DEEP VESSEL TRACKING: A GENERALIZED PROBABILISTIC APPROACH VIA DEEP LEARNING

Aaron Wu, Ziyue Xu**, Mingchen Gao, Mario Buty, Daniel J. Mollura*

Department of Radiology and Imaging Sciences, National Institutes of Health, Bethesda, MD 20892

ABSTRACT

Analysis of vascular geometry is important in many medical imaging applications, such as retinal, pulmonary, and cardiac investigations. In order to make reliable judgments for clinical usage, accurate and robust segmentation methods are needed. Due to the high complexity of biological vasculature trees, manual identification is often too time-consuming and tedious to be used in practice. To design an automated and computerized method, a major challenge is that the appearance of vasculatures in medical images has great variance across modalities and subjects. Therefore, most existing approaches are specially designed for a particular task, lacking the flexibility to be adapted to other circumstances. In this paper, we present a generic approach for vascular structure identification from medical images, which can be used for multiple purposes robustly. The proposed method uses the state-of-the-art deep convolutional neural network (CNN) to learn the appearance features of the target. A Principal Component Analysis (PCA)-based nearest neighbor search is then utilized to estimate the local structure distribution, which is further incorporated within the generalized probabilistic tracking framework to extract the entire connected tree. Qualitative and quantitative results over retinal fundus data demonstrate that the proposed framework achieves comparable accuracy as compared with state-of-the-art methods, while efficiently producing more information regarding the candidate tree structure.

Index Terms— Vascular Structure, Deep Learning, Principal Component Analysis, Nearest Neighbor Search, Generalized Probabilistic Tracking

1. INTRODUCTION

Accurate analysis of biological vasculatures is of significant importance in several medical image analysis tasks. For instance, cardiac and pulmonary diseases are the two leading causes of mortality, both of which are closely related to tubular anatomical structures, i.e. coronary artery, airway, and pulmonary vessels. Imaging techniques, such as optic imaging and computed tomography (CT), are commonly utilized for such investigations. In order to make reliable diagnostic and prognostic judgments, it is often necessary to extract precise quantitative information regarding the related vessel geometry. Therefore, accurate and robust image segmentation methods are highly needed. However, there are two major challenges towards an efficient and rigorous system. First, manual identification of vascular trees is often infeasible for clinical practice due to the extreme complexity of biological vasculatures, which makes it too time-consuming and tedious. Second, the appearance of vascular structures in medical images has great variance across modalities and subjects. Therefore, it is quite difficult to have a general description of the vasculatures. Consequently, most conventional methods rely on certain assumption for the appearance of a candidate structure, and are specially designed for a particular task with exclusive models [1]. These models are often carefully crafted so that they can capture the specific target, and may need to be modified according to structure size [2]. Due to the structure complexity and appearance variability, these methods are often limited in their robustness, and need further processing steps for better results [3].

In this paper, we present a generic approach for vascular structure identification from medical images, which can be applied for multiple tasks robustly. In order to address the challenges described above, we used deep CNNs [4] under a probabilistic tracking framework, designed with generalized particle filter [5]. It has been demonstrated in many applications that deep learning is capable of capturing the features of complex shape and appearance patterns, and that it has superior performance over most state-of-the-art methods. By using deep learning, we shift the difficulty in mathematically modeling the appearance to the collection of proper training data, which in most cases is more resolvable. The effectiveness of CNN for vessel identification is illustrated in previous study [4], whereas the efficiency, as well as the capability of forming a tree structure, can be a challenge for such algorithms. On the other hand, probabilistic tracking methods have been investigated in multiple studies for its robustness and efficiency [2]. Typically, the Monte Carlo method is used to predict the progression of a tracking path. Under this framework, most research focused on better sampling strate-

---

*Corresponding author: ziyue.xu@nih.gov. This research is supported by the Center for Infectious Disease Imaging, an intramural research program of the National Institute of Allergy and Infectious Diseases (NIAID) and the National Institute of Biomedical Imaging and Bioengineering (NIBIB).

