Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
May 2024 5(5), 141–148

ASSESSING THE PERFORMANCE OF CATARACT NET AND OTHER DEEP LEARNING SYSTEMS FOR AUTOMATED CATARACT DETECTION

Pragya Shrivastava [1], Chandra Shekhar Gautam [1], Sajal Kumar Kar [1]

[1] Department of Computer Science and Engineering AKS University SATNA

Corresponding Author

Pragya Shrivastava,
pspragyashri@gmail.com

DOI
10.29121/shodhkosh.v5.i5.2024.3612

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

The normal lens of the eye, which is located behind the iris and pupil, becomes clouded when a cataract develops. Normally clean, the lens aids in focusing light onto the retina, enabling sharp vision. The formation of a cataract results in an opaque or clouded lens, which distorts or blurs vision. Although aging is frequently linked to cataract development, additional causes include heredity, trauma, certain drugs, or underlying medical disorders like diabetes. The usual symptoms are progressive loss of vision, heightened susceptibility to light, blurred or yellowed colors, and difficulties seeing at night. A thorough eye exam that includes slit-lamp and visual acuity tests is typically used to diagnose cataracts.

Keywords: Machine learning, Deep learning, Convolutional Neural Network (CNN), ResNet

1. INTRODUCTION

The World Health Organization (WHO) and several ophthalmology research indicate that corneal disorders are the primary cause of visual impairment in the world. Roughly 12.7 million people worldwide are believed to be blind due to corneal disease, making up a sizeable portion of the visually impaired population. Corneal blindness is more common and is primarily caused by trauma, infections, and inadequate access to healthcare. For example, corneal blindness can be a major cause of visual impairment in various regions of Sub-Saharan Africa and South Asia.

Machine learning (ML) is transforming the field of corneal diagnostics through its ability to improve image processing, forecast the course of disease, and customize treatment. Machine learning algorithms are able to accurately identify disorders like keratoconus and corneal dystrophies by analyzing corneal pictures from devices such as topographers and OCTs. By analyzing patient data, they are able to forecast the course of a disease and evaluate its response to treatment. Additionally, machine learning (ML) facilitates remote screening, lowers diagnostic mistake rates, and combines various data sources for thorough evaluations. With regard to corneal health, this technology streamlines treatment regimens, enables early intervention, and enhances patient care overall. We will compare machine learning algorithms for corneal detection in this research.

2. METHODOLOGY

We are considering two papers “Cataract Detection using Deep Learning” by Saroj Kailash Panda and Nikhil Panjwani as paper 1 and “CataractNet: An Automated Cataract Detection System Using Deep Learning for Fundus Images” by Masuam Shah Junayed, MD Baharul Islam, Arezoo Sadeghzadeh and Saimunur Rehman as paper2.Both the papers are having their own designed models and their model is compared by the already existing model such as VGG16, VGG19, ResNet. Paper, one has designed their own model and the name given to that model is “Ourmodel” and paper two has designed their own model and its name is given as “CataractNet”.

We are going to compare the models on the basis of performance metrics like Accuracy, Precision, Recall and F1 Score.

2.1. EXISTING DATASET

- **OURMODEL**

The Indian diabetic retinopathy image (IDRiD) dataset, the fundus image registration (FIRE) [13], the ACHIKO-I fundus image dataset, the color fundus image database, the digital retinal images for vessel extraction (DRIVE) [14] database, and the high resolution fundus dataset HRF [15] are among the datasets used in this study. are merged, rearranged, and subjected to pre-processing in order to generate a cataract dataset. Then, it is enlarged to a substantial number of photographs using a data augmentation technique.

- **CatracactNet**

A cataract dataset, also known as the high-resolution fundus (HRF), is gathered, rearranged, and pre-processed from many standard datasets of fundus pictures published throughout the past 20 years [15]. image archive, ACHIKO-I fundus image dataset [16], Indian diabetic retinopathy imaging dataset (IDRiD) [17], fundus image registration (FIRE) [16] dataset, and color fundus image database [18].as well as digitized retinal pictures for the vascular extraction database (DRIVE) [19]. Then, via the data augmentation procedure, it is expanded to a sizable number of photos.

- **Introduction of Some** Existing Model used in CNN for FUNDUS Images

VGG16

A deep convolutional neural network architecture called VGG16 was unveiled by the University of Oxford's Visual Geometry Group (VGG). It is renowned for being straightforward and efficient in identifying intricate patterns in photos, and it was created with picture classification in mind.

