Article

Assessing Machine Learning Models for Predicting Age with Intracranial Vessel Tortuosity and Thickness Information

Hoon-Seok Yoon [ID], Jeongmin Oh [ID] and Yoon-Chul Kim *[ID]

Division of Digital Healthcare, College of Software and Digital Healthcare Convergence, Yonsei University, Wonju 26493, Republic of Korea; hoonseok123@yonsei.ac.kr (H.-S.Y.); ohjeongmin@yonsei.ac.kr (J.O.)
* Correspondence: yoonckim@yonsei.ac.kr; Tel.: +82-33-760-2318

Abstract: This study aimed to develop and validate machine learning (ML) models that predict age using intracranial vessels' tortuosity and diameter features derived from magnetic resonance angiography (MRA) data. A total of 171 subjects' three-dimensional (3D) time-of-flight MRA image data were considered for analysis. After annotations of two endpoints in each arterial segment, tortuosity features such as the sum of the angle metrics, triangular index, relative length, and product of the angle distance, as well as the vessels' diameter features, were extracted and used to train and validate the ML models for age prediction. Features extracted from the right and left internal carotid arteries (ICA) and basilar arteries were considered as the inputs to train and validate six ML regression models with a four-fold cross validation. The random forest regression model resulted in the lowest root mean square error of 14.9 years and the highest average coefficient of determination of 0.186. The linear regression model showed the lowest average mean absolute percentage error (MAPE) and the highest average Pearson correlation coefficient (0.532). The mean diameter of the right ICA vessel segment was the most important feature contributing to prediction of age in two out of the four regression models considered. An ML of tortuosity descriptors and diameter features extracted from MRA data showed a modest correlation between real age and ML-predicted age. Further studies are warranted for the assessment of the model's age predictions in patients with intracranial vessel diseases.

Keywords: machine learning; magnetic resonance angiography; intracranial artery; age prediction; medical image analysis; feature extraction

1. Introduction

Aging and its related diseases affect individuals during their life spans. Due to improvements in healthcare and medical technology, life expectancy and the proportion of elderly people have increased worldwide recently [1]. As a measure of brain-related aging, predictions of age from brain image data have been increasingly investigated by numerous research groups [2–5]. A great deal of studies has been conducted using structural magnetic resonance imaging (MRI) data for brain age prediction. Brain image data from healthy subjects have been used to develop a machine learning (ML) model for predicting chronological age [6,7]. T1-weighted structural brain MRI data have been used to train and validate a deep convolutional neural network (CNN) architecture [8]. Region-of-interest (ROI) volumes extracted from T1-weighted structural data have been used as features for ML regression models [9]. T1-weighted brain MRI and computed tomography (CT) data were used in a past study [10]. Multimodal MRI images, including T1-weighted imaging, T2* relaxometry, and diffusion tensor imaging (DTI) have been used to extract features and predict brain age using a multiple linear regression analysis [11]. T1-weighted, fluid-attenuated inversion recovery (FLAIR), and susceptibility-weighted imaging (SWI) images have been used to train a three-dimensional (3D) CNN architecture [12]. There have also been association studies between the morphometric features available from MRI

image data and chronological age. Gray matter volume and fractional anisotropy tend to decrease with age [13]. High frequencies of white matter hyperintensities, microbleeds, or lacunar infarcts have been associated with aging [3].

MR angiography (MRA) or computed tomography angiography (CTA) provide insights different from those provided by commonly used structural T1-weighted brain image data. MRA or CTA image data facilitate the noninvasive assessment of vascular aging. Several studies have investigated the association between vascular tortuosity, kinking, and coiling and age [14–16]. A more detailed quantitative analysis of the cerebral vasculature can help assess chronological age [17]. The oldest age group tends to show a larger mean diameter of the major intracranial vessel than the youngest age group [17]. A lower number of unbranched small vessels has been observed in older age groups [17,18]. The length and meandering of bifurcating branches have been shown to increase with age [19], although the study did not involve an ML-based age prediction. Deep CNN models trained on healthy subjects' MRA have been demonstrated to estimate the biological aging of intracranial vessels in patients suspected of having vessel-related diseases. Nam et al. proposed a deep CNN model for predicting chronological age using healthy subjects' brain time-of-flight (TOF) MRA data [20]. Mouches et al. demonstrated high accuracy in brain age prediction by combining MRA data with T1-weighted data in deep CNN architectures [21]. They identified the basilar artery, the middle cerebral artery M2 segments, and the left posterior cerebral artery as the artery regions that mainly contribute to age prediction. It has been reported that arterial tortuosity is a risk factor for aneurysms or other vessel-aging-related diseases [14,22].

