Derek Burrows
Sam Powell
Adina Stoica

MIDDLE CEREBRAL ARTERY ANEURYSM CLASSIFICATION

1. INTRODUCTION

Nearly 6 million, or 1 out of every 50, people in the United States suffer from an unruptured cerebral aneurysm, a blood-filled sac resulting from a weakening in the wall of a blood vessel in the brain. If left untreated, cerebral aneurysms may rupture, resulting in severe hemorrhaging that can in turn cause permanent neurological damage and even lead to death. Treatment of a cerebral aneurysm entails an open-brain surgery in which a neurosurgeon places a number of highly-specialized clips of varying types near the base of an aneurysm, thereby cutting off blood-flow, removing the risk of rupture, and restoring normal blood flow through the surrounding blood vessels. At the moment, decisions about the necessary clipping strategy are only made during the surgery, i.e. after the aneurysm has been exposed by the neurosurgeon.

Recent work by Dr. Ralph Dacey of the Washington University School of Medicine seeks to shift this decision making process prior to surgery via the visual analysis of three-dimensional aneurysm models segmented from brain angiograms. His work seeks to establish cerebral aneurysm equivalence classes, where aneurysms are deemed equivalent, i.e. assigned to the same class, if they require the similar surgical clipping strategies to be sufficiently resolved. We propose to automate this classification process via the use of shape analysis techniques paired with machine learning over a small set of manually categorized aneurysm meshes.

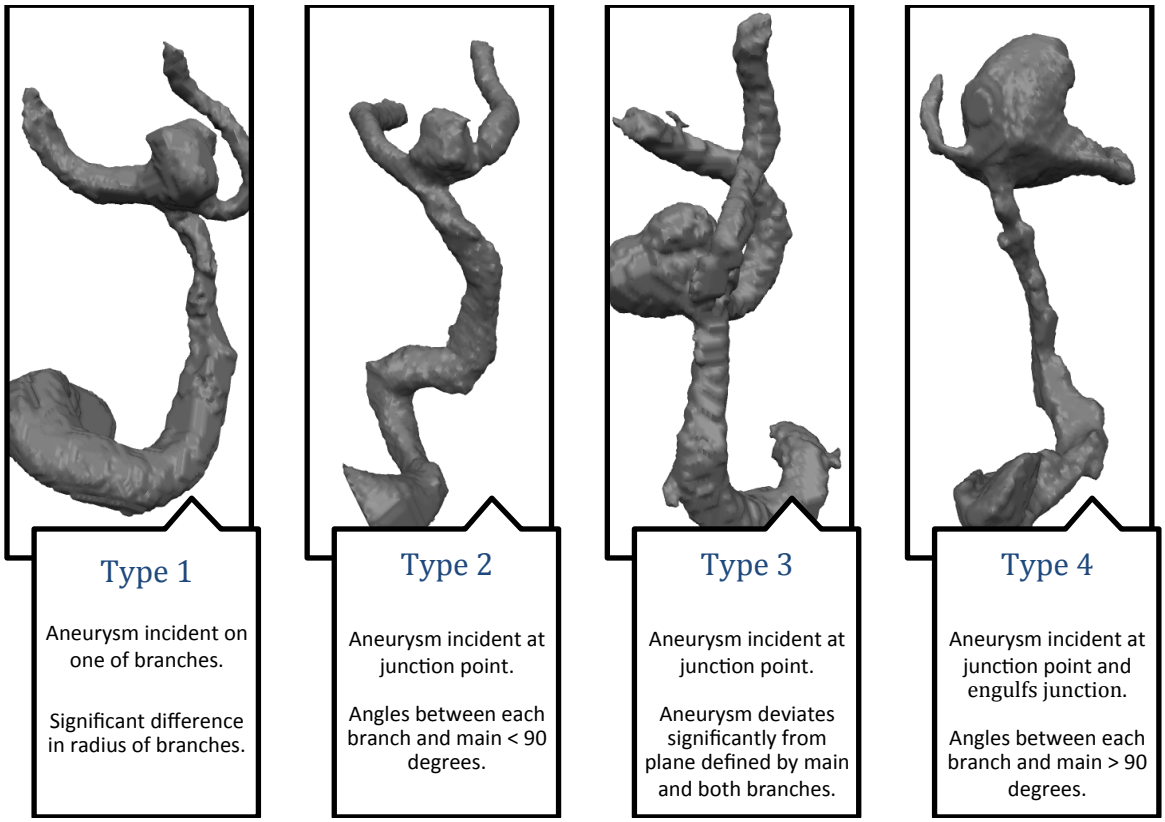

Figure 1: Aneurysm Classes and Verbal Feature Descriptions

2. METHODS

2.1 Feature Extraction

To perform classification of a given three-dimensional aneurysm mesh, we extract 5 scale-invariant metrics which were developed based on verbal feature descriptions made by domain-experts during visual analysis and manual classification. An overview of the aneurysm classes and verbal feature descriptions is provided in **Figure 1**. A description of the main and branch annotations (M1, M2, M3) is provided in **Figure 2**. The features are as follows:

