Speech-to-Text Interface to MammoClass

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

Ricardo Sousa Rocha
Inês Dutra
Outline

• MammoClass

• Development of Speech to Text Interface to MammoClass

• Web Speech API applied to Mammoclass

• Conclusions and Future Work
Outline

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

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

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


http://cracs.fc.up.pt/~nf/mammoclass/
Forms to enter the findings

**Enter Data**

- Patient's age
- Mass size
- Breast Composition
- Mass shape
- Mass clockface location
- Mass margins (1)
- Mass margins (2)
- Mass margins worst
- Mass density
- Side
- Quadrant
- Depth

[Select a value]  
[Select a value]  
[Select a value]  
[Select a value]  
[Select a value]  
[Select a value]  
[Select a value]  
[Select a value]  
[Select a value]  
[Select a value]  

Predict  Reset
Forms to enter the findings and Results

Enter Data

- Patient's age: 46
- Mass size: 0
- Breast Composition: Almost entirely fat
- Mass shape: Round
- Mass clockface location: 4.0
- Mass margins (1): Microlobulated
- Mass margins (2): Indistinct
- Mass margins worst: Mass Margins (1)
- Mass density: iso/Low
- Side: Left
- Quadrant: Lower Outer
- Depth: Middle

Result

Predicted mass density: iso (99%)
Prediction: mass benign with a probability of 88%. 

Results provided to fill out the forms with some data.
Outline

• MammoClass

• Development of Speech to Text Interface to MammoClass

• Web Speech API applied to MammoClass

• Conclusions and Future Work
Development of Speech to Text Interface to MammoClass
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
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
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
# Table of comparison

<table>
<thead>
<tr>
<th>Software</th>
<th>Free</th>
<th>Price</th>
<th>Languages</th>
<th>Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free Voice to Text</td>
<td>Yes</td>
<td>0$</td>
<td>English, Spanish, French and Japanese</td>
<td>Windows</td>
</tr>
<tr>
<td>Talking Desktop</td>
<td>No</td>
<td>47$</td>
<td>English, Spanish, French and German</td>
<td>Windows</td>
</tr>
<tr>
<td>Dragon Naturally Speaking Home</td>
<td>No</td>
<td>199$</td>
<td>English</td>
<td>Windows</td>
</tr>
<tr>
<td>Freesr Speech Recognition</td>
<td>Trial</td>
<td>NA</td>
<td>English</td>
<td>Windows</td>
</tr>
<tr>
<td>Simon</td>
<td>Yes</td>
<td>0$</td>
<td>English</td>
<td>Linux, Windows</td>
</tr>
<tr>
<td>Web Speech API</td>
<td>Yes</td>
<td>0$</td>
<td>Portuguese and many more</td>
<td>All</td>
</tr>
<tr>
<td>Voice Note</td>
<td>Yes</td>
<td>0$</td>
<td>Portuguese and many more</td>
<td>All</td>
</tr>
</tbody>
</table>
What tool to choose?

Our idea is that the tool should:

• Be Free
• Support the Portuguese language
Candidates tools

Web Speech API VS VoiceNote
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 apreseneto aspectos demográficos normais.
Web Speech API Vs VoiceNote

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

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

Voice Note: Não consigo visualizar imagens no solares que sugiro malignidade microcalcificações suspeitas outras alterações significativas em qualquer um dos lados.
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.
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
# Terms BI-RADS tested with Web Speech API

- **86 Terms**

<table>
<thead>
<tr>
<th>Number of hits</th>
<th>Percentage of hits</th>
<th>Number of wrong</th>
<th>Percentage of wrong</th>
</tr>
</thead>
<tbody>
<tr>
<td>63</td>
<td>73.26%</td>
<td>23</td>
<td>26.74%</td>
</tr>
</tbody>
</table>

*Tests done with my voice*
Things to consider

- Results may not be reliable due to being 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.
Outline

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

Enter Data

Patient's age
Mass size
Breast Composition
Mass shape
Mass clockface location
Mass margins (1)
Mass margins (2)
Mass margins worst
Mass density
Side
Quadrant
Depth

Speech-to-Text
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:

Web Speech API applied to Mammoclass - Permission

http://www.alunos.dcc.fc.up.pt/ wants to use your microphone. Learn more

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:


You must enable Google Chrome access the microphone
Sound translated into text by API

Text sent to the server

Server call a parser that extracts the relevant information from the text

Server sends table with the information to the client

JavaScript fill in the fields with the extracted information
MammoClass - Speech to Text Interface

To obtain a prediction in terms of malignancy for a certain mass is only necessary to provide the values of the findings, annotated through the Breast Imaging Reporting and Data System (BIRADS), in the form below. It is also possible to get a prediction of the attribute mass density, in case this feature is not known.

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:


Enter Data

- Patient's age
- Mass size

Speech-to-Text

Available at:
Outline

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