Deep learning for identifying and classifying retinal diseases

Mohamed BERRIMI
Department of computing sciences
University of Ferhat Abbas 1
Setif, Algeria
mohamedberrimi0@gmail.com

Abdelouaheb MOUSSAOUI
Department of computing sciences
University of Ferhat Abbas 1
Setif, Algeria
moussaoui.abdel@gmail.com

***Abstract*—Vision and eye health are one of the most crucial things in human life, it needs to be preserved to maintain the life of the individuals. Eye diseases such as CNV, DRUSEN, AMD, DME are mainly caused due to the damages of the retina, and since the retina is damaged and discovered at late stages, there is almost no chance to reverse vision and cure it, which means that the patient will lose the power of vision partially and maybe entirely. Optical Coherence Tomography is an advanced scanning device that can perform non-invasive cross-sectional imaging of internal structures in biological tissues by measuring their optical reflections. which will help the ophthalmologists to take a clear look on the back of the eye and determine at early stages the damage caused to the retina, macula, and optic nerve. The aim of this study is to propose a novel classification model based on deep learning and transfer learning to automatically classify the different retinal diseases using retinal images obtained from Optical Coherence Tomography (OCT) device. We propose a deep CNN architecture and compared the obtained results with pretrained models such as Inception V3 and VGG-16, our proposed CNN architecture gave an accuracy of 98.5% and Inception V3 model gave an accuracy up to 99.27% on the test set while VGG-16 gave only 53% we modified VGG-16 architecture by adding more convolution layers and regularization terms to obtain a result up to 93.5%.**

***Index Terms*—Convolutional neural networks, Optical coherence Tomography, Transfer learning, medical imaging,image classification**

I. Introduction

OCT stands for Optical Coherence Tomography is an advanced scanning device that can perform noninvasive cross-sectional imaging of internal structures in biological tissues by measuring their optical reflections. [1], which will help the ophthalmologists to take a clear look on the back of the eye and determine at early stages the damage caused to to the retina, macula, and optic nerve. Diagnosing or predicting these pathologies at early stages can increase the chance of curing the patients and restore vision ability, it has been also proved [5] that eye diseases dont just affect the retina and the patients vision, but it has also a the relation between heart diseases, hypertension, which means that predicting and diagnosing eye diseases at early stages can also save the health of the patient's heart.

Four main types of eye diseases are treated in this study:

- **DME** Diabetic Macular Edema is a type of eye disease due to the damage of blood vessels in the retina. When left untreated, DME will cause the build-up of liquid in the macula further leading to a swollen area on the retinal layer and consequently irreversible eye blindness.
- **AMD** Age-elated Macular Degeneration is a damage affects the macula (small are at the center of the retina) , leads to center-blind, and its a blinding disease with no cure at present. This disease doesnt just affect the vision of the patient, but it can also cause heart diseases and hyper-tensions, according to [5]. The number of people living with macular degeneration is expected to reach 196 million worldwide by 2020 and increase to 288 million by 2040 [6]
- **DRUSEN** One the first signs of AMD pathology is called DRUSEN. Drusen are yellow deposits under the retina, they are not symptoms of eye diseases, but the the appearance of large number of them can lead to AMD and vision loss. [16] Ophthalmologists these days uses OCT imaging scanning to detect DRUSEN and define their types if they are serious and can lead to AMD, or if they can disappear [9], which gives them the opportunity to make prior decisions.
- **CNV** Choroidal neovascularization is a very common vision-threatening disease that leads to vision loss. involves the growth of new blood vessels that originate from the choroid through a break in the Bruch membrane into the subretinal pigment epithelium or subretinal space. [7] [15] any damages the Bruch membrane can be complicated by CNV.

Related works

Machine learning is nowadays used in different fields of medical imaging, computer-aided diagnosis [19], image segmentation [18] and image-guided therapy, this means that there are multiple areas in medicine, where machine learning methods can be applied to improve patients' health care. Different studies have been proposed to classify OCT images. Md Akter Hussain et al. [31] proposed a classification model based on random forest classifier, with 15-fold cross-validation tests, to detect (AMD) or Diabetic Macular Edema (DME) using retinal features from Spectral

978-1-7281-5467-1/20/$31.00 ©2020 IEEE.

Domain Optical Coherence Tomography (SD-OCT) images. the dataset contained 251 (59 normal, 177 AMD and 15 DME) images, and obtained the accuracy of 95 % while testing and 96 %in train set.

Venhuizen et al. [42] proposed an architecture to detect AMD diseases using Random Forest with a maximum of 100 trees is used for the classifier. The method achieved an AUC of 0.984 with a dataset of 384 (269 AMD, 115 control) OCT volumes. Shenghua He et al. [36] have proposed a very deep CNN network architecture contained 9 CONV blocks and 2 FC layers. 2 max-pooling layers were placed after the 3rd and the 6th CONV blocks, respectively. 1 global pooling layer was placed after the 9th CONV block. The spatial support of the filter in each of the CONV layers was set as 3 x 3 pixels. The number of the filters in the first three Conv layers was set to 32.In order to compensate for the information loss caused by max pooling, the number of filters in 2nd three CONV layers and 3rd three CONV layers were set to 64 and 128, respectively. Two FC layers followed the global pooling layer. The first FC layer included 512 neurons and the second one included 5 neurons. One dropout was set between these two FC layers with a dropout ratio of 0.5 to further avoid overfitting. A softmax layer was placed at the end of the classifier.

The experiments on 269 OCT images showed that the average prediction accuracy of the CNN-based method was 0.866. The test set accuracy wasnt mentioned in their paper, we believe that such very deep architecture on a few data (269 images) can lead to overfitting. one trick could have been done is to use data augmentation [35] to get more data for training.

Muhammad Awais et al. [30] have worked on the classification of SD-OCT images using VGG-16 pretrained model, to detect DME diseases on a dataset consisting of 32 OCT volumes (16 DME and 16 normal cases). Each volume contains 128 Bscans with resolution of 1,024 px 512 px. They did many experiments on the dataset combining CNN with other classifiers (kNN and Decision trees). the best configuration was obtained 93.5 %. One of their configurations gave 100 % accuracy, by setting the k = 1 in kNN classifier, and 2 FCL1 , hence this leads also to overfitting cause of the bad choice of parameters. Cecilia S. Lee et al. [33] published a paper on classifying AMD diseases using a modified version of the VGG16 convolutional neural network on a total of 80,839 images (41,074 from AMD, 39,765 from normal) were used for training and 20,163 images (11,616 from AMD, 8,547 from normal) were used for validation. The training was then performed using multiple iterations each with a batch size of 100 images with a starting learning rate of 0.001 with stochastic gradient descent optimization. At each iteration, the loss of the model was recorded, and at every 500 iterations, the performance of the neural network was assessed using cross-validation with the validation set. The training was stopped when the loss of the model decreased and the accuracy of the validation set decreased. Accuracy in the area under the ROC curve (AUROC) was 92.77 %.

This is a very interesting paper indeed, due to the use of pretrained model and regularization techniques to speed the learning phase and avoid overfitting.
Another paper on classifying OCT images using deep learning methods by Parmita M. et al. [34] for multilabel multiclass classification for OCT retinal images to diagnose patients who may exhibit multiple pathologies, the dataset consists of 36,150 images, applied data augmentation [35] and modified version of Inception Resnet V2 pre-trained model by removing the output activation layer softmax and replaced it with sigmoid. The goal of their study was to compare transfer learning with CNN from scratch the resulted work is accurate and exact match for transfer learning was 74.5% and 30.14% compared to 85.23% and 64.3% for direct learning. A paper was published recently that used the dataset that we are using in this study published by Kuntoro Adi Nugroho [32] comparison of Handcrafted and Deep Neural Network Feature Extraction for Classifying Optical Coherence Tomography (OCT) Images the deep neural network-based methods outperformed the handcrafted feature with 88% and 89% accuracy for DenseNet and ResNet compared to 50% and 42% for HOG and LBP respectively. The issue with this study is that they have split the dataset into 50% for training and 50% for validation. In this study, we are proposing different split of the data and different approaches

II. PROPOSED ARCHITECTURES

We have proposed two convolutional neural networks CNN [40] architectures and fine tuned two pretrained models and a modified version of VGG-16, all have been trained on google Colab GPU [43], which is Tesla K8 GPU, the architectures details are as follows :

1) **First CNN architecture** the first CNN architecture is composed of 3 Convolution layers each conv layer is followed by Max pooling layer. then followed by one dense layer and ouptut layer composed of 4 neurons corresponding to our number of classes.

Parameter	Value
Dropout rate	/
Batch size	80
Number of epochs	15
optimizer	Adam
Number of learnable parameters	399,396
Early stopping patience factor	10
Learning rate	0.001
One epoch training duration	335s

TABLE I
FIRST CNN ARCHITECTURE HYPER-PARAMETERS

2) **Second CNN architecture** We added to the first CNN architecture one dropout layer [39] and one Batch normalization layers [**?**]. The hyperparameters to compile and run the model are illustrated in table II.

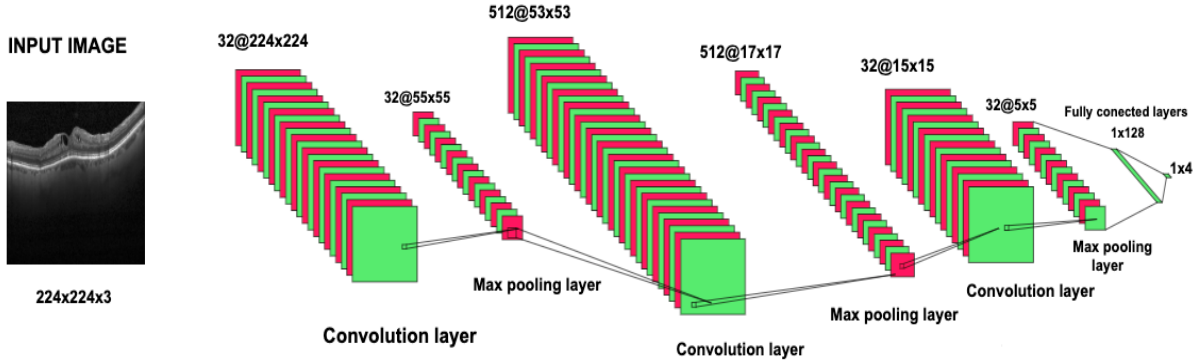

Fig. 1. Our proposed CNN architecture.

Parameter	Value
Dropout rate	0.1
Batch size	80
Number of epochs	15
optimizer	Adam
Number of learnable parameters	399,460
Early stopping patience factor	10
Learning rate	0.001
One epoch training duration	335s

TABLE II
SECOND CNN ARCHITECTURE HYPER-PARAMETERS.

The number of learnable parameters decreased comparing the first table, due to the adding of regularization parameters (Dropout and BatchNormalization)

Architecture interpretation : The model is sequential which allows us to create the model layer-by-layer. The architecture consists of convolutional layers, max-pooling layers, dropout layers, and fully connected layers.

The first layer is a convolutional layer with 32 filters each of size 3 x 3. We are also required to specify the input shape in the first layer, which is 224 x 224 x 3 in our case.

We will be using the Rectified linear unit (**ReLU**) activation function for all the layers except the final output layer. ReLU is the most common choice for activation function in the hidden layers and has shown to work pretty well.

The second layer is a pooling layer. The pooling layers are used to reduce the dimension. Max Pooling with a 4x4 window only considers the maximum value in the 4x4 window. A dropout layer with dropout [39] rate of 0.1 means 10% of the neurons will be turned off randomly. This helps prevent overfitting by making all the neurons learn something about the data and not rely on just a few neurons. Randomly dropping neurons during training means other neurons will have to do the work of the turned-off neurons, thus generalizing better and prevent overfitting The third layer is again a convolutional layer of 512 filters each of size 3 x 3 followed by another max-pooling layer of 3x3 window. Usually, the number of filters in the convolutional layer grows after each layer. The first layers with a lower number of filters learn simple features of the images whereas the deeper layers learn more complex features. The next layers are again convolutional layers with 32 filter size. followed by a Batch normalization layer. We need to flatten the 3D feature map output from the convolutional layer to 1D feature vectors before adding in the fully connected layers. This is where the flattening layer comes in. The following dense layer (fully connected layer) has 128 neurons. The final output layer is another dense layer that has a number of neurons equal to the number of classes. The activation function is softmax because it is a multi-class classification problem.

3) **Inception V3 Model** We used Inception V3 [45] pre-trained weights to classify our images, we removed the first three top layers and changing the output layer to have four neurons corresponding to our number of classes, the pooling operation was set to average and the optimizer is rmsprop.
4) **VGG-16 Model** We used VGG-16 [44] pre-trained weights to classify our images same as we did with Inception V3 model, we removed the first three top layers and changing the output layer to have four neurons corresponding to our number of classes, the pooling operation was set to average.
5) **Modified version of VGG-16 Model** We added

two more Convolution layers in the top of VGG16 pretrained model, we changed the Flatten layer with *Global Average Pooling* [37], and the ouptut layer to have 4 neurons corresponding to four classes.

Compiling and training the models

All the models are compiled with categorical cross-entropy loss function and the Adam as an optimization algorithm, except the InceptionV3 model where we tried both Adam and rmsprop, the last one gave better results. The accuracy metric is used to evaluate the model. Training the model using a GPU speed up the training process, We set the number of epochs to 15 and use a regularization method Early Stopping [38], In our case, we tell EarlyStopping to monitor validation-accuracy and if it does not improve for 10 epochs continuously, stop the training process. The model checkpoint is used to save the model.

The monitor parameter allows us to set a metric that we want to keep an eye on. In our case, we only save the model when the validation accuracy is the max. We save the best model to be used later to make predictions and thus evaluate the models performance

III. DATASET OVERVIEW AND PREPROCESSING

We have used the OCT images dataset from [27], the dataset is organized into 3 folders (train, test, val) and contains subfolders for each image category (NORMAL,CNV,DME,DRUSEN). There are **84,495** images (JPEG) and 4 categories (NORMAL,CNV,DME,DRUSEN).

Our dataset was taken from different research labs, what makes the sizes of the images varies **(496, 768, 3), (496, 1024, 3), (496, 512, 3), (496, 1536, 3), (512, 512, 3)** the first two values refers to the width and height of the image, and the thrid one refers to the image channels , meaning in this case that the images are in RGB [1].

- **Data Rescaling**: Importing the images with the original sizes will lead to use big part of hardware resources and the time of processing the images will highly increase, we reduced the image sizes to 224 X 224 pixels, same as Imagenet [41] dataset images sizes because we aim to do experiment on pretrained models.

- **Data Resampling**: The dataset is splitted into 3 folders as explained in the previous section, with only 8 images per class for validation and 242 images for test and the rest for training. as detailed in table This split is not efficient and can lead to extream overfitting. We made another split of 80% for training , 20% for validation (table 3.2) and after constructing our model we tested our model on 968 images.

IV. RESULTS AND DISCUSSIONS

After compiling our first CNN architecture we obtained an accuracy of **97.7%** in validation set and **95.6 %** when testing on our test set data.

[1]Red Green Blue channels

	Training	Testing	Validation	
CNV	37205	242	8	**37455**
DME	11348	242	8	**11598**
DRUSEN	8616	242	8	**8866**
NORMAL	26315	242	8	**26565**
	83484	**968**	**32**	**84484**

TABLE III
DATASET ORIGINAL SPLIT

	Training	Testing	Validation	
CNV	33,039	242	4,174	**37455**
DME	7,182	242	4,174	**11598**
DRUSEN	4,450	242	4,174	**8866**
NORMAL	22,149	242	4,174	**26565**
	66,788	**968**	**16,696**	**84484**

TABLE IV
DATASET DISTRIBUTION AFTER RESAMPLING

After the second experiment of our CNN architecture (after adding the regularization terms)we obtained an accuracy up to **98.75%** and accuracy of **98%** while testing. meaning when we added the two regularization terms the accuracy did enhance for both validation and testing as shown in figure 6.

Our first experiment on the VGG16 model, the training time was 16 minutes per epoch, we obtained an accuracy of 53.75% on validation and in testing the model. the accuracy in the validation part didn't improve the 10 first epochs until early stopping algorithm stopped the training process.

After our second experiment on VGG-16 model where we added more Convolution layers are Batch normalization layers at the top of the model, the training time did enhance and the accuracy raised to **93.5** in validation compared to the previous experiment where VGG-16 gave only **53** %.

The results of our experiment on the InceptionV3 model were exciting, after 15 epochs of training and validation, we obtained an accuracy up to **99.03%** in training and **100%** in the validation part, and accuracy of **99.27 %** in 986 testing images as illustrated in figure 9. This experiment outperformed all the other results.

The plot below explains the change of accuracy value during the training epochs also the loss value for both training and validation is given. The red line is showing how well the model is learning by each epoch(full iteration on the dataset), the blue line in the other side is the learning curve calculated from a hold-out validation dataset that gives an idea of how well the model is generalizing, in this case, the model is doing a good job since theres no big difference between the training accuracy and the validation.

CNN - accuracy ON TEST SET : 0.9865702479338843

	precision	recall	f1-score	support
Normal	0.98	0.98	0.98	242
CNV	1.00	0.97	0.98	242
DME	0.99	0.99	0.99	242
DRUSEN	0.98	1.00	0.99	242
micro avg	0.99	0.99	0.99	968
macro avg	0.99	0.99	0.99	968
weighted avg	0.99	0.99	0.99	968

Fig. 2. Matrix showing the precision and recall for each classed obtained by second CNN architecture .

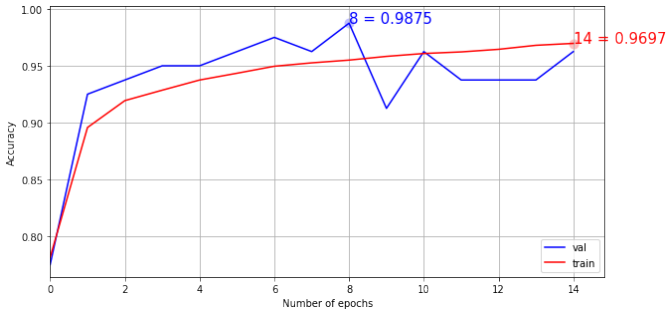

Fig. 3. Model's accuracy obtained by the second CNN architecture.

Inception V3- Rmsprop optimizer - accuracy ON TEST SET : 0.9927685950413223

	precision	recall	f1-score	support
Normal	1.00	0.98	0.99	242
CNV	0.98	1.00	0.99	242
DME	1.00	1.00	1.00	242
DRUSEN	1.00	0.99	0.99	242

Fig. 4. Model's accuracy obtained by the Inception model while testing

The performance on the test data is consistent with the performance on the training data. DME , DRUSEN and Normal images have a great precision value of 100%, in all the evaluation metrices. CNV have lower precision value compared to other classes.

Fig. 5. Comparaisons between the results obtained by our experiments.

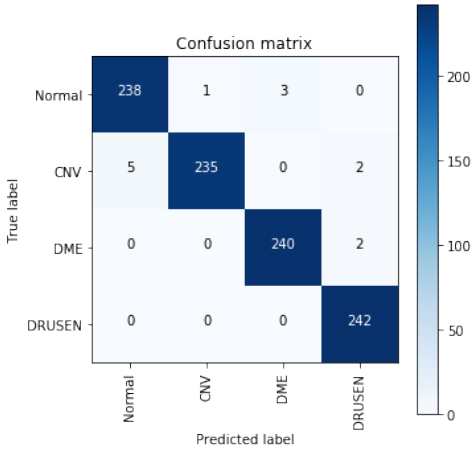

Fig. 6. Confustion matrix obtained by the second CNN architecture.

Predictive labels on the X-axes and true labels on the Y-axes, we have used 968 images to test our model, the blue cells shows the correct predicted images meaning the predicted lables matches the actual labels of the images, in this case the model correctly predicted 955 images out of 968 images where the DRUSEN images were 100% correctly predicted. and the wight cells show the incorrectly predicted labels which 13 in this case.

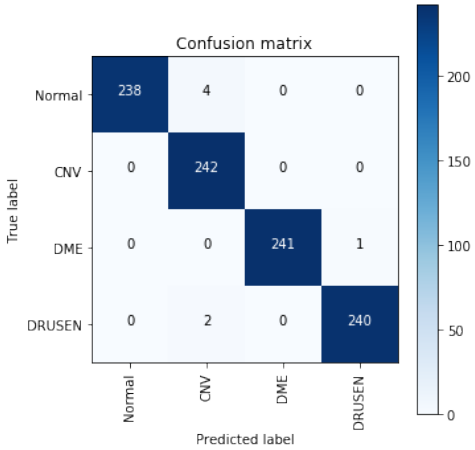

Fig. 7. Confustion matrix obtained by InceptionV3 model.

This model was able to correctly predict 961 out of 968 images, meaning only 7 images were incorrectly classified, this shows better results compared to the previous confusion matrix where the 2nd CNN models miss-classified 13 images. The CNV category is better classed here compared to the previous confusion matrix of the 2nd CNN architecture, and the test accuracy improved as well. , all the instances of CNV class were correctly classified.

V. Conclusion

We proposed a novel classification model based on deep neural networks to classify retinal images obtained from OCT scanning device, the results has shown that the use of Convolutional neural networks [40] can give very interesting results in image classification and can be used to asses doctors in medical diagnose and opens up to a new, simple and effective method for early CNV, DME, and DRUSEN detection.
We have also seen that the use of some pre-trained models can enhance the results in time and model effectiveness.
Our proposed CNN architecture and the Inception V3 model outperformed the results shown in [32].

References

[1] Al-Mujaini, Abdullah et al. Optical coherence tomography: clinical applications in medical practice. Oman medical journal vol. 28,2 (2013): 86-91. doi:10.5001/omj.2013.24

[2] Huang, D et al. Optical coherence tomography. Science (New York, N.Y.) vol. 254,5035 (1991): 1178-81.

[3] Weston Eye Center. Weston Eye Center, www.westoneyecenter.com/Glaucoma/.

[4] Chong GT, Lee RK. Glaucoma versus red disease: imaging and glaucoma diagnosis. Curr Opin Ophthalmol 2012;23:79-88.

[5] Jie Wang, Yangjing Xue, Saroj Thapa, Luping Wang, Jifei Tang, and Kangting Ji, Relation between Age-Related Macular Degeneration and Cardiovascular Events and Mortality: A Systematic Review and Meta-Analysis, BioMed Research International, vol. 2016, Article ID 8212063, 10 pages, 2016.

[6] Pascolini D, Mariotti SP Global estimates of visual impairment: 2010 British Journal of Ophthalmology 2012;96:614-618.

[7] Chan W, Ohji M, Lai TYY, et al Choroidal neovascularisation in pathological myopia: an update in management British Journal of Ophthalmology 2005;89:1522-1528.

[8] Natural History of Choroidal Neovascularization in Degenerative Myopia Avila, Marcos P. et al. Ophthalmology , Volume 91 , Issue 12 , 1573 - 1581

[9] Spaide, Richard F, and Christine A Curcio. Drusen characterization with multimodal imaging. Retina (Philadelphia, Pa.) vol. 30,9 (2010): 1441-54. doi:10.1097/IAE.0b013e3181ee5ce8

[10] 1- Kenji Suzuki, Pingkun Yan, Fei Wang, and Dinggang Shen. Machine learning in medical imaging. Int J Biomed Imaging, 2012:12, 2012.

[11] Weston Eye Center. Weston Eye Center, www.westoneyecenter.com/Glaucoma/.

[12] Panwar, Nishtha et al. Fundus Photography in the 21st Century–A Review of Recent Technological Advances and Their Implications for Worldwide Healthcare. Telemedicine journal and e-health : the official journal of the American Telemedicine Association vol. 22,3 (2016): 198-208. doi:10.1089/tmj.2015.0068

[13] B-Scan (Ophthalmic Ultrasound), eyehealthconsultants.com/how-we-help-you/hi-tech-eye-diagnostics/17-hi-tech-eye-diagnostics/117-b-scan-ophthalmic-ultrasound.html.

[14] "Smartphone Funduscopy-How to Use Smartphone to Take Fundus Photographs." EyeWiki, 2 Mar. 2018, eyewiki.aao.org/SmartphoneFunduscopy-How-to-use-smartphone-to-take-fundus-photographs.

[15] Choroidal Neovascularization (CNV). Practice Essentials, Background, Pathophysiology, 26 Feb. 2019, emedicine.medscape.com/article/1190818-overview.

[16] Drusen in Age-Related Macular Degeneration Abdelsalam, Ahmed et al. Survey of Ophthalmology , Volume 44 , Issue 1 , 1 - 29

[17] Tarunika, K & Pradeeba, R.B & P, Aruna. (2018). Applying Machine Learning Techniques for Speech Emotion Recognition. 1-5. 10.1109/IC-CCNT.2018.8494104.

[18] S. Badea, M & Felea, Iulian & Florea, Laura & Vertan, Constantin. (2016). The use of deep learning in image segmentation, classification and detection.

[19] Ali, Liaqat & Khelil, Khaled & Kanwal et al. (2017). Machine learning based computer-aided diagnosis of liver tumours. 139-145. 10.1109/ICCI-CC.2017.8109742.

[20] Ioffe, Sergey & Szegedy, Christian. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.

[21] Dumoulin, Vincent & Visin, Francesco. (2016). A guide to convolution arithmetic for deep learning.

[22] CS231n Convolutional Neural Networks for Visual Recognition, cs231n.github.io/.

[23] Scherer, Dominik & Mller, Andreas & Behnke, Sven. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. 92-101. 10.1007/978-3-642-15825-4-10.

[24] Gupta, Dishashree, and Dishashree. Fundamentals of Deep Learning - Activation Functions and Their Use. Analytics Vidhya, 22 Feb. 2019, www.analyticsvidhya.com/blog/2017/10/fundamentals-deep-learning-activation-functions-when-to-use-them/.

[25] Janney, Bethanney & Meera, G & Uma Shankar, G & Divakaran, Sindu & Abraham, S. (2015). Detection and classification of exudates in retinal image using image processing techniques. Journal of Chemical and Pharmaceutical Sciences. 8. 541-546.

[26] Banerjee, Sreeparna & Roy Chowdhury, Amrita. (2015). Case Based Reasoning in the Detection of Retinal Abnormalities Using Decision Trees. Procedia Computer Science. 46. 402-408. 10.1016/j.procs.2015.02.037.

[27] https://data.mendeley.com/datasets/rscbjbr9sj/2

[28] S. Kermany et al. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell. 172. 1122-1131.e9. 10.1016/j.cell.2018.02.010.

[29] Donges, Niklas. The Random Forest Algorithm. Towards Data Science, Towards Data Science, 22 Feb. 2018, towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd.

[30] Awais, Muhammad & Mller, Henning & Meriaudeau, Fabrice. (2017). Classification of SD-OCT images using Deep learning approach. 10.1109/ICSIPA.2017.8120661.

[31] Hussain, Md. Akter & Bhuiyan, Alauddin et al (2018). Classification of healthy and diseased retina using SD-OCT imaging and Random Forest algorithm. PLOS ONE. 13. e0198281. 10.1371/journal.pone.0198281.

[32] Nugroho, Kuntoro. (2018). A Comparison of Handcrafted and Deep Neural Network Feature Extraction for Classifying Optical Coherence Tomography (OCT) Images. 1-6. 10.1109/ICICOS.2018.8621687.

[33] Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images Lee, Cecilia S. et al. Ophthalmology Retina , Volume 1 , Issue 4 , 322 - 327

[34] Mehta, Parmita & Lee, Aaron & Lee, Cecilia & Balazinska, Magdalena & Rokem, Ariel. (2018). Multilabel multiclass classification of OCT images augmented with age, gender and visual acuity data. 10.1101/316349.

[35] Romero Aquino, Nelson & Gutoski, Matheus & Hattori, Leandro & Lopes, Heitor. (2017). The Effect of Data Augmentation on the Performance of Convolutional Neural Networks.

[36] He, Shenghua & Zheng, Jie et al. (2018). Convolutional neural network based automatic plaque characterization for intracoronary optical coherence tomography images. 107. 10.1117/12.2293957.

[37] Qiu, Suo. (2018). Global Weighted Average Pooling Bridges Pixel-level Localization and Image-level Classification.

[38] Prechelt, Lutz. (2000). Early Stopping - But When?. 10.1007/3-540-49430-83.

[39] Srivastava, Nitish & Hinton, Geoffrey & Krizhevsky, Alex & Sutskever, Ilya & Salakhutdinov, Ruslan. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. 15. 1929-1958.

[40] Bengio, Y & Lecun, Yann. (1997). Convolutional Networks for Images, Speech, and Time-Series.

[41] Deng, Jia & Dong, Wei & Socher, Richard & Li, Li-Jia & Li, Kai & Li, Fei Fei. (2009). ImageNet: a Large-Scale Hierarchical Image Database. IEEE Conference on Computer Vision and Pattern Recognition. 248-255. 10.1109/CVPR.2009.5206848.

[42] G. Venhuizen, Freerk & van Ginneken, Bram et al. (2017). Automated Staging of Age-Related Macular Degeneration Using Optical Coherence Tomography. Investigative Opthalmology & Visual Science. 58. 2318. 10.1167/iovs.16-20541.

[43] https://colab.research.google.com

[44] Simonyan, Karen & Zisserman, Andrew. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556.

[45] Szegedy, Christian & Vanhoucke, Vincent & Ioffe, Sergey & Shlens, Jon & Wojna, ZB. (2016). Rethinking the Inception Architecture for Computer Vision. 10.1109/CVPR.2016.308.