Automatically labeling Circle of Willis vessels

Li Chen

Task description

The Circle of Willis (CoW): A set of cerebral arteries (9 types of vessels connecting as a ring) located at the center of brain (Fig 1).

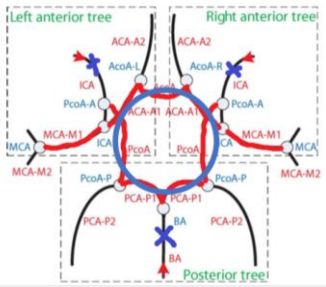

Fig 1 Anatomical structure of CoW vessels (along the blue circle) with vessel names (red labels), and bifurcation point names (blue labels) [1]

<u>Benefits</u> of automatically labeling CoW vessels

- Insightful knowledge of vascular structures on brain health
- Time and labor saving

<u>Difficulty</u> of labeling vessels

- Anatomical variation: vessels are irregular, missing or duplicated
- Unbalanced label (nine bifurcation types) and most type of "No label"
- Expensive data (Each scan cost ~$1000, preprocess needs ~1 hour)

Approaches

Approach 1: Classify bifurcation points based on bifurcation features

<u>Features</u>:

- *Bifurcation position:* normalized (offset due to different scan area) 3D position of bifurcations
- *Directional fitness*: angle fitness (maximum sum of inner product of all directions in a bifurcation) with 9 model bifurcations

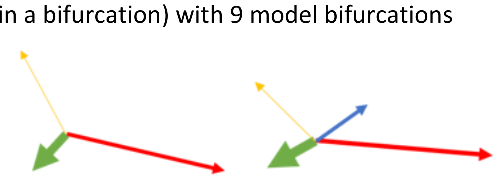

Fig 2 Illustration for finding corresponding vectors(each color is corresponding result). Searching from the pair of vectors with maximum product, remove the pair then search maximum product again until one bifurcation has no more vectors

- *Graph topological features*: a structural index NSI[2], which consists of node degree, radius, vessel likelihood (Frangi[3] filtered image intensity) as well as their nearby neighbor nodes

<u>Classifier</u>

Support vector machine (SVM) and logistic regression (LR) were used as classifiers. Class weight inversely proportional to class frequencies.

Inverse of regularization strength was found to be 2.0 for best accuracy in LR. Penalty parameter of error term of 13 and linear kernel were found to have best accuracy in SVM.

Approach 2: Find most probable bifurcation for each type of label.

<u>Features:</u>

- *Positional likelihood*: Assume Gaussian distribution of position of each bifurcation type and calculate model positions of bifurcations.
- *Directional fitness*: Angle fitness for certain bifurcation type
- *Graph topological likelihood:* Assume Gaussian distribution of NSI

<u>Sequential Labeling:</u>

Calculate the combined likelihood of each bifurcation to be each label type.

Start labeling the type with highest possibility, update offset between model and test position based on labeled bifurcation, refresh positional likelihood.

Label the bifurcation with highest likelihood if the possibility is greater than a threshold, which tested to be 0.0001.

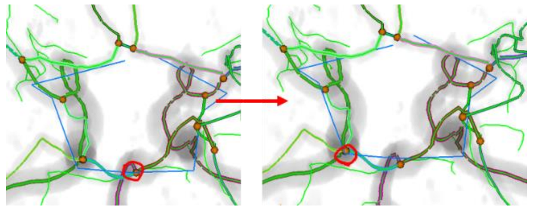

Fig 3 Update position after each label (blue lines are model positions, red balls are possible bifurcations, red circles show position after offset during each labeling)

Approach 3: Find branches by sum of possibility in connected bifurcations

In addition to features in approach 2, also include possibility from Gaussian distribution of branch length and average radius. Labeling based on sum of maximum possibility from two end bifurcations.

Experiments

Data:

Images from vascular imaging lab of UW, 24 patients (aging 16-82, mean 59, most have severe vascular disease, bad condition of CoW) with twice scan

CoW vessels have been labeled manually as ground truth.

Software:

Vessel graphs were extracted from raw MRI using iCafe, a tool I have previously developed. Python with scikit-learn was used for training.

Evaluation metric:

Due to limited data, leave-one-out cross-validation was used.

Four metrics were used to evaluate the label results.

- *Accuracy*: The percentage of the predicted bifurcation type agreed with ground truth.

To have a better idea of how the label was incorrect, disagreed situations were evaluated with three metrics:

- *Missing rate*: The percentage of the bifurcations having labels in ground truth, but predicted as "No label".
- *Wrong label rate*: The percentage of the bifurcations predicted to have a label but actually they have "No label" in ground truth.
- *Mistake rate*: The percentage of the bifurcations labeled in ground truth, but predicted as another type of label.

Results

<u>Metric evaluation:</u>

Approach	Accuracy	Missing rate	Wrong label rate	Mistake rate
1 SVM	96.70%	0.51%	1.74%	1.00%
1 LR	97.24%	0.51%	1.65%	0.59%
2	98.51%	0.82%	0.41%	0.25%
3	97.92%	0.98%	0.66%	0.43%

Fig 4 Evaluation of three approaches

<u>Feature selection:</u>

Approach	with NSI	without NSI
1 SVM	96.70%	95.96%
1 LR	97.24%	96.30%
2	98.51%	98.48%
3	97.92%	97.71%

<u>Mistake Area:</u>

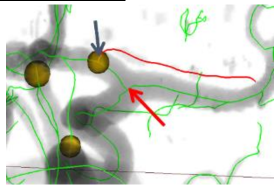

Fig 5 (left) Correctness with/without graph topological features (NSI)
Fig 6 (right) Mistakes occur where region, direction and topology are all similar

<u>Compare with other research:</u>

Only limited research for labeling CoW, none labeling mainly diseased cases.

One research reported bifurcation accuracy of 95% and 58% (29/50) success rate for completely labeling the entire CoW[1]

While my success rate is 2.1% (1/48) using approach 2

<u>Label result display:</u>

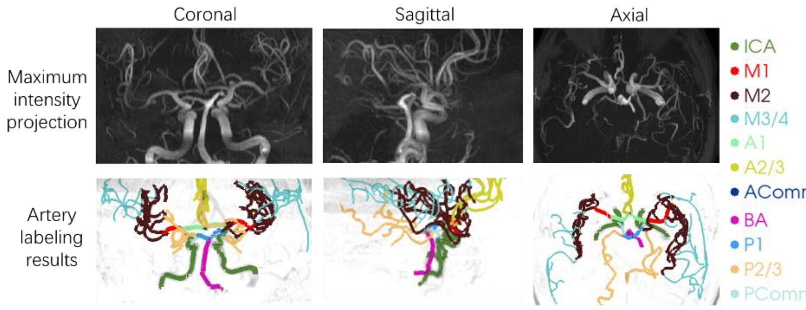

Fig 7 Labeling result for a success case

Summary of findings

The probability approach of finding each bifurcation type has highest accuracy. Graph topological features are useful in labeling bifurcations.

Reference:

[1] H. Bogunovic, *IEEE TMI*, 2013. [2] L. Chen, *J. Theor. Biol.*, 2015.

[3] A. F. Frangi, *Medical Image Comput.*, 1998.