Detecting Glaucoma with SegNet

SHANGHAI JIAO TONG UNIVERSITY

Hongyi Guo, Yuqiao He

For more details, please visit our github repository.

Introduction

- Glaucoma (青光眼) is the second leading cause to blindness and has no early symptoms. So, detecting Glaucoma via computer vision approaches can be very helpful to prevent this horrible disease.
- In this project, we implemented SegNet and achieved fine image segmentation accuracy and IU on eye photos. Based on our segmentation, we get a pretty precise cup-to-disc ratio, which is a key part of glaucoma detection. And we did some experiments on different kind of loss functions to find which one is the best for eye photos' segmentation. Last but not least, we implemented transfer learning, to be exact, ADDA (Adversarial Discriminative Domain Adaptation) to segment eye photos with different styles, such as different brightness.

Background

- Glaucoma (青光眼)
  - Glaucoma is a group of eye diseases which result in damage to the optic nerve and vision loss. This disease gradually reduces your peripheral vision. It has no early symptoms and therefore is called the "Silent Thief of Sight". Glaucoma is one of the most common chronic eye diseases for seniors and is the leading cause of blindness for those over the age of 64. Half the people with glaucoma do not know they have it.
- Cup-to-disc ratio (杯盘比)
  - The cup-to-disc ratio (often notated **CDR**) is a measurement used in ophthalmology and optometry to assess the progression of glaucoma.

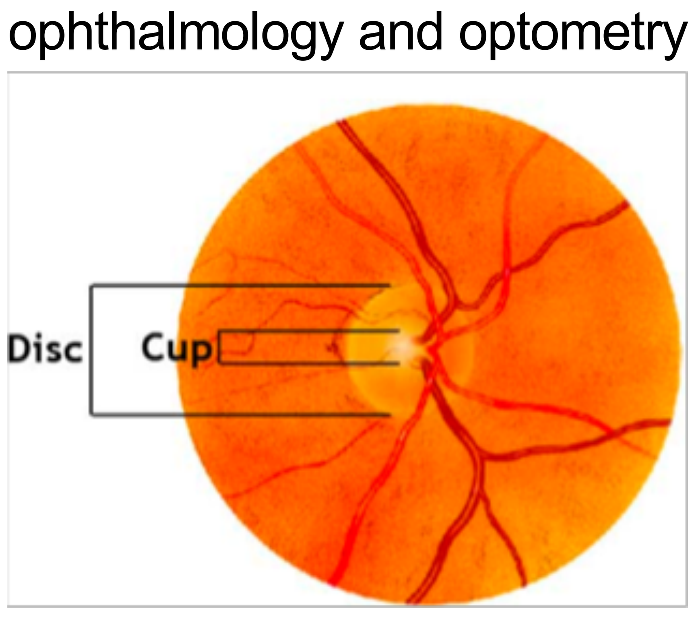


Fig 1: Demonstration of Cup and Disc

- Glaucoma, which is in most cases associated with an increase in intraocular pressure, often produces additional pathological cupping of the optic disc. As glaucoma advances, the cup enlarges until it occupies most of the disc area.

Algorithm

- SegNet
  - SegNet is a deep encoder-decoder architecture for multi-class pixelwise segmentation. The architecture consists of a sequence of non-linear processing layers (encoders) and a corresponding set of decoders followed by a pixelwise classifier.
  - Typically, each encoder consists of one or more convolutional layers with batch normalization and a ReLU non-linearity, followed by non-overlapping maxpooling and sub-sampling. The sparse encoding due to the pooling process is upsampled in the decoder using the maxpooling indices in the encoding sequence (see the figure below).

