Rupture Status Classification of Intracranial Aneurysms Using Morphological Parameters

Uli Niemann, Philipp Berg, Annika Niemann, Oliver Beuing, Bernhard Preim, Myra Spiliopoulou, Sylvia Saalfeld

Otto-von-Guericke University Magdeburg, Germany

Corresponding email: {uli.niemann, sylvia.saalfeld}@ovgu.de

Abstract—Intracranial aneurysms are pathologic dilations of the vessel wall, which bear the risk of rupture and of fatal consequences for the patient. Since treatment may be accompanied by severe complications as well, rupture risk assessment and thus rupture risk prediction plays an important role in clinical research. In this work, we investigate the potential of morphological features for rupture risk status classification in 100 intracranial aneurysms. We propose a pipeline for morphological feature extraction and rupture status classification with subsequent feature ranking and inspection. Our classification setup involves training separate models for each aneurysm type (sidewall or bifurcation) with multiple learning algorithms. We report on the classification performance of our pipeline and examine the predictive power of each morphological parameter towards rupture status classification. Further, we identify the most important features for the best models and study their marginal prediction.

Keywords-Medical Image Analysis; Intracranial Aneurysm; Morphological Parameters; Rupture Status Classification;

I. INTRODUCTION

Intracranial aneurysms are pathologic dilations of the intracranial vessel wall. They bear the risk of rupture and thus subarachnoidal hemorrhages with often fatal consequences for the patient. Since treatment may cause severe complications as well, substantial research was carried out to characterize the patient-specific rupture risk based on various parameters, including whether the aneurysm is asymptomatic or symptomatic, age, gender, size and location [1].

Numerous studies identified parameters such as aspect ratio, undulation index and nonsphericity index as statistically significant to the aneurysm rupture status [2], [3]. However, although these studies allow for a retrospective analysis, the clinician needs further guidance in case an asymptomatic aneurysm was detected and the rupture risk should be determined. Therefore, we study to what extent the rupture status can be predicted with supervised classification models learned on morphological parameters that are available to the clinician. We extract common parameters that were found in other studies to correlate with rupture risk. We train separate classifiers for each aneurysm type (sidewall or bifurcation) using multiple learning algorithms. We juxtapose the impact of each feature towards the best models' predictions and study their marginal predictions.

II. MATERIALS

Our pipeline is illustrated in Fig. 1. The preprocessing steps comprise the segmentation and neck curve extraction of the medical image data, which is explained in this section.

A. Data Acquisition

The presented methods were developed for a database comprising 74 patients (age: 33-85 years, 17 male and 57 female patients) with 100 intracranial aneurysms at the university hospital of Magdeburg, Germany. Each patient had 3D rotational angiography on an Artis Q (Siemens Healthineers, Forchheim, Germany) as part of the necessary clinical work-up performed. The dataset acquisition and the subsequent analysis steps are in accordance with the guidelines of the local ethics committee.

B. Segmentation and Neck Curve Extraction

Aneurysms and vessels are segmented with a threshold-based approach [4] from digital subtraction data reconstructed from 3D rotational angiographies. Afterwards, we extract the vessel's centerline with the vascular modeling toolkit (VMTK, vmtk.org) [5]. As prerequisite for feature extraction, the aneurysm has to be delineated from the parent vessel. For this purpose, we employ the automatic ostium detection of Saalfeld et al. [6].

