

# Analyzing Optimal Image Preprocessing Techniques for Automated Retinal Disease Diagnosis

Lakshmi Chakka

Received September 08, 2023

Accepted December 16, 2023

Electronic access December 31, 2023

Machine learning has made remarkable strides in the field of disease diagnosis, revolutionizing patient treatment and care. By interpreting medical images, machine learning techniques have the potential to improve diagnostic accuracy in a briefer span. And while the central area of interest in diagnosis has been cancer detection, retinal diseases have also gained significant attention over the years. Diagnosing retinal diseases efficiently is critical to preserving vision, enhancing patient results, and aligning with the overall goal of preserving public health. Using OCT scans of retinal images sourced from Kaggle, this study hopes to identify the ideal preprocessing methodology (among normalization, standardization, rescaling, and RGB to BGR conversion) in the diagnosis of retinal diseases DME, CNV, and Drusen. The original hypothesis of this study was that normalization would be the ideal preprocessing technique; however, by employing the ResNet-50 model, the findings of this study reveal that rescaling images, a preprocessing technique, has proved to be the most efficient in accurately diagnosing retinal diseases.

## Introduction

From creating visually appealing images to enhancing the quality of biometric data, preprocessing techniques have played a significant role in our world. These techniques have become essential in fields such as artificial intelligence, data analytics, and computer vision. Currently, researchers continue to explore novel image preprocessing methods to address diagnosis of diseases through medical imaging. For example, preprocessing techniques are being refined to enhance the analysis of mammography, a screening tool used for early detection of breast cancer<sup>1</sup>.

While state-of-the-art research has focused on image preprocessing techniques in many other fields, there has not been clear research regarding the effect of different preprocessing techniques in retinal disease diagnosis<sup>2-4</sup>. Focusing on medical imaging, this study seeks to bridge the significant gap in research by investigating the impact of various image preprocessing techniques specifically within the domain of retinal disease (diseases that cause harm to the retina, a light-sensitive layer of tissue found on the back of an eyeball). By analyzing methodologies applicable to diseases related to the retina, this research endeavors to reveal optimal approaches that could greatly refine the overall accuracy of diagnosis.

Addressing retinal diseases through appropriate treatment and image enhancement may help mitigate the complexity of diagnosing these diseases<sup>5</sup>. Extracting relevant features from retinal images (e.g., blood vessels, optic disc, lesions, and other anatomical structures) after utilizing preprocessing techniques such as noise reduction (which reduces distortions) and image

clean-up may aid in having a better visualization of retinal abnormalities and the diagnosis of retinal diseases.

## Image Preprocessing

Image preprocessing techniques aim to enhance the quality of images and therefore improve the analysis process<sup>6-8</sup>. The image preprocessing techniques analyzed in this study include normalization, rescaling, converting RGB images to BGR, and image standardization, all of which stand as pivotal steps in refining image data.

Normalization is a preprocessing technique used to modify pixel intensities to a common range. Studies have demonstrated that normalization simplifies image analysis by ensuring that differences in brightness or contrast between images do not dominate the analysis process<sup>9</sup>.

Rescaling is the process of changing the dimensions of images while maintaining the image's content integrity. It is a technique that establishes uniformity and adjusts the image size so the content may be further analyzed.

Converting RGB images to BGR is a common color space transformation used in image preprocessing applications which essentially rearranges pixel values for each channel. An RGB pixel, for example, with values (R, G, B) would be converted to BGR, resulting in a pixel with values (B, G, R). The RGB to BGR conversion is significant in that studies have shown that BGR is preferable when performing preprocessing tasks with certain libraries<sup>8</sup>.

Image standardization focuses on changing pixel values to have a specific mean and standard deviation to boost the con-

---

sistency of the image. Studies have proven standardization's significance in which it has been established that the preprocessing technique can help reduce noise and may focus on relevant patterns of the image<sup>7</sup>.

In this research, image preprocessing techniques were applied to OCT (Optical Coherence Tomography) images which were utilized to provide valuable insights into anatomical composition, enabling easier diagnosis of certain medical conditions. OCT scans use light waves to capture high-resolution, cross-sectional images of biological tissues in the eye such as the retina<sup>10</sup>.

## Application of Deep Learning in Medical Imaging

Many researchers have investigated the application of deep learning in disease diagnosis with dental image analysis, breast cancer, neurodegenerative diseases, radiology, and many others<sup>11,12</sup>.