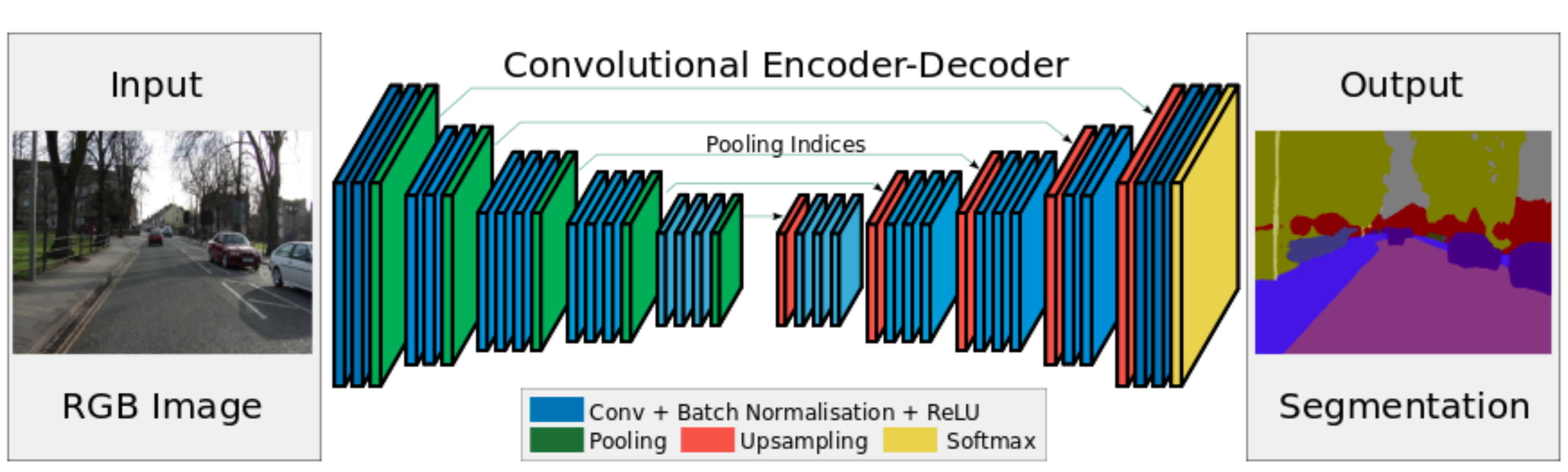


Fig 2: Architecture of SegNet

- Transfer Learning
  - Domain Adaption
    - Domain adaptation is a field associated with machine learning and transfer learning. This scenario arises when we aim at learning from a source data distribution a well performing model on a different (but related) target data distribution.
  - ADDA
    - The adversarial discriminative domain adaptation (ADDA) considers independent source and target mappings by untying the weights, and the parameters of the target model are initialized by the pre-trained source one.
    - This is flexible because of allowing more domain specific feature extractions to be learned. ADDA minimizes the source and target representation distances through iteratively minimizing these following functions, which is most similar to the original GAN:

Experiment

- Base Experiment
  - See the figure below

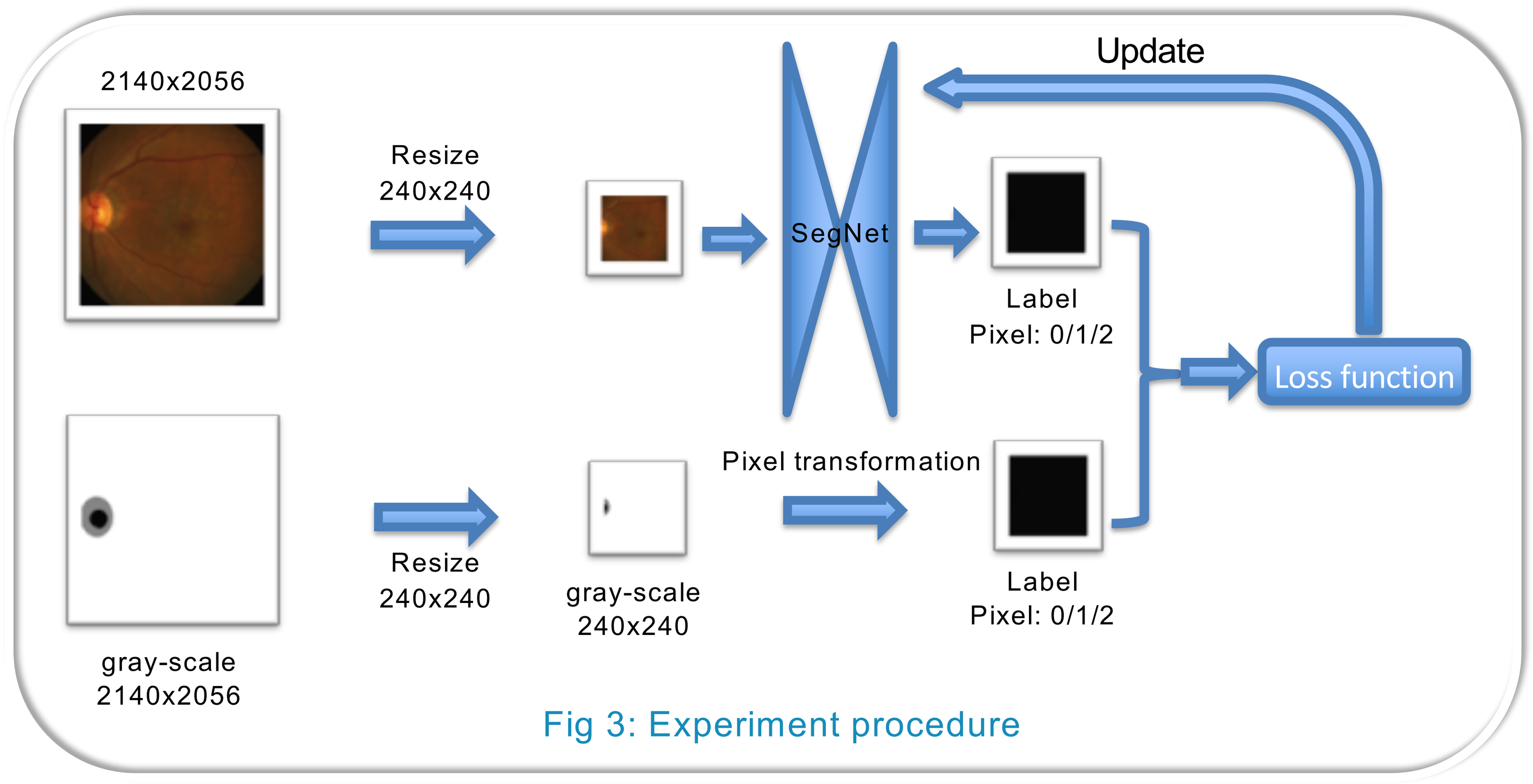


Fig 3: Experiment procedure

- Then we did some experiment on loss function
  - Normal cross entropy

$$-\sum_{c=1}^{M} y_{o,c} \log(p_{o,c})$$

Note

- M - number of classes (dog, cat, fish)
- log - the natural log
- y - binary indicator (0 or 1) if class label c is the correct classification for observation o
- p - predicted probability observation o is of class c

- Weighted cross entropy
  - Total loss is defined as the weighted summation of the cross entropy loss of each class.
  - The weight is related to the proportion of each class. The larger the proportion is, the smaller the weight. In our case, the loss weight of cup area, disc area and other area is 7.3, 2.6 and 0.03.
- Dice coefficient loss
  - The dice coefficient loss of each class is defined as

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

  - X and Y are our output and ground truth label.
  - Our experiment results as following shows that despite the accuracy are almost the same in the end, dice loss (blue) can achieve a better IU and its loss curve is more reasonable. And in fact, dice loss computes faster than others.

