MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation
P. Otero et al. (Eds.)

doi:10.3233/SHTI220166

Automatic Diagnosis of Diabetic Retinopathy from Fundus Images Using Neuro-Evolutionary Algorithms

Diego Aquino-Brítez[a], Jordan Ayala Gómez[a], José Luis Vázquez Noguera[a], Miguel García-Torres[ab], Julio César Mello Román[ac], Pedro E. Gardel-Sotomayor[ag], Veronica Elisa Castillo Benitez[d], Ingrid Castro Matto[d], Diego P. Pinto-Roa[ae], Jaques Facon[f], Sebastian Alberto Grillo[a]

[a] Department of Computer Engineer, Universidad Americana, Asunción, Paraguay
[b] Division of Computer Science, Universidad Pablo de Olavide, Seville, Spain
[c] Facultad de Ciencias Exactas y Tecnológicas, Universidad Nacional de Concepción, Concepción, Paraguay
[d] Department of Ophthalmology, Hospital de Clínicas, Facultad de Ciencias Médicas, Universidad Nacional de Asunción, San Lorenzo, Paraguay
[e] Facultad Politécnica, Universidad Nacional de Asunción, San Lorenzo, Paraguay
[f] Department of Computer and Electronics, Universidade Federal do Espírito Santo, São Mateus, Brazil
[g] Facultad de Ciencias y Tecnología, Universidad Católica Nuestra Señora de la Asunción, Campus Alto Paraná, Paraguay

Abstract

Due to the presence of high glucose levels, diabetes mellitus (DM) is a widespread disease that can damage blood vessels in the retina and lead to loss of the visual system. To combat this disease, called Diabetic Retinopathy (DR), retinography, using images of the fundus of the retina, is the most used method for the diagnosis of Diabetic Retinopathy. The Deep Learning (DL) area achieved high performance for the classification of retinal images and even achieved almost the same human performance in diagnostic tasks. However, the performance of DL architectures is highly dependent on the optimal configuration of the hyperparameters. In this article, we propose the use of Neuroevolutionary Algorithms to optimize the hyperparameters corresponding to the DL model for the diagnosis of DR. The results obtained prove that the proposed method outperforms the results obtained by the classical approach.

Keywords:

Diabetic Retinopathy, Neuroevolution, Deep Learning, Evolutionary Algorithms

Introduction

Diabetes mellitus (DM) [20], also known as diabetes, consists of a heterogeneous group of metabolic disorders, mainly caused by the presence of high blood glucose levels. Over time, diabetes can lead to a visual impairment called Diabetic Retinopathy (DR) [12], which is a degenerative eye disease that progressively damages the retina. Therefore, early diagnosis is a key factor in order to prevent the disease progression and permanent damage to the vision system.

In this context, diabetic retina screening [9] is crucial to diagnose this eye disease in its early stages. Thus, increases the probabilities to halt DR degenerative progression. To this end, Ophthalmologists often use the Retinography method [10] to obtain and display color images of the fundus eye, including the retina and other structures. This method allows to verify and analyze the presence of DR-related lesions [30] such as hemorrhages, microaneurysms, hard and soft exudates.

Currently, DR is considered the leading cause of visual loss in the world [37], with approximately 93 million people diagnosed. Considering the increasing worldwide prevalence of DR, there is a need to optimize DR detection procedures. Therefore, it makes sense to propose the optimization of computer-assisted early diagnosis of DR.

DR consists of two types and four stages [39]. The two types correspond to non-proliferative (NPDR) and proliferative (PDR) Diabetic Retinopathy. NPDR corresponds to the initial stages, while PDR is an advanced stage. The four stages of DR correspond to:

1. Mild NPDR: is associated with the occurrence of microaneurysms.
2. Moderate NPDR: consists of a reduction in the blood supply capacity of the vessels that feed the retina.
3. Severe NPDR: occurs due to an increased number of blocked blood vessels.
4. PDR: is the most advanced stage of the disease, which involves retinal detachment and permanent vision loss.

In recent decades, computational power has increased considerably with the development of new hardware and software technologies [18]. In consequence, new algorithms and computational paradigms have been proposed and implemented.

Bio-inspired algorithms [6] such as Deep Neural Networks (DNNs) and Evolutionary Algorithms (EAs) have emerged and proven their worth with successful results that have reached and even surpassed the state of the art of many classical techniques in related fields.

DNNs are brain-inspired models [24] used for many tasks such as image, audio, speech processing, medical image analysis, etc. Deep Learning (DL) is a particular type DNN that has been very successful in image classification tasks [28] allowing the generation of new applications in the field of medical image processing [29]. In particular, DL techniques have been applied for the analysis and classification of fundus images. Research works in this field, include blood vessel segmentation [22], DR classification [21], DR grading [23] and interpretative visual maps [19].

Optimal performance of DNNs models depends on the correct hyperparameters settings [27], which is a major challenge due to the usually large number of hyperparameters present. Moreover, considering the non-existence of an explicit methodology for hyperparameters optimization (HPO). Nevertheless, some of the most popular used HPO methods [38] are random, manual and grid search, as well as Bayesian, Gradient-based and Evolutionary Optimization procedures.

EAs are based on evolutionary biology [2], which are used to solve problems of searching and optimization. In this regard, Genetic Algorithms (GA) [26] mainly consist of an iterative process to find optimal solutions.

The Genetic Algorithm was inspired from the Darwinian theory of evolutionary, in which the survival of fitter creature and their genes were simulated [26]. GA process starts with an initial random population, where each individual is evaluated by a fitness function, the best individuals are enabled for the crossover step, and then certain genetic mutations are applied to the offspring in order to achieve a better adaptation to their environment. The process ends when the stop criterion is reached.

Neuro-evolution [14] is a term that refers to the use of EAs to optimize the configuration and training of DNNs. Many previous works in the literature use EAs to train or optimize DNN Architectures [3, 11, 38]. From our understanding, the use of Neuro-evolutionary Algorithms (NA) has not previously been used to improve DL models for DR diagnosis.

In this work, our proposal is a methodology using an evolutionary optimization procedure for the HPO of DNN with the aim of improving the binary classification rates of fundus images, for automatic diagnosis of DR.

Methods

This section includes details about the data set used in this research, as well as descriptions of the implemented algorithms and the proposed optimization procedure.

Dataset

The data set used [4] in this research corresponds to 1,135 color fundus images collected at the Department of Ophthalmology, Hospital de Clínicas, Facultad de Ciencias Médicas (FCM), Universidad Nacional de Asunción (UNA), Paraguay. The acquisition of the Retinographies was carried out using a Zeiss brand camera, model Visucam 500. Each image was captured, classified and labeled by experienced ophthalmologists. In Figure 1, can be seen some images obtained from the database.

The set of fundus images was divided into two independent data sets of images, the training set including 1,021 images and the test set with 144 images. In this work, 2 classes are considered: patients with DR and healthy patients (NO DR). In Table 1, the number of samples is detailed by classes.

Table 1 – Number of Samples

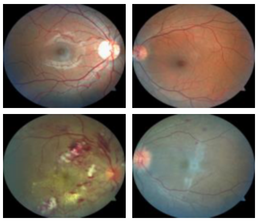
Classes	Train-Validation	Test	Total
DR	513	57	570
NO DR	508	57	565
Total	1021	144	1135

Residual Neural Networks

Deep Convolutional Neural Networks have demonstrated their effectiveness in tasks related to feature extraction and classification [28]. Nevertheless, one of the main drawbacks related to very deep networks is the vanishing gradient. To avoid this, Residual Networks (ResNet) have been proposed in [16].

ResNet can basically be defined as a type of neural network that applies identity mapping, which means that the input from some layer is passed directly or as a shortcut to another layer. In this way, it has been possible to increase the performance of neural networks with a very high number of layers.

Figure 1 – Images of retinal fundus obtained from the dataset. (a, b) correspond to healthy and (c, d) unhealthy.



Transfer Learning

The Transfer Learning (TL) [36] technique is a very useful approach based on leveraging knowledge acquired by pre-trained models for similar tasks. Commonly used in the fields of Computer Vision and Natural Language Processing.

The main advantage of the TL is time saving and the possibility to achieve useful results without the need for a large training data set.

Fine tuning

Training a neural network with a small amount of data for training can affect the generalization capabilities of the model, which can lead to a situation known as overfitting. Basically, it means that the generated model dramatically decreases its performance with unknown data.

The fine tuning (FT) technique [35] consists of freezing a certain number of layers of the network on the training step. Taking into account that the frozen layers have previously learned to extract universal features, allowing their reuse for similar problems. This proposal is very useful in order to avoid overfitting.

Evolutionary Optimization

Nowadays, facing with optimizations problems is very common [33]. Heuristic algorithms are a valid alternative with successful results for solving these problems. In this sense, EAs are inspired by the process of natural selection with the objective of obtaining the optimal solutions for the problem at hand.

Optimization problems may involve one or more objectives that may be subject to certain constraints. When more than one objective is considered, these objectives are often in conflict at the same time [15].

In this study, the EA implemented for the optimization process is a Genetic Algorithm [5]. This choice is motivated by the fact that only one objective (classification accuracy) is considered for this proposal.

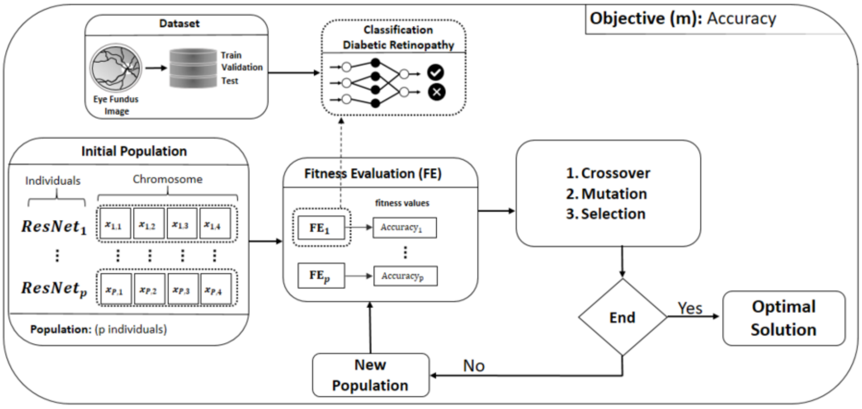
The encoding of chromosomes, and genetics operators are specific to this optimization approach.

The Optimization Procedure

This research involves supervised learning, where the main objective is the optimization of ResNet architecture to achieve higher binary classification accuracies for DR diagnosis. Figure 2 shows the main components of the proposed model.

The essential component corresponds to a GA, which is the EA selected to perform the single-objective optimization.

Figure 2 – Schema of the Neuro-evolutive Algorithm for DR detection

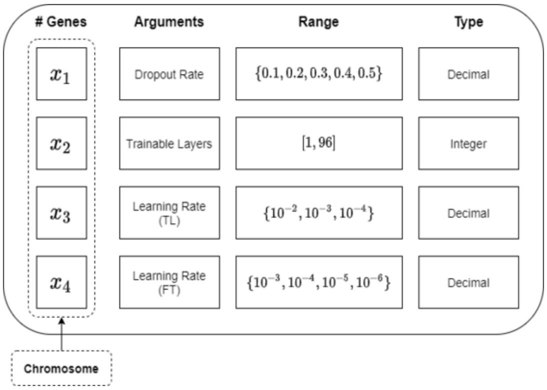


The encoding of the Chromosome corresponds to 4 genes:

- **Dropout rate** [34]**:** consist of the rate of neurons in the network that are randomly deactivated during the training process.
- **Number of Trainable Layers:** indicates the layer number from which the weights will be readjusted during the training process.
- **Learning Rate (TL/FT):** determines the step size at each iteration of the gradient descent during the transfer learning process (TL) and during the fine-tuning process (FT) respectively.

All of the genes are of numeric type and within a limited range of alternatives, which allows speeding up the search process. Figure 3 shows the chromosome encoding process.

Figure 3 – Chromosome Encoding



The Fitness Evaluation corresponds to training of the Residual Neural Network, where the output value is given the binary classification accuracy.

Table 2 lists the parameters used for the GA algorithm.

The Genetic Algorithm optimization process ends when the total number of generations is reached or if the stop criterion is met. The stop criterion [31] is given when the maximum fitness value has not changed in at least seven consecutive generations.

Table 2 – GA Parameters

Parameters	Value
Population size	100
Number of generations	30
Chromosome length (genes)	4
Mutation probability	0.05
Crossover probability	0.75

Results

In this section, we present the results obtained for the ResNet 50v2 Optimization.

Experiment Setup

The implementation code was written in Python 3.7.10. In addition, several libraries available for this language have been used, such as TensorFlow [1] and DEAP [13].

The experiments were performed on the Google Cloud platform, which provides a NVIDIA Tesla K80 GPU, two Intel Xeon Skylake virtual CPUs and 13 GB of RAM.

The dataset was divided into three sets, training, validation and testing. The training dataset has been used to the neural network, while the validation set was used to evaluate the loss and the accuracy through each epoch. The transfer learning process was performed first, by readjusting the weights of the last layer of the base network. After the transfer learning process has ended, the testing dataset was used to evaluate the performance with unseen images.

Afterwards, the fine-tuning process has been performed. In this step, the learning rate is lowered and multiple layers of the neural network are retrained.

Both of these processes have been run multiple times during the evaluation of each individual by the evolutionary algorithm.

The base neural network model employed was the ResNet 50v2 [17], with average pooling layer and an added dropout layer for regularization. The optimized hyperparameters are then the dropout rate, the learning rate of the transfer learning process, the learning rate of the fine-tuning process and the number of trainable layers (used during the fine-tuning process).

Experiment Results

Neuro-evolutive optimization experiments have been conducted using the ResNet 50v2 as a starting point for the optimization process.

Tables 3 and 4 show the results obtained with different neural networks and the optimized solutions provided by our proposal, where it can be seen the improvement obtained in the accuracy from the optimized models with respect to the original (baseline) ResNet 50v2 [17], ResNet 101v2 [7], Xception [8], Mobilenet v2 [32]. The ResNet 50v2 was chosen as a baseline as it performed well and the training time was one of the lowest among the tested networks.

Table 5 shows the optimized parameters that were obtained once the Evolutionary Algorithm reached the stop criterion. Both solutions obtained good results with a dropout rate of 0.4 and with a learning rate of 0.00001. The selected optimal solution required more training time as it had a lower learning rate during the transfer learning and it retrained 90 layers in the fine-tuning process.

Table 3 – Binary Classification results I

Model	Accuracy (%)	Time (s)
ResNet 50v2 (baseline) [17]	82.46	29
ResNet 101v2 [7]	87.72	49
Xception [8]	83.33	30
Mobilenet v2 [32]	78.07	22
Optimized Solution 1	93.86	56
Optimized Solution 2	94.73	51

Table 4 – Binary Classification results II

Model	Sensitivity (%)	Specificity (%)
ResNet 50v2	81.48	86.67
ResNet 101v2	84.31	85.71
Xception	83.33	81.82
Mobilenet v2	90.00	78.38
Optimized Solution 1	91.67	96.30
Optimized Solution 2	93.10	94.64

Table 5 – Optimized Hyperparameters

Hyperparameter	Solution 1	Solution 2
Dropout	0.4	0.4
Trainable Layers	90	64
Learning Rate (TL/FT)	0.001/0.00001	0.01/0.00001

Discussion

Both of the optimized solutions show a substantial improvement in classification performance over the baseline and the other three different neural networks tested.

According to the standard criteria [25] established in the UK on tests of retinal screening for diabetics, the minimum required for sensitivity is 80% and specificity is 95%.

The optimized solution 2 performs better when taking into account the accuracy percentage. However, in order to fit the UK audit standard, only the optimized solution 1 fits requirements for sensitivity and specificity.

Conclusions

In this research a procedure based on neuro-evolutionary algorithms for the optimization of hyperparameters of Deep Neural Networks is proposed in order to improve DR detection in retinographies. For this purpose, a reduced set of fundus images has been used in the application of the transfer learning and fine-tuning processes.

The results obtained by the proposed method demonstrate its validity in the detection of Diabetic Retinopathy (DR) by achieving an accuracy of 93.8% while also complying with the UK audit standards for screening tests with a sensitivity of 91.6% and a specificity of 96.3%. It is important to emphasize that the dataset is part of one of the open databases obtained at Hospital de Clínicas de Asunción, Paraguay.

Future work will focus on multiclass classification and optimization of other neural network architectures, thus allowing a future comparison between different architectures in different areas.

Acknowledgements

This research was funded by CONACYT, Paraguay, grant number PINV18-846.

References

[1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, and others, Tensorflow: Large-scale machine learning on heterogeneous distributed systems, *ArXiv Prepr. ArXiv1603.04467*. (2016).

[2] M. Abdel-Basset, L. Abdel-Fatah, and A.K. Sangaiah, Metaheuristic algorithms: A comprehensive review, *Comput. Intell. Multimed. Big Data Cloud with Eng. Appl.* (2018) 185–231.

[3] D. Aquino-Brítez, A. Ortiz, J. Ortega, J. León, M. Formoso, J.Q. Gan, and J.J. Escobar, Optimization of Deep Architectures for EEG Signal Classification: An AutoML Approach Using Evolutionary Algorithms, *Sensors*. **21** (2021) 2096.

[4] V.E.C. Benítez, I.C. Matto, J.C.M. Román, J.L.V. Noguera, M. García-Torres, J. Ayala, D.P. Pinto-Roa, P.E. Gardel-Sotomayor, J. Facon, and S.A. Grillo, Dataset from fundus images for the study of diabetic retinopathy, *Data Br.* (2021) 107068.

[5] D. Whitley, A genetic algorithm tutorial, Statistics and computing **4.2** (1994) 65-85.

[6] L.N. de Castro, Fundamentals of natural computing: an overview, *Phys. Life Rev.* (2007). doi:10.1016/j.plrev.2006.10.002.

[7] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A.L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, *IEEE Trans. Pattern Anal. Mach. Intell.* **40** (2017) 834–848.

[8] F. Chollet, Xception: Deep learning with depthwise separable convolutions, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017: pp. 1251–1258.

[9] T.A. Ciulla, A.G. Amador, and B. Zinman, Diabetic retinopathy and diabetic macular edema: pathophysiology, screening, and novel therapies, *Diabetes Care*. **26** (2003) 2653–2664.

[10] L.P. Cunha, E.A. Figueiredo, H.P. Araújo, L.V.F.

Costa-Cunha, C.F. Costa, J. de M.C. Neto, A.M.F. Matos, M.M. de Oliveira, M.G. Bastos, and M.L.R. Monteiro, Non-mydriatic fundus retinography in screening for diabetic retinopathy: agreement between family physicians, general ophthalmologists, and a retinal specialist, *Front. Endocrinol. (Lausanne).* **9** (2018) 251.

[11] T. Domhan, J.T. Springenberg, and F. Hutter, Speeding up automatic hyperparameter optimization of deep neural networks by extrapolation of learning curves, in: Twenty-Fourth Int. Jt. Conf. Artif. Intell., 2015.

[12] D.S. Fong, L. Aiello, T.W. Gardner, G.L. King, G. Blankenship, J.D. Cavallerano, F.L. Ferris, and R. Klein, Retinopathy in diabetes, *Diabetes Care*. **27** (2004) s84--s87.

[13] F.-A. Fortin, F.-M. De Rainville, M.-A.G. Gardner, M. Parizeau, and C. Gagné, DEAP: Evolutionary algorithms made easy, *J. Mach. Learn. Res.* **13** (2012) 2171–2175.

[14] E. Galván, and P. Mooney, Neuroevolution in deep neural networks: Current trends and future challenges, *IEEE Trans. Artif. Intell.* (2021).

[15] J. Handl, S.C. Lovell, and J. Knowles, Multiobjectivization by decomposition of scalar cost functions, in: Int. Conf. Parallel Probl. Solving from Nat., 2008: pp. 31–40.

[16] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016: pp. 770–778.

[17] K. He, X. Zhang, S. Ren, and J. Sun, Identity mappings in deep residual networks, in: Eur. Conf. Comput. Vis., 2016: pp. 630–645.

[18] T. Hwang, Computational Power and the Social Impact of Artificial Intelligence, (2018).

[19] H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma, and W. Qian, An interpretable ensemble deep learning model for diabetic retinopathy disease classification, in: 2019 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2019: pp. 2045–2048.

[20] W. Kerner, and J. Brückel, Definition, classification and diagnosis of diabetes mellitus, *Exp. Clin. Endocrinol. \& Diabetes*. **122** (2014) 384–386.

[21] X. Li, T. Pang, B. Xiong, W. Liu, P. Liang, and T. Wang, Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification, in: 2017 10th Int. Congr. Image Signal Process. Biomed. Eng. Informatics, 2017: pp. 1–11.

[22] K. Liskowski Pawełand Krawiec, Segmenting retinal blood vessels with deep neural networks, *IEEE Trans. Med. Imaging*. **35** (2016) 2369–2380.

[23] F.J. Martinez-Murcia, A. Ortiz, J. Ram\'\irez, J.M. Górriz, and R. Cruz, Deep residual transfer learning for automatic diagnosis and grading of diabetic retinopathy, *Neurocomputing*. (2020).

[24] W.S. McCulloch, and W. Pitts, A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.* (1943). doi:10.1007/BF02478259.

[25] A. Mead, S. Burnett, and C. Davey, Diabetic retinal screening in the UK, *J. R. Soc. Med.* **94** (2001) 127–129.

[26] S. Mirjalili, Genetic algorithm, in: Evol. Algorithms Neural Networks, Springer, 2019: pp. 43–55.

[27] G.B. Orr, and K.-R. Müller, Neural networks: tricks of the trade, Springer, 2003.

[28] W. Rawat, and Z. Wang, Deep convolutional neural networks for image classification: A comprehensive review, *Neural Comput.* **29** (2017) 2352–2449.

[29] M.I. Razzak, S. Naz, and A. Zaib, Deep learning for medical image processing: Overview, challenges and the future, *Classif. BioApps*. (2018) 323–350.

[30] C. Sabanayagam, R. Banu, M.L. Chee, R. Lee, Y.X. Wang, G. Tan, J.B. Jonas, E.L. Lamoureux, C.-Y. Cheng, B.E.K. Klein, and others, Incidence and progression of diabetic retinopathy: a systematic review, *Lancet Diabetes \& Endocrinol.* **7** (2019) 140–149.

[31] M. Safe, J. Carballido, I. Ponzoni, and N. Brignole, On stopping criteria for genetic algorithms, in: Brazilian Symp. Artif. Intell., 2004: pp. 405–413.

[32] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, Mobilenetv2: Inverted residuals and linear bottlenecks, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018: pp. 4510–4520.

[33] Y. Sawaragi, H. NAKAYAMA, and T. TANINO, Theory of multiobjective optimization, Elsevier, 1985.

[34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.* **15** (2014) 1929–1958.

[35] N. Tajbakhsh, J.Y. Shin, S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway, and J. Liang, Convolutional neural networks for medical image analysis: Full training or fine tuning?, *IEEE Trans. Med. Imaging*. **35** (2016) 1299–1312.

[36] L. Torrey, and J. Shavlik, Transfer learning, in: Handb. Res. Mach. Learn. Appl. Trends Algorithms, Methods, Tech., IGI global, 2010: pp. 242–264.

[37] J.W.Y. Yau, S.L. Rogers, R. Kawasaki, E.L. Lamoureux, J.W. Kowalski, T. Bek, S.-J. Chen, J.M. Dekker, A. Fletcher, J. Grauslund, and others, Global prevalence and major risk factors of diabetic retinopathy, *Diabetes Care*. **35** (2012) 556–564.

[38] S.R. Young, D.C. Rose, T.P. Karnowski, S.-H. Lim, and R.M. Patton, Optimizing deep learning hyper-parameters through an evolutionary algorithm, *Proc. Work. Mach. Learn. High-Performance Comput. Environ. - MLHPC '15*. (2015) 1–5. doi:10.1145/2834892.2834896.

[39] W.L. Yun, U. Rajendra Acharya, Y. V Venkatesh, C. Chee, L.C. Min, and E.Y.K. Ng, Identification of different stages of diabetic retinopathy using retinal optical images, *Inf. Sci. (Ny).* **178** (2008) 106–121. doi:https://doi.org/10.1016/j.ins.2007.07.020.

Address for correspondence

José Luis Vázquez Noguera, Universidad Americana, Brasilia 1100, Asunción, Paraguay; E-mail: jose.vazquez@ua.edu.py.