Contributing to the field of imaging significantly, CNNs (Convolutional Neural Networks) have emerged as a beneficial tool for processing visual data. The structure of this system, resembling the neural network found in the visual cortex of animals, consists of weights akin to connectivity in neurons<sup>13</sup>.

The fundamental elements of CNNs include convolutional layers, activation functions, pooling, and fully-connected layers as portrayed in Figure 1. The convolutional layers of CNN models apply convolution operations to input data where filters extract various features such as edges or patterns. Activation functions introduce non-linearities, aid in learning complicated patterns, and form the network's output characteristics. Pooling layers manage the amount of parameters and computation in the network while retaining important information. Fully-connected layers aid in making predictions or classifications based on features that have been learned and produce output probabilities.

To carry out this research, it was chosen to utilize the CNN deep learning architecture and ResNet-50 model (pretrained with ImageNet). The ResNet-50 model is primarily used for image classification and recognition assignments. Its architecture, utilizing residual functions, enables the network to learn with respect to layer inputs.

A previously published study has focused on the different preprocessing techniques used for dental images<sup>14</sup>. In this research, different preprocessing techniques were designed to improve radiographic images for the segmentation and classification of dental images. The findings of this study indicate that specific preprocessing methods, including a hybrid metaheuristic algorithm, bi-level histogram equalization, and the application of a high-pass filter, consistently yielded superior performance metrics when compared to alternative enhancement approaches.

Researchers in the past have also investigated image preprocessing techniques for breast cancer detection in which the

model incorporates various methods such as image preprocessing, feature selection, feature extraction, and numerous machine learning algorithms<sup>15</sup>. Various image preprocessing techniques such as denoising, segmentation, and image enhancement were explored. Ultimately, it was concluded that the geometric mean filter remained most effective. The proposed method used for this study offers benefit through accurate detection of breast cancer through image analysis.

## Application of Deep Learning in Retinal Disease Diagnosis

In addition to other areas of medical imaging, researchers have also delved into the implementation of machine learning in retinal disease diagnosis<sup>16,17</sup>.

A study advocated the detection of two of the most common retinal diseases, DME (Diabetic Macular Edema) and AMD (Age-Related Macular Degeneration)<sup>18</sup>. This study reviews various deep learning methods (including VGG-16, MobileNet, ResNet-50, Inception V3, and Xception) for classifying OCT images of retinal diseases. ResNet-50 is shown to be the most accurate with a 96.21% testing accuracy.

Furthermore, a review has comprehensively examined the utilization of deep learning techniques in the context of diabetic retinopathy, focusing on their application for the classification of retinal landmarks and diseases<sup>19</sup>. The deep learning techniques used by many different researchers include using a CNN model while implementing classification techniques.

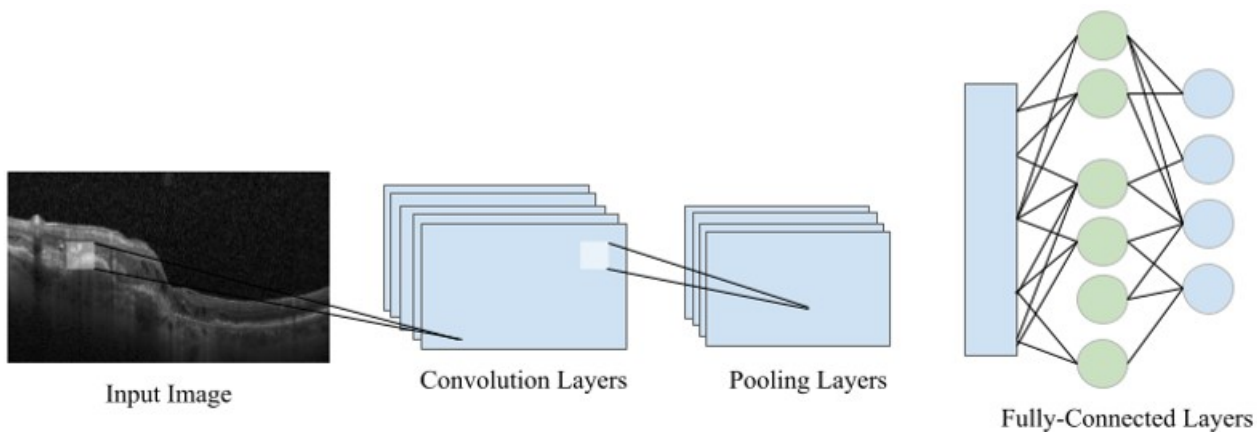
An alternative analysis inspects the implementation of an artificial intelligence-based decision-making and classification system for abnormalities in OCT retinal images<sup>20</sup>. In this investigation, it has been stressed that incorporating effective preprocessing techniques such as reducing speckle noise as well as proper segmentation methodology can lead to efficient health-care management.

