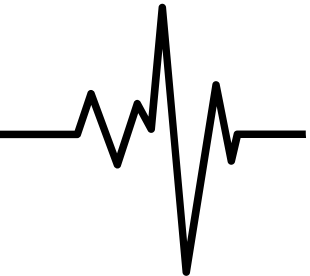
UNESP BCC | TCC 2019

RECONHECEDOR E SEPARADOR DE INSTRUMENTOS MUSICAIS

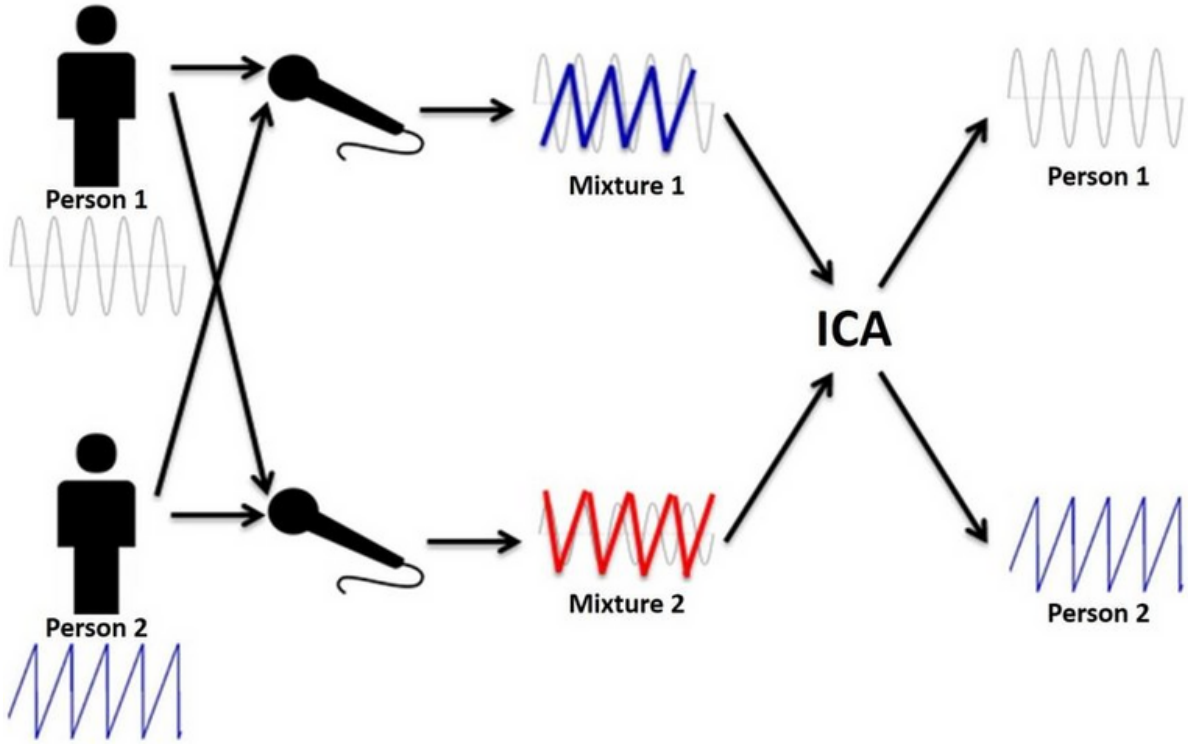
ORGANIZAÇÃO

- Introdução
- Fundamentação teórica
- Bases de dados
- Desenvolvimento das soluções
- Resultados



