

Classification of Optical Coherence Tomography using Convolutional Neural Networks

A. A. Saraiva^{2,6}^a, D. B. S. Santos¹^b, Pimentel Pedro¹^c, Jose Vigno Moura Sousa¹^d,
N. M. Fonseca Ferreira^{3,4}^e, J. E. S. Batista Neto⁶, Salviano Soares³^f and Antonio Valente^{2,5}^g

¹UESPI - University of State Piauí, Piripiri, Brazil

²University of Trás-os-Montes and Alto Douro, Vila Real, Portugal

³Coimbra Polytechnic, ISEC, Coimbra, Portugal

⁴Knowledge Engineering and Decision-Support Research Center (GECAD) of the Institute of Engineering, Polytechnic Institute of Porto, Portugal

⁵INESC-TEC Technology and Science, Porto, Portugal

⁶University of São Paulo, São Carlos, Brazil

Keywords: OCT, CNN, Classification, K-fold, Labeled Optical Coherence Tomography.

Abstract: This article describes a classification model of optical coherence tomography images using convolution neural network. The dataset used was the Labeled Optical Coherence Tomography provided by (Kermany et al., 2018) with a total of 84495 images, with 4 classes: normal, drusen, diabetic macular edema and choroidal neovascularization. To evaluate the generalization capacity of the models k-fold cross-validation was used. The classification models were shown to be efficient, and as a result an average accuracy of 94.35% was obtained.

1 INTRODUCTION

An examination known as optical coherence tomography (OCT) has gained ground in the latest complementary clinical tests for the diagnosis of retinal and vitreous disease (Preti et al., 2018).

This technology was developed by Fujimoto at the Massachusetts Institute of Technology, applied in the ophthalmological diagnosis by Puliafito. The use of this examination has become fundamental in the diagnosis, on evolution and postoperative control of multiple macular conditions (Dimitrova et al., 2017).

According to (Swanson and Fujimoto, 2017) approximately 30 million procedures of optical coherence tomography (OCT) images are performed per year, the analysis and interpretation of these images,

consumes a significant amount of time. OCT helped patients prevent or minimize vision loss by detecting retinal diseases in the early stages of treatment (Swanson and Fujimoto, 2017).

According to (Sivaprasad and Moore, 2008), the growth of new choroidal blood vessels is known as choroidal neovascularization (CNV). These new vessels come from a rupture in the Bruch membrane that is located in the subretinal pigment epithelium. According to (Baxter et al., 2013) CNV occurs in about 2 to 3% of cases of posterior uveitis.

Diabetic macular edema (DME) is a complication of diabetes caused by fluid accumulation in the macula, or central portion of the eye, that causes the macula to swell (Wells et al., 2016). The macula is filled with cells responsible for direct vision that aid in reading and directing (Wells et al., 2016).

When the macula begins to fill with fluid and swell, the capacity of these cells is impaired, causing blurred vision (Bressler et al., 2016). The DME is diabetic retinopathy, in which the blood vessels of the eye are damaged, allowing the fluid to escape, this type of disease can also be diagnosed through the

^a  <https://orcid.org/0000-0002-3960-697X>

^b  <https://orcid.org/0000-0003-4018-242X>

^c  <https://orcid.org/0000-0002-5291-0810>

^d  <https://orcid.org/0000-0002-5164-360X>

^e  <https://orcid.org/0000-0002-2204-6339>

^f  <https://orcid.org/0000-0001-5862-5706>

^g  <https://orcid.org/0000-0002-5798-1298>

OCT (Gill et al., 2017).

Drusen are small accumulations of yellow or white extracellular material that accumulate between Bruch's membrane and the retinal pigment epithelium of the eye, which can be detected by OCT (Gaier et al., 2017). The presence of small Drusen is normal with advancing age, and most people over 40 have some Drusen (Alten et al., 2017). However, the presence of larger and more numerous Drusen in the macula is an early sign of age-related macular degeneration (DMRI) (Schlanitz et al., 2017).

In this way, he inspired the design an OCT image classifier, in order to classify three types of pathologies visible in OCT: CNV, DME, and DRUSEN, in an automated and fast way. The method chosen and implemented constitutes in the classification of OCT images of retinas of living patients, from this, to identify whether or not they have any of the diseases mentioned above. For the classification was used the dataset Labeled Optical Coherence Tomography with total 84,495 images provided by (Kermany et al., 2018).