## Retinal Disease

The retinal diseases investigated in this study include Diabetic Macular Edema (DME), Choroidal Neovascularization (CNV), and Drusen.

DME is an occurrence of retinal thickening and is often the underlying reason for vision loss among patients with diabetes<sup>21</sup>. Retinal drusen are yellow spots that accumulate under the retina, composed of proteins and lipids. CNV is a condition in which abnormal blood vessels appear under the retina. The OCT images of these retinal diseases can be referred to in Figure 12. The model was trained with images of each retinal disease as well as normal eye images.

This study essentially investigates image preprocessing techniques, how different methods may affect the performance of



**Fig. 1** Structure of a CNN model

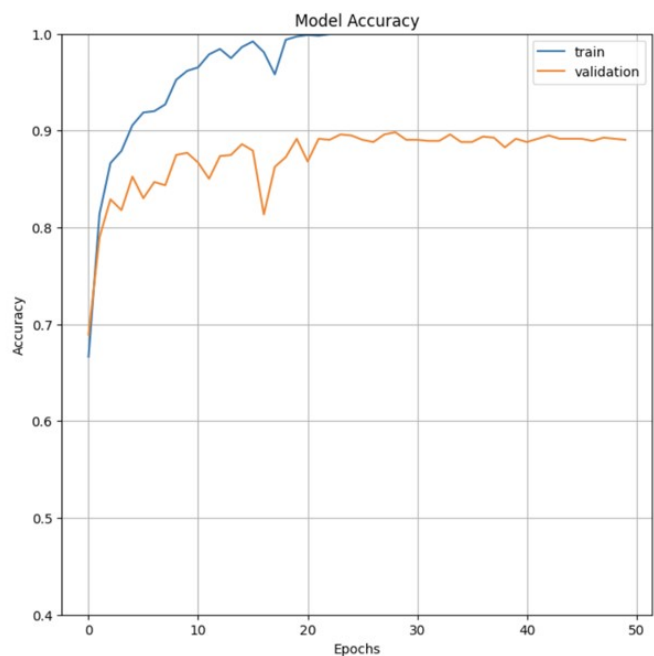
a deep learning model, and what may come to be the preprocessing technique with the highest efficiency for retinal disease diagnosis.

## Results

Through the ResNet-50 model, the optimal preprocessing technique was identified. The x-axis of each graph represents the number of epochs, illustrating the model's progress over a series of training iterations.

In contrast to Figure 2 (accuracy with no preprocessing), Figure 3 (accuracy with standardization) seemed to perform with reduced instances of underfitting, however, both are similar in terms of the training peak. Figure 4 (accuracy with RGB to BGR conversion) surpasses the training set accuracy achieved through image standardization as it reaches an accuracy of 100% before 20 epochs. Figure 5 (accuracy with rescaling) shows better accuracy than that of image standardization and RGB to BGR conversion with no signs of underfitting or overfitting. Finally, Figure 6 (accuracy with normalization) displays the highest level of accuracy among all figures after 50 epochs, achieving a validation accuracy exceeding 0.9, however, it has exhibited indications of underfitting, making rescaling the most optimal method of preprocessing OCT scans.

When the precision of the models are examined, it can be seen that most models matched or exhibited inferior performance in comparison to the Figure 7 (the original model). Figure 10 (precision of the model with image standardization being applied) shows promising results in contrast to Figure 8 (precision of the model with the normalization method applied to images). Figure 11 (precision when RGB to BGR conversion is implemented) portrays lower signs of overfitting/underfitting. This is while Figure 9 (precision when images are rescaled) exhibits the best

