A Comparative Study on the Automated Detection and Diagnosis of Diabetic Retinopathy

Muhammad Zain Amin
Erasmus +: Erasmus Mundus Joint Master Degree in Medical Imaging and Applications
University of Bourgogne, Le Creusot, France
Muhammad-Zain_Amin@etu.u-bourgogne.fr

Abstract

Diabetic Retinopathy (DR) is a chronic degenerative eye disease that particularly occurs in diabetic patients. Diabetic retinopathy is caused by the diabetes mellitus, where high blood glucose levels if not controlled can cause lesions on the retina that effect vision and ultimately lead to permanent blindness. DR is an irreversible disease and early detection of DR is crucial to maintain the patients vision. Manual inspection of DR detection is a very time-consuming process and requires retinal examination by ophthalmologist. This raises the need for using the technology to detect DR at an early stage. Hence, an automated detection system for DR diagnosis is proved to be a very helpful tool for eye care professionals. From the last few years, the utilization of classification and segmentation methods have shown promising results in assessing the retinal fundus images and facilitated the early diagnosis of DR disease. In this paper, we will analyze the state-of-the-art methods recently used for DR detection in detail. Also, we will examine the research gaps in the DR domain and address the critical challenges that require further investigation.

1. Introduction

Diabetic Retinopathy detection is one of the most challenging tasks in the medical imaging field despite the significant evolution of multiple approaches and network architectures in the last decade. DR is a chronic retina disease and is the leading cause of blindness globally. According to the report of the World Health Organization, diabetes cases have been rapidly growing over the last few decades. In 1980, there were nearly 108 million diabetic incidents globally, which increased to 422 million in 2014 [1]. The number of diabetes patients is further expected to reach 643 million by 2030, according to a report by International Diabetes Federation [2]. Regular Retina checkup of diabetes patients is necessary to detect the DR disease at an early stage. 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)* - is the initial stage of diabetic retinopathy usually emerges as a small red round dot on the retina due to vessel wall weakness. The red round dots have clear margins and are less than 125 micrometers in size.
- *Haemorrhages (HM)* - emerges as larger spots on the retina with a size of more than 125 µm. HM have irregular margins, contrary to MA. There are two types of HM, classified as: flame with superficial spots and, blot with deeper spots.
- *Hard exudates (HE)* - appears due to the problem of plasma leakage and are visible on the retina as yellow spots. Hard exudates have sharp margins and can be observed in the outer layer of the eye retina.
- *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 both are red lesions, and hard exudates and soft exudates appear as bright lesions. Figure 1. shows the appearance of different lesions on the retina. In 2003, several DR experts and researchers created the International Clinical Disease Severity Scale for Diabetic Retinopathy to simplify the classification process [3]. They classified DR disease into five different stages depending on the presence of lesions as shown in table 1.

Table 1. DR Stages based on the International Clinical Disease Severity Scale for Diabetic Retinopathy [3].

Diabetic Retinopathy Stages	**Retinal Lesions**
No Diabetic Retinopathy (**No DR**)	No Lesions Visible
Mild Non-Proliferative Diabetic Retinopathy (**Mild NPDR**)	Localized swelling of the small blood vessels in the retina (Lesion Visible: Microaneurysms)
Moderate Non-Proliferative Diabetic Retinopathy (**Moderate NPDR**)	Mild NPDR 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)

The automated DR diagnosis systems are time and cost effective and are more efficient as compared to manual diagnosis. The manual diagnosis of DR requires a plenty of time and effort and is extremely prone to misdiagnosis. In this study, we survey the most recent papers related to diabetic retinopathy segmentation and classification based on deep learning and machine learning. Overall, the focus of this paper is to discuss, compare, and identify the pros and cons of recent automated DR diagnosis methods and their overall impact on the future research directions.

