



An overview of big data in Diabetic Retinopathy

Zaid Bin Tariq Baig¹, Sami Rehman², Eman Tariq³

PhD Student, Control and Computer Science, North China Electric Power University, Beijing, China¹

Master Student, Control and Computer Science, North China Electric Power University, Beijing, China²

Abstract: An extensive review is conducted in the perspective of comprehending key strategies interested in creating diabetic retinopathy algorithms throughout this study. The research reveals indicated specialists use retina computer vision algorithms in conjunction with numerical learning techniques to identify abnormalities in eyes, and that they have used control criteria including such vascular areas to do so. We conducted a systematic review of approaches for mechanically identifying and classifying diabetic retinopathy throughout this work. This reviewed in the previous out though those writers use essentially three ways for diagnosing diabetic retinopathy, and a detailed look at methods for evaluating diabetic retinopathy is provided. Significantly the most frequent outcomes of diabetes are diabetic retinopathy. Unfortunately, many patients are unaware of any symptoms until it is too later to treat them efficiently. A pathway will be built for early retinal image analysis and prognosis during the field of counselling through analyzing the retina's, optical nerves, and optical brain Centre's evoked applying suitable. Diabetic retinopathy is among the most devastating chronic diseases in the world, and is also one of the main causes of preventable vision. Timely identification of diabetic retinopathy allows for prompt treatment, and in order to achieve this, Screening programs will require a large amount of effort., particularly automated screening systems. A representative fundus picture database is essential for automated screening tools to function properly. We present a novel diabetic retinopathy database in this study, as well as a review of existing accessible datasets. Our database is the first and only database we are aware of that contains diabetic retinopathy abnormalities and significant fundus structures documented for every image in the database, making it ideal for the creation and assessment of existing and new diabetic retinopathy treatments.

Keywords: Big data in diabetic retinopathy, Diabetic retinopathy and big data, Healthcare and big data, diabetic retinopathy.

I. INTRODUCTION

Diabetes causes diabetic retinopathy, which is an eye disease. Such a syndrome is caused by damage to the coronary arteries of the light-sensitive tissue at the back of the eye (retina). At first, diabetic retinopathy may cause no symptoms or very moderate vision changes. It, on the other hand, has the potential to induce blindness. Anyone with type 1 or type 2 diabetes can develop the illness. This eye issue is more likely to develop the longer you have diabetes and the less well your blood sugar is controlled. When high blood glucose levels damage the blood vessels in the brain, this is known as diabetic retinopathy. These blood arteries have the tendency to rupture and leak. They also could close, limiting blood flowing through. On the retina, aberrant new blood vessels might occur. All of these changes have the propensity to blur your vision. [1] Diabetic retinopathy usually has no symptoms in the early stages. Some participants describe changes in their vision, also including difficulties reading or seeing objects that are far away. These modifications can arise at any age. Arteries in the retina tend to bleed into the vitreous in the terminal stages (the gel-like fluid that fills your eye). You may notice dark, floating dots or streaks that resemble cobwebs if this happens. The spots may clear up on their own, but it's imperative to seek early treatment. Without therapy, the bleed may recur, aggravate, or develop in scarred. [2]. Diabetes can induce high blood sugar levels, which can contribute to DR. Over time, too many amounts of sugar can harm your retina, which detects light and sends information to your brain via a nerve in the corners of your eyes. Diabetes has a deleterious impact on the blood vessels in the body. Sugar impedes the tiny blood vessels that supply your retina, causing them to leak fluid or bleed, resulting in eye injury. To compensate for the clogged blood vessels, new blood vessels form in your eyes that do not work properly. These developing blood vessels have a high risk of bursting or spilling blood. There are frequently no symptoms in the early stages of diabetic retinopathy. Symptoms frequently become more noticeable as the condition progresses Diabetic retinopathy often affects both eyes. The following are some of the disorder's clinical symptoms:

- eyesight problems
- colour vision impairment
- Floaters are translucent patches and black strings that float in front of a person's eyes and shape the future of their vision.
- Peripheral vision is poor.
- a dark or empty patch in the vision's core
- a complete and abrupt loss of vision