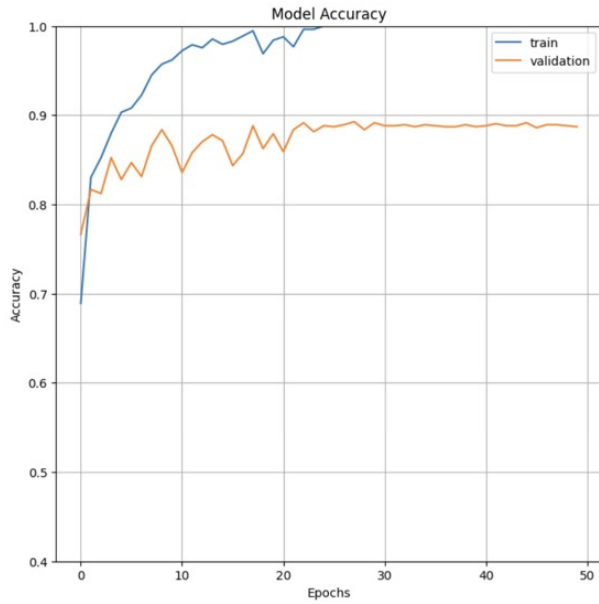


**Fig. 2** The accuracy of the model with no image preprocessing techniques being applied.

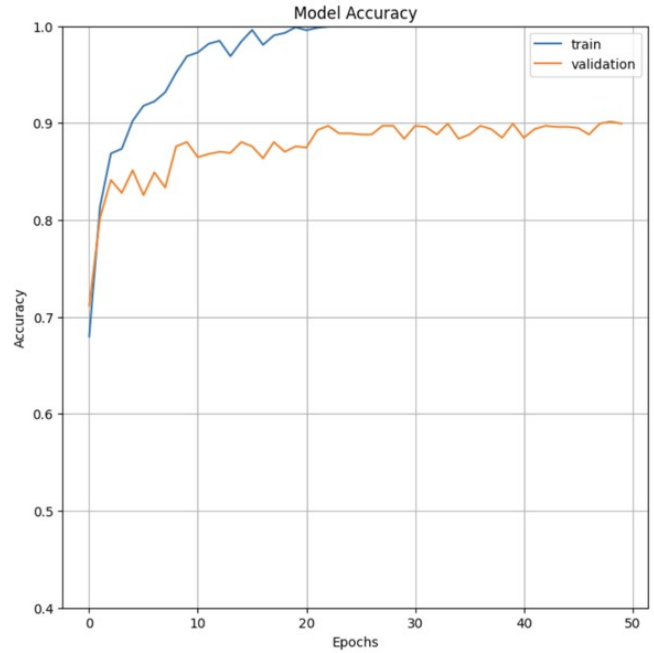
results in terms of precision when compared to other model precisions.

## Discussion

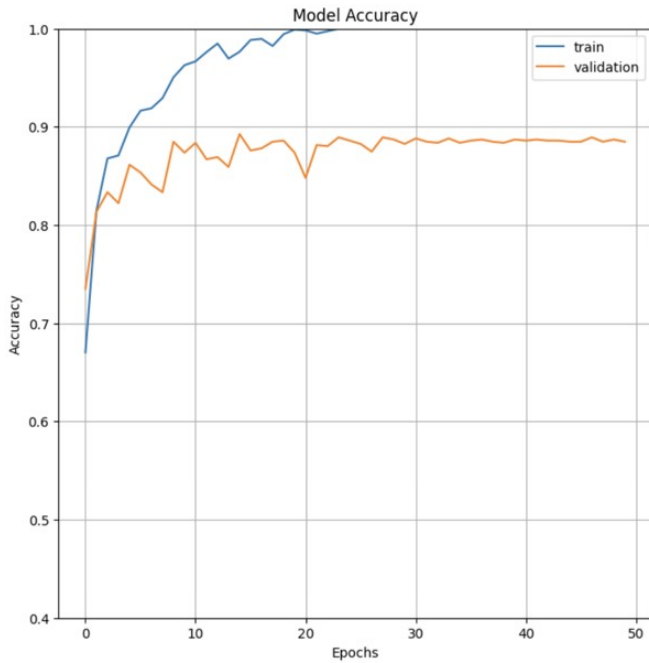
In this study, four different preprocessing techniques were analyzed, all of which have exceeded the accuracy of the original model. This is while no other preprocessing technique other than



