Speech-to-Text Interface to MammoClass

2nd Breast Cancer Workshop 2015 – April 7th 2015 Porto, Portugal

Ricardo Sousa Rocha

Inês Dutra

CRACS INESCTEC

Outline

- MammoClass
- Development of Speech to Text Interface to MammoClass
- Web Speech API applied to Mammoclass
- Conclusions and Future Work

CRACS INESCTEC

Outline

- MammoClass
- Development of Speech to Text Interface to MammoClass
- Web Speech API applied to Mammoclass
- Conclusions and Future Work

CRACS INESCTEC

MammoClass

Classification of a mammogram based in a reduced set of mammography findings

http://cracs.fc.up.pt/~nf/mammoclass/

CRACS INESCTEC

How is it done?

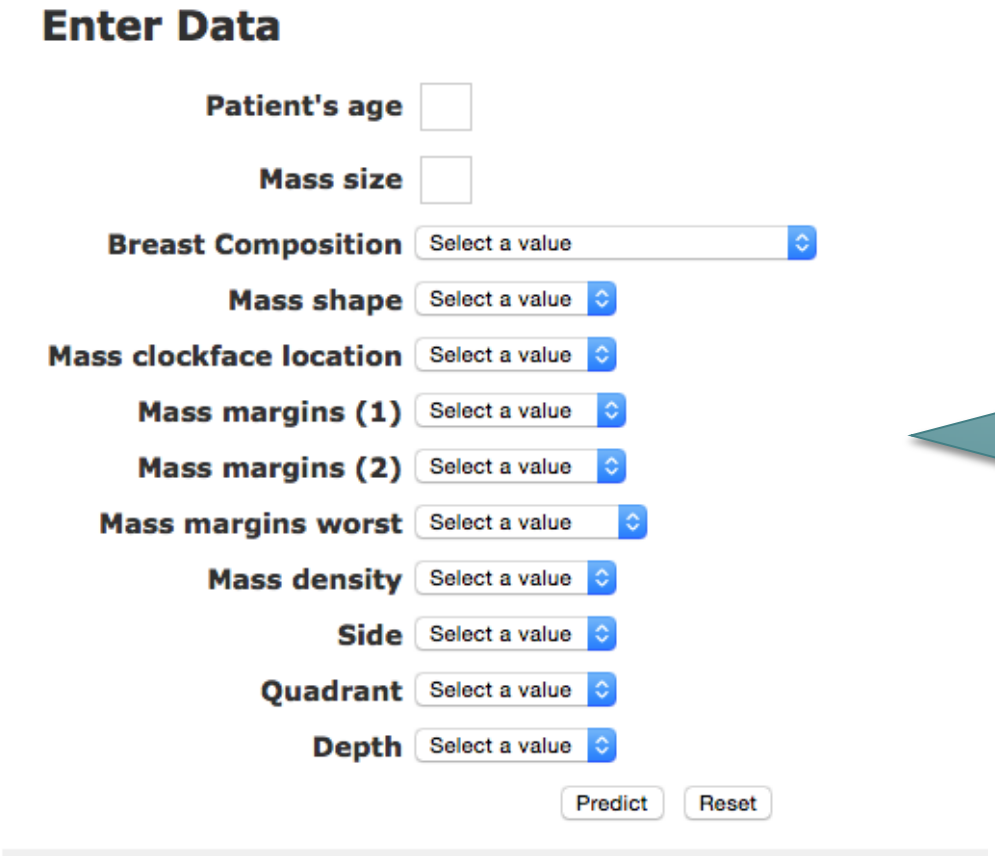
- To obtain a prediction in terms of malignancy for a certain mass is only necessary to provide the values of the findings through forms.
- The output will indicate the probability of a certain mass being benign or malignant. In the latter case it is suggested that the patient should perform a biopsy. The probabilities are computed using machine learning models built as described in:

