

# Convolutional Neural Network Transfer for Automated Glaucoma Identification

José Ignacio Orlando<sup>a,b</sup>, Elena Prokofyeva<sup>c,d</sup>, Mariana del Fresno<sup>b,e</sup> and Matthew B. Blaschko<sup>f</sup>

<sup>a</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina;

<sup>b</sup>Pladema Institute, Gral. Pinto 399, 7000 Tandil, Argentina;

<sup>c</sup>Scientific Institute of Public Health (WIV-ISP), Brussels, Belgium;

<sup>d</sup>Federal Agency for Medicines and Health Products (FAMHP), Place Victor Horta 40/40,  
1060 Brussels, Belgium;

<sup>e</sup>Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC-PBA),  
Buenos Aires, Argentina;

<sup>f</sup>Center for Processing Speech and Images, Dept. Elektrotechniek, KU Leuven,  
Kasteelpark Arenberg 10, 3001 Leuven, Belgium.

## ABSTRACT

Most current systems for automated glaucoma detection in fundus images rely on segmentation-based features, which are known to be influenced by the underlying segmentation methods. Convolutional Neural Networks (CNNs) are powerful tools for solving image classification tasks as they are able to learn highly discriminative features from raw pixel intensities. However, their applicability to medical image analysis is limited by the non-availability of large sets of annotated data required for training. In this article we present results of analysis of the viability of using CNNs that are pre-trained from non-medical data for automated glaucoma detection. Two different CNNs, namely OverFeat and VGG-S, were applied to fundus images to generate feature vectors. Preprocessing techniques such as vessel inpainting, contrast-limited adaptive histogram equalization (CLAHE) or cropping around the optic nerve head (ONH) area were explored within this framework to evaluate the improvement in feature discrimination, combined with both  $\ell_1$  and  $\ell_2$  regularized logistic regression models. Results on the Drishti-GS1 dataset, evaluated in terms of area under the average ROC curve, suggests the viability of this approach and offer significant evidence of the importance of well-chosen image pre-processing for transfer learning when the amount of data is not sufficient for fine-tuning the network.

**Keywords:** Fundus imaging, Glaucoma, Convolutional Neural Networks,  $\ell_1$  and  $\ell_2$  regularized logistic regression

## 1. INTRODUCTION

Glaucoma is a chronic and irreversible neuro-degenerative condition that is one of the leading causes of preventable blindness in the world.<sup>1</sup> In European countries glaucoma is the second main cause of blindness.<sup>2</sup> The prevalence of blindness due to glaucoma continued to increase in Latin America over the last decade.<sup>3</sup> This disease is a consequence of an accumulation of aqueous humor in the eye due to a defect of its drainage system. This condition progressively elevates the intra-ocular pressure (IOP), affecting the optic nerve and resulting in permanent blindness if left untreated.<sup>4</sup> As glaucoma may be asymptomatic in its early stages, at least half of patients with this eye pathology remain undiagnosed, while more than half of those who are undergoing treatment do not have the disease.<sup>2</sup> As glaucoma is a chronic life long disease, which if not timely diagnosed and treated can lead to blindness, the major challenge is to be able to screen for glaucoma to detect the large number of undiagnosed people.<sup>2</sup> Current screening practice includes the examination of optic nerve head (ONH) through fundus imaging and measurement of the cup to disc ratio (CDR).<sup>5</sup> The CDR is one of the most commonly used diagnostic criteria, even though it can be influenced by factors such as sex, age or race, etc.<sup>6</sup> Inspection of other

---

Further author information: (Send correspondence to J.I.O.)

J.I.O.: E-mail: jiorlando@conicet.gov.ar, Telephone: +54 249 4439690

retinal structures—including nerve fiber defects or vessels, using such imaging modalities as optical coherence or Heidelberg Retinal tomographies—is usually required to confirm the diagnosis.<sup>7</sup>

Automated screening systems for glaucoma detection from fundus photographs are extremely valuable as they facilitate the early diagnosis of the disease, which is critical to prevent its progression.<sup>8</sup> However, most recent strategies are based on analysing properties of the ONH,<sup>9,10</sup> blood vessels<sup>11–13</sup> or the nerve fiber layer,<sup>14,15</sup> for which it is essential to previously detect and segment these regions. As segmentation-based features are known to be significantly influenced by the precision of the underlying segmentation methods,<sup>16</sup> an increasing scientific effort in the field of glaucoma detection is devoted to the development of new strategies based on overall image properties. Such approaches are based on transformations of the image intensities,<sup>17</sup> texture analysis<sup>18</sup> or combinations of image derived features with clinical and genetic information.<sup>19</sup> An important limitation of these strategies, however, is that they rely on hand-crafted features, which require a significant engineering effort to develop,<sup>20</sup> or in data that is too complex to obtain under certain clinical settings.

Convolutional Neural Networks (CNNs) have emerged as a powerful technique for solving computer vision tasks such as image classification<sup>21</sup> and semantic segmentation.<sup>22</sup> These methods are robust enough to learn deep properties of the images that are usually ignored but are relevant for the objective task, avoiding intermediate steps such as segmentation or feature design and selection. However, this advantage is achieved at the cost of learning from extremely large, annotated training sets.<sup>23</sup> Thus, their applicability is restricted to domains that have large amounts of labelled data, which is normally difficult to obtain as image acquisition and labelling require experts and costly devices.<sup>24</sup> To the best of our knowledge, CNNs were only used once for glaucoma detection, in the work by Chen *et al.*<sup>25</sup> Such a method requires one to previously segment the ONH and preprocess the resulting image, and by means of training the CNN on an augmented version of a large local data set, the authors were able to automatically learn relevant features that were suitable for glaucoma identification.