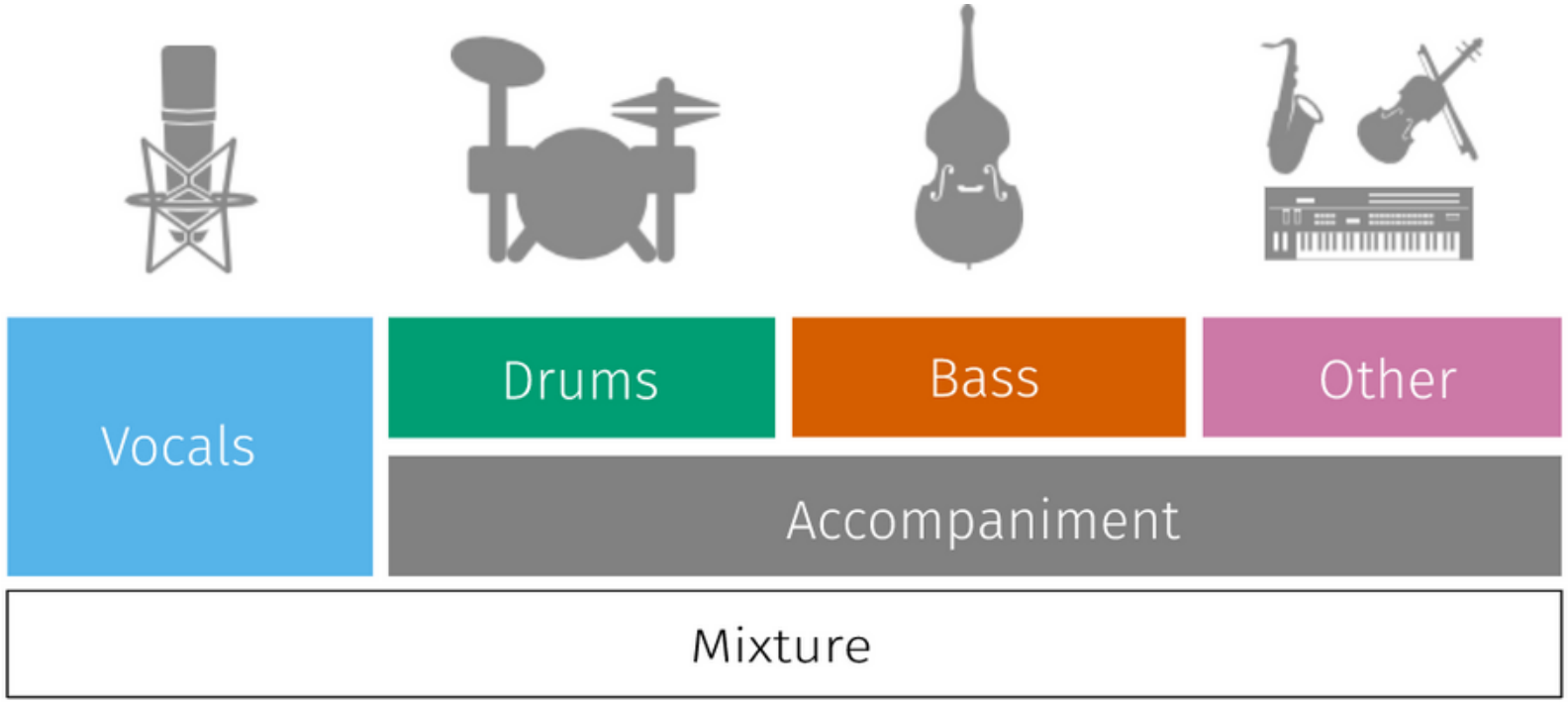
INTRODUÇÃO

The cocktail party



INTRODUÇÃO

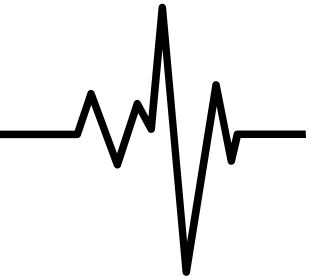
Separação cega de fontes



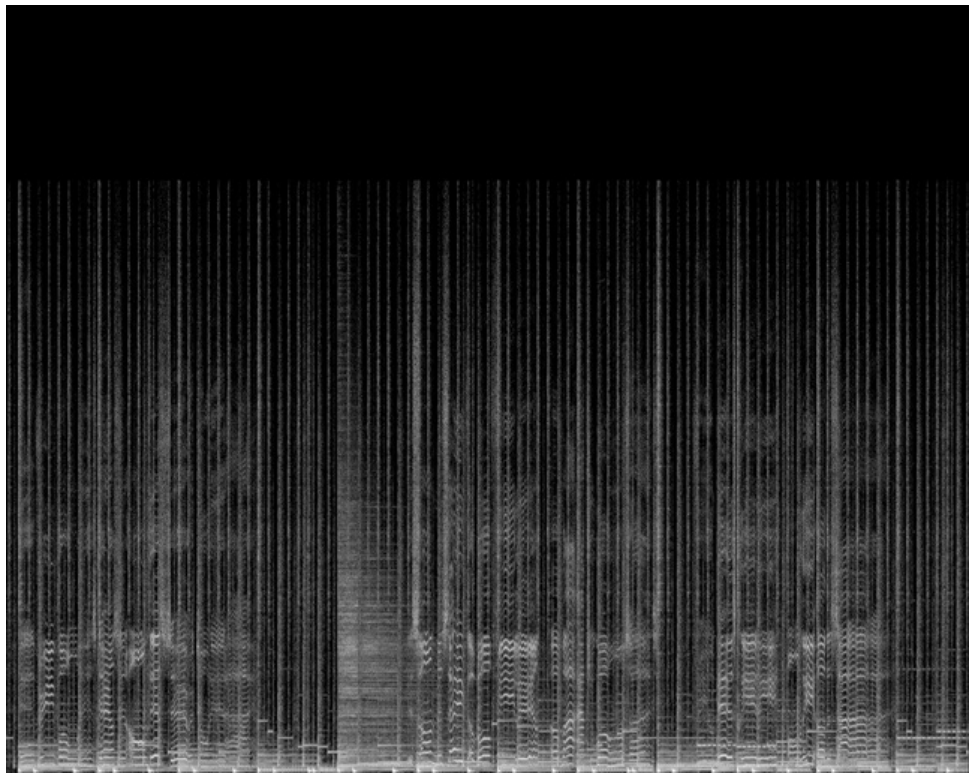
INTRODUÇÃO

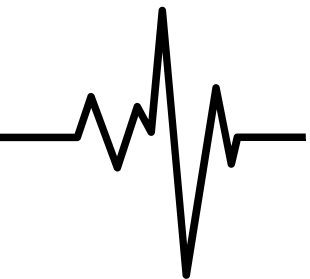
Aplicações

- KARAOKÊ
- ACAPELLA
- MASHUPS
- SAMPLES
- EDIÇÂO

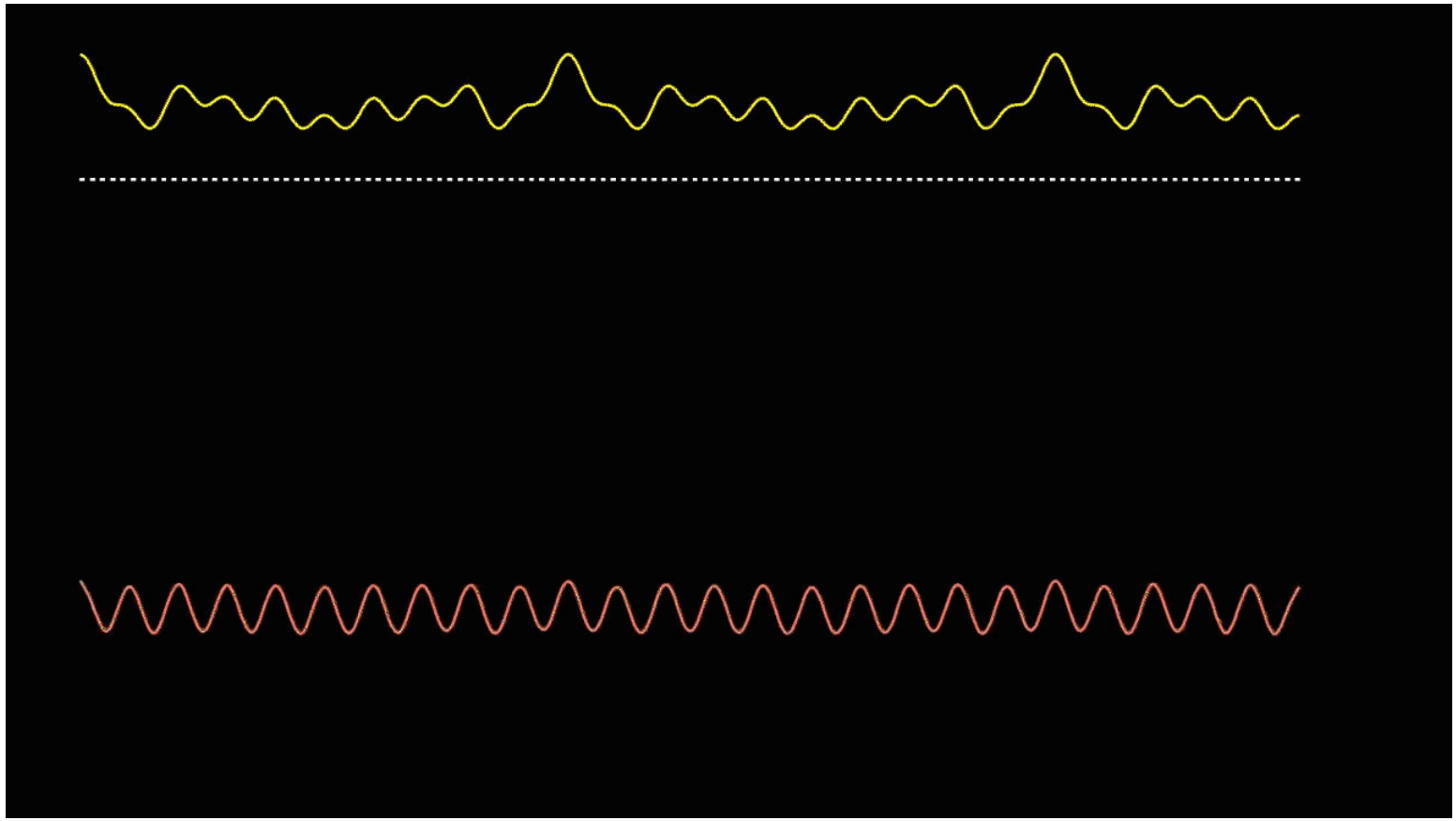


FUNDAMENTAÇÃO TEÓRICA

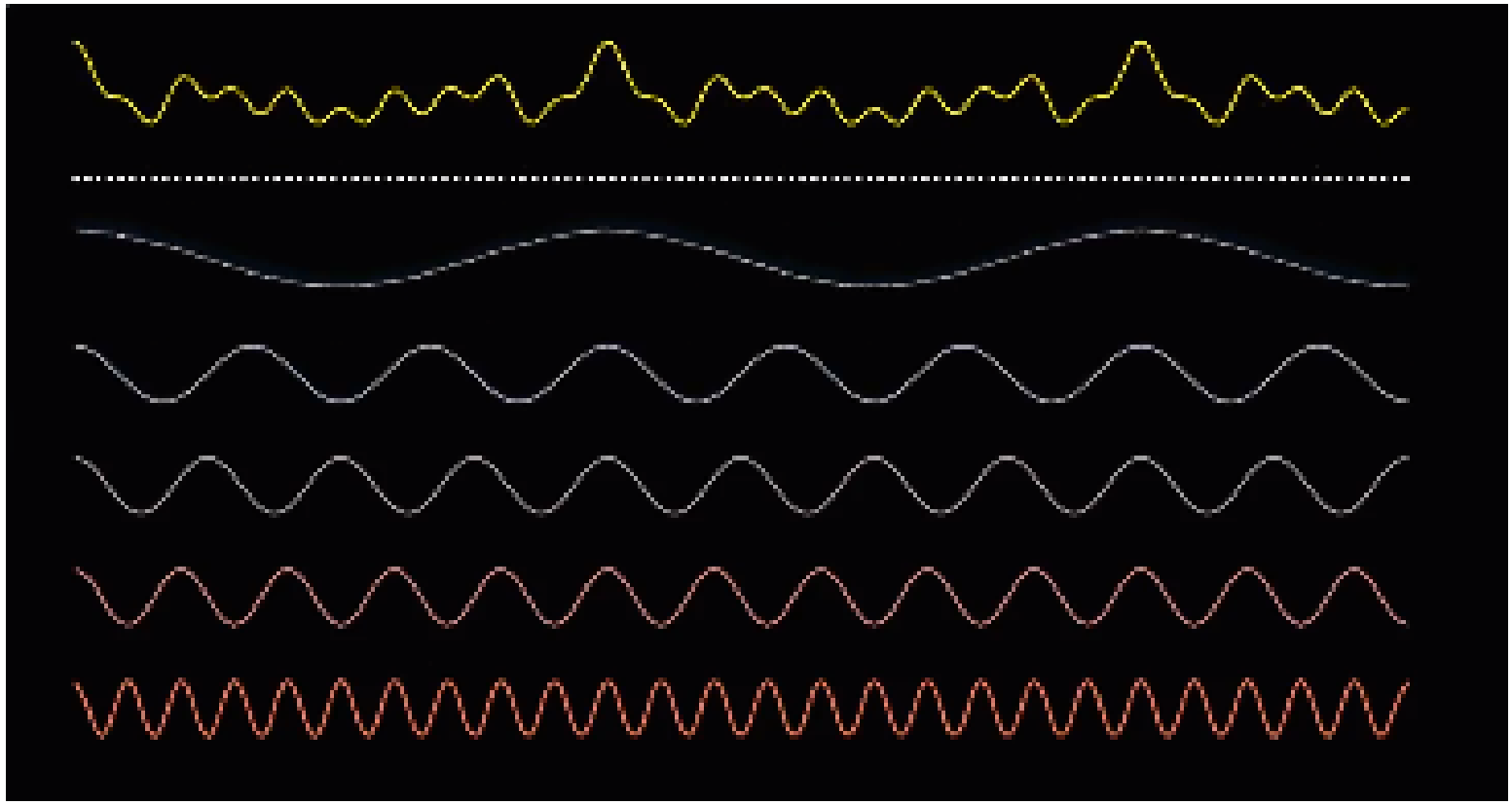




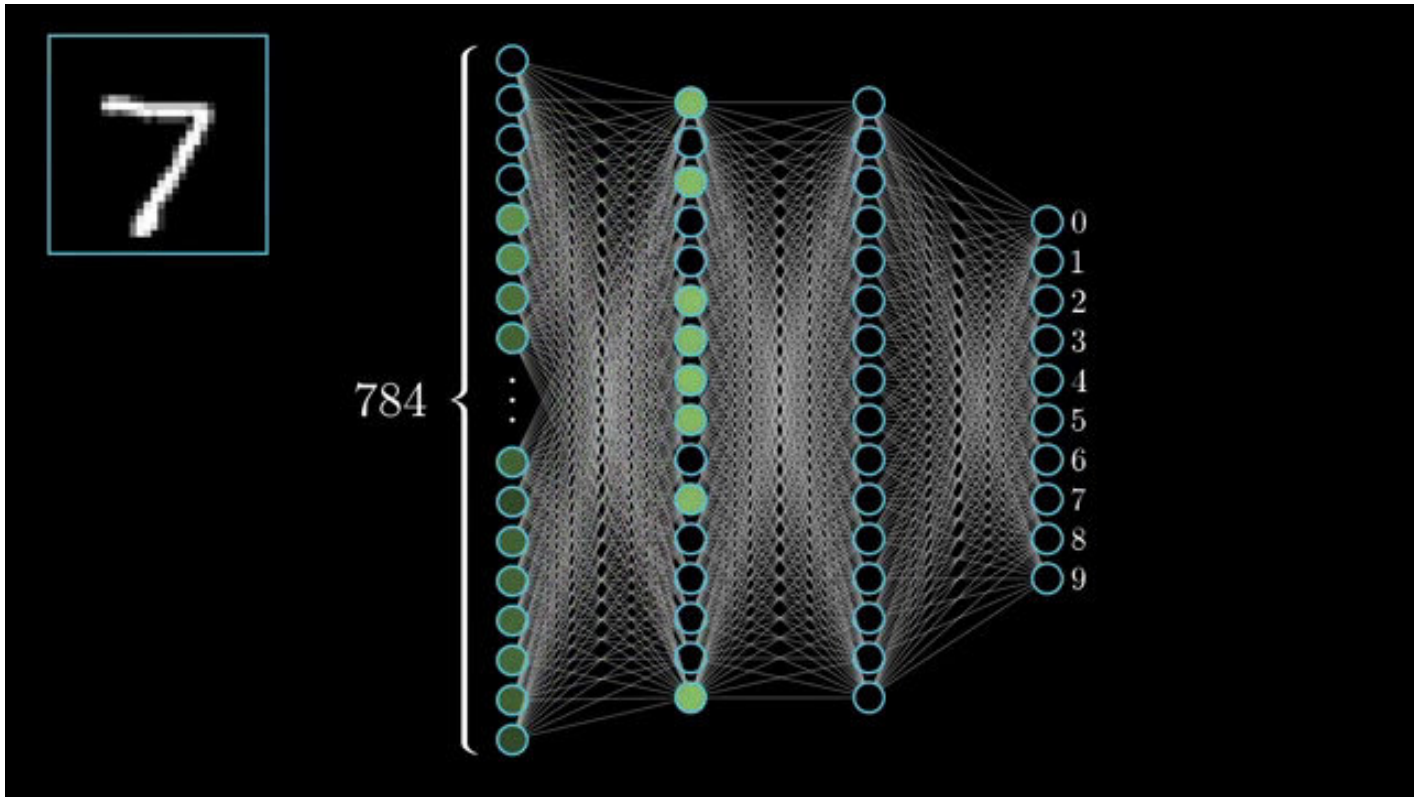