The extraction of morphometric features with a vessel analysis tool has been investigated to evaluate aging and disease conditions from the intracranial vasculatures [23]. With a similar purpose in mind, instead of using end-to-end deep CNN architectures, we proposed the use of tortuosity and diameter features extracted from a vessel segment in the region of the circle of Willis arteries from 3D TOF MRA image data, and we predicted subjects' ages using an ML regression model. We also investigated what image features would contribute the most to the age predictions in some regression models.

2. Materials and Methods

2.1. Data and Preprocessing

Three-dimensional TOF MRA data and age information were available in the IXI datasets (https://brain-development.org/ixi-dataset (accessed on 25 October 2023)), which were obtained from normal and healthy subjects. After 87 subjects' data were discarded, the data from 171 subjects were considered for our analysis. The data exclusion criteria were: (1) the image data were of poor quality, (2) the imaging field of view (FOV) was not sufficient along the superior-inferior direction to cover the basilar artery (BA) and internal carotid arteries (ICAs), or (3) the Dijkstra's path-finding result was incorrect in either the left or right ICA. The data with the insufficient FOV were discarded in our analysis because the selection of the two endpoints needed to be consistent anatomically for the feature extraction. Also, the data with incorrect path-finding results were discarded because incorrect path-finding would result in incorrect feature extraction, which would negatively affect age prediction. The incorrect path-finding result was attributed to the fact that the Dijkstra algorithm finds the shortest path in a centerline, and a highly tortuous vessel such as the ICA may produce false positive centerlines which would be shorter than the correct path. The 171 subjects had a mean age of 49.3 years with a standard deviation of 16.5 years (age range = 20–83). A total of 85 subjects were male, and the remaining 86 subjects were female.

The overall process for the age predictions, including image preprocessing, is summarized in Figure 1. A three-dimensional seeded region-growing was performed to segment the vessels. The morphology.skeletonize function of the Scikit-Image library [24] was used to obtain the 3D centerlines of the vessels. The Plotly (Plotly Technologies Inc., Montreal, QC, Canada) Python library (https://plotly.com/python (accessed on 25 October 2023))

was used to manually annotate the landmarks in the circle of Willis arteries [25]. The Dijkstra algorithm was used to find a path between two manually annotated endpoints. The vessel segments were labeled with colors, as shown in Figure 2, and they included the anterior communicating artery (ACOMM), left and right anterior cerebral arteries (L-A1 and R-A1, respectively), left and right middle cerebral arteries (L-M1 and R-M1, respectively), left and right internal carotid arteries (L-ICA and R-ICA, respectively), left and right posterior communicating arteries (L-PCOMM and R-PCOMM, respectively), left and right posterior cerebral arteries (L-P1, R-P1, L-P2, and R-P2, respectively), and basilar arteries (BA). In Figure 2, ACOMM and L-PCOMM are not shown since the centerlines of these vessel segments were not shown after the vessel segmentation followed by the skeletonization. The resulting segments of the centerlines were colored differently, and they were saved in .html files, which were opened to verify the accuracy of the manual annotations. The points along the centerline underwent smoothing via spline interpolation, as indicated by the red line in Figure 3.
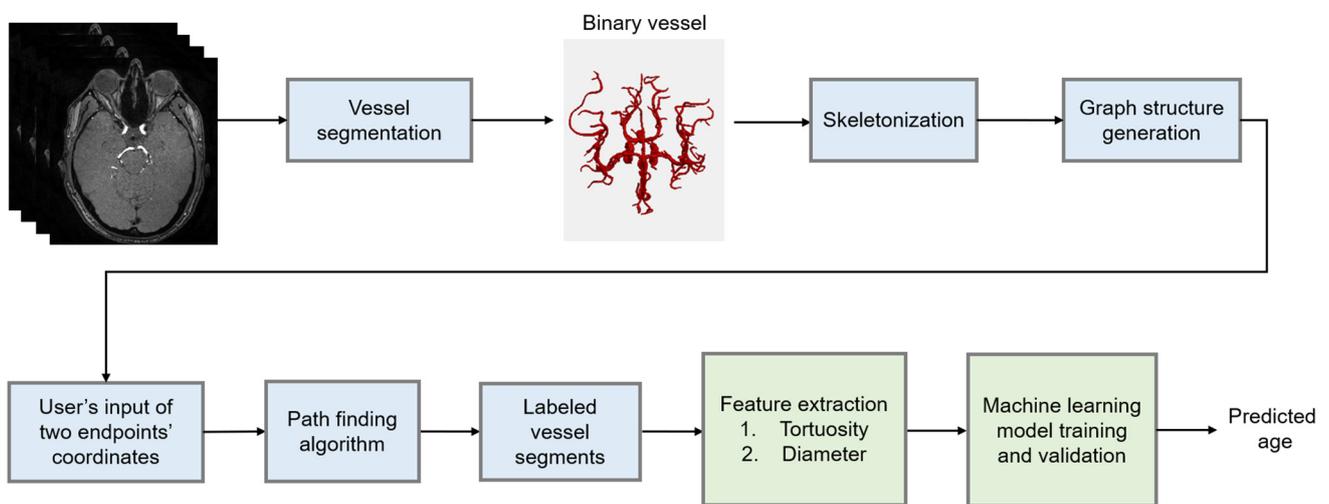