As an alternative to an expensive and dedicated training task, different authors have recently proposed to use off-the-shelf CNNs to solve problems different from those for which they were originally trained.<sup>24,26–28</sup> That is, CNNs that were pre-trained using a large amount of data from a different problem are transferred to extract high-dimensional but discriminative feature vectors from images from a completely different domain. These features are afterwards used to train a dedicated classifier designed to solve the new task. By means of such an approach it is possible to solve different computer vision problems using the same, single model, learned from a sufficiently large, off-the-shelf training set.<sup>29</sup> As the CNN has already learned discriminative features, the dedicated classifier can be trained from a smaller, application-related set. Despite the fact that this strategy looks promising and was already explored for several specific computer vision applications,<sup>26</sup> only few studies describe the application of this concept for medical image classification,<sup>24,27–29</sup> and none for fundus image analysis.

In this article we analyze the viability of using features extracted with pre-trained, off-the-shelf CNNs in the context of glaucoma detection in limited-size fundus image data sets. Different levels of cropping around the ONH area combined with state-of-the-art preprocessing strategies, such as vessel inpainting and contrast enhancement, are explored as inputs to a CNN in order to evaluate its capability to better characterize the disease. Thus, instead of using manually engineered features, we propose to use only simple preprocessing methods to transfer the already trained CNN to this other task.  $\ell_1$  and  $\ell_2$  regularized logistic regression classifiers are trained with these features to obtain a likelihood indicating the presence or the absence of glaucoma in each image. Results on a publicly available set of fundus images suggest that this approach yields significant discrimination between glaucomatous and healthy eyes.

The remainder of this paper is organized as follows. The strategy followed to extract features using the transferred CNN is analyzed in Section 2. Details about the data set and the CNNs used for evaluation are provided in Section 3, while the results obtained are presented in Section 4. Finally, Section 5 concludes the paper.

## 2. METHODS

In this article we analyze the capability of CNNs pre-trained from non-medical data for glaucoma detection in fundus images. For this purpose, we explore how to transfer such networks to this task without relying on an expensive retraining using large amounts of data. In general, this problem is known as transfer learning,<sup>30</sup> and

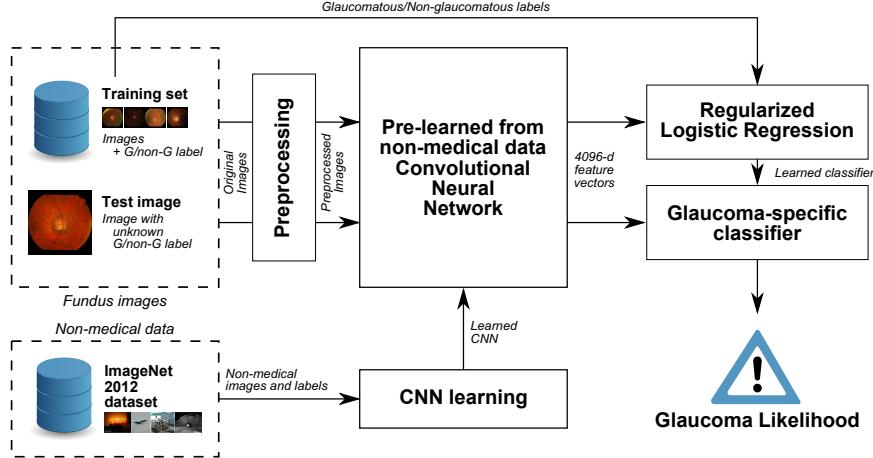


Figure 1: Schematic representation of our method for transferring CNNs pre-trained from non-medical data to glaucoma detection in fundus images.

can be tackled by following one of these two different approaches. One consists of feeding the pre-trained network with the new images, retrieving the outputs of the first fully connected layer as feature vectors,<sup>26</sup> and using them to train a new classifier explicitly devoted to the new task. The second strategy is to not only replace and retrain the classifier on top of the CNN on the new data set, but to also fine-tune the weights of the pretrained network by continuing the backpropagation process. Depending on the new task, it is possible to fine-tune all the layers of the net or to keep some of the earlier ones fixed, restricting the fine-tuning to some of the higher-level layers of the network.<sup>29</sup> This is motivated by the fact that the earlier features of a CNN contain generic features that might be useful to many tasks (such as edge, color or texture detectors), but later layers become progressively more specific to the details of the classes contained in the original dataset. However, fine-tuning requires a sufficiently large sample of data to avoid overfitting, a setting that cannot be accomplished with current publicly available sets for glaucoma detection.

The alternative approach of using the output features of the pre-trained network (known as *CNN codes*) and train an offline classifier based on them usually suffers if the images significantly differ from the original ones. In order to reduce the effect of this issue, we propose to analyze the usage of different preprocessing strategies before feeding the network with the images. If the preprocessing strategy enhances image characteristics that are related to the presence of glaucoma, it is expected that this approach can contribute to the computation of better features. A schematic diagram of our approach is illustrated in Figure 1. We describe both the preprocessing methods and the classifier in the sequel.

## 2.1 Image preprocessing

Fundus images are preprocessed with state of the art techniques before extracting CNN codes. Figure 2 shows examples of all the operations we consider. In particular, we evaluate the contribution of cropping around the area of the ONH, applying contrast enhancement (CLAHE), and including or removing the vessels.