III. METHODS

With our pipeline, we extract morphological features and learn classification models that distinguish between unruptured and ruptured intracranial aneurysms.
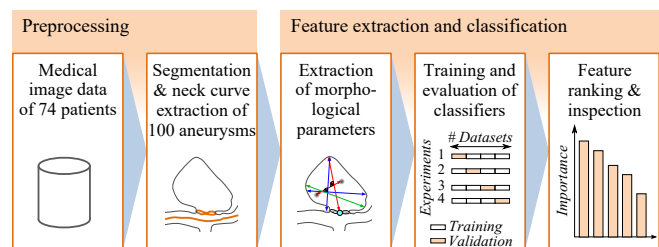

Figure 1. Pipeline of our presented approach: aneurysm segmentation from raw image data as well as subsequent automatized extraction of the centerline and the neck curve was carried out. Afterwards, the morphological parameters were extracted. Classifiers were trained and evaluated based on these parameters. Finally, feature ranking and inspection was conducted.
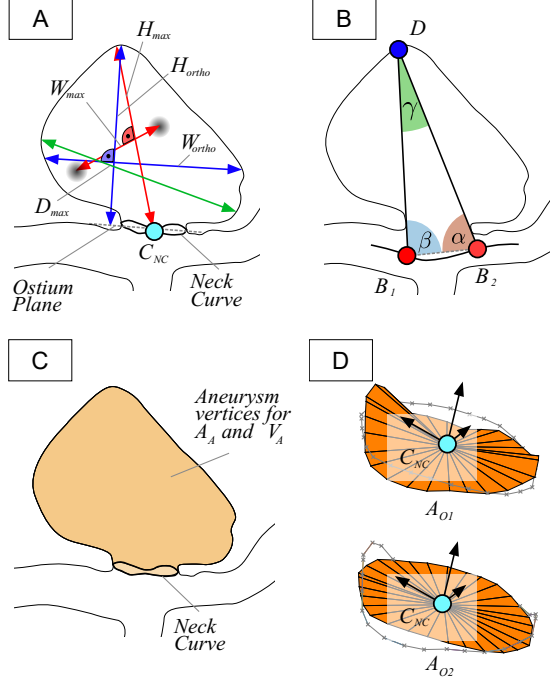
Figure 2. Illustration of the extracted morphological features H_{max}, W_{max}, H_{ortho}, W_{ortho} and D_{max} (A). The angles α, β and γ are extracted based on B_1, B_2 and the dome point D (B). Separating the aneurysm from the parent vessel based on the neck curve yields A_A and V_A (C). The area of the ostium and the projected ostium, i.e., A_{O1} and A_{O2}, are shown in (D), where C_{NC} denotes the center of the neck curve.

A. Extraction of Morphological Features

For each 3D surface mesh, we obtain the neck curve, the dome point D and the two base points B_1 and B_2 as described in our previous work [6]. Hence, B_1 and B_2 are approximated as points on the centerline with largest distance where the rays from B_1 and B_2 to D are not intersecting the surface mesh. In Fig. 2, the extracted parameters are illustrated, where H_{max}, W_{max}, H_{ortho}, W_{ortho} and D_{max} describe the aneurysm shape [2], [7]. The angles α, β and γ are extracted based on B_1, B_2 and D, and $\Delta_{\alpha\beta}$ denotes the absolute difference of α and β.

Separating the aneurysm from the parent vessel based on the neck curve allows for extraction of the surface area A_A and the volume V_A of the aneurysm. We provide two measurements for the surface area of the ostium: A_{O1} and A_{O2} (recall Fig. 2-D). The parameter A_{O1} denotes the surface area of the ostium, i.e., the surface of the triangulated ostium area obtained by connecting the neck curve points with their mean C_{NC}. The parameter A_{O2} denotes the surface area of the neck curve when projected into a plane, see also [6]. Hence, A_{O2} is extracted to obtain a comparable parameter to other studies that often employ a cutting plane to determine the ostium.

In case of a severely lobulated aneurysm, our method gets stuck in a local optimum and considers only one of the
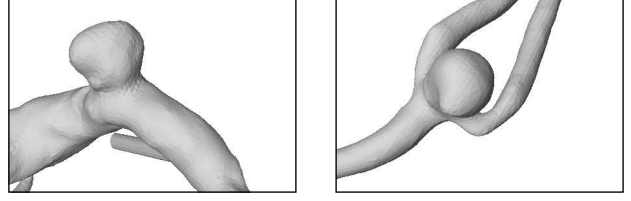

Figure 3. Illustration of a sidewall aneurysm at the side of the parent vessel wall (left) and a bifurcation aneurysm at a vessel bifurcation (right).

multiple dome points. Although the extracted positions of B_1 and B_2 could vary slightly, the neck curve extraction can be carried out and the parameters can be extracted.