FUNDAMENTAÇÃO TEÓRICA



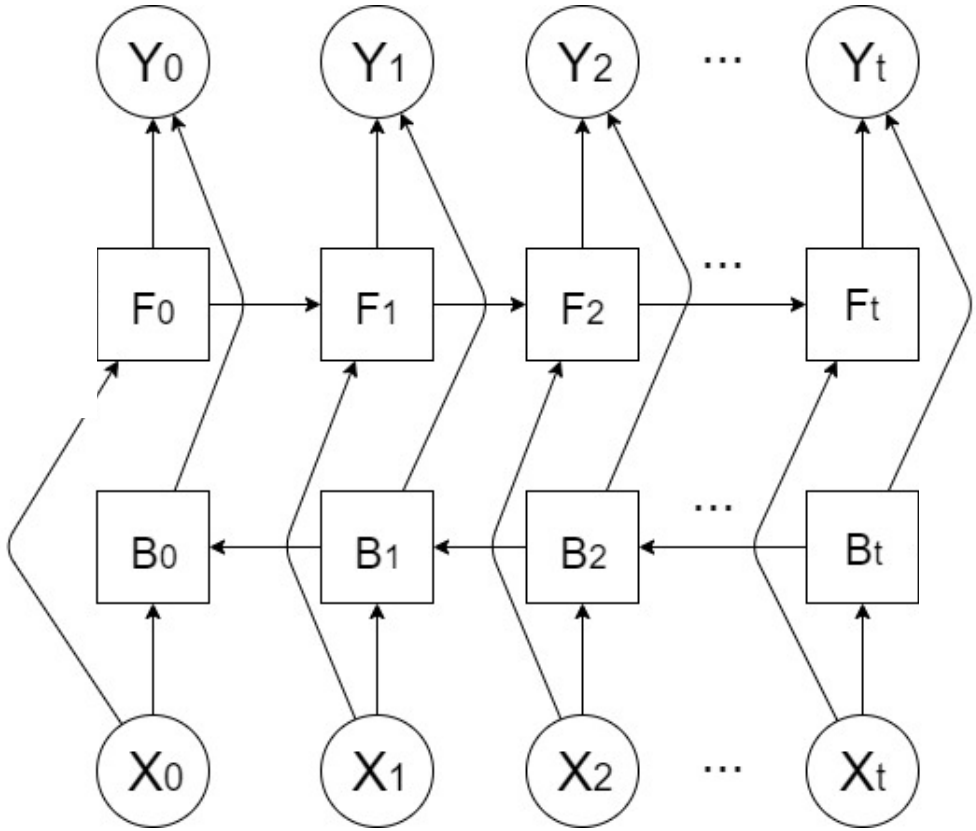
FUNDAMENTAÇÃO TEÓRICA



FUNDAMENTAÇÃO TEÓRICA



FUNDAMENTAÇÃO TEÓRICA



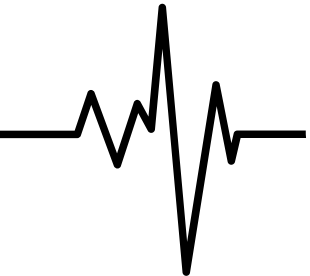
BASE DE DADOS

SOUND OF PIXELS:

YouTube

SISEC:

STEMS



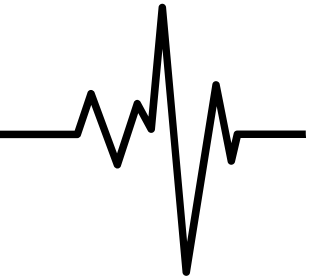
BASE DE DADOS

SOUND OF PIXELS:

- 536 solos de 11 instrumentos
- 149 duetos de 9 composições

SISEC:

- 150 arquivos que equivalem a 600 solos



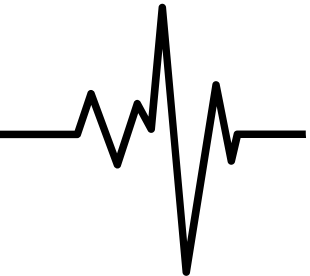
