AI-Driven Vocal Tract Segmentation for Speech Disorders Analysis

Project Workshop 2 - Neuroengineering 2024/25

Professor: Cerveri Pietro
Tutor: Cavicchioli Matteo

Baggini Guia
Burinato Federica
De Carlo Federico
Martuscelli Sara
Missana Matteo

What happens in the **vocal tract** of patients with **non-fluent variant primary progressive aphasia**?

Two kinds of **aphasia**

Dysarthria

Consistent speaking patterns

Apraxia of speech

Inconsistent speech patterns

How Doctors Diagnose Aphasia

Electropalatography

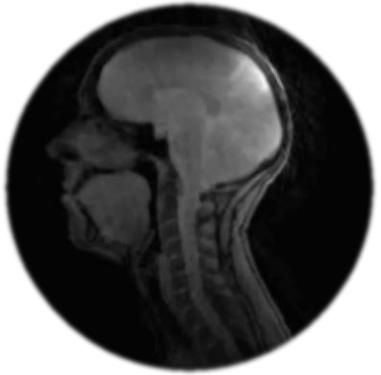
✕ Low spatial coverage

Cine X-Ray

✕ Invasive

Electromagnetic Articulography

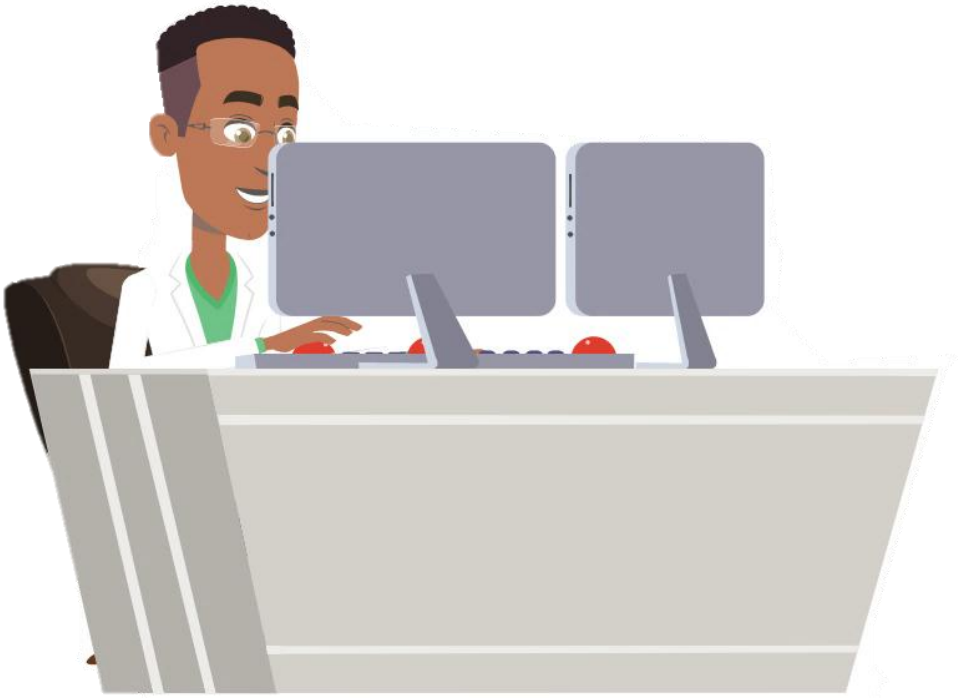
✕ Alter articulator cinematics



dsMRI

✓ Better frame rate and temporal resolution

dsMRI



Vocal Tract Segmentation

Gold Standard:

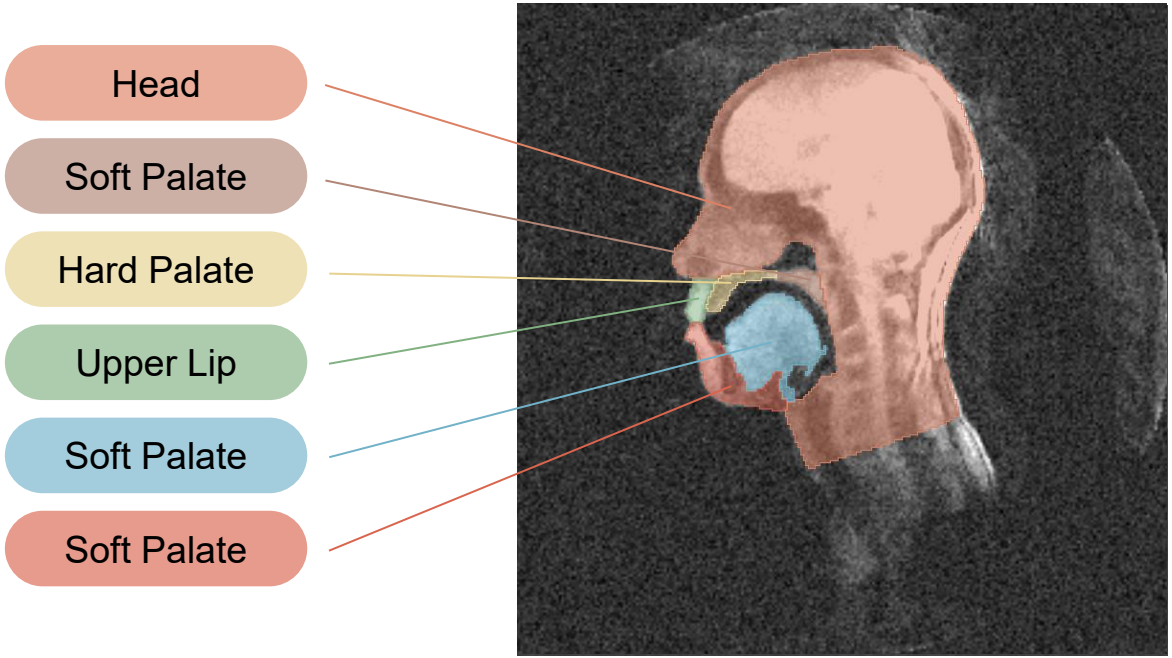
Expert-driven segmentation

Extra tool:

Automatic segmentation with AI

Our **Goal**

Develop a **NN** for **Automatic segmentation** of articulators from **dsMRI images** of the vocal tract



