

RetiNet - Feature Extractor for Learning Patterns of Diabetic Retinopathy and Age-Related Macular Degeneration from Publicly Available Datasets

Henok E. Ghebrechristos *,
Gita Alaghband, PhD
Computer Science Department
University of Colorado-Denver
Denver, CO 80204
henok.ghebrechristos@ucdenver.edu

Richard Y. Hwang, MD PhD
Department of Ophthalmology
University of Colorado School of Medicine
Aurora, CO 80048

Abstract— Diabetic Retinopathy (DR) and Age-related Macular Degeneration (AMD) are two common vision threatening eye conditions. In a large-scale screening environment DR and AMD can be assessed by detecting specific retinal findings in fundus images. In this paper, we introduce a new deep learning based feature extractor for automatic classification of DR and AMD from fundus images. We used a small dataset containing 60000 images with four severity levels of DR and two classes of AMD to design and fine-tune a deep learning model called RetiNet. This dataset, which consisted of two publicly available datasets (MESSIDOR and Kaggle), was augmented and employed to evaluate RetiNet. RetiNet can achieve diagnosis performance comparable to retina experts on the MESSIDOR dataset with cross-dataset testing (i.e., the feature extractor was trained on an independent dataset and tested on MESSIDOR). Our algorithm obtained an average accuracy of 88% on the validation set.

Keywords—Age-related Macular Degeneration (AMD), Artificial Neural Network (ANN), Convolutional Neural Network(CNN), Deep Learning, Diabetic Retinopathy(DR)

I. INTRODUCTION

Applying deep learning automation to medical image analysis such as discovering patterns of symptoms of unhealthy medical conditions and distinguishing these patterns from those that are healthy, requires domain specific and expert level understanding of the medical condition as well as thoughtful consideration of the manual procedures experts use to diagnose patients. Since its conception, deep learning has received considerable attention from the medical community and thus far has been effective in several application areas such as the prognosis of Alzheimer's disease and cognitive impairment, organ segmentation and detection [1], detecting signs of diabetic retinopathy [2] and age-related macular degeneration [3] from retinal fundus images. In this paper, we discuss our approach to automated diagnosis of eye conditions that utilizes recent advances in deep learning. Our strategy takes both the characteristics of domain data and the manual steps involved

into account. We present a deep neural network that achieves an exceptional performance in fundus image diagnosis in an experimental setting. In ophthalmology, fundus photos of the retina can be used to make diagnoses of diabetic retinopathy and macular degeneration. Although the focus of our study is abnormality detection in fundus images, the study was crafted to understand the steps required to develop a functional deep learning algorithm for medical use and steps highlighted here apply to any other medical domain in which deep learning is a potential solution. In addition, we present both the success and failures of all the approaches we took and the pros and cons of certain design and architectural considerations specific to retinal image analysis.

The goals and contributions presented in this paper are:

1. We present the design and analysis of a deep convolutional neural network (CNN, ConvNet) called RetiNet (acronym for retinal fundus image feature extraction network) and discuss the experimental results and lessons learned during the study. Fundus images and different disorders of the eye that manifest themselves in the retina make up the dataset used for the design and development of RetiNet. In ophthalmology, fundus photos of the retina can be used to make diagnosis of diabetic retinopathy and macular degeneration.
2. Several research papers [2], [3] have been published discussing the application of deep learning to the medical domain mentioned above. However, many use transfer learning (i.e. transferring parameters from a network trained on a different dataset) and fine-tuning the network to the domain task. In this regard, the contribution of this paper is to present our network design approaches that enabled our network to achieve exceptional performance using small publicly available datasets as well as to present a deep learning model, fully trained and validated from scratch, that can be used as a feature extraction base for future deep learning research on retinal fundus image classification.
3. We present the steps required to develop a functional deep learning algorithm for medical use. In addition, we present both the success and failures of approaches we took and the pros and

cons of certain design and architectural considerations specific to retinal image analysis.