3. ARCHITECTURE

- **Layers:** There are three fully connected layers and thirteen convolutional layers among the sixteen learnable parameters that make up VGG16
- **Convolutional Layers** It downsamples and captures features using 2x2 max-pooling layers and tiny 3x3 convolutional filters, which aids in the detection of intricate patterns.
- **Fully Connected Layers:** Three completely linked layers at the conclusion of the network assist in categorizing the extracted features.

VGG19

VGG19 is a variant of the VGG architecture, extending the original VGG16 model. Here’s a concise overview of its architecture and features:

Architecture:

- Layers: VGG19 includes 19 layers with learnable parameters—16 convolutional layers and 3 fully connected layers.
- Convolutional Layers: It employs 3x3 convolutional filters, similar to VGG16, but with more layers to increase depth. It also includes 5 max-pooling layers with 2x2 pooling windows to reduce spatial dimensions.
- Fully Connected Layers: At the end, there are 3 fully connected layers that perform the classification.

ResNet

ResNet (Residual Network) is a deep learning architecture designed to address the challenges of training very deep neural networks. Introduced by Microsoft Research in 2015, ResNet features a novel approach that significantly improves performance in image classification and other tasks.

Architecture:

Residual Blocks: The concept of residual learning is introduced by ResNet. Residual blocks learn the residual (or difference) between the input and output rather than the desired mapping directly. Gradients can go across the network more efficiently during training thanks to shortcut connections that bypass one or more layers.

Layers: ResNet architectures vary in depth. Common versions include ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152, where the numbers indicate the total number of layers

4. PROPOSED MODEL

Our Model:

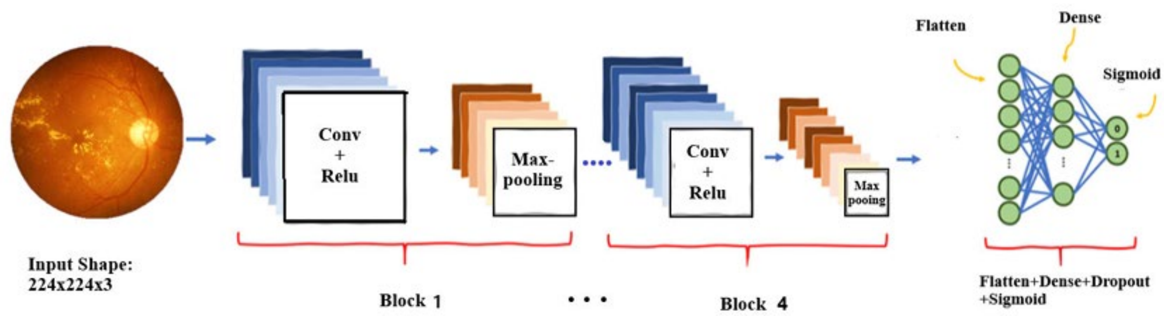
OURMODEL has SIXTEEN layers total, with 50% of those layers being in the first four blocks and the remaining layers being for grading. The RGB pictures (224×224) and 32 filters with Ks(3×3) make up the first block of inputs.Max pooling layers are used, with a stride of two.in an effort to save space. containing blocks with ReLu activation features. An identical block is used as the third block, except it has 64 liters, while the fourth block has 128 liters. The output from each of these is sent to the remaining levels, which are fully coupled to one another and consist of the drop out and dense layers. Since this is a binary classification model, a sigmoid function is being used. In order to look into how block numbers impact the accuracy of classification, cataract

Catract Net

Convolutional neural networks (CNNs) are deep neural networks that use non-linear activation functions, convolutional and pooling layers, and other techniques to build complicated hierarchies of features [11], [12]. While these two phases are separated in manual feature extraction methods, they are merged in deep learning-based approaches throughout the feature extraction phase and during the classification process. To overcome the drawbacks of manual feature extraction and lower the computational cost, a brand-new deep learning model called CataractNet is put out. It has sixteen layers total. Four blocks, each with two levels, make up half of the layers, with the remaining layers being used for classification. The inputs for the remaining block consist of 32 filters with kernel and RGB (three input channels) images with a size of 224 224.