P.Ferreira, N. A. Fonseca, I. Dutra, R. Woods, and E. Burnside, ***Predicting Malignancy from Mammography Findings and Surgical Biopsies****, submitted.*

http://cracs.fc.up.pt/~nf/mammoclass/

CRACS INESCTEC

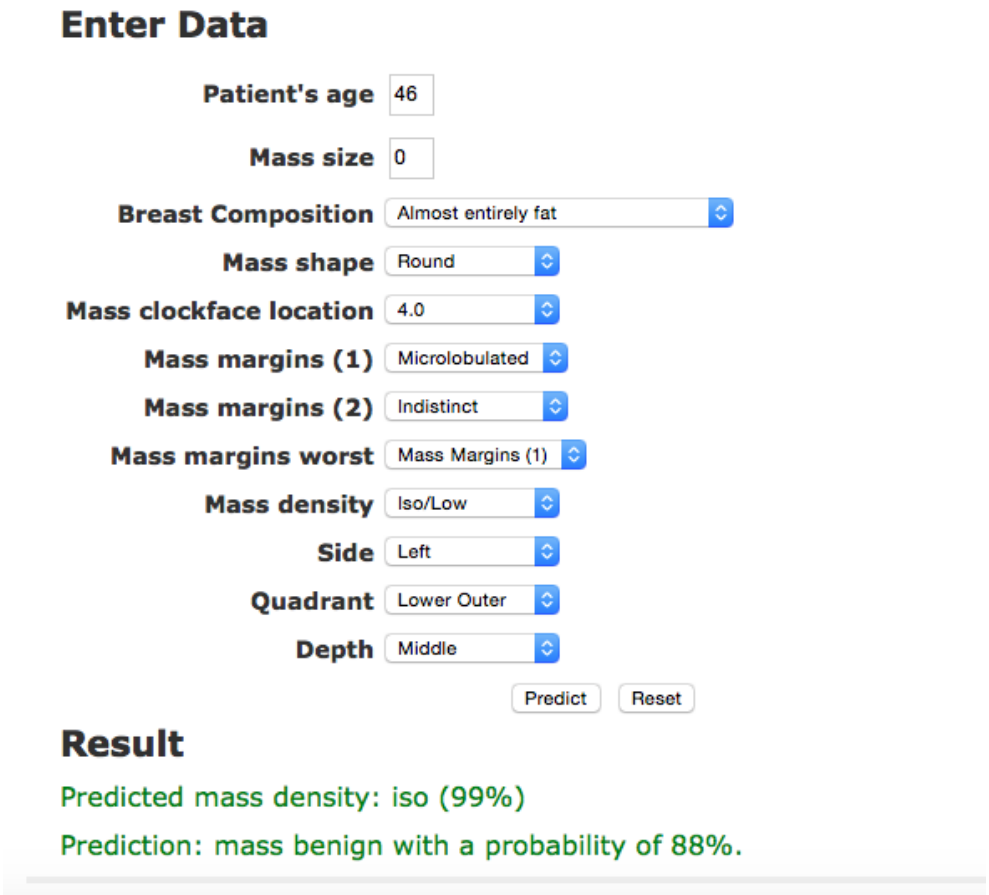
Forms to enter the findings



Empty Forms

CRACS INESCTEC

Forms to enter the findings and Results



Results provided to fill out the forms with some data

CRACS INESCTEC

Outline

- MammoClass
- Development of Speech to Text Interface to MammoClass
- Web Speech API applied to Mammoclass
- Conclusions and Future Work

CRACS INESCTEC

Development of Speech to Text Interface to MammoClass

CRACS INESCTEC

What is Speech to Text

- Speech-to-text software is a type of software that effectively takes audio content and transcribes it into written words in a word processor or other display destination. This type of speech recognition software is extremely valuable to anyone who needs to generate a lot of written content without a lot of manual typing. It is also useful for people with disabilities that make it difficult for them to use a keyboard.
- Speech-to-text software may also be known as voice recognition software.

http://www.techopedia.com/definition/23767/speech-to-text-software

CRACS INESCTEC

Tested Tools

- Free Voice to Text (1) - Can be used to send emails and documents just dictating . It supports the following languages: English, Spanish, French and Japanese.
- Talking Desktop (2) - In addition to making text recognition, it has functions to dictate times and meteorological warnings. Seems to present problems of a few controls and slow reaction time . It supports English, Spanish, French and German
- Dragon Naturally Speaking Home (Premium) (3) - Through research seems quite accurate, and works very well. However only supports the English language.

(1)http://download.cnet.com/Free-Voice-to-Text/3000-7239_4-76115951.html
(2) http://voice-recognition-software-review.toptenreviews.com/talkingdesktop-review.html
(3)http://www.nuance.com/for-business/by-product/dragon/product-resources/edition-comparison/index.htm

CRACS INESCTEC

Tested Tools

- Freesr Speech Recognition (4) - has the ability to assign a number to each of the windows and dictate to each of them. Only supports English language .
- Simon (5) - Open source software available for windows and linux but only in English language
- Web Speech API (6) - Google API that allows the programmer to obtain a translation of voice to text, has the advantage of the Portuguese language, as well as many others.
- Voice Note (7) - Extension for google chrome, it support the Portuguese language, as well as many others.