**Zooming around the ONH** - Glaucoma manifests specifically on the ONH region.<sup>17</sup> As images are usually downscaled before feeding a CNN, it can be hypothesized that the network might not be able to capture enough information about the existence of the disease due to the reduced size of the ONH. Or, on the contrary, it might be possible that the CNN can better characterize glaucoma by capturing information from other non-ONH structures of the retina, such as the vascular tree. Several experiments considering different zooms of the original fundus images, centered around the ONH, were carried out to validate one of these hypotheses. 4 different zooms were considered (Fig. 2): the original image as it is, including the blank regions outside the field of view (FOV); a squared crop of the area inside the FOV; the peripapillary area (PPA), which involves the ONH and its surrounding zone; and the ONH itself.

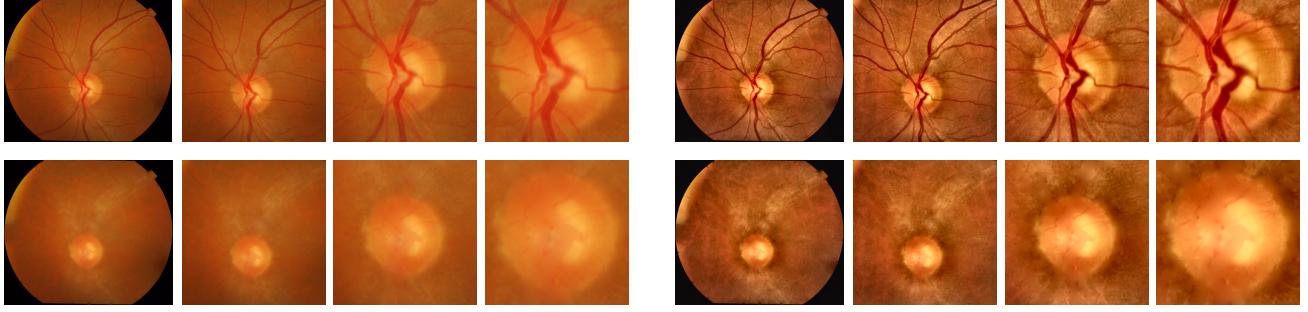


Figure 2: Preprocessing strategies evaluated. First group: without CLAHE. Second group: with CLAHE. From left to right: original image, cropped FOV, peripapillary area (PPA), and ONH. First row: original images. Second row: images after vessel inpainting.

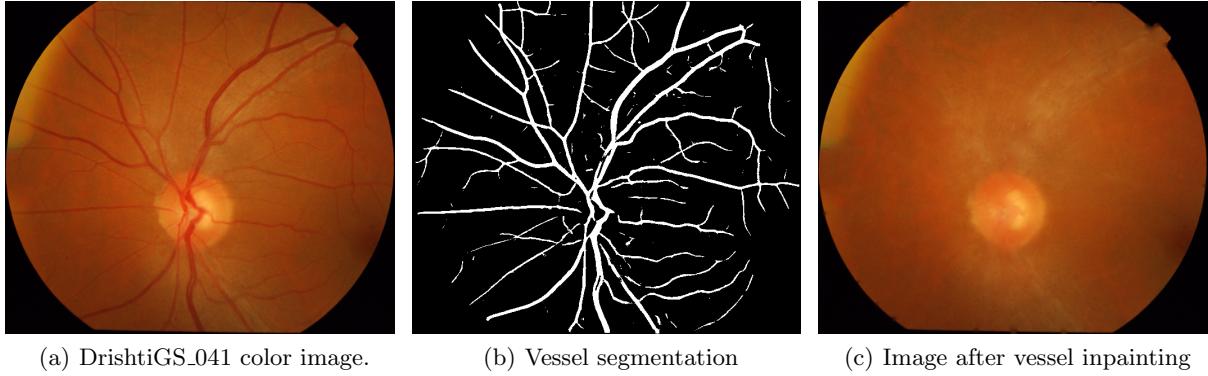


Figure 3: Vessel subtraction.

**CLAHE enhancement** - Contrast-Limited Adaptive Histogram Equalization (CLAHE) is an enhancement operation that is widely applied in retinal image analysis due to its capability to improve the contrast and illumination of the fundus<sup>31</sup> (Fig 2). Additionally, it is known to benefit retinal nerve fiber characterization.<sup>32,33</sup> Performances with and without applying CLAHE were evaluated.

**Vessel subtraction** - Though there are publications indicating a certain correlation between glaucoma and changes in the vascular structures,<sup>13</sup> other authors suggest to remove them before extracting features.<sup>8,17,34,35</sup> Based on such assumptions, we obtain features from images with and without the vessels inpainted. The vascular tree is segmented using a modified version of a method based on fully-connected conditional random fields.<sup>36,37</sup> Pixel level features are extracted from the green band of the RGB images, after CLAHE enhancement, as this color band exhibits a better contrast between vessels and the remaining parts of the retina compared with the red and blue channels.<sup>37</sup> Instead of using the feature set originally proposed, we use responses to line detectors<sup>38</sup> and to B-COSFIRE filters<sup>31</sup> as unary features, and Zana and Klein vessel enhanced images<sup>39</sup> as pairwise features. We observed that this combination of features empirically achieved better binary representations of the vascular trees. Vessels are removed from the original color image by applying inward interpolation on each of the color bands: new intensities are smoothly interpolated inward from the pixels values on the outer boundary of the vascular binary mask. As this strategy is affected if the segmentation method underestimated the actual vessel calibre (because the iterative inpainting might reintroduce the vasculature to the image instead of removing it), the vascular masks are morphologically dilated using disks of size 3. Fig 3 illustrates a segmentation of the vascular tree and the resulting image after removing it.