Diabetic retinopathy will sometimes culminate in a shattered retina. If inflammatory process drags the retina out from the rear of the eye, this problem can occur. Floating specks in the individual's peripheral vision, blinking lights, and profound vision loss are classic causes. If a detached retina is not treated, a person faces a considerable danger of losing their vision permanently. Glaucoma happens whenever the normal fluid in the eye gets blocked while new blood vessels form. The blockage raises the chances of optic nerve injury and vision loss by causing an accumulation of eye pressure. [3]

II. LITERATURE REVIEW

Serum glucose and blood hypertension treatment are essential in preventing the initiation of DR. In individuals with severe late - onset and proliferative retinopathy, pan-retinal laser photocoagulation reduces the probability of mild and severe retinal detachment by 50%. In eyes with visual loss, focal laser photocoagulation reduces the likelihood of substantial eye problems by 50% to 70%. Patients with proliferative retinopathy and significant macular bleeding benefit from early vitrectomy. When standard treatment fails, route of administration steroid injections may be considered in eyes with chronic vision loss. On the incidence or progression of DR, there is insufficient evidence on the efficacy or safety of serum therapy, medicinal therapies, or activities are undertaken artery bioactive molecules. [4]. The majority of our everyday tasks have evolved significantly in the previous decade. And ever synergy involving new technological developments, innovation, and digital communication also exemplified in digital health. Due to machine learning, we are no longer limited to a descriptive analysis of data, since we may acquire more value by recognizing and anticipating patterns that emerge from inductive reasoning. [5]

Vision loss caused by DR is on the downturn in high-income countries (e.g., the United States and the United Kingdom), due to a number of concerted public health efforts, high knowledge, regional- or nationwide DR inspection, and the accessibility of system for providing, such as anti-vascular endothelial growth factor (anti-VEGF) agents. The actual number of people with DR is predicted to rise as the prevalence of diabetes rises and the global population grows and ages. [6] Moreover, due to a lack of planning, eye-care resources, and restricted access to quality and cheap eye-care, low- and middle-income countries, which suffer the brunt of the diabetes crisis, are unlikely to show similar trends of reduced DR vision to those seen in strong ones. Machine learning software platforms that reveal the reasoning behind a prediction allow for "how much" models, which can be used to understand if and how adjusting particular characteristics could improve outcomes, leading for the detection of the observed value. [7] The diagnosis criteria for T2DM were based on the American Diabetes Association's guidelines from 2003. 13 DR was diagnosed using a macula-centered 45° fundus image and indirect ophthalmoscopy with pupils dilated, according to the International Clinical Diabetic Retinopathy Severity Scale. [14] Two expert ophthalmologists read the fundus photographs and performed the examinations. Microaneurysms, more than 20 intraretinal hemorrhages in each of the four quadrants, definite venous beading in 2+ quadrants, notable intraretinal microvascular anomalies in 1+ quadrant, neovascularization, or vitreous/preretinal hemorrhage were all included in the DR group. Patients with DR were classified as positive samples, whereas those without DR were classified as negative samples. The training and validation sets were generated using the random under-sampling method due to the imbalance in the distribution of positive and negative data. The prediction models were trained using a training set that contained 90% positive samples and 10% negative samples. The rest of the samples were utilized in the validation set to test the machine learning models' ability to predict DR in diabetic patients. [8]. Overall global diabetes number is expected to rise from 285 million in 2010 to 439 million before 2030. • Diabetes continues to be the major cause of legal blindness in those aged 25 to 65 in the Western world. • It is responsible for 1.8 million of the world's 37 million affects an estimated (4.9 per cent). • Diabetes mellitus is evident in 33% of individuals with type 2 diabetes. That is the most precise way to predict diabetic retinopathy. People diagnosed with DM just before age of 3 decades have a 50% chance of developing DR after 10 years, and a 90% chance after 30 years. • Approximately 99 per cent of individuals with type I DM and 60 percent of patients with type II DM have diabetic retinopathy after 20 years of diabetes. Because PDR is driven by abnormally high average blood glucose levels, which are more common in type I diabetes than type II diabetes, this is the case. Maculopathy, on the other hand, is more common in type II diabetes than type I diabetes for unclear reasons. DR rarely develops within about five years of diabetes maybe before adolescents, even though about 5% of type II diabetics develop it at the time of diagnosis. [9] Diabetic retinopathy is the major cause of visual impairment in men between the ages of 25 and 74 in the developed world. Over fifteen years with diabetes, it affects three out of every four diabetic patients. Chronic hyperglycemia is the most common cause of diabetic retinopathy and significant complications. In important clinical studies such as the Diabetes Control and Complications Trial (DCCT) [1] as well as the UK Observational Diabetes Research, the value of long-term glycemic control has been decisively proven (UKPDS). [2,3] Nevertheless, basic processes that cause hyperglycemia to cause diabetic retinopathy and the anatomical abnormalities that are seen histopathologically are still unknown. [10] This is critical to detect the tissue when attempting to diagnose DR. We may evaluate whether patients have DR based on evidence regarding the precise location and thickness of the vessels if the vessels' position can be accurately determined. Nevertheless, since the retina primary issues several other components beyond vessels, vessel tracing is indeed a difficult task. Arterial



