Semi-supervised Learning for Biomedical Image Segmentation via Forest Oriented Super Pixels(Voxels)

Lin Gu[1]([✉]), Yinqiang Zheng[1], Ryoma Bise[2], Imari Sato[1], Nobuaki Imanishi[3], and Sadakazu Aiso[3]

[1] National Institute of Informatics, Tokyo, Japan
`ling@nii.ac.jp`
[2] Kyushu University, Fukuoka, Japan
[3] Keio University, Tokyo, Japan

Abstract. In this paper, we focus on semi-supervised learning for biomedical image segmentation, so as to take advantage of huge unlabelled data. We observe that there usually exist some homogeneous connected areas of low confidence in biomedical images, which tend to confuse the classifier trained with limited labelled samples. To cope with this difficulty, we propose to construct forest oriented super pixels(voxels) to augment the standard random forest classifier, in which super pixels(voxels) are built upon the forest based code. Compared to the state-of-the-art, our proposed method shows superior segmentation performance on challenging 2D/3D biomedical images. The full implementation (based on Matlab) is available at https://github.com/lingucv/ssl_superpixels.

Keywords: Image segmentation · Semi-supervised learning · Random forest · Super pixels(voxels)

1 Introduction

To collect massive biomedical data is becoming relatively easy, yet to annotate them correctly remains very challenging, since it necessitates specific knowledge from doctors and biomedical experts. This has turned out to be a bottleneck issue when applying supervised machine learning methods on biomedical image segmentation and recognition tasks, which usually require a large amount of labelled samples. Semi-supervised learning (SSL) [6,7,13] is promising in exploiting huge unlabelled data to improve classification performance. However, these SSL algorithms either depend on the kernel density estimation (KDE) [7] or low density assumption in [6], both of which rarely hold true in biomedical imaging applications.

In this paper, we focus on semi-supervised learning for biomedical image segmentation, without relying on those restrictive assumptions in [6,7]. Our key observation is that there usually exist some homogeneous connected areas of low

confidence in biomedical images, which tend to confuse the classifier trained with limited labelled samples. For example, the pathological and optic disk regions in Fig. 1(a) have hindered the standard random forest for vessel segmentation, as shown in Fig. 1(b). We propose a novel forest oriented super pixel(voxel) method, named FOSP(FOSV), to be discriminant to the segmentation task. Our forest oriented super pixels(voxels) are built upon the forest based code, rather than the pixel intensity/color in the existing sense [2,8]. Our proposed FOSP(FOSV) will segment the initial estimation image into atomic regions (Fig. 1(c)) of consistent estimation to pick up the low confidence samples. By leveraging these atomic regions, we train a semi-supervised learning algorithm through a Bayesian strategy that successfully suppresses the confusing homogeneous areas. Compared to the state-of-the-art, our proposed method shows superior vessel segmentation performance on challenging 2D retinal and X-ray images. We have also verified that our method can be applied seamlessly to 3D volume data by constructing supervoxels instead.
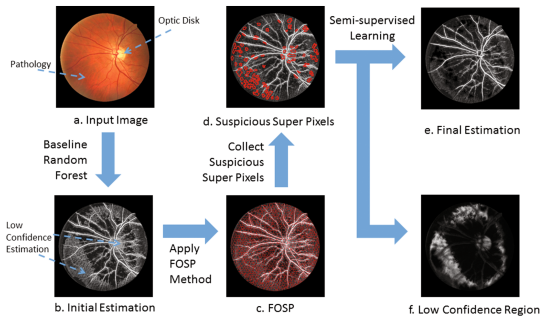
Fig. 1. The pipeline of our proposed semi-supervised learning method on the basis of forest oriented super pixels (FOSP).

Our core contributions can be summarized as that: 1. We propose FOSP (FOSV) to capture the complementary information of random forest; 2. With forest oriented super pixels(voxels), we succeed in unsupervised prediction of the low confidence regions that would otherwise confuse the classifier; 3. Our semi-supervised learning method shows outstanding segmentation performance on challenging 2D/3D biomedical images. 4. Our method's prediction of low confidence region, which is often related to pathology, raises the potential in areas such as unsupervised disease diagnose.