DESENVOLVIMENTO

REDES DE RECONHECIMENTO:

- 1025 nós de entrada
- 256 nós LSTM intermediários
- 15/19 nós de saida
- Unidirecional

REDES DE SEPARAÇÃO:

- 1025 nós de entrada
- 512 nós LSTM intermediários
- 1025 nós de saida
- Bidirecional



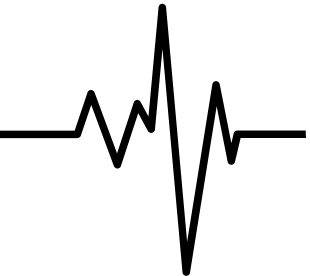
DESENVOLVIMENTO

REDES DE RECONHECIMENTO:

- Solos
- Duetos

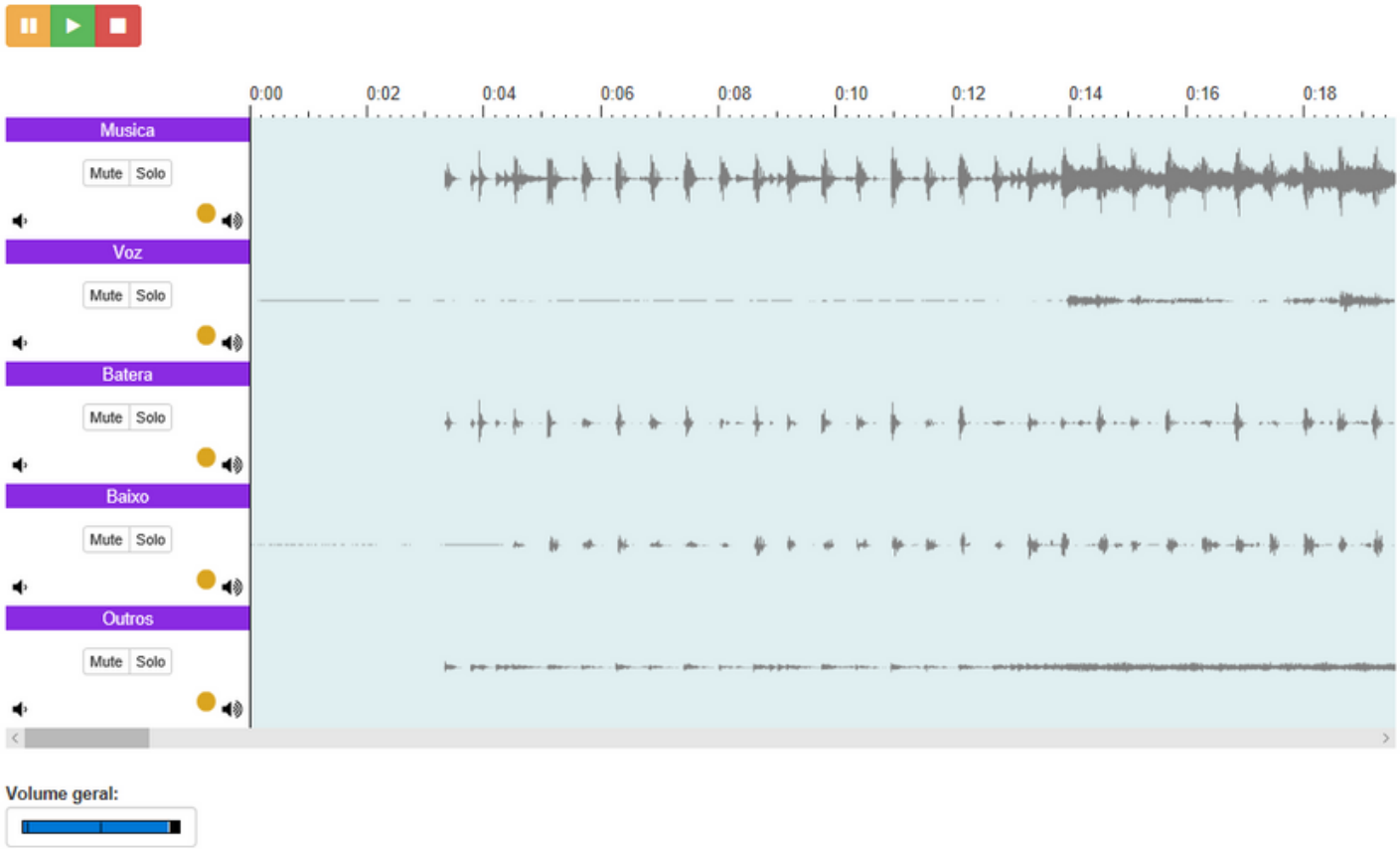
REDES DE SEPARAÇÃO:

- Voz
- Bateria
- Baixo
- Acompanhamento (violão, piano e etc.)