4. A major goal of our study is to initiate an open source development for deep learning solutions for eye exams without monopolizing the data used and without involving hospitals or clinics for the study. To start with, we took most of the publicly available datasets and defined good model images for each class with the help of experts. Next, we applied label preserving augmentation, and preprocessing techniques to enhance our dataset both in volume and variety. This is key since we considered only publicly available datasets for the initial study.

5. Qualification of the network for clinical use is not part of this study but is part of our end goal which will involve not only more experts in the field but also validation through clinical use in clinical setting. In this regard, we take an uncommon approach to build a network that is suitable for the intended domain by making our development process data driven. Hence, another contribution of this paper is to share our approach to inspire deep learning researchers to consider our approach to relieve themselves of the burden of data when considering a similar research.

This paper continues with a description of the dataset and targeted problem domain in Section 2. Section 3 discusses the architecture of the automation algorithm followed by the training and validation methodologies used to deploy RetiNet in Section 4. Section 5 presents the dataset, preprocessing, and augmentation techniques employed to increase the volume and variety of the training set. In Section 6 we present results for the diagnosis on the various classes of diabetic retinopathy and age-related macular degeneration of the test set. Concluding remarks and discussion of future work are covered in Section 7.

II. PROBLEM DOMAIN AND SAMPLE SET DESCRIPTION

Ophthalmology and optometry are the two professions where an automation algorithm like RetiNet can expedite the traditional manual process experts employ to diagnose a patient. Since the dataset used contains retinal fundus images, the focus of the design and development of RetiNet is to build a common feature extraction base for future research involving fundus images and to elevate the ever-growing disadvantage of binding deep learning research to datasets by providing a pre-trained RetiNet to be used as an extraction base for our ongoing researches. Hence, we detail our work on characteristics of fundus images as the driver for the design of RetiNet.

Deep neural networks learn robust features from large amount of dataset and the ability to learn distinguishing factors increases when there is enough volume and variety of samples in the dataset. The dataset we used is small compared to typical datasets for deep learning models. To tackle this limitation, we have studied the image characteristics of each class in the dataset and developed algorithms for augmenting and preprocessing the dataset that results in an exceptional performance. For this publication, we consider two common eye conditions; diabetic retinopathy (DR) and age-related

macular degeneration (AMD) – the two leading causes of vision loss around the world [2] [3].

A. Diabetic Retinopathy

Diabetic retinopathy (DR) is a diabetes complication that affects the eye [10] [11]. It causes progressive damage to the retina and it can be a sight-threatening complication if left untreated. Its primary cause is the damage to blood vessels of the light-sensitive tissue in the retina. In its early stages, DR may exhibit no symptoms or only cause mild vision problems. If the disease is not diagnosed early, it can eventually cause blindness [12]. The risk of developing retinopathy increases after puberty. Twenty years after the diagnosis of diabetes, 80% of type 2 diabetics and nearly all type 1 diabetics show some signs of retinopathy [13]. Thus, most guidelines recommend annual screening for those with no retinopathy, 9 months for mild diabetic retinopathy, repeat examination in 6 months for moderate diabetic retinopathy, and closer follow up for any stage more severe than moderate diabetic retinopathy or clinically significant macular edema.

B. Age-related Macular Degeneration

Age-related macular degeneration (AMD) is the leading cause for irreversible blindness around the world [6]. AMD is caused by the deterioration of the macula [8]. Macula is a region of the retina which is responsible for focusing central vision in the eye. AMD can be classified into two main categories: exudative (wet) AMD or non-exudative (dry) AMD based on presence or absence of abnormal neovascularization. Dry AMD is characterized by drusen or abnormalities in the retinal pigment epithelium. Wet AMD is characterized by the presence of choroidal neovascularization (CNV), which can disrupt or destroy the retina or its supporting tissues. Clinically this can manifest as subretinal or intraretinal lipid, fluid or blood. Sometimes the gray-green CNV can be seen directly [9].

We designed a deep learning based automated grading algorithm that distinguishes between different severity levels of DR and classes of AMD. We discuss the details of the algorithm and training procedure in section III and IV respectively. The different severity classes of DR and conditions of AMD included in this study are summarized in Table 1.

