Automated Detection & Diagnosis of Diabetic Retinopathy (DR)

By **Muhammad Zain Amin**

University of Bourgogne, Le Creusot, France

Muhammad-Zain_Amin@etu.u-bourgogne.fr

Diabetes Facts

- Diabetes is a chronic disease that occurs either when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces. Insulin is a hormone that regulates blood glucose.

- 422 Million people affected worldwide
- 25.8 M only in USA (8.3% of population)
- Projection for 2030: 643 Million people worldwide according to International Diabetes Foundation.

- Many complications for various organs : Retina, Nervous System, Heart, Kidneys
- In our study, we focus on the Retina based Diabetic Retinopathy disease.

Diabetic Retinopathy

- DR is a chronic eye retina disease, where high blood glucose levels if not controlled can cause lesions on the retina that effect vision and ultimately lead to permanent blindness.
- 28.5% of adults with diabetes (> 40 years old) have DR.

What can be done?

- Prompt diagnosis
- Early detection and screening of DR is crucial to avoids vision loss in 50% of the patients.

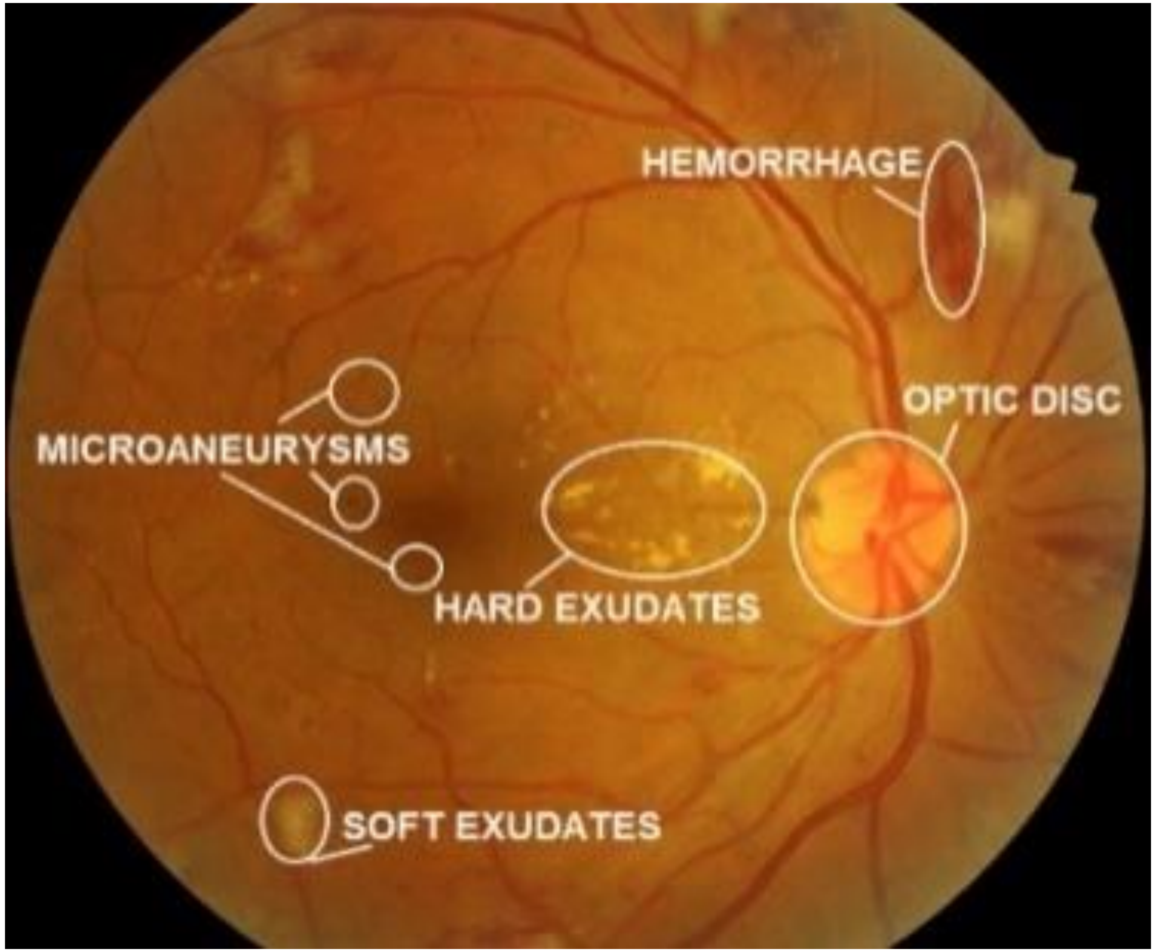
DR Detection Problem?