2 FOSP(FOSV) Based Semi-supervised Learning

2.1 Overview

Given an input image Fig. 1(a) and limited number (*e.g.*, 500) of labelled training samples, we could train a random forest classifier to estimate the vessel(Fig. 1(b)). Random forest and some other supervised learning methods often

fail when encountering data beyond the representative training set. For example, here classifiers trained with limited data get confused by unseen pathology and optic disk regions, and return an ambiguous estimation in Fig. 1(b). Observing that the low confidence regions are often connected areas of homogeneous estimation, we propose a forest oriented super pixel (FOSP) to segment the estimation image into atomic regions of consistent classifier prediction (Fig. 1(c)). With the super pixels in hand, we could thus pick the suspicious super pixels (Fig. 1(d)) to train the semi-supervised classifier, which is able to predict the low confidence region (Fig. 1(f)). Our semi-supervised learning algorithm finally produces a segmentation via Bayesian estimation with low confidence prediction as prior knowledge. As shown in Fig. 1(e), our method successfully suppress the influence of confusing pathology and optic disk areas compared to the baseline method (Fig. 1(b)) with limited training data. The details of each step will be presented in the following subsections.

2.2 Tree and Forest Based Code

Random forest [3] is a widely employed classifier with several attractive characteristics, such as high efficiency and robustness to over-fitting. Its tree based code is also well studied [9] to reveal complementary information of tree structure that is not apparent in the final class prediction.

Following the definition in [9], we reformulate the forest's prediction into a compact form. The structure of individual tree could be regarded as a mapping function that maps the data of pixel p into a tree based code $\phi(p)$, which is a binary vector of the same number of leaf nodes. Each element of a tree's $\phi(p)$ corresponds to a leaf node and is set to 1 if that leaf node contains pixel p and 0 if not. The forest based code is thus composed by stacking individual tree's code together. Meanwhile we pack the accumulated posterior class probability of leaf vector into a leaf matrix w with each column corresponding to a leaf node. So the class prediction $E(p)$ would become as $E(p) = w\phi(p)$. In the following, we measure the distance of random forest prediction via the forest based vector $\phi(p)$ to finally construct forest oriented super pixels(voxels).

2.3 Forest Oriented Super Pixels(Voxels)

Super pixels(voxels) are usually obtained by clustering the pixels into meaningful atomic regions [2,8], which have served as the building block for several computer vision and biomedical imaging applications [12]. Unlike existing methods that based on unsupervised colour space, we propose our FOSP(FOSV) which defines the distance on forest based code.

We describe the algorithm of FOSP(FOSV) in Procedure 1. Given either a 2D or 3D biomedical image and pre-trained random forest classifier from limited labelled data, individual pixel p is associated with an estimation score $E(p)$ and its forest based code $\phi(p)$.

Procedure 1. Algorithm of Generating Forest Oriented Super Pixels(Voxels)

Input: Estimation score $E(p)$ and the tree based code $\phi(p)$ of individual pixel p of a 2D or 3D image, the initial number of super pixel(voxels) K, the maximal number of iterations $iter_{max}$ and the stopped threshold ϵ

Output: Forest Oriented Super Pixels(Voxels)

 1: Evenly distribute K seeds at regular grids with a fixed step $\S = \sqrt{N/K}$ where N is the number of pixels in the image.
 2: Set label $l(p) = -1$ and distance $d(p) = \infty$ for each pixel p
 3: Set the residual error $err = \infty$ and $iter = 1$
 4: **while** $err > \epsilon$ and $iter < iter_{max}$ **do**
 5:     **for** Each cluster centre c_i of super pixel(voxel) **do**
 6:         For each pixel(voxel) p in the region $2S \times 2S$ or $2S \times 2S \times 2S$, Compute the forest oriented distance D_{co} between c_i and p using Equation 1.
 7:         **if** $D_{co} < d(p)$ **then**
 8:             Set $d(p) = D; l(p) = i$
 9:         **end if**
10:     **end for**
11:     Merge similar super pixels(voxels)
12:     **for** Each super pixel(voxel) s_i **do**
13:         **for** Each super pixel(voxel) s_j in the neighbours of current s_i **do**
14:             Calculate the variance $v(\phi(\omega(S_j) \cup \omega(S_i)))$ of the random forest code on the union region belonging to super pixel(voxel) s_j and s_i.
15:             **if** $v(\phi(\omega(S_j) \cup \omega(S_i))) < \epsilon_m$ **then**
16:                 Merge the super pixel(voxel) s_j to s_i
17:                 Update the cluster centre c_i of super pixel(voxel) s_i
18:             **end if**
19:         **end for**
20:     **end for**
21:     **for** Each seed s_i **do**
22:         Update the cluster centre c_i as the centroid of region ω_i belonging to super pixel(voxel) s_i .
23:         Calculate residual error E
24:     **end for**
25: **end while**

