STCI: Screening Tool for Cognitive Impairment

DSR-Batch 41- Demo Day
April 15th, 2025

Alfonso Hernández

Motivation

What is cognitive impairment?

- Age-Associated Cognitive Decline (AACD).
- Mild Cognitive Impairment (MCI) .
- Dementia: Alzheimer’s Disease (AD) among others.

Typical screening: Mini-Mental State Examination (MMSE):

- A 30-point questionnaire.
- Takes between 5-10 min.
- Useful for moderate-to-severe impairment detection.
- Lacks sensitivity to mild impairment and often fails to detect MCI.
- Educational background of the patient can affect score.
- Trained personnel is needed to supervise conduct the test and interpret the results.

Motivation

Is early screening and diagnosis beneficial?
Cognitive decline can be slowed down with timely and proper pharmacological treatment, targeted mental stimulation, and management of comorbidities.

Prevalence of MCI :

- Europe: **5.1 to 24.5%**
- US**: 6 to 20%**
- Latin America: **6.8% up to 25.5%**
- Worldwide: **~15.56%**

Can AI help?

Automated screening tool concept (1/2)

- Use a deep learning model to analyze speech patterns and classify subjects based on the audio created reading a reference text.
- Based on a Pre-trained Audio Neural Network (**PANN**).
- The PANN was previously trained on Audioset, consisting of 2,084,320 human-labeled 10-second sound clips from **Youtube**

