SEUSL

Sri Lankan Journal of Technology

Research

SLJoT
2022 Sp Issue 25 - 31
ISSN 2773-6970

A Multi-label Classifier with Transfer Learning Model for Detecting Multiple Ocular Diseases in Fundus Images

A.M. Aslam Sujah[1*], A. Fathima Sharfana[1] and A. Mohamed Sazni[2]

[1]Department of Information and Communication Technology, South Eastern University of Sri Lanka, Sri Lanka
[2]Department of Electronic and Telecommunication Engineering, University of Moratuwa, Sri Lanka

[*]Corresponding Author: ameersujah@seu.ac.lk ‖ ORCID: 0000-0002-2582-831X

Received: 01-10-2022 * Accepted: 10-11-2022 * Published Online: 30-11-2022

***Abstract*—**Early diagnosis of ophthalmic diseases can save vision. Fundus images provide better opportunities for screening various fundus diseases such as diabetic retinopathy, age-related macular degeneration, cataract, glaucoma, hypertension, etc. With its better learning abilities with a small dataset, the deep learning network is achieving remarkable success in ophthalmic disease diagnosis. These architectures can assist doctors in screening with higher reference values through better processing and analysis of fundus images. Fundus images may depict different ophthalmic diseases simultaneously. However, previous studies in this field focused on diagnosing any single disease. In this study, we focus on detecting multiple fundus diseases by analyzing fundus images with the use of deep neural network architecture. We present a model with a multi-label classifier for diagnosing multiple fundus diseases present in a fundus image. To address the inadequacy of data, we make use of transfer learning as the feature extractor. ResNet50 proves to be the best feature extractor for this task. The study utilizes the Ocular Disease Intelligent Recognition – 2019 dataset for testing this model. The accuracy of the proposed model outperforms the manual convolutional neural network, EfficientNet-v3, Inception-v3, and MobileNet-v2.

***Keywords*—**Ocular disease, Fundus imaging, Deep Learning, Transfer Learning

I. Introduction

Any abnormality of the macula or posterior pole was referred to as a macular disease, except for macular alterations brought on by diabetic retinopathy (DR). It included conditions including retinal holes, myopic maculopathy, and age-related macular degeneration. Macular illnesses, especially age-related macular degeneration, are a major cause of vision loss in senior persons, with a sharp rise in frequency after age 75, according to population-based research conducted in every region of the world and meta-analyses of these studies (Jonas *et al*., 2014). One of the most prevalent illnesses populations in the world today is people with diabetes. The most frequent side effect of diabetes is DR. Early in the course of the illness, there are no unusual symptoms, but it will eventually cause blindness. It is one of the four main disorders that cause blindness (Leasher *et al*., 2016). There is proof that patients with DR who receive therapy and early identification have better visual results at a lower cost (Walton *et al*., 2016) (Fong *et al*., 1999) (Bloomgarden, 2007) (Jones and Edwards, 2010). If not treated early, it will result in permanent visual impairment (Bressler, 2004) (Ye *et al*., 2014).

Fundus imaging is the method of using reflected light to create a two-dimensional (2D) picture of the retinal tissues that are projected in three dimensions onto the imaging plane. Therefore, a process that creates a 2D image and whose image intensities reflect the amount of reflected light is called fundus imaging (Abràmoff and Kay, 2012). The most apparent use of retinal screening is the identification of retinal diseases such as age-related macular degeneration, glaucoma, retinopathy of prematurity, and most significantly the early detection of DR which involves imaging the patient's retinas through telemedicine.

It has been demonstrated that using fundus images for DR screening would increase the frequency of retinal screening and make it easier for medical professionals to assess patients with potentially blinding diseases (Paz *et al*., 2006). The traditional method of manual picture interpretation in screening for diabetic tele-retinal disease may be replaced by one that uses computer algorithms (Walton *et al*., 2016). A lot of patients with fundus illnesses will benefit from the integration of artificial intelligence and ophthalmology medical care in terms of meeting their practical demands.

Over recent years, Deep learning algorithms have improved object recognition, segmentation, and classification in medical images (Litjens *et al*., 2017). As opposed to the traditional feature extraction methods, deep learning

techniques do not rely on the manual process of defining features by the experts using their domain knowledge. Deep Convolutional Neural Networks (CNN) can learn features to distinguish images. However, a common problem in deep learning networks is due to the demand for large-size data, and huge computational time for training.