The unique definition of distance D_{co} is our key contribution since it is defined on the forest based code which accounts for the global complementary information [9] of random forest. Thus our FOSP(FOSV) is discriminant to segmentation task and attempts to gather pixels with similar ϕ codes (thus) leading to similar predictions. An additional benefit is that the samples who shares more leaf nodes would be more likely to be send to same super pixel(voxel), offering an advantage in improving the semi-supervised learning. Note that this distance also depends on the iteration number, which enables the super pixels(voxels) to quickly move to the high or low score region.

$$D_{co}(p_i, p_j, iter) = \lambda_{co}\|\phi(p_i) - \phi(p_j)\| + \lambda_s\|p_i - p_j\| + \lambda_g d_g,$$
$$d_g(p_i, p_j) = \|E(p_j) - (1 - e^{iter-iter_{max}})E(p_i) + e^{iter-iter_{max}}g(E(c_i))\|, \quad (1)$$
$$g(x) = \begin{cases} 1 , x > 0.5 \\ 0 , x \le 0.5 \end{cases}$$

We also introduce an efficient merging operation to greedily merge the neighbouring super pixels(voxels) if the variance of ϕ of their pixel is small enough. In this way, content sparse super pixels(voxels) with uniform and similar prediction would be merged together.

2.4   Collect the Low Confidence Super Pixels(Voxels)

We now attempt to find the super pixels(voxels) with least confidence to classifier. As shown in Fig. 1(b), the low confidence regions that confuse the random forest often show less variance. Therefore, we define each super pixel(voxel)'s confidence score be $S_c(s_i) = v(E(\omega(s_i)))m(|E(\omega(s_i)) - 0.5|)$ where $v(E(\omega(s_i)))$ is the variance of estimation score and $m(|E(\omega(s_i)) - 0.5|)$ is the mean value of confidence value [5,14] belonging to s_i. In practice, we set a threshold to collect the small confidence score as low confidence candidates.

Collect the super pixels(voxels) whose confidence scores $S_c(s)$ are small enough to be the candidate regions of low confidence.

2.5   Semi-supervised Learning

Given the low confidence super pixels(voxels) obtained from above, we then proceed to leverage them to improve the supervised learning. Simply combining the collected low confidence super pixels(voxels) as negative samples into existing often limited training samples would not help due to severe imbalance between positive and negative samples *e.g.*, 30000 negatives vs. 500 positive. Instead, we propose to take a Bayesian strategy to train a pair of random forests. The first random forest aims to provide a prior probability of the suspicious area. Specifically, we take the samples of suspicious super pixels(voxels) as positive training data and the rest of image as negative ones to train a classifier, which generate $E_s(p)$ to predict the prior probability of pixel p to be misclassified. Then we train a standard random forest with labelled data to give the prediction $E_l(p)$. Finally, we could obtain an estimation through the Bayesian theorem $E(p) = (1 - E_s(p))E_l(p)$.
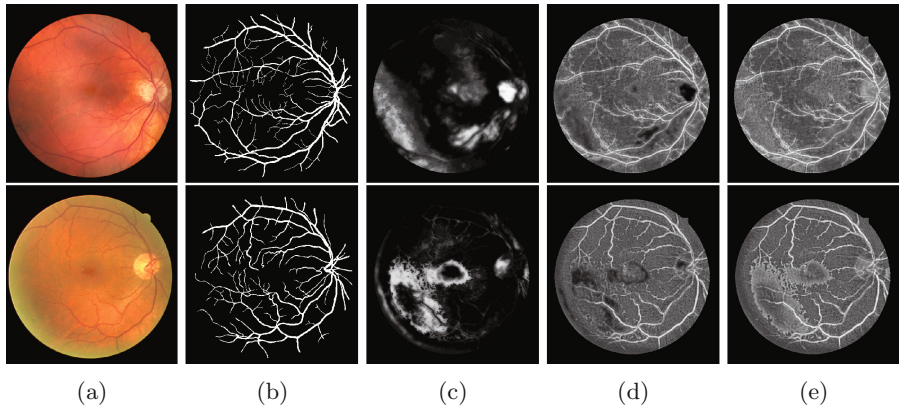
3   Experiments