## 2.2 $\ell_1$ and $\ell_2$ regularized logistic regression

Glaucoma detection is a binary classification task that we estimate by means of a supervised learning approach known as regularized logistic regression.<sup>40</sup> Let  $S$  be a training set composed of  $n$  training instances  $\{(\mathbf{x}_i, y_i), i =$

$1, \dots, n\}$ . Each  $\mathbf{x}_i \in \mathbb{R}^d$  is a  $d$ -dimensional feature vector extracted from the  $i$ -th image using a CNN, and  $y_i \in \{-1, +1\}$  represents its corresponding class label,  $+1$  indicating glaucomatous and  $-1$  non-glaucomatous. Logistic regression can be written as a regularized risk minimization with logistic loss. The objective function to be minimized is:

$$\hat{\beta} = \arg \min_{\beta} \lambda \Omega(\beta) + \sum_{i=1}^n \log(1 + e^{-y_i \langle \beta, \mathbf{x}_i \rangle}) \quad (1)$$

where  $\beta$  is a coefficient vector of the linear discriminant function,  $\lambda \geq 0$  is a scalar parameter controlling the degree of regularization by the regularizer  $\Omega : \mathbb{R}^d \mapsto \mathbb{R}_+$  and  $\langle \cdot, \cdot \rangle$  is the canonical inner product in  $\mathbb{R}^d$ . Two well-known regularizers were used as  $\Omega$ , the  $\ell_1 = \|\cdot\|_1$  and  $\ell_2 = \|\cdot\|_2$  norms.  $\ell_1$  imposes the sum of the absolute values of the parameters  $\beta$  to be small, encouraging sparse parameter vectors.<sup>41</sup> This setting makes this regularizer suitable for simultaneous learning and feature selection when the size of the training set  $n$  is small and the size of the feature vector  $d$  is high, as in this case. By contrast,  $\ell_2$  regularization may not be effective when only a few features are relevant.<sup>41</sup> Both norms were experimentally validated to evaluate if the high dimensional CNN codes are discriminative by themselves, or if only some of their components are relevant for glaucoma identification. The likelihood of a given image having glaucoma or not is obtained by the linear discriminant function  $\hat{y} = \langle \beta, \mathbf{x}_i \rangle$ . An existing implementation of k-support regularized logistic regression<sup>42, 43</sup> was used to minimize the objective function, assigning  $k = 1$  for the  $\ell_1$  norm and  $k = d$  for the  $\ell_2$  norm.

### 3. MATERIALS

#### 3.1 Convolutional neural networks

Two different CNNs were used in our experiments, namely OverFeat<sup>21\*</sup> and VGG-S<sup>44†</sup>. Both networks were initially trained on the *ImageNet 2012* training set, which comprises 1.2 million natural, non-medical color images labeled with 1000 classes, with the second one reporting the highest performance in the PASCAL-VOC competition.

OverFeat architecture is inspired by the work of Krizhevsky *et al.*,<sup>45</sup> with minor changes described by Sermanet *et al.*<sup>21</sup> We use the accurate version of their feature extractor, which is composed of 6 convolutional layers, with filters sizes from  $7 \times 7$  to  $3 \times 3$  containing from 96 to 1024 kernels. Techniques such as max-pooling, rectification non-linearities and dropout are used, as described in the original reference. A similar architecture was introduced by Chatfield *et al.*,<sup>44</sup> known as VGG-S. It comprises 5 convolutional layers instead of 6, with filters sizes similar to those in OverFeat but with 512 kernels in the last 3 convolutional layers. The last three layers of both CNNs are fully connected, and they are responsible of scoring the likelihood of each of the classes. As our purpose is to use these CNNs as feature extractors, the output of the first fully-connected layer is used here to retrieve 4096-d feature vectors from input images. Images have to be rescaled to  $231 \times 231$  or  $224 \times 224$  before feeding Overfeat and VGG-S, respectively.

#### 3.2 Data sets

Experiments were carried out on the data set Drishti-GS1.<sup>46</sup> It consists of 101 fundus images, taken with eyes dilated, centered on the ONH, at  $30^\circ$  FOV, with  $2896 \times 1944$  pixels and saved as uncompressed PNG images. All images were originally collected at the Aravind Eye Hospital, India. Glaucoma patient selection was done by clinical investigators based on clinical findings during examination. Selected patients were 40-80 years of age, with a roughly equal number of males and females. Patients undergoing a routine refraction test and not found to be glaucomatous were chosen to represent the normal class. The diagnostic opinion for each image (normal or glaucomatous) was obtained from 4 glaucoma experts based only on the images, and the provided gold standard was derived as the majority opinion, i.e. 3 out of the 4 experts.

For vessel inpainting, vascular trees were segmented with a model trained on the DRIVE<sup>47</sup> training set and calibrated using the DRIVE test set. After model selection, parameters were set to  $C = 10^2/c$ , with  $c$  equal to the number of training pixels, and  $\theta_p = 5$  (refer to the original papers for additional details<sup>36, 37</sup>). Feature

---

\*<http://cilvr.nyu.edu/doku.php?id=software:overfeat:start>

†<http://www.vlfeat.org/matconvnet/pretrained/>

parameters were fixed to the values provided in the original references. As features were originally designed for DRIVE and depend on the resolution of the images—which significantly differs between data sets—Drishti-GS1 images were downsampled by a factor of 2.96 before segmentation so vessels appear at a similar resolution to DRIVE. As images are then further downsampled before applying CNNs, this process does not negatively affect the end result.