Table I provides a brief description, summary statistics and a visualization of the distribution for each extracted parameter. Additionally, the statistical significance of multiple parameters w.r.t. rupture risk status, including D_{max}, H_{max}, H_{ortho}, AR_1, AR_2, β and γ, is shown in Table I.

B. Classification & Evaluation Setup

We model our classification problem with a two-class target feature (*unruptured* vs. *ruptured*), using the 100 samples described in subsection II-A and the 22 morphological parameters from Table I as input features. As motivated by the results reported by Baharoglu et al. [8], we learn distinct models for the subset of sidewall aneurysms (SW; 9 ruptured out of 24) and for the subset of bifurcation aneurysms (BF; 29 ruptured out of 62). See Fig. 3 for an illustration of sidewall and bifurcation aneurysms. Further, we train our classifiers on a combined group (ALL; 43 ruptured out of 100), which includes 14 additional samples that could not be clearly identified as either sidewall or bifurcation aneurysms.

For classification, we employ 10 algorithms: three decision tree variants (CART [9], C4.5 [10] and its successor C5.0), Naïve Bayes (NBayes), a feed-forward neural network with a single hidden layer (NNET [11]), generalized partial least squares (GPLS) [12], random forest (RF) [13], k-nearest neighbor classifier (KNN), support vector machine with linear kernel (SVMLin) and gradient boosted trees (GBT) [14]. Some of the chosen classifiers may struggle because of the different scale of the variables or the dataset's high dimensionality. Therefore, we optionally engage three preprocessing transformations to the dataset: the "range" transformation scales the data to be within [0,1], the "z-score" transformation subtracts the mean and divides by the standard deviation, and the "pca" transformation performs "z-score" and a principal component analysis and retains only the first principal components where the cumulative explained variance first exceeds 95 %.

We evaluate our approach using 5 times repeated 10-fold stratified cross-validation. The classifier's hyperparameters are tuned by means of a grid search using accuracy (ratio of correctly labeled samples) as quality measure.

