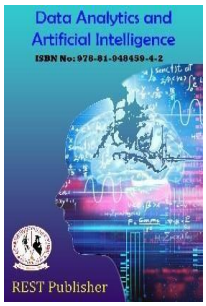
Data Analytics and Artificial Intelligence

Vol: 4(2), 2024

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/

DOI: https://doi.org/10.46632/daai/4/2/8



Detection of Diabetic Retinopathy Using KNN & SVM Algorithm

E. Kamalanaban, J. Senthil Murugan, M. Gayathri, M. Kathambari, V. Mancy Arokiya Mary

Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College

*Corresponding Author Email id: j.senthilmurugan@velhightech.com

Abstract: *Diabetic Retinopathy (DR) is a medical condition caused by diabetes. The development of retinopathy significantly depends on how long a person has had diabetes. Initially, there may be no symptoms or just a slight vision problem due to impairment of the retinal blood vessels. Later, it may lead to blindness. Recognizing the early clinical signs of DR is very important for intervening in and effectively treating DR. Thus, regular eye check-ups are necessary to direct the person to a doctor for a comprehensive ocular examination and treatment as soon as possible to avoid permanent vision loss. Nevertheless, due to limited resources, it is not feasible for screening. As a result, emerging technologies, such as artificial intelligence, for the automatic detection and classification of DR are alternative screening methodologies and thereby make the system cost-effective. People have been working on artificial- intelligence-based technologies to detect and analyze DR in recent years. This study aimed to investigate different machine learning styles that are chosen for diagnosing retinopathy. Thus, a bibliometric analysis was systematically done to discover different machine learning styles for detecting diabetic retinopathy. The data were exported from popular databases, namely, Web of Science (WoS) and Scopus. These data were analyzed using Biblioshiny and VOS viewer in terms of publications, top countries, sources, subject area, top authors, trend topics, co-occurrences, thematic evolution, factorial map, citation analysis, etc., which form the base for researchers to identify the research gaps in diabetic retinopathy detection and classification.*

Keywords: *Diabetic Retinopathy (DR), Artificial Intelligence (AI), Bibliometric Analysis, Web Of Science (WOS), Biblioshiny, VOS viewer, thematic evolution, factorial map, citation analysis.*

1. INTRODUCTION

Diabetic retinopathy is an ever-increasing problem. Early screening and timely treatment of the same can reduce the burden of sight threatening retinopathy. Any tool Which can aid in quick screening of this disorder and minimize requirement of trained human resource for the same would probably be a boon for patients and Ophthalmologists. DR is an eye disease known to cause moderate to severe visual loss and is the leading cause of blindness in working-age people suffering with long standing diabetes. The health burden is accentuated by the huge per capita cost. This has further increased since the introduction of anti VEGF agents. Very often the disease does not show over t symptoms until it reaches an advanced stage; however, if detected early on, vision impairment can be averted by early intervention which is also the most cost-effective option. In view of the alarming increase in the number of people with diabetes and dearth of trained retinal specialists and ophthalmologists, a computer-based analysis of the fund us images by an automated approach would lessen the burden of the health systems in screening for DR and offer a near ideal system for its management.

Therefore, screening will be valuable at any stage of the disease and will also be helpful in avoiding blindness among 90% patients. Clinical and photographic methods were used to assess retinopathy during the examinations of diabetic patients enrolled in the Early Treatment Diabetic Retinopathy Study (ETDRS). Inanalyzing available data from eyes randomly selected for deferral of treatment, the authors compare the clinical detection (including contact lens biomicroscopy) with photographic detection (30° stereoscopic color fundus photographs) of diabetic macular edema. Based on clinical detection, 53% (1778 patients) had hard exudates within 1 discdiameter (DD) of the center of macula, ,56% (1868 patients) had retinal thickening with in this region, and 31% (1027 patients)

had thickening at the center of macula. These analyses show agreements of 83, 78, and 83% between retinal specialists and photographic graders when assessing these three characteristics, respectively. Agreement was 81 % in the detection of macular edema for which treatment is indicated (clinically significant macular edema).

