Cataract Detection Via near infrared Eye Images

Vikalp Kumar Tripathi
DSAI, IIIT Naya Raipur
vikalp20101@iiitnr.edu.in

Anand Kumar Sahu
DSAI, IIIT Naya Raipur
anand20102@iiitnr.edu.in

Abstract
As we all know, age-related blindness problems and eye diseases are mainly caused due to cataract and also according to the National Blindness and Visual Impairment Survey of India**[1]**, cataract is the primary cause of blindness after the age of 50 years, which contributes to 66.2% of blindness cases and 80.7% of severe visual impairment cases. This percentage will surely increase, and more than our traditional treatment methodology will be required to detect every case. The need for more resources and the unavailability of a sufficient number of ophthalmologists will burden the healthcare system. To deal with this emerging problem, we need to automate the system, which will have the capability to detect cataracts in the patient. We provide a noble solution that uses NIR (Near Infrared images) to detect cataracts. For this, we have developed a dataset that contains only masked images of all the available images of the dataset(Periocular Dataset), with the help of which we are going to perform segmentation with the help of Unet and DeepLabV3 and classification of the presence of cataracts using ResNet50 and make a comparison between their working accuracy. The suggested segmentation approach quickly and economically detects imperfect eye borders, and the classification network performs exceptionally well on the cataract dataset.

Keywords: Cataract, Ophthalmologist, NIR, DeepLabV3, Unet, Deep Learning, ResNet50

1. Introduction

As we all know, cataract is an age-related condition where a foggy coating forms over the eye lens, blurring it and resulting in poor vision. The classic technique of cataract detection uses a slit lamp or an ophthalmoscope to capture pictures of the eyes, and the ophthalmologist analyses and tests the patient's eyes to determine whether a cataract is present. However, there aren't enough professionals in such a large community to see all the patients promptly. To solve this issue more effectively, we're developing a model that uses segmentation of the dataset's eye pictures to identify the existence of cataracts and categorize each eye as healthy or unhealthy. To do this, we have created a dataset consisting of only masked images of the eyes of different humans from different angles from the periocular dataset as no publicly masked images dataset was available, so we have manually labeled all the images using label studio.

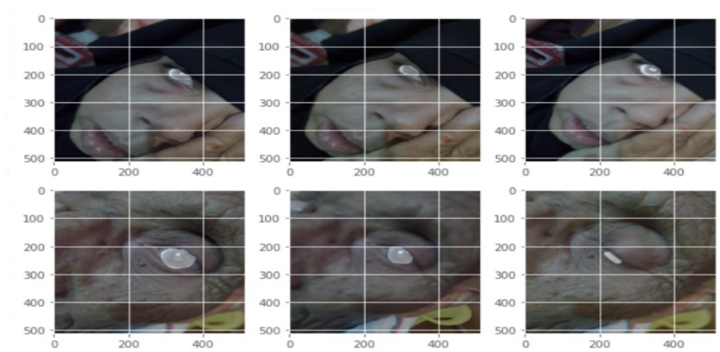

Fig1: Dataset with mask

The primary motivation for solving this problem is the lack of a sufficient number of specialists is a concern in healthcare in this age of AI and technology when we are in a race for advancement. And the issue we're focusing on is a major one since, if not treated on time, it affects nearly everyone until old age and will result in irreversible blindness. Therefore, why not utilize our expertise to make all processes automatic and accurate at the beginning when we only need to determine whether a cataract is present or not. This will assist in obtaining the proper therapy at the appropriate moment, which can solve their blindness problem and other eye diseases.

Our main contributions are:

- Creating a dataset that contains only masked images of eyes with the help of which segmentation can be done quickly among various cases.
- Identifying which method will be better to replace old cataract detection manual methods by comparing various old and new deep learning methods.