- Manual inspection of DR is a very time-consuming process and extremely prone to misdiagnosis. Also, requires retinal examination by ophthalmologist. This raises the need for using the automated DR detection and screening systems.

How to detect DR?

DR is detected and classified by the shape and appearance of various types of lesions on a retina. Commonly, the four types of lesions are:

- Microaneurysms (MA) - First detectable sign of DR, less than 125 micrometers in size
- Haemorrhages (HM) - Size of more than 125 µm
- Hard exudates (HE) - appears due to the problem of plasma leakage
- Soft exudates (SE) - appears as white oval or round shape on the retina due to problem caused by swelling of nerve fiber
- HM and MA are red lesions, and HE and SE are bright lesions.



DR Stages with Corresponding Lesions Appearance

- In 2003, DR experts created the International Clinical Grading Scale to detect the severity of DR. They classified DR disease into five stages.

Diabetic Retinopathy Stages	Retinal Lesions
No Diabetic Retinopathy (No DR)	No Lesions Visible
Mild Non-Proliferative Diabetic Retinopathy (Mild NPDR)	Microaneurysms (Localized swelling of the small blood vessels in the retina)
Moderate Non-Proliferative Diabetic Retinopathy (Moderate NPDR)	Microaneurysms, plus small bleeds (dot and blot Haemorrhages), leaks (Hard Exudates) or closure (cotton wool spots) of small blood vessels
Severe Non-Proliferative Diabetic Retinopathy (Severe NPDR)	Moderate NPDR plus further damage to blood vessels (intraretinal Hemorrhages, venous beading, intraretinal microvascular abnormalities)
Proliferative Diabetic Retinopathy (PDR)	One or more of: - New vessel formation, Vitreous/pre-retinal hemorrhage, tractional retinal detachment)

Automated Detection and Diagnosis of Diabetic Retinopathy

How to automate the process of DR detection and Diagnosis?

In the scope of an automated DR diagnosis, there are two main approaches:

- Classification, and Segmentation.
- The classification approach is for grading the DR severity into five different stages :- No DR, Mild Non-Proliferative DR, Moderate Non-Proliferative DR, Severe Non-Proliferative DR, Proliferative DR.
- And the second approach is the segmentation of DR-associated lesions such as hard exudates, soft exudates, microaneurysms, and hemorrhages.

Various Deep Learning and Machine Learning based DR classification and Segmentation methods have been proposed by several researchers over the last few years.

Deep Learning and Machine Learning

- Deep learning (DL) is a branch of machine learning (ML), based on the creation of multi-layered neural networks.
- DL methods such as Convolutional Neural Networks have been extensively used by researchers for the DR segmentation, and classification tasks.

Difference between Deep Learning and Machine Learning

- The classification performance of DL methods increases when the size of training data increases, resultant in a greater number of learned features. But, if the size of the training data is small, DL methods can also lead to overfitting problem.
- ML methods require hand-crafted feature extraction. ML techniques require domain experts to detect most of the features from images in order to reduce data complexity and assist ML classifiers by providing data with more visible patterns.
- DL methods didn’t require domain expertise and eliminate the need of handcrafted feature extraction. CNNs usually require less preprocessing, and they can easily identify the patterns at the pixel level.

The massive success of Deep CNNs architectures is because of their ability to extract and learn features from huge datasets.