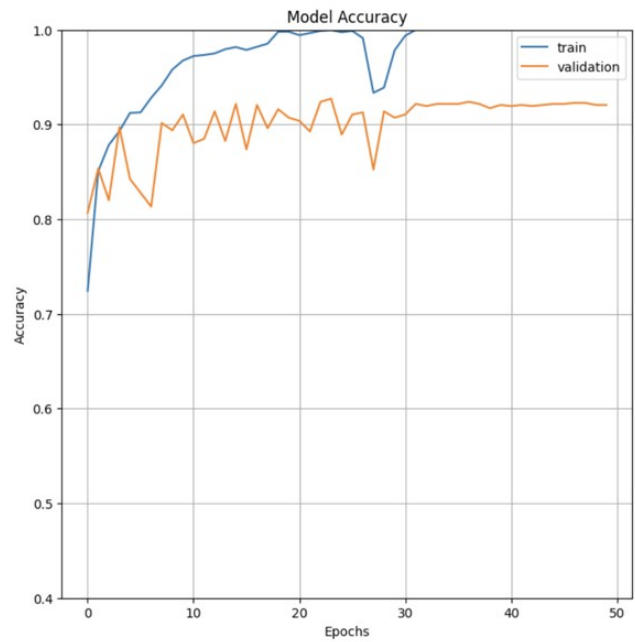
**Fig. 3** The accuracy of the model when utilizing the image standardization technique.



**Fig. 5** The accuracy of the model when the images are rescaled.



**Fig. 4** The accuracy of the model when the images are converted from RGB to BGR images.

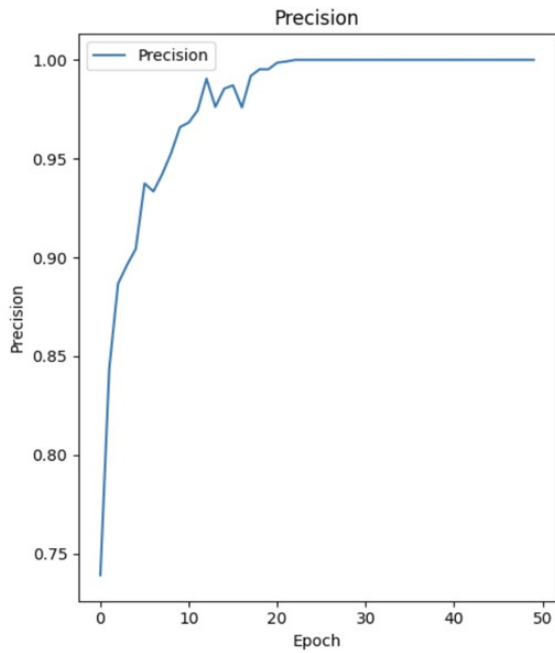


**Fig. 6** The accuracy of the model when the images are normalized.

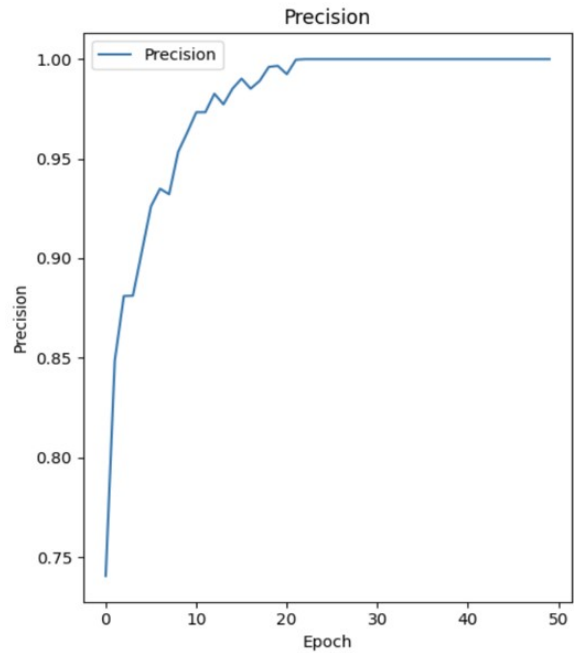
the rescaling technique has outperformed the original model in terms of precision.

Figure 3 (accuracy with standardization) showed reduced underfitting compared to Figure 2 (accuracy with no preprocess-

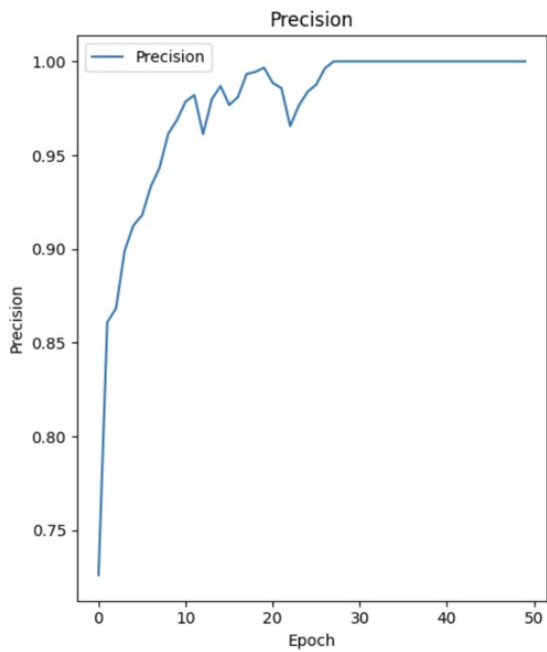
ing), suggesting that the model accuracy portrayed in Figure 2 struggles to learn from the training data and consequently performs poorly. Figure 3, in comparison, indicates standardization has helped the model better learn the patterns present in the