DATASET

- S1_MICROSCOPIC
- S1_SEGREGATION
- S1_TOPCOP

- S2_MICROSCOPIC
- S2_SEGREGATION
- S2_TOPCOP

- S4_MICROSCOPIC
- S4_PATAKA
- S4_WELCOME

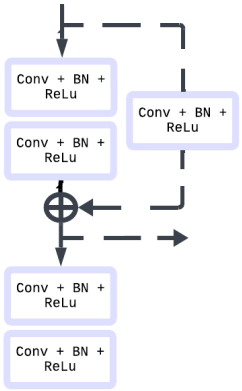
Training Set → **Data Augmentation**

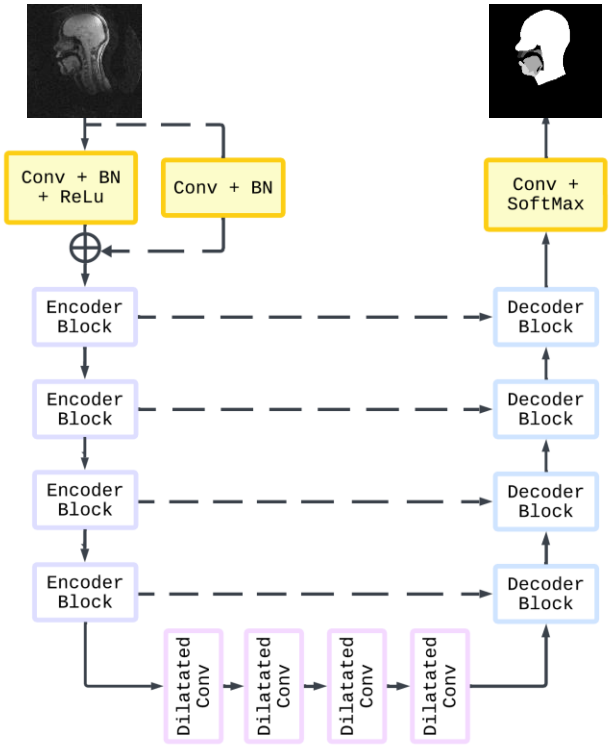
Horizontal flip
Brightness
Contrast

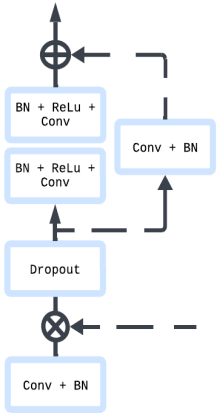