The second block is the same block with the same parameter values. The third block is then an identical block with 64 filters this time. The number of filters is raised to 128 in the fourth block. The four blocks' aggregate outputs are fed into the fully connected layers as a feature map. The purpose of these layers—the dropout, thick, and flatten layers—is to detect cataracts. To gather the filtered cataract features, three sets of dense and dropout layers are built, with the dense layers being characterized by 64, 128 and 256 flattened neurons. Moreover, to stop the model from overtightening by setting 40%, 40%, and 50%, three dropout layers are set to 0.4, 0.4, and 0.5.

Figure 1: Architecture of the CataractNet Model



4.1. PERFORMANCE COMPARISON ON THE BASIS OF PERFORMANCE METRICES:

Sno.	Train:test	Category	Result from Paper A				Result from Paper B				Conclusion
			Accuracy%	Precision %	Recall%	F1 score	accuracy%	Precision %	Recall%	F1 score	
	90-10	VGG16	91.35	86.19	91.92	89.94	91.35	86.19	91.92	89.54	F1 score is different
1		VGG19	91.88	86.68	91.76	89.97	92.88	86.68	91.76	89.97	accuracy is different
2		ResNet	91.11	85.69	90.43	86.52	91.18	85.69	90.43	86.52	accuracy is different
3		ourmodel	92.92	86.37	95.43	89.59					
4		CatractNet					91.92	86.37	95.43	89.51	accuracy & F1 score different
	80-20	VGG16	91.35	86.19	91.93	90.54	91.35	86.19	91.92	90.54	recall is different
1		VGG19	92.88	86.68	91.92	89.97	92.88	86.68	91.76	89.97	recall is different
2		ResNet	97.41	96.75	97.39	97.04	97.66	96.75	97.38	97.04	accuracy & recall is different
3		ourmodel	99.13	99.08	99.17	99.07					
4		CatractNet					99.13	99.08	99.07	99.07	recall is different
	70-30	VGG16	94.45	86.19	96.49	94.34	94.89	86.19	92.86	90.54	all different
1		VGG19	95.63	86.68	93.68	96.13	95.63	86.68	93.76	95.97	recall & f1 score different
2		ResNet	96.21	95.62	97.43	96.78	96.26	95.62	97.43	98.56	accuracy & f1 score is different
3		OURMODEL	98.96	98.24	99.08	98.15					
4		CatractNet					98.96	98.24	99.08	98.97	f1 score is different

Figure 2: Performance Comparison on the basis of Performance metrices Of Paper1 and Paper2

VGG16 when Train-Test Percentage is 90-10

Accuracy percentage is 91.35% which is same for both the paper A and B. Precision is 86.19 % which again same for both the papers. Recall is 91.92% which same in both the papers. F1 score is 89.94% in paper A and 89.54% is paper B. So, all the parameters are same except F1 score and the difference is 0.4%.

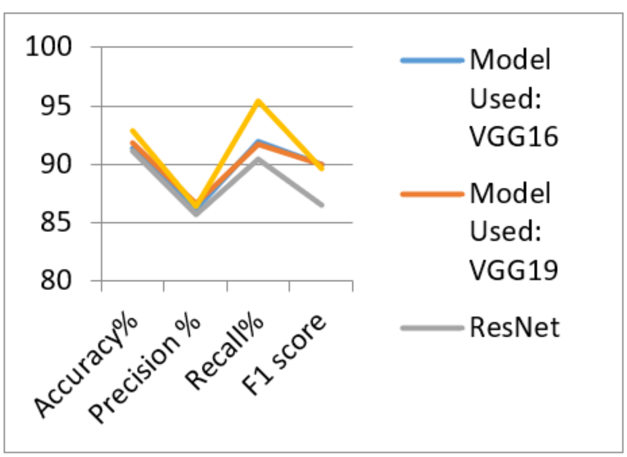


Figure3 Comparison of models in paper A (90-10)