The classification stage consists of two sub-steps, where the first is carried out to the training of a deep learning known as Convolutional Neural Networks (CNN), the second sub-step is the validation of the model, that is, the tests with unknown images by CNN (Saraiva. et al., 2019b), (Saraiva. et al., 2019a). The method covered ensures a robust coverage in image recognition, under certain assumptions that will be clarified throughout the text.

The paper is divided into 5 sections, in which section 2 is characterized by the description of the methodology applied, followed by the validation metrics in section 3. The results after the application of the proposal and the final conclusions are presented in sections 4 and 5, respectively.

The data set is organized into 4 categories: NORMAL, CNV, DME, DRUSEN, totaling 84,495 OCT images. The OCT images were selected from retrospective cohorts of adult patients at the Shiley Eye Institute of the University of California at San Diego, the California Retina Research Foundation, the Medical Center Ophthalmology Associates First People's Hospital and the Beijing Tongren Eye Center between July 1 2013 and March 1, 2017.

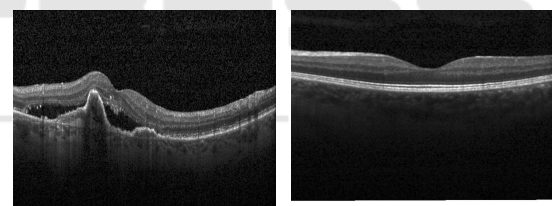
Each image underwent a step-by-step classification system consisting of several layers of trained classifiers with experience for checking and correcting image labels. Each image in the data set began with a label that matches the patient's most recent diagnosis. The first level of classifiers consisted of undergraduate medical students who had completed and passed a review of the OCT course of interpretation.

This first level of classifiers conducted initial quality control and excluded OCT images containing noises or significant reductions in image resolution. The second level of classifiers consisted of four ophthalmologists who independently classified each image that passed the first level. Finally, a third level of two senior independent retinal specialists, each with more than 20 years of experience in clinical retina, checked the labels for each image (Kermany et al., 2018).

1.1 CNV

CNV generally reaches individuals under 50 years of age, and its early diagnosis is extremely important for the prompt institution of treatment, which may prevent the occurrence of fibrosis and consequent permanent central visual acuity in this economically active population (Roy et al., 2017).

The main symptoms are central scotoma and metamorphopsia, but the patient may be asymptomatic, especially when the affected eye already has low visual acuity prior to the presence of a central or pericentral scar or granuloma. In the figure 1 it is possible to visualize the OCT image, divided into normal and CNV.



(a) Image of CNV

(b) Normal image

Figure 1: Example of OCT images divided into normal and with CNV.

1.2 DME

In DME, an accumulation of liquid and proteins occurs in the macula region (Parhi et al., 2017). The retina becomes swollen, and the vision is greatly impaired. This accumulation of liquid and proteins begins because of the excess of blood sugar in a prolonged way, which damages the blood vessels (Parhi et al., 2017).

DME, defined as a retinal thickening involving or approaching the center of the macula, is the most common cause of vision loss in patients affected by diabetes. In the figure 2 you can view the OCT image, divided into Normal and DME.

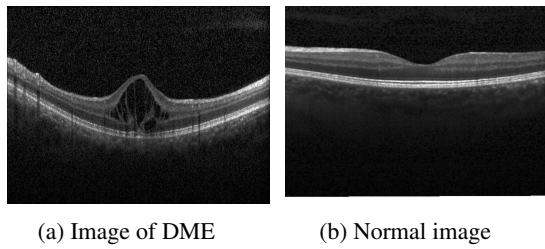


Figure 2: Example of OCT images divided into normal and with DME.

1.3 DRUSEN

Optical disc drusen are calcified deposits of extruded mitochondria that appear in the upper part of the optic nerve in approximately 2% of the population (Gaier et al., 2017). In the figure 3 you can view the OCT image, divided into normal and DRUSEN.

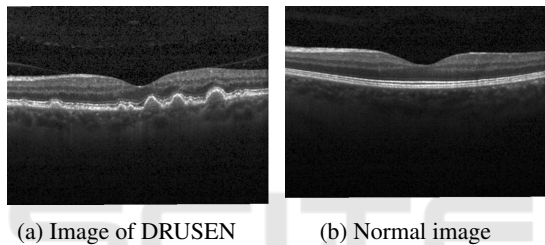


Figure 3: Example of OCT images divided into normal and with DRUSEN.

