1st Latin American Music Information Retrieval (LAMIR) workshop
December 9-11, 2024 – Rio de Janeiro

MUSIC SOURCE SEPARATION IN NOISY BRAZILIAN CHORO RECORDINGS

Pedro Donadio de T. Júnior[1,2], Martín Rocamora[3,4], Luiz Wagner P. Biscainho[2]

[1]Department of Electronics and Computation, Universidade Federal do Amazonas, Brazil
[2]Universidade Federal do Rio de Janeiro, Brazil
[3]Music Technology Group, Universitat Pompeu Fabra, Spain
[4]Facultad de Ingeniería, Universidad de la República, Uruguay

1. MOTIVATION

- Choro is considered the first musical style to originate in Brazil, dating back to the 1870s;
- Some historical recordings from the early 20th century include noise inherent to the process of recording and playing shellac records;
- This work investigate the instrument separation task applied to historical recordings of this Brazilian music genre, using models originally trained on clean tracks.

2. BACKGROUND

- In recent years, deep learning approaches have achieved state-of-the-art performance in music source separation;
- We chose to base the investigation of this work on the Hybrid Demucs model;
- The model has been trained (from scratch) on a choro music data set containing 10 albums for training and 2 albums for validation.

Number	Songbook	# Songs
1	Altamiro Carrilho	13
2	Benedicto Lacerda	12
3	Chiquinha Gonzaga	12
4	Choro Meets Bach	14
5	Ernesto Nazaré 1	11
6	Ernesto Nazaré 2	11
7	Ernesto Nazaré 3	11
8	Inéditos	14
9	Jacob do Bandolim 1	12
10	Jacob do Bandolim 2	12
11	Pixinguinha	12
12	Roda de Choro	12
13	Severino Araújo	12
14	Waldir Azevedo 1	12
15	Waldir Azevedo 2	12
16	Zequinha da Abreu	12

3.EXPERIMENTS

- The experiments were conducted to test the separation models on simulated historical choro recordings. To do that, 20 tracks were carefully selected from the test set of the choro dataset to ensure they were free of leakage;
- To simulate historical 78 RPM noisy recordings, we combine the clean tracks of choro test set and the 5 pre conditioned noise tracks according to:

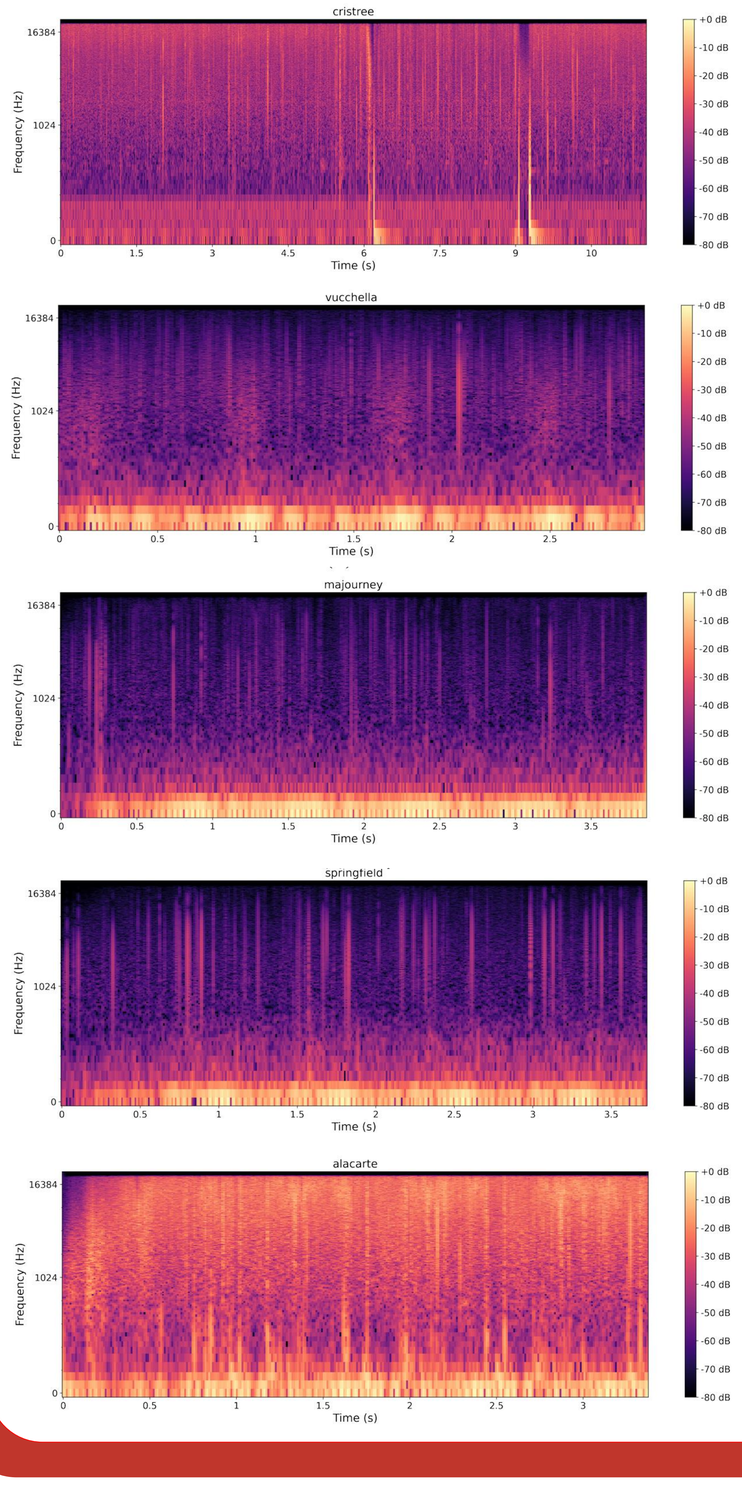
$$C_{\text{noisy}} = \beta.(G.N_{pp} + C_{\text{clean}})$$