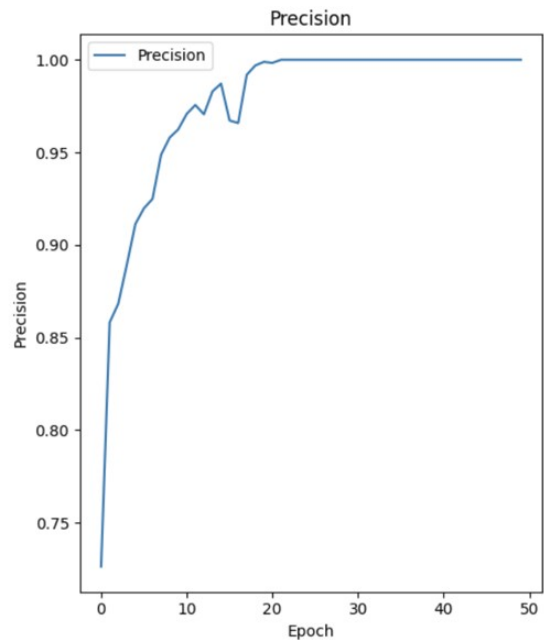
**Fig. 7** The precision of the model with no image preprocessing techniques being applied.



**Fig. 9** The precision of the model when images are rescaled.



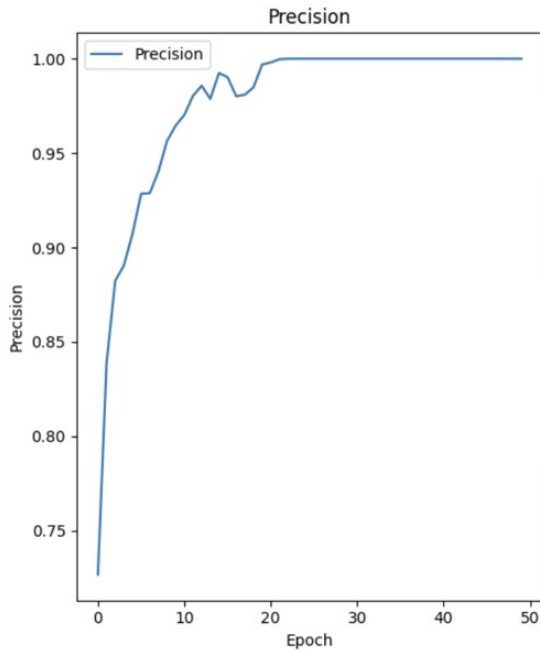
**Fig. 8** The precision of the model when normalization technique is being applied.



**Fig. 10** The precision of the model when utilizing the image standardization technique.

training data. This is while Figure 4 (accuracy with RGB to BGR conversion) exhibited faster convergence, reaching 100% accuracy early. This acceleration might suggest that certain

features essential for diagnosis are emphasized through this color space transformation. Figure 5 (accuracy with rescaling) outperformed standardization, RGB to BGR conversion, and normalization in both accuracy and precision while maintaining



**Fig. 11** The accuracy of the model when the images are converted from RGB to BGR images.

Metrics
Accuracy: $\frac{TP + TN}{TP + FP + TN + FN}$
Precision: $\frac{TP}{TP + FP}$
TP = true positives TN = true negatives FP = false positives FN = false negatives

Image Preprocessing
$X_{\text{normalized}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$
Standardized Pixel Value = $\frac{\text{Original Pixel Value} - \text{Mean}}{\text{Standard Deviation}}$
Rescaled Width = Original Width $\times$ scale_width Rescaled Height = Original Width $\times$ scale_height scale_width = 1.5 scale_height = 1.5
$BGR_{ijc} = RGB_{ij(2-c)}$ BGR <sub>ijc</sub> : The pixel value at position (i, j) for color channel c in the BGR image RGB <sub>ijc</sub> : The pixel value at position (i, j) for color channel c in the RGB image c: color channel (red, blue, green)

sible instances of exaggeration in minor fluctuations or noise present in the dataset. When applied, normalization can have the potential to magnify the impact of outliers in the dataset.

