International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-09, Issue-08, Nov 2023

Automated Papilledema Detection: Harnessing Machine Learning

Jay Sawant, Student, Vidyalankar Institute of Technology, Mumbai, India, jay.sawant@vit.edu.in

Niti Chikhale, Student, Vidyalankar Institute of Technology, Mumbai, India, niti.chikhale@vit.edu.in

Priyanka Shrivastava, Assistant Professor, Vidyalankar Institute of Technology, Mumbai, India, priyanka.shrivastava@vit.edu.in

Abstract - **Vision and eye health have wide-ranging and significant effects on the economy, sustainable development, health, and many other facets of life. Eye diseases and vision impairments can affect an individual's mental and physical health to a large extent if timely treatment is not provided. Papilledema is one of the eye ailments and is known as the swelling of the optic disc caused due to an aberrantly high intracranial pressure. Papilledema can be rendered harmless if it is detected at an early stage. Nowadays, Artificial Intelligence is playing a vital role in the early detection of various types of diseases due to which early diagnosis and treatment is provided to the patient making his recovery even quicker. Machine Learning algorithms have been deployed in many areas of the medical field to increase the efficiency of recognition of diseases. This approach involves the use of machine learning for the classification of Papilledema. The dataset utilized in this approach consists of eye MRI images segregated into three classes. The images are processed, features are extracted using a pre-trained ResNet-18 network and are fed into four classifiers namely Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest Classifier (RFC), and Gradient Boosting Classifier (GBC). A comparative analysis between the results of the four models has been drawn. The RFC model gave the highest accuracy of 97.22% for the validation dataset whereas the KNN model outperformed the others with an accuracy of 97.87% on the test dataset.**

Keywords - Artificial Intelligence, Classification, Machine Learning, Papilledema

I. Introduction

According to the statistics presented by the World Health Organization (WHO), about 285 million people with visual impairment live in the world out of which 39 million are blind and 246 million suffer from low vision [1]. Prevention and treatment of eye diseases are very important as the human eye is a vital sensory organ. Some of the diseases that affect the eye are cataracts, glaucoma, papilledema, and so on. Papilledema is the optic disc swelling due to an abnormally high intracranial pressure which is caused due to intracerebral mass lesions, cerebral hemorrhage, head trauma, and many other causes [2]. On the other hand, Pseudo- Papilledema which is the ostensible elevation of the optic disc is a fairly common finding in ophthalmic diagnosis and can be misleading causing serious consequences [3]. Papilledema's detection especially when there is the presence of Pseudo-Papilledema is vital as Papilledema can sometimes indicate a serious underlying condition. It can be an alarming sign for disease entities that cause increased intracranial pressure such as brain tumors, cerebrospinal inflammation, and idiopathic intracranial hypertension [4]. Papilledema can be diagnosed by using digital ocular fundus photography that obtains optic-disk images for detection of the disease [5]. Another method is to subject a particular patient to lumbar puncture and MRI.

Healthcare experts find it simple to accurately diagnose diseases at an early stage and assess symptoms based on radiographic images, but human error is unavoidable. In contrast, supervised machine learning (ML) algorithms have demonstrated a remarkable ability to outperform the conventional method for disease detection and support health professionals in the early identification of high-risk conditions [6]. Pattern recognition and machine learning hold the potential to enhance the accuracy of disease approach and detection in the biomedical field [7]. There are many more reasons why Image Processing and Machine Learning are used in the diagnosis of diseases on such a

 DOI : 10.35291/2454-9150.2023.0226

wide scale in today's world. This approach involves the classification of optical eye MRI images into Papilledema, Pseudo Papilledema, and Normal. The dataset was in the form of raw images which were pre-processed and augmented to avoid class imbalance problems. The processed images were then fed into a pre-trained ResNet-18 neural network which was responsible for extracting the features from the images. When features of an image dataset are extracted using a ResNet-18 and fed into classifiers, the accuracy of classification generally increases [8]. These features were then trained using four machine learning models namely Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest Classifier (RFC), and Gradient Boosting Classifier (GBC). Respective accuracies yielded by the models were calculated and a comparative analysis between the models was drawn.