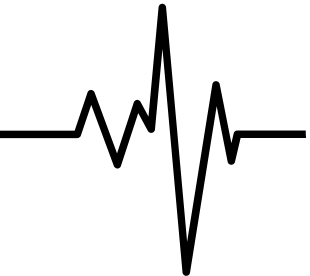
DESENVOLVIMENTO

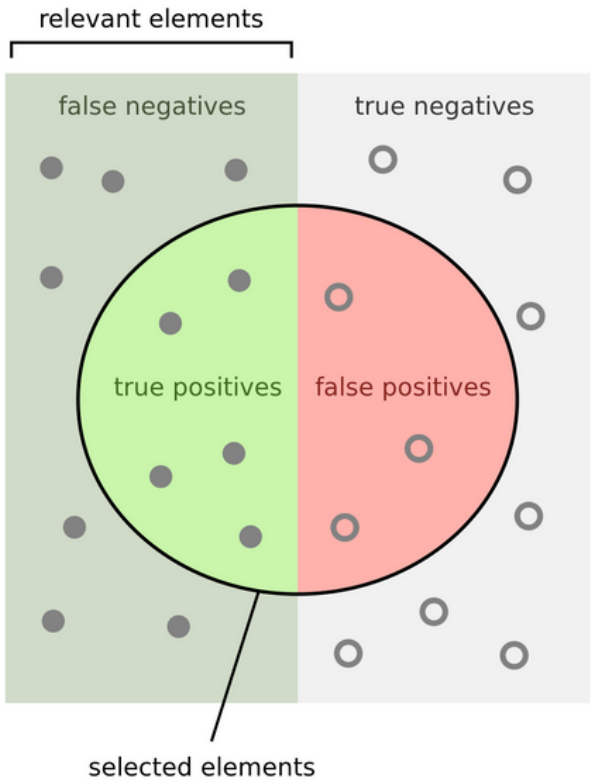
INTERFACE



RESULTADOS

REDE DE RECONHECIMENTO DE SOLOS



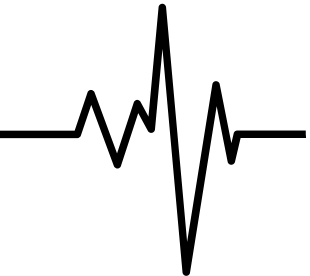


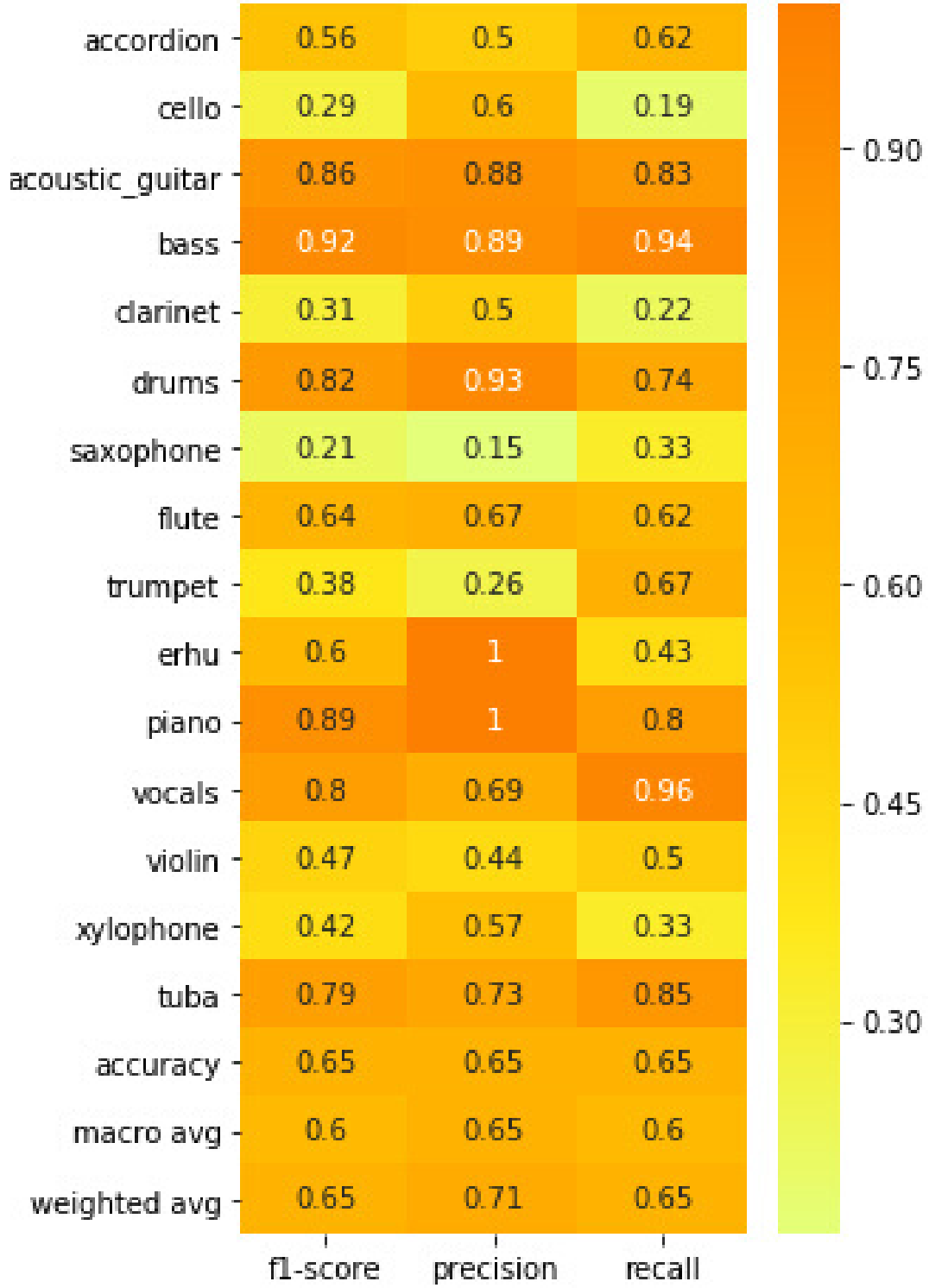
How many selected items are relevant?