Figure 1. A flowchart of the overall process. The light green boxes are primarily related to the machine learning predictions of age.
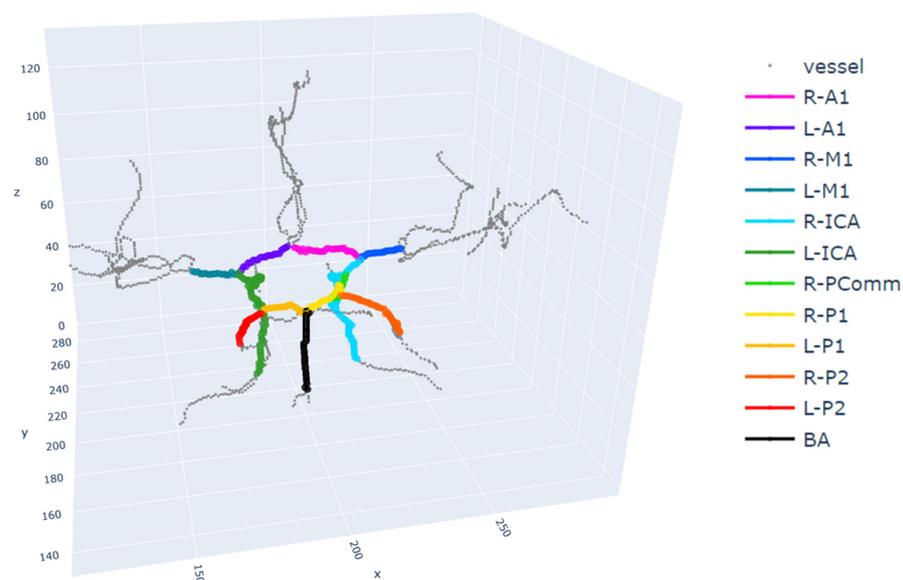