There are 2,084,320 YouTube videos containing 527 labels

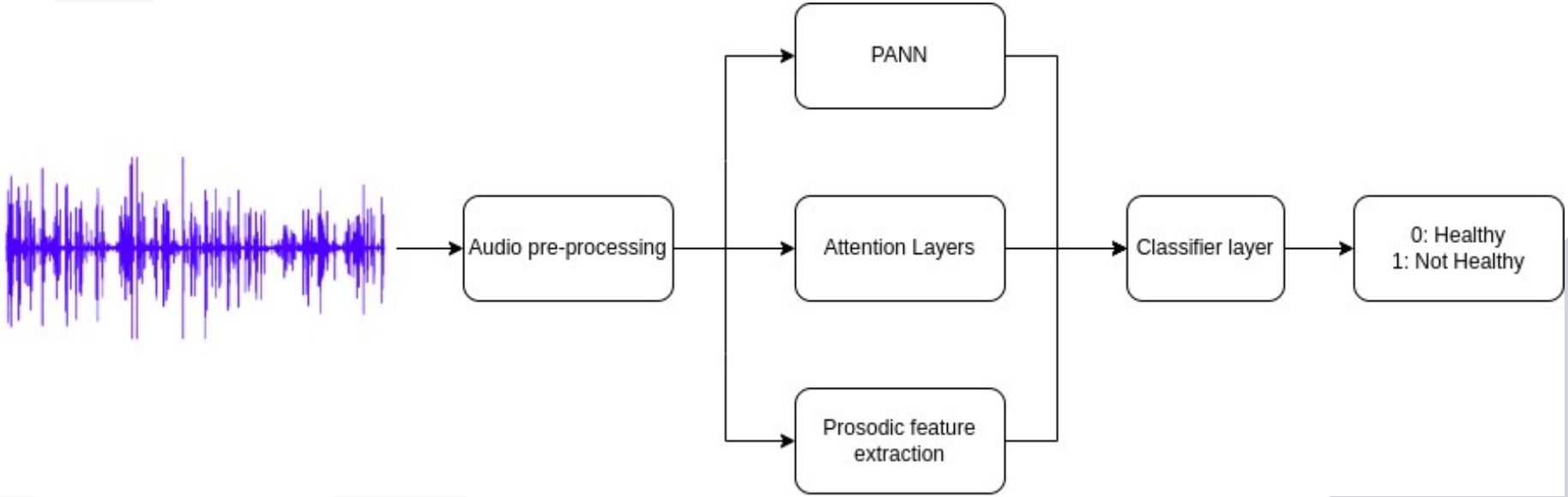
speech ×

Show detailed breakdown? ☐

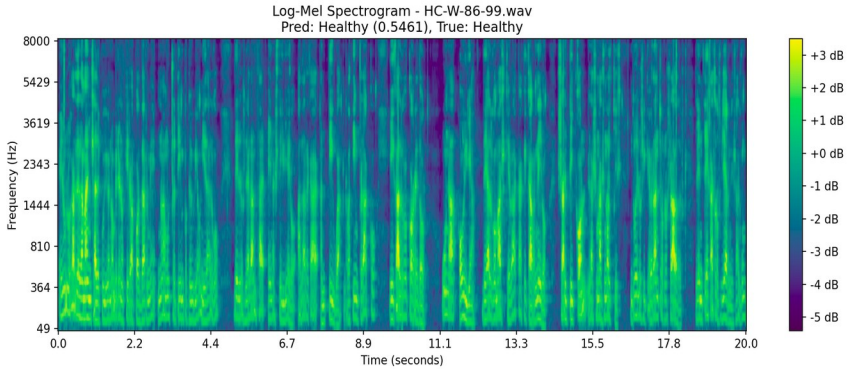
Label	Quality estimate? ▾	Number of videos
Speech	100%	1,010,480
Child speech, kid speaking	100%	11,816
Female speech, woman speaking	100%	8,513
Male speech, man speaking	90%	17,716
Hubbub, speech noise, speech babble	80%	1,480
Speech synthesizer	20%	1,713

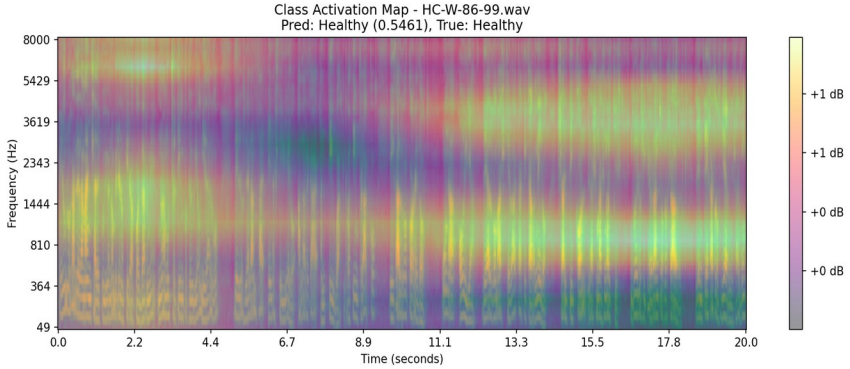
Automated screening tool concept (2/2)

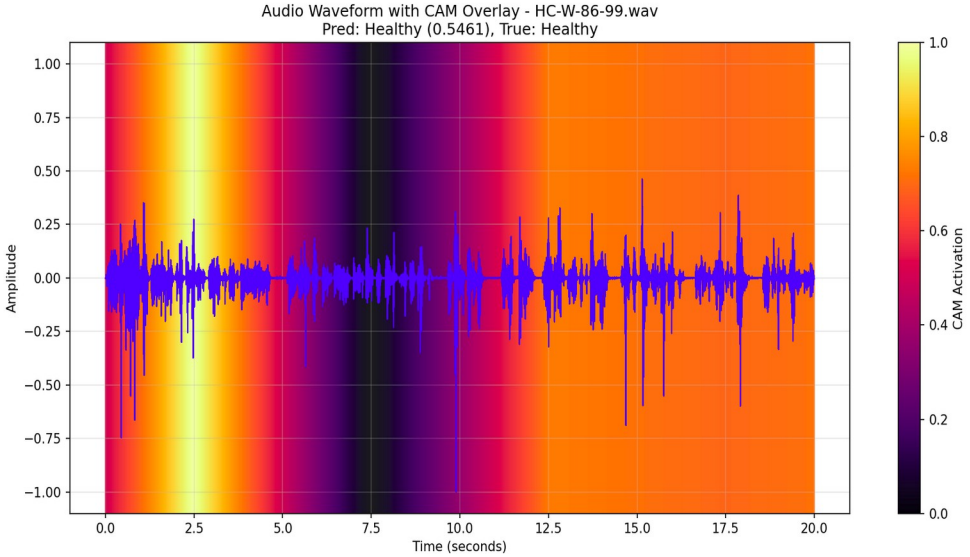
- Use the pre-trained PANN to extract spectral characteristics from visual representations of the audio (Spectrograms).
- Use an Attention path that focus on discovering and capturing time-dependent features (ie. changes in volume, sequences, repetitions).
- Use prosodic features extracted from the speech itself to guide the classification (phonation time, number of pauses, length of pauses).
- Classify an audio between as Healthy(HC) or Not Healthy (MCI or AD)



