Classifying Diabetic Retinopathy Images using Induced Deep Region of Interest Extraction

Ashutosh Kushwaha* and P. Balamurugan†

* akushwaha9808@gmail.com, † balamurugan.palaniappan@iitb.ac.in

IEOR, Indian Institute of Technology Bombay.

Main Contributions

- We consider the task of classifying diabetic retinopathy images coming from a small real-world data set.
- We propose a **semi-automatic two-stage approach** for diabetic retinopathy image classification.
  - **Regions of interest (RoI) are identified by segmentation task.**
  - **RoI fuel the classification task with additional discriminative information.**
- We also devise **new loss functions** which improve the quality of segmentation tasks in diabetic retinopathy.

Introduction

Diabetic Retinopathy (DR) related to progression of diabetes.

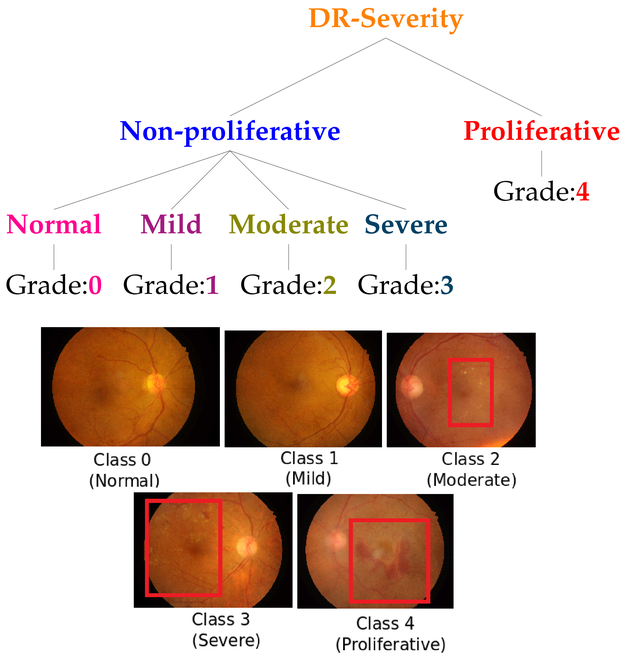

Figure 1: DR Severity Levels

Severity levels of DR associated with lesions, *e.g.* haemorrhages, hard exudates, micro-aneurysms.

- **Goal:** Given a retinal fundus image $x \in \mathscr{X}$, classify it into a DR grade $y \in \mathscr{Y} = \{0,1,2,3,4\}$.
- **Idea:** Use deep learning to learn a multi-class classification map $h : \mathscr{X} \to \mathscr{Y}$.

Proposed Two-step Approach

- **Segmentation**
  - Two U-Nets used for segmentation of haemorrhages and hard exudates.
  - Training done on different segmentation data.
  - New loss functions proposed for segmentation of hard exudates.
- **Classification**
  - For an image x in classification data set, find haemorrhage and hard exudate masks of x.
  - Haemorrhage and hard exudate masks of x superimposed over grayscale channel of image x.
  - ResNet 101 used for training synthesized images.

New Modified Tversky Segmentation Loss

$$T_M(\alpha_1,\alpha_2,\beta_1,\beta_2) = T_P + T_N \text{ where}$$

$$T_P = \frac{\sum_{i=1}^{N} p_{0i}g_{0i}}{\sum_{i=1}^{N} p_{0i}g_{0i} + \alpha_1 \sum_{i=1}^{N} p_{0i}g_{1i} + \beta_1 \sum_{i=1}^{N} p_{1i}g_{0i}}$$

$$T_N = \frac{\sum_{i=1}^{N} p_{1i}g_{1i}}{\sum_{i=1}^{N} p_{0i}g_{0i} + \alpha_2 \sum_{i=1}^{N} p_{0i}g_{1i} + \beta_2 \sum_{i=1}^{N} p_{1i}g_{0i}}$$

α_1, α_2, β_1, β_2 are hyperparameters. p_{ji} and g_{ji} denote class j probability of pixel i in predicted and ground truth binary images.

New Focal Modified Tversky Loss

$$FMT = (1 - T_P)^{\frac{1}{\gamma_1}} + (1 - T_N)^{\frac{1}{\gamma_2}}$$

FMT focuses more on hard negatives as compared to the easily predicted lesion pixels (γ_1 and γ_2 used to down-weight the importance of easily classified lesion pixels).

Please check our paper and code to know more on experimental settings.