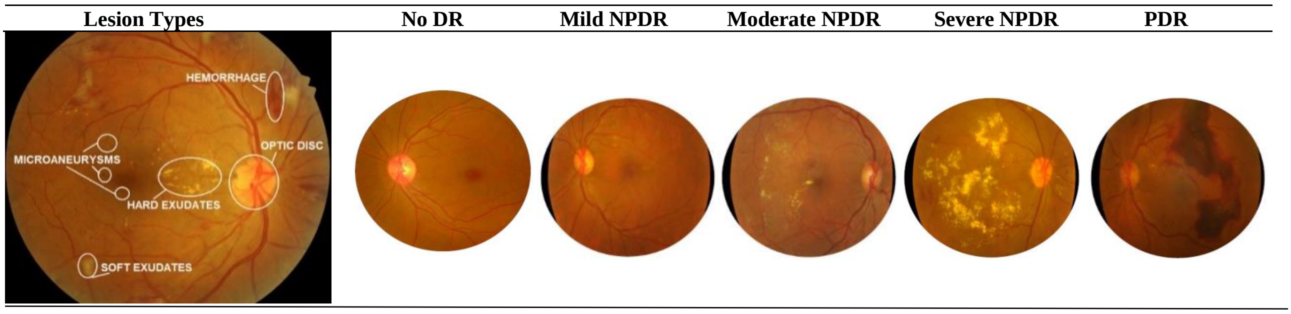

Figure 1. Lesion Types with Diabetic Retinopathy Stages Visuals

2. Deep Learning and Machine Learning

Deep learning (DL) is a branch of machine learning (ML), and it is mainly based on the creation of complex, multi-layered neural networks. In DL, neural networks allow data to be passed between the nodes in a highly connected manner. There are several DL methods such as Convolutional Neural Networks (CNNs), Auto Encoders, and Boltzmann Machine. Multiple applications of deep learning for medical imaging analysis include detection, classification, segmentation, and information retrieval.

From the last few years, DL methods have been extensively used by researchers for the DR segmentation, and classification tasks. DL methods have the ability to easily learn features from the input data. The major advantage of deep learning over machine learning is that the performance of DL methods increases when the size of training data increases, resultant in a greater number of learned features. Contrary to it, if the size of the training data is small, DL methods usually causes overfitting problem. In case of dealing with smaller datasets, low-complexity ML methods such as support vector machine, random forest, decision tree, k-nearest neighbours, logistic regression etc. typically show better performance during the classification tasks.

One of the main limitations of ML methods is that they require hand-crafted feature extraction. ML techniques require domain experts to detect the majority of the features from images in order to reduce data complexity and assist ML classifiers by providing data with more visible patterns. On the other hand, DL methods didn’t require domain expertise and eliminate the need of hand-crafted feature extraction. One of the most popular DL techniques namely, Convolutional Neural Networks has been widely used in medical imaging segmentation and classification problems [4]. CNNs usually require less preprocessing, and they can easily identify the patterns at the pixel level. They are also less affected by distortions and geometrical transformations. The massive success of Deep CNNs architectures is because of their ability to extract and learn features from huge datasets [5].

3. Diabetic Retinopathy Datasets

The fundus images are captured by fundus cameras to photograph interior, back surface of an eye. The optic disc, macula, and central and peripheral retina represent the primary structures commonly observable on the fundus image. Multiple fundus image datasets are publicly available and used by researchers to detect DR. These datasets are used to train, validate, and test the systems, as well as to assess how well a system performs compare to other systems. Table 2 shows various datasets used in our study.

Table 2. Retina Fundus Image Datasets.

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 [6]	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 [7]	30244	520 x 520	Image Level	DR grading	Varies

4. Diabetic Retinopathy Diagnosis Methods

Multiple researchers have implemented automated diagnosis methods for DR severity classification and lesions segmentation. Several performance metrices such as accuracy, sensitivity, F1 measure, area under the curve (AUC) and, specificity scores are used to evaluate the performance of DR classification and segmentation models. The performance details of these models are shown in table 3.

4.1. Diabetic Retinopathy Stages Classification

In this section, we demonstrate the specifications and performance of various DR severity classifications model.