Diabetic Retinopathy Datasets

- In the scope of an automated DR diagnosis, the fundus image databases with well-defined annotations and tasks are needed to train the DR classification and segmentation models.
- For DR severity classification, image level annotations are needed to train the model.
- For DR lesions segmentation, pixel level annotations are needed to train the model.

Dataset	Image Count	Resolution (px)	Annotations	Task	Camera used
Messidor	1200	1440 x 960 to 2304 x 1536	Image Level	DR grading + DME grading	45° wide view
Kaggle EyePACS	35126	Varies	Image Level	DR grading	Varies
IDRiD	516	4288 x 2848	Image & Pixel Level	DR grading + DME grading + Lesion Segmentation	50° wide view
Private Dataset – DDR	12,522	Varies	Image & Pixel Level	DR grading + Lesion Segmentation	45° wide view
E-Ophtha	463	Varies	Pixel Level	Exudates and Microaneurysms Detection	50° wide view
Private Dataset	30244	520 x 520	Image Level	DR grading	Varies

Deep Learning for Automatic Preprocessing of Fundus Images

DR Lesions Segmentation

- For DR lesions segmentation task, only few pixel level annotated datasets are available, and most of them only contains annotations for only one or two lesions.
- In 2018, The IDRiD Diabetic Retinopathy segmentation and grading challenge was introduced by the **University of Bourgogne's** renowned Professor Fabrice Meriaudeau along with Indian DR researchers.
- The IDRiD challenge introduced the IDRiD database. The IDRiD database is the only publicly available one, that contains the pixel-level annotations of four major types of lesions. This dataset also includes image-level annotations to provide DR severity information.
- So, this makes the IDRiD database perfect for development and evaluation of image analysis algorithms for early detection of diabetic retinopathy.

DR Lesions Segmentation

- The manual segmentation of DR lesions stay difficult on fundus images and require strict rules.
- In our study we have studied three different automated approaches used for DR lesions segmentation.

Custom CNN approach [1]

- Guo et al. proposed a custom CNN segmentation approach. A pretrained CNN architecture namely VGG-16 was used as a backbone model for feature extraction. The model has the capability to segment all the lesions at the same time.

U-Net Segmentation approach [2]

- Yan et al. proposed an efficient mutual Global-Local U-Net based lesion segmentation method. Usually due to the small size of lesion regions in the fundus images, the lesion segmentation becomes very difficult to perform. Commonly down sampling the fundus images helps us in the segmentation procedure but at the same time resultant in the loss of detailed and context information. So, they built an effective network that is mutually trained on the entire fundus image as well as on its patches.

Collaborative Learning approach [3]

- Zhou et al. developed a collaborative learning approach to mutually enhance the performance of DR severity classification and lesions segmentation with an attention mechanism approach. The model first performed the semantic segmentation task with multi lesion mask generational model using a small set of pixel-level annotated data. After, a lesion attentive disease grading model is utilized to improve the severity classification based on the initially predicted lesions map for image-level annotated data. At the same time, the lesion attentive model can enrich the lesions maps by utilizing the DR class specific details to improve the segmentation model.

- **Does the problem solve?**
- Globally ….. Yes.

DR Lesions Segmentation Approaches Comparison

DR LESIONS SEGMENTATION						
Custom CNN Segmentation Approach			**Lesions**	**AUC Scores**	**Lesions**	**AUC Scores**
[1]	Custom CNN	IDRiD (413)	Microaneurysms	0.463	Hard Exudates	0.795
	-	-	Haemorrhages	0.637	Soft Exudates	0.711
[1]	Custom CNN	DDR (12,522)	Microaneurysms	0.105	Hard Exudates	0.555
	-	-	Haemorrhages	0.359	Soft Exudates	0.265
[1]	Custom CNN	e-ophtha (463)			Hard Exudates	0.417
U-Net Segmentation Approach						
[2]	Custom U-Nets	IDRiD (413)	Microaneurysms	0.525	Hard Exudates	0.889
	-	-	Haemorrhages	0.703	Soft Exudates	0.679
Collaborative Learning Approach						
[3]	Custom U-Net	IDRiD (413)	Microaneurysms	0.9828	Hard Exudates	0.9935
	-	-	Haemorrhages	0.9779	Soft Exudates	0.9936

DR Lesions Segmentation Approaches Conclusion

- The Custom CNN segmentation approach proposed by Guo et al. and Custom U-Net segmentation approach proposed by Yan et al. show better AUC scores for Hard Exudates and Soft Exudates Lesions as compared to Microaneurysms and Haemorrhages Lesions.