monitoring, matching filtration, morphology analysis, distortion modelling, & deep learning are all five types of vessel strategies that have already been suggested. [11]

- Matched flirting
- Morphological processing
- Deformation model
- Vascular tracking
- Deformation models

For medicine, big data is shifting beyond a promise to actuality. Ai algorithms could now diagnose pneumonia using chest x-rays or diabetic retinopathy with fundoscopic images, with results augmenting but sometimes surpassing human patients diagnosed. Artificial intelligence algorithms can now diagnose pneumonia from chest x-rays¹ and diabetic retinopathy from fundoscopic pictures, thanks to recent advances in computer power and deep learning, which have resulted in well-publicized achievements in clinical practice. Predictive analytics has proven to be especially useful in making sense of the gigabytes of data generated by medical records (EHRs). Predictive analytic algorithms have been demonstrated to anticipate & occasionally prevent significant occurrences including cardiac arrest readmissions⁴, chronic obstructive lung disease, and newborn infection. Cancer, with its abundance of data, appears to be a good fit for predictive analytics. Despite the need for better projections of life expectancy, acute care utilization, side effects, and genetic and molecular risk, data analysis application in cancer are limited. Current predictive analytic treatments, we think, could close significant gaps in oncology risk classification methodologies. Predictive analytics applications in pathological interpretation, medication discovery, and population health management are gaining traction, paving the door for future technologies to enter clinical practice. However, physicians, developers, and politicians must address the research, technical, and legal constraints that prevent analytics from being used in oncology in order to realize this promise. Blindness affects 1.5 billion people worldwide, with 0.4 million people suffering from DR. Despite the fact that global blindness and visual impairment have decreased, DR-related blindness has grown from 0.2 million to 0.4 million, with moderate–severe visual impairment increasing from 1.4 million to 2.6 million between 1990 and 2015. Despite the fact that a combination of social, nutritional, and medical assistance has averted or slowed the advancement of DR, it remains a global issue due to the pandemic growth of DM, which has a 25-fold higher risk of vision loss. Due to a lack of funds and skills, screening and treatment of DR is more difficult in underdeveloped nations. Among the Asia–Pacific area, DR is the major cause of new cases of blindness in middle-aged and older people. It is thought to be the cause of 51 percent of blindness and 56 percent of visual impairment cases worldwide, yet only 28 percent to 84 percent of DM patients are aware of the disease. ¹⁴ In Africa, DR affects 7% to 62.4 percent of the population, with 15% having severe DR. Ethiopia is one of the top four nations in Sub-Saharan Africa with the highest adult diabetic populations (3.8 percent), yet there aren't enough studies, screening protocols, uniform referral criteria, or retinal photocoagulation. As a result, the goal of this study was to determine the prevalence of DR and its causes among diabetic patients in this location.[12]

III. DIABETIC RETINOPATHY AND BIG DATA

Big data is a term for massive data collections (structured or unstructured) that are hard to evaluate, manage, or manage using leading technologies or methods due to the quantity, diversity, and velocity. ¹ To make big data more trustworthy, we must consider the validity and quality of the data, since this result in more dependable & useful information. If it comes to maintaining such a large amount of data, the key is to figure out how to make sensible and smart use of everything. It entails assessing, analyzing, integrating organizing data in a manner that allows for greater judgement, reduced costs, and so more effectively and actively in healthcare systems.

