AIRI × Университет Сириус

Voice Filter

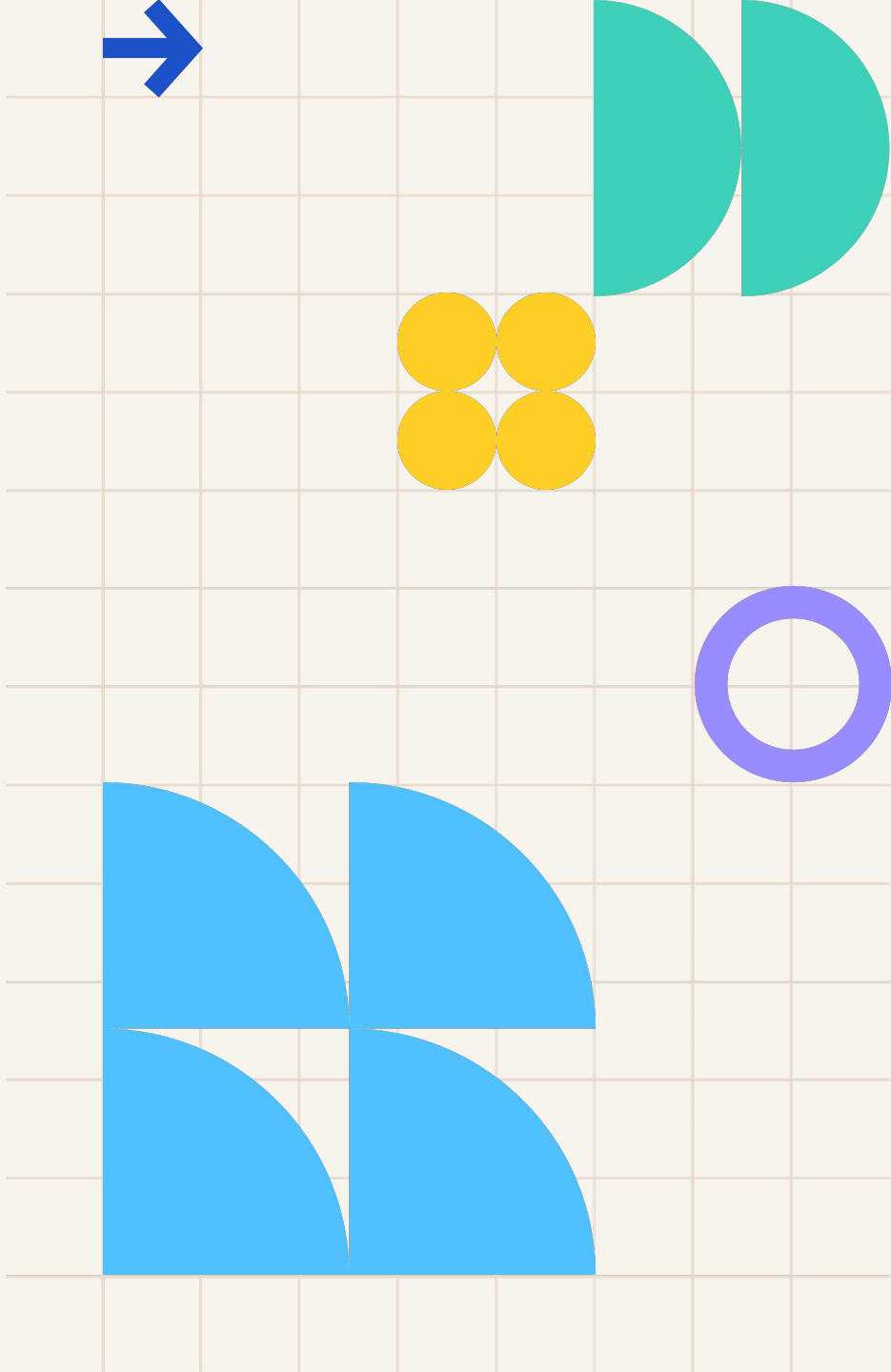
Maksim Ivanov
MSU

Dmitrii Korzh
Skoltech

Makhneva Elizaveta
Skoltech

Fedotovskikh Evgenii
Polytech

Mentor: Egor Shvetsov



AGENDA

01 Problem description

02 VoiceFilter

03 Results

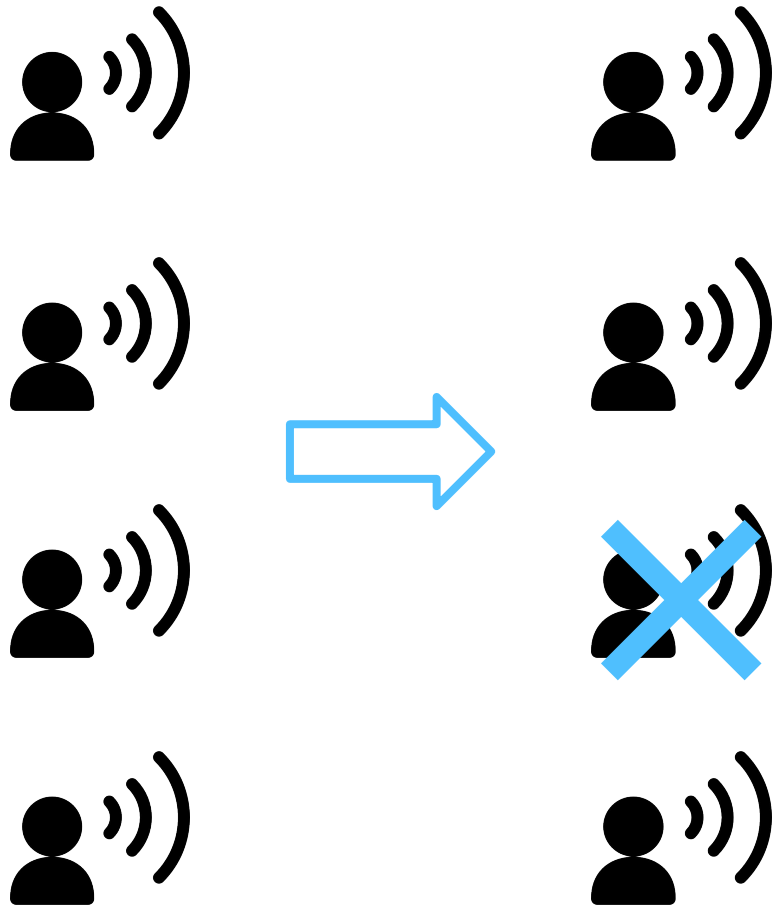
01

Problem description

Problem description

It is possible to remove all but one speaker's voice from an audio sample given his speech embedding. Such solutions are commonly used with speech recognition devices.

However, it is also possible to solve reverse problem, remove speaker's voice except all others voices or sounds given his embedding. Such solution can be used in audio editing tools.

