RETINAL IMAGE ANALYSIS USING LEARNED IMAGE PRIORS

Dwarikanath Mahapatra

Department of Computer Science, ETH Zurich, Switzerland.
*dwarikanath.mahapatra@inf.ethz.ch

ABSTRACT

We present a Field of Experts (FoE) model to segment the optic cup and disc from retinal fundus images. FoE learns a prior model of the cup and neuro-retinal rim region in the form of filters, without relying on hand crafted features. FoE filter responses are integrated into a Random forest (RF) classifier that outputs probability maps for the test image and finally segments the cup and rim. Experimental results show our method achieves significantly better performance than current state-of-the-art.

Index Terms— Field of Experts, glaucoma, optic disc, optic cup, segmentation, CDR.

1. INTRODUCTION

Glaucoma is a chronic disease affecting the optic nerve, resulting in its progressive damage and elongation of the optic cup. Its timely detection can check disease progress. Hence it is essential to have a fast and robust method for detecting onset of glaucoma, as most patients are unaware of their condition until considerable damage has been done [1]. Ophthalmologists identify optic cup-to-disc ratio (CDR) as an important indicator of glaucoma progression. However manual delineation of the disc and cup boundary (for calculating CDR) and subsequent assessment is subjective, time consuming and cost prohibitive. This necessitates the development of automated techniques for cup and disc segmentation.

Automatic CDR measurement involves: 1) optic disc localization and segmentation; and 2) optic cup segmentation. Current state-of-the-art methods for disc segmentation use morphological features [2] and active contours [3]. Their performance depends upon initialization and ability to identify weak edges. Machine learning methods [4] have gained importance as they provide a powerful tool for feature classification. Success of learning based methods depends on hand designed features which are not always stable for different datasets. Therefore recent research in other domains have focused on methods to learn discriminative features from training data [5, 6]. Optic cup detection without depth information is challenging. However different methods have used level sets [7] and superpixels [8, 4] for cup segmentation. Hand crafted features in each case limits their applicability to different datasets.

We propose a Field of Experts (FoE) model for learning image priors of the optic cup and neuro-retinal rim. FoE extends the traditional Markov random field (MRF) models by learning potential functions over extended neighborhoods, without explicitly defining them. FoE priors do not rely on hand designed features and learn spatial filters that are characteristic of the different regions of the retina such as cup, rim and background. This approach overcomes stability issues of hand designed features. The response of learned filters to patches is used to train a Random forest (RF) classifier that generates probability maps for the test image and the final segmentation is obtained using graph cuts.

2. METHODS

2.1. Field Of Expert Priors

FoE learns rich image priors using MRFs and sparse coding. MRFs are limited in capturing image statistics because of hand crafted clique potentials and small neighborhood systems. The FoE model [9] is a more general representation of image priors and generalizes beyond the training data. The MRF has overlapping cliques whose potentials are modeled as filters represented as a Product of Experts (PoE) [10].

Sparse coding methods represent an image patch as a linear combination of learned filters or bases $J_i$ as

$$\min_{a, J} E(a, J) = \sum_c \left\| x_c - \sum_i a_{i,c} J_i \right\|^2 + \lambda \sum_{i,c} S(a_{i,c}) \tag{1}$$

where $x_c$ is an image patch and $S$ is a sparseness prior that penalizes non-zero coefficients $a_{i,c}$. The limitation of this formulation is for an image patch with $n$ pixels it is impossible to find $n$ fully independent linear components.

PoE models overcome this limitation by modeling high dimensional probability distributions with the product of several expert distributions that individually model a low dimensional subspace. The PoE model comprising individual ex-
pertains or functions \( \phi_f \) with parameter \( \alpha_f \) is given by

\[
\phi_e(x_c; J_f, \alpha_f) = \prod_{f=1}^{F} \phi_f(x_c; J_f, \alpha_f).
\]  

We postulate a Student-T distribution for each expert since the color histograms of retinal images closely follow this form. Each expert is thus defined by

\[
\phi_f(x_c; J_f, \alpha_f) = \left( 1 + \frac{1}{2} (J_f, x_c)^2 \right)^{-\alpha_f}.
\]

For positive \( \alpha \) the above expression has high value when the patch \( x_c \) and filter \( J_f \) are less coincident, which is the opposite of typical definitions for potential functions. However, this is not a problem since the parameters are learned accordingly, i.e., smaller \( \alpha \)'s are associated with those filters corresponding to more likely configurations.