One important tool for the reliable evaluation and comparison of medical computer vision applications is a database that contains a selected group of high-quality medical images that are indicative of diabetic retinopathy and have been confirmed by professionals. Along with the images, data about the medical results is required. The ground truth is the knowledge gathered from the results. An appropriate algorithm will take an image as input and produce the results or report that matches the ground truth. The consistency of the algorithms is assessed and compared during the examination. Sensitivity and specificity are the most critical performance measures for medical diagnostic procedures. Even though the fact that diabetic retinopathy results do have geographical placements in the fundus, detection limits are established on an image basis - an image either contains a certain finding or it doesn't. Computer vision researchers must guarantee is critical for computer vision researchers to guarantee that the information extraction diabetic retinopathy findings logically conform to the findings identified by experts, in other terms, that they present at the same point in the image. As a result, the more extensive expert ground truth includes the description of how diabetic retinopathy findings appear visually. Diabetic retinopathy has been the major cause of visual impairment in the developed world's working-age population. Over 93 million people in the world are severely impacted. Diabetes affects 29.1 million individuals worldwide, and 347 million patterns are expanding, as per the Centers for Disease Control and Preventative measures., and according to World Health Organization. Diabetic retinopathy (DR) is an eye condition caused by long-term diabetes. Around 40% to 45 per cent of diabetics in the U. S. must be in some stage of the condition. If DR is diagnosed early



enough, vision deterioration can be reduced or avoided.[13] However, this can be complicated so the condition typically has some manifestations until it's too early to cure. Diagnosis DR is now a night before going to bed and painstaking process that includes a knowledgeable clinician assessing and analyzing digital retinal fundus pictures of the retina. By the time human readers give final reviews, which is usually a day or two later, the delayed results mean missed follow-up, miscommunication, and delayed treatment. The occurrence of lesions correlated with the disease's vascular disorders can help clinicians recognize DR. While this strategy is effective, it necessitates a large number of resources. In locations where the frequency of diabetes in the local population is high and DR detection is most needed, the essential competence and equipment are frequently inadequate. The infrastructure intended to avoid blindness due to DR will become more and more insufficient as the international diabetes federation remains high. The requirement for a systematic and automated DR aspects to consider has long been recognized, and prior efforts using image classification, pattern recognition, including algorithms have made significant progress. The purpose of this competition is to push an automated detection system to its utmost using retinal fundus photography as input, presumably producing in designs with realistic testing is commonly. To clarify the effect of the competing models on strengthening DR monitoring, services will be public contributed.[14]

Under this research, we provide a Hadoop framework-based Diabetic Retinopathy (DR) detection technique that can detect early symptoms of diabetes from ocular retinal pictures. The retinal pictures are divided into five classes in the proposed approach.

- No Diabetic Retinopathy (DR)
- Mild DR
- Moderate DR
- Severe DR
- Proliferative DR

Several eye-related diseases can indeed be detected using retinal or fundus imaging. Retinal scans can show blood vessels, the optic disc, the macula, and thus the fovea, among other features. Such visual picture structures provide information that is utilized to diagnose & treat for a variety of retina illnesses. According to the Preventive Care Eye Disease Study (ETDRS), diabetic retinopathy can be characterized as either non-proliferative or proinflammatory diabetic retinopathy (PDR). The early stage of the disease is known as NPRD, and it could be classified as weak, intermediate, extreme, or life threatening. PDR, but in the other hand, is just a more important level of condition that will be divided into early and advanced stages.[15]