(4) http://freesr.org
(5) https://simon.kde.org
(6) https://dvcs.w3.org/hg/speech-api/raw-file/tip/speechapi.html
(7) https://voicenote.in

CRACS INESCTEC

Table of comparison

Software	Free	Price	Languages	Platform
Free Voice to Text	Yes	0$	English, Spanish, French and Japanese	Windows
Talking Desktop	No	47$	English, Spanish, French and German	Windows
Dragon Naturally Speaking Home	No	199$	English	Windows
Freesr Speech Recognition	Trial	NA	English	Windows
Simon	Yes	0$	English	Linux, Windows
Web Speech API	Yes	0$	Portuguese and many more	All
Voice Note	Yes	0$	Portuguese and many more	All

CRACS INESCTEC

What tool to choose?

Our idea is that the tool should:

- Be Free
- Support the Portuguese language

CRACS INESCTEC

Candidates tools

Web Speech API

VS

VoiceNote

CRACS INESCTEC

Web Speech API Vs VoiceNote

Relatorio: A pele e o tecido celular subcutâneo apresentam aspectos mamográficos normais.

WS API: a pele e o tecido celular subcutâneo apresentam aspectos demográficos normais

Voice Note: a pele e do tecido celular subcutâneo apresento aspectos demográficos normais.

CRACS INESCTEC

Web Speech API Vs VoiceNote

Relatório: Não se individualizam imagens nodulares que sugiram malignidade, micro-calcificaçoes suspeitas ou outras alteraçoes significativas, em qualquer dos lados.

WS API: não consigo visualizar imagens nodulares que sugiro malignidade microcalcificaçoes suspeitas outras alteraçoes significativas em qualquer dos lados

Voice Note: Não consigo visualizar imagens no solares que sugiro malignidade microcalcificaçoes suspeitas outras altraçoes significativas em qualquer um dos lados.

CRACS INESCTEC

Web Speech API Vs VoiceNote

Relatório: No actual estudo, observamos padrão mamográfico de densidades fibroglandulares dispersas, pela pequena quantidade de parênquima mamário.

WS API: no atual estudo observamos pedro mamográfico de densidades fibroglandular dispersas pela pequena quantidade de parênquima mamário.

Voice Note: No actual estudo observamos pedro monográfico de densidades fibroglandular dispersas pela pequena quantidade parênquima mamário.

CRACS INESCTEC

Results

The results are very similar, which leads me to believe that the VoiceNote was built using the WebSpeech API.

The chosen tool to use was **Web Speech API**.

Because:

- allows greater freedom since it is an API
- can be integrated easy way in any element of a web page

CRACS INESCTEC

Terms BI-RADS tested with Web Speech API

86 Terms

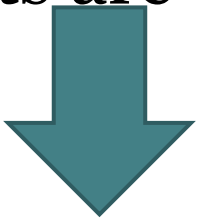
Number of hits	Percentage of hits	Number of wrong	Percentage of wrong
63	73,26%	23	26,74%

Tests done with my voice

CRACS INESCTEC

Things to consider

- Results may not be reliable due to be carried out only with my voice
- Results may vary since the API does not make any voice learning, unlike paid tools
- Some of the results are wrong only on the word genre



Possible future solution

Test the API and find patterns that can be corrected from the obtained text.