As all these data sets do not include FOV masks, which are necessary for further processing, we automatically generate them by thresholding the luminosity plane of the CIELab version of the RGB images at a value of 0.01.<sup>31</sup> If the resulting binary mask is all ones, an alternative approach is applied to estimate the FOV mask, where the RGB bands are summed up and the resulting image is thresholded at 150 level. To smooth borders and reduce noise, the mask is postprocessed with a median filter using square windows of side 5, and only its largest connected component is preserved.

## 4. RESULTS

All the preprocessing strategies were analyzed in combination with the OverFeat network with the purpose of determining which of them is the best. Due to the relatively small size of the available data set, each configuration is evaluated with 200 trials as follows. First, images are randomly divided into training (70%) and test sets (30%), as suggested in the literature.<sup>40</sup> 30% of the training images are randomly selected and separated as the validation set, which is used to determine the value of the regularization parameter  $\lambda = 10^i$  of the logistic regression classifier, with  $i \in \{-5, -4, \dots, 6\}$ . The area under the ROC curve (AUC) is used as a numerical indicator of the overall model performance during model selection. Two different techniques for augmenting the training sets are evaluated: images are rotated at two different angles ( $90^\circ$  and  $45^\circ$ ) until reaching  $360^\circ$ , and horizontally flipped, incrementing the size of the sets by a factor of 8 and 16, respectively. Performance of each configuration is measured in terms of the area under the average ROC curve, obtained with the Fawcett's method.<sup>48</sup> Results are presented in Figure 4 and analyzed in the sequel.

If vessels are not inpainted, OverFeat codes achieve the best results with images cropped in the FOV region, independently of the regularizer used. By contrast, if vessels are removed from the images, then both norms report the best performance when images are zoomed in the PPA. In any of these cases it is not possible to determine which is the best augmentation technique, as it varies from one configuration to another. In the case of non-inpainted cropped FOV images, the best AUC (0.7626) is obtained by OverFeat features with images augmented using  $90^\circ$  rotations; when using inpainting, the largest value (0.7212) is achieved without augmentation. CLAHE enhancement is ill advised in most of the analyzed configurations, as it significantly decreases the AUC values. It only demonstrates a positive contribution under vessel inpainting, in the cropped FOVs and the ONHs, and when using the  $\ell_1$  norm with features obtained from the original images. The  $\ell_2$  norm performs equally or better than the  $\ell_1$  for almost all the settings. With regards to vessels, their subtraction does not offer any improvement in results. In fact, most of the AUC values are decreased when this operation is used. The only exception can be observed when focusing on the ONH and applying the CLAHE operation. An additional experiment, not included in the figure, was made feeding the classifiers with OverFeat codes extracted from the raw binary segmentations of the vessels, but AUCs obtained did not significantly differ from chance performance (the highest was 0.5303 in the  $90^\circ$  augmented version).

An additional experiment was performed using VGG-S, feeding it with images preprocessed with the best strategy we found when using OverFeat (RGB images cropped inside the FOV without CLAHE, without inpainting,  $90^\circ$  augmentation; and RGB images cropped around the peripapillary area without CLAHE, with vessels inpainted and without data augmentation). The purpose of this analysis was to determine if slight changes in the architecture of the CNN might affect the performance of the transferred network. In both experiments we used  $\ell_2$  regularized logistic regression due to its better results while evaluating preprocessing techniques. We observed that OverFeat features performed better than VGG-S both in images with (AUC= 0.7212 vs. AUC= 0.6655, respectively) and without vessels inpainted (AUC= 0.7626 vs. AUC= 0.7180, respectively).

Extensive comparison with respect to other state of the art strategies is not feasible as most of them used their own private data sets. To the best of our knowledge, only the work by Chakrabarty *et al.*<sup>20</sup> was evaluated on DRISHTI-GS1, reporting an AUC value of 0.78. Such an approach is based on a classification framework that