Figure 2. An example of the annotated vessel segments. A1, anterior cerebral artery (ACA) A1; M1, middle cerebral artery (MCA) M1; ICA, internal carotid artery; PComm, posterior communicating artery; P1, posterior cerebral artery (PCA) P1; P2, PCA P2; BA, basilar artery; R, right; L, left.
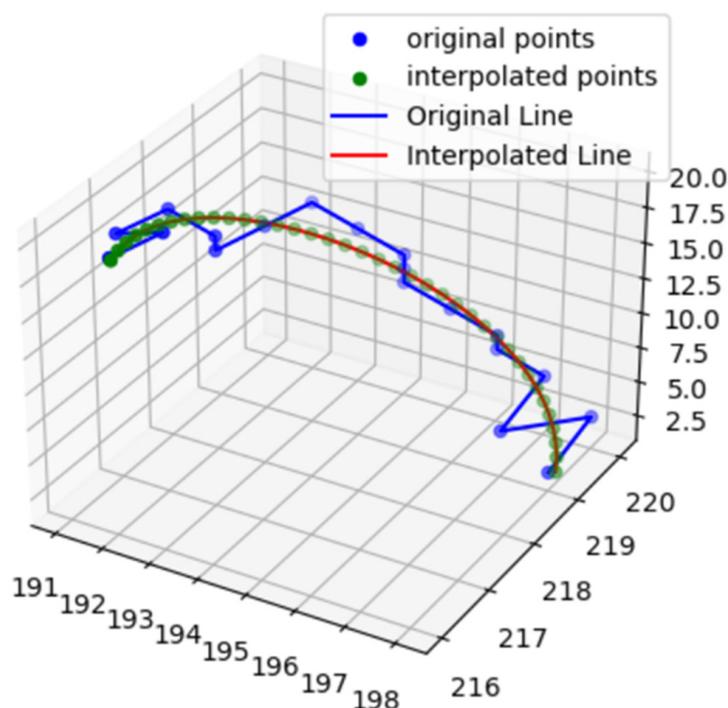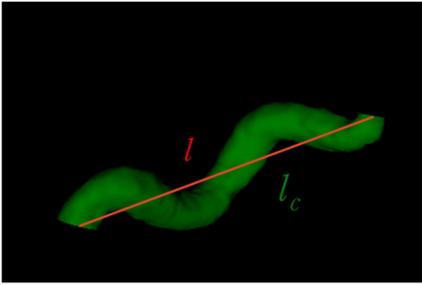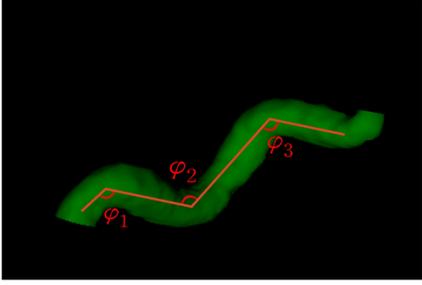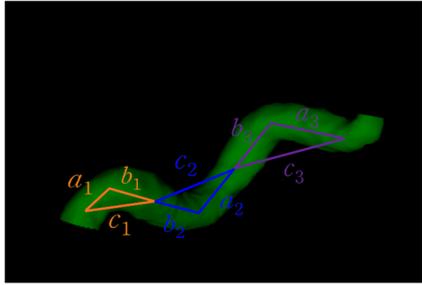
Figure 3. The spline interpolation for smoothing the centerline in a basilar artery. The skeleton is discretized in the 3D coordinate space, and hence, when displayed, it shows a jagged appearance, as indicated by the blue line. A spline interpolation results in the smooth curve which is indicated by the red line.

2.2. Feature Extraction

The extracted features for each segment consisted of the curve length (l_c), Euclidean length (l), relative length (RL), sum of angle metrics (SOAM), product of angle distance (PAD), and triangular index (TI). Table 1 shows the descriptions and mathematical formulations for the tortuosity-related features along with image illustrations. In addition, diameters were calculated along the centerline points using the ndimage.distance_transform_edt function in the SciPy library [26] on a 3D binary vessel mask. The mean, minimum, maximum, standard deviation, 25th percentile, 50th percentile, and 75th percentile values of the diameters were extracted for each segment. These features for the arterial segments in all subjects considered were saved as .xlsx files. Some subjects' data had missing values in certain arterial segments. In our study, the ACOMM, L-PCOMM, and R-PCOMM segments had more missing data than the remaining vessel segments.

To develop the ML models for age estimation, we took features from a subset of the artery segments. The subset consisted of the left and right internal carotid artery (ICA) tortuosity descriptors and basilar artery (BA) tortuosity descriptors, as well as the left and right ICA and BA diameter features. We determined two endpoints in a left or right ICA vessel segment by annotating one endpoint at the bifurcation of the ACA, ICA, and MCA and annotating the other endpoint at the highest curvature of the C3 segment. We also determined two endpoints in a BA vessel segment by annotating one endpoint at the bifurcation of the BA and PCAs and annotating the other endpoint at the bifurcation of the BA and vertebral arteries (VAs). The BA and ICAs were chosen because there were no missing data in their feature values across all subjects, and they had relatively greater vessel lengths and thicknesses than the other vessel segments, facilitating a more reliable quantitative image analysis than the small and thin vessel segments whose vessel diameters could be close to the width of one or two voxels. Hence, the total number of features per subject was $3 \times 13 = 39$, which is equivalent to three vessel segments multiplied by 13 features (i.e., six features related to tortuosity and seven features related to vessel diameter statistics).

Table 1. List of tortuosity features per vessel segment.