Each method has its advantages but in general there was close agreement between these methods, particularly for clinically significant macular edema, which supports there liability of each method. Regardless of the type of diabetes, all individuals diagnosed with DM need regular and repetitive annual retinal screening for timely detection and apt treatment of diabetic retinopathy (DR).Conventionally, retinopathy screening is done by fundus examination by ophthalmologists or with the help of color fundus photography using conventional fundus cameras (mydriatic or non-mydriatic)by trained eye technicians or optometrists. The primary issue is the grading of the retinal images by ophthalmologists or trained persons, whose numbers are very scarce compared to the load of patients requiring screening. Second, some of these patients are based in rural areas and can't visit an eye care provider. Thirdly, as such follow ups are required for years together, the attitude, and/or behavioral aspects negatively impact the patients practice despite knowledge of consequences. These issues can be solved with provision of an automated imaging system within easy reach of the patient. Hence, there has been an increasing interest in the development of automated analysis software using computer machine learning/artificial intelligence (AI) for analysis of retinal images in people with diabetes thus solving at least some part of the problem.

2. RELATED WORKS

Diabetic retinopathy occurs due to long-term diabetes with changing blood glucose levels and has become the most common cause of vision loss world wide. It has become a severe problem among the working-age group that needs to be solved early to avoid vision loss in the future. Artificial intelligence-based technologies have been utilized to detect and grade diabetic retinopathy at the initial level. Early detection allows for proper treatment and, as a result, eyesight complications can be avoided. The in-depth analysis now details the various methods for diagnosing diabetic retinopathy using blood vessels, micro aneurysms, exudates, macula, optic discs, and hemorrhages. In most trials, fund us images of the retina are used, which are taken using a fundus camera. This survey discusses the basics of diabetes, its prevalence, complications, and artificial intelligence approaches to deal with the early detection and classification of diabetic retinopathy. The research also discusses artificial intelligence-based techniques such as machine learning and deep learning. New research fields such as transfer learning using generative adversarial networks, domain adaptation, multitask learning, and explainable artificial intelligence in diabetic retinopathy are also considered. A list of existing datasets, screening systems, performance measurements, biomarkers in diabetic retinopathy, potential issues, and challenges faced in ophthalmology, followed by the future scope conclusion, is discussed. To the author, no other literature has analyzed recent state-of-the-art techniques considering the PRISMA approach and artificial intelligence as the core. Diabeticretinopathy is a leading cause of blindness among working-age adults. Early detection of this condition is critical for good prognosis. In this paper, we demonstrate the use of convolutional neural networks (CNNs) on color fundus images for the recognition task of diabetic retinopathy staging. Diabetic Retinopathy (DR) is an eye disease associated with chronic diabetes. DR is the leading cause of blindness among working aged adults around the world and estimated it may affect more than 93 million people. Progression to vision impairment can be slowed or controlled if DR is detected intime, however this can be difficult as the disease often shows few symptoms until it is too late to provide effective treatment. Diabetic Retinopathy (DR) is a common complication of diabetes mellitus,which causes lesions on the retina that effect vision. If it is not detecte dearly, it can lead to blindness. Unfortunately, DR is not are versible process, and treatment only sustains vision. DR early detection and treatment can significantly reduce the risk of vision loss. The manual diagnosis process of DR retina fund us images by ophthalmologists is time, effort, and cost consuming and proneto misdiagnosis unlike computer-aided diagnosis systems. For this article, the recent state-of-the-art methods of DR color fund us images detection and classification using deep learning techniques have been reviewed and analyzed.