2. Literature Review

For a long time, ophthalmic images of eyes have been used rigorously to assess the severity of cataracts. They have been proven very accurate and helpful in recognizing and classifying cataract disease**[2]**. For detecting cataracts generally, we use six types of ophthalmic images(fundus images**[3]**, tomography images**[4]**, retro illumination images**[5]**, digital camera images**[6]**, ultrasonic images**[7]**, and slit-lamp images**[8]**). Out of these six, the two imaging techniques most frequently employed in cataract identification are slit-lamp images and fundus images. Rural residents, however, may find it difficult to get slit-lamp and fundus imaging equipment. But the digital camera is easily accessible in these places compared to the equipment required for ophthalmic images for cataract detection **[9]**. Therefore, a technology that can detect the presence of cataracts with the help of digital camera images is required.

Nayak**[10]** used SVM to perform the tasks of classification while selecting features from the pupil region of images utilizing image preprocessing methods like big ring area and edge pixel count. Fuadah et al. **[11]** manually converted all the images of the digital cameras into grayscale. The extracted features are contrast, uniformity, and differences with the help of a co-occurrence matrix of gray level. At last, KNN was used to classify the eye images into healthy or not healthy, that is, whether a cataract is present or the eye is normal. Kumar**[12]** developed a new algorithm for cataract detection known as a texture features-based algorithm that works on the intensity of the occurrence of cataracts from the images captured by digital cameras.

An app was designed by Agarwal et al.**[13]** for detecting the presence of cataracts using the smartphone's camera as a medium to take images of an individual's eye. Single-layer perceptron was used by Sigit et al. **[14]** to develop a smartphone system for detecting cataracts to reduce the workload of ophthalmologists. His system accuracy was very high while performing classification. Yusuf et al.**[15]** emphasized that transfer learning that was learned on ImageNet had an impact on the classification accuracy and based on this emphasis, he proposed a website to detect presence of cataracts with the help of CNN by just uploading the eye images taken by the digital cameras.

3. Proposed Solution

We have proposed a model in which we directly do segmentation of the input images which will create the masked of the input images and then it will perform classification of the obtained images into healthy or non-health category.

But, for the segmentation model we are using Unet and DeepLabV3 and for doing the training of them we are going to convert the periocular dataset images into masked images manually with the help of Label Studio and after this training our segmentation model will be able to directly generate masked images and can successfully classify that image. For the classification we are going to use the pre-trained model ResNet50.

3.1 Dataset Creation

For doing perfect segmentation of the input images we need to train our segmentation model with the help of masked images. But for doing this no publicly dataset is available according to our knowledge which consists of both images as well as their corresponding masked images. So for doing this we have used Cataract Mobile Periocular Database (CMPD) provided by IITJ which consists of pre and post cataract surgery images of individuals. This dataset consists of raw images of both eyes of patients suffering from cataract before and after surgery but does not contain their respective masked images. So, we manually created masked images of all the images present in CMPD with the help of label studio and created a dataset of only masked images which will be used for training our segmentation models.

3.2 Model Architecture

Our architecture consists of twosubparts segmentation as well as classification. For segmentation we are going to use Unet and DeepLabV3 and for classification we are going to use ResNet50.

3.2.1 Segmentation

In this part of our model architecture we are going to take the input from the user as the NIR image of the eye and will pass into our segmentation model to convert the corresponding image into its mask and then will be provided as input to the next model. For segmenting we are going to use Unet and DeepLabV3.

3.2.1.1 Unet

Deep learning world was introduced with Unet architecture in the year 2015. This architecture used to segment various kind of images such as microscopy images which are transmitted by light and electron microscopic stacks of some neuronal structures are some examples of the situations where this architecture can be used. People have done various modification in this architecture because of which it has various variants now and out of all these variants we are going to use the architecture shown in Fig2:

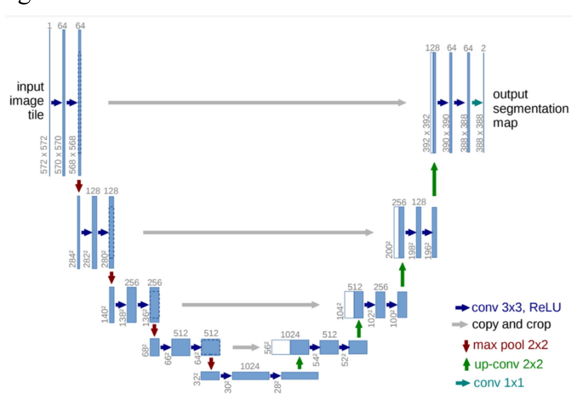

Fig2: Unet Architecture[16]

An input image is processed before going through a few convolutional layers using the ReLU activation function, as shown by the model's architecture. As can be seen, the image size drops from 572X572 to 570X570 and eventually to 568X568. This reduction in dimensionality was brought about by the employment of unpadded convolutions, which is what caused it. We can also see an encoder block on the left and a decoder block on the right in addition to the Convolution bricks. Strides 2's max-pooling layers help the encoder block keep the size of the images consistently decreasing. When we reach the decoder section, we observe that the number of filters in the convolutional layers begins to decrease and that the next layers steadily upsample until we reach the top. We also watch as skip connections are used to connect the decoder blocks' layers to earlier outputs.For the loss from the earlier layers to reflect more strongly on the overall values, this skip connection is a crucial idea.They have also been shown scientifically to produce superior results and speed model convergence. A few additional convolutional layers are placed after the last convolution layer in the final convolution block. To show the output, a filter of 2 with the appropriate function is placed in this layer.

3.2.1.2 DeepLabV3

DeepLabv3, a semantic segmentation architecture, makes a number of improvements over DeepLabv2. The problem of segmenting objects at many sizes is addressed by modules that employ atrous convolution in cascade or parallel to capture multi-scale context by adopting numerous atrous rates. In order to boost efficiency and encapsulate global context, image-level features from the Atrous Spatial Pyramid Pooling module of DeepLabv2 were also included. The Fig3 shows the architecture we are going to use in our case.

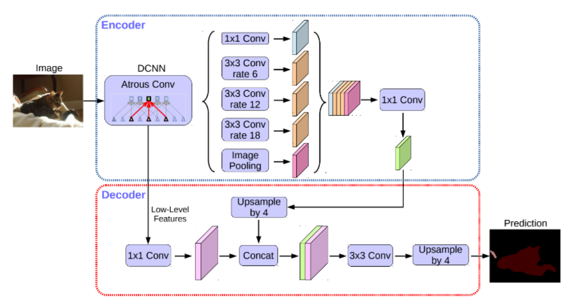

Fig3: DeepLabV3 Architecture[17]

In this architecture, while the decoder module polishes the segmentation outcomes at object boundaries, the encoder module handles multiscale contextual information by using dilated convolution at several scales. Dilated convolution allows us to maintain a consistent stride with a bigger field of view as we move further into the network without adding more parameters or processing power. Additionally, it permits bigger feature maps to be generated, which is helpful for semantic segmentation. It has been demonstrated that when the sampling rate increases, there are less valid filter weights—that is, weights that are applied to the valid feature region rather than padding zeros—which is why dilated spatial pyramid pooling is used. After being bilinearly upsampled by a factor of 4, the encoder features are concatenated with the comparable low-level features of the network backbone that have the same spatial resolution.

3.2.2 Classification

After performing segmentation and obtaining the masked image of the input image we are going to perform the classification by passing the masked image into our pretrained model (here we are going to use ResNet50) which will provide us with the output that whether the eye is healthy or unhealthy.

3.2.2.1 ResNet50

A convolutional neural network with 50 layers is called ResNet-50. Many computer vision applications are built on the conventional neural network known as ResNet, or Residual Networks. ResNet's primary innovation was the ability to train extremely complex neural networks with more than 150 layers.Backpropagation drastically reduces the gradient value, which results in very little change in weights. This is circumvented by ResNet. It utilises the "SKIP CONNECTION" feature. The architecture we are going to use is shown in Fig3.

