Learning Deep Representations of Medical Images using Siamese CNNs with Application to Content-Based Image Retrieval

6.869 Advances in Computer Vision / Final Project

Yu-An Chung / Wei-Hung Weng

Dec 12, 2017

Accepted and Presented in *NIPS 2017 Machine Learning for Health Workshop*

Massachusetts Institute of Technology

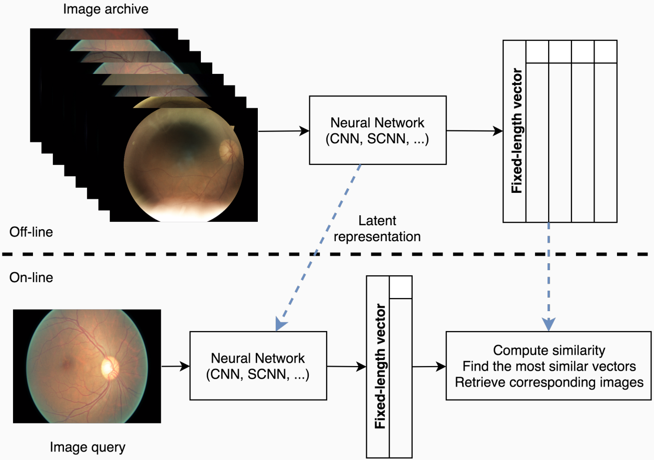
CSAIL

Learning Representation

- Learning effective data representation are key factors of successful medical imaging machine learning tasks [Litjens 2017]
- But expert annotation is usually the bottleneck.

Content-Based Medical Image Retrieval (CBMIR)

- Help decision making by retrieving similar images
- Similar case identification and knowledge discovery



Related Works

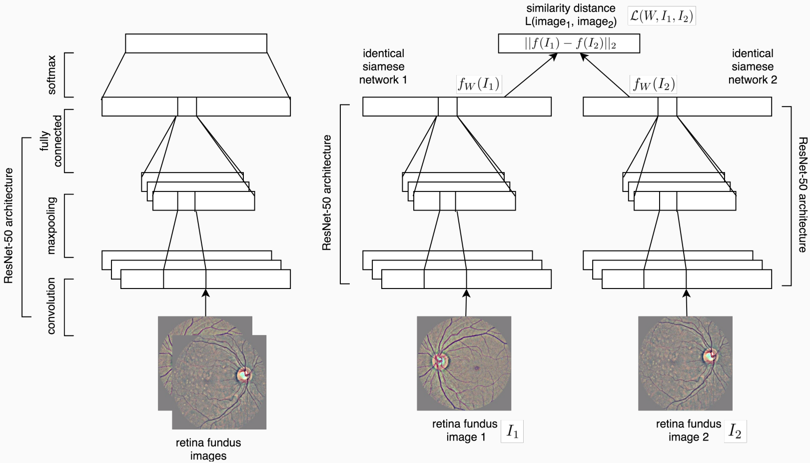
- Deep learning in Medical Imaging
    - Less handcrafted features
    - Classification [Esteva 2017] / segmentation [Havaei 2017] / image generation [Nie 2017] / captioning [Shin 2015] / CBMIR
- Deep learning in CBMIR
    - lung CT [Sun 2017] - ResNet
    - prostate MRI [Shah 2016] - CNN + hashing-forest
    - X-ray image [Anavi 2016, Liu 2016] - five-layered pre-trained CNN, extracted the fully-connected layer, integrated textual metadata and fed into an SVM classifier / combined three-layer CNN with Radon barcodes
    - Combining single pre-trained CNN structure with other techniques but without an end-to-end approach.
    - But relying heavily on expert labeling!