Deep CNN models such as EfficientNetv2 (Tan and Le, 2021), MobileNetv2 (Howard *et al*., 2017), Inception v3 (Szegedy *et al*., 2016), and ResNet50(He *et al*., 2016) are the deeper networks that can extract more robust and complex features. They are pre-trained on the largest labelled dataset such as ImageNet (Russakovsky *et al*., 2015). The low-level features of images are learnt by the initial filters of these networks, and the top layers are trained for specific tasks. Transfer learning allows us to transfer the knowledge of a model which was trained on a different dataset to our problem domain instead of learning something new from scratch (Shaha and Pawar, 2018; Hussain, Bird and Faria, 2019; Krishna and Kalluri, 2019). The top layers of the network will be fine-tuned for a specific task with a smaller dataset will deliver the best accuracy (Shaha and Pawar, 2018).

A deep convolutional neural network was trained by the Google research team utilizing a retrospective development data set of 128 175 retinal pictures that had been rated three to seven times for diabetic retinopathy (DR), diabetic macular oedema, and image gradeability by a by a committee of 54 US-licensed ophthalmologists and senior people in ophthalmology. The performance of the model is comparable to professional ophthalmologists (Gulshan *et al*., 2016).

(Qiao *et al*., 2017) in 2017 used deep learning architecture for the diagnosis of cataracts present in fundus images. They conclude that the feature extracted from the deep learning network, compared to other machine learning models performs better. Based on the time-invariant feature cup-to-disc ratio and anisotropic dual-tree complex wavelet transform features, (Kausu *et al*., 2018) offered a new approach to glaucoma detection. In comparison to previous efforts, the findings demonstrate that the suggested technique attained an accuracy rate of 97.67% with 98% sensitivity by employing a multilayer perceptron model. (Christopher *et al*., 2018) evaluated identifying glaucomatous optic neuropathy (GON) in fundus photos using deep learning architectures. They employed multiple deep learning architectures and transfer learning models to categorize a huge collection of fundus images as GON or healthy. Such transfer learning models are VGG16, Inception, and ResNet50. They also trained their models to diagnose moderate-to-severe GON and achieved a higher sensitivity of 88% and specificity of 95%. Their results are evident that deep learning-based assessment can be used in clinical diagnostics of ophthalmic disease, and also the transfer learning models outperform their native counterparts on the testing data. (Govindaiah *et al*., 2018) The efficiency of the deep convolutional neural network in identifying age-related macular degeneration (AMD) was investigated by Govindaiah and his team and showed that the neural network which is explicitly trained with a larger fundus image dataset could produce much better results than that of pre-trained transfer learning networks. (Dai *et al*., 2020) trained a small convolutional neural network with a deep learning technique called Gradient-weighted Class Activation Mapping (Grad-CAM) to detect hypertension. The model identified the red patchy areas in arterial/venous bifurcations which is a clear sign of blood pressure elevation.

(Zhang *et al*., 2019) proposed a system called DeepDR to directly detect DR from fundus images via transfer and ensemble learning and achieved the best sensitivity of 97.5% and specificity of 97.7%. The system performed well in comparison with human evaluation metrics. (Islam *et al*., 2019) proposed a CNN-based model to detect eight types of ocular disorders and achieved a validation accuracy of 87.6% which is a pretty decent accuracy compared to the other works. However, the work of (Islam *et al*., 2019) classifies eight classes of ocular diseases which are novel and the reason for the decent accuracy achieved in this work. (Wang *et al*., 2020) solved the task of multi-label fundus images by adopting different strategies. Pre-trained EfficientNet was fine-tuned with fundus images and two weak classifiers were integrated to get the final results. They achieved reasonably best results for an eight-classification problem by fine-tuning the pre-trained model with a very limited training dataset.

Unfortunately, CNN still has difficulties when used for the study of fundus imaging. First of all, multi-label fundus image categorization is a more common and practical issue since a realistic fundus picture in the real world is likely to have different fundus illnesses. Second, it may be difficult to locate sufficient examples of accurate fundus images, particularly those pertaining to some unusual fundus illnesses. Third, it is challenging to train a single model to successfully achieve high disease diagnosis accuracy due to the limited fundus imaging data and unavoidable image noise (Wang *et al*., 2020).

