CS5098 MSc Dissertations
MS Teams Meetings

All meetings were conducted via MS Teams during the development of the common deep learning pipeline (from 24/06 to 09/07).

Team Members

- Adam Jaamour
- Ashay Patel
- Shuen-Jen Shen

Legend

Colour	Task	Attendance
	Task completed	Attended meeting
	Task incomplete	Missed meeting

24/06/2020 - MS Teams Meeting 1

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	N/A
Ashay	Yes	N/A
Shuen-Jen	Yes	N/A

Issues discussed & Tasks carried out

- Set out tasks on what to implement for the next meeting for a very basic DL pipeline with data pre-processing, a pre-trained CNN model and some output visualisation.
- Setup common GitHub repo and working environment in separate branches.
- Queried St Andrews FixIt service for downloading CBIS-DDMS dataset on lab machine.

Tasks set for 29/06

Name	Tasks
Adam	Data pre-processing using mini-MIAS dataset for now: ● Initial pre-processing (CSV, script to organise images into class folders) ● Create usable dataset (images mapped to labels) ● Resize images for VGG19 CNN model input and split datasets ● (investigate data augmentation) ● (investigate how to remove artefacts)
Ashay	Implement VGG19 through Keras (on any data for now). Try 2 models: ● VGG19 on its own (using downsampled images) ● VGG19 with extra convolution/spooling layers using 512px image
Shuen-Jen	Take output from any template neural network from TF/Keras and visualise the output metrics using matplotlib and seaborn Output metrics to create: ROC, AUC, Overall accuracy to compare with other papers, Confusion matrix for classification
All	Download CBIS-DDSM dataset (send email to FixIt to either download after connecting to VPN, or to copy from Shuen-jen's lab computer).

29/06/2020 - MS Teams Meeting 2 (code review #1)

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	Yes
Ashay	Yes	Yes
Shuen-Jen	Yes	Yes

Issues discussed & Tasks carried out

- Merged all branches to master
- Combined the output of data pre-processing (Adam) to the input for the VGG19 model (Ashay)
- Combined the output of the CNN model (Ashay) to the input of the result visualisation function (Shuen-jen)
- Applied mini-MIAS dataset transformation for Ashay and Shuen-Jen and tested/debugged full pipeline on GPU, refactored and reorganised some of the code.

Tasks set for 01/07

Name	Tasks
Adam	Reach out to IT about CBIS-DDSM dataset. Code improvements (normalise input images, include model name in output figure filename). Update README instructions for running the code.
Ashay	Run training over more epoch for wednesday meeting
Shuen-Jen	Formatting the Confusion Matrix
All	Research GPU memory usage for multiple training cycles

30/06/2020 - MS Teams Meeting 3 (code review #2)

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	Yes
Ashay	Yes	Yes
Shuen-Jen	Yes	Yes

Issues discussed & Tasks carried out

- Stuart replied that the DDSM dataset is available on BigTMP and CBIS-DDSM is still downloading and will be uploaded right after.
- Merged code and ensured it worked fine
- Checked that the DDSM dataset was accessible from BigTMP

Tasks set for 01/07

Name	Tasks
Adam	Research gpu memory usage for multiple training cycles Investigate github commits author issue
Ashay	Transforms code & shuffle data when splitting
Shuen-Jen	Rotate labels (e.g. 45 degrees)
All	N/A

01/07/2020 - MS Teams Meeting 4

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	Yes
Ashay	Yes	Yes
Shuen-Jen	Yes	Yes

Issues discussed & Tasks carried out

- CBIS-DDSM dataset successfully downloaded on BigTMP.
- Carry out changes recommended by David:
    - Do not use test dataset
    - Use same train/test/validation splits as other papers
    - Use validation output in confusion matrices
- Presentation:
    - Created presentation slides
    - Rehearsed presentation

No task set

02/07/2020 - MS Teams Meeting 5

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	N/A
Ashay	Yes	N/A
Shuen-Jen	Yes	N/A

Issues discussed & Tasks carried out

- Gave our presentation for team DHB.
- Discussed tasks for next week: need to process large CBIS-DDSM dataset rather than mini-MIAS. Model and outputs are good enough for now and don't need to be touched.
- Both CBIS-DDSM and DDSM datasets have been downloaded and available at */cs/tmp/datasets/CBIS-DDSM*.

Tasks set for 07/07

Name	Tasks
Adam	Create dummy data using the mini-MIAS data set (similar CSV to the one Shuen-Jen is creating) Research data generators in Keras and batch data processing
Ashay	
Shuen-Jen	Generate CSV file with all paths that we need to use: ● 1st column is path to the image and the 2nd other column is the label ● Use the CSV files that are already split in training/testing sets ● (need to append the */cs/tmp/datasets/CBIS-DDSM* to the path) ● https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM#5e40bd1f79d64f04b40cac57ceca9272
All	N/A

07/07/2020 - MS Teams Meeting 6 (code review #3)

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	Yes
Ashay	Yes	Yes
Shuen-Jen	Yes	Yes

Issues discussed & Tasks carried out

- Merged code running on CBIS-DDSM dataset.
- Fix predictions for binary tasks (probability instead of class) and test the whole pipeline on both CBIS-DDSM and mini-MIAS datasets.
- Problems to consider for future code developments:
    - Maintain aspect ratio when resizing
    - Figure output saving, differentiate between binary and multi
    - Monitor accuracy/loss changes for CBIS-DDSM
    - Investigate out of memory warnings (doesn't appear with small batch size = 5 on mini-MIAS)

Tasks set for 08/07

Name	Tasks
Adam	Run basic model on CBIS-DDSM
Ashay	Run advanced model on CBIS-DDSM
Shuen-Jen	Run basic model on CBIS-DDSM
All	Saved outputs to google drive

08/07/2020 - MS Teams Meeting 7

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	Yes
Ashay	Yes	Yes
Shuen-Jen	Yes	Yes

Issues discussed & Tasks carried out

- Meeting with David & Lewis
- Can ignore TF warnings
- Improvements to consider (based on David's suggestions):
    - Future: combine images from the same cases into 2 channels?
    - Instead of resizing image to VGG size, use conv layers at the beginning to resize (still need to make it square). Resize to 2048x2048px with padding to maintain aspect ratio, use 2 conv layers with stride 2 to get to 512x512px.
    - Complete the list of papers we are comparing with, including section of image used (ROI or whole image), type of CNN, dataset used and what type of image used (everything or just calc/mass/...), image size (original or resized), training/test/validation split.

No tasks set

09/07/2020 - MS Teams Meeting 8

Attendance & Task Completion

Name	Present	Task completed?
Adam	Yes	N/A
Ashay	Yes	N/A
Shuen-Jen	Yes	N/A

Issues discussed & Tasks carried out

- Merged code refactoring (adam) and large image sizes (ashay) and tested the results.
- Refactored output file names for better organisation.
- Use epoch size 100/50
- Upload all results so far to the shared “results” directory on our shared Google Drive.

Tasks set for 10/07

Name	Tasks
Adam	Run CBIS-DDSM basic model with large images (batch size 2)
Ashay	Run CBIS-DDSM advanced model with large images (batch size 2)
Shuen-Jen	Run CBIS-DDSM basic model with large images (batch size 1) Add 2 papers to the excel sheet of papers
All	Setup private repositories to individual development