For the categorization of diabetic retinopathy pictures, the suggested scheme uses three distinct steps: feature extraction, feature reduction, and image classification. The Histogram of Oriented Gradients is utilized as a feature descriptor to represent each of the Diabetic Retinopathy images in the early stages of the method. HOG characteristics are dimensionally reduced using Principal Component Analysis (PCA). The K-Nearest Neighbors classifier is utilized in the last stage of the algorithm to categorize the retinal images into distinct classes in a distributed system. Tests were conducted on a large number of high-resolution retinal pictures captured under various imaging settings.[16] Every subject has visuals for both the left and right eyes. In terms of some common performance evaluation platform is utilized with MapReduce and Mahout framework for programming, Hadoop has been determined to deliver superior results than existing ways in handling such massive datasets. The proposed scheme's outcomes are compared to some of the most competitive state-of-the-art approaches. The proposed method assesses.[17]

There are therapy alternatives available for the vast majority of diabetics who want to avoid or delay the start of diabetic retinopathy and vision loss. According to the DCCT and the UKPDS, glycemic and blood pressure control can help patients with diabetes avoid or delay the development of eye issues. A large proportion of patients with severe NPDR and PDR, as well as retinal oedema, can benefit from prompt laser photocoagulation therapy to avert vision loss.[18] Because a considerable number of individuals with vision-threatening illnesses might not always display signs, continuous retinopathy screening is indeed an essential method. A correct recognition of diabetic retinopathy using retinal fundus images has been a hot topic in the medical image processing field for a long time. [19]The high potential for new goods in the healthcare profession, as well as the possibilities for lower medical costs, justify the interests. Unfortunately, thanks to the shortage of a globally acknowledged and significant picture dataset besides confirmed real data and tight evaluation phase, the maturity of state-of-the-art algorithms cannot be assessed. The database's challenges and issues are examined from the medical, image processing, and security perspectives in this paper. The assessment method is proposed depending on the debate, and a prototype image database containing contextual information is demonstrated.[20]

IV. CONCLUSION

Diabetic retinopathy is incurable. Usually performed even before the retina is harmed, laser analysis (photocoagulation) is usually highly effective in preventing visual impairment. The surgical removal of translucent gel can enhance eyesight if the retina has not been severely damaged (vitrectomy). In some cases of proliferative diabetic retinopathy, an anti-inflammatory or anti-vascular endothelial growth factor drug injection can aid in the vascular endothelium. Because symptoms do not appear until the disease has progressed to the point of becoming fatal, early detection through routine screening is critical. Warning symptoms of DR can be found in non-proliferative diabetes mellitus, even though it is crucial to diagnose and assess DR through its beginning stages. Whenever a diabetic has proper eye care on a regular basis and therapy, when necessary, DR will occasionally result in total vision. Throughout this DR research, a significant amount of time is spent looking for hemorrhages, microaneurysms and exudates, diabetic macular edema, and aberrant fresh faces arteries in image data as indicators of the presence of retinopathy. This research aids with in timely identification of retinopathy, which can lead to chronic visual loss if not treated promptly. The report discusses the scientists' studies for detecting diabetic retinopathy. Tech people and researchers who need to leverage continuing studies in this field may benefit from our effort.