- S5_COUNT
- S5_KA
- S5_PA

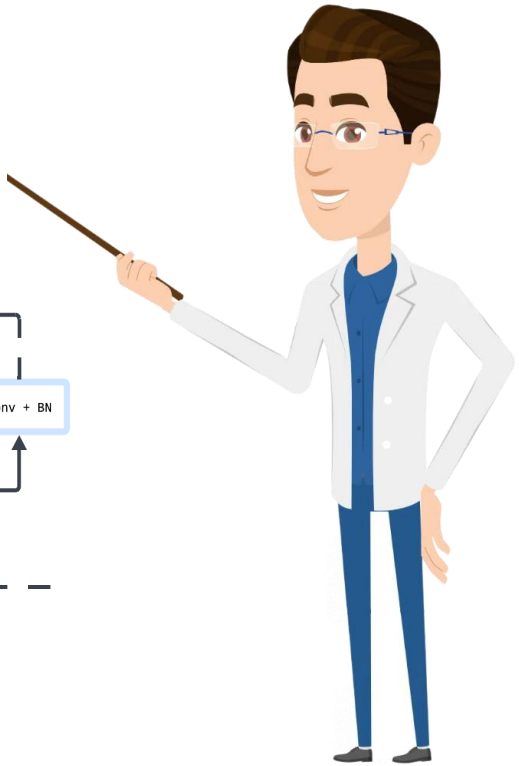
Validation & Test Set

Paper's Architecture

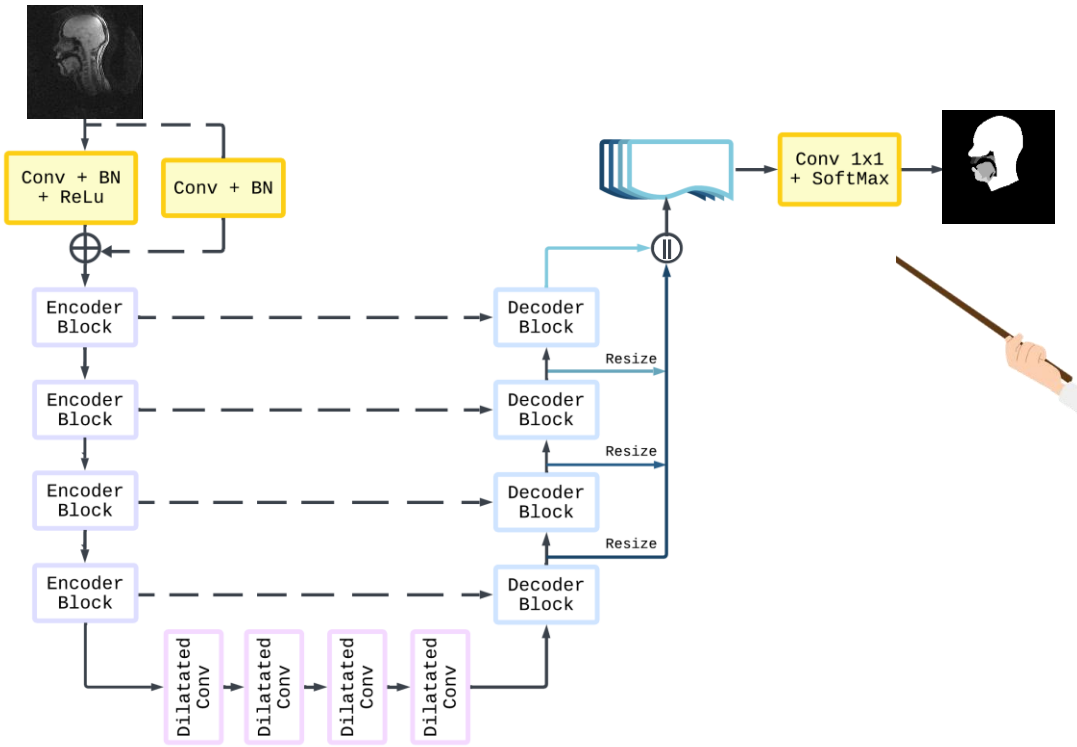








Our Architecture

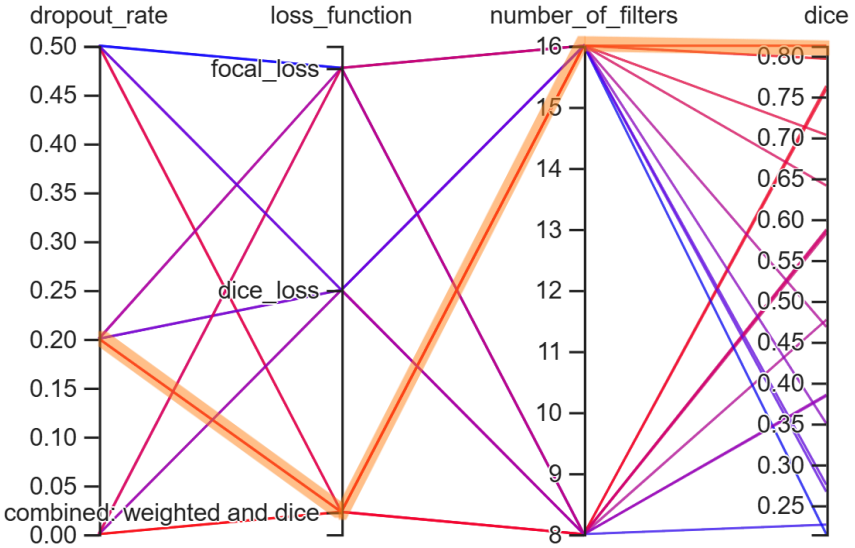


Grid Search

Dropout Rate

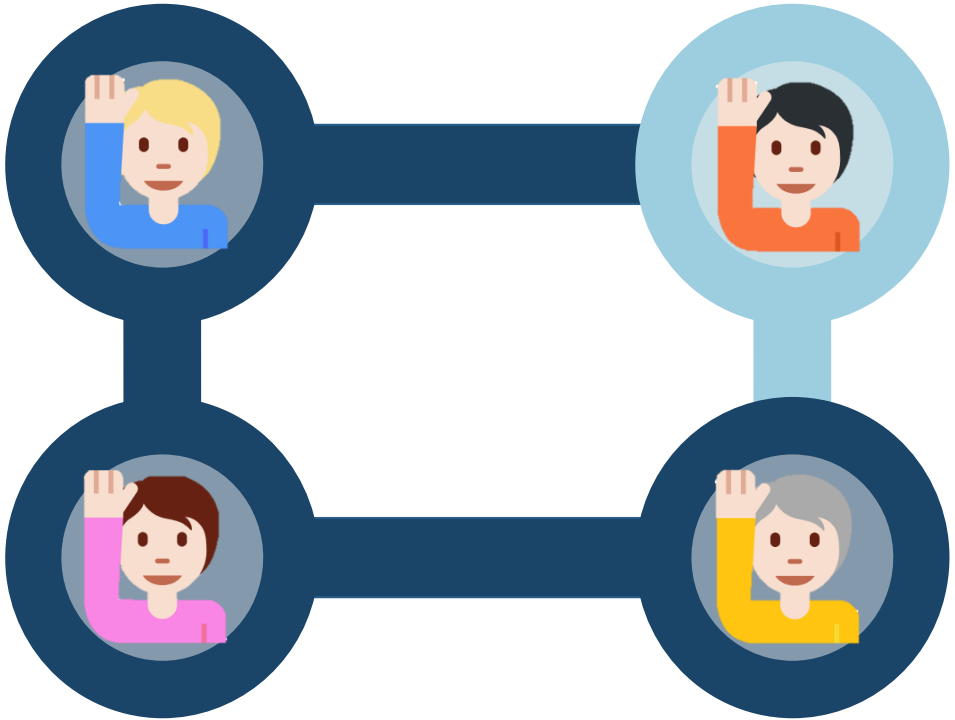
Loss Function

Number of filters



