Retinal Image Classification via Vasculature-guided Sequential Attention

Mengliu Zhao and Ghassan Hamarneh
Medical Image Analysis Lab
School of Computing Science
Simon Fraser University, Canada
{mengliuz, hamarneh}@sfu.ca

Abstract

Age-related macular degeneration and diabetic retinopathy are diseases of increasing prevalence globally in recent years. Traditionally, diagnosing these diseases relied on manual visual inspection by experts, which was costly, time-consuming and laborious as it required closely examining high-resolution color fundus images. More recently, deep learning networks have shown great potential in predicting diseases from retinal images. However, being purely data-driven, these networks are susceptible to overfitting and their training requires large annotated data. In this paper, we propose to enrich deep learning-based fundus image classifiers with prior knowledge on special structures in the retina implicated with the disease. In particular, we leverage vessel priors to guide the attention mechanism of deep learning architectures. In addition, we leverage a bi-directional dual-layer LSTM module to learn the inter-dependencies between a sequence of prior-guided attention maps deployed across the depth of the disease classification network. Results on the clinical datasets show the proposed method could bring performance improvement by as much as 8%.

1. Introduction

1.1. Motivation

The prevalence of eye diseases has been on the rise during the past years, both globally and regionally. According to a recent Lancet publication [3], more than 216 million people suffer from moderate to severe visual impairment. The fact sheet from the US National Eye Institute shows around 1.3 million of Americans are blind and the figure is expected to rise to 2.2 million by 2030 [14]. Two of the leading causes of blindness are age-related macular degeneration (AMD) and diabetic retinopathy (DR) [14]. Early diagnosis and treatment of these diseases are crucial in vision preservation, which makes automatic and accurate classification of retinal images extremely important [18, 19].

1.2. Machine (and deep) learning for retinal image classification

Limited attempts to build automatic classification systems have been made using traditional machine learning methods that relied on hand-crafted features. In the work of Roychowdhury et al. [17], AdaBoost was used for feature reduction in a two-step hierarchical binary DR classification (with DR or without) approach albeit with low specificity (53%) according to their reported results. Wang et al. [22] proposed to combine multi-scale features and feature selection algorithms for AMD classification, but their work focused on optical coherence tomography images, and use a fairly small dataset with only 45 patients, while multiple images come from the patient at different scans.

The recent success of deep learning-based visual recognition for numerous applications has sparked renewed interest in addressing the task of retinal disease classification and grading from fundus images. In the work of Gulshan et al. [8], the Inception network was used for retinopathy grading (5 levels). Pratt et al. [16] used a 13-layer convolutional neural network (CNN) for retinopathy grading (5 levels). In the work of Gargeya et al. [7], a deep network with 5 residual blocks was constructed for feature generation, then the output feature is input into a decision tree with other metadata information for binary retinopathy classification.

Previous fundus imaging-based deep learning methods for AMD and DB classification and grading relied on a purely data-driven tuning of network architectures yet lacked any disease-specific customization to encode existing prior knowledge, such as anatomical structural changes associated with the progression with of specific diseases. One form of AMD, the wet AMD has been known to be associated with abnormal growth of blood vessels in the eyes [5]. By examining retinal photography images, McGowan et al. [12] discovered correlations between AMD and, not only blood vessels around the macular area but

also the blood vessel caliber across the whole fundus image. Also, a recent study by Jackson et al. [9] discovered a high prevalence of vascular abnormalities in conjunction with AMD. However, to the best of our knowledge, automatic deep learning AMD classification methods completely ignored any vascular priors. On the other hand, DR is caused by retinal blood vessel changes due to diabetes and, to a certain extent, could also be linked to vessel overgrowth on the retina [6, 10]. This literature shows how the development of AMD and DB is highly correlated with changes in retinal blood vessel structures and suggests that automatic methods, deep learning or otherwise, could leverage such prior information.