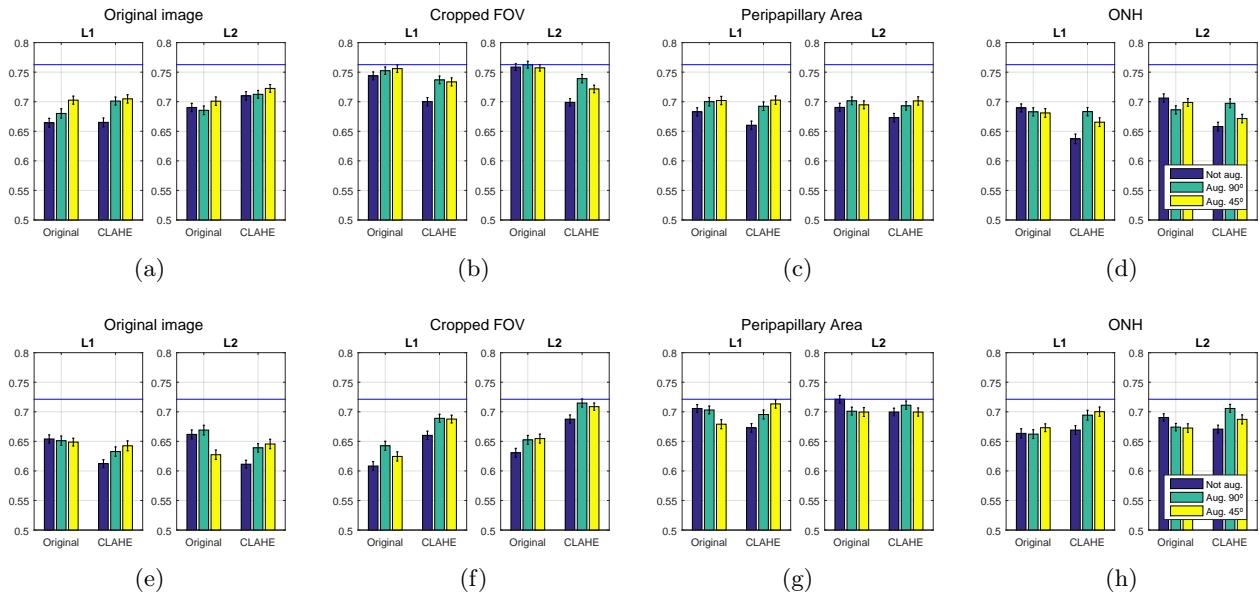


Figure 4: Area under the average ROC curves obtained from 200 trials<sup>48</sup> using OverFeat features with  $\ell_1$  or  $\ell_2$  regularized logistic regression, respectively. First row: using images without vessel inpainting. Second row: using images with vessel inpainting. The blue line corresponds to the best AUC value for the row.

fuses both overall image and segmentation based features within a co-training based semi-supervised setting. The main difference with our work is that we only rely on features from the cropped version of the original fundus image, and we do not integrate any information extracted from ONH segmentations. Additionally, a larger training set was used for learning the classification model reported in that work.

## 5. DISCUSSION

In this work we have analyzed the capability of transferring pre-trained, off-the-shelf CNNs for glaucoma detection using fundus photographs. A variety of state-of-the-art preprocessing techniques, including different crops around the ONH region, were used to feed the neural networks and to recover features that are then used in two different regularized logistic regression models. Performance on cropped FOVs was higher than in other crops when using OverFeat, indicating that such a network might be able to recover valuable information from other regions than the ONH. This evidence is also supported by the low AUC values achieved when vessels are inpainted. These results contradict other studies that focused on the extraction of features only in the PPA or in the ONH after removing the vessels.<sup>17, 49</sup> The negligible AUC values obtained using only the vasculature, however, suggest that vessels are not useful by themselves but only when combined with other sources of information in the fundus. CLAHE enhancement, which is known to improve the contrast of the fundus and is widely applied, increases the AUC values under some settings. However, in most of the non-inpainted cases, it is not helpful. By contrast, when vessels are removed, CLAHE benefits results achieved on cropped FOVs and in the ONH, and also in the PPA when  $\ell_1$  regularization is used. This might be associated with an improvement in the discrimination of the optic cup and the ONH, though such an enhancement is not competitive with results achieved using other settings.  $\ell_1$  regularization is known as a suitable tool for simultaneous feature selection and classification, as it favors sparse coefficient vectors, a feature that is not provided by  $\ell_2$  regularization, which tends to use all the coefficients. As the  $\ell_2$  norm performs equal to or better than the  $\ell_1$  norm in most of the experiments, it can be conjectured that all the coefficients in the feature vectors are useful for this task. However, it is worth mentioning that the improvement obtained combining CLAHE with the  $\ell_1$  norm can be associated with an improvement of certain coefficients in the feature vector. Additionally, we observed that in this data set VGG-S features do not perform better than OverFeat ones.

In conclusion the preprocessing strategy, although being simple, supports transfer learning from small data sets where fine-tuning via retraining is not possible. In terms of the overall performance of our method, we achieved a competitive area under the ROC curve score with respect to other existing strategies,<sup>20</sup> but with the advantage of not requiring the segmentation of the optic cup and the optic disc regions, and training from a smaller data set. Further experiments will be made in the future in order to mix transferred CNN codes with ONH features. New experiments with other preprocessing techniques for the automated detection of other retinal diseases such as diabetic retinopathy will also be explored. Additional evaluation using larger data sets is also encouraged. Finally, we are pleased to release our Matlab code for preprocessing and experiments at <https://github.com/ignaciotorlano/overfeat-glaucoma>.

### Acknowledgments

This work is partially funded by Internal Funds KU Leuven, FP7-MC-CIG 334380 and ANPCyT PICT 2014-1730. J.I.O. is funded by a doctoral scholarship granted by CONICET. We thank the authors that provided us with their code<sup>31, 38, 42, 43</sup> and the CNNs.<sup>21, 44</sup>