Diabetic Retinopathy (DR) is a complication of diabetes that causes the blood vessels of the retina to swell and to leak fluids and blood. DR can lead to a loss of vision if it is in an advanced stage. Worldwide, DR causes 2.6% of blindness. The possibility of DR presence increases for diabetes patients who suffer from the disease for a long period. This paper systematically reviews the recent progress in diabetic retinopathy screening. It provides an integrated overview of the current state of knowledge of emerging techniques using artificial intelligence integration in national screening programs around the world. Existing methodologic all approaches and research insights are evaluated. An understanding of existing gaps and future reactions is created. Over the past decades, artificial intelligence has emerged into the scientific consciousness with breakthroughs that are sparking increasing interest among computer science and medical communities. Diabetes is a chronic end organ disease that occurs when the pancreas does not secrete enough insulin or the body is unable to process it properly. Overtime, diabetes affects the circulatory system, including that of there tina. Diabetic retinopathy is a medical

condition where the retina is damaged because fluid leaks from blood vessels into the retina. Ophthalmologists recognize diabetic retinopathy based on features, such as blood vessel area, exudes, hemorrhages, microaneurysms and texture. The IDP has high sensitivity and specificity to detect RDR. Computer analysis of retinal photographs for DR and automated detection of RDR can be implemented safely into the DR screening pipeline, potentially improving access to screening and health care productivity and reducing visual loss through early treatment. Increasing health care productivity is a prerequisite to improve health care affordability.

Automation has improved productivity in many sectors of the economy, whereas in health care, productivity has remained stagnant in the last 20 years.1 Regular eye examinations are necessary to diagnose diabetic retinopathy (DR) at an early stage, when it can be treated with the best prognosis and visual loss delayed or deferred.2-4 In 2010, US eye care practitioners examined less than 60% of the estimated 23 million people with diabetes, leaving millions of people at risk for potentially preventable visual loss and blindness. The International Clinical Diabetic Retinopathy (ICDR) severity scale was formulated by a consensus of international experts to standardize and simplify DR classification (Table 1) to improve communication and coordination of care among physicians caring for patients with diabetes.13 The ICDR classification simplified the Early Treatment Diabetic Retinopathy Study (ETDRS) classification for non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR) because the latter classification had proved unwieldy in clinical care.14-17 In the present study, 3 fellowship-trained retinal experts (D.P.H., J.D.W., D.F.W.) independently graded retinal images of each eye from people with diabetes using the ICDR severity level scales and a modified definition of macular edema (ME), namely, any retinal thickening, exudate, or microaneurysm within 1 disc diameter of the fovea. The objective of the present study is to determine the sensitivity and specificity of the Iowa Detection Program (IDP) to detect referable diabetic retinopathy (RDR), which we defined as more than mild NPDR as defined by the ICDR and/or ME.

3. PROPOSED SYSTEM AND METHODOLOGY

Machine learning process mainly include two parts, training set followed by validation set. This process occurs by providing large number of training data i.e., thousands of retinal images of varying grades of DR to the machine/system as the training set. In general, using ML methods with a high processing speed, low computational cost, and interpretable decisions is preferred to DCNNs. However, the automatic detection of subtle lesions such as MA did not reach acceptable values. In this review, we collected 2 pure ML-involved models and 6 un-ML methods. As reported in a study by Ali Shah at el. they detected MA using color, Hessian and curvelet-based feature extraction and achieved a SE of 48.2%. Huang et al. focused on localizing NV through using the Extreme Learning Machine (ELM). This study applied Standard deviation, Gabor, differential in variant, and an isotropic filters for this purpose and with the final classifier applying ELM. This network performed as well as an SVM with lower computations. Support Vector Machine is a supervised machine learning algorithm which is extensively used for both classification and regression day to day problems. It is mostly used in classification problems rather than regression problems. In the SVM algorithm, we will have a number of features. We can plot these each data item has is a features as a point in n-dimensional space where the value of each feature represent the value of a particular coordinate in the n dimensional space. Then we classify the plotted data points into n classes by means of a hyper plane. The k-nearest neighbors (KNN) algorithm is a simple and it is easy-to-implement focused on supervised machine learning algorithm. It is mainly used to solve both classification and regression problems. A supervised machine learning algorithm is one that pointed on labelled input data from user-dataset, directed to learn a function. The function produces an appropriate output when a new unlabelled data is feed on the algorithm. KNN captures the idea of similarity which is often called distance / proximity / closeness. Here we are calculating the distance between points on a graph. This distance is used to classify the given data. That is less distance with data point suggests that higher similarity.