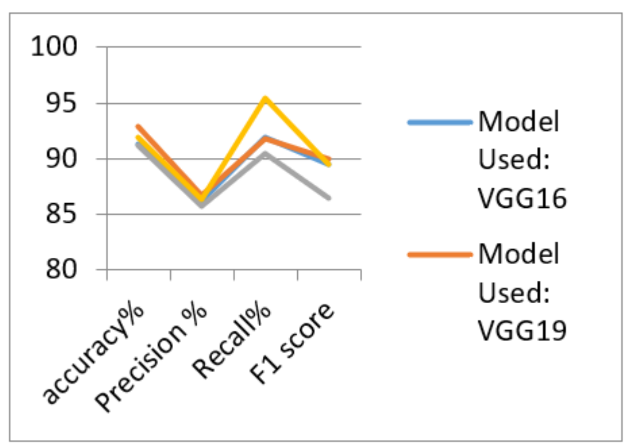
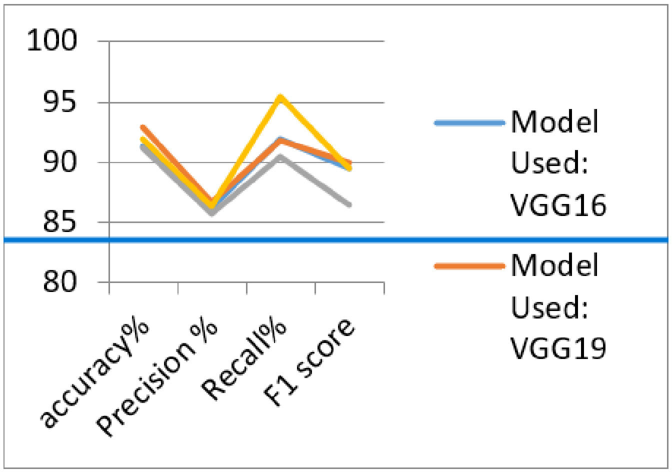


Figure 4 Comparison of paper B (90-10)



VGG16 when Train-Test Percentage is 80-20

Accuracy percentage is 91.35% which is same in both the papers A and B. Precision is 86.19% same on both the papers. Recall is 91.93% in paper A and 91.92% in paper B.F1 score is 90.54% in both the papers. So , in the case when they are using 80-20 percentage of train test then accuracy, precision, F1 score are same only Recall is different by 0.1%.

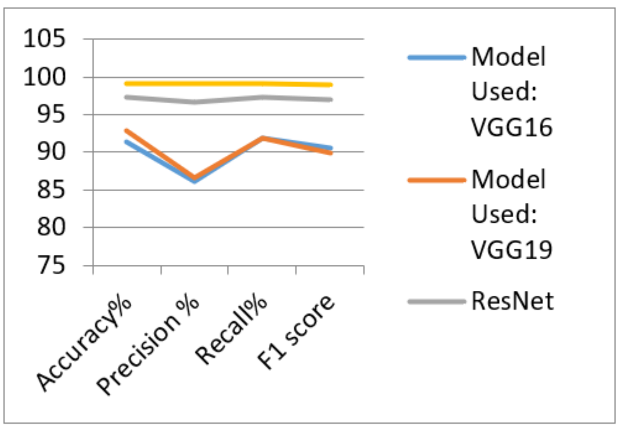


Figure 5 Comparison of paper A (80-20)

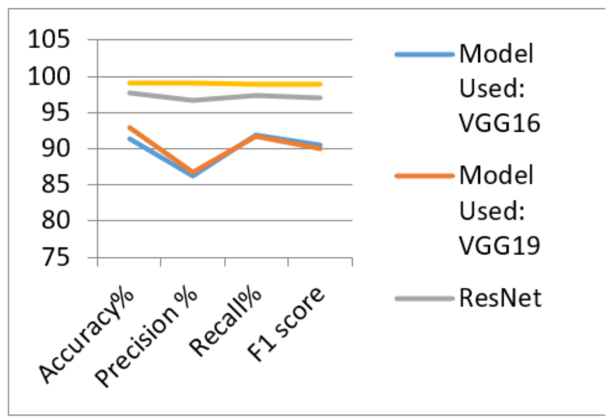


Figure 6 Comparison of paper B (80-20)

VGG16 when Train-Test Percentage is 70-30

Accuracy percentage is 94.45% in paper A and 94.89% in paper B. Precision is 86.19% in paper A and same in paper B. Recall in paper A is 96.49% and that on paper B is 92.86%.

F1 score is 94.34% in paper A and 90.54% in paper B. So, in this case accuracy is different by 0.44%. Recall is different by 3.63%. F1 score is different by the 3.89%. Only one parameter is same that is Precision.