Examples of processed audio







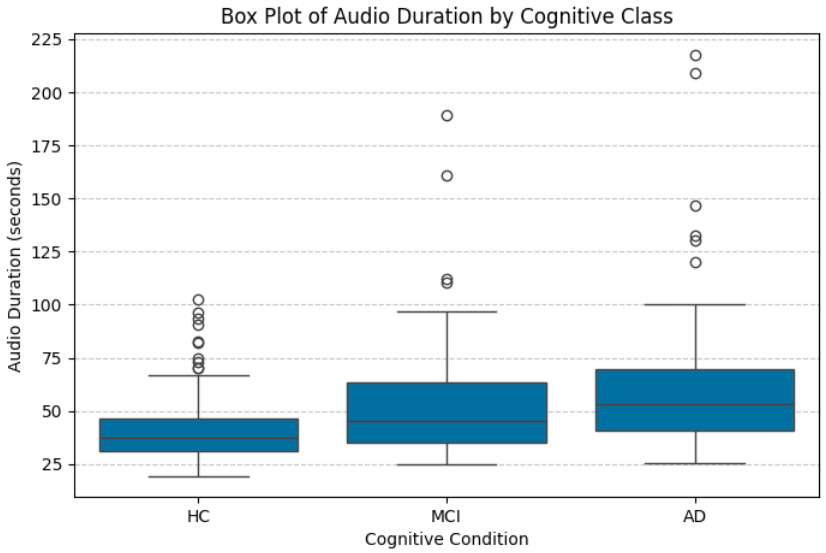
Data use for training

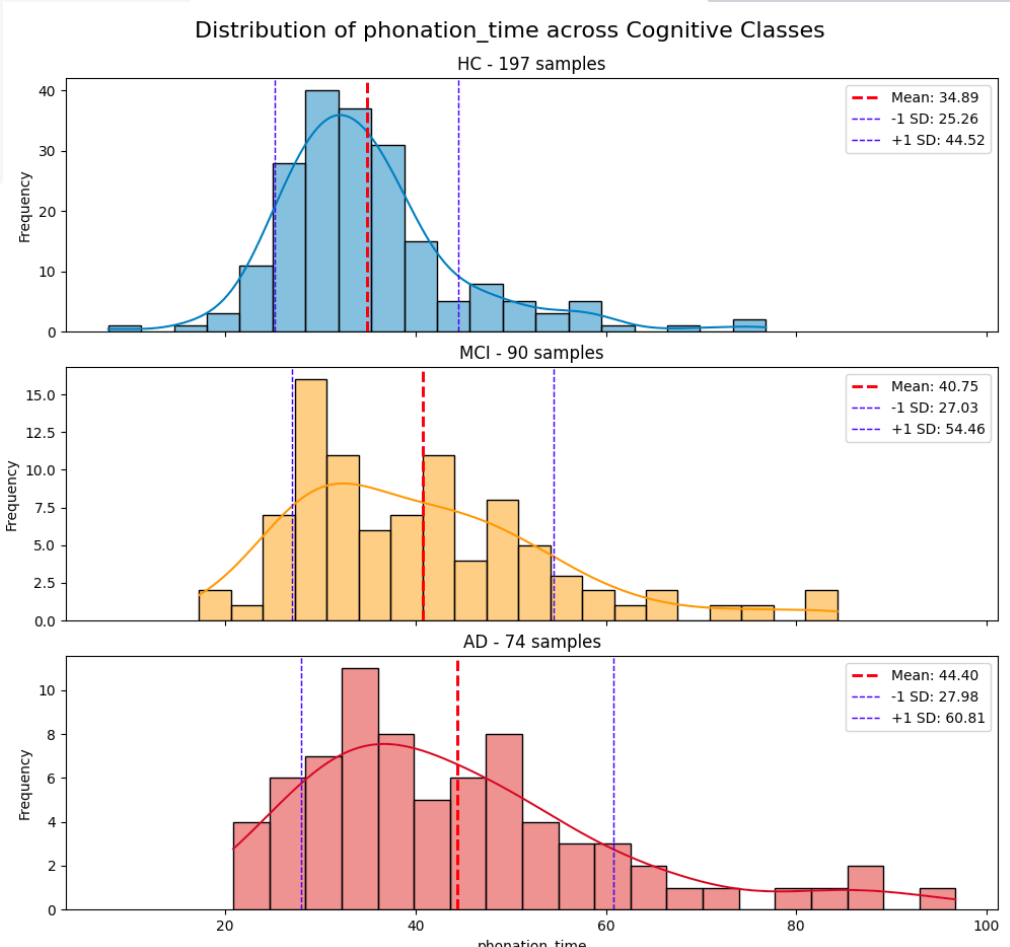
- Dataset: **361** participants:
  - **197 Healthy control (HC)**
  - **90 Minor Cognitive Impairment (MCI)**
  - **74 Alzheimer’s Disease (AD).**
- Fixed reading task (“Don Quixote” passage).
- Controlled variables: age, *education*, no drug/alcohol influence, *MMSE score*, *recording conditions*.
- The only curated corpus for this purpose in Spanish.
- **Discriminating speech traits of Alzheimer’s disease assessed through a corpus of reading task for Spanish language.** Ivanova et all.