**Color Image Prior:** To overcome increased computational demand in training color image priors we: (1) fix the basis filters \( J_f \); and (2) the parameters \( \alpha \) are selected using Maximum Likelihood Estimation (MLE) as in [11]

To learn \( J_f \)'s we randomly select 15,000 training cliques of dimension 15 × 15 × 3 from each of the three classes: optical cup, retinal rim and background. The cliques are always larger than the patch size of 5 × 5 × 3 (which we want to learn) so that multiple overlapping patches are analysed in the training stage. We perform SVD on their covariance matrices and ignore the component with the highest variance (which corresponds to a uniform gray image).

Our method has the following stages: 1) approximate optic disc localization; 2) generating probability maps for the disc; and 3) segmenting the disc into the cup and rim region. Below we discuss each stage in greater detail

### 2.2. Preprocessing and ROI Localization

Before learning class priors the blood vessels are removed using the multi-scale difference-of-closing vessel extraction algorithm of [12] that uses a disk structuring element of radius 16 pixels and 3 pixels for dilation and erosion. We implement the elliptical Hough transform method of [13] to approximately localize the optic disc because it also eliminates the PPA (peri-papillary atrophy) region.

### 2.3. Random Forest Training Of Prior Model

The ROI localization method is deliberately designed so as to obtain a large ROI that includes the entire disc and parts of the neighboring structures. The manually annotated training images identify the optic disc (OD) and optic cup. The neuro-retinal rim is obtained by taking their difference image. To get training patches from the neighboring background, we first calculate the larger diameter of the OD (since it is an ellipse) and determine a square bounding box whose sides are twice the diameter of the OD, and is centered on the OD. Subtracting the OD from this box gives the background.

The FoE model of each class consists of 74 filters, i.e., \( 5 \times 5 \times 3 - 1 \) (ignoring the component with largest variance). Patches of one class give a higher response when filtered with the FoE filters of the same class. However it is difficult to identify the exact filter from each class that would be able to accurately identify the patch label. Therefore we use a RF classifier to distinguish between the three classes.

Each training patch is filtered with all the 74 filters of the PoE models of optic cup, neuro-retinal rim, and background to get \( 74 \times 3 = 222 \) features. We ensure that the number of samples to train the RF are approximately equal for each class and taken equally from each patient image. This avoids unbalancing of the trained RF.

For a given test image, the ROI is identified using the elliptical Hough transform method previously described. A \( 5 \times 5 \times 3 \) patch is extracted around each ROI pixel and filtered with the FoE models of each class. The feature vectors are input to the trained RF classifier which generates probability maps for the ROI. Although the RF can output a class label for each pixel, it may not ensure spatial consistency. Hence we use the probability maps in a second order MRF cost function which is optimized using graph cuts.

### 2.4. Final Segmentation using Graph Cuts

A second order MRF energy function is written as

\[
E(L) = \sum_{s \in P} D(L_s) + \lambda \sum_{(s,t) \in N} V(L_s, L_t),
\]

where \( P \) denotes the set of pixels and \( N \) is the set of neighboring pixels for pixel \( s \). \( \lambda \) is a weight that determines the relative contribution of penalty cost \( (D) \) and smoothness cost \( (V) \). \( D(L_s) \) is given by

\[
D(L_s) = -\log\left( Pr(L_s) + \epsilon \right),
\]

where \( Pr \) is the likelihood (from probability maps) previously obtained using RF classifiers and \( \epsilon = 0.00001 \) is a very small value to ensure that the cost is a real number.

**Smoothness Cost:** \( V \) ensures a spatially smooth solution by penalizing intensity discontinuities over a pixel's 8 neighbors. \( V \) is given by

\[
V(L_s, L_t) = \begin{cases} 
\frac{(L_s - L_t)^2}{2\sigma^2}, & ||s-t|| = 1, \\
0, & L_s \neq L_t, \\
\frac{1}{||s-t||}, & L_s = L_t.
\end{cases}
\]

Optimizing Eqn 4 using graph cuts gives the final labels.

### 3. EXPERIMENTAL RESULTS

#### 3.1. Dataset Description

Our proposed method for optic disc and cup segmentation was validated on the DRISHTI-GS dataset [14] which con-
Quantitative evaluation is based on the overlap measure of region overlap and absolute pointwise localization error (Eqn. 4), which was the value fixed for our experiments. The difference is also statistically significant since $p < 0.01$ (from Student-t tests) for all methods compared to FoE. Segmenting the optic cup is more challenging than the disc due to absence of distinguishing depth information. While pallor is one factor, it is not always reliable due to similar intensity profiles of neighboring regions. FoE obtains high segmentation accuracy than hand crafted features by learning image priors for the cup region.

Since [4] is a superpixel based approach, pixels from different classes may be grouped in one superpixel which affects its performance. [3] uses a modified Chan-Vese model, which finds it challenging to segment the optic disc using only intensity information. [2] uses only morphological features which is good enough for disc segmentation, but does not perform as well for cup segmentation. [8] was designed specifically for cup segmentation and hence performs well. However FoE outperforms all these methods.