a) Difference in angle between the main (M1) and first branch (M2)
b) Difference in angle between the main (M1) and second branch (M3)
c) Angle of deviation between the junction plane and the aneurysm
d) Ratio of the average radii between the aneurysm and the main (M1)
e) Ratio of the average radii between the smaller of the first and second branches (M1/M2) and the larger of the first and second branches (M1/M2)

The procedures for calculating these metrics are explained below.

2.1.1 Skeletonization

Given an input aneurysm mesh, we first perform skeletonization using the algorithm proposed by [Liu 2010]. The algorithm yields an approximation of the medial axes of the mesh, i.e. the set of all points with more than one closest point on the mesh boundary. The medial axes are given by a polyline, consisting of a set of vertices and their connectivity. Each element in the polyline is also ascribed two properties, a removal time and an isolation time which respectively reflect the radius/thickness (maximum distance to the mesh boundary perpendicular to the medial axis) and the elongation/length (minimum distance to the mesh boundary parallel to the medial axis) at the location of the element. A third property, reflecting the tubularity (degree of difference between elongation/length and radius/thickness) at the location of a given element, can also be obtained by simply calculating the difference between the removal time and isolation time and then normalizing that difference. For more details about the algorithm and medial axis properties, we refer the reader to the referenced paper.

2.1.2 Skeleton Segmentation and Pruning

The next step of our algorithm identifies the point on the medial axis where the main blood vessel (M1) begins to branch. We do this by simply selecting all of the points with a degree of three, i.e. those points connected to three neighbors. Of the degree-three points, we then select the one whose immediately connected edges have a maximum combined tubularity property. We designate this point as the junction point.

Using this junction point, we then segment the medial axis into three smaller polylines, one corresponding to the main blood vessel (M1) and two corresponding to each of the branch blood vessels (M2 and M3). We do this by performing three traces that respectively start at each of the immediately connected neighbors of the junction point. Each trace collects vertices and edges as it iterates over the medial axis by repeatedly selecting edges with the maximum elongation property. These segmented polylines are shown in **Figure 2**.

With these segmented polylines, we are then able to calculate the average radius of main vessel (M1) and branch vessels (M2/M3) by simply taking an unweighted average of the removal times over all of the edges in the appropriate polyline.

We then identify the point on the medial axis where the aneurysm begins and perform a trace, in a similar fashion as before, to extract a polyline for the aneurysm.

2.1.3 Vector and Junction Plane Calculation

Our method then calculates four vectors corresponding to the principal direction of divergence of the main vessel, the branch vessels, and the aneurysm from the junction area. We do this by first calculating, for each of the segments, a set of vectors between each pair of immediately connected points in the segment within some user-specified threshold of the junction plane. We then take, again for each segment, an unweighted average of these pairwise vectors to obtain the final divergence vector. For each segment, we used a threshold of 10% of the number of points in the segment to calculate the pairwise vectors.

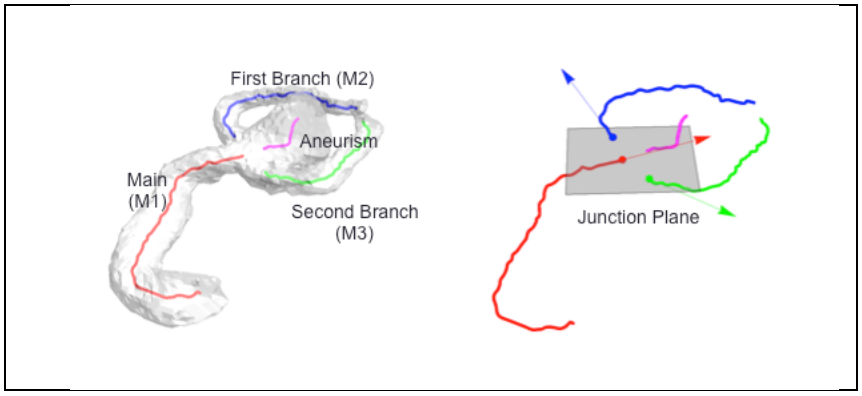

Figure 2: Annotated Aneurysm and Vector and Junction Plane Visualization

Using the vectors corresponding to the main vessel and each of the branch vessels, we estimate the best-fitting plane containing those vectors, which we call the junction plane. An example aneurysm mesh and the calculated vectors and junction plane are shown in **Figure 2**.