4.1.1. Traditional CNN approach:

CNN methods have been extensively utilized in many DR detection tasks. In the study [8], Mobeen-ur-Rehman et al. classified the fundus images using well-known CNN models such as AlexNet, VGG-16, SqueezeNet and reported an accuracy score of 93.46, 91.82 and 94.49 % respectively. They also built their own 5-layered CNN architecture with each layer having their own specifications. The first layer combined 4 different kernels for feature extraction and contained a pooling layer. The pooling layer reduces the size of the convolutional layer, and thus reduces the number of parameters to learn and the amount of computation time. After pooling, the four outputs of the kernels are convolved with 16 kernels in the next convolutional layer. Again, the pooling layer is utilized to reduce the output matrix. The last three layers of CNN model contains fully connected neural network layers with 100, 50 and 10 neurons respectively. In their proposed model, contrary to the Deep CNNs, they have used only two convolutional layers to avoid the overfitting problem. The custom model shown better performance as compared to AlexNet, VGG-16 and SqueezeNet models and reported an accuracy of 98.15%, sensitivity of 98.94%, and specificity of 97.87%.

4.1.2. Multipath CNNs with Machine Learning approach:

Since the last decade, multipath neural networks have been used by researchers for solving various kind of classification tasks. Gayathri et al. [9] 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. The performance of their model is tested by using three publicly available datasets, Messidor, EyePACS, and IDRiD. The best reported accuracies on all the datasets is generated by J48 classifier with multipath CNN network.

4.1.3. Machine Learning approach:

Jorge de la Calleja et al. [10] implemented a DR detection model by using local binary patterns for the feature extraction phase and utilized support vector machine (SVM), random forest (RF), and artificial neural network (ANN) for the classification purposes. Different binary patterns were used such as uniform patterns, rotation invariant uniform patterns, and rotation invariant patterns to evaluate the performance of the model. The model was tested on Messidor database. From the original Messidor dataset of 1200 images, a subset of 100 fundus images was selected in the first experiment. The experiment results show the best classifier was support vector machines with an accuracy and F1 score of 78.8%, and 0.7917 respectively using rotation invariant uniform patterns. 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% and an F1-score of 0.974% using uniform patterns and rotation invariant uniform patterns.

4.1.4. Transfer Learning approach:

In the study [11], S. Wan et al. analyzed the performance of pre-trained CNN models, namely, AlexNet, VGGNet-16, VGGNet-19, VGGNet-s, GoogLeNet, and ResNet for DR severity classification task. These models were trained on the Kaggle EyePACS dataset to detect the five stages of diabetic retinopathy. The dataset contains 35,126 retinal fundus images, and these images were cropped, normalized, and augmented during the preprocessing step. Transfer Learning was implemented on all the above-mentioned CNN models by fine-tuning the hyperparameter and last fully connected layer. The VGGNet-s model performed better as compared to other models and achieved a higher accuracy and specificity scores of 95.68% and 97.43% respectively.

4.1.5. Attention Modules approach:

X. Li et al. [12] proposed a cross-disease attention network for mutual grading of DR and Diabetic Macular Edema (DME) disease. The proposed model namely CANet, find the relationship between DME and DR disease using image level annotations. The primary structure of the network composed of a disease specific attention module to learn features for individual disease and a disease dependent attention module to find the relationship between DR and DME diseases. Both of these attention modules were integrated in a neural network to understand disease specific and disease dependent features. The CANet model classified the fundus images into referable and non-referable categories using the Messidor database, while the IDRiD database was used to classify DR into five stages and DME into three stages. ResNet-50 architecture was used to extract features and the features are used as an input to the first two attention modules. The first two attention modules include multiplication, concatenation, max pooling, convolution, and the fully connected layers. The last two attention modules only contain the fully connected and multiplication layer. Resizing and data augmentation were performed on the fundus images before feeding the images to the ResNet-50 model. The proposed CANet model achieved an overall accuracy of 92.6 % on the Messidor database.

4.1.6. Ensemble Learning approach:

Ensemble learning is the process of combining several deep learning models into an ensemble model to obtain their collective performance and to improve the overall performance of individual models. In ensemble learning, individual models are stacked together. The DR severity classification method proposed by Jiang et al. [7] integrated several deep learning models such as Inception V3, ResNet152, Inception-ResNet-V2 and use Adaboost algorithm to classify the retinal fundus images as referable DR or non-referable DR. The proposed ensemble deep learning model was tested on a private dataset containing 30,244 retinal fundus images and obtained an accuracy and specificity scores of 88.21% and 90.85% respectively.