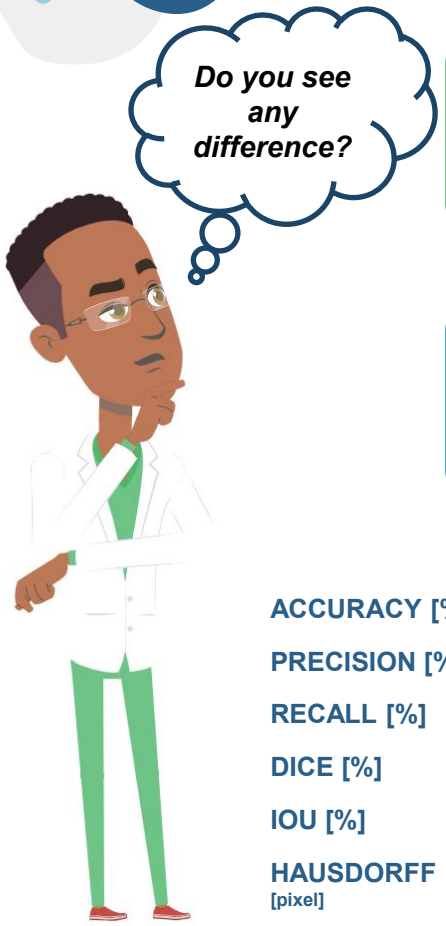
Leave-one-patient-out Cross-Validation

Training Set

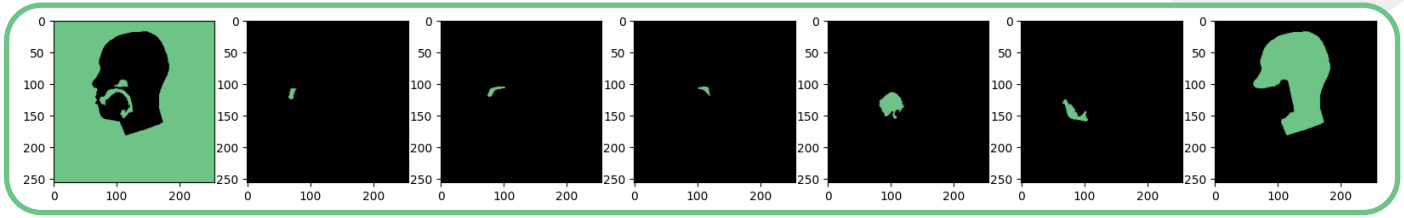


Validation & Test Set

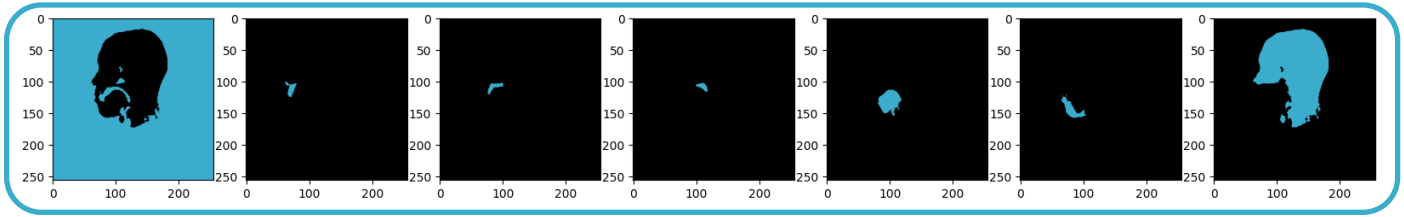
Test Set Results



Ground truth

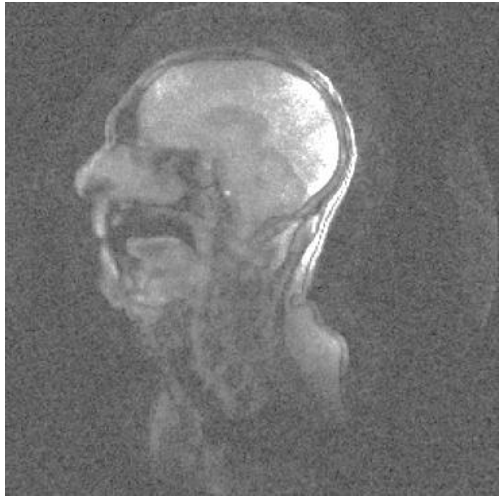


Predictions

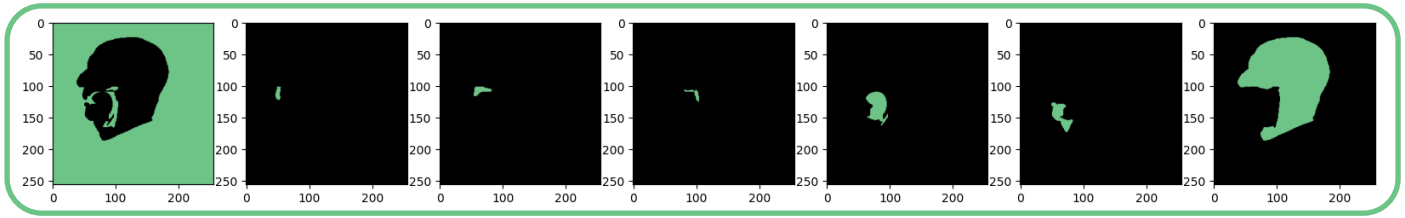


	Background	Upper Lip	Hard Palate	Soft Palate	Tongue	Lower Lip	Head
ACCURACY [%]	98,0	99,9	99,9	99,9	99,8	99,8	98,0
PRECISION [%]	98,0	71,7	73,7	82,9	97,5	90,7	97,3
RECALL [%]	99,5	94,2	79,1	82,9	88,4	86,6	91,4
DICE [%]	98,7	81,2	76,2	82,6	92,7	88,5	94,2
IOU [%]	97,5	68,4	62,6	70,6	86,4	79,5	89,1
HAUSDORFF [pixel]	7,04	7,33	2,94	2,39	4,87	5,84	12,93