Challenges

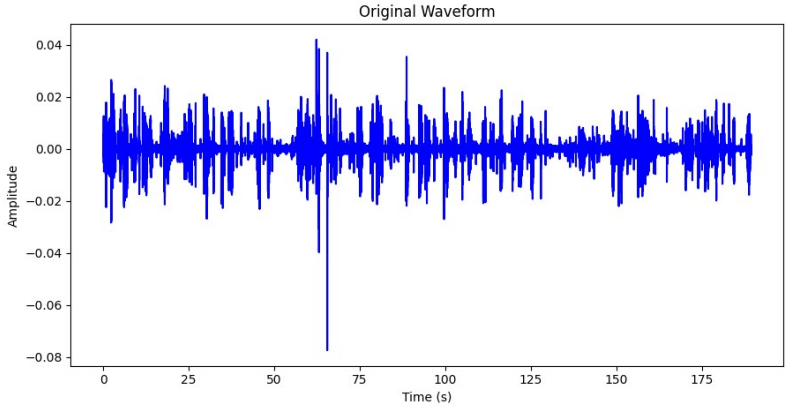
Problem	Solution
Size of the dataset.	Data augmentation via SpecAugment
Class imbalance of the dataset.	Selective data pruning based on cosine similarity of the samples (inter and intra-class). Changes to the loss function (Focal Loss or label smoothing), class weights, threshold optimization.
Quality of the dataset	Noise filtering, RMS normalization, Audio/noise separation (Demucs) & voice activity detection (pyAnnote)

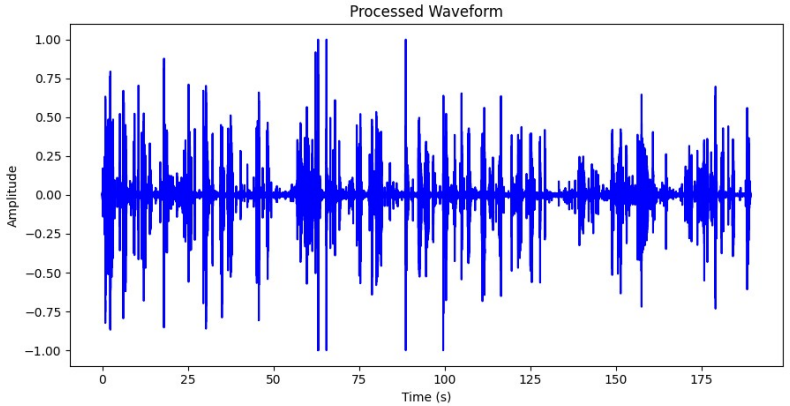
Class distribution of a sample of features