Although most preprocessing techniques improved accuracy, precision did not consistently improve across all methods. Notably, Figure 9 (rescaling) showcased the best precision among the techniques studied. This suggests that while some techniques enhance overall accuracy, they may not significantly impact precision.

Through this research and results shown by Figure 5 and Figure 9, it was concluded that rescaling is a highly effective preprocessing method for retinal disease diagnosis due to its ability to stabilize pixel intensities, enhance convergence, and extract only the significant aspects of the image<sup>22,23</sup>. One study, aiming to provide analysis on the different preprocessing methods, reveals that rescaling essentially reduces variations in pixel values and maintains the fundamental structure of the image unlike the other preprocessing methods that may alter the visual features of the image<sup>23</sup>. With the reduction of abnormal dissimilarity in pixel intensities and improved convergence, the efficiency and overall accuracy of the model is enhanced.

Additionally, research consistently demonstrates that rescaling techniques do not differentiate between individual pixels and are applied uniformly across the entire image, maintaining the overall consistency in visual elements<sup>24</sup>. The even modification of rescaling methodology is crucial in that it maintains the relative relationships between the diverse components within the image.

Healthcare professionals (specialists who focus on the retina) may utilize the results of this study while utilizing automated retinal disease diagnosis as a screening tool in identifying early signs of retinal diseases, specifically DME, CNV, and Drusen.

a balanced convergence with low signs of overfitting or underfitting. Figure 6 (accuracy with normalization) has displayed significant instances of underfitting and overfitting due to pos-

The implementation of the rescaling technique holds substantial promise in the clinical realm as it allows for a much more consistent analysis and detection of retinal abnormalities.

Future perspectives may include examining a wider range of preprocessing techniques (such as histogram equalization, contrast enhancement, etc) beyond what has already been inquired. Not only this, but the application of many effective preprocessing techniques may be compared to the application of only one efficient preprocessing technique in terms of their impact on the final performance metrics of the model being developed.

It also may be suggested to explore different deep learning architectures or models apart from a custom CNN and ResNet-50. The efficacy of transfer learning with other pre-trained models combined with the rescaling technique may also be further investigated. Exploring the scope and depth of deep learning applications can bring forth several advantages, including heightened model performance of disease diagnosis and more precise medical interventions.

In addition to exploring different preprocessing methodologies and architectures, it may be proposed to train the model with Fundus Photography (such that users are not limited to solely OCT scans) and observe whether results differ based on the medical imaging.

## Methodology

To conduct this research, a dataset sourced from Kaggle, and the Python programming language was utilized to construct the model within the Google Colab environment. Previous studies have shown positive outcomes in utilizing these sources<sup>25</sup>.

Using a dataset derived from Kaggle, a well-known platform for open datasets, the data consists of OCT images of the retina<sup>26</sup>. These OCT images are categorized into Diabetic Macular Edema (DME), Drusen, Choroidal Neovascularization (CNV), and normal eye images. The dataset consisted of 1120 images each of DME, Drusen, CNV, and normal eye diseases (4480 images large). The machine learning model was trained with images of each retinal disease as well as normal eye images. OCT scans of each retinal disease and differences in their appearance may be observed in Figure 12. The author would like to emphasize that Kaggle promotes accessibility and transparency in data usage, ensuring that contributors adhere to ethical standards.



**Fig. 12** Representative Optical Coherence Tomography Images

In addition to the existing layers presented by ResNet-50, additional layers were added in order to further strengthen the model's accuracy. These layers included the Flatten layer and two Dense layers. Flatten was used to reshape the input tensor into a single-dimensional array without modifying the data. The Dense layers were fully connected layers responsible for feature extraction and transformation, allowing the network to comprehend complex relationships between features. The first Dense layer utilizes the ReLU (Rectified Linear) activation function, ultimately aiding quicker convergence. The second Dense layer applies the Softmax activation function which serves as the output layer for classification tasks.

After the model was completely built, the program has plotted both the accuracy and precision of the model that was further analyzed. Accuracy would signify how often the model correctly diagnoses a retinal disease based on the used preprocessing technique, while precision would measure the model's ability to avoid false positives when diagnosing DME, CNV, or Drusen, both of which are crucial to investigate.

