A Speech-to-Text Interface for MammoClass

Ricardo Sousa Rocha, MSc[1]; Pedro Ferreira, MSc[2]; Inês Dutra, PhD[1,2]; Ricardo Correia, PhD[3]; Rogerio Salvini, PhD[4]; Elizabeth Burnside, MD, MPH, MS[5].

[1]DCC-FC, University of Porto, Portugal
[2]CRACS-INESC TEC, Porto, Portugal
[3]CINTESIS and CIDES-FM, University of Porto, Portugal
[4]Institute of Informatics, Federal University of Goiás, Brazil
[5]University of Wisconsin, Madison, USA

Outline

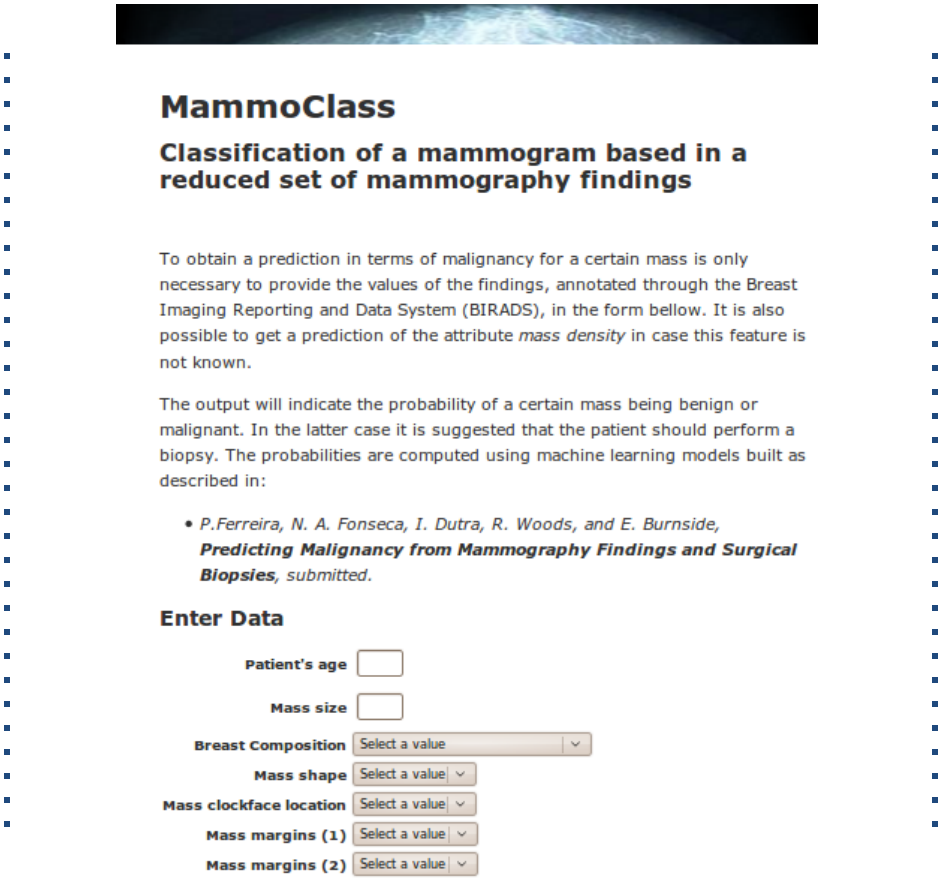
- MammoClass
- Development of STT Interface for MammoClass
- Web Speech API applied to MammoClass
- Conclusions and Future Work

Outline

- MammoClass
- Development of STT Interface for MammoClass
- Web Speech API applied to MammoClass
- Conclusions and Future Work

MammoClass

Classification of a mammogram based in a set of mammography findings



MammoClass

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

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 bellow. 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:

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

Enter Data

Patient's age

Mass size

Breast Composition Select a value

Mass shape Select a value

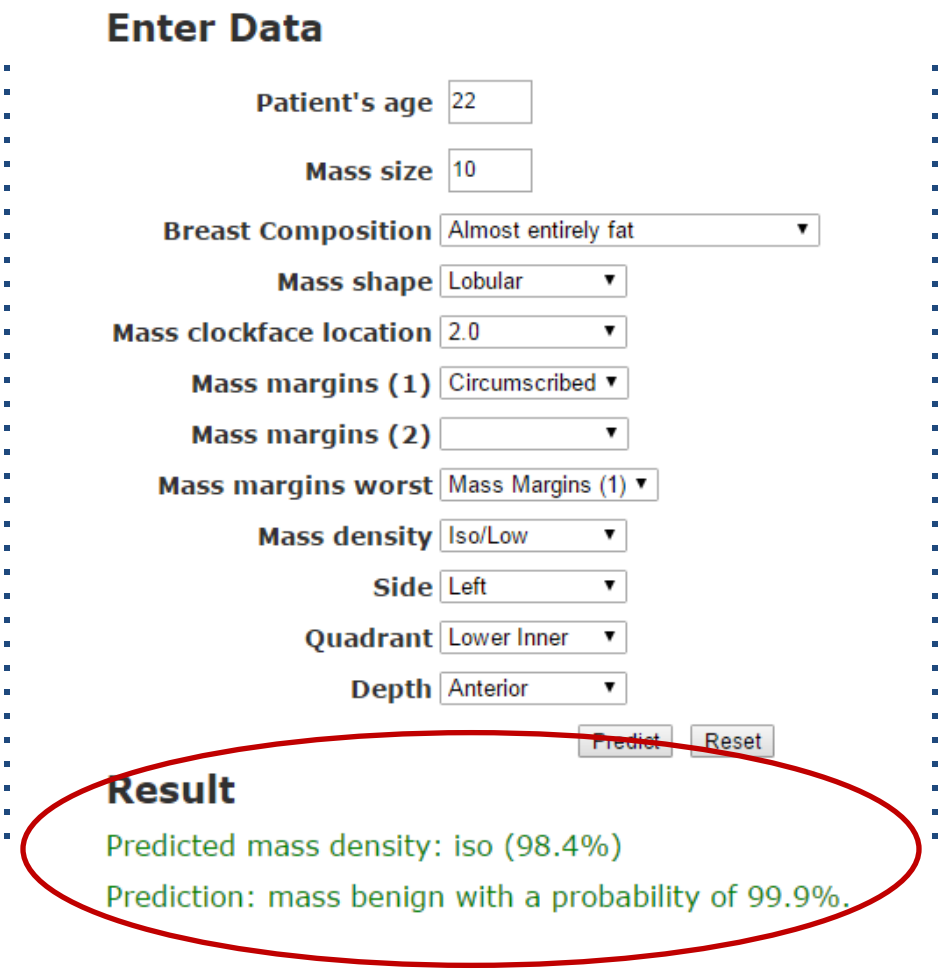
Mass clockface location Select a value

Mass margins (1) Select a value

Mass margins (2) Select a value

MammoClass

Classification of a mammogram based in a set of mammography findings



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:

Ferreira, P., Fonseca, N.A., Dutra, I., Woods, R., Burnside, E.:

Predicting Malignancy from Mammography Findings and Image-Guided Core Biopsies.

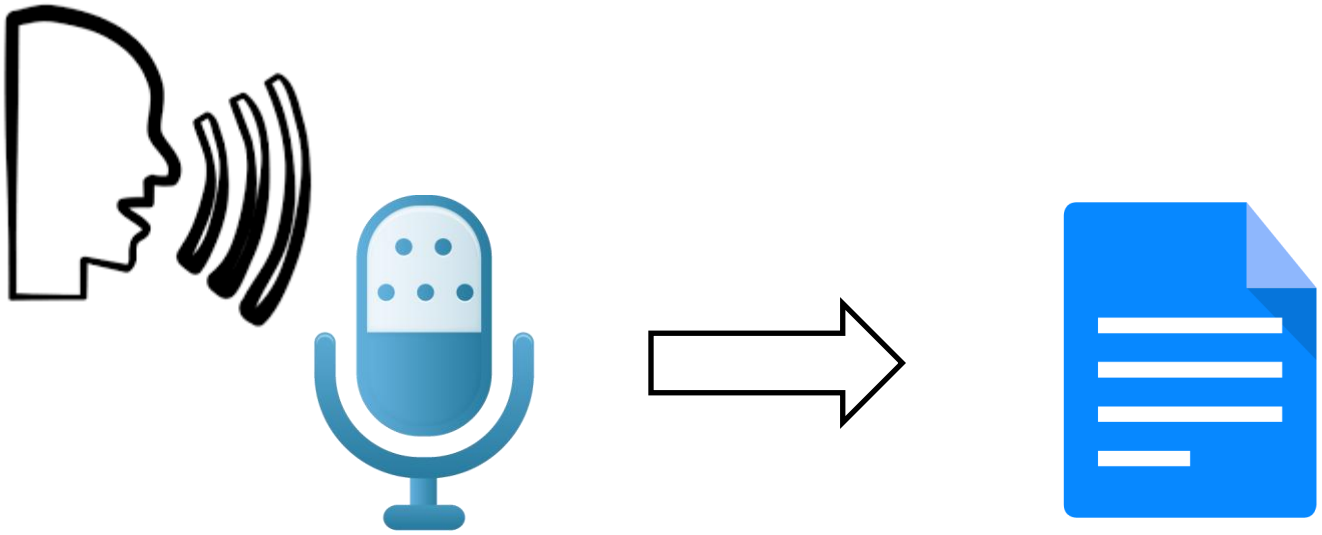
In: Int. Journal of Data Mining and Bioinformatics, 2015.