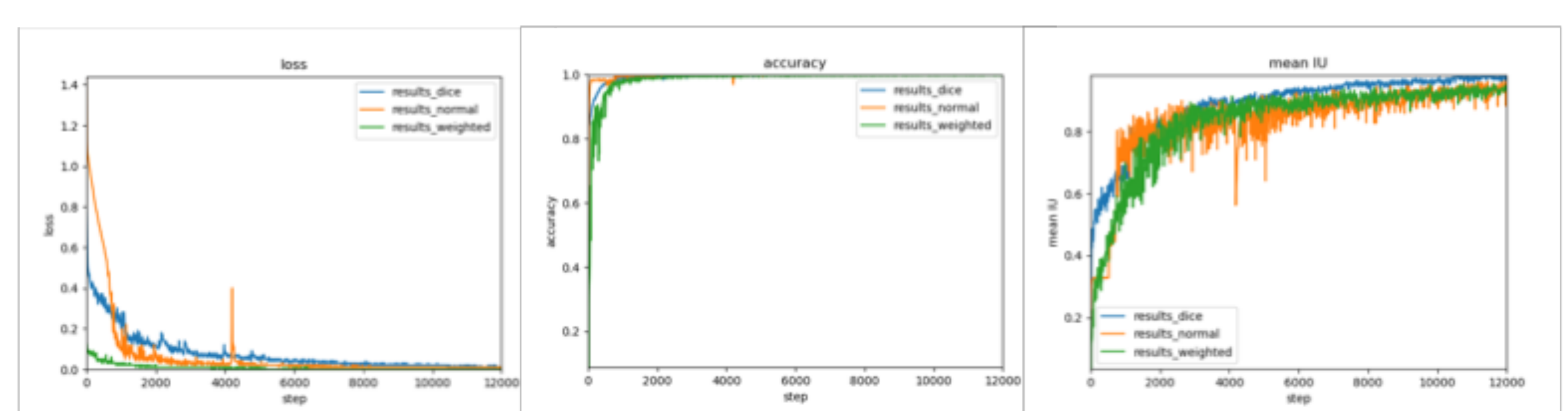


Fig 3: Experiment procedure

- Transfer Learning Experiment
  - We implemented ADDA (Adversarial Discriminative Domain Adaptation) to do transfer learning on a dataset with larger brightness, which has a very pool segmentation effect if we don't add domain adaption.
  - ADDA will learn a target encoder (as Generator) with GAN. The Discriminator judges the encode result is from the source encoder or target encoder. Note that the target decoder and classifier are not trained and use parameters from source SegNet.
  - The whole process is described as :

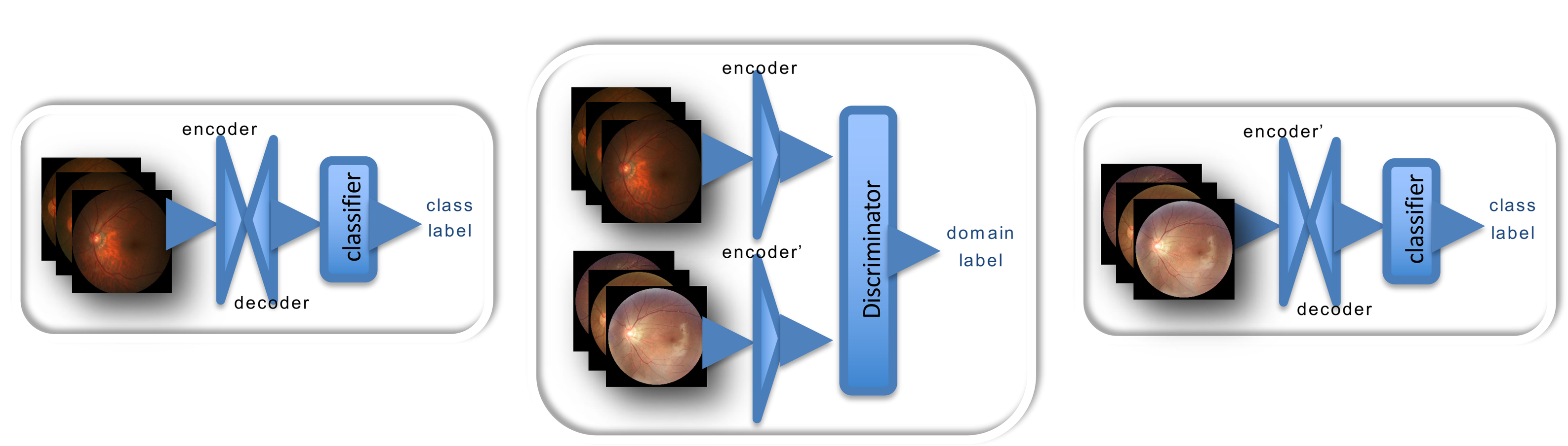
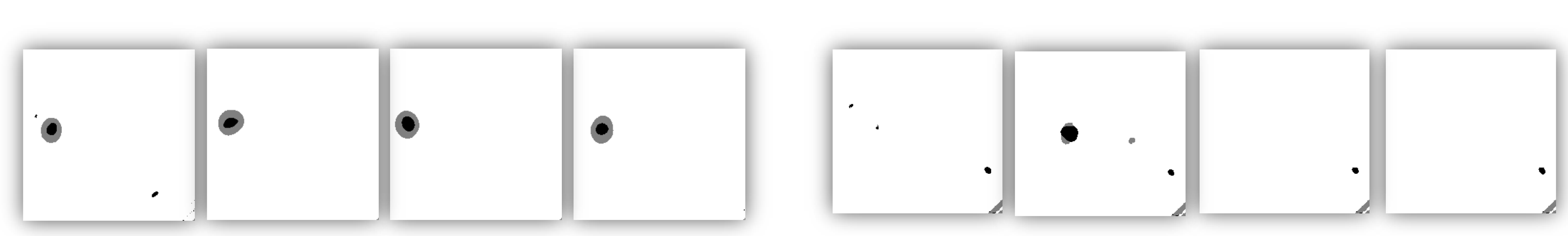


Fig 4: demo of ADDA

  - We achieved good results



After ADDA

Before ADDA