4. ARCHITECTURE DIAGRAM

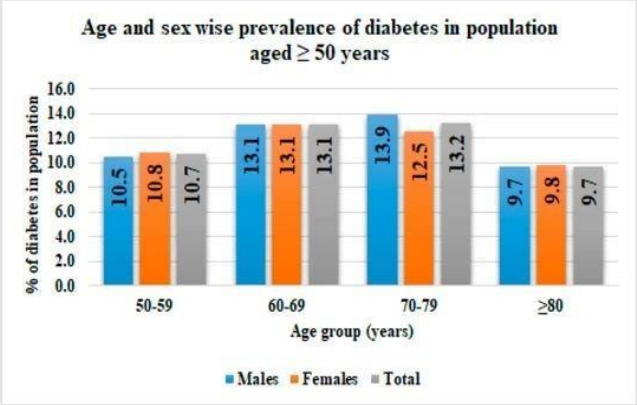


FIGURE 1

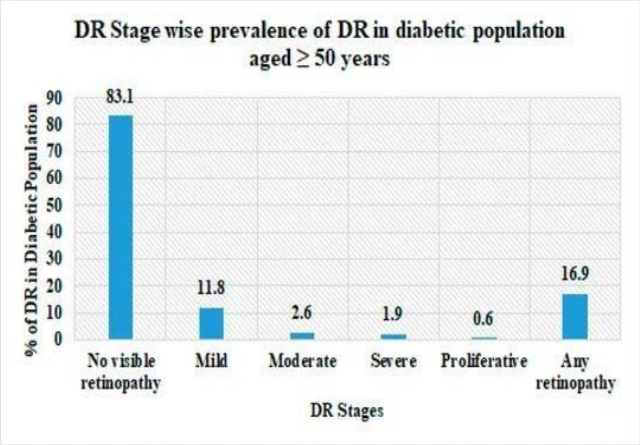


FIGURE 2

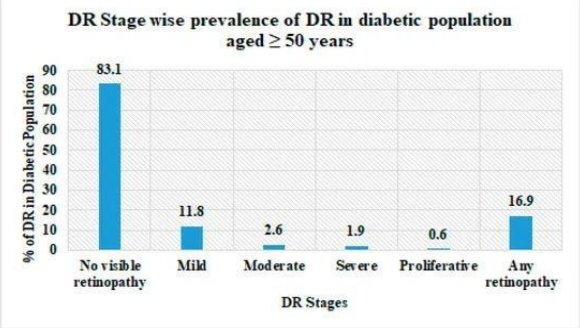


FIGURE 3

5. HARDWARE & SOFTWARE REQUIREMENTS

GPU (Graphics Processing Unit):

- Graphics processing unit, a specialized processor originally designed to accelerate graphics rendering. GPUs can process many pieces of data simultaneously, making the useful for machine learning, video editing, and gaming applications.
- A graphics processing unit (GPU) is an electronic circuit that can perform mathematical calculations at high speed. Computing tasks like graphics rendering, machine learning (ML), and video editing require the application of similar mathematical operations on a large dataset.
- How does a GPU work? GPUs work by using a method called parallel processing, where multiple processors handle separate parts of a single task. AGPU will also have its own RAM to store the data it is processing.

Memory (RAM):

- RAM is a common computing acronym that stands for random-access memory. Sometimes it's called PC memory or just memory. In essence, RAM is your computer or laptop's short-term memory. It's where the data is stored that your computer process or need storun your applications and open your files.
- RAM is a temporary memory bank where your computer stores data it needs to retrieve quickly. RAM keeps data easily accessible so your processor can quickly find it without having to go into long- term storage to complete immediate processing tasks.
- Although all RAM basically serves the same purpose, there are a few different types commonly in use today: Static RAM (SRAM) Dynamic RAM(DRAM) Synchronous Dynamic RAM (SDRAM).