Outline

- MammoClass
- Development of STT Interface for MammoClass
- Web Speech API applied to MammoClass
- Conclusions and Future Work

What is Speech-to-Text?

Type of software that **takes audio content** and **transcribes it** into written **words** in a word processor or other display destination



What is Speech-to-Text?

- **Advantages:**

- Valuable to anyone who **needs to generate a lot of written content** without a lot of manual typing

- Useful for **people with disabilities** that make it difficult for them to use a keyboard

Speech-to-Text Interface for MammoClass

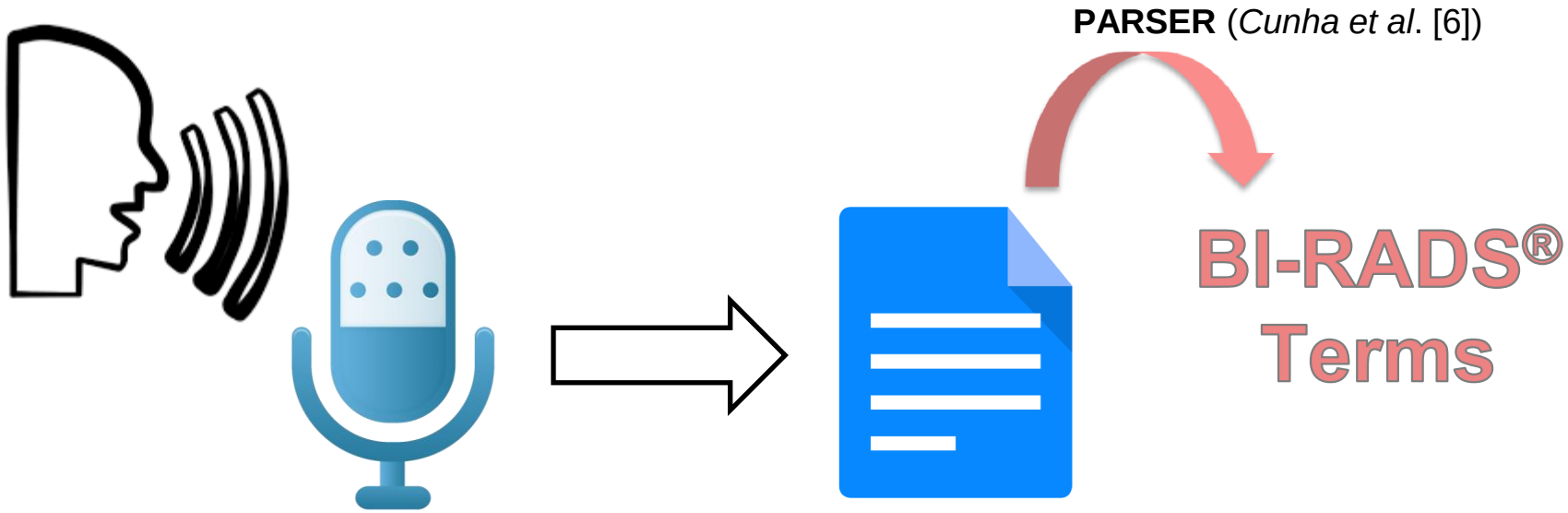
Speech-to-Text Interface for MammoClass

- Some works in the literature do not favor the use of speech recognition technology in the area of radiology and report a high error rate on the resulting recognized texts [1], [2], [3], [4], [5]

- These **works focus** on the **text itself rather than on relevant words** that could be **extracted** from the text to build structered data for posterior automatic studies

Speech-to-Text for MammoClass

What's the innovation?



