

# Initial plan: Analysis and diagnosis of histology images of tumours

by Marley Sudbury (1838838)  
<SudburyMS@Cardiff.ac.uk>

Supervisor: Prof. Paul Rosin  
<RosinPL@Cardiff.ac.uk>

CM3203  
One Semester Individual Project  
40 credits



# Contents

<b>1</b>	<b>Project Description</b>	<b>1</b>
<b>2</b>	<b>Ethics</b>	<b>2</b>
<b>3</b>	<b>Project Aims and Objectives</b>	<b>3</b>
<b>4</b>	<b>Work Plan</b>	<b>4</b>
	<b>References</b>	<b>7</b>



# 1 Project Description

In this project, I will develop a system which can analyse histological images of lymph nodes in order to detect any tumours present, and diagnose the type of tumour. The purpose of this is to validate and assist with the existing manual process, which is time intensive and requires review to counteract any potential human error (Vestjens et al. 2012). In order to assist with this process, the system should show which parts of the input image have lead to it's decision. This would make it an instance of explainable AI.

The aims and motivations for this project are similar to the paper *Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning* (Wang et al. 2021). The findings of this paper suggest that deep learning models can be comparable to experienced pathologists in detecting and classifying tumours. In this project, I will explore and compare several different machine learning models.

The data set I will be primarily using in this project is a set of histological images of lymph nodes provided by Dr Damian Farnell and Dr Adam Jones from the Cardiff University School of Dentistry. They will also provide advice relating to the problem throughout the project. There are similar data sets available on the internet, which may be used to supplement the main data set in training, validation or testing.

As well as the deep learning approach, I also aim to explore the use of other computer vision techniques to approach this problem, and compare these with the deep learning approach. Research by Gheisari et al. (2018) suggests that techniques such as scale-invariant feature transform (SIFT) can also be effective in classifying tumours. Some computer vision techniques may also be required in pre-processing the data set.

## 2 Ethics

I have considered the ethics of this project in discussion with my supervisor. The main ethical concern is the data set being used to train the deep learning models. The data used in this research is a collection of anonymised histology images of lymph nodes. It is secondary data, and not available to the public. However, my supervisor and I are unsure if these images constitute human material. If they don't, then no ethical review is required according to the framework for the Ethical Review of Research using Secondary Data and/or Publicly Available Information Only (Cardiff University 2021). However, if they do then ethical review is required. I have asked the School Research Ethics Committee (SREC) for advice on this and will proceed accordingly.

If ethical review is required, I will complete and submit the School of Computer Science and Informatics Application for Ethical Review.

### 3 Project Aims and Objectives

- Create an image classification system which can:
  - Detect presence of a tumour in a histological image of lymph node
  - Categorise the tumour(s)
  - Highlight the detected tumours to assist in manual diagnosis
  - Label each pixel in the image with the determined tissue type
- Implement multiple appropriate deep learning models (currently the following, subject to change):
  - ResNet
  - VGG
  - EfficientNet
- Implement a non deep learning approach
- Evaluate each approach taken in terms of the ways in which they differ structurally, and the differences between them in performance
- Compare results of the evaluation with results presented by others in similar research
- Write up the results of the project into a report

## 4 Work Plan

In addition to the plan below, I will be meeting with my supervisor every Monday at 11:00 to discuss progress on the project. Where a scheduled review meeting is indicated below, it will be in the same slot as the regular meetings.

- Week 1 (w/c 31/01):
  - Initial meeting with supervisor
  - **Milestone:** write and submit initial plan by deadline (23:00, 07/02/2022)
  - Background reading
- Week 2 (w/c 07/02):
  - Background reading
  - Implement simple deep learning image classifier to become familiar with fundamentals
  - Explore available computational resources (lab machines<sup>1</sup>, AR-CCA<sup>2</sup>, OpenStack<sup>3</sup>)
  - Send enquiry email to ethics committee as per section 2
- Week 3 (w/c 14/02):
  - Examine data set
  - Investigate and implement stain normalisation
  - Complete and submit ethics form if required

---

<sup>1</sup><https://docs.cs.cf.ac.uk/notes/advice-on-using-cuda-in-the-linux-lab/>

<sup>2</sup><https://www.cardiff.ac.uk/advanced-research-computing/>

<sup>3</sup><https://docs.cs.cf.ac.uk/notes/using-openstack/>



- Week 4 (w/c 21/02):
  - **Milestone:** train deep learning models on data set so that they can identify and diagnose tumours
  - Fine tune deep learning models based on initial evaluation
- Week 5 (w/c 28/02):
  - Implement non deep learning approach to tumour detection and diagnosis
- Week 6 (w/c 07/03):
  - **Scheduled review meeting (11:00, 07/03)**
  - Evaluate performance of systems
- Week 7 (w/c 14/03):
  - Continue evaluation
  - Write report section ‘Results and Evaluation’
- Week 8 (w/c 21/03):
  - Write report section ‘Background’
  - Write report section ‘Approach’
- Week 9 (w/c 28/03):
  - Write report section ‘Future Work’
  - Write report section ‘Conclusions’
  - Write report section ‘Reflection on Learning’
- Easter recess (02/04–24/04)
  - If any of the previous work takes longer than planned, the Easter recess can be used to catch up
- Week 10 (w/c 25/04):
  - **Scheduled review meeting (11:00, 25/04)**
  - Write report section ‘Introduction’
  - Write report section ‘Abstract’
  - Write report section ‘Acknowledgements’

- Week 11 (w/c 02/05):
  - Redraft report and ensure that glossary, table of abbreviations, appendices and references are all present and correct
- Week 12 (w/c 09/05):
  - **Milestone:** finalise report and submit by deadline (23:00, 13/05/2022)

# References

- Cardiff University. 2021. *Research ethics*. Available at:  
<https://intranet.cardiff.ac.uk/students/study/postgraduate-research-support/integrity-and-governance/research-ethics> [Accessed:  
3 February 2022].
- Gheisari, S., Catchpoole, D. R., Charlton, A., Melegh, Z., Gradhand, E.  
and Kennedy, P. J. 2018. Computer Aided Classification of  
Neuroblastoma Histological Images Using Scale Invariant Feature  
Transform with Feature Encoding. *Diagnostics (Basel)* 8(3), doi:  
10.3390/diagnostics8030056.
- Vestjens, J. H. M. J. et al. 2012. Relevant impact of central pathology  
review on nodal classification in individual breast cancer patients.  
*Annals of Oncology* 23(10), pp. 2561–2566. doi:  
10.1093/annonc/mds072.
- Wang, X., Chen, Y. and Gao, Y. 2021. Predicting gastric cancer outcome  
from resected lymph node histopathology images using deep learning.  
*Nature Communications* 12, doi: 10.1038/s41467-021-21674-7.