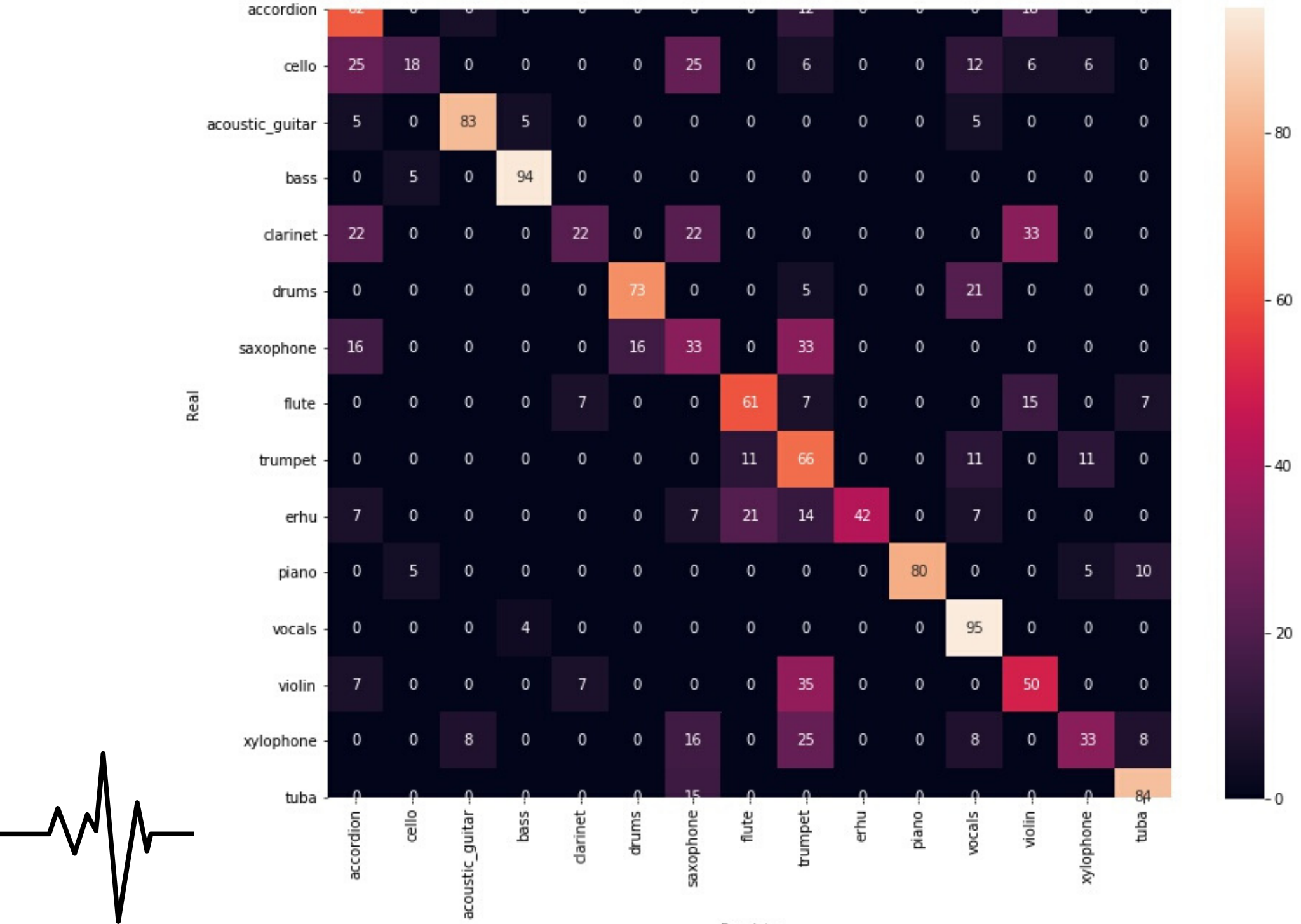
Precision =

How many relevant items are selected?

Recall =

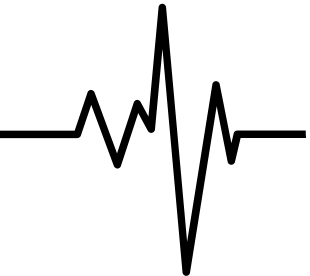


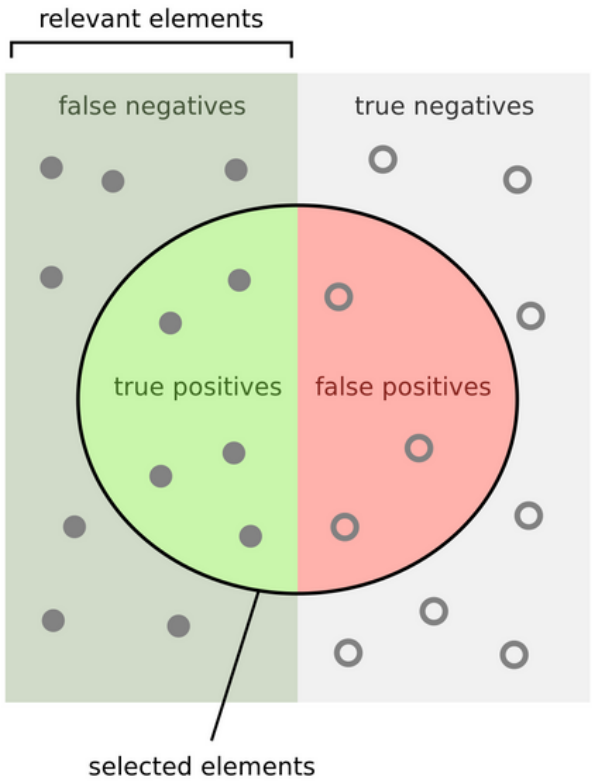



accordion
cello
acoustic_guitar
bass
clarinet
drums
saxophone
flute
trumpet
erhu
piano
vocals
violin
xylophone
tuba
Real
80
60
40
20
0

RESULTADOS

REDE DE RECONHECIMENTO DE DUETOS





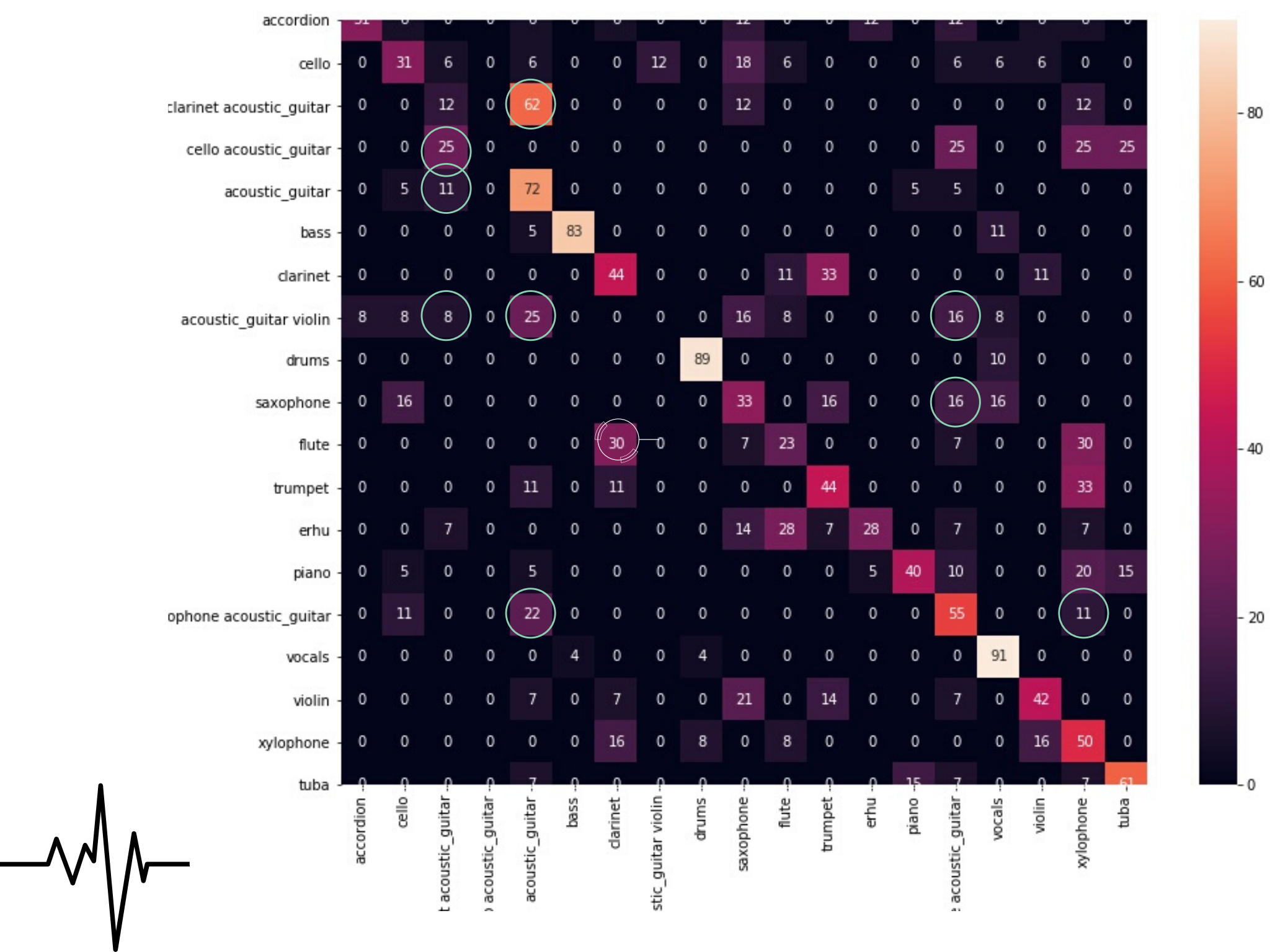
How many selected items are relevant?