Feature Name	Description	Illustration
Curve length (l_c)	The length along the curve	
Euclidean length (l)	The distance between starting and ending points of the curve	
Relative length (RL)	$\mathrm{RL} = \frac{l}{l_c}$	
Sum of angle metrics (SOAM)	$\frac{\sum_{i=1}^{n}(180° - \varphi_i)}{l_c}$ φ_i: the angle between two adjacent lines connecting the center and two endpoints for the ith center	
Product of angle distance (PAD)	$\mathrm{PAD} = \frac{\mathrm{SOAM}}{\mathrm{RL}}$	
Triangular index (TI)	$\mathrm{TI} = \frac{\sum_{i=1}^{n} \frac{a_i + b_i}{c_i}}{n}$	

2.3. Model Development

The ML model development was performed using the Scikit-learn library [27]. We compared the following six ML models for age prediction: (1) random forest regression, (2) linear regression, (3) AdaBoost regression, (4) gradient boosting regression, (5) Bayesian ridge regression, and (6) XGBoost regression.

2.3.1. Random Forest

Random forest is an ensemble-based ML method that uses the predictions made by multiple decision trees. Each decision tree is trained on data samples randomly selected with replacements [28]. Moreover, a subset of features is randomly chosen for a decision tree during training. In regression, a test sample goes through all the trained decision tree models individually, and then the final output is the average of the predictions made by the decision tree models. Random forest is popular due to its simplicity for training.

2.3.2. Linear Regression

In a linear regression model, the output is given as a linear combination of the input features with a bias term added. Linear regression is simple and interpretable because it can describe how the inputs affect the output. Training is performed based on parameter estimation on a linear model with a certain optimization method. The least squares method is the most popular method for the parameter estimation [29]. The linear model parameters are used to predict an outcome with test data. With regards to prediction performance,

linear regression sometimes can be superior to other sophisticated nonlinear ML models in cases with small numbers of training samples or low signal-to-noise ratio samples.

2.3.3. Bayesian Ridge Regression

A Bayesian ridge regression model is a model that utilizes a posterior probability distribution in its weight parameters. Bayesian ridge regression adds the l2-norm regularization terms to the objective function of the linear regression model. The model parameters in the linear model are assumed to follow a Gaussian distribution [30]. Bayesian ridge regression with l2-norm regularization not only prevents overfitting but also is robust against outliers.

2.3.4. AdaBoost

AdaBoost is an acronym for adaptive boosting algorithm [31]. The boosting first generates many weaker ML classifiers during training, and then it creates a final model by constructing a fine-tuned and stronger classifier with all the weaker models. It trains a base classifier from an initial training dataset and then modifies the weights of the initial training data samples based on the classifier's performance. The weights of the misclassified samples become larger, and a new base classifier will be trained with the samples whose weights are changed. Among ML models, AdaBoost is less prone to overfitting and can be used to improve the accuracy of weaker classifiers.

2.3.5. Gradient Boosting

Gradient boosting is an ML approach that first generates a decision tree to approximate a non-linear relationship between the input features and the output, and then it uses boosting as described in the AdaBoost method. It works by iteratively fitting a new model to the residual errors of a previous model. Unlike AdaBoost, gradient boosting updates weights using gradient descent. The method can reduce bias, but overfitting may occur. It theoretically performs better than AdaBoost, but it takes computationally longer and has a memory disadvantage.

2.3.6. XGBoost

XGBoost is an acronym for extreme gradient boosting [32]. It is based on gradient boosting, and importantly, it is designed to support parallel computing for time-efficient predictions from large scale datasets. It learns a tree model during each training iteration to minimize residual errors. XGBoost includes regularization to prevent overfitting and is widely adopted in medical data analysis due to its improved performance in terms of speed, scalability, and accuracy.

2.3.7. Training and Validation

A four-fold cross-validation was performed for our evaluation. The 171 subjects were randomly assigned to one of the four groups for the cross-validation. For each fold, hyperparameters were tuned on the training data using a randomized search [33] after the determination of a candidate hyperparameter set in the scikit-learn library. Table 2 shows the hyperparameter values used to train each of the following five regression models: random forest regression, AdaBoost regression, gradient boosting regression, Bayesian ridge regression, and XGBoost. An evaluation based on the validation data was performed using the metrics described in the next section.

Table 2. Hyperparameter settings of the regression models.