Camera/ Retinal Imaging Device:

- ❖ Retinal imaging is a diagnostic test that creates high-quality digital images of the inner, back surface of your eye. It allows the diagnosis of many eye conditions like diabetes-related retinopathy, glaucoma and macular degeneration. Your provider will tell you how often you need retinal imaging.
- ❖ There are two common types of eye exame quipment used for retinal imaging: tabletop cameras and handheld cameras.

TensorFlow, PyTorch, or scikit-learn:

- ❖ Tensor Flow can be used to develop models for various tasks, including natural language processing, image recognition, handwriting recognition, and different computational-based simulations such as partial differential equations.
- ❖ PyTorch is a fully featured framework for building deep learning models, which is a type of machine learning that's commonly used in applications like image recognition and language processing. Written in Python, it's relative lyeasy for most machine learning developers to learn and use.
- ❖ Scikit-Learn, also known as sklearn is a python library to implement machine learning models and statistical modelling. Through scikit-learn, we can implement various machine learning models for regression, classification, clustering, and statistical tools for analyzing these models.

OpenCV:

- ❖ OpenCV is a great tool for image processing and performing computer vision tasks. It is an open-source library that can be used to perform tasks like face detection, objection tracking, landmark detection, and much more.
- ❖ OpenCV is the huge open-source library for computer vision, machine learning, and image processing and now it plays a major role in real- time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even the handwriting of a human.

PyCharm:

- ❖ PyCharm is a dedicated Python Integrated Development Environment (IDE) providing a wide range of essential tools for Python developers, tightly integrated to create a convenient environment for productive Python, web, and data science development.
- ❖ It offers a wide range of features suitable for beginners, such as code completion, syntax highlighting, and integrated debugging tools. PyCharm also provides support for popular web frameworks like Django and Flask, making it an excellent choice for web development beginners.

Django:

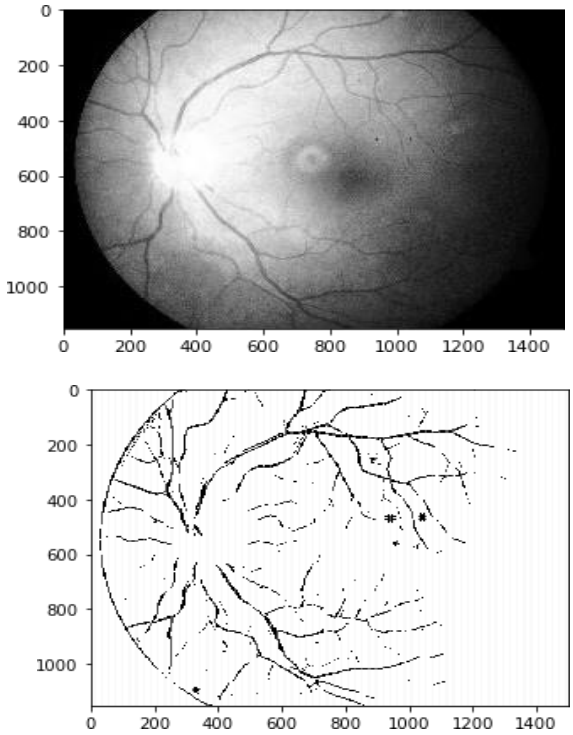
- Django can be (and has been)used to build almost any type of website from content management systems and wikis, through to social networks and news sites. It can work with any client-side framework, and can deliver content in almost any format (including HTML, RSS feeds, JSON, and XML).
- Django, a high-level Python web framework, is renowned for its rapid development capabilities, security features, and scalability. Its modular design facilitates code reuse, while its security measures, scalability, large community support, and extensive libraries contribute to its dominance.

Experimental Results:

```
from scipy import misc
from PIL import Image
from skimage import exposure
from sklearn import svm

import scipy
from math import sqrt,pi
from numpy import exp
from matplotlib import pyplot as plt
import numpy as np
import glob
```

```
import matplotlib.pyplot as pltss
import cv2
```



Future enhancements:

The proposed approach is suitable even for small datasets. New techniques based on deep learning are data hungry but show impressive performances in different classification tasks including DR. Future work includes benchmarking the performance of deep learning techniques and the proposed texture-based features in a small dataset. And could furthermore implemented into an application.