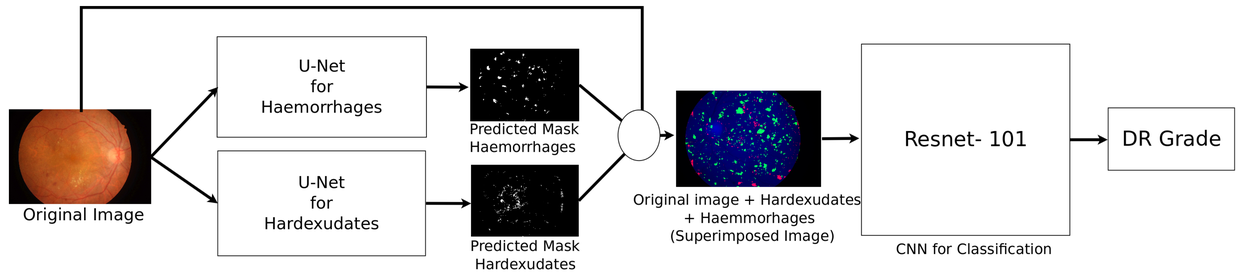

Figure 2: Our proposed approach

Loss Function Used	Train/Test	F1 Score	F2 Score	Precision	Recall	Pixel accuracy	IOU	Specificity
Weighted Binary Cross Entropy (3:1)	Train	.8595	.8602	.8588	.8608	.9942	.7551	.9970
	Test	.7750	.7770	.7788	.7811	.9959	.6397	**.9980**
Weighted Binary Cross Entropy (5:1)	Train	.9215	.9215	.9216	0.9215	0.9967	0.8549	0.9983
	Test	0.8038	0.7960	0.8263	0.7926	**0.9967**	0.6832	0.99852
Unweighted Binary Cross Entropy	Train	0.7260	0.7286	0.7228	0.7305	0.9886	0.5729	0.9940
	Test	0.67788	.6703	.7009	.6681	.9941	.5236	.99735
Tversky	Train	.8797	.9129	.8298	.8797	.9946	.7861	.9958
	Test	.7948	.8128	.7752	.8274	.9961	.6674	.9975
Modified Tversky	Train	.8835	.9136	.8377	.9349	.9949	.7921	.9961
	Test	.7984	.8317	.7545	.8597	.9959	.6724	.9971
Focal Modified Tversky	Train	.9109	.9255	.8654	.9421	.9956	.8221	.9968
	Test	**.8146**	**.8332**	**.7888**	**.8471**	**.9965**	**.6980**	**.9977**

Table 1: Segmentation Results for Hard exudates (Best results for test set in bold)

Dataset Used	Classification Loss Function	Train Accuracy	Validation Accuracy	Test Accuracy
Original Images	Tversky	79.3%	53.6%	38.8%
	Modified Tversky	75.1%	26.8%	33.1%
	Focal Modified Tversky	78.3%	48.8%	35%
Original Images + Hard exudates + Haemorrhages	Tversky	74.7%	61%	48.5%
	Modified Tversky	72.2%	65.9%	49.5%
	Focal Modified Tversky	77.4%	65.9%	49.5%

Table 2: Classification Results (synthesized data with original image + hard exudate segments + haemorrhage segments show consistent improvement over original images)

Conclusions

- Two-step pipeline improves overall classification performance.
- Proposed losses useful for hard exudate segmentation.
- **Limitations:**
  - Overfitting is an issue in classification which needs to be resolved by better methods.
  - Semi-automatic method needs to be made fully end-to-end.

References

[1] N. Abraham and N. M. Khan. *A Novel Focal Tversky Loss Function with improved Attention U-Net for Lesion Segmentation*. arXiv:1810.07842, 2018.

[2] K. He, X. Zhang, S. Ren and J. Sun. *Deep Residual Learning for Image Recognition*. CVPR, 2016.

[3] T-Y. Lin, P. Goyal, R. Girshick, and K. He, and Dollár, P. *Focal Loss for Dense Object Detection*. ICCV, 2017.

[4] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe and F. Meriaudeau. *Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research*. Data, 2018.

[5] O. Ronneberger, P. Fischer and T. Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. MICCAI, 2015.

[6] S. S. Salehi, D. Erdogmus and A. Gholipour. *Tversky Loss Function for Image Segmentation using 3D Fully Convolutional Deep Networks*. IWMLMI, 2017.

[7] T. Y. Wong, C. M. G. Cheung, M. Larsen, S. Sharma and R. Simó. *Diabetic Retinopathy*. Nature Reviews Disease Primers, vol. 2, 2016.

[8] Latex template courtesy: https://www.latextemplates.com

Second author's work partially supported by SEED Grant from IIT Bombay.