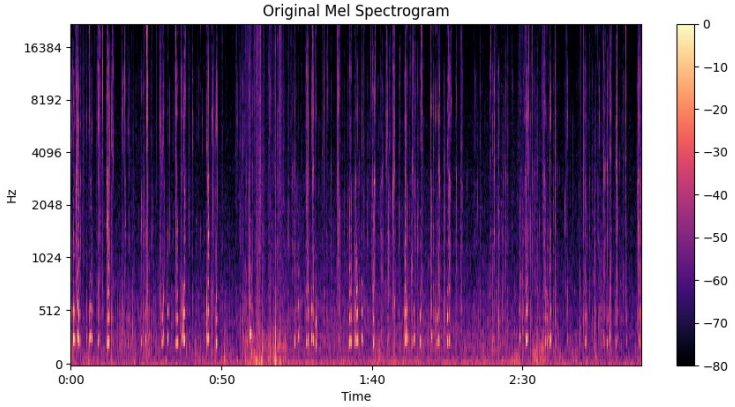


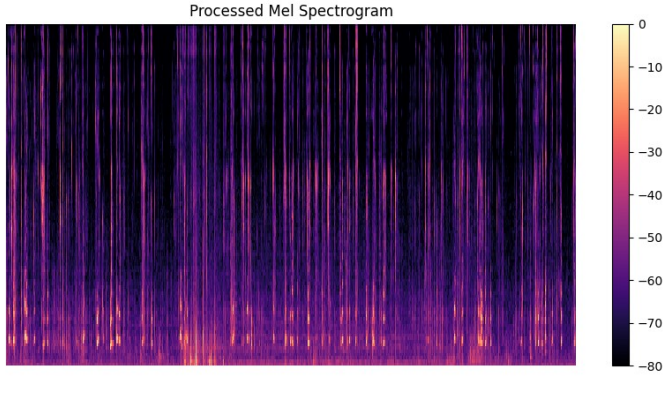


Effect of noise reduction and RMS normalization



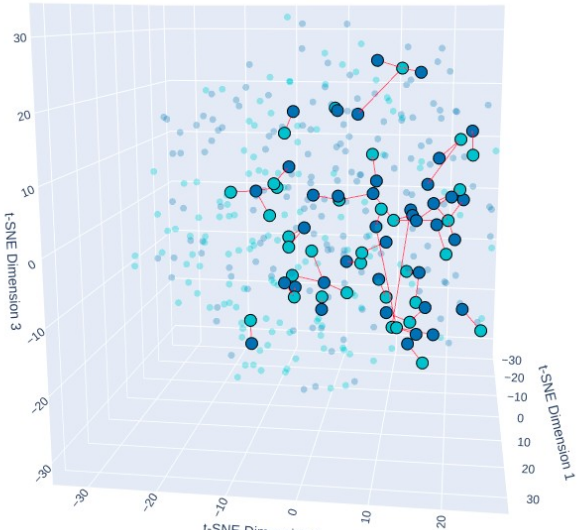




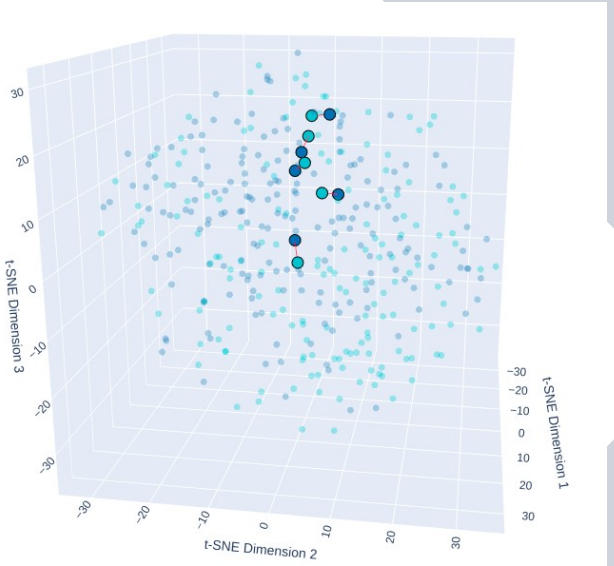


Change in sample discrimination based on different architectures

DenseNet (trained on natural images)



PAAN (trained on Youtube audio)