2.2 Classification

To perform classification using the aforementioned features, we selected the "random forests" machine-learning algorithm first introduced in [Breiman 2001]. The algorithm uses an ensemble of bagged decision trees. This algorithm is particularly appealing for our application due to the high-variance inherent in our model, i.e. a small number of features and a small set of unevenly distributed training examples, as well as the algorithm's ease of performing multi-class classification. Further, the algorithm allows for the calculation of an out-of-bag estimate of the test, i.e. read-world error, meaning a split of the dataset into training and test sets is unnecessary and all examples in the dataset may be used to train the classifier.

For the random forest algorithm, we used a tree count of 1000 and did not place constraints on the tree depth, i.e. trees were grown until all training data was correctly classified.

We originally wrote our own random forests classifier in C++ for MATLAB, which ran a lot faster than MATLAB's TreeBagger class. Unfortunately, we wrote our binary trees to be balanced (for ease of use with MATLAB), which meant that we could never drive our training error to 0. We didn't have time to modify the code to use unbalanced trees, and instead decided to use the MATLAB program to run our data. We are providing the C++ code in the archived project files also, but in order to get the reported results you should run the file useMATLAB.m.

3. DATASET AND RESULTS

We were provided a total of 48 aneurysm meshes that had been manually classified by domain-experts. The distribution of manual classifications of the meshes is shown in the second column of **Table 1**. Using random forests on this dataset, we achieved a training error of 0 for each of the 4 types, i.e. we

were able to correctly classify all of the training data. The out-of-bag estimates of the test error are shown in the third column of **Table 1**.

Aneurysm Class	Number of Examples	Out-of-Bag Error
Type 1	8	0.56
Type 2	22	0.33
Type 3	7	0.84
Type 4	11	0.48
Overall	**48**	**0.31**

Table 1: Dataset by Class and Results

4. DISCUSSION AND FUTURE WORK

The random chance achieves a classification accuracy of 0.25. For all of the aneurysm classes except Type 3, we were able to achieve an out-of-bag error better than random chance (0.75). Due mostly to the uneven distribution of the number of training examples (specifically Type 2), we were able to achieve an overall test accuracy of approximately 70%.

In order to increase this accuracy in the future, we believe first that additional training examples are needed for the deficient classes, 1 and 3. The additional examples will help to reduce the high variance and rectify the uneven training distribution. Additionally, we believe that more discriminative features are necessary. After examining the current feature set, we realized that many, taken alone, do not allow for perfect splitting between classes. Specifically, there are many aneurysms classified as Type 2 in the training dataset that exhibited characteristics of a Type 1 aneurysm, i.e. significant difference in branch radii. Because the ratio of branch radii is the predominant feature used for classifying Type 1 aneurysms against other classes, we experience a large error for that type. Further, we realized that there are many ambiguous cases, like the aforementioned example, where an aneurysm exhibits characteristics of multiple types. We believe that having domain-experts reclassify the training data in a fuzzy manner, i.e. providing multiple types in ambiguous cases, would be beneficial not only for test error evaluation but could be incorporated into the classifier training, as well.

5. CONCLUSIONS

In closing, we propose an automated algorithm for classification of aneurysm meshes using a combined shape analysis and machine learning approach. We obtain an overall classification accuracy of 0.7, while doing better on some aneurysm types for which we have more training data and better features, and worse on the ones that do not meet these criteria. We hope that in the future, with the addition of more examples, better classified/fuzzy classified examples and better features, we will be able to increase our accuracy and provide an accurate method for automatically deciding the aneurysm clipping procedure before the actual surgery, using 3D aneurysm models obtained from angiograms.

6. REFERENCES

1. [Breiman 2001] Breiman, Leo. "Random forests." *Machine learning* 45.1 (2001): 5-32.
2. [Liu 2010] Liu, Lu, et al. "A simple and robust thinning algorithm on cell complexes."*Computer Graphics Forum*. Vol. 29. No. 7. Blackwell Publishing Ltd, 2010.
3. http://www.bafound.org/node/124