## Acknowledgements

I would like to express my gratitude towards Lumiere Education who supported me every step of the way. Special thanks are due to my mentor, Pat (Bingqing) Zhang, for her exceptional mentorship and support.

## References

- 1 I. Lbachir, R. Es-Salhi, I. Daoudi and S. Tallal, *A New Mammogram Preprocessing Method for Computer-Aided Diagnosis Systems*, <https://ieeexplore.ieee.org/document/8308280>.
- 2 I. Ebtehaj, H. Bonakdari, M. Zeynoddin, B. Gharabaghi and A. Azari, *Evaluation of preprocessing techniques for improving the accuracy of Stochastic rainfall forecast models - international journal of environmental science and technology*, SpringerLink.
- 3 S. Masoudi, S. Harmon, S. Mehralivand, S. Walker, H. Raviprakash, U. Bagci, P. Choyke and B. Turkbey, *SPIE Digital Library*.
- 4 A. Suganya and A. S. L., *Current medical imaging*, 10 2174 1573405620666230328113721.
- 5 C. Mpyet, *Journal of the West African College of Surgeons*.
- 6 S. Alam and N. Yao, *The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis - computational and Mathematical Organization theory*, SpringerLink.
- 7 K. Arya and M. Saraswat, *Automated microscopic image analysis for leukocytes identification: a survey*.
- 8 D. Pitaloka, A. Wulandari, T. Basaruddin and D. Liliana, *Enhancing CNN with preprocessing stage in Automatic Emotion Recognition*, Procedia Computer Science.

- 
- 9 G. Kumar and P. Bhatia, *Analytical Review of preprocessing techniques for offline handwritten*, [https://www.researchgate.net/publication/255707843\\_Analytical\\_Review\\_of\\_Preprocessing\\_Techniques\\_for\\_Offline\\_Handwritten\\_Character\\_Recognition](https://www.researchgate.net/publication/255707843_Analytical_Review_of_Preprocessing_Techniques_for_Offline_Handwritten_Character_Recognition).
  - 10 L. Chen, C. Tang, Z. Huang, M. Xu and Z. Lei, *Journal of the Optical Society of America. A, Optics, image science, and vision*, **38**, 973–984.
  - 11 S. Masoudi, S. Harmon, S. Mehralivand, S. Walker, H. Raviprakash, U. Bagci, P. Choyke and B. Turkbey, *SPIE Digital Library*.
  - 12 D. Murcia-Gómez, I. Rojas-Valenzuela and O. Valenzuela, *Applied Sciences*, **12**, 11375.
  - 13 D. Sarvamangala and R. Kulkarni, *Convolutional neural networks in medical image understanding: A survey*, SpringerLink.
  - 14 A. Sukanya, K. Krishnamurthy and T. Balakrishnan, *Current medical imaging*, **16**, 776–780.
  - 15 V. Jasti, A. Zamani, K. Arumugam, M. Naved, H. Pallathadka, F. Sammy, A. Raghuvanshi and K. Kaliyaperumal, *Security and Communication Networks*.
  - 16 S. Muchuchuti and S. Viriri, *MDPI*.
  - 17 B. Rashed and N. Popescu, *Systematic Literature Review. Sensors (Basel, Switzerland)*, **22**, 7065.
  - 18 A. Elsharif and S. Naser, *International Journal of Academic Engineering Research*.
  - 19 P. Uppamma and S. Bhattacharya, *Journal of healthcare engineering*, 2728719.
  - 20 V. K., *Biomedicine*.
  - 21 O. Musat, C. Cernat, M. Labib, A. Gheorghe, O. Toma, M. Zamfir and A. Boureanu, *Romanian journal of ophthalmology*.
  - 22 K. Maharana, S. Mondal and B. Nemade, *Global Transitions Proceedings*.
  - 23 A. Bhandari, *Analytics Vidhya*.
  - 24 M. Anderson, *Better Machine Learning Performance through CNN-based image resizing*, Unite.AI.
  - 25 A. Chatterjee, M. Gerdes and S. Martinez, *A machine learning overview. A Machine Learning Overview*, <https://pubmed.ncbi.nlm.nih.gov/32403349/>.
  - 26 N. Eladawi, M. Elmogy, M. Ghazal, O. Helmy, A. Aboelfetouh, A. Riad, S. Schaal and A. El-Baz, *Frontiers in bioscience (Landmark edition)*, **23**, 247–264.