### REFERENCES

- [1] Abràmoff, M. D., Garvin, M. K., and Sonka, M., “Retinal imaging and image analysis,” *Biomedical Engineering, IEEE Reviews in* **3**, 169–208 (2010).
- [2] Prokofyeva, E. and Zrenner, E., “Epidemiology of major eye diseases leading to blindness in Europe: A literature review,” *Ophthalmic Research* **47**(4), 171–188 (2012).
- [3] Leasher, J. L., Lanssingh, V., Flaxman, S. R., Jonas, J. B., Keeffe, J., Naidoo, K., Pesudovs, K., Price, H., Silva, J. C., White, R. A., et al., “Prevalence and causes of vision loss in Latin America and the Caribbean: 1990–2010,” *British Journal of Ophthalmology* (2014).
- [4] Tham, Y.-C., Li, X., Wong, T. Y., Quigley, H. A., Aung, T., and Cheng, C.-Y., “Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis,” *Ophthalmology* **121**(11), 2081–2090 (2014).
- [5] Spaeth, G. L., “Appearances of the optic disc in glaucoma: A pathogenetic classification,” in [*Symposium on Glaucoma. Transactions of the New Orleans Academy of Ophthalmology*], 114–152 (1981).
- [6] Hoffmann, E. M., Zangwill, L. M., Crowston, J. G., and Weinreb, R. N., “Optic disk size and glaucoma,” *Survey of Ophthalmology* **52**(1), 32–49 (2007).
- [7] Kotowski, J., Wollstein, G., Ishikawa, H., and Schuman, J. S., “Imaging of the optic nerve and retinal nerve fiber layer: An essential part of glaucoma diagnosis and monitoring,” *Survey of Ophthalmology* **59**(4), 458–467 (2014).
- [8] Haleem, M. S., Han, L., van Hemert, J., and Li, B., “Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: A review,” *Computerized Medical Imaging and Graphics* **37**(7), 581–596 (2013).
- [9] Joshi, G. D., Sivaswamy, J., and Krishnadas, S., “Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment,” *Medical Imaging, IEEE Transactions on* **30**(6), 1192–1205 (2011).
- [10] Köhler, T., Bock, R., Hornegger, J., and Michelson, G., “Computer-aided diagnostics and pattern recognition: Automated glaucoma detection,” in [*Teleophthalmology in Preventive Medicine*], 93–104, Springer (2015).
- [11] Koh, V., Cheung, C. Y.-l., Zheng, Y., Wong, T. Y., Wong, W., and Aung, T., “Relationship of retinal vascular tortuosity with the neuroretinal rim: The Singapore Malay eye study,” *Investigative Ophthalmology & Visual Science* **51**(7), 3736–3741 (2010).
- [12] Xu, Y., Liu, J., Lin, S., Xu, D., Cheung, C. Y., Aung, T., and Wong, T. Y., “Efficient optic cup detection from intra-image learning with retinal structure priors,” in [*Medical Image Computing and Computer-Assisted Intervention—MICCAI*], 58–65, Springer (2012).
- [13] Wu, R., Cheung, C. Y.-L., Saw, S. M., Mitchell, P., Aung, T., and Wong, T. Y., “Retinal vascular geometry and glaucoma: The Singapore Malay eye study,” *Ophthalmology* **120**(1), 77–83 (2013).

- [14] Muramatsu, C., Hayashi, Y., Sawada, A., Hatanaka, Y., Hara, T., Yamamoto, T., and Fujita, H., “Detection of retinal nerve fiber layer defects on retinal fundus images for early diagnosis of glaucoma,” *Journal of Biomedical Optics* **15**(1), 016021 (2010).
- [15] Odstrcilik, J., Kolar, R., Jan, J., Gazarek, J., Kuna, Z., and Vodakova, M., “Analysis of retinal nerve fiber layer via Markov random fields in color fundus images,” in [*Systems, Signals and Image Processing (IWSSIP), 19th International Conference on*], 504–507, IEEE (2012).
- [16] Bekkers, E. J., Zhang, J., Duits, R., and ter Haar Romeny, B. M., “Curvature based biomarkers for diabetic retinopathy via exponential curve fits in  $SE(2)$ ,” in [*Proceedings of the Ophthalmic Medical Image Analysis Second International Workshop, OMIA*], Trucco, E., Chen, X., Garvin, K., M., Liu, J. J., and Frank, X. Y., eds. (2015).
- [17] Bock, R., Meier, J., Nyúl, L. G., Hornegger, J., and Michelson, G., “Glaucoma risk index: Automated glaucoma detection from color fundus images,” *Medical Image Analysis* **14**(3), 471–481 (2010).
- [18] Mohammad, S. and Morris, D. T., “Texture analysis for glaucoma classification,” in [*BioSignal Analysis, Processing and Systems (ICBAPS), International Conference on*], 98–103, IEEE (2015).
- [19] Xu, Y., Duan, L., Wong, D. W. K., Wong, T. Y., and Liu, J., “Glaucoma detection by learning from multiple informatics domains,” in [*Proceedings of the Ophthalmic Medical Image Analysis Second International Workshop, OMIA*], Trucco, E., Chen, X., Garvin, K., M., Liu, J. J., and Frank, X. Y., eds. (2015).
- [20] Chakravarty, A. and Sivaswamy, J., “Glaucoma classification with a fusion of segmentation and image-based features,” in [*Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*], **1**, 689–692, IEEE (2016).
- [21] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., and LeCun, Y., “Overfeat: Integrated recognition, localization and detection using convolutional networks,” in [*International Conference on Learning Representations*], (2014).
- [22] Shakeri, M., Tsogkas, S., Ferrante, E., Lippe, S., Kadoury, S., Paragios, N., and Kokkinos, I., “Sub-cortical brain structure segmentation using F-CNN’s,” in [*Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*], 269–272 (2016).
- [23] Jaderberg, M., Vedaldi, A., and Zisserman, A., “Speeding up convolutional neural networks with low rank expansions,” in [*Proceedings of the British Machine Vision Conference*], BMVA Press (2014).
- [24] Bar, Y., Diamant, I., Wolf, L., and Greenspan, H., “Deep learning with non-medical training used for chest pathology identification,” in [*SPIE Medical Imaging*], 94140V (2015).
- [25] Chen, X., Xu, Y., Yan, S., Wong, D. W. K., Wong, T. Y., and Liu, J., “Automatic feature learning for glaucoma detection based on deep learning,” in [*Medical Image Computing and Computer-Assisted Intervention-MICCAI*], 669–677, Springer (2015).
- [26] Sharif Razavian, A., Azizpour, H., Sullivan, J., and Carlsson, S., “CNN features off-the-shelf: An astounding baseline for recognition,” in [*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*], 806–813 (2014).
- [27] Ciompi, F., de Hoop, B., van Riel, S. J., Chung, K., Scholten, E. T., Oudkerk, M., de Jong, P. A., Prokop, M., and van Ginneken, B., “Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box,” *Medical Image Analysis* **26**(1), 195–202 (2015).
- [28] van Ginneken, B., Setio, A. A., Jacobs, C., and Ciompi, F., “Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans,” in [*Biomedical Imaging (ISBI), IEEE 12th International Symposium on*], 286–289, IEEE (2015).
- [29] Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., and Summers, R. M., “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *Medical Imaging, IEEE Transactions on* **35**(5), 1285–1298 (2016).
- [30] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H., “How transferable are features in deep neural networks?,” in [*Advances in Neural Information Processing Systems*], 3320–3328 (2014).
- [31] Azzopardi, G., Strisciuglio, N., Vento, M., and Petkov, N., “Trainable COSFIRE filters for vessel delineation with application to retinal images,” *Medical Image Analysis* **19**(1), 46–57 (2015).

