ELEC 4010N Final Project

Cheng Tsun Hei, Kang Ting Hao Howard

Topics

1. Project 1: Semi-supervised Classification
2. Project 2: Domain Generalization on Fundus Images

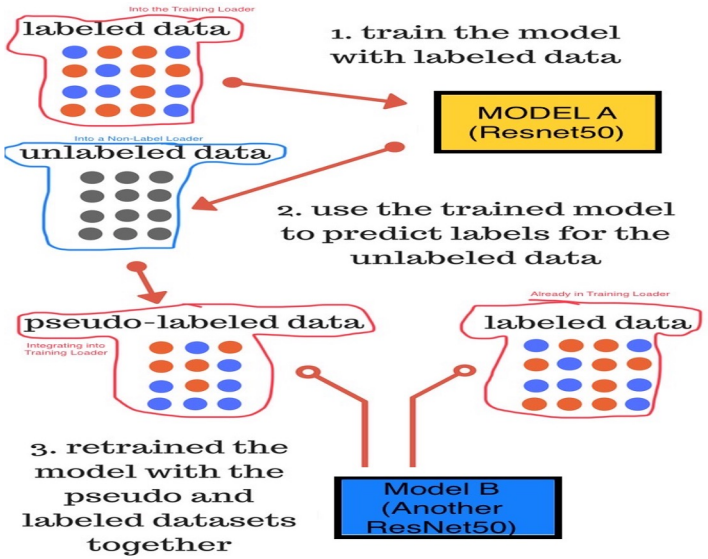
Project 1: Background Introduction

In real cases,, the majority of training data will be unlabelled: hence we need semi-supervised training

The algorithm is based on Pseudo Labelling

- Labelled Sample (Training): random 270
- Labelled Sample (Validation): random 90
- Unlabelled Sample (Training): random 540
- Other testing case

Project 1: Method Development



A. Pseudo Labeling

- A1. Build training and test loaders for Pseudo Labeling
  [ ] 1 cell hidden
- A2. Build the network for Pseudo Labeling
  [ ] 1 cell hidden
- A3. Train and test the model for Pseudo Labeling
  [ ] 1 cell hidden
- A4. Plot the training and test curves
  [ ] 2 cells hidden

B. Classification

- B1. Build training and test loaders for Classification
  [ ] 1 cell hidden
- B2. Build the network for Classification
  [ ] 2 cells hidden
- B3. Model Training and Testing
  [ ] 1 cell hidden
- B4. Plot the training and test curves
  [ ] 2 cells hidden

Image src: https://www.analyticsvidhya.com/blog/2017/09/pseudo-labelling-semi-supervised-learning-technique/

Project 1: Result Analysis

Experiment	#Epoch (A, B)	AUC (Last epoch of A)	Test AUC for B	Test ACC for B
Benchmark	(N/A, 10)	N/A	77.9056%	78.8889%
1	(10, 10)	74.9251	65.7343%	84.4444%
2	(10, 10)	53.2828%	57.3232%	73.3333%
3	(5, 10)	63.3488%	57.9475%	78.8889%
4	(5, 10)	71.3675%	73.8248%	85.5556%

Project 1: Conclusion

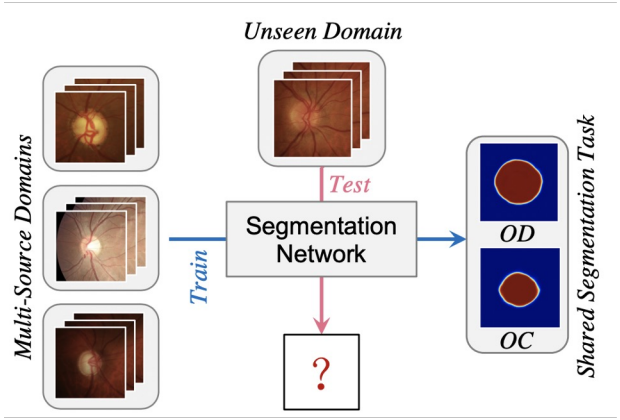
- For simplicity, no regularization parameter is given:
  - Performance may improve with regularization.

- Overfitting issue is severe.
  - May require more data or make a easier model

- The final result is similar between supervised and semi-supervised model

Project 2: Background Introduction

- Segmentation on Fundus Images
  - Optic disc and cup
- Domain Generalization
  - Perform well in unseen domain
  - Useful in real world
- Data (With different scanners):
  - Domain 1: Drishti-GS (Train: 50, Test: 51)
  - Domain 2: RIM-ONE-r3 (Train: 99, Test: 60)
  - Domain 3: REFUGE (Train) (Train: 320, Test: 80)
  - Domain 4: REFUGE (Test) (Train: 320, Test: 80)



Project 2: Method Development

- Domain-oriented Feature Embedding (Dofe)
- Four experiment conducted:
  - Epoch: 80, learning rate: 0.001, model: DeeplabV3+ (include encoder and decoder)

Train	Test
Domain: 1, 2, 3	Domain: 4
Domain: 1, 2, 4	Domain: 3
Domain: 1, 3, 4	Domain: 2
Domain: 2, 3, 4	Domain: 1

Project 2: Result Analysis

- Segmentation Performance on Optic Cup:

Test:	Domain 1	Domain 2	Domain 3	Domain 4	Average
Dice Coefficient	0.8484	0.8290	0.8614	0.8808	0.8799
95% of Hausdorff Distance	30.12	22.22	19.84	14.46	21.66
Average surface Distance	15.44	11.56	9.852	6.627	10.87

Project 2: Result Analysis

- Segmentation Performance on Optic Disc:

Test:	Domain 1	Domain 2	Domain 3	Domain 4	Average
Dice Coefficient	0.9640	0.9154	0.9322	0.9332	0.9362
95% of Hausdorff Distance	13.91	21.84	18.55	16.23	17.63
Average surface Distance	6.305	12.31	9.506	7.47	8.898

Project 2: Conclusion

- Generalization
  - By using domain generalization, it would be getting better performance for unseen datasets/domain.
- Complexity increased
  - More domain models need to train
- Performance Trade-off
  - Need to fufil all domains