Why?

- Question arises, might be due to the class imbalances in the dataset, which usually effect the segmentation models, **Answer is NO**.

The main reason is:-

- Because the hard and soft exudates lesions are bigger in size as compared to microaneurysms and haemorrhages lesions. So, it effects the segmentation performance of these models.

So, how to manage wrong segmentations in case of DR lesions?

- Use of deep learning for the segmentation of retinal fundus images with anatomical guarantees.
- Use of Collaborative Learning Approach based on Attention Mechanism

Zhou et al. collaborative learning approach is reasonable for the use in clinical practice. How?

- Because their approach segmented all types of lesions almost equally and performed extremely well with an AUC scores range from 0.9779 ± 0.9936.

DR Classification

- For DR severity grading, various approaches have been covered in our study.

Traditional CNN approach [4]

- CNN methods have been extensively utilized in many DR classification tasks. Mobeen-ur-Rehman et al. classified the fundus images using CNN models such as AlexNet, VGG-16, and SqueezeNet. They also built their own 5-layered CNN architecture with each layer having their own specifications and reported higher accuracy as compared to above mentioned models.

Machine Learning approach [5]

- Jorge de la Calleja et al. implemented a DR classification model by using local binary patterns for the feature extraction phase and utilized support vector machine, random forest, and artificial neural network for the classification purposes.

Multipath CNNs with Machine Learning approach [6]

- Gayathri et al. proposed a DR severity classification system based on multipath CNN and ML classifiers. The multipath CNN network is designed for the extraction of features from the retinal fundus images. Then, ML classifiers such as support vector machine, random forest, and J48 was used to categorize the inputs based on the DR severity. The multipath CNN network is composed of two feed-forward paths. First path is similar to traditional CNN. The second path is designed for extraction of features along multiple paths.

DR Grading Classification Approaches Comparison

Ref.	Methods	Dataset (Size)	Performance Measures for DR Severity Classification Models				
			Accuracy	Sensitivity	Specificity	F1-Score	AUC
DR Severity Grading Classification							
Traditional CNN Approach							
[4]	CNN (AlexNet)	MESSIDOR (1200)	93.46	92.38	94.53	-	-
[4]	CNN (VGG-16)	MESSIDOR (1200)	91.82	93.47	88.54	-	-
[4]	CNN (SqueezeNet)	MESSIDOR (1200)	94.49	94.47	94.54	-	-
[4]	Custom CNN	MESSIDOR (1200)	98.15	98.94	97.87	-	-
Multipath CNNs with Machine Learning Approach							
[6]	Custom M-CNN + SVM	MESSIDOR (1200)	94.16	94.2	96.8	94.1	-
[6]	Custom M-CNN + RF	MESSIDOR (1200)	91.83	91.8	95.0	91.6	-
[6]	Custom M-CNN + J48	MESSIDOR (1200)	99.75	99.8	99.9	99.7	-
[6]	Custom M-CNN + SVM	Kaggle (35126)	96.18	96.2	95.1	95.8	-
[6]	Custom M-CNN + RF	Kaggle (35126)	96.99	97.0	92.1	96.4	-
[6]	Custom M-CNN + J48	Kaggle (35126)	99.9	99.9	99.9	99.9	-
[6]	Custom M-CNN + SVM	IDRiD (413)	85.47	85.5	94.5	-	-
[6]	Custom M-CNN + RF	IDRiD (413)	78.93	78.9	91.7	75.8	-
[6]	Custom M-CNN + J48	IDRiD (413)	99.03	99.0	99.7	99.0	-
Machine Learning Approach							
[5]	Local Binary Patterns + ANN	MESSIDOR (100)	77.6	-	-	78.06	-
[5]	Local Binary Patterns + RF	MESSIDOR (100)	74.2	-	-	73.74	-
[5]	Local Binary Patterns + SVM	MESSIDOR (100)	78.8	-	-	79.17	-
[5]	Local Binary Patterns + ANN	MESSIDOR (71)	94.9	-	-	94.96	-
[5]	Local Binary Patterns + RF	MESSIDOR (71)	97.4	-	-	97.40	-
[5]	Local Binary Patterns + SVM	MESSIDOR (71)	95.7	-	-	95.80	-