In this study, to address the issues mentioned above, we focused on developing a deep learning framework for the multi-label classification of fundus disease identification using a improved transfer learning model as a feature extracting algorithm and multi-label classifier for disease categorization. The dataset consists of fundus images for more than eight types of eye diseases with a small amount of photos in each category. Therefore, transfer learning models are utilized to improve the accuracy of feature extraction with a limited number of images. We compare and analyze the performance of CNN with transfer learning models to recommend the best model for the future use of deep learning in ophthalmic disease diagnosis.

II. Methodology

Building a framework for the automatic diagnosis of multi-label fundus disorders is the goal of this project, which will be accomplished through the design of a corresponding model based on analyzing several transfer learning models. As of multiple components, each of which is a two-

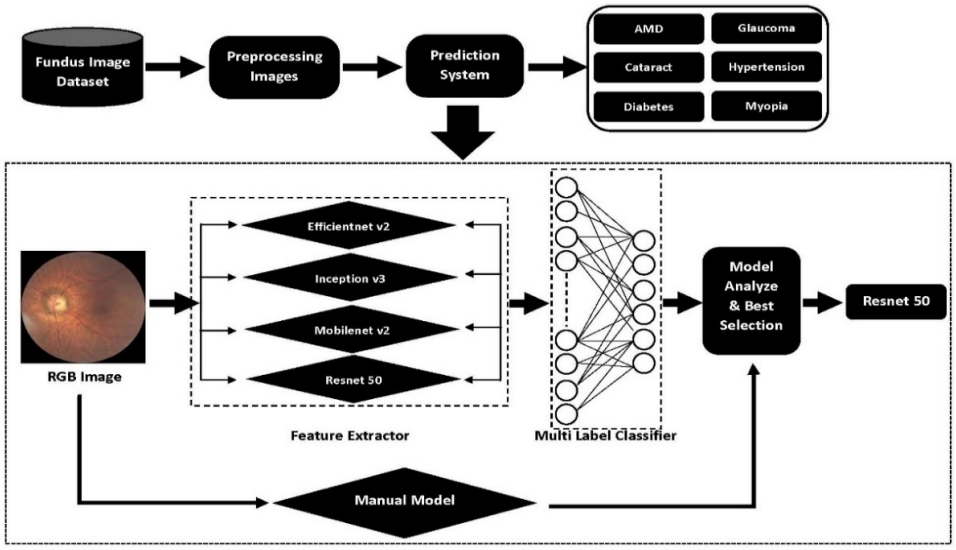

Figure 1: Architecture of the automatic recognition of multi-label fundus diseases

part neural network. The first part is a feature extractor, which includes a pretrained transfer learning model without the upper layer. A multi-label classifier in the second part makes predictions based on the preceding characteristics. This neural network was created from scratch and trained to fulfill certain requirements. Following this, the exact implementation phases of the framework will be presented in depth.

A. Dataset

The dataset originates from Peking University's "International Competition on Ocular Disease Intelligent Recognition". The dataset consists of authentic patient information acquired by Shanggong Medical Technology Co., Ltd. The training set is a structured ophthalmic database consisting of 3008 patients with color fundus images from the left and right eyes, as well as diagnostic keywords from the doctor. The testing set consists of 736 color fundus photographs of patients. Varied cameras on the market capture fundus images, resulting in various image resolutions. Figure 2 illustrates the six labelled patient dataset classes utilized for model development, including age-related macular degeneration (AMD), diabetes, glaucoma, cataract, hypertension, and pathological myopia. One patient may contain a single label or numerous labels.

B. Preprocessing

Before resizing the images to 300 × 300 pixels, the original data is separated into a training set and a validation set with a ratio of 8:2. As depicted in Figure 3, we have utilized data augmentation techniques, such as Random Flip (Horizontal and Vertical), Random Rotation, Random Zoom, etc., to increase the amount of the data set while preserving the original image's essential characteristics but not a complete copy. Also, rescale input image's RGB values in the range [0, 255] to the range [0, 1].