Precision =

How many relevant items are selected?

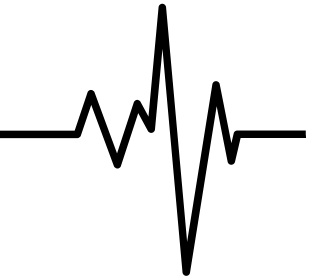
Recall =

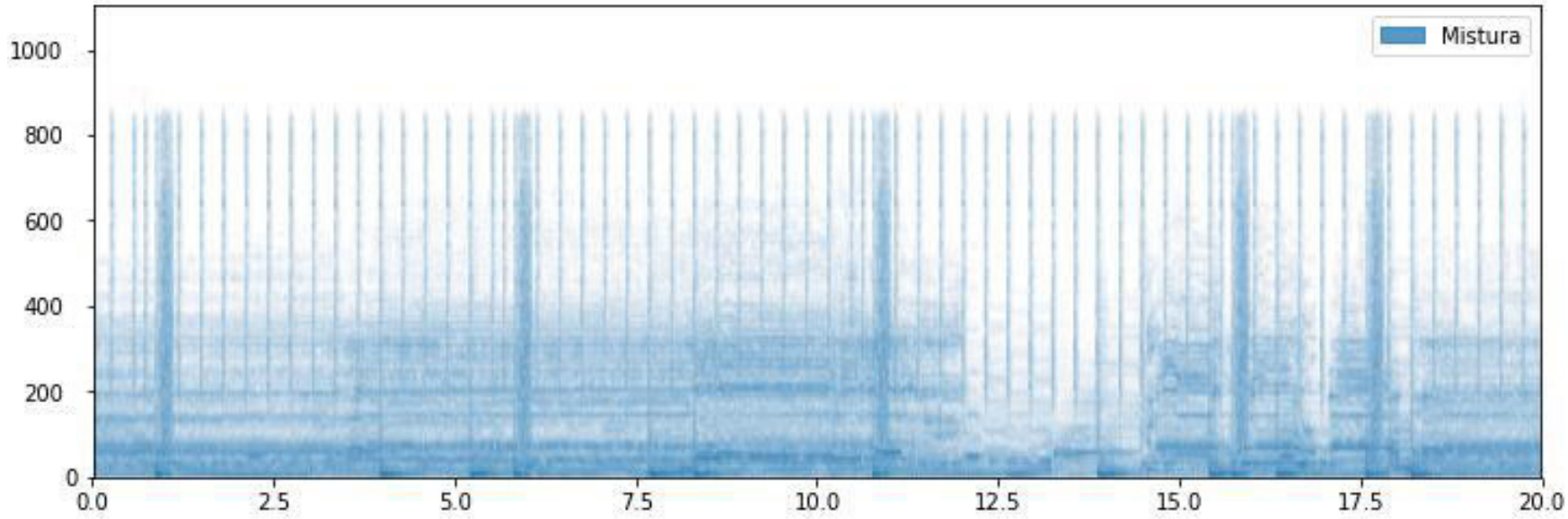
	f1-score	precision	recall
accordion	0.48	1	0.31
cello	0.24	0.33	0.19
clarinet acoustic_guitar	0.15	0.2	0.12
cello acoustic_guitar	0	0	0
acoustic_guitar	0.5	0.4	0.67
bass	0.91	1	0.83
clarinet	0.56	0.56	0.56
acoustic_guitar violin	0	0	0
drums	0.89	0.94	0.84
saxophone	0.083	0.056	0.17
flute	0.44	0.43	0.46
trumpet	0.5	0.45	0.56
erhu	0.35	0.44	0.29
piano	0.55	0.89	0.4
saxophone acoustic_guitar	0.29	0.19	0.56
vocals	0.86	0.79	0.96
violin	0.52	0.67	0.43
xylophone	0.46	0.35	0.67
tuba	0.62	0.62	0.62
accuracy	0.51	0.51	0.51
macro avg	0.44	0.49	0.45
weighted avg	0.51	0.58	0.51

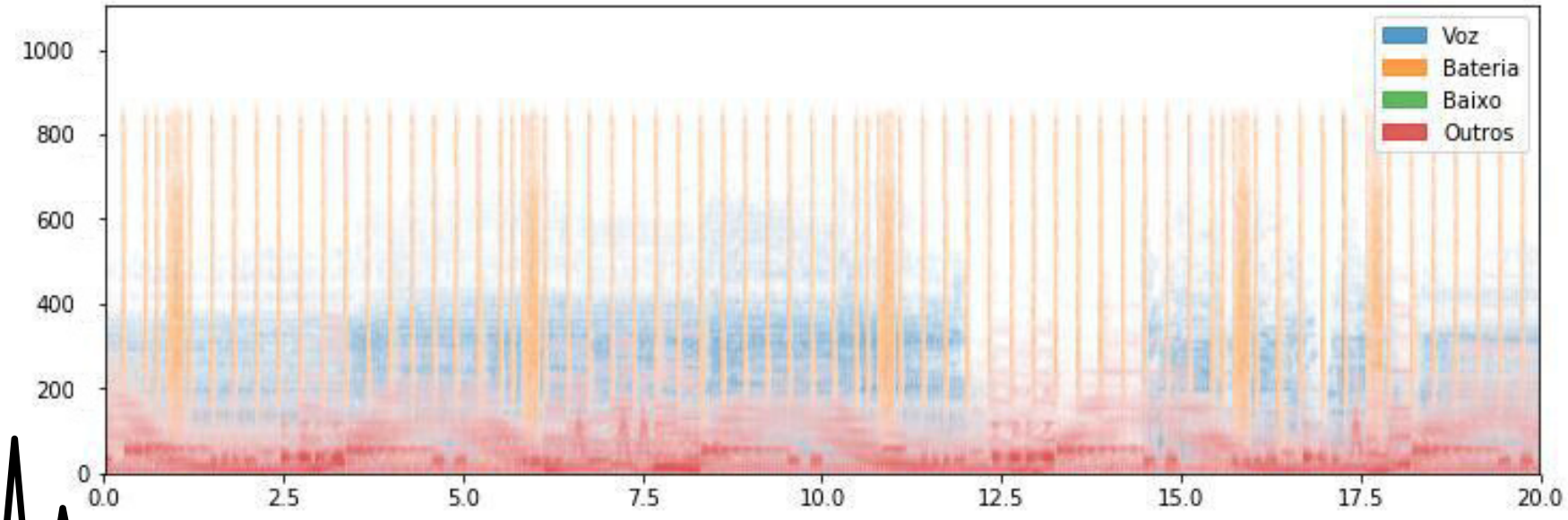

accordion
cello
clarinet acoustic_guitar
cello acoustic_guitar
acoustic_guitar
bass
clarinet
acoustic_guitar violin
drums
saxophone
flute
trumpet
erhu
piano
ophone acoustic_guitar
vocals
violin
xylophone
tuba
80
60
40
20
0

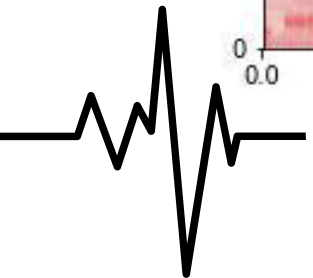
RESULTADOS

REDES DE SEPARAÇÃO




Mistura
1000
800
600
400
200
0
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0


Voz
Bateria
Baixo
Outros
1000
800
600
400
200
0
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0