4.2. Diabetic Retinopathy Lesion Segmentation:

In this section, we discussed and investigate different segmentation techniques used for the detection of different lesion types.

Table 3. Performance details of different DR severity classification and lesions segmentation approaches.

Ref.	**Methods**	**Dataset (Size)**	**Performance Measures for DR Severity Classification Models**				
			Accuracy	**Sensitivity**	**Specificity**	**F1-Score**	**AUC**
DR Severity Classification							
Traditional CNN Approach							
[8]	CNN (AlexNet)	MESSIDOR (1200)	93.46	92.38	94.53	-	-
[8]	CNN (VGG-16)	MESSIDOR (1200)	91.82	93.47	88.54	-	-
[8]	CNN (SqueezeNet)	MESSIDOR (1200)	94.49	94.47	94.54	-	-
[8]	Custom CNN	MESSIDOR (1200)	98.15	98.94	97.87	-	-
[7]	CNN (Inception V3)	Private Dataset (30244)	87.91	84.35	91.46	-	0.935
[7]	CNN (ResNet-152)	Private Dataset (30244)	87.20	84.76	89.63	-	0.940
[7]	CNN (Inception-ResNet-V2)	Private Dataset (30244)	86.18	83.94	88.41	-	0.943
Transfer Learning Approach							
[11]	CNN (AlexNet)	Kaggle (35126)	89.75	81.27	94.07	-	0.9342
[11]	CNN (VGGNet-s)	Kaggle (35126)	95.68	86.47	97.43	-	0.9786
[11]	CNN (VGG-16)	Kaggle (35126)	93.17	90.78	94.32	-	0.9616
[11]	CNN (VGG-19)	Kaggle (35126)	93.73	89.31	96.49	-	0.9684
[11]	CNN (ResNet-50)	Kaggle (35126)	90.40	88.78	95.56	-	0.9365
[11]	CNN (GoogLeNet)	Kaggle (35126)	93.36	77.66	93.45	-	0.9272
Attention Modules Approach							
[12]	CNN (ResNet-50)	MESSIDOR (1200)	92.6	92.0	-	91.2	96.3
Multipath CNNs with Machine Learning Approach							
[9]	Custom M-CNN + SVM	MESSIDOR (1200)	94.16	94.2	96.8	94.1	-
[9]	Custom M-CNN + RF	MESSIDOR (1200)	91.83	91.8	95.0	91.6	-
[9]	Custom M-CNN + J48	MESSIDOR (1200)	99.75	99.8	99.9	99.7	-
[9]	Custom M-CNN + SVM	Kaggle (35126)	96.18	96.2	95.1	95.8	-
[9]	Custom M-CNN + RF	Kaggle (35126)	96.99	97.0	92.1	96.4	-
[9]	Custom M-CNN + J48	Kaggle (35126)	99.9	99.9	99.9	99.9	-
[9]	Custom M-CNN + SVM	IDRiD (413)	85.47	85.5	94.5	-	-
[9]	Custom M-CNN + RF	IDRiD (413)	78.93	78.9	91.7	75.8	-
[9]	Custom M-CNN + J48	IDRiD (413)	99.03	99.0	99.7	99.0	-
Ensemble Learning Approach							
[7]	Ensemble Deep Learning Model	Private Dataset (30244)	88.21	85.57	90.85	-	0.946
Machine Learning Approach							
[10]	Local Binary Patterns + ANN	MESSIDOR (100)	77.6	-	-	78.06	-
[10]	Local Binary Patterns + RF	MESSIDOR (100)	74.2	-	-	73.74	-
[10]	Local Binary Patterns + SVM	MESSIDOR (100)	78.8	-	-	79.17	-
[10]	Local Binary Patterns + ANN	MESSIDOR (71)	94.9	-	-	94.96	-
[10]	Local Binary Patterns + RF	MESSIDOR (71)	97.4	-	-	97.40	-
[10]	Local Binary Patterns + SVM	MESSIDOR (71)	95.7	-	-	95.80	-
			Performance Measure for DR Lesions Segmentation Models				
			MA & HM Lesions	**AUC Score**	**Exudates Lesions**	**AUC Score**	
DR Lesions Segmentation							
Custom CNN Segmentation Approach							
[13]	Custom CNN	IDRiD (413)	Microaneurysms	0.463	Hard Exudates	0.795	
	-	-	Haemorrhages	0.637	Soft Exudates	0.711	
[13]	Custom CNN	DDR (12,522)	Microaneurysms	0.105	Hard Exudates	0.555	
	-	-	Haemorrhages	0.359	Soft Exudates	0.265	
[13]	Custom CNN	e-ophtha (463)			Hard Exudates	0.417	
U-Net Segmentation Approach							
[14]	Custom U-Nets	IDRiD (413)	Microaneurysms	0.525	Hard Exudates	0.889	
	-	-	Haemorrhages	0.703	Soft Exudates	0.679	
Collaborative Learning Approach							
[15]	Custom U-Net	IDRiD (413)	Microaneurysms	0.9828	Hard Exudates	0.9935	
	-	-	Haemorrhages	0.9779	Soft Exudates	0.9936	