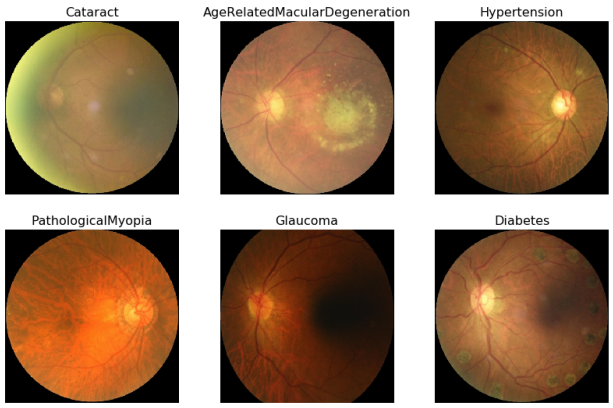

Figure 2: Sample Image for Each Disease Class

C. Feature Extraction

To train a robust feature extraction, we must select an outstanding CNN to extract features from every fundus source image and construct a compact feature vector as a preliminary step. CNN's present development is mostly driven by fixed development resources. If there is enough processing power, the network will continue to grow. We proposed a fundus image classification approach with multiple labels utilizing ResNet50 as the feature extraction framework because of the excellent efficiency of ResNet50 (He *et al*., 2016). In addition, for comparison research, we utilized various CNN transfer learning approaches as feature extractors. Examples are EfficientNetv2 (Tan and Le, 2021), MobileNetv2 (Howard *et al*., 2017), Inception v3 (Szegedy *et al*., 2016), and ResNet50.

We remove all the upper CNN layers because of connecting to the next classifier. We studied two distinct initialization procedures in this paper. To train the model, we randomly initialize the network parameters. Second, we pre-train the network's parameters for exchanging data across domains and assignments. Previously, the CNN network was trained on ResNet50 is utilized as a feature extraction framework, parameters are initialized, and the complete network is then retrained with data from the new domain.

D. Classifier

The multi-label classifier uses the feature extractor's output as its input, which uses a multi-label classifier based on a customized neural network with six output neurons by the specific data distribution. The classifier uses the same network for different feature extractors. It helps to analyze the different transfer learning models and find out the best one with the same configurations. Simultaneously, we built a CNN network manually without any transfer learning strategies and evaluated the effectiveness and prediction accuracy. Table I shows the full organizational structure of the manual CNN network.

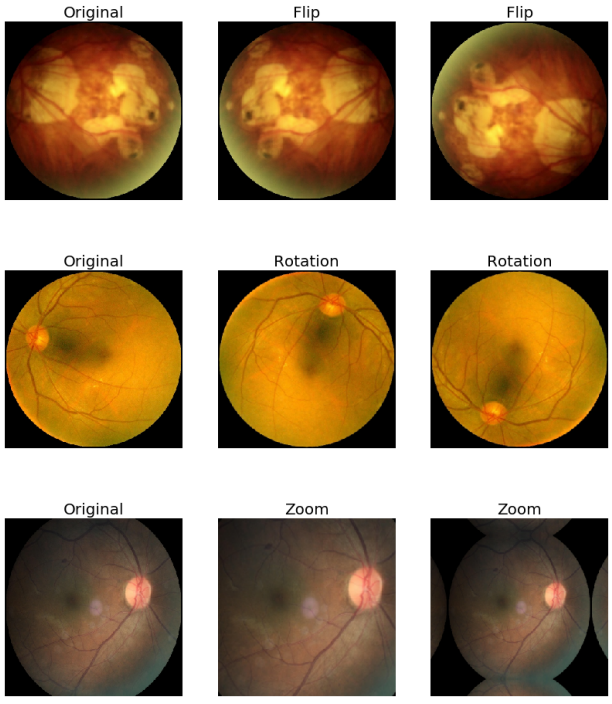

Figure 3: Image Augmentation

III. Results and Discussion

The percentage of samples that are correctly classified is known as accuracy, and it is the most basic evaluation metric for classification tasks. This study evaluates the accuracy of different transfer learning models with the traditional CNN model and selects the best model with an efficient feature extractor and multi-label classifier. Finally, it applies the best model to automatically diagnose multi-label fundus disorders.