REFERENCES

- [1] U. Behera and A. Brar, "Socioeconomic status and diabetic retinopathy in India," *Indian J. Ophthalmol.*, vol. 69, no. 11, p. 2939, 2021, doi: 10.4103/ijo.ijo_1508_21.
- [2] A. Das, G. Prashanthi, T. Das, R. Narayanan, and P. Rani, "Clinical profile and magnitude of diabetic retinopathy: An electronic medical record-driven big data analytics from an eye care network in India," *Indian J. Ophthalmol.*, vol. 69, no. 11, p. 3110, 2021, doi: 10.4103/IJO.IJO_1490_21.
- [3] "Diabetic Retinopathy | National Eye Institute." <https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy> (accessed Jan. 29, 2022).
- [4] R. Rajalakshmi et al., "Correlation between markers of renal function and sight-threatening diabetic retinopathy in type 2 diabetes: a longitudinal study in an Indian clinic population," *BMJ open diabetes Res. care*, vol. 8, no. 1, May 2020, doi: 10.1136/bmjdr-2020-001325.
- [5] J. B. Jonas and C. Sabanayagam, "Epidemiology and Risk Factors for Diabetic Retinopathy," pp. 20–37, 2019, doi: 10.1159/000486262).
- [6] S. Qummar, F. G. Khan, S. Shah, A. Khan, A. Din, and J. Gao, "Deep Learning Techniques for Diabetic Retinopathy Detection," *Curr. Med. imaging*, vol. 16, no. 10, pp. 1201–1213, Feb. 2020, doi: 10.2174/1573405616666200213114026.
- [7] N. Musacchio et al., "Artificial Intelligence and Big Data in Diabetes Care: A Position Statement of the Italian Association of Medical Diabetologists," *J. Med. Internet Res.*, vol. 22, no. 6, Jun. 2020, doi: 10.2196/16922.
- [8] W. Li et al., "Predictive model and risk analysis for diabetic retinopathy using machine learning: a retrospective cohort study in China," *BMJ Open*, vol. 11, no. 11, p. e050989, Nov. 2021, doi: 10.1136/BMJOPEN-2021-050989.
- [9] "Diabetic retinopathy ."
- [10] K. Landau and M. Kurz-levin, "Retinal disorders," *Handb. Clin. Neurol.*, vol. 102, pp. 97–116, 2011, doi: 10.1016/B978-0-444-52903-9.00010-8.
- [11] Y. H. Li, N. N. Yeh, S. J. Chen, and Y. C. Chung, "Computer-Assisted Diagnosis for Diabetic Retinopathy Based on Fundus Images Using Deep Convolutional Neural Network," *Mob. Inf. Syst.*, vol. 2019, 2019, doi: 10.1155/2019/6142839.
- [12] M. Tilahun, T. Gobena, D. Dereje, M. Welde, and G. Yideg, "Prevalence of Diabetic Retinopathy and Its Associated Factors among Diabetic Patients at Debre Markos Referral Hospital, Northwest Ethiopia, 2019: Hospital-Based Cross-Sectional Study," *Diabetes, Metab. Syndr. Obes. Targets Ther.*, vol. 13, pp. 2179–2187, Jun. 2020, doi: 10.2147/DMSO.S260694.
- [13] F. Bandello et al., "Diabetic Macular Edema," pp. 102–138, 2017, doi: 10.1159/000455277).
- [14] Kaggle.com, "Diabetic Retinopathy Detection Challenge," 2015. <https://www.kaggle.com/c/diabetic-retinopathy-detection> (accessed Jan. 29, 2022).
- [15] V. E. Castillo Benítez et al., "Dataset from fundus images for the study of diabetic retinopathy," *Data Br.*, vol. 36, p. 107068, Jun. 2021, doi: 10.1016/J.DIB.2021.107068.
- [16] A. Pak, A. Ziyaden, K. Tukeshev, A. Jaxylykova, and D. Abdullina, "Comparative analysis of deep learning methods of detection of diabetic retinopathy," <http://www.editorialmanager.com/cogenteng>, vol. 7, no. 1, Jan. 2020, doi: 10.1080/23311916.2020.1805144.
- [17] C. Gilbert, I. Gordon, C. Mukherjee, and V. Govindhari, "Guidelines for the prevention and management of diabetic retinopathy and diabetic eye disease in India: A synopsis," *Indian J. Ophthalmol.*, vol. 68, no. 13, pp. S63–S66, Feb. 2020, doi: 10.4103/ijo.IJO_1917_19.
- [18] "MillennialEYE | Big Data in Ophthalmology." <https://millenniaeye.com/articles/nov-dec-19/big-data-in->



ophthalmology/ (accessed Feb. 15, 2022).

- [19] R. B. Parikh, A. Gdowski, D. A. Patt, A. Hertler, C. Mermel, and J. E. Bekelman, “Using Big Data and Predictive Analytics to Determine Patient Risk in Oncology.,” Am. Soc. Clin. Oncol. Educ. book. Am. Soc. Clin. Oncol. Annu. Meet., vol. 39, no. 39, pp. e53–e58, Jan. 2019, doi: 10.1200/EDBK_238891.
- [20] D. S. Fong et al., “Retinopathy in Diabetes,” Diabetes Care, vol. 27, no. suppl_1, pp. s84–s87, Jan. 2004, doi: 10.2337/DIACARE.27.2007.S84.