978-1-4799-2349-6/16/$31.00 ©2016 IEEE

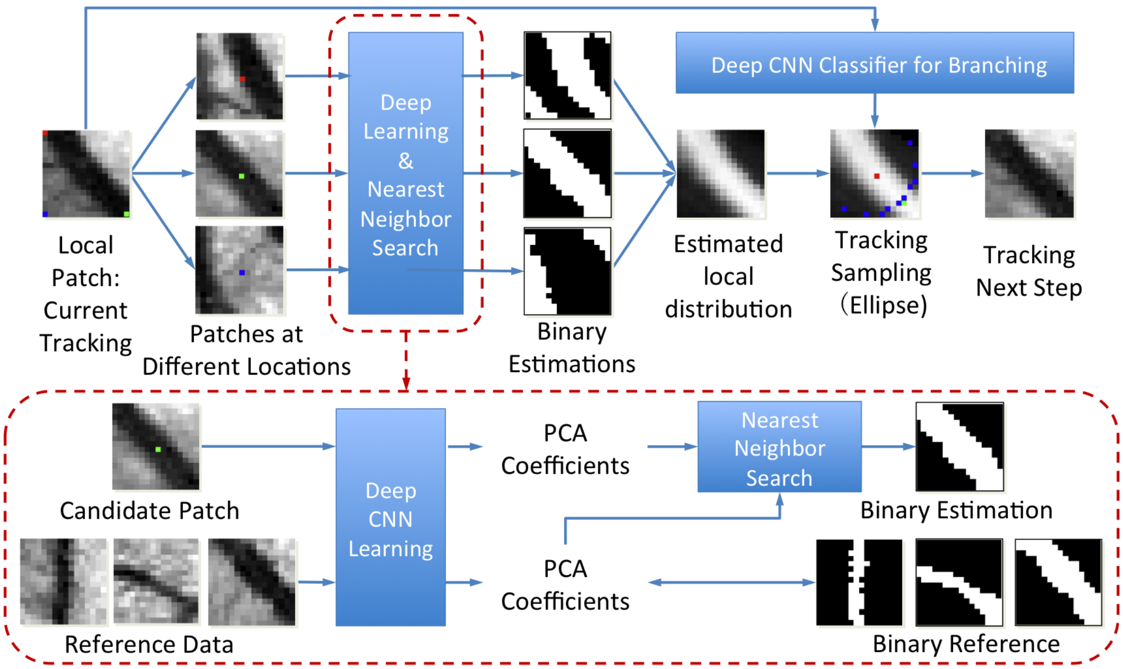

Fig. 1. Flowchart of the vessel tracking method with deep CNN and generalized particle filter.

gies, smarter particle designs, or more accurate observation models. However, this approach may not be optimal for medical image analysis tasks due to its inherent requirement of intensity models. Recently, a generalized particle filter has been proposed in [5]. It no longer estimates the target distribution with random Monte Carlo samples; instead, a continuous Gaussian measure is utilized for the evaluation. Here, we follow a similar idea and designed a generalized probabilistic tracking method in the discrete space of images using the PCA and nearest neighbor search concepts from N4 Field study [4]. This enables us to combine the advantages of both tracking and deep learning methods. A flowchart of the proposed framework is shown in Fig.1. In the next section, it is presented in detail.

2. METHODS

As shown in Fig.1, the proposed framework consists of two main techniques: a learning system trained to generate the estimation of local vessel probability map, and a tracking framework based on this map. Specifically, a deep CNN was first trained to transform an image patch into its corresponding binary vessel segmentation. Then, the overall vessel probability distribution was estimated by combining all binary annotations according to their locations, instead of using random Monte Carlo samples. Once the probability map was generated, the tracking was performed via edge sampling, and with proper learning, structures with significantly different appearances are converted to vessel probability maps.