where β represents the global track gain, G is the gain applied to the pre-processed noise track N_{pp} to produce the desired SNR, and C_{clean} is the track extracted from choro test set.

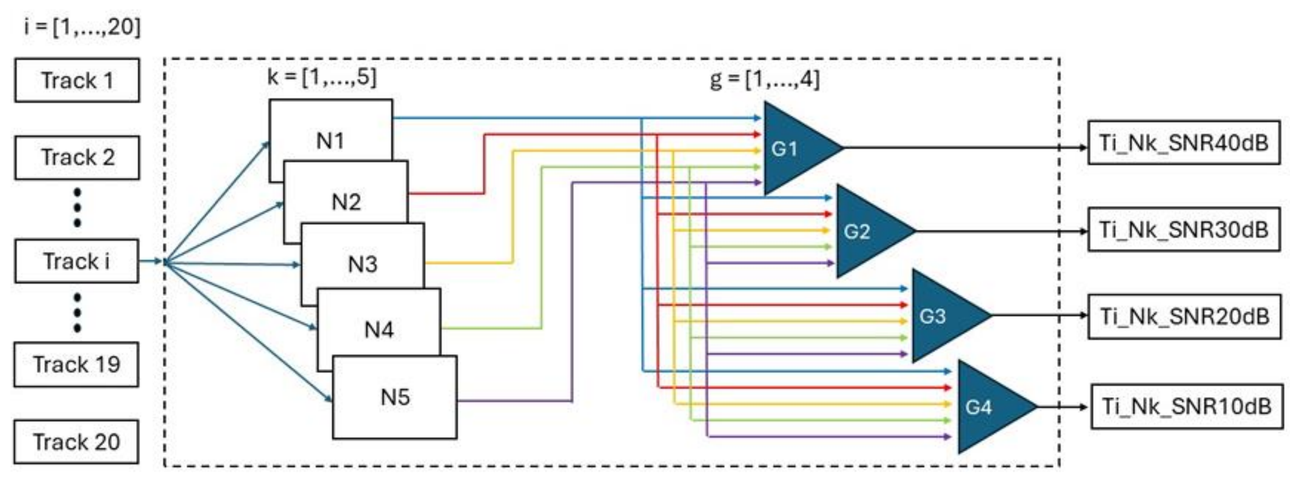
- Noise gains were adjusted to induce SNR values of 10, 20, 30, and 40 dB for each track. In total, 400 noisy mixtures were generated according to this scheme (Figure Section 4.)
- To assess the quality of separation, three objective metrics are widely adopted in the literature: signal-to distortion ratio (SDR), signal-to-interference ratio (SIR) and signal-to-artifact ratio (SAR) (Tables Section 4).

4. RESULTS

Noise spectrograms



Preparing the noisy tracks of the test set



Noise SNR	STRINGS			WIND			PERCUSSION		
	SDR	SIR	SAR	SDR	SIR	SAR	SDR	SIR	SAR
20 dB	14.210	28.248	16.104	14.134	24.596	15.590	5.657	16.918	5.234
30 dB	15.777	28.298	17.878	15.365	25.130	17.112	7.176	19.231	7.380
40 dB	16.128	28.230	18.503	15.582	25.280	17.414	7.453	19.452	8.233
Clean	16.130	28.373	18.399	15.700	25.407	17.576	7.6172	19.340	8.334

Average value of SDR, SIR and SAR for each SNR (considering the 20 noisy tracks in the test set)

Noise category	STRINGS			WIND			PERCUSSION		
	SDR	SIR	SAR	SDR	SIR	SAR	SDR	SIR	SAR
alacarte	12.749	27.398	14.309	14.169	24.699	15.617	-0.016	6.042	-6.685
cristree	11.880	27.291	14.128	14.035	25.083	16.023	0.597	7.475	-6.396
majourney	6.264	27.656	9.166	9.545	21.626	10.154	6.373	17.835	6.850
springfield	5.663	27.465	8.090	9.935	21.005	10.774	6.077	17.126	6.395
vucchella	5.102	26.898	7.091	10.484	20.101	11.447	5.631	16.241	6.122

Average value of SDR, SIR and SAR computed for the 20 noisy tracks in test set with an SNR of 10 dB

Conclusions

- The results demonstrate that the system is robust when dealing with tracks containing additive noise, even though it has been pre trained on clean recordings. Some families, such as percussion, face greater challenges in separation at lower SNRs, while others, such as wind instruments, show good results across all SNR levels. Some separation results, as well as the noises used, are available for listening at https://www02.smt.ufrj.br/~pedro.donadio/index_Lamir.html.

Listen to the noises and results here