Class	Description
0	Normal: No Diabetic Retinopathy or Macular Degeneration
1	Mild Nonproliferative DR
2	Moderate Nonproliferative DR
3	Severe Nonproliferative DR or Proliferate Diabetic Retinopathy
4	Dry AMD
5	Wet AMD

Table I: Percentage of clinical rating of diabetic retinopathy and macular degeneration

III. RETINAL FUNDUS IMAGE FEATURE EXTRACTION NETWORK (RETINET)

The RetiNet automation algorithm is based on deep ConvNet. Due to space limitations, a familiarity with Deep Convolutional Networks is assumed. [13] [17] are excellent resources for detailed coverage. Deep ConvNets are a category of Deep Neural Networks that consist of several processing layers including convolutional layers. The layers consist of several parameters that a training algorithm tunes using training dataset and ground truth label of each example in the dataset. Convolutional layer and the associated convolution filtering are two core building blocks of ConvNets. Convolution filtering modifies the special frequency characteristics of the input at each layer and construct feature maps using responses from local regions of the input. The convolution filter consists of learnable parameters and is used as a feature detector that encodes the type of feature in a feature map – an output a convolutional filter. Our network applies small filters in each layer to extract the different abnormalities from our dataset and uses a fully connected layer to distinguish between 5 classes based on the amount and severity of these abnormalities. Most deep retinal diagnosis publications have focused on transfer learning due to the lack of availability of big datasets. Those networks that rely on fine tuning use selective training approaches to reduce overfitting and to increase accuracy. Our approach attempts to characterize network relationship to data and we adhere to training from scratch without transfer learning and selective sampling.

At the start of our study, we adopted AlexNet [16]– a deep convent architecture that, for the first time, outperformed traditional machine vision techniques at visual recognition tasks–as the base model for the development of RetiNet. AlexNet is a large deep convent trained to classify about 1.3 million high-resolution images into 1000 different classes. In an annual ImageNet Large Scale Visual Recognition Challenge (ILSRVC), AlexNet achieved an error rate of 39.7%, which was considerably better than the previous top results. AlexNet has 60 million parameters and 500000 neurons. The architecture consists of five convolutional layers, some of which are followed by max pooling layer and a fully connected layer with 1000-way classification output.

During the development and validation of RetiNet we modified AlexNet architecture to fit our dataset. RetiNet consists of about 12.9 million parameters and 134000 neurons.

A. Model Architecture

RetiNet architecture consists of 5 convolutional layers followed by rectified linear unit (ReLU) non-linearity for response normalization. Each convolution layer applies a 3x3 filter to the input. As depicted in Fig. 6, there are two variations in how convolution is applied to the input data. Inspired by AlexNet [16], RetiNet uses Net A and B in alternation to extract features of interest from the input. Two stacks of nets A and B are employed in RetiNet for maximum classification accuracy. The first convolutional layer of Net A filters the 512x512 input image using a 3x3 filter with a stride of 1 pixels. We

experimented with different values of stride and the most stable results were achieved when the stride values of all convolutional filters were set to 1. The second convolution layer takes as input a pooled and response-normalized feature maps of the first layer and filters them with a kernel of size 3x3. The outputs of the second convolution layer are then pooled and passed to the first convolution layer of Net B which performs two convolutional passes with a kernel size of 3x3 and outputs 192 feature maps. The second convolution layer of Net B convolves the feature maps generated by the first layer with a kernel size of 3x3 at a stride 1. The final convolution layer outputs a pooled feature map convolved with a kernel size 3x3 and passes the 256 feature maps to the classification network.