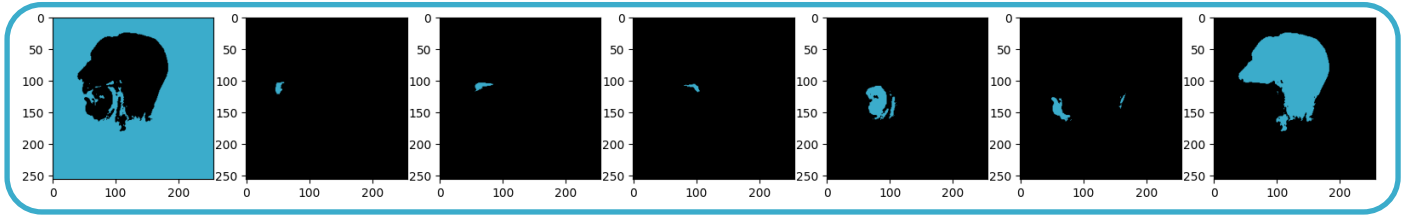
Pathological Patient



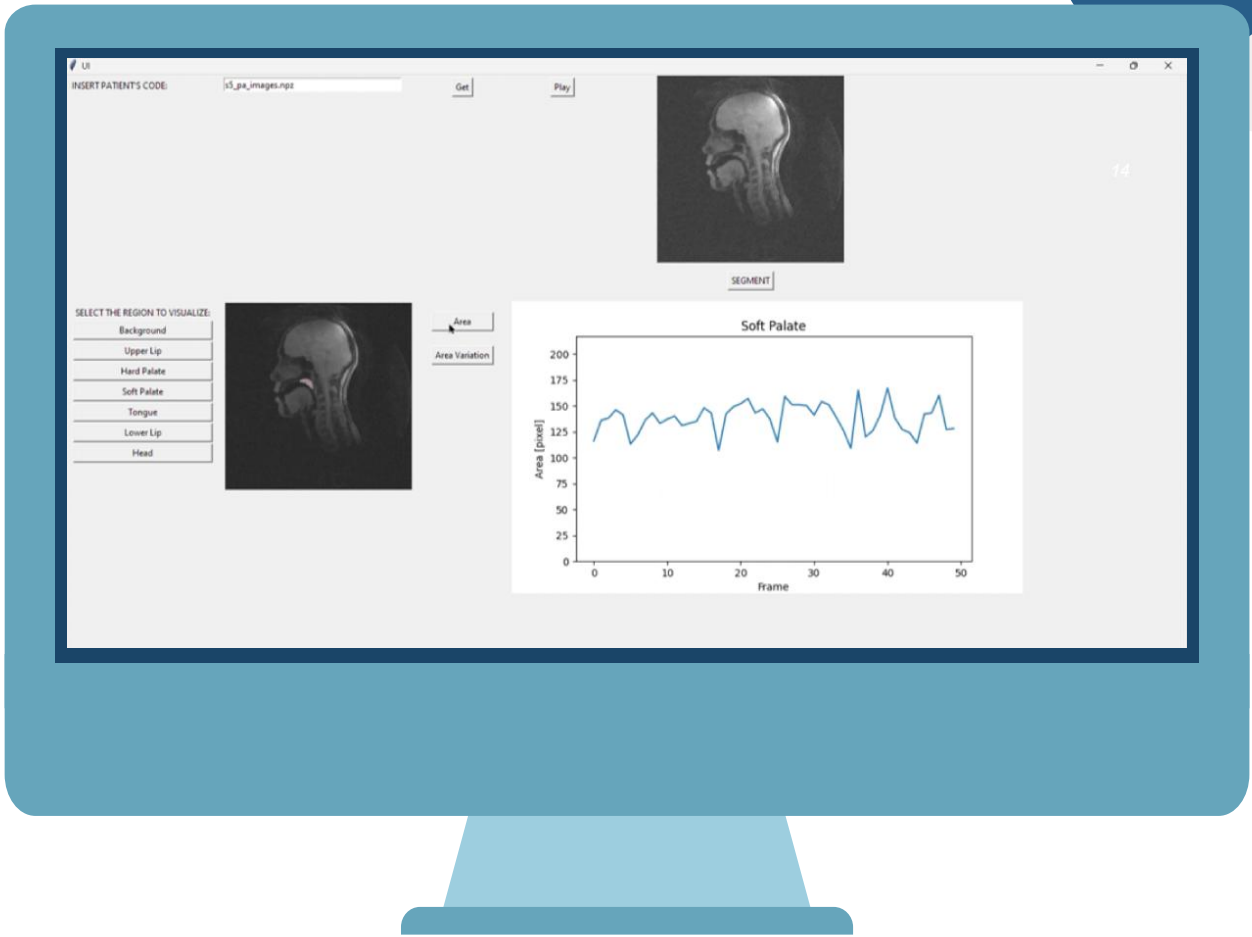
Ground truth



Predictions



User-friendly GUI



Limitations

Automatic segmentation cannot substitute human experts

Small dataset size

Overfitting on the Gaussian Noise

Future Steps

Automatic distinction between apraxia and dysarthria

Improvement of the assistive tool to help doctors in make the right diagnosis

Thank you for your kind attention!

Sara Martuscelli

Matteo Missana

Federica Burinato

Federico De Carlo

Guia Baggini