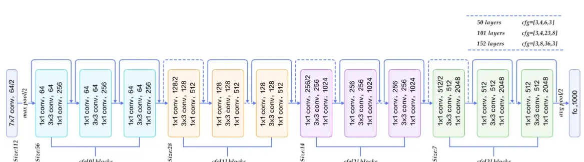

Fig4: ResNet50 Architecture[18]

The ResNet architecture adheres two design rules. First, there are the same number of filters in each layer regardless of the output feature map's size. Second, even though the size of the feature map is cut in half, it contains twice as many filters to maintain the time complexity of each layer. In the 50-layer ResNet, the bottleneck construction block is utilised. A bottleneck residual block, often known as a "bottleneck," reduces the number of parameters and matrix multiplications by using 11 convolutions. This greatly accelerates the training of each layer.

4. Results

We have combined both segmentation methods with classification method and tested our dataset on both these cases and obtained accuracy of 73% in the case of Unet+ResNet and accuracy of 72% in the case of DeepLabV3+ResNet and shown their respective plots and comparison table below.

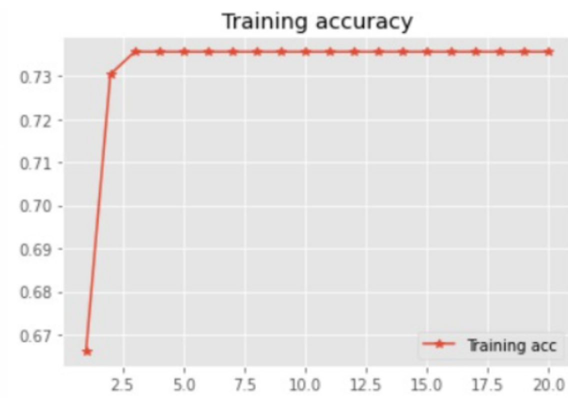

Fig5: Unet+ResNet Accuracy

Fig6: Unet+ResNet Loss

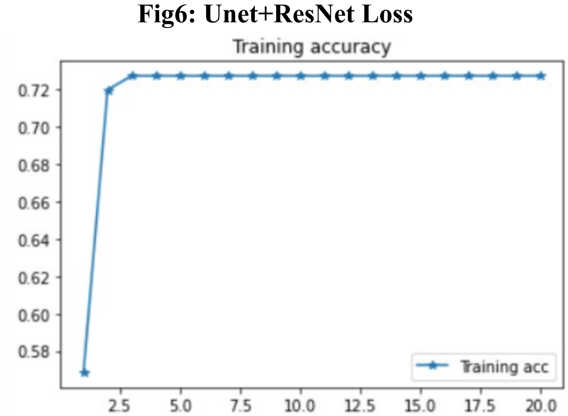

Fig7: DeepLabV3+ResNet Accuracy

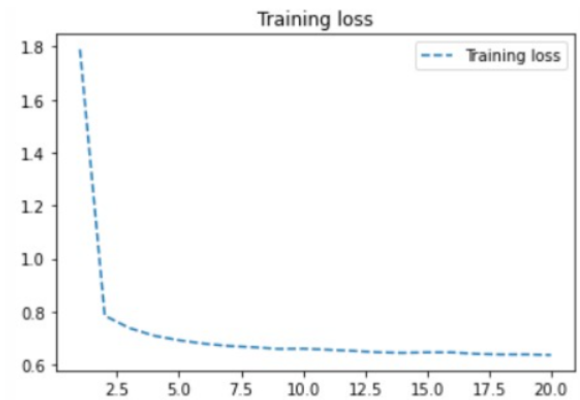

Fig8: DeepLabV3+ResNet Loss

Model	Accuracy	Loss
Unet + Resnet50	73.5%	0.6245
DeepLab v3+ Resnet50	72%	0.64

Table1: Results Comparison

5. Conclusion

The most frequent surgical procedure is cataract surgery, which is one of the main causes of vision impairment in the world. The prognosis, ongoing evaluation, and choice of whether to proceed with surgery are typically left largely up to the ophthalmologist's judgement. It is crucial to have a clinical decision-support system to increase the sensitivity and specificity of cataract identification and monitoring in settings with limited resources and experts. A deep learning technique for cataract diagnosis is presented in this paper and compared two of the algorithms Unet and DeepLabV3 and found out that Unet is performing better than DeeLabV3 and can be further improved by combining different pretrained model other than ResNet50 in future for getting better results than the outcomes we are getting now.