Table I: Structure of the manual CNN network

Layer (type)	Output Layer Shape	Param #
Sequential Layer_1	(32, 300, 300, 3)	0
Sequential Layer_2	(32, 300, 300, 3)	0
Convolutional Layer_1	(32, 298, 298, 32)	896
Max Pooling Layer_1	(32, 149, 149, 32)	0
Convolutional Layer_2	(32, 147, 147, 64)	18496
Max Pooling Layer_2	(32, 73, 73, 64)	0
Convolutional Layer_3	(32, 71, 71, 64)	36928
Max Pooling Layer_3	(32, 35, 35, 64)	0
Convolutional Layer_4	(32, 33, 33, 64)	36928
Max Pooling Layer_4	(32, 16, 16, 64)	0
Convolutional Layer_5	(32, 14, 14, 64)	36928
Max Pooling Layer_5	(32, 7, 7, 64)	0
Convolutional Layer_6	(32, 5, 5, 64)	36928
Max Pooling Layer_6	(32, 2, 2, 64)	0
Flatten Layer	(32, 256)	0
Dense Layer_1	(32, 64)	16448
Dense Layer_2	(32, 6)	390

Table II: Configuration Parameters

Configuration	Value
Batch Size	32
Epoch	20
Augmentation (Random Flip)	Horizontal and Vertical
Augmentation (Random Rotation)	0.2
Augmentation (Random Zoom)	0.2
Batch Normalization	True
Optimization Function	Adam

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

The dataset consists of 80% training data, 20% verification data, and an extra 300 test datasets. Table II shows the hyperparameter configuration. Multiple sets of comparative tests were designed to confirm the efficacy of our methodology.

A. Comparison Experiment of Manual and Transfer Learning Model

The outcomes of training CNN from the bottom to the top are unsatisfactory, as seen in Figure 4. Because the model's results on the validation dataset are much better than those on the testing dataset due to the restrictions of the dataset,

it appears that the training procedure may have been over fitting. We attempted to introduce batch normalization to address this issue. A parameterized learnable network layer is present at each tier of the network input.

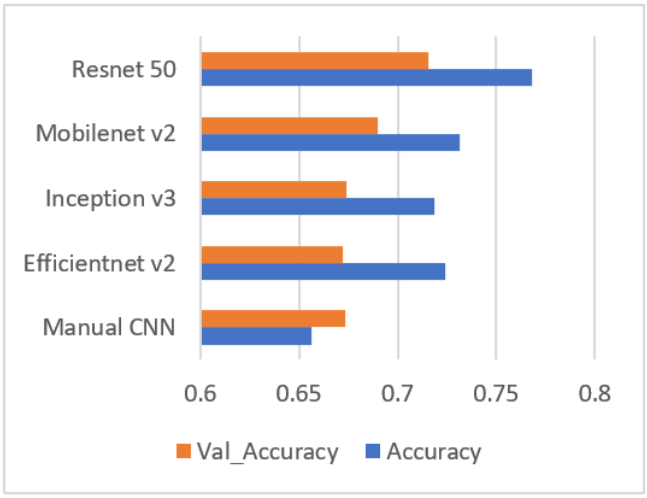

Figure 4: Accuracy and Validation Accuracy Comparison

B. Comparison Experiment of Manual CNN and Resnet Model with Training Dataset

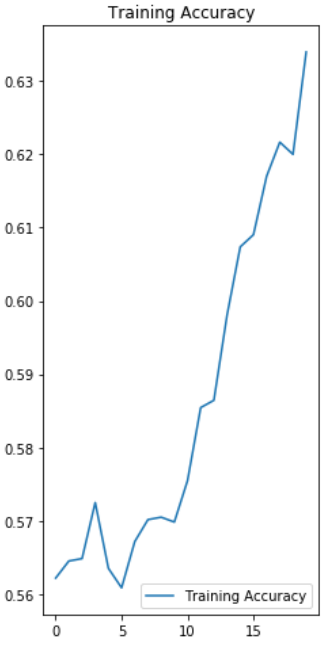

Figure 5: Training Accuracy of Manual CNN

Figure 5(a) represents the flow of training accuracy of manual CNN architecture where Figure 5(b) shows the flow of training accuracy of Resnet 50 transfer learning architecture concerning a number of epochs (20). Analyzing the charts helps to get an idea of the efficiency and fast performance of Resnet 50 than the manual CNN.