2 MATERIALS AND METHODS

In this section, the structure of the adopted systems will be presented to solve the classification of OCT images, classifying them as, CNV, DME, NORMAL, DRUSEN, will also be presented the entire structure of the algorithms as well as the evaluation metrics.

2.1 Structure of the System

The system was divided into three stages, the first consisting of the division and normalization of the dataset, according to the figure 4. In the second one the training was carried out and finally the data validation was done. Pre-processing consists of normalizing the data, the images are in grayscale, all pixels are divided by 255, to convert them into floating points. Represented in the figure 4 in yellow.

In figure 4 you can see the process of construction and training of artificial neural network model. In the test data prediction step, the test images that were separated by the k-fold algorithm are entered,

so accuracy is collected. The process is repeated 5 times, changing the test and training images after the k-fold calculation.

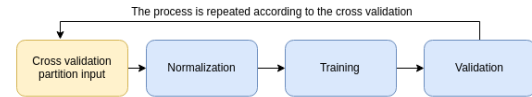


Figure 4: Construct, training and validate of the models.

2.2 CNN

CNNs are similar to traditional neural networks, both are composed of neurons that have weights and bias that need to be trained. Each neuron receives some inputs, applies the scalar product of inputs and weights in addition to a non-linear function (Chen et al., 2017).

A CNN assumes that all inputs are images, which allows you to encode some properties in the architecture. Traditional neural networks are not scalable for images, since they produce a very high number of weights to be trained (Esteva et al., 2017).

A CNN consists of a sequence of layers as can be seen figure 6, in addition to input layer, which is usually composed of an image with width and height, there are three main layers: convolutional layer, pooling layer and fully connected layer. In addition, after a convolutional layer it is common an activation layer, normally a linear rectification unit function (ReLU) equations 1, 2. These layers, when sequenced (or stacked), form an architecture of a CNN (Salamon and Bello, 2017).

$$f(x) = x^+ = \max(0, x) \quad (1)$$

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (2)$$

2.2.1 Convolutional Layer

The convolutional layer is the most important layer of the network, where it carries out the heaviest part of computational processing. This layer is composed of a set of filters (kernels) capable of learning according to a training (Ustinova et al., 2017). The kernels are small matrices that in this case was used the size 3x3 to obtain a better precision in the time to go through the matrix of the images, composed by real values that can be interpreted as weights.

Given a two-dimensional image, I , and a small array, K of size $h \times w$ (kernel), the convoked image, $I * K$, is calculated by overlapping the kernel at the top of the image of all possible shapes, and recording the

sum of the elementary products between the image and the kernel equation 3.

$$(I * K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{ij} \cdot I_{x+i-1, y+j-1} \quad (3)$$

The kernels are convolved with the input data to get a feature map. These maps indicate regions in which specific features in relation to kernels. The actual values of the kernels change throughout the training, causing the network to learn to identify significant regions to extract characteristics from the data set (Maggiori et al., 2017), in this way, each filter results in an output of a three-dimensional array. In the convolution results matrices the ReLU activation function, equations 1, 2 are applied in each element of the convolution result.

2.2.2 Pooling Layer

After convolution layer exists a pooling layer. The pooling technique is used to reduce the spatial size of the resulting convolution matrices, according to the figure 5. Consequently, this technique reduces the amount of parameters to be learned in the network, contributing to the control of over-fitting, ie avoiding the condition when a trained model works very well in training data, but does not work very well in test data (Yu et al., 2017). The pooling layers operate independently on each of the channels of the convolution result. In addition, you must first determine the size of the filter to perform pooling.

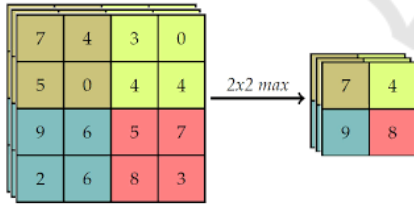


Figure 5: Example max-pooling with a 4x4 image.

The maximum pool operation reduces the size of the resource map, this operation can be described by the equation 4. Let S be the value of the passed and $Q \times Q$ the shape of the feature map before the maximum grouping and p determines the clustering max-pooling size (Havaei et al., 2017). The output of the max-pooling operation would be $D \times D$ size.

$$D = (Q - p) / S + 1 \quad (4)$$

2.2.3 Fully Connected Layer

The fully connected layer comes after a convolutional or pooling layer, it is necessary to connect each ele-

ment of the convolution output matrices to an input neuron. The output of the convolutional and pooling layers represent the characteristics extracted from the input image. The purpose of fully-connected layers is to use these characteristics to classify the image in a pre-determined class.

The last two layers of the network use the sigmoid function as the activation function, equation 5. This function takes a real value and "transforms" it into the interval between 0 and 1. In particular, large negative numbers become 0 and large positive numbers become 1 (Zaheer and Shaziya, 2018). The sigmoid function has seen frequent use historically since it has a good interpretation like the firing rate of a neuron: from not firing (0) to a fully saturated firing at a presumed maximum frequency (1) (Zaheer and Shaziya, 2018).

$$f(x) = \text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

The technique known as dropout is also used in the fully connected layer to reduce training time and avoid over-fitting. This technique consists in randomly removing a certain percentage of neurons from a layer at each training iteration, re-adding them to the next iteration (Kovács et al., 2017).

2.2.4 CNN Architecture

In the figure 6 the CNN architecture is displayed, it has 12 layers, where the first ten convolutional layers and the last two without convolution with the sigmoid activation function. The input of the network receives a 150x150 pixel image, each the convolutional layer has the activation function ReLUs. For the convolution kernel, the 3x3 size was adopted, because this way it is possible to have a greater precision in the time to go through the entire image.

After two convolutional layers a Max-pooling layer is used, this reduces the size of the matrices resulting from the convolution. With this layer it is possible to reduce the amount of parameters that will be learned by the network, this way it is done over fitting control.

In the latter two the sigmoid activation function is used, this function is responsible for making the probabilistic distribution of the input image belong to each of the classes in which the network was trained. To reduce the training time and to avoid over-fitting is used dropout in the layer, ie it is randomly removed at each training interaction, a certain percentage of the neurons of a layer, re-adding them in the following iteration.

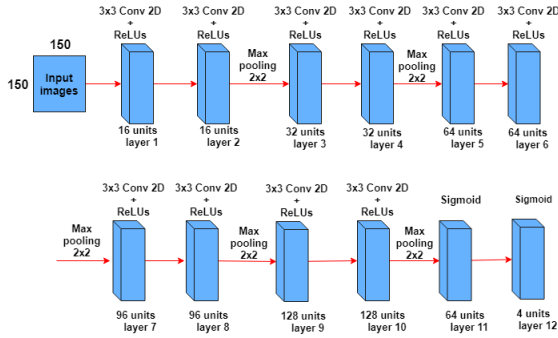


Figure 6: Construction of the CNN training mode.

3 METRICS OF THE EVALUATION

3.1 Cross Validation

Cross-validation is an evaluation technique on the ability of generalization models, from a dataset, is widely used in problems where the object is the modeling and prediction (Vehtari et al., 2017). With this it is possible to estimate how precise the model is, that is, its accuracy with data that it does not know.

The k-fold cross-validation method consists of dividing the total set into k subsets of the same size. One subset is used for testing, and the other $k-1$ subsets for training. This process is repeated by k times, if circularly changing the subset of tests (Grimm et al., 2017).

The final precision of the model is estimated by equation 6, at where Ac_f is the sum of the differences between the actual value y_i and the predicted value $\hat{y}_i$ e k is the amount of k-fold divisions. With this it is possible to infer the generalization capacity of the network.

$$Ac_f = \frac{1}{k} \sum_{i=1}^k (y_i - \hat{y}_i) \quad (6)$$

3.2 Confusion Matrix

As a statistical tool we have the confusion matrix that provides the basis for describe the accuracy of the classification and characterize the errors, helping refine the ranking (Saraiva et al., 2018). The confusion matrix is formed by an array of squares of numbers arranged in rows and columns that express the number of sample units of a particular category, inferred by a decision rule, compared to the category current field.

Usually below the columns is the set reference data that is compared to the product data of the classification that are represented along the lines. The figure 7 shows the representation of an array of confusion. The elements of the main diagonal in bold indicate the level of accuracy, or agreement, between the two sets of data.

Classification	Reference data				Total lines
	1	2	...	a	
1	x_{11}	x_{12}	...	x_{1a}	x_{1+}
2	x_{21}	x_{22}	...	x_{2a}	x_{2+}
...	...	...	...	...	...
a	x_{a1}	x_{a2}	...	x_{aa}	x_{a+}
Total columns	x_{+1}	x_{+2}	...	x_{+a}	n

Figure 7: Example matrix of confusion.

The measures derived from the confusion matrix are: the total accuracy being that chosen by the present work, accuracy of individual class, producer precision, user precision and Kappa index, among others. The total accuracy is calculated by dividing the sum of the main diagonal of the error matrix x_{ii} , by the total number of samples collected n . According to the equation 7.

$$T = \frac{\sum_{i=1}^a x_{ii}}{n} \quad (7)$$

4 RESULTS

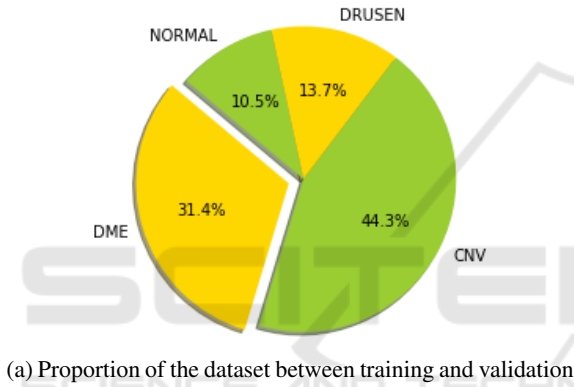
In this section will be presented the classification performance results of the training model. The metrics used to evaluate the results are: The average accuracy of the cross validation, specificity and sensitivity, given by the ROC curve.

In the table 1 the results obtained by the training network are presented. through the table it is possible to extract information such as: False positives, False negatives, True positives, True Negative and accuracy of each interaction of cross validation. The average accuracy of the network was 94.35%.

In the figure 9 it is possible to check the data relating to iteration 5 of the table 1. In it is shown a confluence matrix, a graph of the ROC curve and the precision recall curve. Where it is possible to visualize the sensitivity and specificity of each class. It is possible to visualize in the figure 8 the graphs of proportions, of each disease and of the test and training.

Table 1: Table of iterations and classifications.

**		CNV	DME	NORMAL	DRUSEN	ACC
1	CNV	5274	10	19	10	
1	DME	38	7247	77	83	
1	NORMAL	2017	63	2045	5	
1	DRUSEN	193	229	9	1343	
1	ACC					94.44%
2	CNV	5087	14	168	44	
2	DME	9	7329	47	106	
2	NORMAL	7	96	2148	6	
2	DRUSEN	126	212	22	1413	
2	ACC					94.55%
3	CNV	5252	12	22	27	
3	DME	17	7348	19	107	
3	NORMAL	176	118	2014	12	
3	DRUSEN	138	172	2	1461	
3	ACC					95.13%
4	CNV	5253	9	8	43	
4	DME	58	7130	56	247	
4	NORMAL	390	71	1831	27	
4	DRUSEN	160	79	7	1527	
4	ACC					93.16%
5	CNV	5275	13	19	06	
5	DME	21	7213	48	209	
5	NORMAL	171	101	2043	4	
5	DRUSEN	206	123	6	1438	
5	ACC					94.51%



(b) Demographics Proportion Chart

Figure 8: Demographics Proportion Chart.

It is worth noting that the test and training data were separated proportionally to their total quantity.

Neural network training was performed using an NVIDIA GTX 1060 video card, which features 1280 CUDA cores (processors), 6 GB of dedicated memory, 12 GB of RAM and a fourth-generation Core i5 Processor with time of training of 29 minutes.

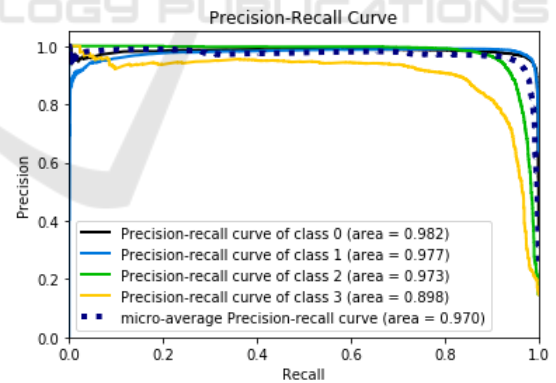
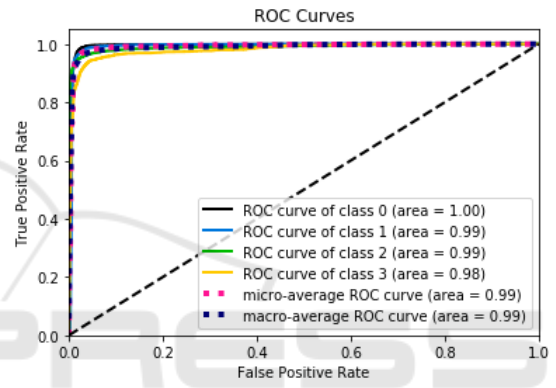
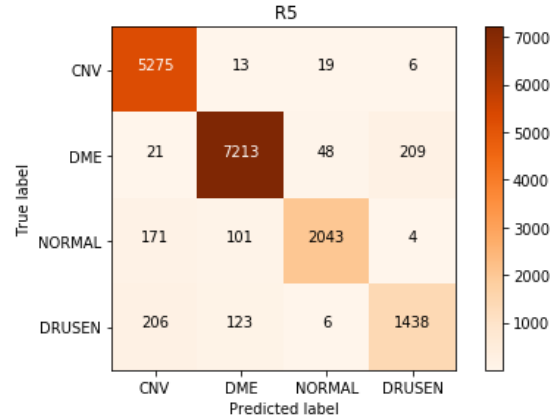


Figure 9: Interaction test 5 table 1, Precision recall curve and ROC curves.

5 CONCLUSION

In this work, a model was presented to classify visible pathologies in OCT, the classes are CNV, DME, DRUSEN and NORMAL. For the validation of the models, cross validation was performed, where it is

possible to verify the generalization capacity. The classification model evaluated in this work was shown to be efficient, obtaining an average of 94.35 % accuracy. Being able to reach a high accuracy, even with the unbalanced dataset and with the iterations obtained in the cross validation being proxies.

ACKNOWLEDGMENTS

The elaboration of this work would not have been possible without the collaboration of the Engineering and DecisionSupport Research Center (GECAD) of the Institute of Engineering, Polytechnic Institute of Porto, Portugal and FAPEMA.

REFERENCES

- Alten, F., Lauermann, J., Clemens, C., Heiduschka, P., and Eter, N. (2017). Signal reduction in choriocapillaris and segmentation errors in spectral domain oct angiography caused by soft drusen. *Graefes's Archive for Clinical and Experimental Ophthalmology*, 255(12):2347–2355.
- Baxter, S. L., Pistilli, M., Pujari, S. S., Liesegang, T. L., Suhler, E. B., Thorne, J. E., Foster, C. S., Jabs, D. A., Levy-Clarke, G. A., Nussenblatt, R. B., et al. (2013). Risk of choroidal neovascularization among the uveitides. *American journal of ophthalmology*, 156(3):468–477.
- Bressler, S. B., Glassman, A. R., Almukhtar, T., Bressler, N. M., Ferris, F. L., Googe Jr, J. M., Gupta, S. K., Jampol, L. M., Melia, M., Wells III, J. A., et al. (2016). Five-year outcomes of ranibizumab with prompt or deferred laser versus laser or triamcinolone plus deferred ranibizumab for diabetic macular edema. *American journal of ophthalmology*, 164:57–68.
- Chen, Y.-H., Krishna, T., Emer, J. S., and Sze, V. (2017). Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks. *IEEE Journal of Solid-State Circuits*, 52(1):127–138.
- Dimitrova, G., Chihara, E., Takahashi, H., Amano, H., and Okazaki, K. (2017). Quantitative retinal optical coherence tomography angiography in patients with diabetes without diabetic retinopathy. *Investigative ophthalmology & visual science*, 58(1):190–196.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115.
- Gaier, E. D., Rizzo III, J. F., Miller, J. B., and Cestari, D. M. (2017). Focal capillary dropout associated with optic disc drusen using optical coherence tomographic angiography. *Journal of Neuro-ophthalmology*, 37(4):405–410.
- Gill, A., Cole, E. D., Novais, E. A., Louzada, R. N., Carlo, T., Duker, J. S., Waheed, N. K., Bauman, C. R., and Witkin, A. J. (2017). Visualization of changes in the foveal avascular zone in both observed and treated diabetic macular edema using optical coherence tomography angiography. *International journal of retina and vitreous*, 3(1):19.
- Grimm, K. J., Mazza, G. L., and Davoudzadeh, P. (2017). Model selection in finite mixture models: A k-fold cross-validation approach. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(2):246–256.
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., and Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35:18–31.
- Kermany, D., Zhang, K., and Goldbaum, M. (2018). Labeled optical coherence tomography (oct) and chest x-ray images for classification. *Structural Equation Modeling: A Multidisciplinary Journal*.
- Kovács, G., Tóth, L., Van Compernelle, D., and Ganapathy, S. (2017). Increasing the robustness of cnn acoustic models using autoregressive moving average spectrogram features and channel dropout. *Pattern Recognition Letters*, 100:44–50.
- Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P. (2017). Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2):645–657.
- Parhi, K. K., Reinsbach, M., Koozekanani, D. D., and Roychowdhury, S. (2017). Automated oct segmentation for images with dme. In *Medical Image Analysis and Informatics*, pages 115–132. CRC Press.
- Preti, R. C., Govetto, A., Aqueta Filho, R. G., Zacharias, L. C., Pimentel, S. G., Takahashi, W. Y., Monteiro, M. L., Hubschman, J. P., Sarraf, D., et al. (2018). Optical coherence tomography analysis of outer retinal tubulations: sequential evolution and pathophysiological insights. *Retina*, 38(8):1518–1525.
- Roy, R., Saurabh, K., Bansal, A., Kumar, A., Majumdar, A. K., and Paul, S. S. (2017). Inflammatory choroidal neovascularization in indian eyes: Etiology, clinical features, and outcomes to anti-vascular endothelial growth factor. *Indian journal of ophthalmology*, 65(4):295.
- Salamon, J. and Bello, J. P. (2017). Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Processing Letters*, 24(3):279–283.
- Saraiva, A., Melo, R., Filipe, V., Sousa, J., Ferreira, N. F., and Valente, A. (2018). Mobile multirobot manipulation by image recognition.
- Saraiva, A. A., Ferreira, N. M. F., de Sousa, L. L., Costa, N. J. C., Sousa, J. V. M., Santos, D. B. S., Valente, A., and Soares, S. (2019a). Classification of images of childhood pneumonia using convolutional neural networks. In *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2: BIOIMAGING*, pages 112–119. INSTICC, SciTePress.
- Saraiva, A. A., Santos, D. B. S., Costa, N. J. C., Sousa, J. V. M., Ferreira, N. M. F., Valente, A., and Soares,

- S. (2019b). Models of learning to classify x-ray images for the detection of pneumonia using neural networks. In *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2: BIOIMAGING*, pages 76–83. INSTICC, SciTePress.
- Schlanitz, F. G., Baumann, B., Kundi, M., Sacu, S., Baratsits, M., Scheschy, U., Shahlaee, A., Mittermüller, T. J., Montuoro, A., Roberts, P., et al. (2017). Drusen volume development over time and its relevance to the course of age-related macular degeneration. *British Journal of Ophthalmology*, 101(2):198–203.
- Sivaprasad, S. and Moore, A. (2008). Choroidal neovascularisation in children. *British Journal of Ophthalmology*, 92(4):451–454.
- Swanson, E. A. and Fujimoto, J. G. (2017). The ecosystem that powered the translation of oct from fundamental research to clinical and commercial impact. *Biomedical optics express*, 8(3):1638–1664.
- Ustinova, E., Ganin, Y., and Lempitsky, V. (2017). Multi-region bilinear convolutional neural networks for person re-identification. In *Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on*, pages 1–6. IEEE.
- Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5):1413–1432.
- Wells, J. A., Glassman, A. R., Ayala, A. R., Jampol, L. M., Bressler, N. M., Bressler, S. B., Brucker, A. J., Ferris, F. L., Hampton, G. R., Jhaveri, C., et al. (2016). Aflibercept, bevacizumab, or ranibizumab for diabetic macular edema: two-year results from a comparative effectiveness randomized clinical trial. *Ophthalmology*, 123(6):1351–1359.
- Yu, S., Jia, S., and Xu, C. (2017). Convolutional neural networks for hyperspectral image classification. *Neurocomputing*, 219:88–98.
- Zaheer, R. and Shaziya, H. (2018). Gpu-based empirical evaluation of activation functions in convolutional neural networks. In *2018 2nd International Conference on Inventive Systems and Control (ICISC)*. IEEE.