



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 9, September 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Enhancing Ocular Health: Investigating A CNN Model for Cataract Diagnosis Using Eye Fundus Images

Sachet Utekar

Department of Electronics & Computer Science, SIES Graduate School of Technology, Navi Mumbai, India.

ABSTRACT: Recent reports from authoritative sources, including the World Health Organization, highlight the increasing concern over vision-related impairments, with delayed cataract treatment being a leading cause of irreversible blindness in people over 50 years. Cataracts, causing visual loss and hindering daily tasks like driving and reading, require swift and reliable diagnosis and treatment methods. Thankfully, advancements in technology now offer automated, computer-based solutions, reducing time and human error in diagnosis. This research paper presents a Convolutional Neural Network (CNN)-based approach for cataract detection in eye scans, aiming to aid ophthalmologists in accurate and automated diagnosis. Trained on the DR HAGIS dataset (Diabetic Retinopathy, Hypertension, Age-related macular degeneration, and Glaucoma ImageS), the proposed model demonstrated superior performance with 99.918% training accuracy and 97.34% testing accuracy, offering a low-cost diagnostic system for identifying normal and cataractous ocular disease cases from fundus images.

KEYWORDS: Cataract Detection · Eye Fundus Images · Convolution Neural Network · DR HAGIS dataset

I. INTRODUCTION

Cataract is a leading cause of vision impairment worldwide, characterized by clouding of the lens in the eye. Traditional methods of cataract detection rely on manual examination by trained professionals, which can be subjective and time-consuming. The advent of deep learning has opened new avenues for automated cataract detection using image analysis techniques. Vision is a vital human sense, responsible for up to 80% of the brain's processing related to understanding the environment. Any impairment to vision can cause significant disruption and inconvenience. Cataract stands out as a leading cause of vision impairment and, in severe cases, permanent blindness. Cataracts form when a cloudy area develops on the eye's lens due to aging and tissue degradation, leading to blurred vision. They can be categorized into nuclear, cortical, posterior subcapsular, and congenital cataracts.

Different types of cataracts impact vision in specific ways. Nuclear cataracts affect the center of the lens, causing nearsightedness initially and progressing to dense yellow or brown discoloration, severely impeding vision and color perception. Cortical cataracts affect the lens edges, creating white wedge-shaped opacities that gradually move toward the center, affecting vision over time. Posterior subcapsular cataracts occur at the back of the lens, resulting in a tiny opaque area that interferes with reading ability. Congenital cataracts, on the other hand, are genetic and present at birth. Over the years, researchers have proposed various Computer-Aided Detection systems using techniques like Color Fundus Photography and Optical Coherence Tomography, combining image results with demographic data for more accurate diagnostics.

In recent times, Deep Learning-based Neural Network models have garnered attention among researchers for automating and improving precision in ocular diagnostics. These models show promise in detecting and classifying ocular diseases, with Convolutional Neural Networks (CNNs) being particularly extensively studied in this domain. CNNs have been successful in segmenting retinal vessels and classifying specific ocular diseases. However, only a few studies have focused on classifying multiple ocular diseases from fundus images.



II. RELATED WORK

Over the past few years, machine learning-based image understanding systems have seen rapid advancements. Convolutional neural networks (CNNs) have emerged as a mainstream solution for analyzing medical images. They are powerful tools that can extract visual information and perform generative and descriptive tasks, making them effective for image representation and feature classification. Given the scarcity of experts and the complexities of medical images, computer-aided deep learning tools offer great potential to improve diagnosis accuracy and efficiency.

Ocular diseases like diabetic retinopathy, cataract, and age-related macular degeneration are prevalent and can have severe consequences if not detected early. However, early detection can be challenging due to the lack of visible symptoms in the initial stages. Color fundus photography is an economical and effective method for early-stage ocular disease screening. Researchers have explored various deep learning models, including Dense Correlation Network (DCNet) and VGG-16, on datasets like DR HAGIS to extract features and identify ocular diseases with impressive accuracy levels.

In India, where a significant number of people suffer from blindness, Diabetic Retinopathy (DR) and Glaucoma are major causes. To address this issue, a deep neural network model has been proposed to detect the presence of these diseases at early stages, helping patients seek timely medical attention. While the current model achieved an 80% accuracy, further improvements can be made by adding dropout layers and dense layers, which reduce overfitting and enhance feature extraction.

Glaucoma, one of the leading causes of vision impairment, often goes undetected due to the lack of noticeable symptoms. Researchers have proposed a unified MKL framework called MKLclm, which combines demographic data, genome information, and retinal fundus images to detect ocular diseases. This approach has shown promising results for predicting Glaucoma, Age-related macular disease, and Pathological Myopia, outperforming single kernel SVM and MKL models.