1.3. Attention mechanisms in deep learning

Several deep learning methods with attention mechanism have been proposed in the past few years. Wang et al. [21] proposed to add intermediate deconvolution layers to extract attention maps, combining them with the last layer for prediction. In the works of Mnih et al. [13] and Xu et al. [23], attention features were extracted from different locations within an image and stacked into a sequence that is fed into an LSTM framework. Similarly, to handle cancer classification from large histopathology images, BenTaieb and Hamarneh proposed an attention mechanism that adaptively selects only a limited sequence of image locations for further processing [2]. But none of these work leveraged any disease-specific prior information. For diagnosing melanoma, Yan et al. used skin lesion masks to guide attention maps across different layers of the VGG architecture [24], but their work relied on expert-delineated (not automatically-generated) prior masks, nor did they learn the patterns of a sequence of attention maps.

1.4. Contributions

Our work is the first:

1. To leverage anatomical knowledge (in the form of vascular priors) to guide the attention maps for retinal disease classification from fundus images;

2. To automatically extract the attention prior maps (rather than requiring manually-segmented images);

3. To encode the inter-dependency among attention features (deployed across the depth of the network), which we accomplish via a novel bi-directional, dual-layer LSTM.

We perform evaluation on two clinical datasets with cross-validation and an ablation study. The experimental results show that, by using the proposed vasculature priors and the LSTM attention formulation, the results are improved by as much as 8%.

2. Proposed Method

The proposed architecture is illustrated in Figure 1a. In the following we describe the baseline architecture (Section 2.1), the proposed vasculature priors (Section 2.2), LSTM module (Section 2.3) and the corresponding loss functions (Section 2.4).

2.1. Baseline CNN with attention modules

We adopt the baseline CNN architecture proposed by Yan et al. [24], which extends VGG-16 [20] with two additional attention layers and one penultimate global feature vector (obtained via global average pooling). The attention features and global feature vector are combined and input into one dense layer for classification. We denote the intermediate features generated from n different intermediate convolutional layers as $\{F^1, F^2, ..., F^n\}$, where $F^k = \{f_1^k, f_2^k, ..., f_n^k\}$, and f_i^k is the feature of channel i. We further denote the global feature from the last convolutional layer as G, the layer-wise attention maps as $\{M_{attn}^1, M_{attn}^2, ..., M_{attn}^n\}$, and the corresponding attention feature vectors as $\{A^1, A^2, ..., A^n\}$.

2.2. Vasculature priors on attention

A prior image M_{prior} is a binary image mask used to guide the attention features. M_{prior} can be generated using an automatic method or manual delineation. In this work, our goal is to use the vasculature structure to guide the attention feature maps. As expert manual delineation is time-consuming, we set M_{prior} to be the retinal vasculature mask extracted from the input image using the automatic B-Cosfire vessel filtering [1], which is based on calculating the geometric mean of multiple difference of Gaussian filters. Examples of the generated vessel masks could be found in Figure 2.

To guide attention feature maps across different scales, M_{prior} is processed with adaptive average pooling into different sizes as $\{M_{prior}^1, M_{prior}^2, ..., M_{prior}^n\}$ (see details in Yan et al. [24]), corresponding to the sizes of the attention maps $\{M_{attn}^1, M_{attn}^2, ..., M_{attn}^n\}$. M_{attn}^i is guided by M_{prior}^i, for $i = 1 \cdots n$, by maximizing the similarity between the two via a minimizing Dice-based loss (defined in Section 2.4).

The attention feature vector A^k is calculated as follows:

$$\mathcal{G} = \text{Upsample}(W_g \otimes G) \quad (1)$$

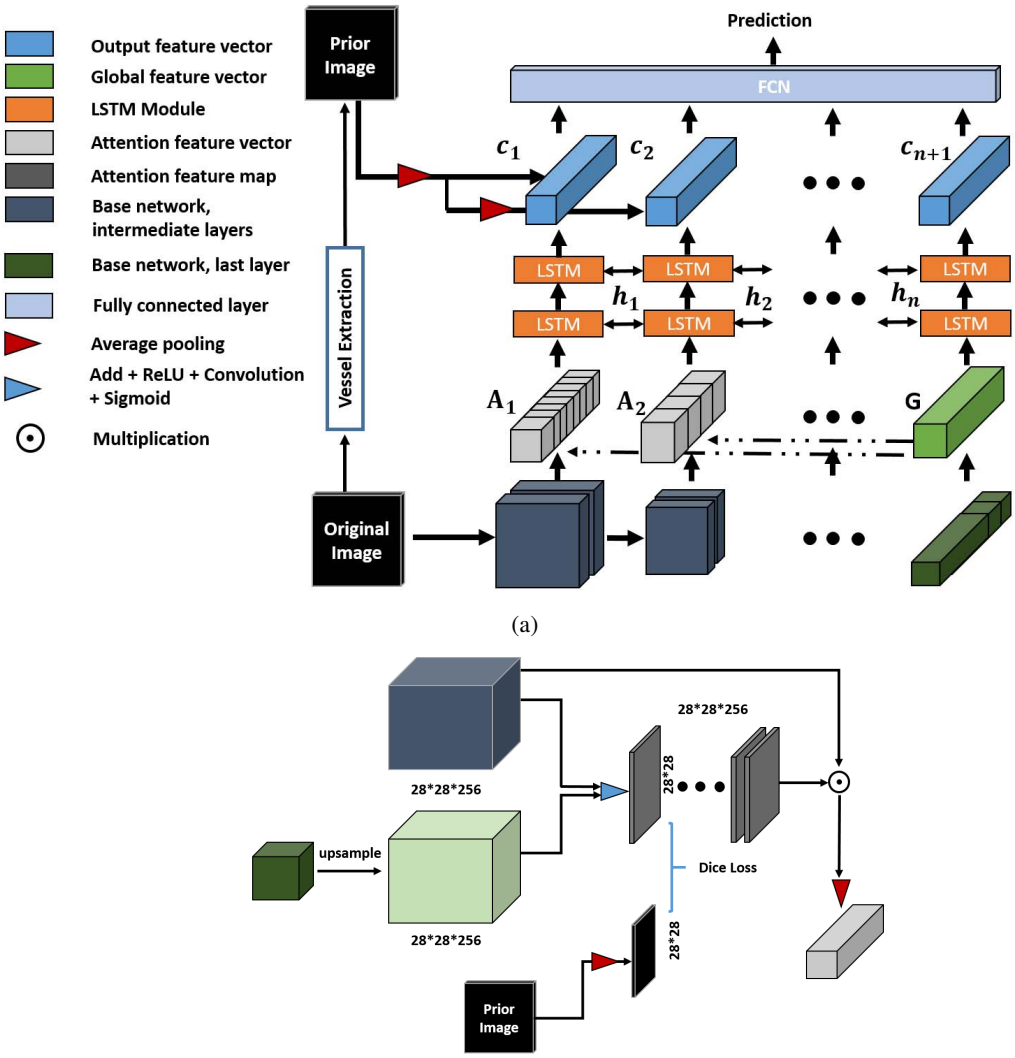
$$\mathcal{F}^k = W_f \otimes F^k \quad (2)$$

$$M_{attn}^k = \sigma(W \otimes ReLU(\mathcal{F}^k + \mathcal{G})) \quad (3)$$

$$a_i^k = M_{attn}^k \odot f_i^k \quad (4)$$

$$A^k = \text{GlobalAveragePool}\{a_1^k, a_2^k, ..., a_n^k\} \quad (5)$$

where W_g, W_f, W are convolutional filter weights, σ is the sigmoid function, and $\otimes$ and $\odot$ are the convolution



(a)

(b)