- [32] Kolar, R., Tornow, R. P., Laemmer, R., Odstrcilik, J., Mayer, M. A., Gazarek, J., Jan, J., Kubena, T., and Cernosek, P., “Analysis of visual appearance of retinal nerve fibers in high resolution fundus images: a study on normal subjects,” *Computational and mathematical methods in medicine* **2013** (2013).
- [33] Odstrcilik, J., Kolar, R., Tornow, R.-P., Jan, J., Budai, A., Mayer, M., Vodakova, M., Laemmer, R., Lamos, M., Kuna, Z., et al., “Thickness related textural properties of retinal nerve fiber layer in color fundus images,” *Computerized Medical Imaging and Graphics* **38**(6), 508–516 (2014).
- [34] Meier, J., Bock, R., Michelson, G., Nyúl, L. G., and Hornegger, J., “Effects of preprocessing eye fundus images on appearance based glaucoma classification,” in [*Computer Analysis of Images and Patterns*], 165–172, Springer (2007).
- [35] Morales, S., Naranjo, V., Angulo, J., and Alcañiz, M., “Automatic detection of optic disc based on PCA and mathematical morphology,” *Medical Imaging, IEEE Transactions on* **32**(4), 786–796 (2013).
- [36] Orlando, J. I. and Blaschko, M. B., “Learning fully-connected CRFs for blood vessel segmentation in retinal images,” in [*MICCAI*], 634–641 (2014).
- [37] Orlando, J. I., Prokofyeva, E., and Blaschko, M. B., “A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images,” *Biomedical Engineering, IEEE Transactions on* (2016).
- [38] Nguyen, U. T., Bhuiyan, A., Park, L. A., and Ramamohanarao, K., “An effective retinal blood vessel segmentation method using multi-scale line detection,” *Pattern Recognition* **46**(3), 703–715 (2013).
- [39] Zana, F. and Klein, J.-C., “Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation,” *Image Processing, IEEE Transactions on* **10**(7), 1010–1019 (2001).
- [40] Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., and Tibshirani, R., [*The elements of statistical learning*], vol. 2, Springer (2009).
- [41] Ng, A. Y., “Feature selection, L1 vs. L2 regularization, and rotational invariance,” in [*Proceedings of the International Conference on Machine Learning*], 78 (2004).
- [42] Sidahmed, H., Prokofyeva, E., and Blaschko, M. B., “Discovering predictors of mental health service utilization with  $k$ -support regularized logistic regression,” *Information Sciences* **329**, 937–949 (2016).
- [43] Argyriou, A., Foygel, R., and Srebro, N., “Sparse prediction with the  $k$ -support norm,” in [*Advances in Neural Information Processing Systems*], 1457–1465 (2012).
- [44] Chatfield, K., Simonyan, K., Vedaldi, A., and Zisserman, A., “Return of the devil in the details: Delving deep into convolutional nets,” in [*Proceedings of the British Machine Vision Conference*], BMVA Press (2014).
- [45] Krizhevsky, A., Sutskever, I., and Hinton, G. E., “ImageNet classification with deep convolutional neural networks,” in [*Advances in Neural Information Processing Systems*], 1097–1105 (2012).
- [46] Sivaswamy, J., Krishnadas, S., and Chakravarty, A., “A comprehensive retinal image dataset for the assessment of glaucoma from the optic nerve head analysis,” *JSM Biomedical Imaging Data Papers* **2**(1), 1004 (2015).
- [47] Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., and Abramoff, M. D., “Comparative study of retinal vessel segmentation methods on a new publicly available database,” in [*Medical Imaging*], 648–656, International Society for Optics and Photonics (2004).
- [48] Fawcett, T., “An introduction to ROC analysis,” *Pattern Recognition Letters* **27**(8), 861–874 (2006).
- [49] Bock, R., Meier, J., Michelson, G., Nyúl, L. G., and Hornegger, J., “Classifying glaucoma with image-based features from fundus photographs,” in [*Pattern Recognition*], 355–364, Springer (2007).