References

[1] National Blindness and Visual Impairment Survey of India.

[2] Litjens, G.; Kooi, T.; Bernard, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. Med. Image Anal. 2017, 42, 60–88.

[3] Xu, X.; Zhang, L.; Li, J.; Guan, Y.; Zhang, L. A hybrid global-local representation CNN model for automatic cataract grading. IEEE J. Biomed. Health Inform. 2019, 24, 556–567.

[4] Zhang, X.; Xiao, Z.; Higashita, R.; Chen, W.; Yuan, J.; Fang, J.; Hu, Y.; Liu, J. A Novel Deep Learning Method for Nuclear Cataract Classification Based on Anterior Segment Optical Coherence Tomography Images. In Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, ON, Canada, 11–14 October 2020.

[5] Zhang, W.; Li, H. Lens opacity detection for serious posterior subcapsular cataract. Med. Biol. Eng. Comput. 2017, 55, 769–779.

[6] Pathak, S.; Kumar, B. A robust automated cataract detection algorithm using diagnostic opinion based parameter thresholding for telemedicine application. Electronics 2016, 5, 57.

[7] Caixinha, M.; Amaro, J.; Santos, M.; Perdigão, F.; Gomes, M.; Santos, J. In-vivo automatic nuclear cataract detection and classification in an animal model by ultrasounds. IEEE Trans. Biomed. Eng. 2016, 63, 2326–2335.

[8] Qian, X.; Patton, E.W.; Swaney, J.; Xing, Q.; Zeng, T. Machine learning on cataracts classification using SqueezeNet. In Proceedings of the 2018 4th International Conference on Universal Village (UV), Boston, MA, USA, 21–24 October 2018.

[9] Zhang, X.; Fang, J.; Hu, Y.; Xu, Y.; Higashita, R.; Liu, J. Machine Learning for Cataract Classification and Grading on Ophthalmic Imaging Modalities: A Survey. arXiv 2020, arXiv:2012.04830.

[10] Nayak, J. Automated classification of normal, cataract and post cataract optical eye images using SVM classifier. In Proceedings of the World Congress on Engineering and Computer Science, San Francisco, CA, USA, 23–25 October 2013.

[11] Fuadah, Y.N.; Setiawan, A.W.; Mengko, T.L.R. Performing high accuracy of the system for cataract detection using statistical texture analysis and K-Nearest Neighbor. In Proceedings of the 2015 International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, Indonesia, 20–21 May 2015.

[12] Pathak, S.; Kumar, B. A robust automated cataract detection algorithm using diagnostic opinion

based parameter thresholding for telemedicine application. Electronics 2016, 5, 57.
[13] Agarwal, V.; Gupta, V.; Vashisht, V.M.; Sharma, K.; Sharma, N. Mobile application based cataract detection system. In Proceedings of the 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 23–25 April 2019.
[14] Sigit, R.; Triyana, E.; Rochmad, M. Cataract Detection Using Single Layer Perceptron Based on Smartphone. In Proceedings of the 2019 3rd International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Indonesia, 29–30 October 2019.
[15] Yusuf, M.; Theophilous, S.; Adejoke, J.; Hassan, A.B. Web-Based Cataract Detection System Using Deep Convolutional Neural Network. In Proceedings of the 2019 2nd International Conference of the IEEE Nigeria Computer Chapter (NigeriaComputConf), Zaria, Nigeria, 14–17 October 2019.
[16] https://blog.paperspace.com/unet-architecture-image-segmentation/
[17] https://keras.io/examples/vision/deeplabv3_plus/
[18] https://blog.devgenius.io/resnet50-6b42934db431