II. RELATED WORKS

Khush Naseeb Fatima et. al. proposed a system for papilledema detection in fundus images using a hybrid feature dataset. The dataset utilized is taken from STARE. Thirteen features that are useful in identifying this disease are extracted from the fundus images after preprocessing and integrated to provide a feature set. Once features are extracted, an SVM classifier is used to classify this feature set. The approach yields a 96.67% accuracy rate [9].

Jin Mo Ahn et. al. proposed a system to analyze how well machine learning distinguishes between ocular neuropathies, pseudo papilledema (PPE), and healthy individuals. The dataset consisted of 295 images of PPE and optic neuropathies and 779 control images. Four machine learning classifiers were compared and their accuracies ranged from 95.89% to 98.63% [10].

Asif Nawaz et. al. designed a deep CNN with minimal memory consumption for multi-class retinal disease detection. standard benchmark dataset of Eye Net having 32 classes of retinal diseases. The proposed CNN model was used for drawing a comparison between precision, recall, and accuracy with different sets of epochs. The suggested method produced a 95% accuracy rate on the Eye-net dataset [11].

Rohit Thanki proposed a system for the classification of retinal fundus images using a deep neural network and machine learning approach. Public datasets such as DRISTHI-GS1 and ORIGA were used for the implementation of this method. 512 deep features of retinal images were extracted using the deep neural network and then classified using various machine learning classifiers k nearest neighbor (KNN), support vector machine (SVM), Decision Tree (DTC), Naïve Bayes (NB), Random Forest (RF), and logistic regression (LR). This method proved that a logistic regression-based classifier improves classification accuracy, sensitivity, and specificity above all other glaucomatous triage systems currently in use [12].

Jin Mo Ahn et. al. proposed a machine learning system to differentiate between optic neuropathies and pseudo papilledema. The dataset used consisted of 779 normal and 295 papilledema and pseudo papilledema images which were augmented later on. A convolutional neural network was constructed to execute the classification. The designed model yielded an accuracy of 100% for the training data, 96.35% for validation data, and 95.89% for the test data [13].

Bisahu Ram Sahu et. al. proposed a machine learning-based ensemble model to carry out skin disease classification. The design put forward was an ensemble model which was a combination of support vector machine (SVM), k-nearest neighbor (KNN), Random Forest (RF), and, Naive Bayes (NB) algorithm. The five classes that the suggested model divided skin diseases into include acne, skin allergies, nail fungus, hair loss, and normal skin. When compared to other models, the suggested ensemble model had the best accuracy rate, at 97.33% [14].

Asmae Ouhmida et. al. devised a method for the classification of Parkinson's disease using nine Machine Learning Algorithms (MLA), namely Support Vector Machine (SVM), Logistic Regression, Discriminant Analysis, K-Nearest Neighbors (KNN), Decision tree, Random Forest, Bagging tree, Naïve Bayes, and AdaBoost. Several evaluation factors were used to determine each classifier's efficiency score after classification algorithms were applied to a Parkinson's dataset consisting of 240 speech measures with 44 features. The KNN classifier produced the greatest F1-score of 97.30% and an accuracy rate of 97.22% [15].

III. PROPOSED METHODOLOGY

The methodology depicted in Figure 1 involves a multi-step process for image classification into three categories: Normal, Papilledema, and Pseudo-Papilledema. Initially, the images undergo augmentation to enhance dataset diversity, followed by passage through an image processing pipeline for quality refinement and feature extraction.

A pre-trained ResNet-18 model is utilized to extract high-level features from the processed images. Subsequently, these features are fed into machine learning classifiers, including Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest Classifier (RFC), and Gradient Boosting Classifier (GBC). These classifiers are used for the final multi-class classification task, contributing to the identification of Normal, Papilledema, and Pseudo-Papilledema images in the dataset.