Feature	Description	Status	$\bar{x} \pm s$	Distribution	p-value
A_A	Area of the aneurysm (without the ostium) $[mm^2]$	Unruptured Ruptured	79.40 ± 86.77 92.26 ± 74.12		0.050
V_A	Volume of the aneurysm $[mm^3]$	Unruptured Ruptured	91.24 ± 151.78 91.72 ± 114.92		0.092
A_{O1}	Area of the ostium (variant 1) $[mm^2]$	Unruptured Ruptured	11.34 ± 7.90 10.33 ± 9.56		0.265
A_{O2}	Area of the ostium (variant 2) $[mm^2]$	Unruptured Ruptured	10.45 ± 6.99 9.41 ± 8.29		0.262
D_{max}	Max. diameter of the aneurysm $[mm]$	Unruptured Ruptured	6.24 ± 2.84 7.21 ± 2.77		**0.034***
H_{max}	Max. height of the aneurysm $[mm]$	Unruptured Ruptured	4.74 ± 2.54 5.88 ± 2.59		**0.012***
W_{max}	Max. width of the aneurysm perpendicular to H_{max} $[mm]$	Unruptured Ruptured	5.08 ± 2.66 5.32 ± 2.11		0.119
H_{ortho}	Height of the aneurysm approximated as length of the ray perpendicular to the ostium plane starting from C_{NC} $[mm]$	Unruptured Ruptured	4.26 ± 2.41 5.17 ± 2.41		**0.030***
W_{ortho}	Max. width parallel to the projected ostium plane $[mm]$	Unruptured Ruptured	5.59 ± 2.67 5.94 ± 2.15		0.175
N_{max}	Max. NC diameter, i.e., the max. possible distance between two NC points $[mm]$	Unruptured Ruptured	4.02 ± 1.49 3.75 ± 1.25		0.330
N_{avg}	Avg. NC diameter, i.e., the mean distance between C_{NC} and the NC points $[mm]$	Unruptured Ruptured	3.51 ± 1.20 3.30 ± 1.13		0.237
AR_1	Aspect ratio: H_{ortho}/N_{max}	Unruptured Ruptured	1.08 ± 0.50 1.40 ± 0.56		**0.002****
AR_2	Aspect ratio: H_{ortho}/N_{avg}	Unruptured Ruptured	1.23 ± 0.56 1.60 ± 0.65		**0.003****
V_{CH}	Volume of the convex hull of the aneurysm vertices $[mm^3]$	Unruptured Ruptured	100.36 ± 160.37 108.14 ± 129.38		0.085
A_{CH}	Area of the convex hull of the aneurysm vertices $[mm^2]$	Unruptured Ruptured	92.23 ± 93.64 105.25 ± 81.15		0.079
EI	Ellipticity index: $1 - (18\pi)^{\frac{1}{3}} V_{CH}^{\frac{2}{3}}/A_{CH}$	Unruptured Ruptured	0.27 ± 0.02 0.27 ± 0.02		0.323
NSI	Non-sphericity index: $1 - (18\pi)^{\frac{1}{3}} V^{\frac{2}{3}}/A$	Unruptured Ruptured	0.17 ± 0.25 0.21 ± 0.24		0.067
UI	Undulation index: $1 - \frac{V}{V_{CH}}$	Unruptured Ruptured	0.12 ± 0.35 0.10 ± 0.39		0.686
α	Min. of $\angle DB_1B_2$ and $\angle DB_2B_1$ $[deg]$	Unruptured Ruptured	54.38 ± 12.63 56.90 ± 11.57		0.287
β	Max. of $\angle DB_1B_2$ and $\angle DB_2B_1$ $[deg]$	Unruptured Ruptured	80.70 ± 17.07 92.13 ± 17.29		**<0.001****
γ	Angle at D, i.e. $\angle B_1DB_2$ $[deg]$	Unruptured Ruptured	44.93 ± 19.71 30.98 ± 13.68		**<0.001****
$\Delta_{\alpha\beta}$	Abs. difference between α and β $[deg]$	Unruptured Ruptured	26.32 ± 22.66 35.23 ± 26.05		0.078

IV. Results

In this section, we report on our classification results and elaborate on the features with high model impact.

A. Rupture Status Classification

In Table II, we show the results of classification between *unruptured* aneurysms (negative class) and *ruptured* ones (positive class) on all samples (denoted as ALL), the subset of sidewall aneurysms (SW) and the subset of bifurcation aneurysms (BF). For each subset-algorithm combination, we only show the preprocessing transformation with highest accuracy. In addition to accuracy, we depict two other quality measures: Cohen's kappa and the area under the receiver operating characteristic (ROC) curve. Cohen's kappa measures the model's relative improvement in accuracy in comparison with the expected accuracy of a baseline classifier that randomly predicts the class label. A ROC curve is a graphical representation that juxtaposes sensitivity (true positive rate (TPR)) and false positive rate (FPR) for varying prediction thresholds of a binary classifier. The area under the ROC curve (AUC) takes values between between 0 (0 % TPR, 100 % FPR) and 1 (100 % TPR, 0 % FPR), where a random classifier achieves an AUC of 0.5.

Table II
CLASSIFICATION PERFORMANCE FOR EACH COMBINATION OF DATA SUBSET AND ALGORITHM. PREPR. = PREPROCESSING TRANSFORMATION; ACC. = ACCURACY; SW/BF = SIDEWALL/BIFURCATION SUBSET.