Regression Model	Hyperparameter
Random forest regression	n_estimators = 74, max_depth = 14, min_samples_leaf = 2, min_samples_split = 6, max_features = 'sqrt', criterion = 'absolute_error'
AdaBoost regression	learning_rate = 0.07, loss = 'exponential', and n_estimators = 163
Gradient boosting regression	criterion = 'friedman_mse', learning_rate = 0.11, loss = 'absolute_error', max_depth = 15, n_estimator = 198, and tol = 0.003
Bayesian ridge regression	alpha_1 = 0.0004, alpha_2 = 1×10^{-6}, alpha_init = 10, lambda_1 = 1×10^{-5}, and Tol = 0.01
XGBoost regression	colsample_bytree = 0.95, gamma = 0.22, learning_rate = 0.30, max_depth = 4, n_estimators = 138, and subsample = 0.84

2.4. Evaluation

The performance of the age prediction was quantified using the below metrics. We assumed that x_i was the real chronological age of the *i*th subject and y_i was the predicted age of the *i*th subject. We defined the Pearson correlation coefficient as follows:

$$r = \frac{\sum_{i=1}^{N} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2 \sum_{i=1}^{N}(y_i - \overline{y})^2}}, \tag{1}$$

where $\overline{x}$ is a mean of the x_i values and $\overline{y}$ is a mean of the y_i values. We defined the root mean square error (RMSE) as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{N}}. \tag{2}$$

We defined the mean absolute percentage error (MAPE) as follows:

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left| \frac{x_i - y_i}{x_i} \right|. \tag{3}$$

The coefficient of determination or R^2 was defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{4}$$

Feature importance was calculated for the tree-based models, including the random forest regression, Adaboost regression, gradient boosting regression, and XGBoost regression models with the Scikit-learn library. The calculation was based on the contribution of each feature to the model's prediction performance. The features that led to significant reductions in the Gini impurity or mean square error were considered more important.

Qualitative evaluation was performed by visualizing the scatter plot, which showed correlations between the chronological ages and the ML-predicted ages. Its implementation was made using the Matplotlib Python library [34].

3. Results

After applying the data exclusion criteria provided in the Materials and Methods section, 87 subjects' data were discarded. The remaining 171 subjects were considered for analysis. Table 3 shows the prediction performances of the six ML regression models. The random forest regression model exhibited the lowest average RMSE and the highest

average R^2. The linear regression model exhibited the lowest average MAPE and the highest average Pearson correlation coefficient. The Bayesian ridge regression model showed the poorest performance in all the evaluation metrics considered in this study. The AdaBoost and XGBoost regression models showed similar overall performances. The AdaBoost regression model ranked second for its RMSE, third for its R^2, third for its MAPE, and fourth for its Pearson correlation coefficient. The XGBoost regression model ranked fourth for its RMSE, fourth for its R^2, second for its MAPE, and third for its Pearson correlation coefficient.

Table 3. Evaluation of the models' predictions using the four-fold cross-validation. The bold text indicates the best performances among the models.

Regression Model	Root Mean Squared Error (RMSE)	R^2	Mean Absolute Percentage Error (MAPE)	Pearson Correlation Coefficient
Random forest regression	**14.867 ± 0.515**	**0.186 ± 0.031**	0.319 ± 0.022	0.459 ± 0.058
Linear regression	15.000 ± 1.032	0.162 ± 0.140	**0.290 ± 0.015**	**0.532 ± 0.067**
AdaBoost regression	14.965 ± 0.496	0.175 ± 0.029	0.315 ± 0.023	0.438 ± 0.044
Gradient boosting regression	15.554 ± 1.478	0.106 ± 0.136	0.320 ± 0.034	0.357 ± 0.163
Bayesian ridge regression	16.009 ± 0.419	0.055 ± 0.041	0.347 ± 0.018	0.258 ± 0.089
XGBoost regression	15.207 ± 0.905	0.148 ± 0.062	0.312 ± 0.031	0.456 ± 0.072

Figure 4 shows the scatter plots for the real ages vs. the ML-predicted ages for the exemplary training and the validation data for three regression models, which were (a) the random forest regression model, (b) the linear regression model, and (c) the XGBoost regression model. The random forest model showed overestimation in the young age range while showing underestimation in the old age range, as shown in the training data plot in Figure 4a. The linear regression model showed a more highly scattered plot than the other two models when the left figure in Figure 4b is compared with the left figures in Figure 4a,c. The XGBoost regression model showed almost perfect alignment along the y = x line, as shown in the top figure of Figure 4c, but the validation result showed a more scattered result than the training result, as shown in Figure 4c, implying overfitting in the model training procedure.