DOI : 10.35291/2454-9150.2023.0226

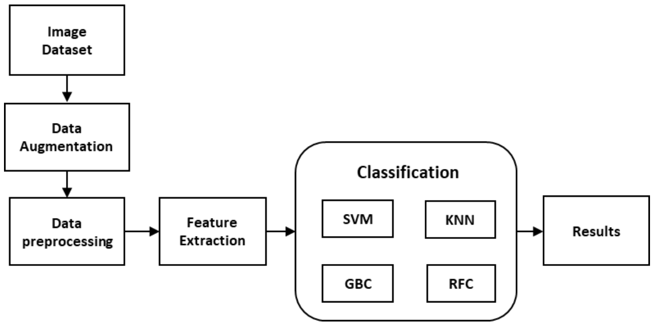

Figure 1: Methodology Block Diagram

A. Dataset Collection and Bifurcation

The dataset utilized in this approach is collected from Kaggle which consists of images segregated into three classes namely Papilledema, Pseudo-Papilledema, and Normal. The dataset was then sorted out into training, testing, and validation, and the segregation of the same is shown in Figure 2 below.

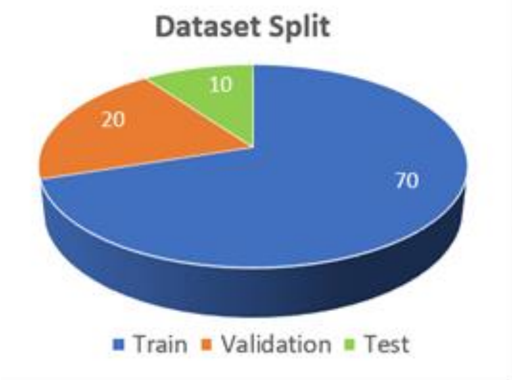

Figure 2: Dataset Bifurcation

B. Data Augmentation

The original dataset available had the likelihood to contribute towards a class imbalance problem as the three classes did not contain an equal number of images. Since typical classifiers seek an accurate performance over a complete range of examples, standard machine learning algorithms often become overwhelmed by the majority class and ignore the minority class [16]. To steer clear of this, traditional augmentation techniques were used. After using them, all three classes comprised an equal number of images which were then used for classification.

C. Data Pre-Processing

In the context of an image-processing problem, pre-processing plays a crucial role in enhancing an image dataset by mitigating undesired distortions and accentuating specific image features, as emphasized in [17]. In this particular methodology, Normalization and Contrast Limited Adaptive Histogram Equalization are employed as image pre-processing techniques. Normalization is applied to standardize pixel values, while Contrast Limited Adaptive Histogram Equalization is utilized to enhance image contrast. The significance of these techniques lies in their collective contribution to refining the dataset. Figure 3 illustrates sample images from each class as they traverse through the proposed pipeline, showcasing the impact of the pre-processing steps on the images.

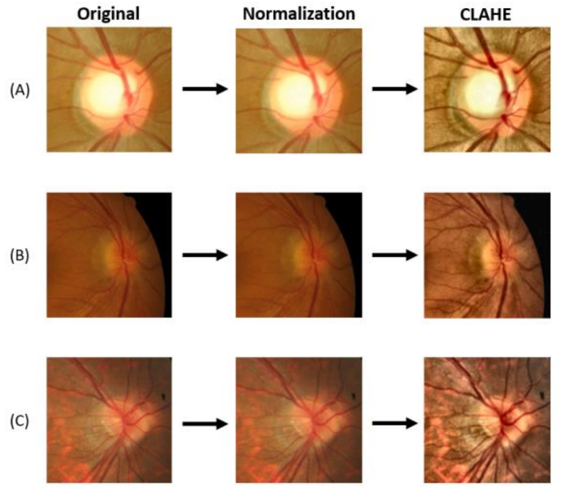

Figure 3: Image Processing Pipeline (A) Normal Image (B) Papilledema Image (C) Pseudo-Papilledema Image

D. Feature Extraction

The augmented and pre-processed images belonging to the three distinct classes were provided as input to a pre-trained ResNet-18 for feature extraction. To perform feature extraction, ResNet-18 is configured by excluding its final fully connected layer responsible for classification. The architecture of the ResNet-18 utilized in this approach is visualized in Figure 4. This modified ResNet-18 architecture is adept at capturing and representing meaningful features from the augmented and pre-processed images, facilitating subsequent stages of the classification process.