4.2.1. U-Net Segmentation approach:

U-Net architectures have performed a key role in the advancement of semantic segmentation in medical imaging field. Various deep learning models have been implemented and proposed for the DR detection and diagnosis tasks, that are based on UNet architecture. In the paper [14], Yan et al. proposed an efficient mutual Global-Local UNet 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. The primary contribution of their model is that it incorporates not only the local information but also take full use of context information. The proposed Global-Local UNet segmentation model reported higher AUC scores on hard exudates, microaneurysms, and haemorrhages lesions in comparison to custom CNN segmentation model L-Seg [13].

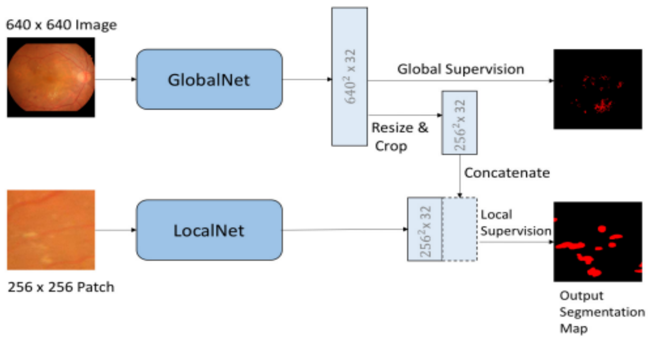

Figure 2. Global-Local UNet segmentation model in [14]

4.2.2. Custom CNN Segmentation approach:

In the study [13], the authors designed an end-to-end model for the segmentation of multiple kind of lesions. In the proposed segmentation model, a pretrained CNN architecture namely VGG-16 was used as a backbone model for feature extraction and learning. The model has the capability to segment hard exudates, soft exudates, microaneurysms, and haemorrhages lesions at the same time. In addition, to improve the learning capability of the proposed segmentation model "L-Seg", they combined all the feature maps by using weighted fusion module. They trained the segmentation model on IDRiD dataset, e-ophtha dataset and a private dataset named DDR. L-Seg model performed way better on IDRiD dataset as compared to other databases and reported an AUC score of 0.795 for hard exudates, 0.711 for soft exudates, 0.463 for microaneurysms, and 0.637 for haemorrhages.

4.2.3. Diabetic Retinopathy Segmentation using Collaborative Learning approach:

Fundus image analysis typically involves two research areas involving lesions segmentation and severity classification. Usually, the researchers studied both areas separately, but the DR severity somehow dependent on the lesions presence. The segmentation of lesions requires pixel-level annotations, and the classification of DR severity usually requires image-level annotations. Zhou et al. [15] 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 that 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. The model reported the higher AUC scores on the IDRiD dataset in comparison to L-Seg model [13] and Global-Local UNet model [14].