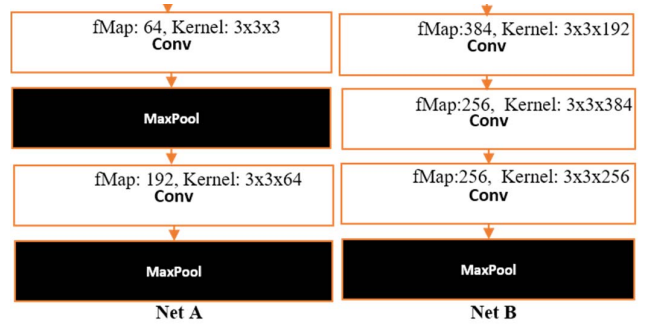


Fig. 6. RetiNet convolutional layer stacks that serve as feature extraction base for DR and AMD classification

The last layer of Net B is connected to a 6-way fully connected (FC) classification network which produced a probabilistic distribution over the six fundus severity classes. RetiNet minimizes training loss by computing the categorical cross-entropy loss between the output of the network and the expected output.

AlexNet applies a 7x7 filter to the input image and uses smaller filters as the input is downsized in lower layers. We attempted this configuration on our dataset and the network stagnated at error rates up to 62%. Unlike AlexNet, RetiNet is designed to capture small and irregular features that distinguish normal retina image from the other classes. For this reason, we used filters of size 3x3 in all convolutional layers with minimal downsizing of the input.

IV. TRAINING AND VALIDATION

RetiNet was trained using stochastic gradient descent (Algorithm 2) with a minimum batch size of 16 examples. The momentum used was equal to 0.9 while weight decay of 0.0005. A small amount of weight decay is important to regularize the network and as well as to reduce training errors. For every training iteration, learning rate, and momentum, RetiNet applies the following update rule for each weight.

$$v_{i+1} = 0.9v_i - 0.0005\alpha (w_i) - \alpha \left\langle \frac{\partial L}{\partial w} |w_i \right\rangle B_i \quad (2)$$

$$w_{i+1} = w_i + v_{i+1} \quad (3)$$

where $\left\langle \frac{\partial L}{\partial w} |w_i \right\rangle B_i$ is the average over the i th training batch of the derivative of the objective function with respect to w , evaluated at w_i [16].

V. DATASET

Deep learning methods are generally effective when trained on large datasets. Due to the lack of publicly available datasets for diagnosing retinal fundus image we have employed several augmentation techniques to increase the amount and variety of learned features for better generalization. A total of 62578 images were used to train and validate the model. The number of images in each set are listed in Table II. The severity classification of each image in the data set is provided in Table I. Since the images were collected from various sources, the dataset contained resolutions of varying degrees which required preprocessing for normalization.

A. Dataset Sources

1) Kaggle

The Kaggle dataset consist of 35,126 training images graded into five DR stages and 53,576 test images with undisclosed DR stage. Images were acquired using multiple fundus cameras and different field of view. Details about image acquisition, such as camera type and field of view, are not revealed. More information about the data can be found in the Kaggle Retinopathy challenge website.

2) Messidor

The publicly available Messidor database consists of 1200 images acquired at three different sites. Images were acquired using a color video 3CCD camera on a Topcon TRC NW6 non-mydratric retinograph with a 45-degree field of view. The images have resolutions of 1440 960, 2240 1488 or 2304 1536 pixels. The Messidor set was exclusively used as an independent set for validating RetiNet performance on DR.

3) UCH-AMD

This dataset was obtained from the Ophthalmology Department at Anschutz Medical Campus. The set consists of 197 images of Wet and Dry AMD as well as fundus images with vein occlusions. Only the AMD samples were used for this study. This study received approval from the Colorado Multiple Institutional Review Board (COMRBB).

Grade	Number of Samples
0	10784
1	10886
2	10584
3	10162
4	10042
5	10120
Total	62578

Table II: Grade distribution in training and validation samples

B. Preprocessing

To increase the number of features to be learned, the images in the data set were preprocessed for noise reduction, to increase contrast between background and retina, and to increase illumination by equalizing color distribution. The techniques for preprocessing include resizing each image to 512x512, applying a digital filter to enhance the image quality by

reducing noise and histogram equalization. In addition, to make the resolution of the retinal fundus image consistent across all examples and to increase the number of samples to avoid overfitting, we employed a subsampling algorithm and augmentation technique discussed below.