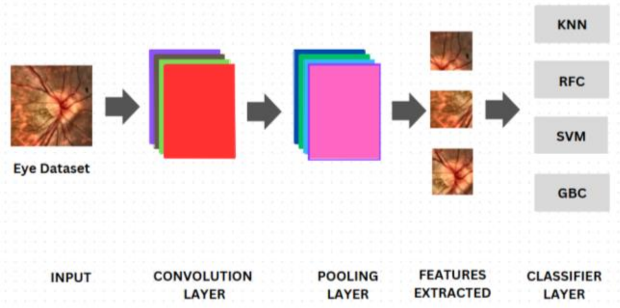

Figure 4: ResNet-18 Architecture for Feature Extraction

E. Classification

The features extracted are supplied to the four machine-learning classifiers namely Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest Classifier (RFC), and Gradient Boosting Classifier (GBC) for multi-class classification. The SVM model is said to be a better technique than neural networks as they have a strong theory of foundation, are less prone to overfitting,

 DOI : 10.35291/2454-9150.2023.0226

and require less memory [18]. A linear SVM algorithm with C=16.0 gave the best results for the eye dataset. A KNN model was used in this approach as it is easy to implement and efficient while working with image datasets. The KNN method assumes that comparable objects will be found nearby [19].

Experimentation on what k value achieves the best accuracy was carried out and it was concluded that k=30 yielded the greatest accuracy for the eye dataset. A Random Forest Classifier (RFC) is superior to other models due to its strong handling of highly non-linearly correlated data, noise resilience, and effective parallel processing [20].

In a similar way as performed for the KNN model, the authors experimented with the RFC model as well and found out that with n_estimators = 500 and random_state = 42, the model gave the best results. The last model used was the Gradient Boosting Classifier (GBC) which is a model readily present in Python. This model is not much used for image classification so, in this approach, it was implemented on the eye dataset to examine its performance. The GBC model yielded optimal results with n_estimators = 500 and random_state = 77.

IV. RESULTS

All the models implemented on the dataset yielded similar accuracies for both testing and validation datasets except for KNN which gave a validation accuracy of 93.15%, which was lower than the rest. However, KNN achieved the greatest testing accuracy of 97.87%. A comparative analysis of the validation and test accuracies for the four models is presented in Table 1, offering a graphical representation of their performance metrics. This table likely provides a visual summary of how well each model generalizes to test data compared to its performance on the validation set.

Model	*Hyperparameter Set*	*Validation Accuracy*	*Test Accuracy*
SVM	Linear, C=16.0	96.79%	97.45%
KNN	k=30	93.15%	97.87%
RFC	n_estimators = 500, Random State = 42	97.27%	97.45%
GBC	n_estimators = 500, Random State = 77	96.79%	95.32%

Table 1: Validation and Test Accuracy Comparison

The precision and recall score across all classes for SVM, RFC, and GBC was nearly the same for both the validation and test datasets which was above 95%. On the other hand, KNN showed variations in these parameters. The precision and recall score calculated concerning the validation dataset for the Normal class was 99% and for Papilledema and Pseudo-Papilledema was less than 95%. For the test dataset, KNN yielded the same statistics as the other classifiers.

V. CONCLUSION AND DISCUSSION

In this approach, an automated system was presented to classify eye MRI images procured from Kaggle into Normal, Papilledema, and Pseudo-Papilledema using image processing, feature extraction, and machine learning algorithms. Out of the four classifiers used, on the validation dataset, the RFC model produced the best accuracy of 97.22%, while on the test dataset, the KNN model fared better than the others with an accuracy of 97.87%. All the models employed performed well on the testing and validation datasets with accuracies of more than 90%. To conclude, this method proved that the combination of an image processing strategy, a neural network, and machine learning models is an optimistic technique for distinguishing between the three classes of eye images in the dataset used.

The approach employed in this study holds significant promise for the automated detection and swift diagnosis of papilledema, offering a valuable contribution to the effective treatment of this harmful medical condition. There is ample room for future advancements in this field, with potential avenues including the application of the developed models to diverse datasets using a variety of feature extraction techniques. Exploring different datasets and employing heterogeneous feature extraction methods has the potential to enhance the efficiency and generalizability of the proposed approach, paving the way for further improvements in the detection and diagnosis of papilledema

ACKNOWLEDGMENT

The authors declare that they received no additional funding for the project.

REFERENCES

[1]. D. Pascolini and S. P. Mariotti, “Global estimates of visual impairment: 2010,” The British Journal of Ophthalmology, vol. 96, no. 5, pp. 614–8, 2012.