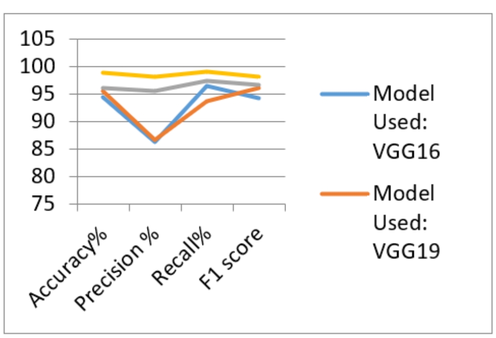


Figure 7 Comparison of paper A (70-30)

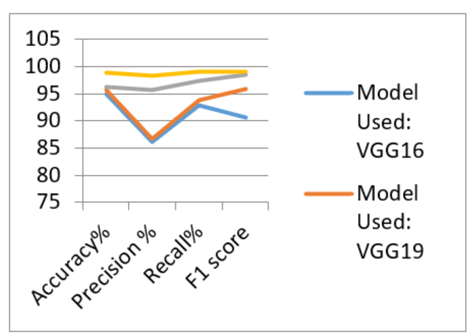


Figure 8 Comparison of paper B (70-30)

VGG19 when Train-Test Percentage is 90-10

Accuracy percentage is 91.88% in paper A and 92.88% in paper B. Precision is 86.69% in paper A and same in paper B. Recall is 91.76% which is again same in both the papers. F1 score is 89.97% which is same in both the papers. All the parameters viz Precision, Recall, F1 score are same in this case only Accuracy is different by 1%.

VGG19 when Train-Test Percentage is 80-20

Accuracy percentage is 92.88% which is same in both the papers. Precision is 86.68% which is again same in both the papers. Recall is 91.92% in paper A and 91.76% in paper B. F1 score is 89.97% which is same in both the papers. In this case Recall is different and the difference is of 0.16%

VGG19 when Train-Test Percentage is 70-30

Accuracy percentage is 95.63% which is same in both the papers. Precision is 86.68% which is again same in both the papers. Recall is 93.68% in paper A and 93.76% in paper B. F1 score is 96.13% in paper A and 95.97 % in paper B. in this case Recall is different by 0.08% and F1 score is different by 0.16%.

ResNet when Train-Test Percentage is 90-10

Accuracy percentage is 91.11% in paper A and 91.18% in paper B. Precision is 85.69% in paper A and same in paper B. Recall is 90.43% which is same in both the papers. F1 score is 86.52% which is same in both the papers. All the parameters viz Precision, Recall, F1 score are same in this case only Accuracy is different by 0.07%.

ResNet when Train-Test Percentage is 80-20

Accuracy percentage is 97.41% in the paper A and 97.66 in paper B. Precision is 96.75% which is again same in both the papers. Recall is 97.39% in paper A and 97.38% in paper B. F1 score is 97.04% which is same in both the papers. In this case accuracy and Recall is different.

Here Accuracy is different by 0.25% and Recall is different by 0.01%.

ResNet when Train-Test Percentage is 70-30

Accuracy percentage is 96.21% in paper A and 96.26% in paper B. Precision is 95.62% which is again same in both the papers. Recall is 97.43% which is same in both the papers. F1 score is 96.78% in paper A and 98.56 % in paper B. In this case Accuracy and F1 score is different.

Here Accuracy is different by 0.05% and F1 score is different by 1.78%.

Our Model when Train-Test Percentage is 90-10

Ourmodel Accuracy percentage is 92.92 %. Precision is 86.37%. Recall is 95.43 %. F1 score is 89.59%.

Our Model when Train-Test Percentage is 80-20

Accuracy percentage is 99.13 %. Precision is 99.08%. Recall is 99.17 %. F1 score is 99.07%.

Our Model when Train-Test Percentage is 70-30

Accuracy percentage is 98.96 %. Precision is 98.24%. Recall is 99.08%. F1 score is 98.15%.

All of the above data derived from paper A only.

CatractNet when Train-Test Percentage is 90-10

Accuracy percentage is 91.92 %. Precision is 86.37%. Recall is 95.43 %. F1 score is 89.51%.

CatractNet when Train-Test Percentage is 80-20

Accuracy percentage is 99.13 %. Precision is 99.08%. Recall is 99.07 %. F1 score is 99.07 %.

CatractNet when Train-Test Percentage is 70-30

Accuracy percentage is 98.96 %. Precision is 98.24%. Recall is 99.08%. F1 score is 98.97%.

All of the above data derived from paper B only.

5. RESULT

On the complete analysis of both the paper we came to the conclusion that the when they are considering train-test ratio as 90-10 then at that case Our model's performance is better than the CataractNet. Here accuracy of the "Ourmodel' is 1% more than the accuracy of the CataractNet. F1 score of the Ourmodel is 0.1% more than the CataractNet. When they are considering train-test ratio as 80-20 then the value of recall of the Ourmodel is 0.1% more than then CataractNet.When they are considering train-test ratio as 70-30 then the value of F1 Score of the CataractNet is 1.82% more than the Ourmodel.