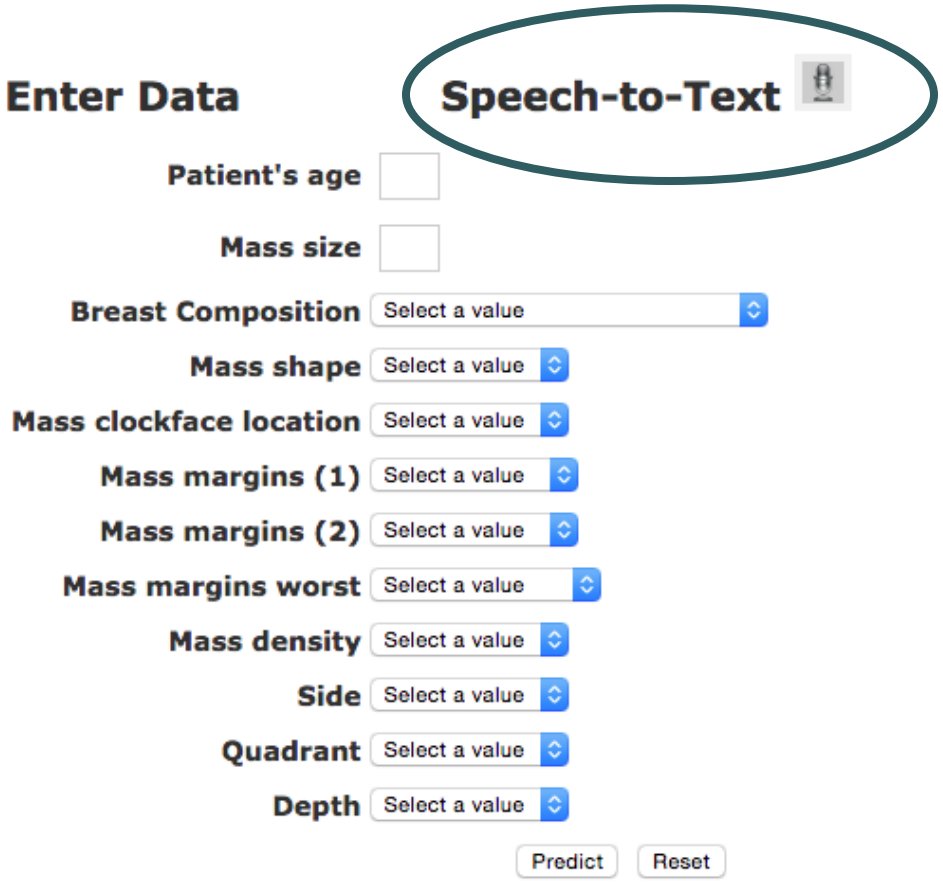
CRACS INESCTEC

Outline

- MammoClass
- Development of Speech to Text Interface to MammoClass
- Web Speech API applied to Mammoclass
- Conclusions and Future Work

CRACS INESCTEC

Web Speech API applied to Mammoclass - Menu



CRACS INESCTEC

Web Speech API applied to Mammoclass – Recording Interface

The output will indicate the probability of a certain mass being benign or malignant. In the latter case it is suggested that the patient should perform a biopsy. The probabilities are computed using machine learning models built as described in:

- *P.Ferreira, N. A. Fonseca, I. Dutra, R. Woods, and E. Burnside,* ***Predicting Malignancy from Mammography Findings and Surgical Biopsies****, submitted.*