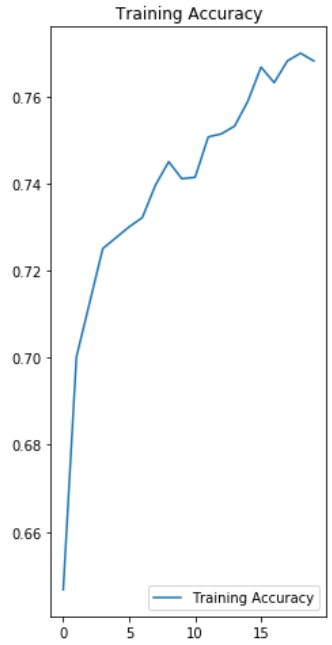

Figure 6: Training Accuracy of Resnet 50

In this research, we employ a range of different training methodologies to address the multi-label classification task of fundus images. The optimum training approach is demonstrated by the results as follows: We begin by resizing the dataset to 224 x 224 and combine the parameters of the transferred neural network's pre-trained Resnet50 transfer learning model. Finally, we construct the finished product by integrating a multi-label classifier. The model can then be tested in real time.

IV. Conclusion

In this study, we have developed a framework for the automatic diagnosis of various ophthalmic diseases by utilizing deep learning models. The framework achieved good results on the ODIR-2019 dataset for a multi-class problem of fundus diseases present in fundus images. Even though deep learning models can be recommended for clinical practices to diagnose multiple fundus.

diseases, there is a high demand for adequate data to train the deep learning models to improve the results to the best. For the variety of fundus diseases, the deep learning model requires more images for better learning.

References

Abràmoff, M. and Kay, C. N. (2012) 'Image Processing', Retina Fifth Edition, 1, pp. 151–176. doi: 10.1016/B978-1-4557-0737-9.00006-0.

Bloomgarden, Z. T. (2007) 'Diabetic retinopathy and diabetic neuropathy', *Diabetes Care, 30*(3), pp. 760–765. doi: 10.2337/dc07-zb03.

Bressler, N. M. (2004) 'From the archives journals', *Journal of the American Medical Association, 291*(15), pp. 1900–1901.

Christopher, M. *et al*. (2018) 'Performance of Deep Learning Architectures and Transfer Learning for Detecting Glaucomatous Optic Neuropathy in Fundus Photographs', *Scientific Reports, 8*(1), pp. 1–13. doi: 10.1038/s41598-018-35044-9.

Dai, G. *et al*. (2020) 'Exploring the effect of hypertension on retinal microvasculature using deep learning on East Asian population', *PLoS ONE, 15*(3), pp. 1–13. doi: 10.1371/journal.pone.0230111.

Fong, D. S. *et al*. (1999) 'Causes of severe visual loss in the Early Treatment Diabetic Retinopathy Study: ETDRS report no. 24', American *Journal of Ophthalmology, 127*(2), pp. 137–141. doi: 10.1016/S0002-9394(98)00309-2.

Govindaiah, A. *et al*. (2018) 'Deep convolutional neural network based screening and assessment of age-related macular degeneration from fundus images', *Proceedings - International Symposium on Biomedical Imaging*, 2018-April(Isbi), pp. 1525–1528. doi: 10.1109/ISBI.2018.8363863.

Gulshan, V. *et al*. (2016) 'Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs', *JAMA - Journal of the American Medical Association, 316*(22), pp. 2402–2410. doi: 10.1001/jama.2016.17216.

He, K. *et al*. (2016) 'Deep residual learning for image recognition', *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, pp. 770–778. doi: 10.1109/CVPR.2016.90.

Howard, A. G. *et al*. (2017) 'MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications'. Available at: http://arxiv.org/abs/1704.04861.

Hussain, M., Bird, J. J. and Faria, D. R. (2019) 'A study on CNN transfer learning for image classification', *Advances in Intelligent Systems and Computing, 840*(June), pp. 191–202. doi: 10.1007/978-3-319-97982-$3_1$6.