1) Histogram Equalization

Histogram Equalization was a preprocessing step taken to adjust the contrast using each channel's intensity distribution. Because of this preprocessing the dark area in the input retina image with low illumination becomes brighter, while the brighter areas are normalized to maintain consistent contrast across the entire region of the image.

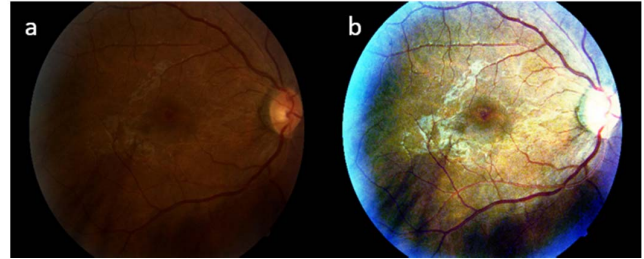


Fig. 7. a) Original sample and b) the sample after histogram normalization

C. Subsampling

As a measure to battle overfitting, both left and right retinal fundus model images were defined to allow filtering of redundant and defective samples that do not fall within a certain threshold. This is a form of in-training selective sampling used in [16] but we perform the subsampling as a preprocessing step. The iterative algorithm for selecting samples follows these steps:

Require: Input dataset

1. Choose a sample x from Dataset
2. **If** x is left $\rightarrow dif = comp(x, model_left)$
3. **If** x is right $\rightarrow dif = comp(x, model_right)$
4. **If** $dif < threshold \rightarrow$ include sample
5. **If** $dif > threshold \rightarrow$ exclude sample

Algorithm 2: Selective sampling procedure to clean and normalize dataset

The function, *comp*, computes the percentage difference between a sample and the model images. Images having more than 20% *diff* value were excluded from training set. This preprocessing step ensures consistency in size and resolution of the retinal fundus relative to the background across all samples in training and test sets.

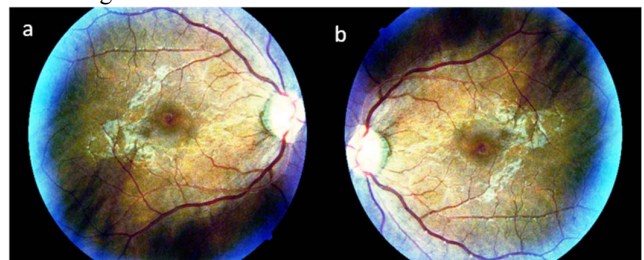


Fig. 8: a) right lateral retinal fundus model image and b) left lateral retinal fundus model image

In addition, to minimize the number of parameters and to avoid unnecessary input to our network from the background of the

images, we cropped a rectangle containing the large patch of the retina. The cropping was performed in four regions of the original image. We made each crop differ from the previous by 3 pixels horizontally and vertically to generate four samples per each sample. This increased the test accuracy of the network by 10% and the volume of the dataset by a factor of four.

D. Augmentation

Performance of RetiNet was poor when trained on our original small dataset which also contained a non-normal distribution of sample labels. The network learned enough features from the normal retinal fundus images which constituted 60% of the dataset and the learning curve stagnated at 60%. We knew the amount of data used was not enough since deep networks need to be trained on a vast number of training images to avoid getting stuck in local minima. To achieve a satisfactory training performance, we applied label-preserving data augmentation techniques. These techniques exploited special invariance of the samples. The techniques – rotation, cropping, and zooming and color perturbation – allow transformed images to be produced from each original sample with very little deviation from the original but with enough distinguishable factors. With this RetiNet could overcome overfitting.

VI. RESULTS

The training and validation of RetiNet was performed on a machine with an Intel Xeon® CPU. It also consisted of one NVIDIA GeForce GTX 1080 with 3840 cores, 320 GB/s memory bandwidths and a core clock of 1126 MHz which was heavily utilized for training and validation. Microsoft’s deep learning framework, Cognitive Toolkit (CNTK) was chosen for the design, training and validation of RetiNet. We used the C# API to evaluate RetiNet’s performance on all Messidor and AMD samples. The dataset containing samples across all DR and AMD conditions is summarized below.