In summary, machine learning techniques, especially CNN-based models, are transforming the field of ocular disease diagnosis by providing accurate and automated solutions. These advancements have the potential to revolutionize the way doctors' approach early detection and treatment of various ocular disorders, thereby improving the quality of life of many individuals.

III. METHODOLOGY

This section covers the dataset description, data preprocessing and the proposed architecture. The CNN layers function as feature extractors and understanding image characteristics.

1.1 Dataset Description

The DR HAGIS database has been created to aid the development of vessel extraction algorithms suitable for retinal screening programs. Researchers are encouraged to test their segmentation algorithms using this database. The DR HAGIS database is free to download for research and educational purposes only. All thirty-nine fundus images were obtained from a diabetic retinopathy screening program in the UK. Hence, all images were taken from diabetic patients. Since patients attending these screening programs exhibit other co-morbidities, the DR HAGIS database consists of following four co-morbidity subgroups:

- Images 1-10: Glaucoma subgroup
- Images 11-20: Hypertension subgroup
- Images 21-30: Diabetic retinopathy subgroup
- Images 31-40: Age-related macular degeneration subgroup

It should be noted that image 24 and 32 are identical, as this fundus image was obtained from a patient exhibiting both

diabetic retinopathy and age-related macular degeneration. Besides the fundus images, the manual segmentation of the retinal surface vessels is provided by an expert grader. These manually segmented images can be used as the ground truth to compare and assess the automatic vessel extraction algorithms. Masks of the FOV are provided as well to quantify the accuracy of vessel extraction within the FOV only.

The images were acquired in different screening centers, therefore reflecting the range of image resolutions, digital cameras and fundus cameras used in the clinic. The fundus images were captured using a Topcon TRC-NW6s, Topcon TRC-NW8 or a Canon CR DGi fundus camera with a horizontal 45-degree field-of-view (FOV). The images are 4752x3168 pixels, 3456x2304 pixels, 3126x2136 pixels, 2896x1944 pixels or 2816x1880 pixels in size. The fundus images are saved as compressed JPEG files with 8 bits per color plane. The ground truth and mask images are saved as binary PNG files.

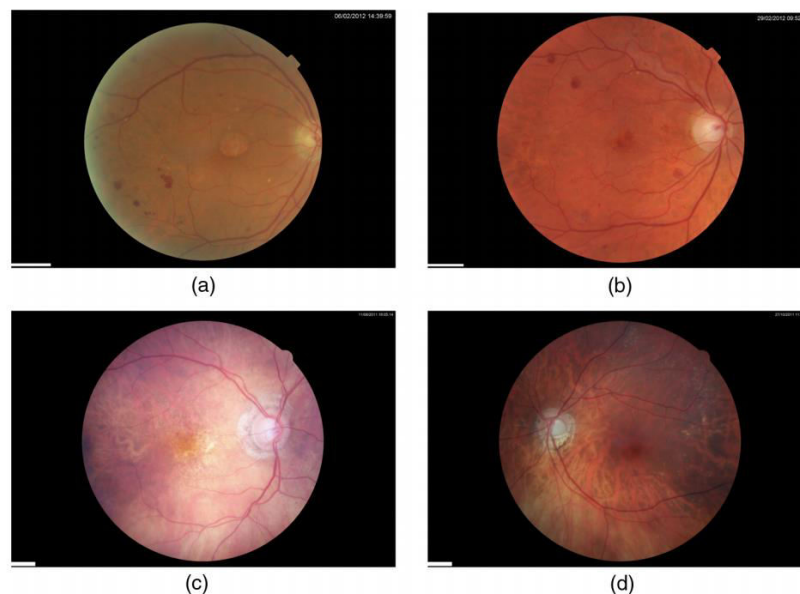


Figure 1: Visualization of sample images in the DR HAGIS dataset. (a) and (b) are normal fundus images and (c) and (d) are cataractous fundus images.

1.2 Dataset Preprocessing

The DR HAGIS dataset is a valuable resource for diabetic retinopathy research, containing retinal images of patients with varying degrees of the disease. Before using this dataset for training a deep learning model, some essential preprocessing steps were performed to ensure data quality and compatibility with the model.