[2]. M. Rigi, S. J. Almarzouqi, M. L. Morgan, and A. G. Lee, “Papilledema: epidemiology, etiology, and clinical management,” Eye and Brain, vol. 7, pp. 47–57, Aug. 2015.

[3]. B. McCafferty, C. M. McClelland, and M. S. Lee, “The diagnostic challenge of evaluating papilledema in the pediatric patient,” Taiwan Journal of Ophthalmology, vol. 7, no. 1, pp. 15–21, 2017.

[4]. R. Dhoot and E. Margolin, “Papilledema,” PubMed, 2021.

 DOI : 10.35291/2454-9150.2023.0226

[5]. D. Milea et al., "Artificial Intelligence to Detect Papilledema from Ocular Fundus Photographs," New England Journal of Medicine, vol. 382, no. 18, pp. 1687–1695, Apr. 2020.

[6]. A. Sharma, J. Pathak, and P. Rajakumar, "Disease Prediction using machine learning algorithms," IEEE Xplore, Apr. 01, 2022.

[7]. P. Hamsagayathri and S. Vigneshwaran, "Symptoms Based Disease Prediction Using Machine Learning Techniques," Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Mar. 31, 2021.

[8]. H. Kibriya and R. Amin, "A residual network-based framework for COVID-19 detection from CXR images," Neural Computing and Applications, Dec. 2022.

[9]. Khush Naseeb Fatima, M. Usman Akram and Shafaat A. Bazaz, "Papilledema Detection in Fundus Images Using Hybrid Feature Set," 5th International Conference on IT Convergence and Security (ICITCS), Oct. 08, 2015.

[10]..J. M. Ahn, S. Kim, K.-S. Ahn, S.-H. Cho, and U. S. Kim, "Accuracy of machine learning for differentiation between optic neuropathies and pseudopapilledema," BMC Ophthalmology, vol. 19, no. 1, Aug, 2019.

[11]..Asif Nawaz, Tariq Ali, Ghulam Mustafa, Muhammad Babar, Basit Qureshi, "Multi-Class Retinal Diseases Detection Using Deep CNN with Minimal Memory Consumption," IEEE Access, pp. 1-1, Jan, 2023.

[12].R. Thanki, "A deep neural network and machine learning approach for retinal fundus image classification," Healthcare Analytics, vol. 3, p. 100140, Nov. 2023.

[13].Jin Mo Ahn, Sangsoo Kim, Kwang-Sung Ahn, Sung-Hoon Cho, Ungsoo S. Kim, "Accuracy of machine learning for differentiation between optic neuropathies and pseudopapilledema," BMC Ophthalmology, Aug. 01, 2019.

[14].Bisahu Ram Sahu, Akhilesh Kumar Shrivas and Abhinav Shukla, "Skin Disease Classification using Machine Learning based Proposed Ensemble Model," 4th International Conference for Emerging Technology (INCET), Jul. 10, 2023.

[15].Asmae Ouhmida, Abdelhadi Raihani, Bouchaib Cherradi, Yasser Lamalem, "Parkinson's disease classification using machine learning algorithms: performance analysis and comparison," 2nd International, Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Mar. 25, 2022.

[16].Xinjian Guo, Yilong Yin, Cailing Dong, Gongping Yang and Guangtong Zhou, "On the Class Imbalance Problem," Fourth International Conference on Natural Computation, Oct, 2008.

[17].M. Sonka, V. Hlavac, and R. Boyle, "Image pre-processing," Image Processing, Analysis and Machine Vision, pp. 56–111, 1993.

[18].N. V. Kumar, P. V. Kumar, K. Pramodh, and Y. Karuna, "Classification of Skin diseases using Image processing and SVM," IEEE Xplore, Mar. 01, 2019.

[19].Roger Singh Chugh, Vardaan Bhatia, Karan Khanna and Vandana Bhatia, "A Comparative Analysis of Classifiers for Image Classification," 10th International Conference on Cloud Computing, Data Science & Engineering, 2020.

[20].A. Sarica, A. Cerasa, and A. Quattrone, "Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review," Frontiers in Aging Neuroscience, vol. 9, Oct. 2017.

 DOI : 10.35291/2454-9150.2023.0226