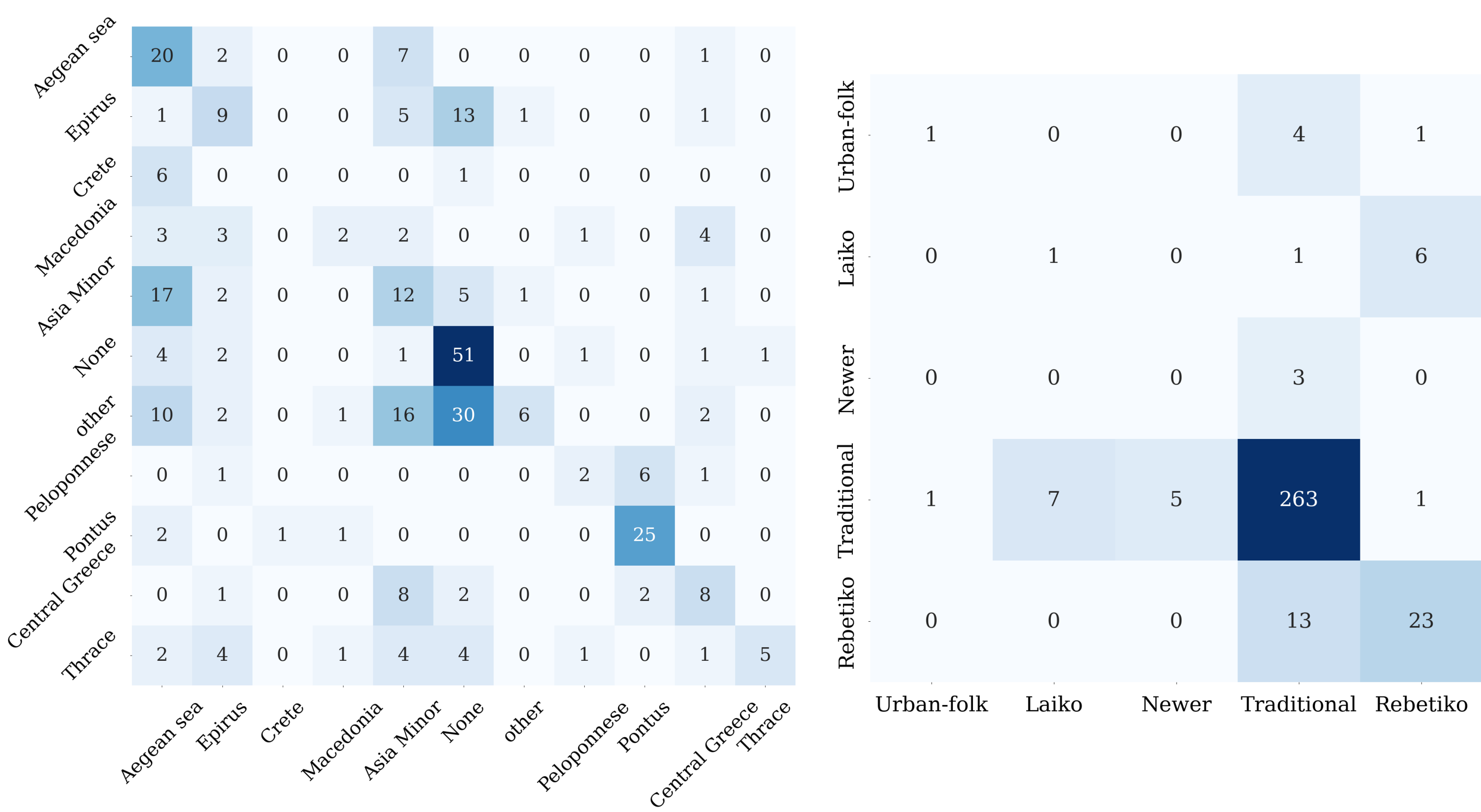
Results

AUC and F1 scores of the models on the test data

Instrument	AUC (%)
voice	68.9
violin	85.2
percussion	95.1
laouto	93.8
guitar	73.5
klarino	90.9

Task	Macro-F1 (%)	Micro-F1 (%)
Genre	39.9	87.2
Place	34.4	42.4

Confusion matrices for “place” and “genre” classes



Discussion

- “instrument” classifiers: the “10-second segments” setup (detect an instrument in each of the segments regardless if it is present on it or not) pushes the models towards a space with latent musical features, where “voice” may not be as discriminative as the rest of the instruments are
- “place” and “genre” classifiers: misclassifications can be related to imbalance between classes as well as statistical correlations across them

Conclusions

Lyra dataset:

- contains rich metadata, hierarchically related in many occasions, that constitute a significant challenge for the current deep learning models to cope with
- includes material that allows MIR tools to be employed for reaching valuable musicological results
- provides the potentiality to combine video and audio signals to experiment with multimodal learning methods