Grade	Number of Samples
0	479
1	133
2	219
3	169
4	30
5	30
Total	1060

Table III: Grade distribution in test set

A. RetiNet Training Performance

During training, performance was measured on the monitoring set during the CNN training process. Fig. 9 shows the accuracy values measured on image level as function of the number of training epochs. An epoch is single forward and backward pass of the network on all training samples [13]. The performance of increased over time and finally converged to a stable maximum accuracy of 99.875%. This maximum performance was achieved after 60 training epochs.

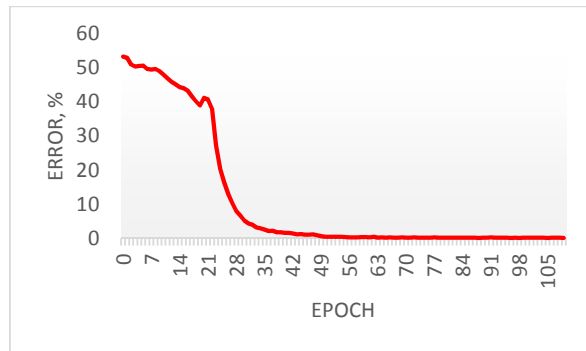


Fig. 9. Training performance of RetiNet

B. Retinopathy Grading Performance

RetiNet model trained for 112 epochs was used to test the prediction performance. Two performance metrics were used to assess the ability of RetiNet to correctly classify retinal fundus images into grades listed in Table I. We computed the sensitivity and accuracy of the network. The sensitivity and accuracy can be computed as:

$$\text{Sensitivity, \%} = \frac{T_P}{T_P + F_N} \quad (4)$$

$$\text{Accuracy, \%} = \frac{T_P + T_N}{T_P + F_N + F_P + T_N} \quad (5)$$

where T_P is the number of abnormal retinal fundus images classified as abnormal, T_N is the number of normal fundus images found as normal, F_P is the number of normal images classified as abnormal and F_N is the number of abnormal images classified as normal. In addition, the confusion matrix of the network over the entire dataset is presented in Table IV. The sensitivity values also correspond with what is known clinically. Categories 0-4 performed with similar sensitivities. This is consistent with the idea that these diagnoses can be made with fundus photography alone. However, category 5, or wet macular degeneration has a lower sensitivity than the other categories. This can be explained because wet AMD can have subtle findings on fundus photography that can be difficult to see because the 3-dimensional picture of the retina can be hard to see from a fundus photograph. In clinical practice, adjunctive testing with fluorescein angiography or optical coherence tomography are often used to help make the diagnosis by imaging CNV or visualizing the 3D structure of the retina.

Output	Grades						Total
	0	1	2	3	4	5	
Positive	450	110	189	150	25	16	940
Negative	29	23	30	19	5	14	120
Total	479	133	219	169	30	30	1060

Table IV: Prediction performance of RetiNet

N = 1060	0	1	2	3	4	5
0	450	0	0	29	0	0
1	18	110	5	0	0	0

2	12	18	189	0	0	0
3	14	0	2	150	3	0
4	5	0	0	0	25	0
5	11	5	0	0	0	16

Table V: Confusion matrix of RetiNet on hold out dataset. The total number of samples are N = 1060

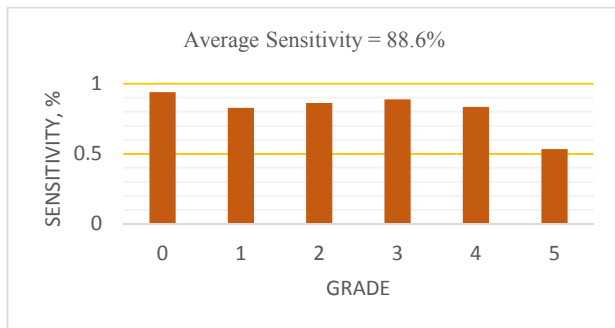


Fig. 10. RetiNet grade sensitivity on test samples.