6. CONCLUSION

The increase in diabetes cases limits the ability of current manual testing. New algorithms for assisted diagnosis are becoming very important today. Early detection of diabetes can help the patients and limits the bad health consequences such as blindness. Using retinal fundus images can help automate the diagnosis. Micro hemorrhages and aneurysms, known as HEM, are the early signs of diabetic retinopathy (DR) and are difficult to identify because of their similarities with normal parts of a healthy human. Other problems such as non-uniform lighting, low contrast, etc. can lead to a bad diagnosis. Texture based techniques for DR detection were proposed in the past. We have implemented automatic detection using ML algorithms SVM and KNN.

a) The SVM algorithm performed best interms of diagnostic accuracy 96.62% in the overall analysis
b) The KNN algorithm was performed on the overall data to be accurated percentage of 94.38%
c) Machine learning automatic disease detection models , especially the SVM, showed good diagnostic accuracy and high sensitivity and have potential for use in screening for diabetic retinopathy

REFERENCES

[1]. K. Jain, Jian chang Mao and K. M. Mohiuddin , "Artificial neural networks :a tutorial," in Computer, vol. 29, no. 3, pp. 31-44, doi: 10.1109/2.485891.
[2]. Emil M Petriu, Professor, University of Ottawa, "Neural Networks: Basics", Carlos Gershenson, "Artificial Neural Networks for Beginners", arxiv.org
[3]. S.SinghandF. Nasoz,"Facial Expression Recognition with Convolutional Neural Networks," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2023, pp. 0324-0328, doi: 10.1109/CCWC47524.2020.9031283.
[4]. Alejandro Martín, Raúl Lara-Cabrera, Félix Fuentes-Hurtado, Valery Naranjo, David Camacho, EvoDeep: A new evolutionary approach for automatic Deep Neural Networks parametrisation, Journal of Parallel and Distributed Computing, Volume 117, 2019.
[5]. M.Alam, M.D.Samad, L.Vidyaratne, A.Glandon, K.M. Iftekharuddin, Survey on Deep Neural Networks in Speech and Vision Systems, Neuro computing, Volume 417.
[6]. N. Carlini and D. Wagner, "Towards Evaluating the Robustness of Neural Networks,"2017IEEE Symposium on Security and Privacy(SP),SanJose,CA,USA, 2017, pp. 39-57,doi: 10.1109/SP.2021.49.

[7]. N. Carlini and D. Wagner, "Towards Evaluating the Robustness of Neural Networks, "2017 IEEE Symposium on Security and Privacy (SP), SanJose, CA, USA, 2017, pp. 39-57, doi: 10.1109/SP.2020.49.
[8]. K. Mohiuddin and J. Mao, "A Comparative Study of Different Classifiers for Handprinted Character Recognition", Pattern Recognition in Practice IV, pp.437-448, 2021
[9]. Y. Le Cun, "Back-Propagation Applied to Handwritten Zip Code Recognition", Neural Computation, vol. 1, pp. 541-551, 1989.
[10]. J. Hertz, A. Krogh and R.G. Palmer, Introduction to the Theory of Neural Computation, Reading, Mass.,, 1991.
[11]. S.Haykin, Neural Networks: A Comprehensive Foundation, NewYork,,1994.
[12]. M. Minskyand S.Papert, Perceptrons: An Introduction to Computational Geometry, Cambridge, Mass.,, 1969.
[13]. ”Application: Facial Expression Recognition" in Machine Learning in Computer Vision. Computational Imaging and Vision, Dordrecht: Springer, vol. 29,
[14]. Charles Robert Darwin, The expression of the emotions in man and animals, London: John Murray London J.Deng, W.Dong, R.Socher, L.-J.Li,K.Li and L.Fei-fei,"Imagenet: A large- scale hierarchical image database", Computer Vision and Pattern Recognition.