Figure 2 shows the comparative results of FoE and the combined disc and cup segmentation methods of [4],[7] and [3]. The CDR is measured by fitting a ellipse to the disc and cup segmentations and calculating the ratio of the larger diameter to the smaller diameter. FoE outperforms all the competing methods as is evident from the higher $F$ and $S$ values, and lower $B$ values. The difference is also statistically significant since $p < 0.01$ (from Student-t tests) for all methods compared to FoE. Segmenting the optic cup is more challenging than the disc due to absence of distinguishing depth information. While pallor is one factor, it is not always reliable due to similar intensity profiles of neighboring regions. FoE obtains high segmentation accuracy than hand crafted features by learning image priors for the cup region.

Since [4] is a superpixel based approach, pixels from different classes may be grouped in one superpixel which affects its performance. [3] uses a modified Chan-Vese model, which finds it challenging to segment the optic disc using only intensity information. [2] uses only morphological features which is good enough for disc segmentation, but does not perform as well for cup segmentation. [8] was designed specifically for cup segmentation and hence performs well. However FoE outperforms all these methods.

3.3. Segmentation Performance

Table 2 summarizes the segmentation performance of different methods. The competing methods are categorized as those which segment only the optic disk ($D$), optic cup ($C$) or both ($D,C$). FoE outperforms all the competing methods as is evident from the higher $F$ and $S$ values, and lower $B$ values. The difference is also statistically significant since $p < 0.01$ (from Student-t tests) for all methods compared to FoE. Segmenting the optic cup is more challenging than the disc due to absence of distinguishing depth information. While pallor is one factor, it is not always reliable due to similar intensity profiles of neighboring regions. FoE obtains high segmentation accuracy than hand crafted features by learning image priors for the cup region.

Since [4] is a superpixel based approach, pixels from different classes may be grouped in one superpixel which affects its performance. [3] uses a modified Chan-Vese model, which finds it challenging to segment the optic disc using only intensity information. [2] uses only morphological features which is good enough for disc segmentation, but does not perform as well for cup segmentation. [8] was designed specifically for cup segmentation and hence performs well. However FoE outperforms all these methods.

Figure 2 shows the comparative results of FoE and the combined disc and cup segmentation methods of [4],[7] and [3]. The CDR is measured by fitting a ellipse to the disc and cup segmentations and calculating the ratio of the larger diameters. FoE as measured from FoE segmentations shows higher correlation coefficient ($r_{FoE} = 0.85$) with manual scores than other methods, i.e $r_{[4]} = 0.80$, $r_{[7]} = 0.78$, and $r_{[3]} = 0.77$.

4. DISCUSSION AND CONCLUSION

We proposed a method to segment the optic disc and cup by learning their prior models using Field of Experts. An ap-
Table 2. Segmentation accuracy in terms of $F$ score, overlap and boundary distance for different methods. $D, C$ indicate if the method segments the optic disc or optic cup or both. $B$ is in pixels; $Time$ is in minutes; $F$-$F$ score; $S$-overlap measure; $B$-boundary error.

<table>
<thead>
<tr>
<th></th>
<th>$F_{oE}$</th>
<th>[2]</th>
<th>[4]</th>
<th>[7]</th>
<th>[3]</th>
<th>[8]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disc</td>
<td>F</td>
<td>95.9</td>
<td>94.0</td>
<td>93.2</td>
<td>92.0</td>
<td>94.6</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>89.5</td>
<td>87.3</td>
<td>86.8</td>
<td>85.2</td>
<td>86.9</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>8.2</td>
<td>9.4</td>
<td>9.9</td>
<td>11.7</td>
<td>9.1</td>
</tr>
<tr>
<td>Cup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optic</td>
<td>F</td>
<td>95.9</td>
<td>94.0</td>
<td>93.2</td>
<td>92.0</td>
<td>94.6</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>86.4</td>
<td>82.1</td>
<td>80.1</td>
<td>82.5</td>
<td>84.2</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>9.3</td>
<td>14.3</td>
<td>15.6</td>
<td>13.2</td>
<td>11.2</td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td>7</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>

proximate ROI encompassing the disc is identified and the cup and rim are segmented using probability maps generated by RF classifiers. Experimental results show our method to outperform other methods in the literature which have been designed to simultaneously or individually segment the disc and cup. FoE’s superior performance is attributed to learning image characteristics without depending on hand crafted features which can be unreliable. In future work we aim to apply our method to the larger MESSIDOR database for optic disc segmentation.

5. REFERENCES