### 3.1   Vessel Segmentation in 2D Biomedical Images

We first evaluate our method for vessel segmentation on the retinal dataset DRIVE [11] and X-ray images of the human hand collected by our Softex C

Series system. For 2D vessel segmentation, we collect three kinds of features: 1. [15 × 15] local patches of raw images; 2. Gabor Wavelet Feature [10], 3. [3 × 3] local patches on [4]. We compare our method with the state-of-the-art random forest. Both use 500 (250 positive and 250 negative) labelled training samples.

The representative results on DRIVE dataset are shown in Fig. 2. We train our method on 20 training images and then evaluate it on the rest of 20 testing images pre-defined by the database. We can observe that our method is able to identify the suspicious region (optical disk and pathology region) as illustrated in Fig. 2(c), which obviously confuses the baseline random forest in Fig. 2(e).

(a)   (b)   (c)   (d)   (e)

Fig. 2. Exemplar experiment results on the DRIVE dataset. (a) Input images; (b) Ground-truth; (c) Estimation of low confidence region by our method; (d) Estimated score of vessel by our method; (e) Estimated score of vessel by baseline random forest.

(a) Input Image   (b) Our Method   (c) Standard RF
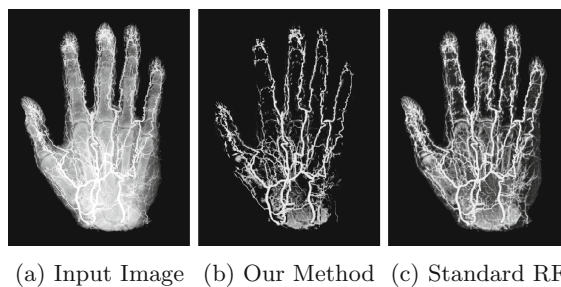
Fig. 3. Exemplar vessel estimation results on X-ray images of hand. (a) Input X-ray image; (b) Estimation by our method; (c) Estimation baseline random forest.

We also present the results on 2D X-ray hand images in Fig. 3. From the input image (a), we can see that the intensity of bone and vessel are relatively close. This makes vessel segmentation extremely challenging in the presence of

bone at background. Our method clearly separates the vessel from background of bone where baseline random forest trained mistakes part of the bone as vessels.

3.2   Quantitative Comparison

Since the groundtruth is provided by DRIVE dataset, we compare the segmentation performance with alternative semi-supervised method: TSVM [6], Robust Node Random Forest [7]. All of the methods are trained with 500 labelled samples. As shown in the Fig. 4(a), our method significantly outperforms the alternatives. In Fig. 4(b) and (e) also demonstrates that our semi-supervised learning method consistently improves the performance from baseline under various sizes of training set (at a p value less than 0.0000005).


(a) PR Curves   (b) F1 vs Labelled Data Size
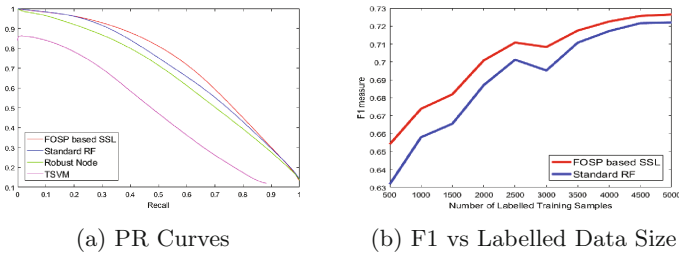
Fig. 4. (a) Precision-Recall curves of various methods. (b) Optimal F1 measure vs. labelled data size

3.3   Interpretation of Low Confidence Regions

Actually, our FOSP(FOSV) based semi-supervised learning method attempts to locate the homogeneous area that assumes less confidence for the random forest. Since the training samples are mainly composed of healthy vessel and background, the pathology area would nevertheless be different and confuse the classification algorithms. As a meaningful by-product, the low confidence region arising from our method actually indicates the pathology region, as shown in Fig. 2(d), without any human annotation.