Subset	Algorithm	Prepr.	Acc.	Kappa	AUC
ALL	GBT	-	.69±.15	.36±.32	.70±.02
	C5.0	-	.66±.16	.32±.31	.68±.04
	GPLS	-	.66±.15	.28±.31	.69±.01
	KNN	range	.66±.16	.27±.34	.63±.01
	CART	-	.65±.12	.29±.26	.61±.03
	NNET	range	.62±.13	.23±.26	.64±.03
	C4.5	-	.62±.15	.24±.31	.64±.03
	NBayes	-	.61±.16	.21±.32	.58±.03
	SVMLin	-	.60±.15	.17±.32	.56±.02
	RF	-	.60±.12	.16±.26	.62±.01
SW	SVMLin	-	.80±.24	.50±.53	.66±.12
	GPLS	range	.79±.26	.49±.56	.73±.03
	C5.0	-	.78±.30	.50±.60	.83±.03
	KNN	pca	.77±.21	.38±.49	.73±.04
	GBT	-	.77±.27	.49±.55	.68±.03
	NNET	range	.75±.28	.42±.58	.69±.05
	CART	-	.75±.34	.47±.66	.63±.06
	RF	-	.74±.28	.40±.55	.69±.05
	NBayes	-	.72±.32	.45±.59	.68±.03
	C4.5	-	.72±.32	.40±.61	.75±.06
BF	GPLS	center, scale	.68±.16	.34±.32	.68±.02
	KNN	pca	.66±.19	.33±.38	.68±.02
	NNET	center, scale	.63±.18	.25±.37	.65±.02
	SVMLin	-	.63±.17	.26±.33	.62±.04
	GBT	-	.62±.15	.24±.29	.59±.04
	NBayes	-	.61±.19	.24±.37	.57±.03
	CART	-	.61±.18	.22±.34	.61±.04
	C4.5	-	.60±.15	.21±.29	.62±.03
	RF	-	.60±.16	.19±.32	.63±.02
	C5.0	-	.59±.16	.19±.32	.61±.02

GBT performs best in accuracy on ALL with 69 %, followed by C5.0, GPLS and KNN with each 66 % (cf. Table II). All classification algorithms work better on the subset of SW aneurysms where SVMLin with a range transformation yields best overall accuracy of 80 % with an AUC of 0.66. For the BF subset, GPLS with a z-score transformation performs best with an accuracy of 68 % and an AUC of .68. Overall, none of the classification algorithms outperforms all others on all three subsets.

We integrated the three best models per data subset in an interactive web-application[1], allowing the user to study how a change in the choice of the values for a feature affects the model confidence. The user may select a sample from the training set, but she is also allowed to freely modify each feature value with the respective slider widget, whereupon the prediction confidence gets immediately updated. A horizontal bar chart visualizes the supporting or contradicting contribution of each input feature to the model's prediction, based on the method of Ribeiro et al. [15].

B. Model Interpretation

For interpretation of our models, we analyze the best model of each subset. We identify the most important features, i.e., the features that have the highest impact on the model prediction (cf. Fig. 4). Feature importance
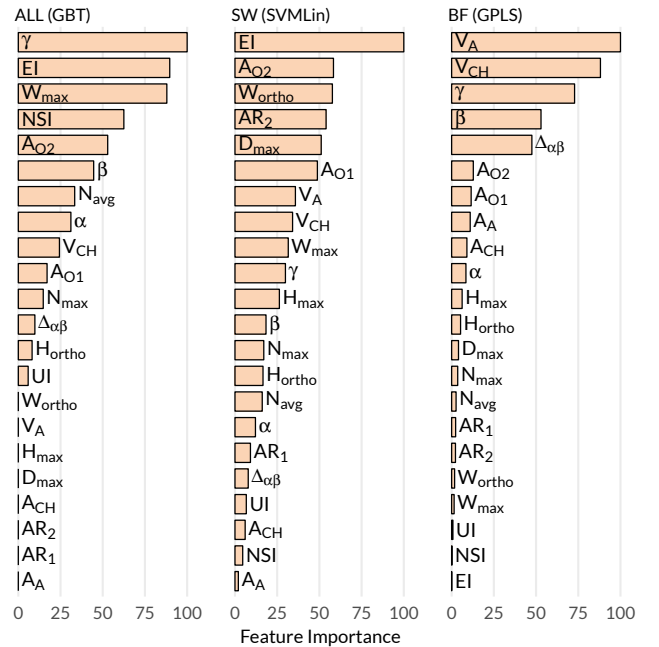
Figure 4. Feature importance for the best model of each data subset. Values are scaled up to 100 according to the highest feature importance.