Looking for the suitable tool…

What tool to choose?

- **Free**
- Support the **Portuguese** language

Tested Tools

- Free Voice to Text

  (+) Can be used to **send emails**, create documents by **just dictating**

  (-) Does **not** support **Portuguese**

Talking Desktop

(+) Makes **text recognition**. Has functions to **recognize dictated text** about **weather** conditions to emit meteorological warnings

(-) Limited controls and **slow reaction time**

(-) Does **not** support **Portuguese**

Dragon Naturally Speaking Home (Premium)

(+) Has a very functional and **user-friendly** interface / Has **specific vocabulary** (eg. medical)

(-) Does **not** support **Portuguese**

Tested Tools (cont.)

Freesr Speech Recognition

(+) Has the ability to **recognize multiple dictated texts**

(-) Does **not** support **Portuguese**

- ## Simon

(+) Open Source software available for Windows and Linux

(-) Does **not** support **Portuguese**

Web Speech API

(+) Google API that allows the programmer to obtain a translation of voice to text

(+) Supports Portuguese!!

VoiceNote

(+) Extension for Google Chrome

(+) Supports Portuguese!!

Tested Tools - Table of Comparison

Software	Free	Price	Languages	Platform
Free Voice to Text	Yes	0$	Eng, Spa, Fre, Jpn	Windows
Talking Desktop		47$	Eng, Spa, Fre, Ger	Windows
Dragon Naturally Speaking Home		199$	Eng	Windows
	 Trial	NA	Eng	Windows
	Yes	0$	Eng	Linux, Windows
	Yes	0$	 + …	All
	Yes	0$	 + …	All

Candidate Tools

Web Speech API

VS

VoiceNote

Candidate Tools

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

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

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

Candidate Tools

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

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

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

Verdict

- Results very **similar**
- We believe VoiceNote was built using the Web Speech API

➢ **<u>Chosen Tool:</u>**

Web Speech API

➢ Allows greater freedom since it is an API

eg. Can be integrated easy way in any element of a Web page

Evaluation WS API

Web Speech API tested with **BI-RADS® Terms**

TABLE II
BI-RADS TERMS AND THEIR TRANSLATION TO PORTUGUESE

mass shape	round oval lobular irregular	arredondada, redonda oval, ovóide, alongada, ovalar lobular, Polilobular irregular
mass margins	circumscribed microlobular obscure indistinct spiculated	circunscrita, bem definida, bem delimitada, regular microlobular obscura, obscurecida indistinta, imprecisa, indefinida, mal definida espiculada
mass density	high equal low	alta, elevada densidade igual, isodensa, homogênea baixa, tênue
breast density	predominantly fatty scattered fibroglandular heterogeneously dense extremely dense	contém gordura fibroglandular heterogênea densa, muito densa, densidade alta da mama

www.alunos.dcc.fc.up.pt/~up201003917/SpeechToText.html

Evaluation WS API

Testing **86** Individual **BI-RADS Terms**

- We classify each returned result as:

- **(C) Correct** – If the original term and the recognized term are **exactly alike**

- **(AC) Almost Correct** – If the original term and the term returned by Web Speech API are **almost identical**

- **(I) Incorrect** - If the original term and the term returned by the API are **completely different**

Evaluation WS API

Testing **86** Individual **BI-RADS Terms**

ambos os lados	ambos os lados	C
distorção arquitectural	distorção arquitectural	C
desorganização arquitectural	desorganização arquitectural	C
ducto dilatado	ducto dilatado	C
carcinoma ductal	carcinoma ductal	C
Gânglio linfático intramamário	Gânglio linfático intramamário	C
Gânglio mamário	Gânglio mamário	C
densidade assimétrica	densidade assimétrica	C
assimetria mamária	assimetria mamária	C
densidade focal	densidade focal	C
foco nodular	foco nodular	C
foco assimétrico	foco assimétrico	C
retracção cutânea	respiração cutânea	I
repuxamento da pele	repuxamento da pele	C
retracção do mamilo	retracção do mamilo	C
inversão mamilar	inversão mamilar	C
espessamento cutâneo	espessamento cutâneo	C
edema	edema	C
eritema	eritema	C
mastite	mastite	C
espessamento trabecular	espessamento trabecular	C
lesão cutânea	lesão cutânea	C
ulceração cutânea	ulceração cutânea	C
adenopatia axilar	adenopatia axilar	C
adenomegalia axilar	adenomegalia axilar	C
cavado axilar	cavado axilar	C
axila positiva	axila positiva	C
Gânglio axilar suspeito	Gânglio axilar suspeito	C