3.4   Neuron Segmentation in 3D Biomedical Images

By constructing super voxels instead, we apply our method to neuron segmentation on the 3D BigNeuron Initiative dataset [1]. For this application, we collect $[7 \times 7 \times 7]$ local patches and the [4] on each slice as features. Again, we use 500 labelled samples as guidance. As illustrated in Fig. 5, our proposed method manages to filter out the out-focused neurons and noises that obviously appear in the results of standard random forest.

(a) Input Image   (b) Our Method   (c) Standard RF
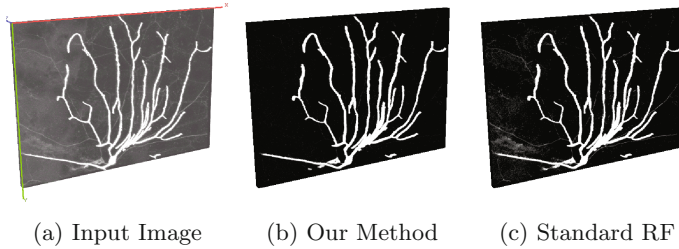
Fig. 5. Exemplar neuron estimation on BigNeuron dataset. (a) Input image; (b) Estimation by our method; (c) Estimation by baseline random forest.

4   Conclusion

In this paper, we propose a Forest Oriented Super Pixel(Voxel) method which aims to capture complementary information of random forest. Our proposed method is able to automatically locate the region that would confuse the classifier such as pathological region. We have also developed a semi-supervised learning through a Bayesian strategy to leverage the collected confusing regions. The superior performance of our method has been demonstrated on various 2D and 3D biomedical image segmentation applications.

References

1. Peng, H., Hawrylycz, M., Roskams, J., Hill, S., Spruston, N., Meijering, E., Ascoli, G.A.: Bigneuron: large-scale 3D neuron reconstruction from optical microscopy images. Neuron **87**(2), 252–256 (2015)
2. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Ssstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Patt. Anal. Mach. Intell. **34**(11), 2274–2282 (2012)
3. Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A.: Classification and Regression Trees. CRC Press, Boca Raton (1984)
4. Coye, T.: A novel retinal blood vessel segmentation algorithm for fundus images. MATLAB Central File Exchange, January 2017
5. Gu, L., Cheng, L.: Learning to boost filamentary structure segmentation. In: ICCV 2015 (2015)
6. Jaakkola, M.S.T., Szummer, M.: Partially labeled classification with Markov random walks. In: NIPS 2002, vol. 14, pp. 945–952 (2002)
7. Liu, X., Song, M., Tao, D., Liu, Z., Zhang, L., Chen, C., Bu, J.: Random forest construction with robust semisupervised node splitting. IEEE Trans. Image Process. **24**(1), 471–483 (2015)
8. Liu, Y.J., Yu, C.C., Yu, M.J., He, Y.: Manifold SLIC: a fast method to compute content-sensitive superpixels. In: CVPR 2016, June 2016

9. Ren, S., Cao, X., Wei, Y., Sun, J.: Global refinement of random forest. In: CVPR 2015 (2015)
10. Soares, J.V.B., Leandro, J.J.G., Cesar, R.M., Jelinek, H.F., Cree, M.J.: Retinal vessel segmentation using the 2-D gabor wavelet and supervised classification. IEEE Trans. Med. Imaging **25**(9), 1214–1222 (2006)
11. Staal, J., Abramoff, M.D., Niemeijer, M., Viergever, M.A., van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. IEEE Trans. Med. Imaging **23**(4), 501–509 (2004)
12. Tian, Z., Liu, L., Zhang, Z., Fei, B.: Superpixel-based segmentation for 3D prostate MR images. IEEE Trans. Med. Imaging **35**(3), 791–801 (2016)
13. Yang, Z., Cohen, W.W., Salakhutdinov, R.: Revisiting semi-supervised learning with graph embeddings. In: ICML (2016)
14. Zhu, X., Goldberg, A., Brachman, R., Dietterich, T.: Introduction to Semi-supervised Learning. Morgan and Claypool Publishers, San Rafael (2009)