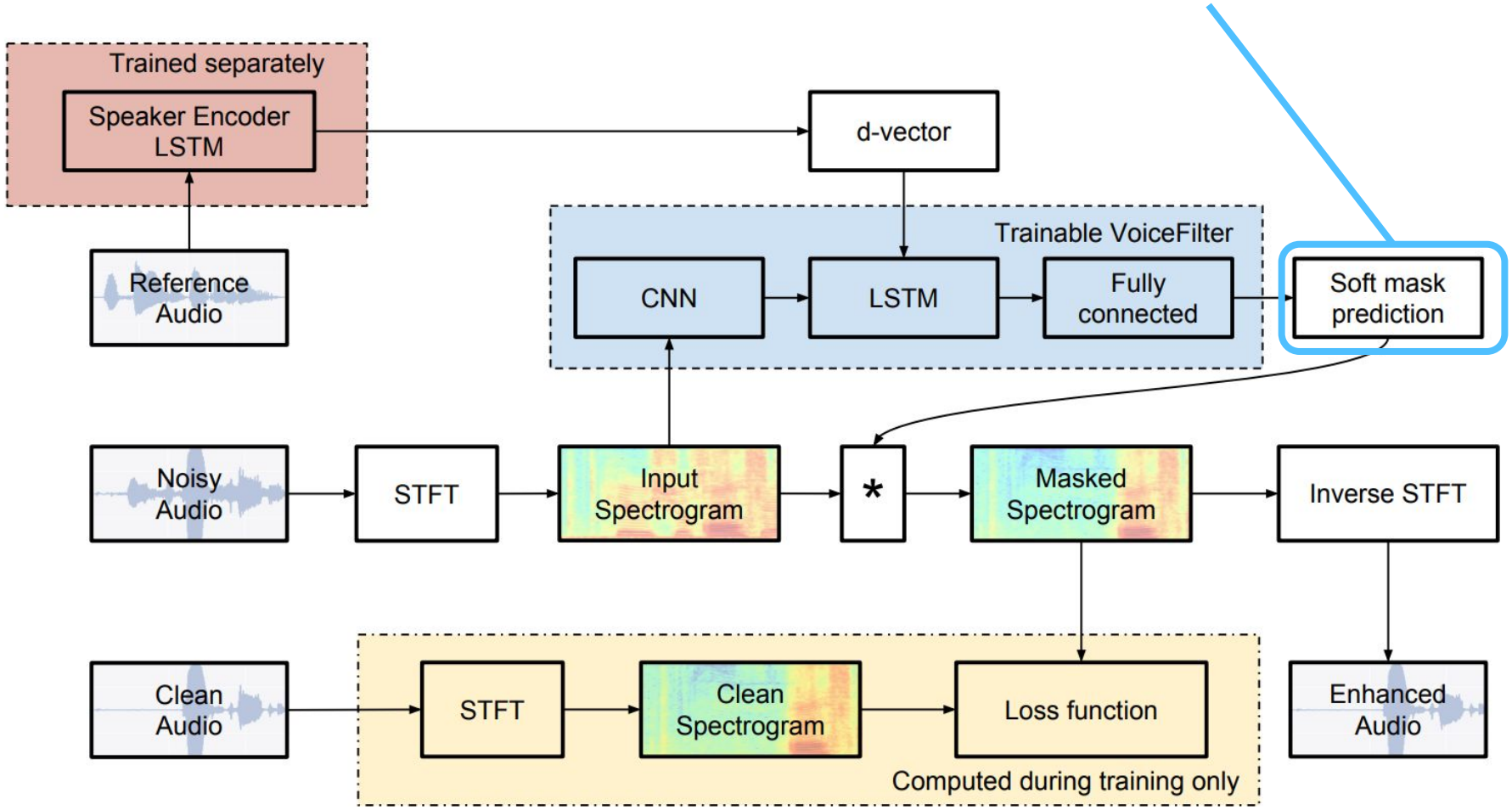


02

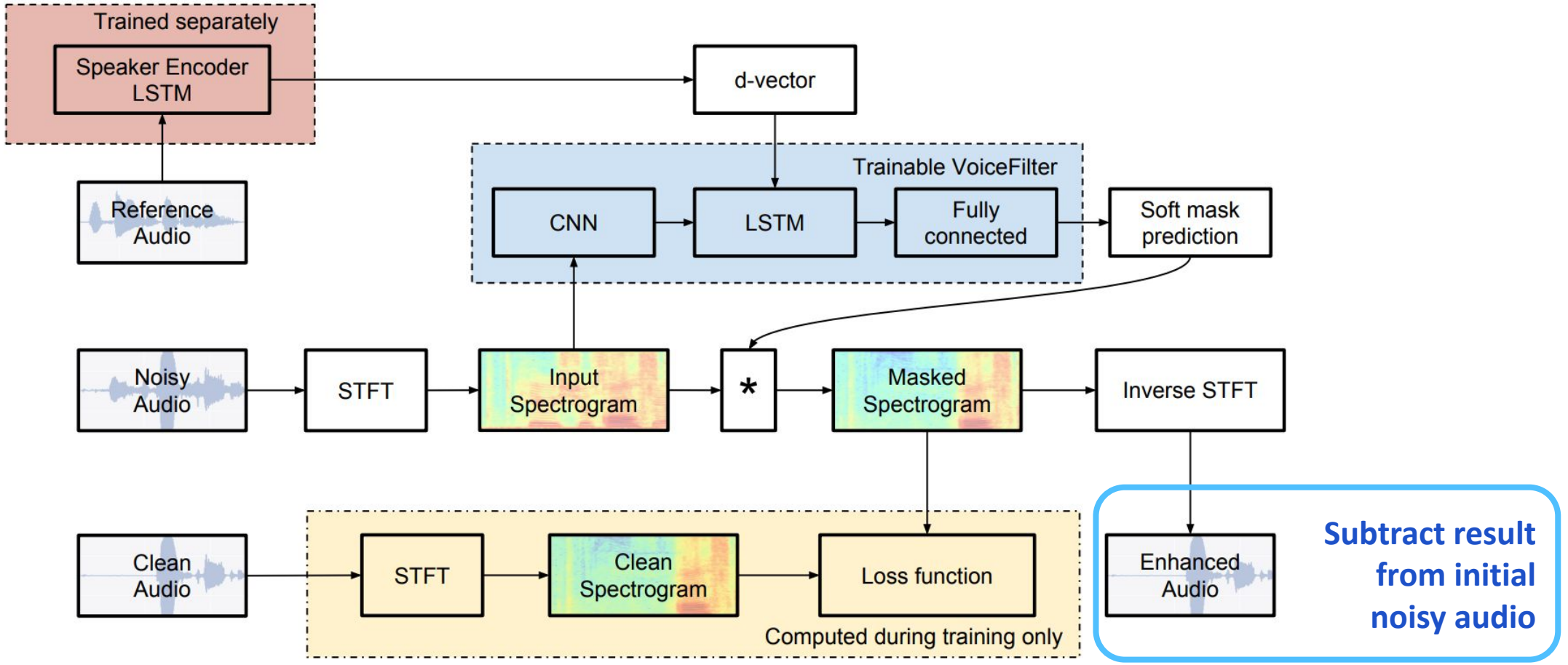
VoiceFilter

VoiceFilter

We learn the mask using speaker's embedding and use it modify original MEL spectrogram to remove the target speaker



Reverse VoiceFilter



03

Results

Better speech recognition

```
1 pipe(x) #orig normalized wav
```

```
{'text': 'HE HAD WRITTEN A NUMBER OF BOOKS HIMSELF'}
```

```
1 pipe(mixed) #mix of w1 and w2
```

```
{'text': 'E HELL IT ON HIS OTHER OT TO SASTO MUSANS NE HE'}
```

```
1 pipe(est_wav) #estimating target from mixed noisy audio
```

```
{'text': 'E TELLS US THAT AT THIS FESTIME SEASON OF THE YEAR'}
```

```
1 pipe(mixed - est_wav)
```

```
{'text': 'HE HAD WRITTEN A NUMBER OF BUILTS IN SHADOWS AMONG THEM A HISTO'}
```

Examples

Synh audio

- Liza + Zenek synh
- Liza Forward
- Liza ReverseBase
- Reverse. Without Liza
- Zenek ReverseBase
- Zenek Forward
- Reverse. Without Zenek

Asynh audio

- Liza, then Zenek
- Liza ReverseBase
- Reverse. Without Liza
- Zenek Forward
- Reverse. Without Zenek

AIRI × Университет Сириус

Discussion of results

- The forward method works well, increases ASR quality;
- Baseline of Reverse (subtraction) works satisfactorily;
- We train reverse modell but its quality is worse then for forward model, because of lack of computational power and time to adjust it;
- One straightforward way to try to improve results is to apply different loss.

Project plan

	Maksim	Dmitrii	Elizaveta	Evgenii
Read the paper	+	+	+	+
Choose the implementation of VoiceFilter	+	+	+	+
Replicate results of the paper	+	+	+	+
Implement reverse algorithm	+	+		
Test algorithm	+	+		
Make dataset			+	+
Check the quality	+	+	+	+

Artificial Intelligence Research Institute

airi.net

- airi_research_institute
- AIRI Institute
- AIRI Institute
- AIRI_inst
- artificial-intelligence-research-institute