Project Details	
Project Code	MRCIIAR24Ca Farnell
Title	Combating Cancer with Data Science: Deep Learning of Histological Images for the Detection of Lymph Node Metastases
Research Theme	Infection, Immunity, Antimicrobial Resistance and Repair
Summary	Lymph nodes are a common site for the spread of cancer. The student will learn state-of-the-art techniques in Data Science involving Deep Learning to detect cancer automatically in lymph-node histology images. This will alleviate the workload of trained histologists in the UK, who struggle to keep up with demand. It will lead to earlier detection and lives saved. Training across two GW4 universities will be immersed in strongly multidisciplinary environments.
Description	Need For This Research: Lymph nodes (LNs) are common sites of metastasis (Jones et al. Frontiers in Oncology 8, 36. 2018). LN cancers are typically diagnosed by histological examination after biopsy. Trained histologists in the UK struggle to cope with demand & this project will alleviate this demand by analysing these images automatically. It will lead to earlier detection of cancers & lives saved. AI methods equal (or surpass) humans in such automated tasks. Originality: Deep Learning (Goodfellow, MIT press, 2016) is a subset of machine learning with its methods based on neural networks. Relatively little research has been carried out Deep Learning has been applied to histological images of lymph tissues, as far as we are aware. Originality will come from applying this promising & powerful technique to a new area & in identifying the most effective network architectures & appropriate (albeit standard) pre-processing steps to analyse these images effectively. This is therefore a good, solid PhD project with potentially much impact in future. Study Design: A previous exploratory undergraduate project was entitled Detecting Metastatic Tissue in Lymph Nodes with Deep Learning (Marley Sudbury, Undergraduate Project Report, Cardiff University 2022). This project used Convolutional Neural Networks (CNNs) via TensorFlow (www.tensorflow.org), which is a free end-to-end open source platform for machine learning that was implemented in the Python programming language. This software will also be used in the proposed PhD. Freely available histological image data & image classification / labels was used from the Head & Neck 5000 (HN5K) study (www.head&neck5000.org.uk) & the Camelyon 16 data set (camelyon16.grand-challenge.org/Data/). The PhD proposed here will refine this use of CNNs via TensorFlow for metastatic cancer tissue detection in histological images for these two freely available datasets. Training will be given in using TensorFlow, Python, & machine learning. In conjunction with supervisors, the student will steer the direction of project (including setting research objectives), as they will solve specific technical problems (below). Initially however, learning outcomes / objectives are: 1. Months 1-3: Successful completion of the 3-month settling in period, including training & starting a literature review. 2. Months 4-15: Student learning on Deep Learning applied to entire images & (separately) via patch-based to classify images into those that contain metastases & those that do not. The student will also learn how to identify optimal network architectures & apply strong validation techniques – both within & between different datasets.

	Evidence of completion of learning shown by the submission of a paper to Journal of Computer Methods & Programs in Biomedicine & of a first-year report. 3. Months 16-27: Learning, implementation, & (again) determination of how to optimise delineation (i.e., segmentation) of tumours in these images via Deep Learning. Evidence of learning shown by submission of a paper to the Journal IEEE Transactions on Image Processing & of a second-year report. 4. Months 28-33: The final learning outcome will be to employ appropriate pre-processing steps (e.g., KERAS pre-processing) that will enhance both image classification & tumour segmentation. Again, evidence of completion will be shown by the submission of a paper to Medical Image Analysis & in a third-year report. 5. Months 34-39: Completion & submission of the PhD thesis. 6. Months 40-45: Application for a transition to independence for the GW4 Biomed2 for the student for further studies along these lines + any exploitation or commercialisation & / or application by the supervision team + student for a post-doctoral research position from (e.g.) EPSRC. 7. Months 46-48: Successful completion of a three-month industrial placement. (Note that this might occur at any point of the PhD, as appropriate).
Supervisory Team	
Lead Supervisor	
Name	Dr Damien JJ Farnell
Affiliation	Cardiff
College/Faculty	College of Biomedical and Life Sciences
Department/School	School of Dentistry
Email Address	farnelld@cardiff.ac.uk
Co-Supervisor 1	
Name	Dr Adam Jones
Affiliation	Cardiff
College/Faculty	College of Biomedical and Life Sciences
Department/School	School of Dentistry
Co-Supervisor 2	
Name	Dr Miranda Pring
Affiliation	Bristol
College/Faculty	Faculty of Health Sciences
Department/School	Bristol Dental School
Co-Supervisor 3	
Name	Professor Paul Rosin
Affiliation	Cardiff
College/Faculty	College of Physical Sciences and Engineering
Department/School	School of Computer Science & Informatics