Evaluation WS API

Testing **86** Individual **BI-RADS Terms**

- **Experiments:**

  Performed by **4 people -** two female and two male
  Each of these people used **3 different devices**:

  - **Laptop** with an **external microphone** NGS brand
  - Same **laptop** with **built-in microphone**
  - **Smartphone**

Evaluation WS API

TABLE III
PERFORMANCE PER EXPERIMENT

Device	Person	C(%)	AC(%)	I(%)	C+AC(%)
Laptop with ext micro	A	67.4	8.2	24.4	75.6
	→ B	77.9	5.8	16.3	83.7
	C	68.6	9.3	22.1	77.9
	D	66.3	7.0	26.7	73.3
	Avg.	70.1	7.5	22.4	77.6
Laptop with int micro	A	68.6	9.3	22.1	77.9
	→ B	74.4	4.7	20.9	79.1
	C	67.4	4.7	27.9	72.1
	D	66.3	3.5	30.2	69.8
	Avg.	69.2	5.5	25.3	74.7
Smartphone	A	69.8	8.1	22.1	77.9
	→ B	74.4	7.0	18.6	81.4
	C	70.9	7.0	22.1	77.9
	D	60.5	5.8	33.7	66.3
	Avg.	68.9	7.0	24.1	75.9

Evaluation WS API

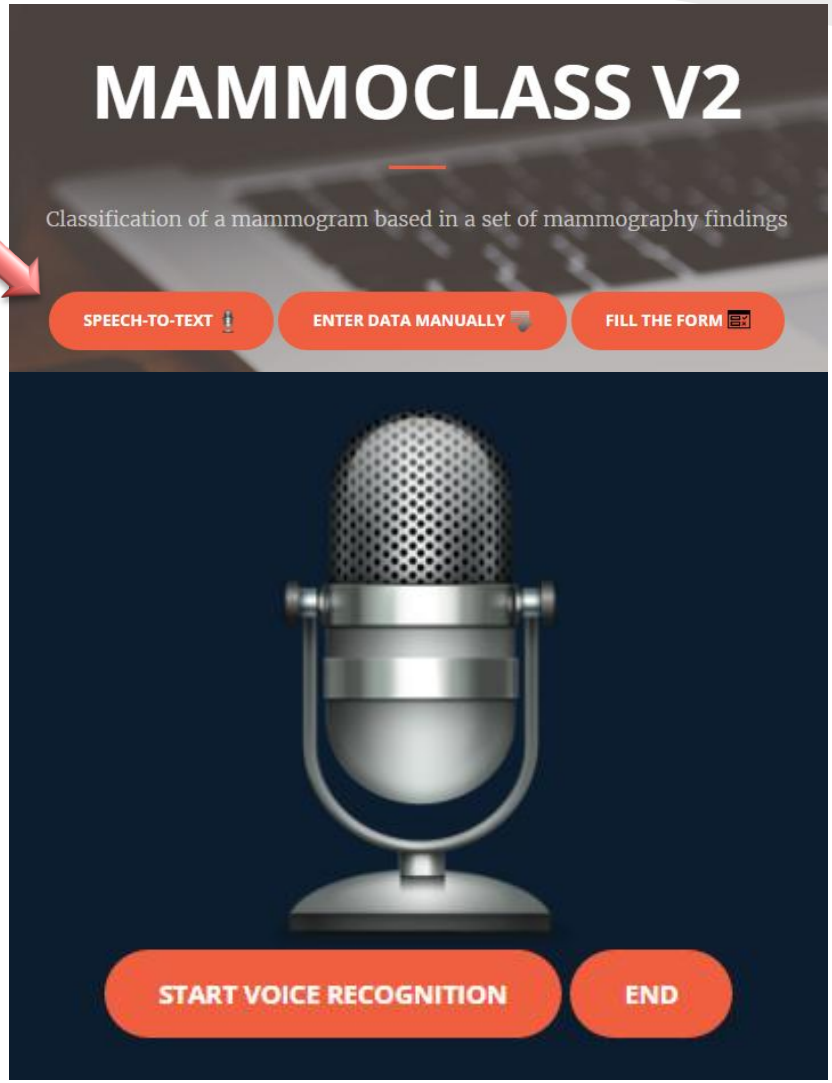
TABLE IV
AVERAGES PER PERSON

Person	C(%)	AC(%)	I(%)	C+AC(%)
A	68.6	8.5	22.9	77.1
B	75.6	5.8	18.6	81.4
C	69.0	7.0	24.0	76.0
D	64.3	5.5	30.2	69.8