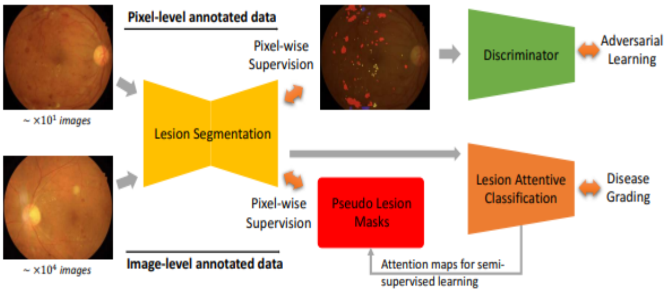

Figure 3. Collaborative Learning Approach in [15]

5. Discussion

This study reviewed nine papers based on different approaches used for automated DR 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 and the second approach is the segmentation of DR-associated lesions such as hard exudates, soft exudates, microaneurysms, and hemorrhages. The key objective of our study is to investigate the performance of various DR classification and segmentation approaches, that will effectively allow future research to be aware of the current progress in the DR domain.

In our study, we have examined multiple DR datasets. These datasets contain fundus images and most of them are publicly available. The number of publicly available DR datasets that constitute the pixel-wise annotations of fundus images is lower as compared to image-wise annotations datasets. To the best of our knowledge, 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. After analyzing the performance of several approaches used for DR diagnosis, we conclude that the performance of segmentation and classification models strongly depends on the size of the training data, as well as the data quality and balance of classes in the dataset. The size of the publicly available dataset such as IDRiD needs to be increased, while the large dataset such as EyePACS dataset needs to be refined because it contains low-quality and miss labeled data.

For DR severity classification, some researchers have built their own custom CNN architecture, while others preferred to use the existing CNN architectures, such as AlexNet, ResNet, VGG and Inception. Use of existing CNNs models with transfer learning approach is simpler and speed up the process of developing new systems. Contrary to it, building a new custom CNN architecture requires a lot of time and effort. But after analyzing the performance of different CNN architectures that are covered in our study, we noticed that the performance of custom CNN architecture is higher as compared to the existing CNN architectures. A custom CNN architecture built by Mobeen-ur-Rehman et al. [8] performed better in comparison to existing architectures. So, this point must be remembered by the researchers during the future research, and more research should be conducted to judge this trend.

We have examined the overall advantage of using ensemble deep learning model over individual DL model. Use of ensemble learning (EL) can help us improve the performance of individual DL models by combining multiple DL models into an integrated ensemble DL model, thus resulting in a more robust and reliable classification of DR severity. An ensemble DL model built by Jiang et al. [7], combined several DL models such as Inception V3, ResNet152, Inception-ResNet-V2 along with the Adaboost classifier for DR severity classification. Their ensemble DL model achieved better performance as compared to individual DL models. Thus, EL can be applied to multiple DL models especially in cases where the individual DL models show low performances as per the expectations. So, in future research, EL approaches should be evaluated for the DR severity classification.

Since the last decade, only few researchers have used the Multipath CNN architectures as feature extractor along with machine learning classifiers for categorizing the DR severity. The process of features extraction is the most critical factor that determines the classification ability of the system. That means, the system with the most efficient feature extractor can have the most accurate classification capability. 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. In the study [9], 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. Thus, most of the shortcomings of traditional CNN methods can be resolved by using Multipath CNN architectures.

For DR lesions segmentation, usually four major types of lesions are detected. Several lesion segmentation models have been developed by researchers [13,14]. But the major concern is that the performance of these lesion segmentation models is not the same across all types of lesions. Specifically, it become noticeable that the AUC scores of hard exudates and soft exudates lesions on the IDRiD dataset are 0.889 ± 0.795 and 0.711 ± 0.679 respectively. Contrary to it, the AUC scores of microaneurysms lesions are 0.525 ± 0.463, while for the haemorrhages lesion the AUC scores are 0.703 ± 0.637. After analyzing these scores, it is clearly evident that the hard exudates lesion is the easiest lesion to be detected during the segmentation process. It can be explained by the reason that the exudates have a much bigger size and have a bright yellow color in comparison to microaneurysms, and haemorrhages lesions that have smaller size, and their color is nearly similar to retina vessels color.