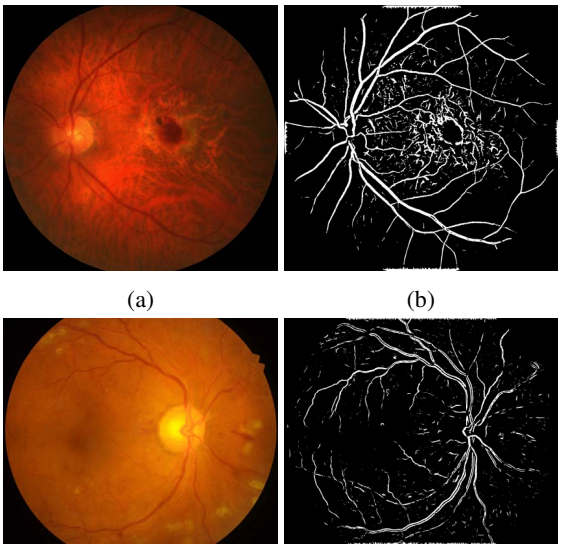
Figure 1: (color figure) (a) Proposed architecture. Attention features (light gray boxes) and global features (green box) are stacked into a sequence and input into a bidirectional dual-layer LSTM, and then input into a classification dense layer. (b) Example of using prior image to guide attention map in layer 4.

and element-wise multiplication operators, respectively. In Equation 1, global feature is convolved by a 256 channel filter and upsampled to match the size of F^k. In Equation 2, intermediate feature F^k is convolved with a 256 channel filter. In Equation 3, the attention map is generated using $\mathcal{G}$ and $\mathcal{F}^k$ from the previous two steps. Then in Equation 4, the attention map is multiplied with intermediate feature f_i^k at channel i on an element-wise way, to preserve the intermediate information. Then the attention vector A^k is obtained by global average pooling of $\{a_1^k, a_2^k, ..., a_n^k\}$.

2.3. Learning the sequence of prior-guided attention maps

In the work of Yan et al. [24], A^1, A^2, G were aligned into a single vector before being input into a dense layer for classification. In contrast, we wish to encode the inter-dependency among all the n learnt attention feature vectors across all layers, A, and the global feature vector G, using a bi-directional dual-layer LSTM model.

In a LSTM module, given inputs x_t, h_{t-1}, c_{c-1} at time step t, $W_{xi}, W_{hi}, W_{xf}, W_{hf}, W_{xo}, W_{ho}, W_{xc}, W_{hc}$ the learnt weights, and b_i, b_f, b_o, b_c the corresponding bi-



(a) (b) (c) (d)

Figure 2: Examples of AMD and retinopathy images and corresponding vessel masks: a) AMD image; b) AMD image vessel mask; c) retinopathy image; d) retinopathy image vessel mask.

ases, the following update scheme is used:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (6)$$
$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (7)$$
$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (8)$$
$$g_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (9)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (10)$$
$$h_t = o_t \odot \tanh(c_t). \quad (11)$$

To this end, $x_1, x_2, ..., x_t$ in Equation 6-9 are replaced by $A^1, A^2, ..., G$. To deal with different sizes of A^k, we tile A^k until they are of the same size as G. The output of the LSTM is then fed into the classification layer (Figure 1a). Illustration of inputs and output of LSTM could be found in Figure 1a)

2.4. Loss functions

Similar to Yan et al. [24], we first use a modified cross entropy loss, called focal loss [11] with $\gamma = 2.0$, to deal with class imbalance:

$$L_F = -(1 - p_i)^\gamma \log(p_i) \quad (12)$$

where p_i is the estimated probability for the class with label i. In a N-class classification problem, p_i has to satisfy $\sum_{i=1}^{N} p_i = 1$.

As described in Section 2.2, we wish to use vessel priors to guide the learning of attention layers (Figure 1b). To this end, we define the following prior loss, based on the Dice similarity coefficient:

$$L^i_{DSC} = 1 - 2\frac{|M^i_{attn} \cap M^i_{prior}|}{|M^i_{attn}| + |M^i_{prior}|}. \quad (13)$$

where M^i_{attn} is the intermediate feature at layer i and M^i_{prior} the corresponding prior mask.