VII. CONCLUSION

Diabetic Retinopathy and Age Related Macular Degeneration are two diseases which can cause vision loss rapidly. In this paper, we presented the design and development of a network which achieved training performance of 43% error rate when trained using of the shelf publicly available datasets. To increase the performance, we improved the quality of the dataset by applying preprocessing techniques such as histogram normalization, cropping, and subsampling. The preprocessing techniques alone were not enough to achieve reasonable performance; thus, the set was augmented to increase in volume and variety. The augmentation techniques included mirroring and other label-preserving image processing techniques. After adding a considerable number of additional samples to the dataset, features were then extracted using deep convolutional network called RetiNet, consisting 5 convolutional filters with 3x3 kernel size. Extracted feature were then used to fine tune a fully connected classification network that minimized a loss function. To measure its effectiveness, we applied RetiNet to a set of unseen DR and AMD samples and achieved 88% correct prediction rate.

VIII. REFERENCES

- [1] J.-Z. Cheng, Dong Ni, Yi-Hong Chou, Jing Qin, Chui-Mei Tiu, Yeun-Chung Chang, Chiun-Sheng Huang, Dinggang Shen, Chung-Ming Chen "Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans," *Sci. Rep.*, vol. 6, p. 24454, Apr. 2016.
- [2] Gulshan, Varun, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, et al. 2016. "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs." *JAMA*, November. doi:10.1001/jama.2016.17216.
- [3] P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson, and N. M. Bressler, "Detection of age-related macular degeneration

via deep learning," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 184–188.

- [4] *Retina and Vitreous*, American Academy of Ophthalmology: The Eye M.D. Association. 2011.
- [5] M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis," *IEEE Trans. Med. Imaging*, vol. 3, pp. 169–208, Jan. 2010.
- [6] A. C. Ho and C. D. Regillo, Eds., *Age-related Macular Degeneration Diagnosis and Treatment*. New York, NY: Springer New York, 2011.
- [7] M. Bhaskaranand Jorge Cuadros, Sandeep Bhat, Muneeswar G. Nittala, "EyeArt+ EyePACS: Automated Retinal Image Analysis for Diabetic Retinopathy Screening in a Telemedicine System," 2015.
- [8] R. Priya and P. Aruna, "Diagnosis of diabetic retinopathy using machine learning techniques," *J. Soft Comput.*, vol. 3, no. 4, pp. 563–575, 2013.
- [9] Giancardo, Luca, Fabrice Meriaudeau, Thomas P. Karnowski, Yaqin Li, Seema Garg, Kenneth W. Tobin, and Edward Chaum. 2011. "Exudate-Based Diabetic Macular Edema Detection in Fundus Images Using Publicly Available Datasets." *ResearchGate* 16 (1): 216–26. doi: 10.1016/j.media.2011.07.004.
- [10] "What Is Diabetic Retinopathy?" *American Academy of Ophthalmology*, 01-Sep-2013. [Online]. Available: <http://www.aao.org/eye-health/diseases/what-is-diabetic-retinopathy>. [Accessed: 17-Feb-2017].
- [11] "Diabetic Retinopathy for Medical Students. EyeRounds.org - Ophthalmology - The University of Iowa." [Online]. Available: <http://webeye.ophth.uiowa.edu/eyeforum/tutorials/diabetic-retinopathy-med-students/index.htm>. [Accessed: 28-Feb-2017].
- [12] R. Klein, B. E. Klein, S. E. Moss, M. D. Davis, and D. L. DeMets, "The Wisconsin epidemiologic study of diabetic retinopathy. III. Prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years," *Arch. Ophthalmol. Chic. Ill 1960*, vol. 102, no. 4, pp. 527–532, Apr. 1984.
- [13] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1153–1159, May 2016.
- [14] M. Lai, "Deep Learning for Medical Image Segmentation," *ArXiv150502000 Cs*, May 2015.
- [15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [16] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1273–1284, May 2016.
- [17] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," *ArXiv12065533 Cs*, Jun. 2012.