scores are calculated as the sum of absolute differences between all pairs of consecutive points of the feature's partial dependence plot. A partial dependence plot visualizes the

[1] Available at https://rbsenzaehler.shinyapps.io/RUSTiC/.

relationship between a feature f and a model's prediction while incorporating the average effect of the remaining features in the model. Thus, by providing an estimation of how the model's prediction changes for each value of f, it is particularly useful for interpreting the models of complex non-parametric classification algorithms. In Fig. 5, we show the partial dependence plots of the five most important features of the best models per subset from Table II.
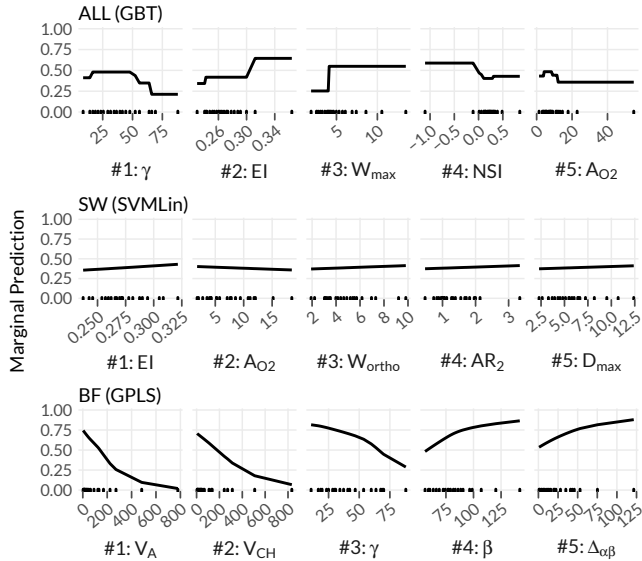

Figure 5. The partial dependence plots show the marginal prediction of the five most important variables for the best model of each data subset.

Fig. 5 illustrates the high importance of the angle γ w.r.t. rupture status classification, since this feature is ranked #1 and #3 for the best models of ALL and BF. Although the curves show slight variations, both models have in common that the marginal probability of predicting the class *ruptured* decreases with increasing values of γ, which means that ruptured aneurysms tend to exhibit lower γ values (cf. Table I). Besides γ, other features, such as EI for ALL and SW, W_{max} for ALL, A_{O2} and W_{ortho}, as well as V_A and V_{CH} for BF exhibit a considerable relevance w.r.t. the model, as shown in Fig. 4 and 5.

V. DISCUSSION

Fig. 5 shows that there is no intersection among the top-5 features found for the sidewall aneurysms (middle subfigure) and for bifurcation aneurysms (bottom subfigure). Although the respective classification algorithms are different, this result might indicate that the interdependencies between each of those input features and the target feature follow different patterns. This is in agreement to Baharoglu et al. [8] who identified differences in the statistical significance of morphological parameters w.r.t. rupture status between sidewall and bifurcation aneurysms.

The importance scores in Fig. 4 indicate some higher-level feature interactions. For example, although mean and standard deviation of EI are equal for both classes (cf. Table I), it is ranked most important in SW and 2nd most important in ALL. Apparently, the model combines EI with other features to generate more robust predictions.