The final loss is the sum of the focal loss and the weighted prior loss, for the total of n attention maps:

$$L = L_F + \sum_{i=1}^{n} w_i \cdot L^i_{DSC}. \quad (14)$$

2.5. Implementation details

For the proposed architecture, we use Adam optimizer with $\beta = (0.9, 0.999)$ and set the initial learning rate to 10^{-4}, weight decay ratio $\epsilon = 10^{-8}$, total epoch $= 20$, and batch size $= 20$. The weight loss w_i in (14) is set empirically to 0.1 for the AMD dataset and 0.001 for the retinopathy dataset. Finally, the parameters for B-Cosfire filter are summarized in Table 1.

3. Experiment

We test the proposed method on two public datasets (iChallenge-AMD and IDRiD)) and evaluate the performance of competing methods using Accuracy, Precision, Recall and F1-score. We also perform ablation studies assessing the value of the vessel priors and the LSTM formulation.

3.1. Datasets:

3.1.1 iChallenge-AMD dataset

We obtained training data from iChallenge-AMD, a recent AMD classification challenge, which contains 398 images: 87 AMD images and 311 non- AMD images, for the purpose of binary classification. All images are color fundus images of resolution $2124 * 2056$. We performed 3-fold cross validation and augmented the training set 2 times by random cropping, scaling and rotation.

3.1.2 IDRiD dataset

We obtained 516 images from this dataset: 413 training and 103 testing data, for the purpose of retinopathy grading. International 5-level diabetic retiopathy (DR) grading is provided: (i) no apparent retinopathy, (ii) mild non-proliferative DR, (iii) moderate non-proliferative DR, (iv) severe non-proliferative DR and (v) proliferative DR. All images are color fundus images of resolution $4288 * 2848$.

	σ		ρ		σ_0		α	
	sym	asym	sym	asym	sym	asym	sym	asym
Retinopathy	5.0	5.0	25	24	1	1	0.1	0.1
AMD	5.0	5.0	20	22	1	2	0.1	0.1

Table 1: Parameters for B-Cosfire vessel filter. Sym: symmetric filter parameters; asym: asymmetric parameters. See corresponding parameter meaning in the the original paper [1].

3.1.3 Dealing with the class imbalance problem

For the AMD dataset, since 311 non-AMD images and 87 AMD images were provided as the training set, we oversample AMD image by 3 times to avoid class imbalance. For retinopathy dataset, we oversample class (ii) seven times, class (iv) two times and class (v) three times.

3.2. Results

3.2.1 Baseline experiments

We carry out baseline experiments using the architecture in the work of Yan et al. [24] without prior information or LSTM, and only with two layers (layer 3 and 4) of attention features.

3.2.2 Assessing the advantage of using the vessel prior

To test the hypothesis that adding vessel prior would help guiding the attention map, we compare the baseline architecture with and without vessel priors.

From Table 2 we see that adding vessel prior to the baseline architecture improves the prediction results by as much as 2%.

Comparing rows $1-2$, $4-5$, $7-8$ and $10-11$ in Table 3 shows that simply adding the vessel prior to the baseline architecture improves the results of almost all evaluation metrics by as much as 6%. The improvement was observed regardless of the number of layers equipped with an attention module.

3.2.3 Assessing the advantage of using the LSTM

Here we set out to evaluate whether the proposed LSTM module could leverage the inter-dependency of the attention sequence.

From Table 2 we see that by adding the LSTM module (proposed), the precision for predicting AMD is improved by as much as 6%.

Comparing rows $2-3$, $5-6$, $8-9$ and $11-12$ in Table 3 shows that, when the length of the attention sequence is no less than 2 (attention from more than 2 layers), by incorporating LSTM, the results improve by at least 3% for all metrics and as much as 6%.

Furthermore, comparing rows $2-3$ from Table 3 shows that, when the attention is limited to only one layer (layer 4), LSTM is no longer effective since there is limited sequential information in the training data.

3.2.4 Overall performance of the proposed method

By combining the vessel prior and the LSTM module for learning inter-dependency from the attention sequence, Table 2 and Table 3 (comparing rows $1-3$, $4-6$, $7-9$ and $10-12$) both show the results are improved by as much as 8%.