Proposed Approach

- Deep Siamese convolutional neural networks (SCNN)
- Learn fixed-length image representation by only image pair information
- Experiment on CBMIR of diabetic retinopathy (DR) fundus images

Deep SCNN [Bromley 1994]

- Relationship and similarity between the image pair
- ResNet-50 + ImageNet pre-trained weight, 25% dropout + batch normalization, ReLU, Adam, Euclidean distance



Loss Function & Evaluation

- Contrastive loss

$$\mathcal{L}(W, I_1, I_2) = \mathbf{1}(L = 0)\frac{1}{2}D^2 + \mathbf{1}(L = 1)\frac{1}{2}[\max(0, margin - D)]^2$$

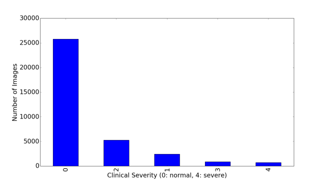
- Evaluation

$$MRR = \frac{1}{Q}\Sigma_{i=1}^{Q}\frac{1}{rank_i}$$

$$MAP = \frac{1}{Q}\Sigma_{i=1}^{Q}AveP$$

Diabetic Retinopathy (DR) Image Dataset

- 35,125 images from Kaggle DR Detection challenge
- 5 severity labels (normal to severe DR) given by experts
- Preprocessing to reduce image variance
    - Rescaling, contrast, brightness, hue normalization, resizing
- Class balancing by augmentation
    - Random cropping / flipping / Gaussian blurring / sharpening / rotation



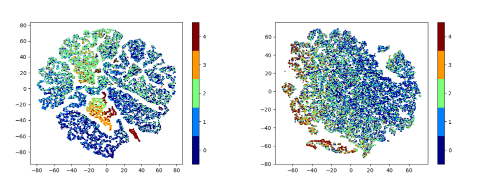
- 70% train, 30% test

CBMIR Performance

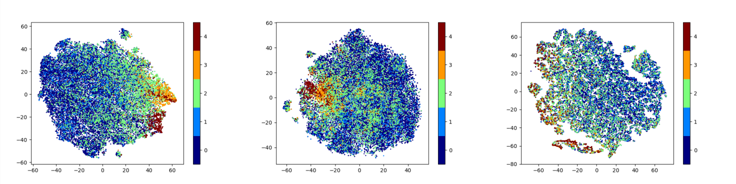
Layer	CNN (3^{rd})	CNN (2^{nd})	CNN (softmax)	SCNN (last)
MAP	0.6209	0.6369	0.6673	0.6492
MRR	0.7608	0.7691	0.7745	0.7737

Visualizing Latent Representations by PCA + tSNE

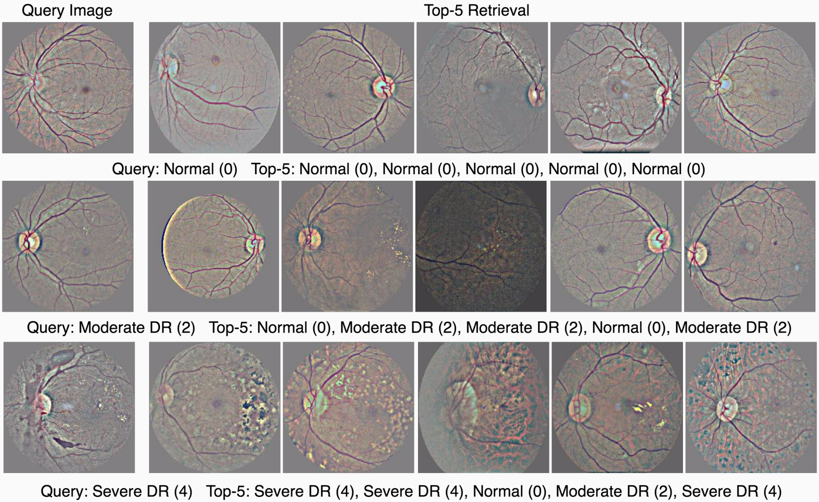
- CNN softmax and SCNN last layer - high performance
- Sliding scale explain the clinical nature of DR progression



- SCNN outperformed other sliding scale representations



CBMIR Examples



Conclusions

- Learning efficient latent representation of DR images by deep SCNN using public dataset
- Only requires image pair information, less expert labeling

- Yu-An Chung (`andyyuan@mit.edu`)
- Wei-Hung Weng (`ckbjimmy@mit.edu`)