Figure 5 shows the ten most important features estimated from the four tree-based models of the random forest regression, AdaBoost regression, gradient boosting regression, and XGBoost regression models. The R-ICA_diam_mean ranked first in the random forest regression and AdaBoost regression models. It ranked second in the XGBoost regression model and sixth in the gradient boosting model. The L-ICA_diam_mean was also one of the important contributing factors in all four models. It ranked second in the random forest regression model, third in the AdaBoost regression model, first in the gradient boosting regression model, and third in the XGBoost regression model. The R-ICA_diam_mean, L-ICA_diam_mean, R-ICA_lc, and R-ICA_diam_std features were in the top-ten list in all four models. The diameter-related features appeared more in the list than the tortuosity-related features for the XGBoost regression model, as shown in Figure 5d.
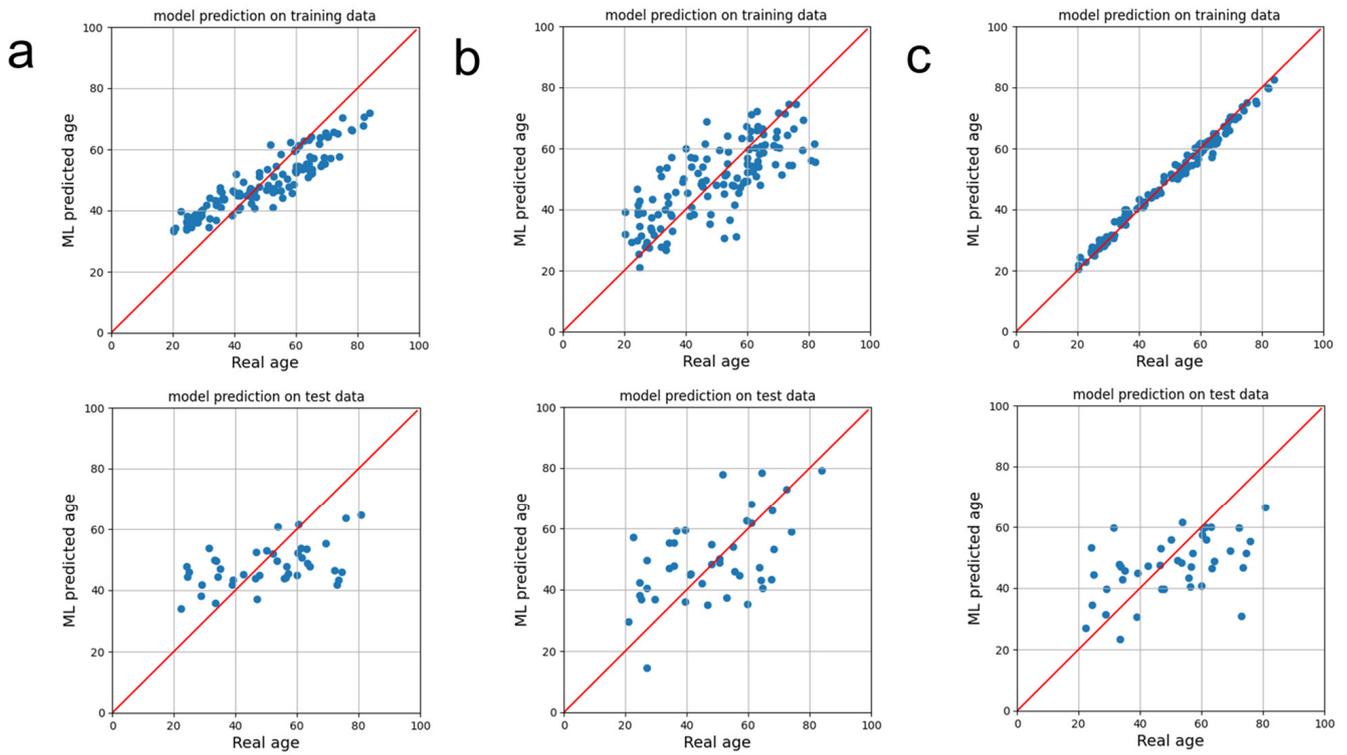
Figure 4. Scatter plots of the age predictions of the (**a**) random forest regression, (**b**) linear regression, and (**c**) XGBoost regression models. The top plots are models' predictions on the training data, and the bottom plots are the models' predictions on the test data. The blue dots are samples, and the red line is the y = x line.
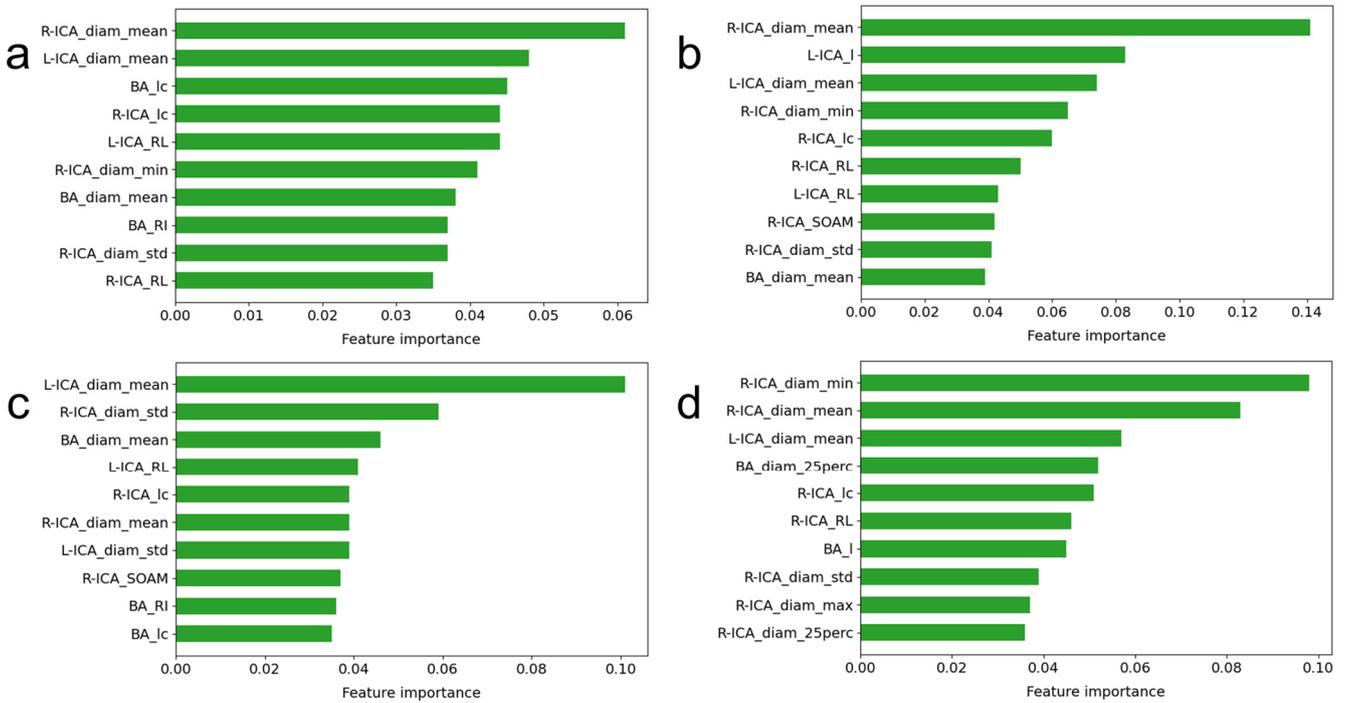