Islam, M. T. *et al*. (2019) 'Source and Camera Independent Ophthalmic Disease Recognition from Fundus Image Using Neural Network', *2019 IEEE International Conference on Signal Processing, Information, Communication and Systems, SPICSCON 2019*, (February 2020), pp. 59–63. doi: 10.1109/SPICSCON48833.2019.9065162.

Jonas, J. B. *et al*. (2014) 'Visual impairment and blindness due to macular diseases globally: A systematic review and meta-analysis', *American Journal of Ophthalmology, 158*(4), pp. 808–815. doi: 10.1016/j.ajo.2014.06.012.

Jones, S. and Edwards, R. T. (2010) 'Diabetic retinopathy screening: A systematic review of the economic evidence', *Diabetic Medicine, 27*(3), pp. 249–256. doi: 10.1111/j.1464-5491.2009.02870.x.

Kausu, T. R. *et al*. (2018) 'Combination of clinical and multiresolution features for glaucoma detection and its classification using fundus images', *Biocybernetics and Biomedical Engineering, 38*(2), pp. 329–341. doi: 10.1016/j.bbe.2018.02.003.

Krishna, S. T. and Kalluri, H. K. (2019) 'Deep learning and transfer learning approaches for image classification', *International Journal of Recent Technology and Engineering, 7*(5), pp. 427–432.

Leasher, J. L. *et al*. (2016) 'Global estimates on the number of people blind or visually impaired by diabetic retinopathy: A meta-analysis from 1990 to 2010', *Diabetes Care, 39*(9), pp. 1643–1649. doi: 10.2337/dc15-2171

Litjens, G. *et al*. (2017) 'A survey on deep learning in medical image analysis', *Medical Image Analysis, 42*, pp. 60–88. doi: 10.1016/j.media.2017.07.005.

Paz, S. H. *et al*. (2006) 'Noncompliance with Vision Care Guidelines in Latinos with Type 2 Diabetes Mellitus. *The Los Angeles Latino Eye Study', Ophthalmology, 113*(8), pp. 1372–1377. doi: 10.1016/j.ophtha.2006.04.018.

Qiao, Z. *et al*. (2017) 'Classification of Cataract Fundus Images', (c), pp. 2–6.

Russakovsky, O. *et al*. (2015) 'ImageNet Large Scale Visual Recognition Challenge', International Journal of Computer Vision, 115(3), pp. 211–252. doi: Russakovsky, O. *et al*. (2015) 'ImageNet Large Scale Visual Recognition Challenge', *International Journal of Computer Vision, 115*(3), pp. 211–252. doi: 10.1007/s11263-015-0816-y.

Shaha, M. and Pawar, M. (2018) 'Transfer Learning for Image Classification', *Proceedings of the 2nd International Conference on Electronics, Communication and Aerospace Technology, ICECA 2018*, (Iceca), pp. 656–660. doi: 10.1109/ICECA.2018.8474802.

Szegedy, C. *et al*. (2016) 'Rethinking the Inception Architecture for Computer Vision', *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.

Tan, M. and Le, Q. V. (2021) 'EfficientNetV2: Smaller Models and Faster Training'. Available at: http://arxiv.org/abs/2104.00298.

Walton, O. B. *et al*. (2016) 'Evaluation of automated telereti-nal screening program for diabetic retinopathy', *JAMA Ophthalmology, 134*(2), pp. 204–209. doi: 10.1001/ja-maophthalmol.2015.5083.

Wang, J. *et al*. (2020) 'Multi-Label Classification of Fun-dus Images with EfficientNet', *IEEE Access, 8*, pp. 212499–212508. doi: 10.1109/ACCESS.2020.3040275.

Ye, H. *et al*. (2014) 'Prevalence of age-related macular degeneration in an elderly urban Chinese population in China: The Jiangning eye study', *Investigative Ophthal-mology and Visual Science, 55*(10), pp. 6374–6380. doi: 10.1167/iovs.14-14899.

Zhang, W. *et al*. (2019) 'Automated identification and grading system of diabetic retinopathy using deep neural net-works', *Knowledge-Based Systems, 175*(May 2020), pp. 12–25. doi: 10.1016/j.knosys.2019.03.016.