Although the partial dependence plot clearly displays the influence of a single feature, interaction effects between two or more features are not represented. The decision tree variant `C5.0` achieves 66 % accuracy on all samples, just 3 % less than the best model (cf. Table II). Being more interpretable than `GBT` and `GPLS` and just slightly less accurate, we opt to study the model closer. In its best run, the algorithm generates multiple classification rules instead of a single tree. A representation of three of those class-characteristic rules is shown in Fig. 6. The left panel shows a rule with a high lift for unruptured samples that satisfy the condition $AR_2 \leq 1.24 \wedge \gamma > 54.16$. This partition (gray box) contains 20 unruptured and 2 ruptured samples. Thus, the lift is 1.6, i.e., in the partition described by the rule, the relative frequency of the class *unruptured* is 1.6 times higher than in the total training set. Two rules for the class *ruptured* are shown in the middle and right panel in Fig. 6 where the interaction effects between H_{ortho} and W_{ortho}, as well as between W_{max} and A_{O2} are highlighted. Fig. 6 also indicates that none of our morphological features alone would reliably predict the rupture risk of the aneurysms. Indeed, building a classifier with only one input feature yields a model with 55 % accuracy at best, which is outperformed by a classifier that predicts the majority class (57 % accuracy).

While we consider our results to be promising, there are arguably some limitations and substantial room for improvement. First, the limited sample size, in particular for the subset of sidewall aneurysms, might lead to overfitting of the classification models. We are keen to evaluate the robustness of our models on a larger number of datasets. The second limitation concerns the quality of the class label due to this specific pathology. Samples that were labeled as unruptured could rupture at a later moment. A further limitation is the limited feature space: our classification models incorporate morphological characteristics only. However, other properties, such as hemodynamic features have been identified to be predictive as well [16], [17]. In future work, we would like to inspect samples with high classification error. Here, our goal is to derive descriptions of groups of aneurysms that are hard to classify, in order to better understand the reasons for misclassification. Further, we would like to study the potential of deep learning on the segmented image data, e.g., using convolutional neural nets. More precisely, our interest is to quantify the positive or negative impact of a careful feature engineering step before rupture risk assessment in comparison with a model that is learned on raw image data.

VI. CONCLUSION

In this paper, we studied the potential of morphological parameters of intracranial aneurysms for rupture status clas-
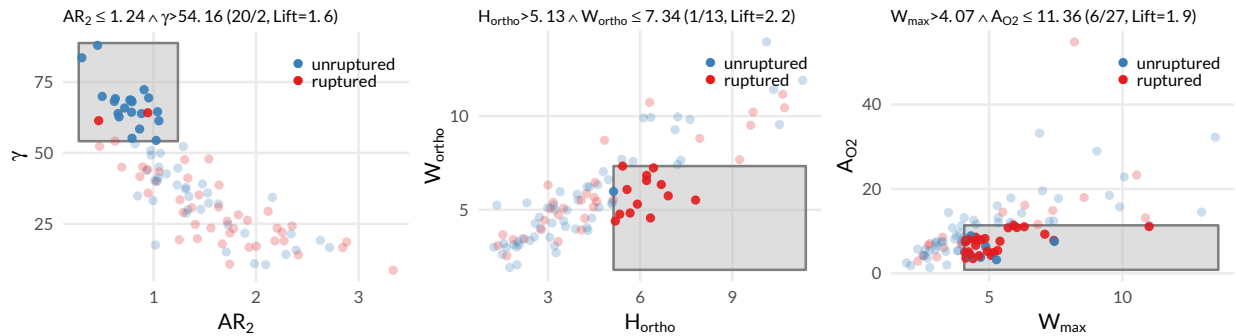
Figure 6. Three classification rules with high lift values extracted from `C5.0` for subset ALL. The title of a subfigure displays the rule's condition. The class counts and lift value of the partition are given in parenthesis. Samples that satisfy the rule condition are shown as opaque points within a gray box.

sification in order to reduce risky treatments in case of low-risk aneurysms. We found that some of the extracted parameters are highly predictive towards the outcome, including the angle between the ostium's base points γ. Although only trained on shape, size and angle features, the performance of our best models is promising. Thus, a natural extension of our work includes incorporating a broader range of feature types to further improve accuracy.