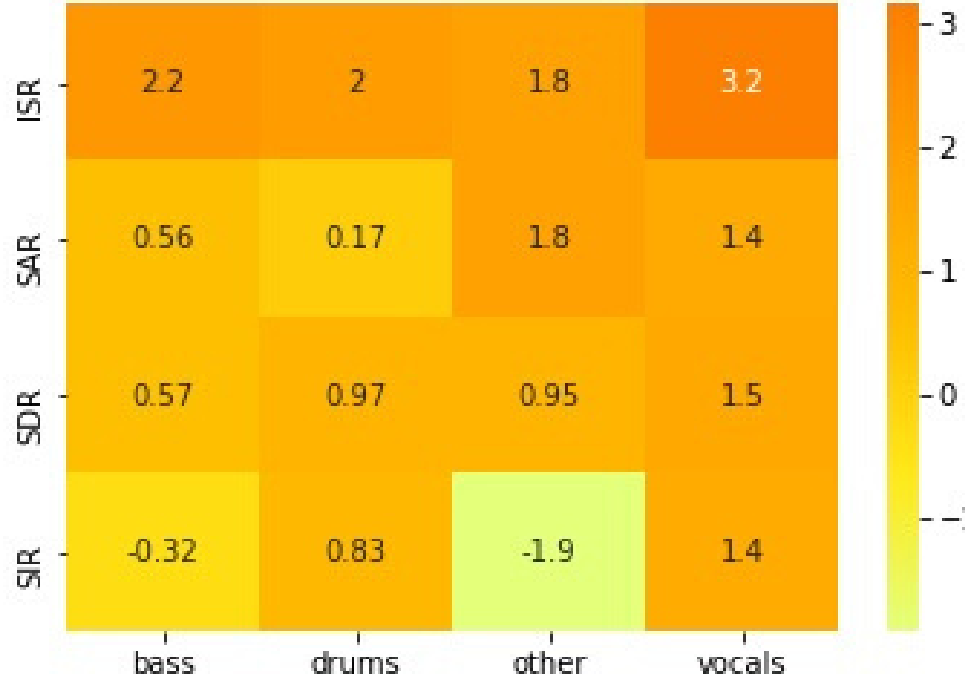
$$\widehat{s(t)} = s_{target}(t) + e_{spat} + e_{interf} + e_{artif}$$

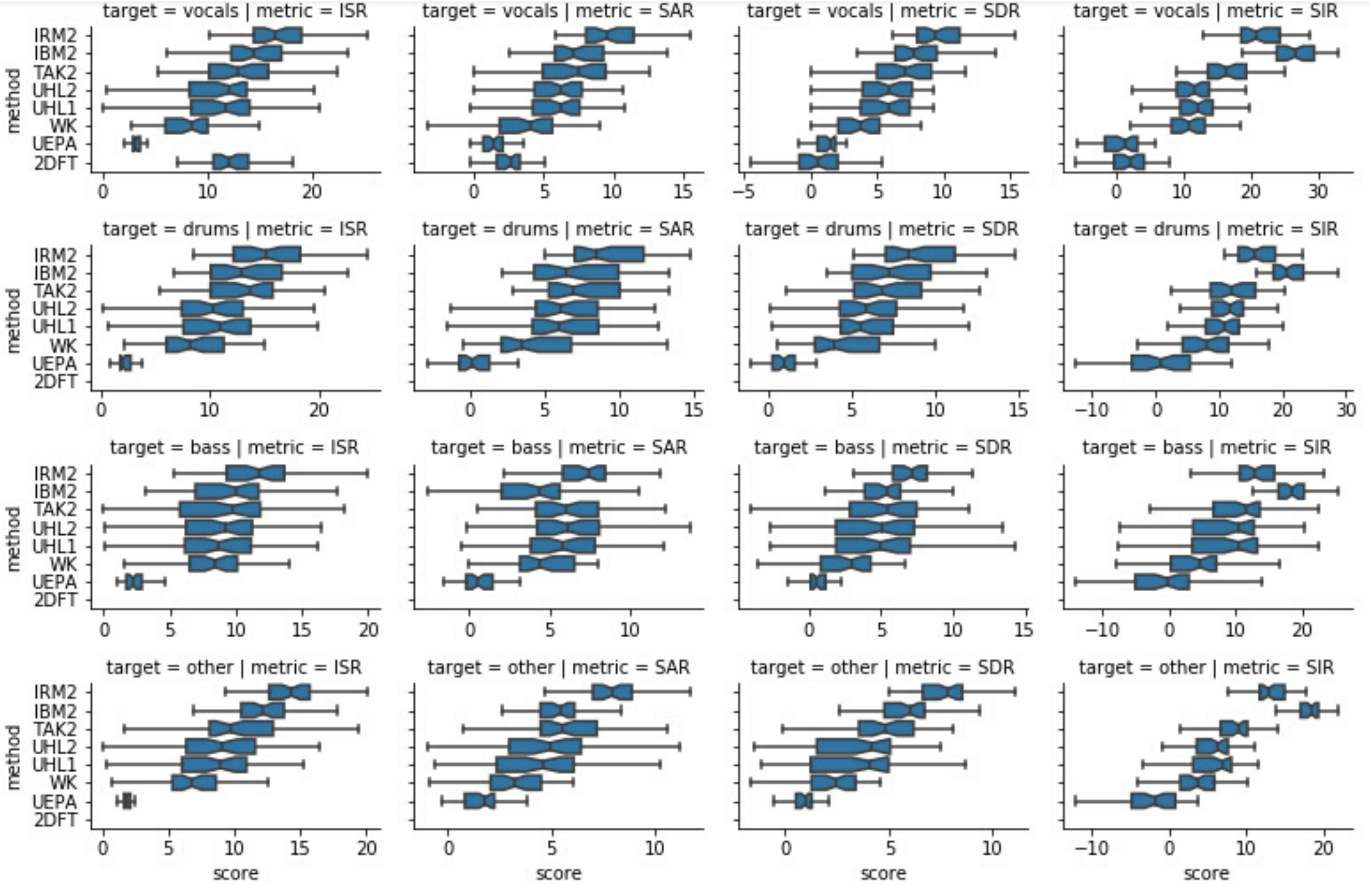
$$SDR = 10\log_{10}\frac{\|s_{target}(t)\|^2}{\|e_{interf} + e_{artif} + e_{spat}\|^2}$$

$$SIR = 10\log_{10}\frac{\|s_{target}(t) + e_{spat}\|^2}{\|e_{interf}\|^2}$$

$$SAR = 10\log_{10}\frac{\|s_{target}(t) + e_{spat} + e_{interf}\|^2}{\|e_{artif}\|^2}$$

$$ISR = 10\log_{10}\frac{\|s_{target}\|^2}{\|e_{spat}\|^2}$$




target = vocals | metric = ISR
target = vocals | metric = SAR
target = vocals | metric = SDR
target = vocals | metric = SIR
target = drums | metric = ISR
target = drums | metric = SAR
target = drums | metric = SDR
target = drums | metric = SIR
target = bass | metric = ISR
target = bass | metric = SAR
target = bass | metric = SDR
target = bass | metric = SIR
target = other | metric = ISR
target = other | metric = SAR
target = other | metric = SDR
target = other | metric = SIR
method
IRM2
IBM2
TAK2
UHL2
UHL1
WK
UEPA
2DFT
score

OBRIGADO

REFERÊNCIAS

HTTPS://GITHUB.COM/VISHWAJEET97/COCKTAIL-PARTY-PROBLEM
HTTPS://GITHUB.COM/NAOMIARO/WAVEFORM-PLAYLIST
HTTP://SISEC17.AUDIOLABS-ERLANGEN.DE/#/RESULTS/1/4/2
HTTP://BASS-DB.GFORGE.INRIA.FR/BSS_EVAL/
HTTPS://WWW.IRISA.FR/METISS/SASSEC07/?SHOW=CRITERIA