Início do reconhecimento

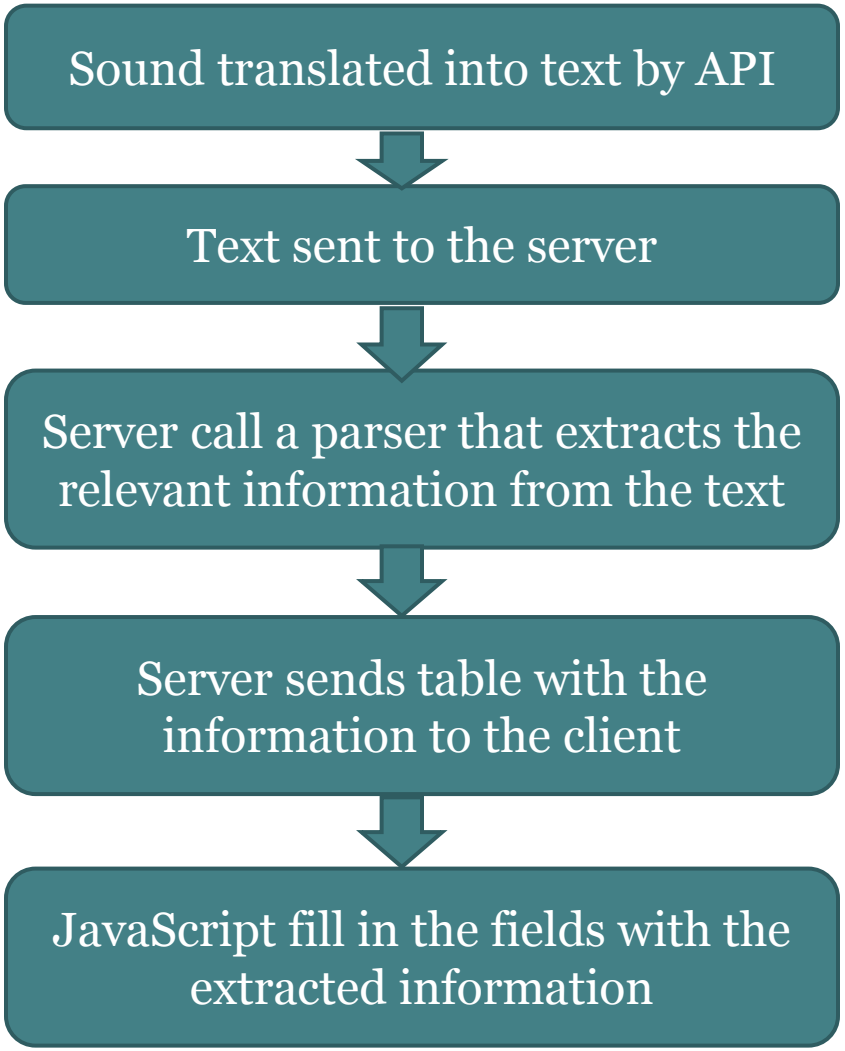
Terminar reconhecimento

CRACS INESCTEC

Web Speech API applied to Mammoclass – Permission

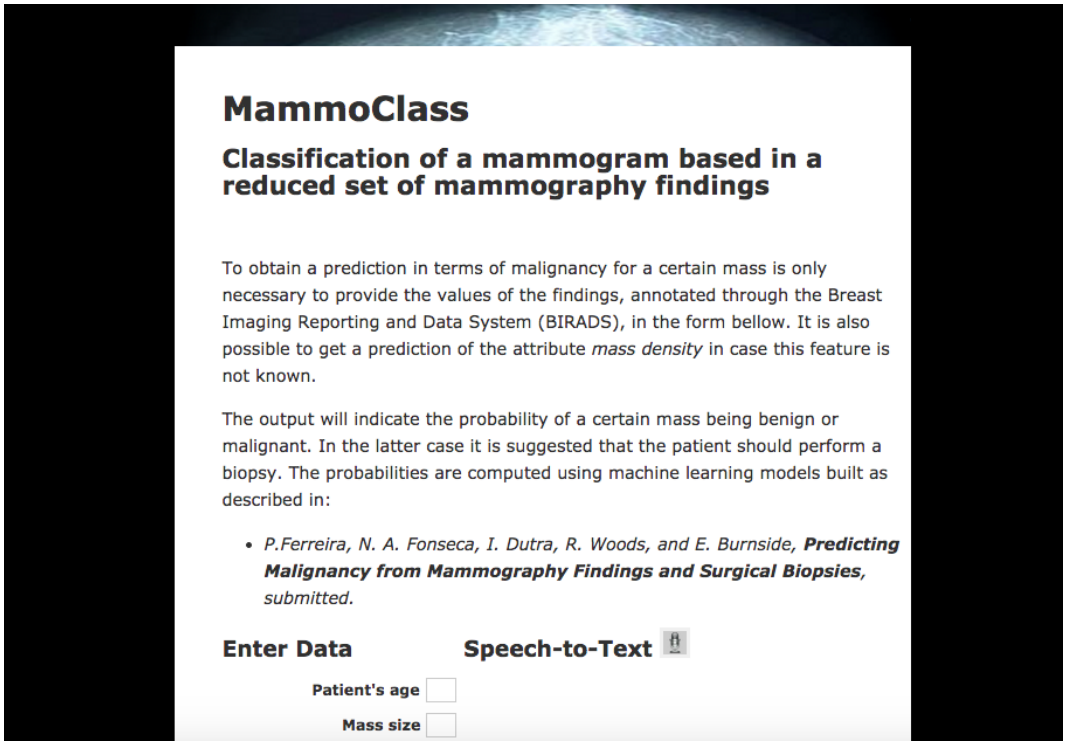
CRACS INESCTEC

Flow chart



CRACS INESCTEC

MammoClass - Speech to Text Interface



- Available at:
  - http://www.alunos.dcc.fc.up.pt/~up201003917/mcwstt/index.html

CRACS INESCTEC

Outline

- MammoClass
- Development of Speech to Text Interface to MammoClass
- Web Speech API applied to Mammoclass
- Conclusions and Future Work

CRACS INESCTEC

Conclusions and Future Work

1) Several Speech to text tools studied.
2) Of all the available we selected two that met the requisites proposed
3) Tests and comparisons were made between these two tools in order to choose the one that best results presented
4) Implementation of speech to text interface, and all the core to handle the API and can send the results to the server

CRACS INESCTEC

Conclusions and Future Work

1. Doing the tests with the BI-RADS terms with other voices beyond mine

2. Find error patterns that can be corrected before sending the sentence to the parser.

Thank you!

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA