
Study of deep learning techniques for prediction of diabetic retinopathy severity

Jordi de la Torre *

Department of Computer Engineering and Mathematics, Universitat Rovira i Virgili
Tarragona, Spain
jordi.delatorre@gmail.com

1 Introduction

Diabetic Retinopathy (DR) is a leading disabling chronic disease and one of the main causes of blindness and visual impairment in developed countries for diabetic patients. Studies reported that 90% of the cases can be prevented through early detection and treatment. Eye screening through retinal images is used by physicians to detect the lesions related with this disease. Due to the increasing number of diabetic people, the amount of images to be manually analyzed is becoming unaffordable. Moreover, training new personnel for this type of image-based diagnosis is long, because it requires to acquire expertise by daily practice.

Deep Learning is a set of Machine Learning techniques for automatically constructing a model using multiple levels of representation from the underlying distribution of a large set of examples, with the final objective of mapping a high-multidimensional input into a smaller multidimensional output ($f: \mathbb{R}^n \mapsto \mathbb{R}^m, n \gg m$). This mapping allows the classification of multidimensional objects into a small number of categories. The model is composed by many neurons that are organized in layers and blocks of layers, using a cascade of layers in a hierarchical way. Every neuron receives the input from a predefined set of neurons. Every connection has a parameter that corresponds to the weight of the connection. The function of every neuron is to make a transformation of the received inputs into a calculated output value. For every incoming connection, the weight is multiplied by the input value received by the neuron and the aggregated value that used by an activation function that calculates the output of the neuron. The parameters are usually optimized using a stochastic gradient descent algorithm that minimizes a predefined loss function. The parameters of the network are updated after backpropagating the loss function gradients through the network. These hierarchical models are able to learn multiple levels of representation that correspond to different

* PhD advisor: Aïda Valls and Domènec Puig

levels of abstraction, which enables the representation of complex concepts in a compressed way [1], [2], [3], [4].

Quadratic Weighted Kappa (QWK) index is used in many medical diagnosis systems because the diseases have different degrees of severity, which are naturally ordered from mild to the most critical cases. If the diagnose is based on image analysis, the classification is even more difficult because in the interpretation of the image data normally is present some level of subjectivity that sometimes makes the conclusions of different experts to differ [5]. Quadratic Weighted Kappa is able to measure the level of discrepancy of a set of diagnosis made by different raters over the same population [6]. The strength of agreement between the raters is evaluated as a function of the distance between the prediction of both raters. For the case of diabetic retinopathy detection, human expert raters report inter-rater values of QWK about 0.80. This index has been used to evaluate the performance of the predictive model, in comparison with the human experts level.

The work done up to now in the thesis has been centered mainly in two studies. First, the construction of a classifier of diabetic retinopathy severity using the information encoded in the images of patient's retina using deep neural networks [7]. Second, the improvement of the classification quality using a learning approach based on ordinal information of the QWK [8].

2 Diabetic retinopathy detection using deep neural networks

The traditional model of pattern recognition has been based on extracting hand-crafted fixed engineered features or fixed kernels from the image and using a trainable classifier on top of those features to get the final classification. Using this scheme the problem of the DR detection has been based on hand engineering the features for the detection of microaneurism, haemorrhages and exudate in retinal images that maximize the performance of the classifier. This type of approach requires a good understanding of the mechanism of the disease, requires a lot of labor time and is very task-specific and thereby not reusable for other different classification problems.

In this first work of the thesis we explore a completely different approach consisting on automatic feature learning. We use a deep convolutional neural network model for predicting the probability of every one of the five standardised DR severity levels [7]. The model is trained using a logarithmic loss function and stochastic gradient descend optimization based algorithms and a set of data augmentation techniques. The training procedure details can be found in [7]. The study was done with the Kaggle dataset of EyePACS. This image set has about 88.000 retina images, which are labeled by expert physicians.

To improve the classification rate, we use a probabilistic combination of the information that can be obtained from both eyes of the same patient. DR

usually affects both eyes, specially when the illness is in high severity stages. The dataset used is big enough to infer from the frequencies of co-occurrence of the classes, the conditional probabilities of having one class in one eye given another class in the other, $P(Left|Right)$ and $P(Right|Left)$. Being $P(Left)$ and $P(Right)$, the probability distributions obtained by our predictive model with the left image and the right image, respectively, we can estimate P_L and P_R using $P_L = P(Left|Right)P(Right)$ and $P_R = P(Right|Left)P(Left)$. To merge the value obtained from the model with the estimation coming from the other eye, we calculate the arithmetic mean. The class with maximum value is the one selected for each eye.

3 Improving the classification rate with QWK loss function

The optimization of the neural networks for multi-class classification is traditionally done using the logarithmic loss. The logarithmic loss has a very robust probabilistic foundation: minimizing it, is the same as minimizing the logarithmic likelihood, that is equivalent to do a Maximum Likelihood Estimation (MLE) or equivalently, to find the Maximum a Posteriori Probability (MAP), given a uniform prior [9]. This loss function is designed to find perpendicular vectors in the output space. This model is suitable when the output classes are independent, but it may not be good in cases where classes are ordered. This is the case of some disease prediction, where an incremental severity scale is present. Normally in those cases a ordinal regression approach is better.

As Quadratic Weighted Kappa index is designed to evaluate a good ordinal rating, we explored the possibility of substituting the log-loss function by the QWK-loss function in deep neural networks training [8]. We defined the optimization procedure in terms of QWK and we showed that, for DR severity prediction, classification improves in more than a 5%. This method is directly generalizable to other multi-class classification problems where there is a prior known information about the predefined ordering of the classes.

4 Conclusions and future work

With the work done up to now we have been able to model the diabetic retinopathy detection using supervised deep learning techniques. Using the new QWK-loss function, we obtained up to a 5% increase in the classification rates over the standard approach. Moreover, thanks to the probabilistic combination of the results of both eyes we have been able to increase even further the results of the model, being able to reach human expert level performance.

The results of our study show that with the direct optimization of the QWK index allow the consecution of better generalization results in different datasets with ordered output classes. Log-loss has to learn the predefined ordering of the classes from data and this seems to be a disadvantage. Results

showed that, depending on the use case, between 6-10% of improvement can be obtained from the direct optimization of QWK. This is a significant improvement that may be worth specially in medical diagnosis, since an accurate detection of the level of severity of a disease usually has great influence on the treatment prescription and the possibility of minimizing bad consequences of the illness.

Future work will be centered on the finding an human-understandable interpretation of the results given by the model and in the usage of unsupervised learning techniques to make the classification.

Acknowledgements

This work is supported by the URV grant 2015PFR-URV-B2-60 and the Spanish research projects PI15/01150 and PI12/01535 (Instituto Salud Carlos III).

References

- [1] Yann Lecun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 5 2015.
- [2] J. Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.
- [3] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(8):1798–1828, August 2013.
- [4] Yoshua Bengio. Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, 2(1):1–127, 2009.
- [5] George Hripcsak and Daniel F Heitjan. Measuring agreement in medical informatics reliability studies. *Journal of biomedical informatics*, 35(2):99–110, 2002.
- [6] Anthony J Viera, Joanne M Garrett, et al. Understanding interobserver agreement: the kappa statistic. *Fam Med*, 37(5):360–363, 2005.
- [7] Jordi de la Torre, Aida Valls, and Domenec Puig. Diabetic retinopathy detection through image analysis using deep convolutional neural networks. In *Artificial Intelligence Research and Development, Frontiers in Artificial Intelligence and Applications*, volume 288.
- [8] Jordi de la Torre, Domenec Puig, and Aida Valls. Weighted kappa loss function for multi-class classification of ordinal data in deep learning. *Pattern Recognition Letters*, pages –, 2017.
- [9] Kevin P. Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012.