CNN based Segmentation Map: Deep CNN provides the capability of learning a set of image properties that allow prediction of a certain aspect or object of an image, i.e. the vasculature for our application. The output is usually a set of labels (classification), or parameters (regression). The aim of this algorithm is to associate an image patch P with its potential vessel segmentation $B(P)$. There are two challenges for this task. First, the feature space is unnecessarily large (N^k for k-dimensional input) for CNN output considering its binary nature; and second, a proper mechanism is needed to produce the estimated binary segmentation. Hence, we adopted the PCA and dictionary searching idea from the N4 Field method [4]. Each training patch P_T has a corresponding binary annotation $B(P_\mathrm{T})$, and PCA is employed to reduce the high-dimensionality of the sparse binary annotation $B(P_\mathrm{T})$ of size $N \times N$ to a lower-dimensional feature vector $\mathrm{PCA}(B(P_\mathrm{T}))$ of size N. Then, the CNN is trained with input-output pairs $\{P_\mathrm{T}, \mathrm{PCA}(B(P_\mathrm{T}))\}$. Using the trained CNN, an estimation of the PCA-reduced feature vector $F(P_t)$ will be produced for any testing input P_t. Further, to find the proper binary annotations for each $F(P_t)$, a dictionary scheme is employed. Pairs of feature vectors and binary segmentations $\{B(P_\mathrm{T}), \mathrm{PCA}(B(P_\mathrm{T}))\}$ are first extracted from subset of the training data. These samples are then used as dictionary $\mathcal{D}$ for recovering binary annotations. For any CNN-predicted feature vector $F(P_t)$, a query for its nearest neighbor among dictionary entries $\mathrm{PCA}(B(P_\mathrm{T})) \in \mathcal{D}$ is performed to find its closest match, then the corresponding $B(P_\mathrm{T})$ is used as the prediction of segmentation for P_t.

Generalized Probabilistic Tracking: With CNN learning, each image patch can be associated with its closest binary vessel annotation. As shown in [4], the local vessel probability at any location can then be estimated by averaging all neighboring patches annotations. Further, the entire scene can be processed by computing the probability for all pixels. However, for our application of vessel structure extraction, such an exhaustive approach suffers from two major drawbacks. First, performing computation across the entire image would be inefficient, with a significant amount of locations within the image irrelevant to candidate vessel structure. This may not be very prominent for 2D applications such as retinal fundus image where vessel-related areas can cover about half of entire scene. However, for 3D applications, such as airways and cardiac vessels, this ratio may well fall down to below 1%. Since the CNN dictionary search computation is fairly expensive, it is more efficient to minimize the number of non-vessel pixels being evaluated. Second, vasculature usually follow a regular connected pattern, and several analysis, including tree generation-wise evaluations and branch length assessment, rely on the construction of a connected tree structure. Therefore, in addition to vessel probability at specific locations, we often want to simultaneously recognize its position at the candidate tree structure. However, with original scheme of exhaustive computation, additional steps will be necessary to extract the tree and remove the isolated detections. Thus a method capable of capturing the connected vascular tree will best fit our needs. Here, we used generalized probabilistic tracking approach to extract the vessels efficiently, and areas not connected to the vessel pathway will not be explored. Vessel tracking aims to estimate the vessel center, directionality, and branching properties from a local image window. Conventionally this relies on testing multiple hypotheses against a model using Monte Carlo random sampling. As mentioned, both model design and sampling strategy need to be carefully crafted, limiting the robustness of the method. By using both CNN based probability map and generalized probabilistic tracking, we addressed these challenges. Under tracking framework, a vessel can be modeled as a sequence of states $S = s_i : i = 1, ..., n$. At a particular tracking step t, local vessel segment is described $s_t = (c_t, v_t, A_t)$ with its center c_t, orientation v_t, and appearance model A_t. Tracking methods try to identify the best fit s_t given the observation data $O_{1:t}$ with Bayesian formulation

$$p(s_t|O_{1:t}) = \frac{p(O_t|s_t)p(s_t|O_{1:t-1})}{p(O_t|O_{1:t-1})}.$$

Under conventional Monte Carlo framework, a sample-based representation of the probability distribution is achieved with M “particles” with hypothesis s_t^i and weight w_t^i as

$$p(s_t|O_{1:t}) = \sum_{i=1}^{M} w_t^i \delta(s_t - s_t^i),$$

where weight is determined following Bayesian rule and against models. See [2] for an example of such framework. Using generalized particle filter, instead of randomly generating M “particles”, a Gaussian mixture model is used to estimate the probability distribution. We hereby adopted similar approximation using the CNN binary results. The vessel probability distribution (replacing A_t) is estimated by considering all binary annotations $B_{x,y}$ associated with neighboring samples within $N \times N$ space $\mathcal{N}$ according to their locations. Consequently, c_t and v_t are estimated by sampling the probability map as shown in Fig.1. An initial point c_0 was first manually placed on the vessel. The vessel probability distribution within the neighborhood window was estimated around this point using the established convolutional neural network and dictionary. Then, the direction of tracking v_0 was determined by finding the eigenvector of the set of point coordinates whose values exceed a certain probability threshold. A new center c_1 was determined by moving a specified distance in the determined direction. This process continued until a branch of vessel is fully tracked, which was determined if the new center's probability dropped below a threshold.