3.2.5 Comparing with the state-of-the-art

Comparing the results of our proposed method to state-of-the-art AMD classification methods [15, 4], and to the DR grading accuracy results reported on the challenge, our proposed method achieves comparable accuracy. However, none of the state-of-the-art methods leverage any attention mechanism, so we expect that their performance will improve with vessel-guided priors and attention sequence modelling, similarly to how our baseline methods improved with these extensions.

Furthermore, as shown by Yiqi et al. [24], including priors can help rendering the regions relevant for classification, thus contributing to a more intuitive and interpretable deep learning model.

4. Conclusion and future work

In this paper we propose a new architecture using vessel prior to guide the attention sequence in deep learning networks. To leverage the inter-dependency among the attention sequence, a bi-directional dual-layer LSTM module is used. Experiments using two clinical dataset, with binary AMD classification and 5-level retinopathy grading tasks, clearly demonstrate the advantages of the proposed architecture. Moreover, our ablation study with both datasets show how the proposed architecture works with varying lengths of the attention sequence, which could be easily extended when there is an even deeper network involved. Numerical results with multiple evaluation metrics are reported and the performance improvement produced by the proposed architecture reaches as much as 8%.

Future work will be applying the vessel prior guided attention sequence mechanism to other applications, such as multi-class AMD grading, as well as other CNN-based architectures with more layers involved.

Acknowledgments. Partial funding for this project is provided by the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors are grateful to the Nvidia Corporation for donating a Titan X GPU used in this research and the support provided by Compute Canada. The authors also thank members of the Medical Image Analysis Lab, Saeed Izadi, KumarAbhishek, and Anmol Sharma for their contributions towards converging on the title of this paper.

References

[1] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov. Trainable cosfire filters for vessel delineation with application to retinal images. *Medical Image Analysis*, 19(1):46–57, 2015. 2, 5

[2] A. BenTaieb and G. Hamarneh. Predicting cancer with a recurrent visual attention model for histopathology images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 129–137. Springer, 2018. 2

[3] R. R. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. *The Lancet Global Health*, 5(9):e888–e897, 2017. 1

[4] 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)*, pages 184–188. IEEE, 2016. 5

[5] A. V. Chappelow and P. K. Kaiser. Neovascular age-related macular degeneration. *Drugs*, 68(8):1029–1036, 2008. 1

[6] D. S. Fong, L. Aiello, T. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, F. L. Ferris, and R. Klein. Retinopathy in diabetes. *Diabetes Care*, 27(suppl 1):s84–s87, 2004. 2

[7] R. Gargeya and T. Leng. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*, 124(7):962–969, 2017. 1

[8] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA: The Journal of the American Medical Association*, 316(22):2402–2410, 2016. 1

[9] T. L. Jackson, R. P. Danis, M. Goldbaum, J. S. Slakter, E. M. Shusterman, D. J. O'Shaughnessy, and D. M. Moshfeghi. Retinal vascular abnormalities in neovascular age-related macular degeneration. *Retina*, 34(3):568–575, 2014. 2

[10] R. Klein, B. E. Klein, S. E. Moss, T. Y. Wong, L. Hubbard, K. J. Cruickshanks, and M. Palta. Retinal vascular abnormalities in persons with type 1 diabetes: the wisconsin epidemiologic study of diabetic retinopathy: Xviii. *Ophthalmology*, 110(11):2118–2125, 2003. 2

[11] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2980–2988, 2017. 4

[12] A. McGowan, G. Silvestri, E. Moore, V. Silvestri, C. C. Patterson, A. P. Maxwell, and G. J. McKay. Retinal vascular caliber, iris color, and age-related macular degeneration in the irish nun eye study. *Investigative Ophthalmology & Visual Science*, 56(1):382–387, 2015. 1

[13] V. Mnih, N. Heess, A. Graves, et al. Recurrent models of visual attention. In *Advances in Neural Information Processing Systems*, pages 2204–2212, 2014. 2

[14] National Eye Institute. Eye disease statistics factsheet. 2014. 1

[15] T. V. Phan, L. Seoud, H. Chakor, and F. Cheriet. Automatic screening and grading of age-related macular degeneration from texture analysis of fundus images. *Journal of Ophthalmology*, 2016, 2016. 5