PAAN after fine tuning

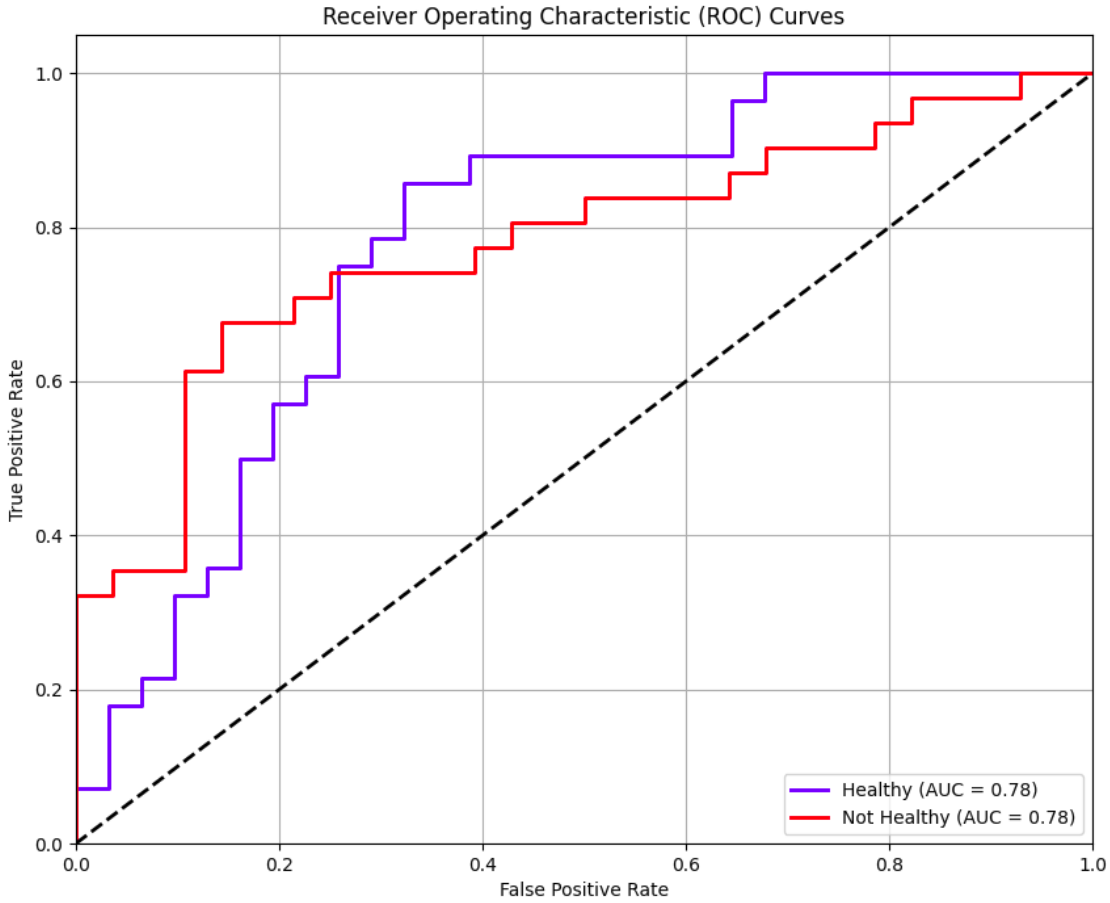
Performance evaluation

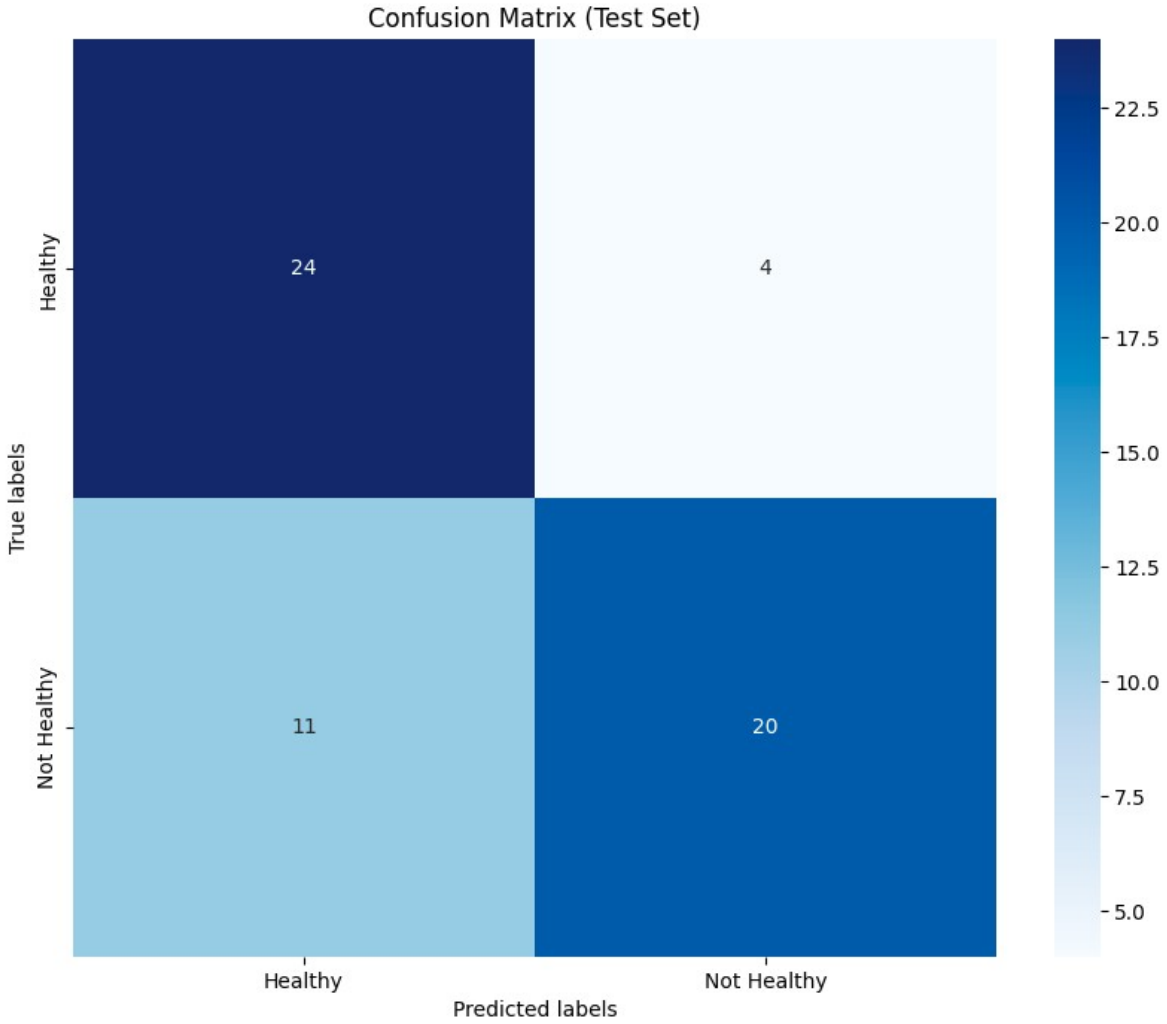
ROC curve: **How well** a test or model can tell the **difference between two groups** by comparing correct hits to false alarms.

AUC: A single number that sums up how good a test or model is at separating two groups; the higher it is (closer to 1), the better

MMSE AUC (for MCI): **0.79**

This model's AUC: **0.78**




Confusion Matrix (Test Set)
Healthy
Not Healthy
True labels
24
4
11
20
Healthy
Not Healthy
Predicted labels
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0

Analysis of misclassified samples

Filename	Predicted	True label	Subgroup	MMSE score	Education [years]
AD-M-67-285.wav	Healthy	Non-Healthy	AD	23	17
AD-M-69-272.wav	Healthy	Non-Healthy	AD	28	12
AD-M-81-280.wav	Healthy	Non-Healthy	AD	18	8
AD-W-78-223.wav	Healthy	Non-Healthy	AD	20	8
HC-M-81-175.wav	Non-Healthy	Healthy	--	28	12
HC-W-69-33.wav	Non-Healthy	Healthy	--	28	14
HC-W-78-217.wav	Non-Healthy	Healthy	--	30	6
HC-W-87-261.wav	Non-Healthy	Healthy	--	30	8
MCI-M-70-102.wav	Healthy	Non-Healthy	MCI	29	12
MCI-M-83-307.wav	Healthy	Non-Healthy	MCI	26	16
MCI-M-90-81.wav	Healthy	Non-Healthy	MCI	28	8
MCI-W-78-240.wav	Healthy	Non-Healthy	MCI	21	6
MCI-W-81-165.wav	Healthy	Non-Healthy	MCI	21	12
MCI-W-82-94.wav	Healthy	Non-Healthy	MCI	29	6
MCI-W-90-251.wav	Healthy	Non-Healthy	MCI	28	8

An MMSE score equal or higher to **24** is considered healthy

Outlook

- Fine-tuning the PAAN for this specific tasks proves the adaptability of the base architecture.
- Multilingual support is guaranteed by the strong training foundation of the PAAN.
- The boundary between HC and mild MCI subjects proved also to be a challenge for the model, however the model’s overall performance could improve with more and better data.
- Improving the AUC to match or exceed the MMSE’s performance on MCI is achievable.
- The development of a mobile or web-based tool based on the trained model is still pending.