Thus if we take the average of the above results then we can see that the performance metrices (acuuracy,F1 score) of the Ourmodel is more. So, Ourmodel is showing better performanve than the CataractNet.

6. CONCLUSION

We provided a comparison of two models in this research. CataractNet and our model. Lightweight deep learning is the foundation of CataractNet, an automated cataract detection system. To improve the dataset to feed the deep network, a cataract dataset consisting of fundus images was first reorganized, pre-processed, and augmented. With the purpose of exploring several layers, the created Cataract Net made use of optimization methods, loss functions, and Deep Learning functions to minimize computing costs without compromising model accuracy. However, our algorithm can identify cataract with a very high accuracy rate from the provided fundus images. This is also lightweight, making it faster to implement and less expensive to compute. To further enhance the model's training, they added to the datasets.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Lavric, Alexandru, et al. "Detecting keratoconus from corneal imaging data using machine learning." IEEE Access 8 (2020): 149113-149121.

Hidalgo, Irene Ruiz, et al. "Evaluation of a machine-learning classifier for keratoconus detection based on Scheimpflug tomography." Cornea 35.6 (2016): 827-832.

Shanthi, S., et al. "Machine learning approach for detection of keratoconus." IOP Conference Series: Materials Science and Engineering. Vol. 1055. No. 1. IOP Publishing, 2021.

Lavric, Alexandru, and Popa Valentin. "KeratoDetect: keratoconus detection algorithm using convolutional neural networks." Computational intelligence and neuroscience 2019 (2019).

Cohen, Eyal, et al. "Use of machine learning to achieve keratoconus detection skills of a corneal expert." International Ophthalmology 42.12 (2022): 3837-3847.

Cao, Ke, et al. "Accuracy of machine learning assisted detection of keratoconus: a systematic review and meta-analysis." Journal of Clinical Medicine 11.3 (2022): 478.

Yoo, Tae Keun, et al. "Adopting machine learning to automatically identify candidate patients for corneal refractive surgery." NPJ digital medicine 2.1 (2019): 59.

Brás, Nuno Miguel Ferreira Vivas. "Characterization and diagnostics of corneal transparency by OCT imaging and machine learning." (2023).

Panda, Saroj Kailash, and Nikhil Panjwani. "Cataract Detection Using Deep Learning." (2023)

Khan, Md Sajjad Mahmud, et al. "Cataract detection using convolutional neural network with VGG-19 model." 2021 IEEE World AI IoT Congress (AIIoT). IEEE, 2021

M. S. Junayed, A. N. M. Sakib, N. Anjum, M. B. Islam, and A. A. Jeny, EczemaNet: A deep CNN-based eczema diseases classi cation, in Proc. IEEE 4th Int. Conf. Image Process., Appl. Syst. (IPAS), Dec. 2020, pp. 174179.

J.-Y. Hung, C. Perera, K.-W. Chen, D. Myung, H.-K. Chiu, C.-S. Fuh, C.-R. Hsu, S.-L. Liao, and A. L. Kossler, A deep learning approach to identify blepharoptosis by convolutional neural networks, Int. J. Med. Informat., vol. 148, Apr. 2021, Art. no. 104402.

Ocular Disease Recognition, Dataset, https://www.kaggle.com/andrewmvd/ocular-disease recognition-odir5k.

J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, 'Ridge-based vessel segmentation in color images of the retina,' IEEE Trans. Med. Imag., vol. 23, no. 4, pp. 501–509, Apr. 2004.

A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, 'Robust vessel segmentation in fundus images,' Int. J. Biomed. Imag., vol. 2013, pp. 1–11, Dec. 2013.

Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, ORIGA-light: An online retinal fundus image database for glaucoma analysis and research, in Proc. Annu. Int. Conf. IEEE Eng. Med. Biol., Aug. 2010, pp. 30653068.

P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, and F. Meriaudeau, Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research, Data, vol. 3, no. 3, p. 25, Sep. 2018.

C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, S. Douma, and A. A. Argyros, FIRE: Fundus image registration dataset, Model. Artif. Intell. Ophthalmol., vol. 1, no. 4, pp. 1628, 2017.

J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE Trans. Med. Imag., vol. 23, no. 4, pp. 501509, Apr. 2004.