Outline

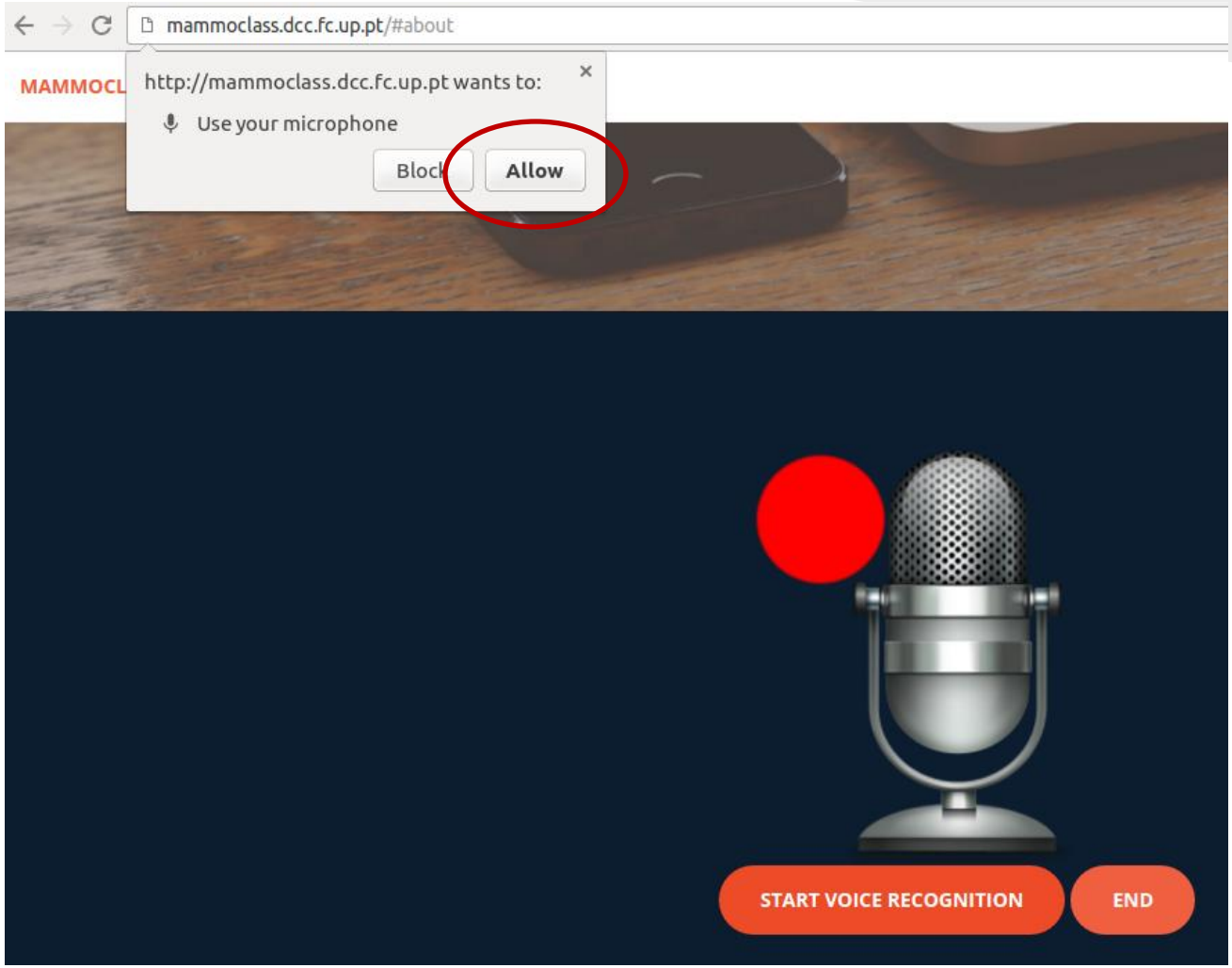
- MammoClass
- Development of STT Interface for MammoClass
- Web Speech API applied to MammoClass
- Conclusions and Future Work

