Can entropy-based image alignment metrics offer improved image aggregation of tissue density for mammographic risk assessment?

Final Report for CS39440 Major Project

Author: Laura Collins (lac32@aber.ac.uk)
Supervisor: Dr. Neil Mac Parthalin (ncm@aber.ac.uk)

4th March 2016
Version: 1.0 (Draft)

This report was submitted as partial fulfilment of a BSc degree in Artificial Intelligence and Robotics (inc Integrated Industrial and Professional Training) (GH7P)

Department of Computer Science
Aberystwyth University
Aberystwyth
Ceredigion
SY23 3DB
Wales, UK

Declaration of originality

In signing below, I confirm that:

- This submission is my own work, except where clearly indicated.
- I understand that there are severe penalties for Unacceptable Academic Practice, which can lead to loss of marks or even the withholding of a degree.
- I have read the regulations on Unacceptable Academic Practice from the University's Academic Quality and Records Office (AQRO) and the relevant sections of the current Student Handbook of the Department of Computer Science.
- In submitting this work I understand and agree to abide by the University's regulations governing these issues.

Name .....

Date .....

Consent to share this work

In signing below, I hereby agree to this dissertation being made available to other students and academic staff of the Aberystwyth Computer Science Department.

Name .....

Date .....

Acknowledgements

I am grateful to...

I’d like to thank...

Abstract

Include an abstract for your project. This should be no more than 300 words.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Background & Objectives

The Project is concerned with the alignment of multiple images using an image-alignment technique called Congealing [4]. The paper describes a method which utilises the Congealing algorithm to align both MNIST handwriting data and MRI Scans by reducing the pixel-wise uncertainty across a collection of images. This project aims to go further, aligning mammography scans and reducing the pixel-wise uncertainty using a number of different fuzzy entropy methods.

1.1 Background

In preparation for my Major Project extensive research was undertaken into the Fuzzy Entropy alignment metrics that would be chosen and implemented in the Congealing algorithm to align mammography scans. S. Al-sharhan's paper 'Fuzzy Entropy: a Brief Survey' [1] was instrumental in a brief comparison between Fuzzy Entropy metrics - allowing a simple way to compare and contrast mathematical differences and implementations.

Researchers have implemented many variations of the Congealing algorithm, with success in varying areas - however very little to no work has been done implementing the algorithm to assess mammography scans. Cox's 'Least Squares Congealing' [2] was quickly disregarded given the project preference to select entropy-based alignment techniques, however did alert to the issue of performance for the project.

The motivation behind choosing mammograms as the input data of choice was an interest into how computer systems and machine learning can help in the medical sector.

1.2 Analysis

1.2.1 Task composition

1.2.1.1 Pixel Membership

From the analysis of the planned Fuzzy Entropy algorithms, one major task to be undertaken would be to calculate the membership of each pixel. Membership stems from Fuzzy set theory **Check this?** which acknowledges that an element can be part of more than one group to a certain

degree, not just one or the other. One common example of this is listing someone as 'Short', 'Average' or 'Tall' in height. If a tall person is someone over 6 feet in height, would a person who measured 5foot 11inches not be classified as tall? In fuzzy set theory, they would be be a certain degree of tall, and a certain degree of average, with the highest membership likely to win out when categorising their height.

There are two common methods to modeling degrees of membership. The first is to manually define the categorty boundaries, so in the case of trapezium functions, the two bases and the two shoulders as in Figure x **include a diagram here showing?**. The other solution would be to iterate over the values you have and to computationally build the an even distribution throughout your membership functions. Whilst this is the preferred method for being dynamic in it's calculations, it is also more computationally expensive as pre-processing of the image would have to be completed before the Congealing algorithm could be run.

Taking this into account, for grey-level pixel values, ranging from 0 (black) to 255 (white), three trapezium functions would be sufficient, therefore modeling 'Low', 'Medium' and 'High' grey-level values. The bases and shoulders would be statically defined by the User, however would be open for editing between running the Congealing code for experimentation purposes. For Non-Probabilistic entropy the highest membership for each pixel from each of the three trapeziums would be taken as the membership degree.

1.2.2 Research questions

- Does the use of Fuzzy Entropy alignment metrics improve the alignment of mammograms?
- Do clinicians / radiographers / mammographers find the output at all useful?
- What advantages / disadvantages does each fuzzy entropy alignment metric entail?

1.3 Research Method

Chapter 2

Experiment Methods

2.1 Algorithms

2.1.1 Non-Probabilistic Entropy

2.1.2 Hybrid Entropy

2.2 Software

2.2.1 Methodology

An adapted Scrum methodology has been undertaken for this project. This has been supported by the tool available at `taiga.io` - a beta web app.