Branch Detection: For identifying bifurcations, another neural network was specifically trained to distinguish between the different image properties of three categories: branches, background, and normal vessel segments. When a new sample window was passed into the network during vessel tracking, the network would first determine which category the window falls into. Then, the vessel probability map is generated. In branching cases, the probability value is sampled along a half-ellipse ahead of the current center point. To determine the direction and center of the two next tracking locations, the intensity profile along the ellipse curve was calculated to detect the number of local maxima. Each local maximum was taken as the center of a branch. The new directions were determined from the current center and the locations of the local maxima. Each branch was then separately tracked. For the background case, tracking was terminated, while for the normal vessel case, tracking was continued as described previously.

3. EXPERIMENTS AND RESULTS

In order to evaluate the performance of the proposed tracking method, we used the 20 retinal fundus images from DRIVE dataset [6]. The CNN follows the model presented by [4, 7]. Three fully connected layers are implemented and half of the samples are dropped out to avoid over-fitting. A Euclidean-loss measure is used to determine error of the resulting prediction from the known feature vector output. After the network converged on an error of 1%, the network was tested on a new set of input samples for validation of the training in order to assess over-fitting. The network was then extracted for use in formation of the dictionary. To generate the dictionary, we sampled local patches and their corresponding manual segmentations. Samples were divided the search space to three categories of background, normal vessels, and branches. The skeleton of the retinal vessel images was first generated and analyzed for locating each of the three cases. Then samples were generated randomly. Note that branch cases here cover both branching and crossing. In order for the resulting learned network to be generalizable, we sampled various types of vessel thickness, orientation, and branch structure. Each type was balanced with an equal number of members. Furthermore, all the acquired samples were reflected on the x-axis and y-axis as well as rotated multiple times to random degrees of rotation. Here, we generated in total 14,400 training samples.

Accuracy: The proposed accurate and robust tracking system is based on N4 Field algorithm; hence, we expect them to have comparable performance. Indeed, the resulting tree from tracking is guaranteed to be connected, while for N4 computation, there could be isolated segments in the final result. These unconnected segments could contribute to slight accuracy gain, but could often be undesirable for clinical applications. Fig.2 illustrates the qualitative results. For comparison, in addition to N4 Field result, we selected a learning-based method of unsupervised fuzzy segmentation [8], and a tracking-based method of multi-scale line tracking [9]. As shown, the proposed method achieves almost identical result as N4 Field, which is better than other previous approaches. ROC curves of both the N4 Field method and the tracking method were shown in Fig.3, against manual segmentations. As listed in Table 1, the quantitative result of the area under the curve (AUC) was found to be 97.01% for the tracking method, which is comparable to full N4 computation, and superior to all other state-of-the-art methods, as listed in [10].

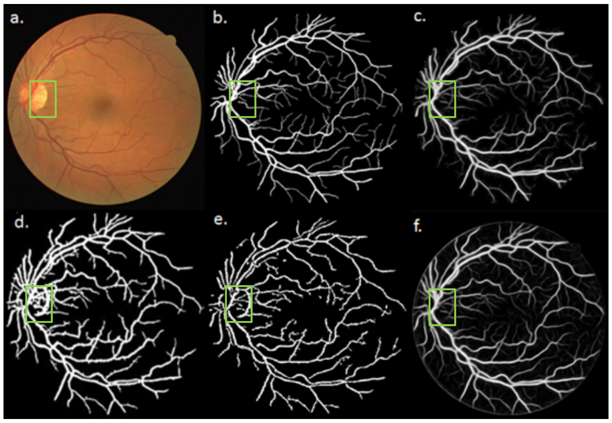

Fig. 2. a) Retinal image, b) Manual segmentation, c) Proposed tracking segmentation, d) Unsupervised fuzzy segmentation [8], e) Multi-scale line tracking [9], f) N4 Field segmentation [4]. Green region shows false positives caused by optic disk.