- **Data Cleaning:** The first step was to clean the dataset by removing any corrupted or incomplete images. This ensured that only high-quality, usable images were included in the training process.
- **Data Augmentation:** To increase the dataset size and improve the model's generalization ability, data augmentation techniques were applied. This involved randomly rotating, flipping, and shifting the images, as well as adjusting their brightness and contrast.
- **Image Resizing:** The images in the DR HAGIS dataset were of varying sizes. To make them uniform and suitable for the model, all images were resized to a standard dimension, preserving their aspect ratio.
- **Normalization:** Normalization was applied to bring pixel values within a specific range (usually between 0 and 1)

to facilitate faster convergence during training.

- **Data Split:** The dataset was divided into training, validation, and testing sets. The training set was used to teach the model, the validation set to tune hyperparameters and prevent overfitting, and the testing set to evaluate the model's performance.
- **Class Balancing:** Diabetic retinopathy can have imbalanced class distributions, with certain severity levels having more samples than others. To address this, techniques like oversampling or undersampling were applied to balance the classes in the training set.

By conducting these preprocessing steps, the DR HAGIS dataset was refined and prepared for feeding into the deep learning model. These measures helped ensure that the model was trained on diverse, high-quality data, which in turn enhanced its ability to accurately detect and classify diabetic retinopathy severity levels in unseen images.

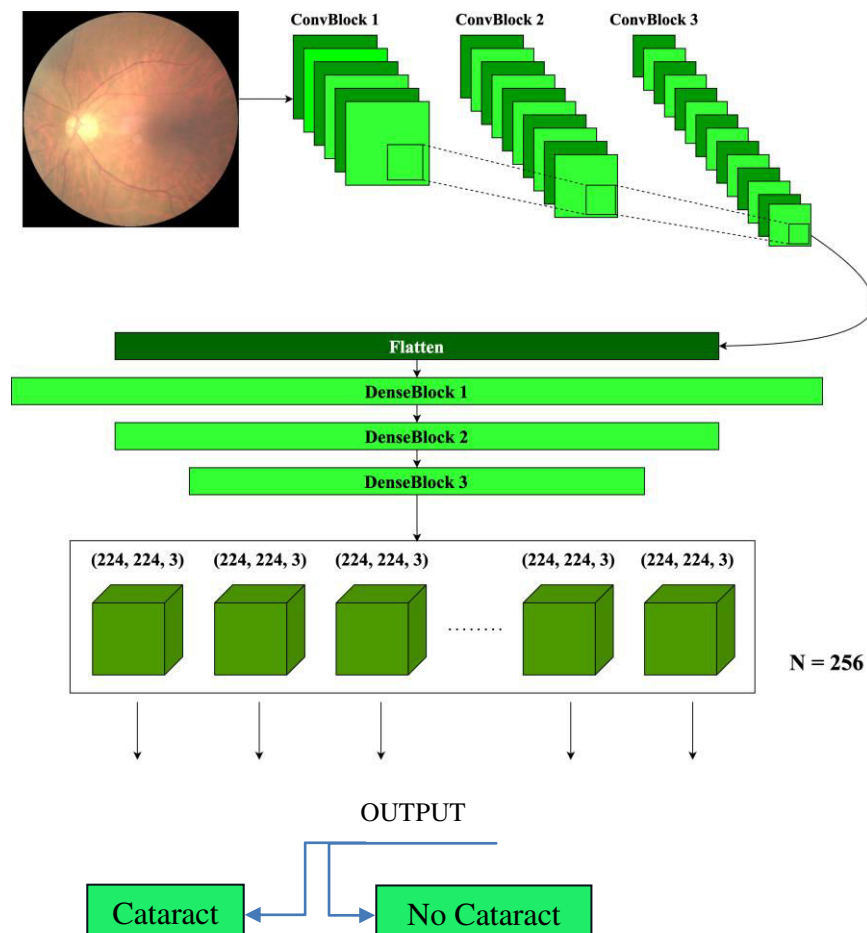


Figure 2: Architecture diagram of the proposed CNN Model.

1.3 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are similar to the regular neural networks, but they are specifically designed to work with two-dimensional images. This characteristic makes them very effective at learning intricate features and attributes in images. CNNs have a wide range of applications, including image classification, segmentation, object

recognition, and even generating new images. In the field of medical imaging, CNNs have proven to be particularly valuable due to their ability to handle these diverse tasks effectively.

A CNN is made up of three main types of layers: convolutional, pooling, and Dense (fully connected) layers. Among these, the convolutional layer plays a pivotal role. It applies a mathematical operation called convolution to the input image, using a set of learned filters or kernels. These filters are adjusted during the training process to identify specific patterns in the image. The result is a smaller convoluted image, which is then passed on to the next layer for further processing.