[16] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng. Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science*, 90:200–205, 2016. 1

[17] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi. Dream: diabetic retinopathy analysis using machine learning. *IEEE Journal of Biomedical and Health Informatics*, 18(5):1717–1728, 2013. 1

[18] H. Safi, S. Safi, A. Hafezi-Moghadam, and H. Ahmadieh. Early detection of diabetic retinopathy. *Survey of Ophthalmology*, 63(5):601 – 608, 2018. 1

[19] R. Schwartz and A. Loewenstein. Early detection of age related macular degeneration: current status. *International Journal of Retina and Vitreous*, 1(1):20, 2015. 1

[20] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 2

[21] W. Wang and J. Shen. Deep visual attention prediction. *IEEE Transactions on Image Processing*, 27(5):2368–2378, 2017. 2

[22] Y. Wang, Y. Zhang, Z. Yao, R. Zhao, and F. Zhou. Machine learning based detection of age-related macular degeneration (amd) and diabetic macular edema (dme) from optical coherence tomography (oct) images. *Biomedical Optics Express*, 7(12):4928–4940, 2016. 1

[23] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio. Show, attend and tell: Neural image caption generation with visual attention. In *International Conference on Machine Learning*, pages 2048–2057, 2015. 2

[24] Y. Yan, J. Kawahara, and G. Hamarneh. Melanoma recognition via visual attention. In *International Conference on*

Methods	Vessel Prior	LSTM	Accuracy	Precision		Recall		F1-score	
				AMD	NAMD	AMD	NAMD	AMD	NAMD
[24]	✗	✗	$94 \pm 2\%$	$89 \pm 6\%$	$\mathbf{95 \pm 2\%}$	$81 \pm 7\%$	$97 \pm 2\%$	$85 \pm 4\%$	$96 \pm 1\%$
[24] + vessel prior	✓	✗	$94 \pm 2\%$	$91 \pm 4\%$	$\mathbf{95 \pm 2\%}$	$\mathbf{82 \pm 5\%}$	$98 \pm 1\%$	$86 \pm 5\%$	$96 \pm 2\%$
Proposed	✓	✓	$\mathbf{95 \pm 3\%}$	$\mathbf{97 \pm 5\%}$	$\mathbf{95 \pm 2\%}$	$82 \pm 10\%$	$\mathbf{99 \pm 1\%}$	$\mathbf{89 \pm 6\%}$	$\mathbf{97 \pm 2\%}$

Table 2: Testing result on AMD dataset with binary classification. AMD: images labeled as age-related macular degeneration; NAMD: without disease. Attention comes from layer $3-4$ of the architecture (see details of the baseline architecture in [24]). Mean $\pm$ standard deviation are reported among the 3 groups for cross-validation.

Row	Methods	Attention	Vessel Prior	LSTM	Accuracy	Mean Precision	Mean Recall	Mean F1
1	[24]	layer 4	✗	✗	56%	54%	56%	55%
2	[24]+vessel		✓	✗	**59%**	**59%**	**59%**	57%
3	**Proposed**		✓	✓	**59%**	58%	**59%**	**58%**
4	[24]	layers 3-4	✗	✗	60%	57%	60%	58%
5	[24]+vessel		✓	✗	62%	60%	62%	60%
6	**Proposed**		✓	✓	**68%**	**63%**	**68%**	**65%**
7	[24]	layers 2-4	✗	✗	65%	65%	65%	63%
8	[24]+vessel		✓	✗	65%	63%	65%	63%
9	**Proposed**		✓	✓	**68%**	**67%**	**68%**	**67%**
10	[24]	layers 1-4	✗	✗	61%	59%	61%	60%
11	[24]+vessel		✓	✗	64%	61%	64%	61%
12	**Proposed**		✓	✓	**67%**	**67%**	**67%**	**65%**

Table 3: Testing result on retinopathy dataset with 5 level grading.

Information Processing in Medical Imaging, pages 793–804.
Springer, 2019. 2, 3, 4, 5, 7