Efficiency for medical image analysis: One major contribution of this work is to significantly reduce the computation complexity of the exhaustive N4 Field with tracking scheme. Theoretical limit of efficiency gain is the ratio between the number of total voxels N_{Full} and voxels visited for vessels N_{Track}. From our experiment of 2D retinal image where this ratio is around 2, the actual tracking is about two times faster, i.e. close to theoretical limit. This speedup will be far greater for 3D medical images, where the ratio reaches beyond 100.

Table 1. AUC statistics for different vessel segmentation methods

Method	N4 Field [4]	**Proposed**	Fuzzy [8]	Tracking [9]
AUC	0.9767	**0.9701**	0.9518	0.9273

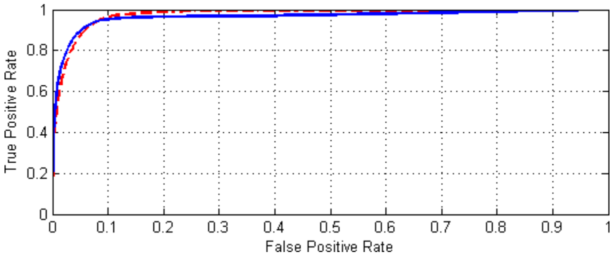

Fig. 3. ROC curves of Full N4 Field segmentation (red dashed) vs. Proposed Tracking scheme (blue solid)

4. DISCUSSION AND CONCLUSION

In this study, we presented a generic approach for vascular structure identification. The proposed algorithm combines the advantages from recent discoveries in computer vision and mathematics for addressing specific accuracy and efficiency challenges in medical image analysis. Specifically, a deep CNN is trained for estimating local vessel probability via PCA and nearest neighbor search, and the resulting map is further utilized within a probabilistic tracking framework to extract the entire connected tree. Performing CNN for all pixels can result in a loss in efficiency; also, no information regarding the tree structure can be simutaneously extracted. Conventional tracking is prone to the inflexibility from model design and sampling strategy. By combining the two, CNN provides probability distribution for tracking efficiently.

Experimental results using 2D retinal data demonstrated its superior performance over most of the existing methods. The proposed method enhanced the efficiency of full image CNN result, losing only minimal unconnected vessels. Also, it provided the tree structure of the candidate vessel map. This method would be more efficient if there are fewer pixels segmented in the tracking method as compared to the full CNN method. Therefore, it would provide feasible solution for 3D medical tasks, which is the next step of this research.

5. REFERENCES

[1] Ziyue Xu, Ulas Bagci, Brent Foster, Awais Mansoor, Jayaram K. Udupa, and Daniel J. Mollura, "A hybrid method for airway segmentation and automated measurement of bronchial wall thickness on CT," *Medical Image Analysis*, vol. 24, no. 1, pp. 1 – 17, 2015.

[2] Fei Zhao and R. Bhotika, "Coronary artery tree tracking with robust junction detection in 3D CT angiography," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, March 2011, pp. 2066–2071.

[3] Ziyue Xu, Fei Zhao, R. Bhagalia, and B. Das, "Generic rebooting scheme and model-based probabilistic pruning algorithm for tree-like structure tracking," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, May 2012, pp. 796–799.

[4] Yaroslav Ganin and Victor Lempitsky, "N4-fields: Neural network nearest neighbor fields for image transforms," in *Computer Vision – ACCV 2014*, vol. 9004 of *Lecture Notes in Computer Science*, pp. 536–551. Springer, 2015.

[5] D. Crisan and K. Li, "Generalised particle filters with Gaussian mixtures," *Stochastic Processes and their Applications*, vol. 125, no. 7, pp. 2643 – 2673, 2015.

[6] J.J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, Eds., pp. 1097–1105. Curran Associates, Inc., 2012.

[8] GiriBabu Kande, P.Venkata Subbaiah, and T.Satya Savithri, "Unsupervised fuzzy based vessel segmentation in pathological digital fundus images," *Journal of Medical Systems*, vol. 34, no. 5, pp. 849–858, 2010.

[9] Marios Vlachos and Evangelos Dermatas, "Multi-scale retinal vessel segmentation using line tracking," *Computerized Medical Imaging and Graphics*, vol. 34, no. 3, pp. 213 – 227, 2010.

[10] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, and S.A. Barman, "Blood vessel segmentation methodologies in retinal images: A survey," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 1, pp. 407 – 433, 2012.