Acknowledgment

References

[1] M. J. Wermer, I. C. van der Schaaf, A. Algra, and G. J. Rinkel, "Risk of rupture of unruptured intracranial aneurysms in relation to patient and aneurysm characteristics," *Stroke*, vol. 38, no. 4, pp. 1404–1410, 2007.

[2] S. Dhar, M. Tremmel, J. Mocco, M. Kim, J. Yamamoto, A. H. Siddiqui, L. N. Hopkins, and H. Meng, "Morphology parameters for intracranial aneurysm rupture risk assessment," *Neurosurgery*, vol. 63, no. 2, pp. 185–197, 2008.

[3] J. Xiang, S. K. Natarajan, M. Tremmel, D. Ma, J. Mocco, L. N. Hopkins, A. H. Siddiqui, E. I. Levy, and H. Meng, "Hemodynamic–morphologic discriminants for intracranial aneurysm rupture," *Stroke*, vol. 42, no. 1, pp. 144–152, 2011.

[4] S. Glaßer, P. Berg, M. Neugebauer, and B. Preim, "Reconstruction of 3D Surface Meshes for Bood Flow Simulations of Intracranial Aneurysms," in *Proc. of the Annual Meeting of the German Society of Computer- and Robot-Assisted Surgery*, 2015, pp. 163–168.

[5] L. Antiga, M. Piccinelli, L. Botti, B. Ene-Iordache, A. Remuzzi, and D. A. Steinman, "An image-based modeling framework for patient-specific computational hemodynamics," *MBEC*, vol. 46, no. 11, p. 1097, 2008.

[6] S. Saalfeld, P. Berg, A. Niemann, M. Luz, B. Preim, and O. Beuing, "Semi-automatic neck curve reconstruction for intracranial aneurysm rupture risk assessment based on morphological parameters," in *Proc. of Computer assisted radiology and surgery (CARS)*, 2018, to appear.

[7] A. Lauric, M. I. Baharoglu, and A. M. Malek, "Ruptured status discrimination performance of aspect ratio, height/width, and bottleneck factor is highly dependent on aneurysm sizing methodology," *Neurosurgery*, vol. 71, no. 1, pp. 38–46, 2012.

[8] M. I. Baharoglu, A. Lauric, B.-L. Gao, and A. M. Malek, "Identification of a dichotomy in morphological predictors of rupture status between sidewall- and bifurcation-type intracranial aneurysms," *Journal of neurosurgery*, vol. 116, no. 4, pp. 871–881, 2012.

[9] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth and Brooks, 1984.

[10] R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.

[11] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*, 4th ed. Springer, 2002.

[12] B. Ding and R. Gentleman, "Classification using generalized partial least squares," *Journal of Computational and Graphical Statistics*, vol. 14, no. 2, pp. 280–298, 2005.

[13] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[14] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[15] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you?: Explaining the predictions of any classifier," in *Proc. of ACM SIGKDD Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.

[16] J. Cebral, F. Mut, J. Weir, and C. Putman, "Quantitative characterization of the hemodynamic environment in ruptured and unruptured brain aneurysms," *American Journal of Neuroradiology*, vol. 32, no. 1, pp. 145–151, 2011.

[17] P. Berg and O. Beuing, "Multiple intracranial aneurysms: a direct hemodynamic comparison between ruptured and unruptured vessel malformations," *International journal of computer assisted radiology and surgery*, vol. 13, no. 1, pp. 83–93, 2018.