Cancer Detection with Deep Learning

Applied Deep Learning

Tallis Shih-Ying Jeng

sj2909

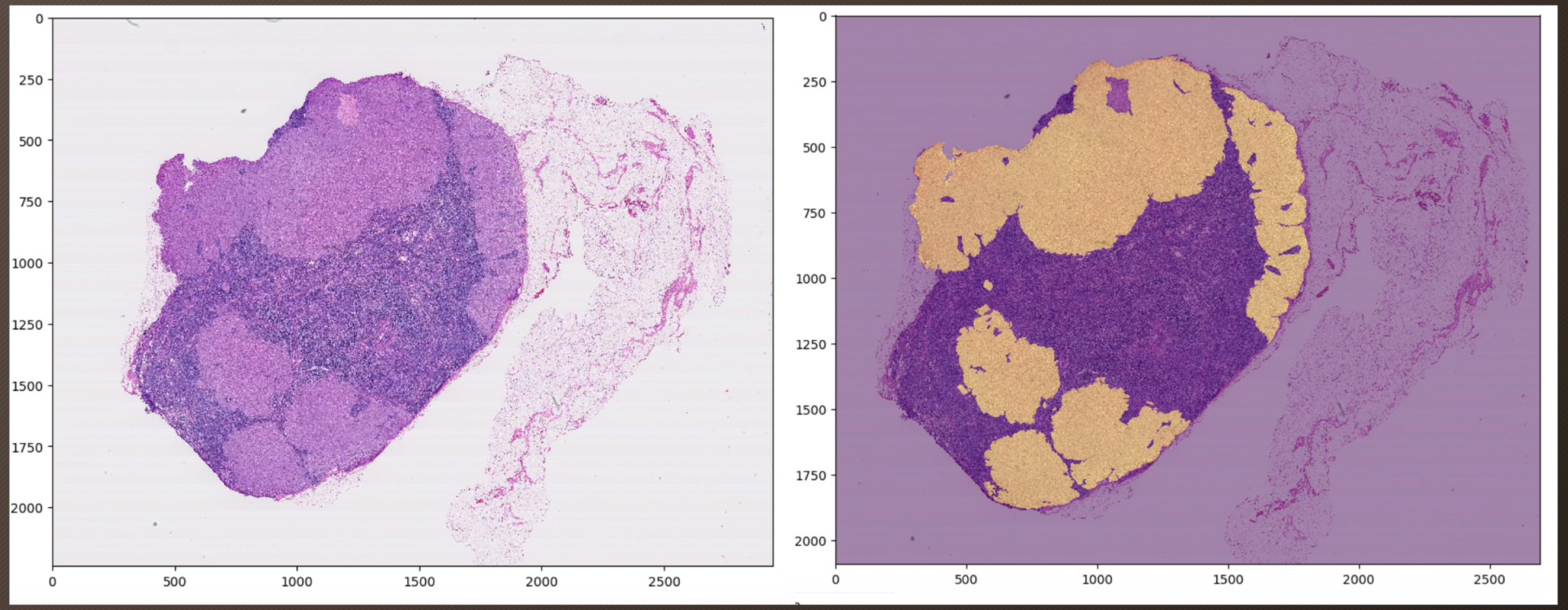
Outline

1. Introduction
2. Data Processing
3. Models & Experiments
4. Results & Analysis
5. Conclusion

1 Introduction

1. Introduction: What

- *A deep learning based system that automatically detect **cancer region in pathology images***



1. Introduction: Why

- *Manual annotation is costly and has **low scalability***
- *Deep learning system is potentially more accurate given the amount of training data the model can learn on in a relatively short period of time*
- *Assistive diagnostic tool provides a second opinion and therefore lowers the risk of misdiagnosis*

1. Introduction: Objectives

- *Build a **Proof-of-concept prototype** for cancer detection*
- *Experiment **multi-scale ensemble models***
  - **Based on the paper** Detecting Cancer Metastases on Gigapixel Pathology Images
- *Design evaluation metrics and analyze experimental results*

1. Introduction: Implementation

- *Models on slide level 5 (32x), models on slide level 7 (128x)*
- *Models for different levels trained independently*
- *Quantitative & Qualitative analysis*
- *Heatmap on slide level 7*

2 Data Processing

2. Data Processing: patch extraction

- *Extract patches (299 x 299) from the slides*
  - *Slides stored on GCS*
  - *Extracted patches stored in Google Drive*

- *Target Label for binary classification*
  - *ROI: Center pixels*
  - *Sum pixels in ROI of the mask -> target class label*

2. Data Processing: Train/Val/Test Split

- *Train & Validation set:*
  - *Randomly extract patches*
  - *Ensure class balanced by under-sampling majority (negative) classes (60 images from each class per slide)*
  - *slides '078', '084', '091', '094', '096', '101'*
- *Test set:*
  - *Slide window across the slide*
  - *Maintain class ratio as it is*
  - *Slide '110'*

3 Models

3. Model: Transfer Learning

- *Transfer Learning VGG16:*
  - *Based on results, better than the result of InceptionV3 and MobileNet*
- *Neural Network*
  - *Dense layer with 512 units, relu activation, RMSprop optimizer*
  - *Dropout (0.5)*
- *Random Forest*
  - *200 estimators, max_depth=8*

https://www.fast.ai/2018/07/02/adam-weight-decay/

3. Model
Transfer Learning + Dense Layer

CNN Transfer Learning Model	
Layer	Component
Input (size x size x 3)	Input
(conv3-64) x2	Block 1
maxpool-2	
(conv3-128) x2	Block 2
maxpool-2 *	
(conv3-256) x3	Block 3
maxpool-2 **	
(conv3-512) x3	Block 4
maxpool-2	
(conv3-512) x3	Block 5
maxpool-2	
Dense-512	Dense
Dropout-0.5	
Softmax	Softmax

3. Experiments

	Transfer VGG16 + Dense layer	Transfer VGG16 + Random Forest
Level 5	Model nn5	Model rf5
Level 7	Model nn7	Model rf7

4 Results

4. Results: Evaluation Metrics

- **Quantitative Evaluation**
  - Accuracy
  - Precision, Recall, F1
  - AUROC
  - Average confidence on erroneous prediction
- **Qualitative Evaluation**
  - Inspect test samples that causes erroneous prediction
  - Heatmap quality

5 Conclusion

5. Conclusion

- *Built proof-of-concept prototype*
- *Multi-scale ensemble not helpful*
- *Ensemble of different model types helpful*

== THE END ==
THANK YOU VERY MUCH

GitHub Repository: https://github.com/JengTallis/adl-cancer-detection
YouTube Video: https://youtu.be/lTjs99j0P14