In simpler terms, imagine the convolutional layer as a powerful tool that scans through the input image, looking for important features like edges, textures, or shapes. These features are captured by the filters and combined to create a more focused representation of the image. This process helps the network understand the essential elements present in the image, making it better equipped to perform tasks like identifying objects or recognizing patterns. CNNs are specialized neural networks that excel at processing images. The convolutional layer is the heart of a CNN, responsible for extracting meaningful features from the input images. This feature extraction process enables CNNs to perform various tasks related to image analysis, making them a valuable tool in fields like medical imaging, where accurate and efficient image processing is crucial

$$S(i, j) = (I \cdot K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n)$$

In the ConvNet, I have used the Pooling layer between consecutive convolutional layers. The main purpose of the pooling layer is to gradually reduce the dimension of the image. It does this by selecting only the most relevant information from the previous layer and discarding the rest. This helps in reducing the number of parameters in the network and makes the model more manageable and efficient. Additionally, pooling layers play a role in preventing overfitting, which is when the model becomes too specialized on the training data and performs poorly on new, unseen data.

There are three types of pooling layers: maximum pooling, average pooling, and global pooling. Maximum pooling picks the highest value from each patch in the feature map, while average pooling calculates the average value for each patch. On the other hand, global pooling condenses the entire feature map into a single value. Among these, maximum pooling is the most commonly used technique because it focuses on extracting the most important features, such as bright regions and edges, from the image.

As we move towards the end of the network, I have typically used the fully connected or dense layer. In this layer, each neuron is connected to every neuron in the previous layer. The dense layer performs matrix-vector multiplication in the background, allowing it to analyze and combine the extracted features from earlier layers. This process helps the network make more informed decisions and produce the final output, such as classifying an image into different categories.

1.4 Proposed Architecture

This paper introduces a novel CNN-based model for classifying ocular fundus images as normal or cataractous. The network comprises 16 layers, starting with an input layer that takes $224 \times 224 \times 3$ -dimensional images. The architecture is organized into three sets, each consisting of a convolutional layer with a 3×3 kernel and ReLU activation, followed by a max pooling layer with a pool size of 2×2 , and a batch normalization layer. After passing through the three sets, a Dropout layer with a 0.2 dropout rate randomly drops 20% of neurons, and then a Flatten layer is applied to transform the 2D representation into a 1D vector.

Next, the 1D vector is fed into a fully connected Dense layer with 256 units and ReLU activation. In the original architecture, two LSTM layers with 256 units each and tanh activation are utilized. However, for this proposed model, the LSTM layers are removed. Instead, the output from the previous Dense layer is directly connected to a final Dense layer with sigmoid activation, responsible for the classification task. Figure 3 illustrates the design of this modified CNN-based model without the LSTM layers.

Moreover, the removal of LSTM layers may have allowed the model to generalize better to a wider range of ocular conditions and diseases. LSTM layers are particularly useful for sequential data and time-series tasks, whereas in this image classification scenario, the temporal aspect may not have a significant impact on the final prediction.

Overall, the modified CNN-based model presented in this paper demonstrates a promising approach for accurate and efficient ocular fundus image classification. Its ability to achieve competitive results without relying on LSTM layers opens up opportunities for scalability and adaptability, paving the way for potential applications in automated diagnostic systems and telemedicine, where fast and reliable classification is crucial for timely medical interventions.

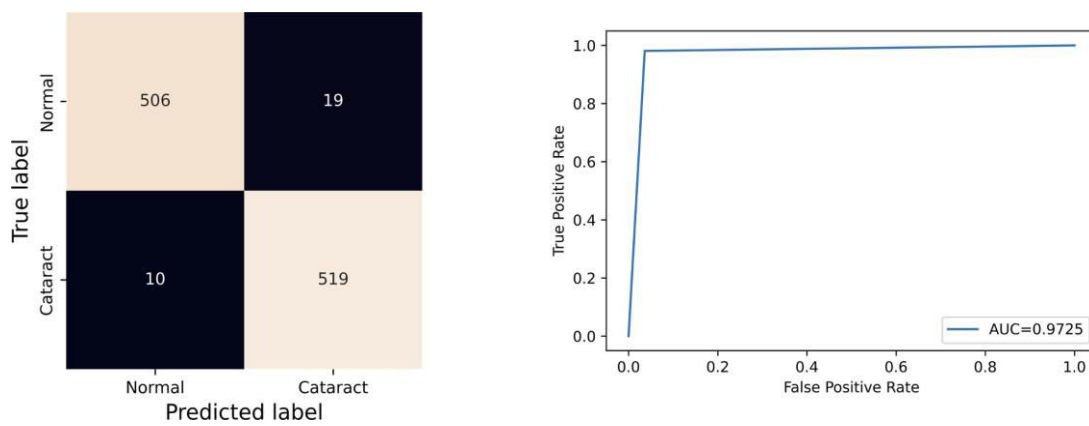


Figure 3: (a) Confusion matrix of the proposed model on the test set. (b) ROC (Receiver Operating Characteristic) Curve of the results on test set.