Web Speech API applied to MammoClass



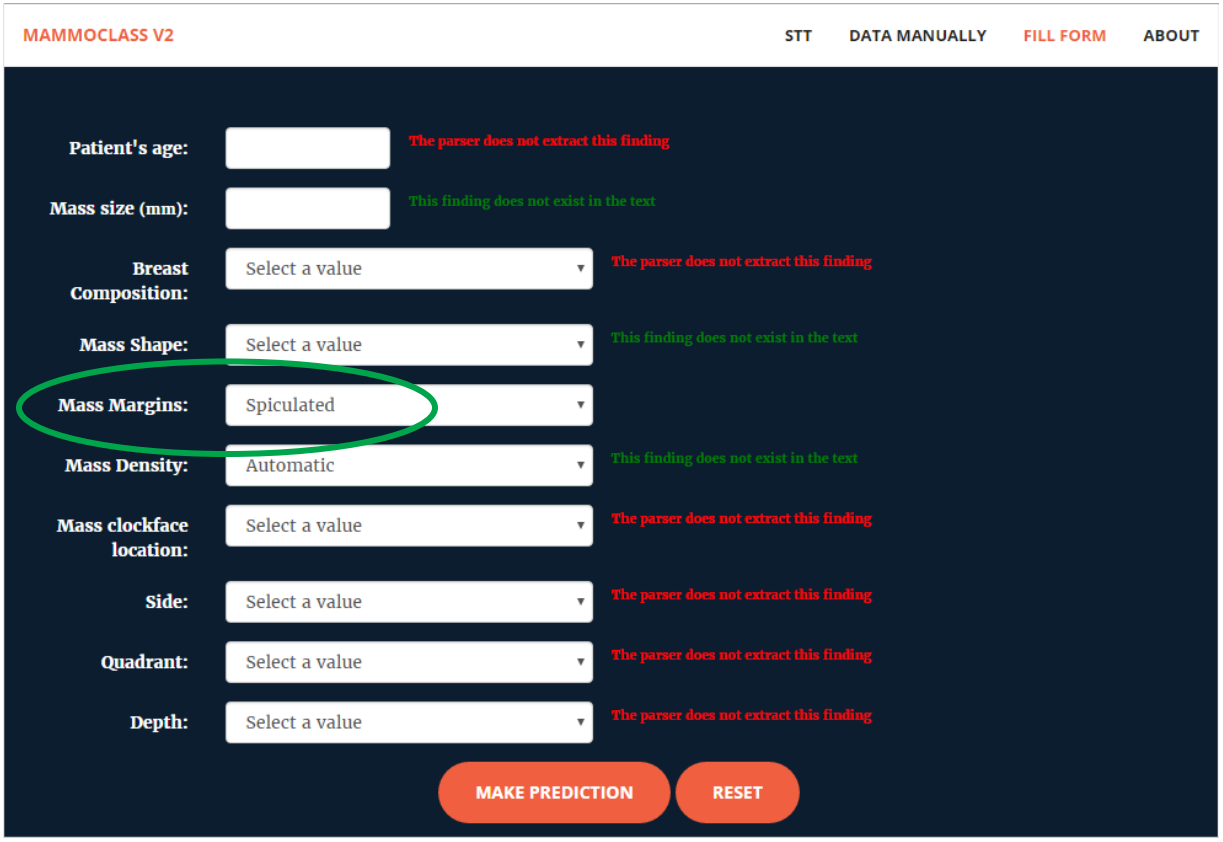
mammoclass.dcc.fc.up.pt

Web Speech API applied to MammoClass

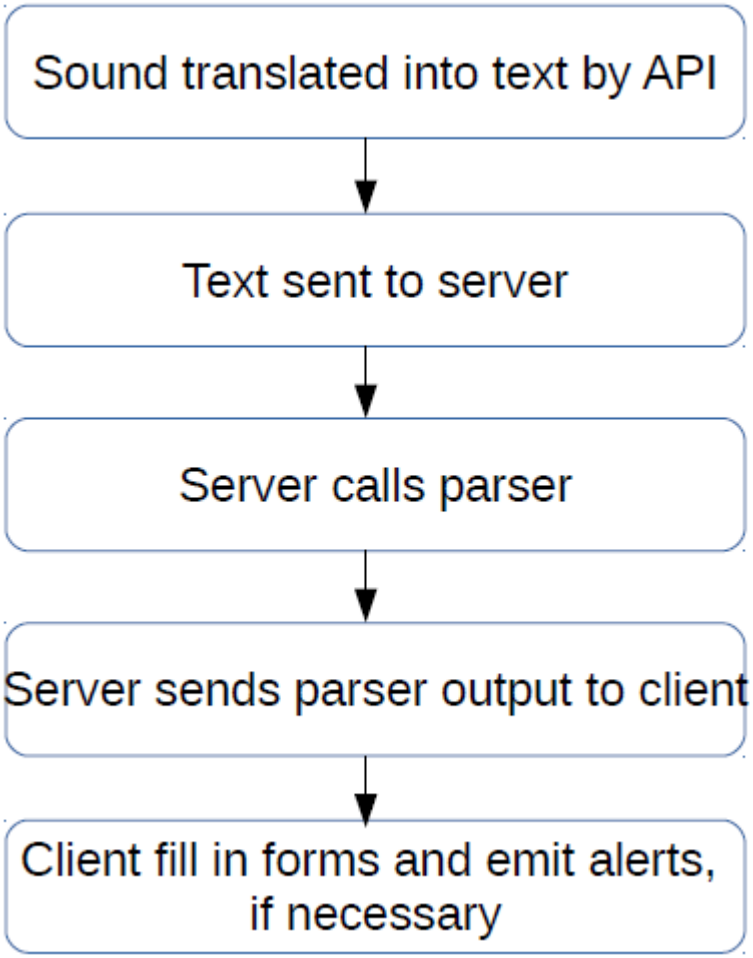


Web Speech API applied to MammoClass

No QSE da mama esquerda observa-se opacidade nodulariforme de contornos **espiculados** e parcialmente obscuros, medindo aproximadamente 2cm, que corresponde no estudo ecográfico complementar a nodulo sólido, com cerca de 1.5cm



Flow Chart – STT API



Conclusions

- We provide to the user an **interface** where **medical reports** can be **dictated** as opposite to input in forms or textboxes

- Although the recognized text sometimes differs from the original written report, the most relevant **BI-RADS terms** are still **recognized**

- **Implementation** of Speech-to-Text interface and all the **core** to handle the **API** and **send** the results to the **server**

Future Work

- Speech interfaces for **long sentences** in Portuguese **need** to be **improved**

- We would like to design and **implement** our **own tools** for recognizing **Portuguese terms**, which could be independent of voice type or entonation, and that could be trained only on the **subset of words** used in the area of **breast cancer**

Thanks

Questions?

Appendix

References

[1] J. du Toit, R. Hattingh, and R. Pitcher, "**The accuracy of radiology speech recognition reports in a multilingual south african teaching hospita**l," BMC Medical Imaging, vol. 15, no. 1, pp. 1–5, 2015. [Online]. Available: http://dx.doi.org/10.1186/s12880-015-0048-1