DR Grading Classification Approaches Conclusion

- **A custom CNN architecture built by Mobeen-ur-Rehman et al. performed better in comparison to existing architectures such as AlexNet, VGG-16 and SqueezeNet.**
- **Why? Due to the problem of overfitting in AlexNet, VGG-16 and SqueezeNet models.**
- **Overfitting indicates that your model is too complex for the problem that it is solving, it means model has too many filters in the case of CNNs, and layers in the case of overall Deep Learning Models. This causes your model to know the example data well but perform poorly against any new data.**
- **The machine learning approach proposed by Jorge de la Calleja was evaluated on the original Messidor database. In the first experiment a subset of 100 fundus images was selected randomly. The approach perform poorly, and the accuracy scores of all the classifiers are in the range 74.2 ± 78.8. In the second experiment, they selected just 71 fundus images due to their higher quality and visual information. The preliminary results of their model in the second experiment observe that random forest obtained the best results demonstrated an average accuracy of 97.4%.**
- **So, question arises, why there is a huge performance difference between 1st and 2nd experiment?**
- **Because, as we have discussed in our previous slide, that the machine learning models require high quality features and require domain experts to detect most of the features from images in order to reduce data complexity and assist ML classifiers by providing data with more visible patterns.**
- **So, what is the best solution then to deal with these issues? A hybrid learning approach**
- **Most of the times the features extracted using traditional CNN architectures suffers from losses in the global structure because of too short or too long forward path. In Multipath CNN architectures, this issue can be resolved by using the shortcut paths. Also, the feature extractors utilized in Multipath CNN architectures can help preserve the losses of global path thus resultant in generating more appropriate local and global structures. Gayathri et al. specifically designed a multipath CNN architecture for the detection of DR features from fundus images. The performance of proposed Multipath CNN architecture along with J48 classifier performed better in comparison to traditional CNN architectures.**

References

[1] Guo, S., Li, T., Kang, H., Li, N., Zhang, Y., & Wang, K. (2019). L-Seg: An end-to-end unified framework for multi-lesion segmentation of fundus images. Neurocomputing, 349, 52-63.

[2] Z. Yan, X. Han, C. Wang, Y. Qiu, Z. Xiong and S. Cui, "Learning Mutually Local-Global U-Nets For High-Resolution Retinal Lesion Segmentation In Fundus Images," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 597-600

[3] Y. Zhou et al., "Collaborative Learning of Semi-Supervised Segmentation and Classification for Medical Images," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 2074-2083

[4] Mobeen-ur-Rehman, S. H. Khan, Z. Abbas and S. M. Danish Rizvi, "Classification of Diabetic Retinopathy Images Based on Customised CNN Architecture," 2019 Amity International Conference on Artificial Intelligence (AICAI), 2019, pp. 244-248

[5] de la Calleja, J., Tecuapetla, L., Auxilio Medina, M., Bárcenas, E., Urbina Nájera, A.B. (2014). LBP and Machine Learning for Diabetic Retinopathy Detection. In: Corchado, E., Lozano, J.A., Quintián, H., Yin, H. (eds) Intelligent Data Engineering and Automated Learning – IDEAL 2014. IDEAL 2014. Lecture Notes in Computer Science, vol 8669. Springer, Cham

[6] Gayathri, S., Gopi, V.P. & Palanisamy, P. Diabetic retinopathy classification based on multipath CNN and machine learning classifiers. Phys Eng Sci Med 44, 639–653 (2021).