- Burn down chart
- User stories
- Retrospectives
- Daily standup

2.2.2 Design

- CRC cards

2.2.3 Implementation

2.2.4 Testing

Chapter 3

Results and Conclusions

Chapter 4

Critical Evaluation

Appendices

Appendix A

Third-Party Code and Libraries

1.1 Congealing Code

The project focused on extending the existing Congealing Code implemented by Learned Miller et al in 2005. A Congealing demo is available on the Congealing website [3] which is open for experimentation. The original demo code was modified and extended to be able to read in mammograms and to work with 2 Fuzzy Entropy algorithms.

Appendix B

Ethics Submission

2.1 Ethics Application Number: 3958

AU Status
Undergraduate or PG Taught

Your aber.ac.uk email address
lac32@aber.ac.uk

Full Name
Laura Collins

Please enter the name of the person responsible for reviewing your assessment.
Reyer Zwiggelaar

Please enter the aber.ac.uk email address of the person responsible for reviewing your application
rrz@aber.ac.uk

Supervisor or Institute Director of Research Department
cs

Module code (Only enter if you have been asked to do so)
CS39440

Proposed Study Title
Entropy based metrics for joint image alignment

Proposed Start Date
25th January 2016

Proposed Completion Date
4th May 2015

Are you conducting a quantitative or qualitative research project?
Mixed Methods

Does your research require external ethical approval under the Health Research Authority?
No

Does your research involve animals?
No

Does your research involve human participants?
Yes

Are you completing this form for your own research?
Yes

Does your research involve human participants?
Yes

Institute
IMPACS

Please provide a brief summary of your project (150 word max)
I will be investigating the use of Congealing multiple MIAS dataset mammograms using several fuzzy entropy alignment metrics. If time permits I plan on speaking to a specialist (radiologist) to determine whether the output mean images of the congealing process are of any significant use to the research into breast cancer detection.

I can confirm that the study does not involve vulnerable participants including participants under the age of 18, those with learning/communication or associated difficulties or those that are otherwise unable to provide informed consent?
Yes

I can confirm that the participants will not be asked to take part in the study without their consent or knowledge at the time and participants will be fully informed of the purpose of the research (including what data will be gathered and how it shall be used during and after the study). Participants will also be given time to consider whether they wish to take part in the study and be given the right to withdraw at any given time.
Yes

I can confirm that there is no risk that the nature of the research topic might lead to disclosures from the participant concerning their own involvement in illegal activities or other activities that represent a risk to themselves or others (e.g. sexual activity, drug use or professional misconduct). Should a disclosure be made, you should be aware of your responsibilities and boundaries as a researcher and be aware of whom to contact should the need arise (i.e. your supervisor).
Yes

I can confirm that the study will not induce stress, anxiety, lead to humiliation or cause harm or any other negative consequences beyond the risks encountered in the participant's day-to-day lives.
Yes

Please include any further relevant information for this section here:

Where appropriate, do you have consent for the publication, reproduction or use of any unpublished material?
Yes

Will appropriate measures be put in place for the secure and confidential storage of data?
Yes

Does the research pose more than minimal and predictable risk to the researcher?
No

Will you be travelling, as a foreign national, in to any areas that the UK Foreign and Commonwealth Office advise against travel to?
No

Please include any further relevant information for this section here:

If you are to be working alone with vulnerable people or children, you may need a DBS (CRB) check. Tick to confirm that you will ensure you comply with this requirement should you identify that you require one.
Yes

Declaration: Please tick to confirm that you have completed this form to the best of your knowledge and that you will inform your department should the proposal significantly change.
Yes

Please include any further relevant information for this section here:

Appendix C

Code Examples

Annotated Bibliography

[1] S. Al-Sharhan, F. Karray, W. Gueaieb, and O. Basir, "Fuzzy entropy: a brief survey," in *Fuzzy Systems, 2001. The 10th IEEE International Conference on*, vol. 3. IEEE, 2001, pp. 1135–1139. [Online]. Available: http://dx.doi.org/10.1109/fuzz.2001.1008855

Paper outlining the different implementations of Fuzzy Entropy, of which 3 will be selected and focused on during this Project.

[2] M. Cox, S. Sridharan, S. Lucey, and J. Cohn, *Least squares congealing for unsupervised alignment of images*, Jun 2008, p. 18.

A disregarded adaption of the Congealing algorithm - however was useful in highlighting performance issues in the original algorithm. Something which was near continuously faced when implementing heavier fuzzy entropy alignment metrics.

[3] E. Learned-Miller, "The congealing page." [Online]. Available: https://people.cs.umass.edu/~elm/congealing/

[4] E. G. Learned-Miller, "Data driven image models through continuous joint alignment," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 2, pp. 236–250, Feb. 2006. [Online]. Available: http://dx.doi.org/10.1109/tpami.2006.34

Learned-Miller's original Congealing method is the basis for this Project - however I am looking to further extend the alignment capabilities using fuzzy entropy metrics, rather than standard Shannon entropy as currently implemented. This paper was extremely useful for understanding of the basic concepts behind it, and will be a good reference guide throughout the project.