The journey to open-sourcing the code and models

FOSDEM 2020

Anis Khlif, Félix Voituret
Who’s been involved

- Romain Hennequin - Lead research scientist
- Laure Prétet - Former intern
- Anis Khlif - Research Engineer
- Félix Voituret - Research Engineer
- Manuel Moussallam - Head of Deezer Research
What is it all about?
Large impact on tech audience

- 9500+ stars on Github
- 200k+ views
- 100k+ read on deezer.io
Myth busting

- Deezer solved source separation
- Spleeter performs better than all other solutions

Spleeter by Deezer
What did we bring?

- State of the art
- Fast
- MIT Licensed
Primer on source separation
Waveform
Time-frequency representation

Hi, Dr. Elizabeth?
Yeah, uh... I accidentally took the Fourier transform of my cat...

Meow!
Magnitude spectrogram
Magnitude spectrogram

Primer on source separation
Magnitude spectrogram

Primer on source separation

In-harmonic Percussive content
Magnitude spectrogram

- Vocal content

Primer on source separation
Learn a **mask** for each instrument!

What fraction of the energy at each time and each frequency bin should be assigned to this instrument.
Magnitude spectrogram
Magnitude spectrogram
Magnitude spectrogram
Spleeter models

2, 4 & 5 stems
A deep learning approach to mask prediction

Deep learning model

Vocal mask

Instruments mask
4-stems

Deep learning model

- Vocal mask
- Drums mask
- Bass mask
- Others mask
5-stems

Deep learning model

Vocal mask
Drums mask
Bass mask
Piano mask
Others mask
Quick introduction to TensorFlow

- Build **computation graph** that represent a parametrized function
- Parameters (or **weights**) can be modified (trained) to fit an optimization function
- A **model** can be run in any tensorflow environment
- Some graph operations can be run very efficiently on GPU
Quick introduction to TensorFlow

\[ \text{model} = \text{computation graph} \ (\text{network architecture}) + \text{weights} \ (\text{parameters}) \]
Overview

Spleeter models

unet

masks

Voice

Instruments

L1 loss

* -

* -
Overview

Example 1

- unet
- masks
- Voice
- Instruments

L1 loss
Overview

Example 1

Parameter update

L1 loss

Spleeter models
Overview

Example 2
Overview

Example 2

Parameter update

L1 loss

Spleeter models
Overview

Example N

Voice

Spleeter models

Instruments

L1 loss

unet

masks
Overview

Example N

Parameter update

L1 loss

Spleeter models

Voice

unet

masks

Instruments
Dataset

- In-house dataset of tracks
- ~24k tracks with stems
- ~80 hours of recording

... that we are not allowed to release!
BUT...
Training

We can release learned weights

- One mask per channel per source.
- 1 branch predicts masks for 1 source, with 2 channels
- ~10M parameters per branch
Open-sourcing Spleeter

Packaging & distribution
Packaging constraints

- Predefined configurations
- On demand model downloading
- Oneliner command
Predefined configurations

- JSON formatted file
- Mostly model related parameters
- Provided as
  - File path
  - Configuration name
Using GitHub releases as model hub

Spleeter → deezer/spleeter

Assets

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>2stems-finetune.tar.gz</td>
<td>206 MB</td>
</tr>
<tr>
<td>2stems.tar.gz</td>
<td>69.7 MB</td>
</tr>
<tr>
<td>4stems-finetune.tar.gz</td>
<td>409 MB</td>
</tr>
<tr>
<td>4stems.tar.gz</td>
<td>140 MB</td>
</tr>
<tr>
<td>5stems-finetune.tar.gz</td>
<td>512 MB</td>
</tr>
<tr>
<td>5stems.tar.gz</td>
<td>174 MB</td>
</tr>
<tr>
<td>checksum.json</td>
<td>248 Bytes</td>
</tr>
</tbody>
</table>
Separate source from command line

Separate with default 2stems configuration

```
$ spleeter separate -i input_file.mp3 -o output_dir
```

Separate with specific embedded configuration

```
$ spleeter separate -i input_file.mp3 -o output_dir -p spleeter:4stems
```
Distribution constraints

Critical dependencies to manage:

- Cross platform
- Cross hardware
- User friendly

Open-sourcing Spleeter
Distribution channels
Continuous integration and delivery
Legal considerations

- No Intellectual Property consensus on **weights**
- We decided to open-source **model**
Bibliography and references

Industrial integrations:

- **AconDigital** plugins
- Various public web applications

~30 projects referenced as using Spleeter on GitHub

Research publications:

Thank you

research@deezer.com