In recent studies, it has been found out that attention mechanisms brought valuable advancements to the way digital images are contextualized and interpreted. Due to the effectiveness of various DL methods, the detection and analysis of DR has become more comprehensive, generalizable, and faster. In study [15], the authors implemented a collaborative learning approach for the segmentation of lesions based on attention mechanism. The model performed the segmentation task with multi lesion mask generational model. Also, a lesion attentive disease grading model is used to improve the DR severity classification. At the same time, the lesion attentive model enhances the lesions maps by using the DR class specific details to improve the lesions segmentation. The collaborative learning model completely outperform the segmentation models proposed in studies [13,14] and reported the highest AUC scores on all types of lesions covered in our study. Researchers can also enhance the future research directions of DR diagnosis by finding the correlation between DR and other leading causes of permanent blindness using attention mechanism. For example, the detection of Diabetic Macular Edema disease is highly likely the possible symptom of developing DR. In the study [12], X. Li et al. proposed a cross disease attention network to grade both the DR and DME by finding the internal relationship between them. Prior to their work, researchers grade either the DR or DME for the clinicians to make patient tailored treatment decisions and totally ignore the internal relationship between DR and DME. Their vital contribution involves the development of disease specific attention module in order to study the important features for diseases, and disease dependent attention module to further find out the relationship between DR and DME diseases. They integrate both the disease specific and disease dependent attention modules into a deep neural network to learn the features and to evaluate the overall performance of the proposed method. Their model reported the best performance results on ISBI 2018 IDRiD challenge dataset and outperforms other multi-task learning methods on the Messidor database. Hence, their research work further illustrates the possibility of future research directions specifically focused on building generalized models in order to effectively find and assess the hidden indicators and symptoms of retina-based diseases that could undoubtedly help the researchers in future.

6. Conclusion

Our study demonstrated the present state of research regarding the detection and diagnosis of DR based on segmentation and classification approaches. By the way, various strategies have laid the foundations for authentic detection and diagnosis of DR. Nonetheless, additional improvements in automated DR diagnosis strategies are still very much required concerning their performance, interpretability, and reliability to meet the needs of ophthalmologists.

References

[1] https://www.who.int/news-room/fact-sheets/detail/diabetes

[2] https://idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html

[3] Wilkinson, C., Ferris, F. L., Klein, R. E., Lee, P. P., Agardh, C. D., Davis, M., Dills, D., Kampik, A., Pararajasegaram, R., & Verdaguer, J. T. (2003). Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology, 110(9), 1677-1682.

[4] Bakator, M. and Radosav, D., 2018. Deep learning and medical diagnosis: A review of literature. Multimodal Technologies and Interaction, 2(3), p.47.

[5] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. Neurocomputing, 187, 27-48. https://doi.org/10.1016/j.neucom.2015.09.116

[6] Li, T., Gao, Y., Wang, K., Guo, S., Liu, H., & Kang, H. (2019). Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. Information Sciences, 501, 511-522. https://doi.org/10.1016/j.ins.2019.06.011

[7] H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma and W. Qian, "An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 2045-2048, doi: 10.1109/EMBC.2019.8857160.

[8] 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, doi: 10.1109/AICAI.2019.8701231.

[9] 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). https://doi.org/10.1007/s13246-021-01012-3

[10] 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. https://doi.org/10.1007/978-3-319-10840-7_14

[11] Wan, S., Liang, Y., & Zhang, Y. (2018). Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers & Electrical Engineering, 72, 274-282. https://doi.org/10.1016/j.compeleceng.2018.07.042

[12] X. Li, X. Hu, L. Yu, L. Zhu, C. -W. Fu and P. -A. Heng, "CANet: Cross-Disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading," in IEEE Transactions on Medical Imaging, vol. 39, no. 5, pp. 1483-1493, May 2020, doi: 10.1109/TMI.2019.2951844.

[13] 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. https://doi.org/10.1016/j.neucom.2019.04.019

[14] 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, doi: 10.1109/ISBI.2019.8759579.

[15] 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, doi: 10.1109/CVPR.2019.00218.