IV. RESULTS AND DISCUSSION

4.1 Evaluation Metrics

In this study, the model's performance is assessed using various metrics to understand how well it performs with the given data. These performance indicators provide valuable insights into the model's efficiency. The proposed CNN architecture is evaluated using six key metrics: accuracy, precision, recall, sensitivity, specificity, and F1 score. These metrics help us measure different aspects of the model's classification ability and give us a comprehensive understanding of its overall performance.

4.1.1 Accuracy

The accuracy of a classification algorithm is a method for assessing how accurately the algorithm can classify individual data points.

4.1.2 Precision

Precision, referred to as positive predictive value, measures the model's ability to accurately predict positive instances.

4.1.3 Recall

The recall of the model evaluates how well it can identify positive samples. When the model can correctly recognize a larger number of positive samples, the recall value increases.



4.1.4 Sensitivity

Sensitivity is a performance metric used to evaluate the model's effectiveness in correctly identifying positive instances. It indicates the number of positive samples the model accurately recognized.

4.1.5 Specificity

Specificity refers to the percentage of true negatives correctly identified by the model.

4.1.6 F1 Score

The F1-score is a single metric that combines a classifier's precision and recall by calculating their harmonic mean.

4.2 Results

The proposed CNN-LSTM combination model achieved remarkable results on the testing set. The model achieved an accuracy, precision, recall, sensitivity, specificity, and F1 score of 97.34%, 96.56%, 98.06%, 96.65%, 98.48%, and 97.31%.

V. CONCLUSION AND FUTURE SCOPE

In this research, I have developed a CNN-based model to classify ocular fundus images as either normal or cataractous. The model was trained on the well-known DR HAGIS dataset and produced remarkable results on both familiar and unfamiliar data. During training, the model achieved an impressive accuracy of 99.918%, while on testing, it maintained a high accuracy of 97.34%. Compared to existing CNN-based and pre-trained models for ocular disease classification, our proposed model demonstrated superior performance while also being more efficient in terms of time and space.

Looking ahead, our future scope involves extending the model to cater to multiple ocular diseases like myopia, glaucoma, etc. My aim is to obtain satisfactory results that would allow us to implement the system on a larger scale. As the world faces increasing complexity in ocular diseases, our system has the potential to offer affordable and portable ocular diagnostic solutions, thereby reducing diagnostic time and enhancing treatment by providing valuable insights.

REFERENCES

1. World Health Organization WHO. Vision impairment and blindness. <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>. [Online: accessed 2023-07-18].
2. G V S Murthy, Sanjeev K Gupta, Neena John, and Praveen Vashist. Current status of cataract blindness and Vision 2020: The right to sight initiative in India - PMC.
3. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2612994>. [Online: accessed 2023-07-16].
4. Mayo Clinic. Cataracts - Symptoms and causes.
5. <https://www.mayoclinic.org/diseases-conditions/cataracts/symptoms-causes/syc-20353790>, Sep 13 2022. [Online: accessed 2023-07-22].
6. Keiron O'Shea and Ryan Nash. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*, 2015.
7. Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9(4):611–629, 2018.
8. Cheng Li, Jin Ye, Junjun He, Shanshan Wang, Yu Qiao, and Lixu Gu. Dense correlation network for automated multi-label ocular disease detection with paired color fundus photographs. In *2020 IEEE 17th international symposium on biomedical imaging (ISBI)*, pages 1–4. IEEE, 2020.
9. Nadim Mahmud Dipu, Sifatul Alam Shohan, and KMA Salam. Ocular disease detection using advanced neural network-based classification algorithms. *Asian Journal For Convergence In Technology (AJCT) ISSN-2350-1146*, 7(2):91–99, 2021.
10. Deepthi K Prasad, L Vibha, and KR Venugopal. Early detection of diabetic retinopathy from digital retinal fundus images. In *2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, pages 240–245.



IEEE, 2015.

11. Yanwu Xu, Lixin Duan, Huazhu Fu, Zhuo Zhang, Wei Zhao, Tianyuan You, Tien Yin Wong, and Jiang Liu. Ocular disease detection from multiple informatics domains. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 43–47. IEEE, 2018.
12. Abien Fred Agarap. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*, 2018.
13. <https://paperswithcode.com/dataset/dr-hagis>



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.423



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com



www.ijirset.com

Scan to save the contact details