Computer "Vision": A Deep-Learning-Based Approach to Detecting Diabetic Retinopathy

Pranav Upadhyayula, Lauren Yang, and Sushil Upadhyayula
{pranavu, lauren11, sushilu}@stanford.edu

Overview

The Problem

- Diabetic Retinopathy (DR) affects 93 million people, most of whom aren't diagnosed until late stage.
- Family eye doctors don't receive training in detecting DR.
- Specialized ophthalmologists still go through a very manual & tedious process to test for DR. [1] [2]

Our Solution

- Use a DenseNet CNN with Softmax output layer to classify images of the eye into 5 stages of DR.
- Use Class Activation Maps (CAMs) to provide interpretability & highlight potential problem areas on image.
- *Results show significant prowess in identifying early-stage DR.*

Real-World Clinical Application

- Interviews with Palo Alto-based optometrists reveal that this would especially help at the low level to be embedded within annual check-ups.
- Estimated 15 minute time savings per patient for specialized ophthalmologists.

Data/Features

Dataset

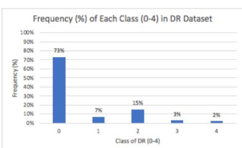
- Kaggle Dataset
- 5,000 RBG images (labeled with true class)
- 80/10/10 train/dev/test split

Pre-processing

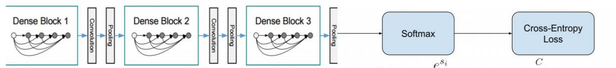
- Class imbalance (see below)
- Images are too high resolution, shrink to 1% of size.

Data Augmentation

- To simulate real world imperfections and generate more images, used:
  - Gaussian blur (random σ between 1.0 and 2.0)
  - Increase in contrast
  - Color shifting + Random Brightening/Darkening



Model



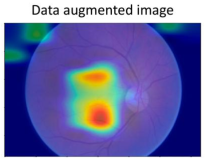
$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}} \qquad CE = -\sum_i^C t_i log(f(s)_i)$$

- DenseNet is based on the finding that CNNs are more accurate and efficient if they contain shorter connections between layers. [3]
- While traditional CNNs only have connections between adjacent layers, DenseNet splits the CNN up into Dense Blocks, where every layer within a dense block is connected to every other layer within that block. [4]
- This reduces the number of parameters by encouraging feature reuse, and solves the vanishing gradient problem. [5]
- DenseNet-121: 121 batch normalization layers, 120 convolutional layers, 121 activation layers, 58 concatenation layers, & 1 global average pooling layer.
- We feed this into a Softmax output layer with 5 classes.
- Mini-Batch Gradient Descent (batch size = 64)

Results

Original post-processed image

Data augmented image



Class Activation Map (CAM)

Comparison of Results

	Sensitivity	Specificity	Train Acc. (n = 8000)	Test Acc. (n = 1000)
Baseline	0.674	0.925	0.979	0.702
DenseNet Class 0	0.938	0.912	0.989	0.938
DenseNet Class 1	0.857	0.988	0.964	0.835
DenseNet Class 2	0.640	0.921	0.915	0.768
DenseNet Class 3	0.346	0.983	0.723	0.321
DenseNet Class 4	0.733	0.996	0.706	0.478
Overall (micro avg)	**0.876**	**0.969**	**0.972**	**0.876**

*In accordance with literature, sensitivity are used rather than recall and precision.

Train vs. Validation Loss Curve

Discussion

Evaluation Metrics

- Baseline: 67.4% sensitivity, 92.5% specificity, 70.6% macro accuracy.
- DenseNet: 87.6% sensitivity, 96.9% specificity, 87.6% macro accuracy.
- Confusion matrix: heavy on diagonals → high true positive rate.

Early-Stage Detection

- DenseNet is very good at classifying Stages 0-2 (No DR, Slight DR).
- Accuracy > 75% for Stages 0-2; high sensitivity & specificity.
- Early detection is most challenging to ophthalmologists, so tremendous potential for computer-aided diagnosis, especially results visualization.

Late-Stage Detection

- DenseNet has some difficulty with Stages 3-4, likely due to class imbalance.
- However, late-stage DR is much easier to manually/visually diagnose, so specialists can use their judgement to make an accurate diagnosis.

Interpretability

- CAMs provide a heatmap visualization of the most useful features. Clinicians can use CAMs for expedited and more accurate diagnosis.

Future Work

Model

- Train on more data: we have 44K more images, but need more computational power.
- Tune the # of frozen layers as a hyperparameter, so that trained weights are more specialized to our eye images

Computer-Aided Diagnosis

- Continue current discussions and collaborations with doctors to gauge how our CNN with CAM visualization could help them improve DR screening.
- Specialists told us they have difficulty detecting DR at early stages, and indicated that this tool could help greatly twofold .
- High-profile companies currently working on such products with doctors.
- Focus on improving CAM interpretability for maximum real-world utility.

Key References

[1] Lee, R., Wong, T. Y., & Sabanayagam, C. (2015). Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss. *Eye and vision (London, England)*, *2*, 17. doi:10.1186/s40662-015-0026-2

[2] Gardner, G. G., Keating, D., Williamson, T. H., & Elliott, A. T. (1996). Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *The British journal of ophthalmology*, *80*(11), 940-4.

[3] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).

[4] Chablani, M. (2017, August 24). DenseNet. Retrieved from https://towardsdatascience.com/densenet-2810936aeebb

[5] Ruiz, P. R. (2018, October 10). Understanding and visualizing DenseNets. Retrieved from https://towardsdatascience.com/understanding-and-visualizing-densenets-7f688092391a