[2] S. Basma, B. Lord, L. M. Jacks, M. Risk, and S. A. M., "**Error rates in breast imaging reports: comparison of automatic speech recognition and dictation transcription**." AJR Am J Roentgenol, vol. 197, pp. 923–927, 2011.

[3] R. Hoyt and A. Yoshihashi, "**Lessons learned from implementation of voice recognition for documentation in the military electronic health record system**," Perspectives in Health Information Management, no.7(Winter):1e, 2010.

[4] S. McGurk, K. Brauer, T. V. Macfarlane, and K. A. Duncan, "**The effect of voice recognition software on comparative error rates in radiology reports**," The British Journal of Radiology, vol. 81, pp. 767–770, 2008.

[5] I. Hammana, L. Lepanto, T. Poder, and M. S. Bellemare, C. Ly, "**Speech recognition in the radiology department: a systematic review**," HIM J., vol. 44, no. 2, pp. 4–10, 2015.

References (cont.)

[6] Cunha et al. H. Nassif, F. Cunha, I. C. Moreira, R. Cruz-Correia, E. Sousa, D. Page, E. S. Burnside, and I. de Castro Dutra, "**Extracting bi-rads features from portuguese clinical texts**," in IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2012, 2012, pp. 1–4.

State of the Art

[7] Kang et al. Use speech recognition technology in surgical pathology and conclude that it is useful in their anatomic pathology workflow and provides a goo return on investment, error reduction, and cost savings.

Boolean Table

Formação nodular hiperdensa com contornos **espiculados** com cerca de **3 cm** na transição dos quadrantes inferiores da mama direita.
Foi detectada tambem uma margem **lobular**. achados imagiológicos muito sugestivos de malignidade - bi-rads - 5.

0 0 **1** 0 0 0 0 0 **1** 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 **30** 0 0