Figure 5. Lists of the ten most important features for the (**a**) random forest regression, (**b**) AdaBoost regression, (**c**) gradient boosting regression, and (**d**) XGBoost regression models.

4. Discussion

We provided the first demonstration of ML-based age prediction and a feature importance assessment based on angiography image data with tortuosity and thickness features extracted from individual vessel segments in the ICA and BA. A variety of ML regression models were trained and validated, and they demonstrated different age prediction scores in terms of RMSE, R^2, MAPE, and Pearson correlation coefficient values. Our findings indicated that the linear regression model showed a prediction performance similar to or higher than the other complex ML models in certain evaluation metrics, although the other complex ML models are known to be well-suited for data showing a nonlinear relationship between features and outcomes. In addition, the ML prediction performance was not as high as expected for the tortuosity and thickness features. This implied that vessel geometry features may not be the determining factors for predicting chronological ages in healthy volunteers. It is noted from the literature that confounding factors such as aerobic activity level can make it difficult to predict chronological age by relying on a vessel's morphometric features alone for ML prediction [35].

We considered the ICAs and BA segments for age prediction. The reason why we chose these segments was that these vessels are usually longer and thicker than other vessel segments such that it could be easier to perform a reliable quantification of vessel tortuosity and diameter measurements. There may be room for improvement in age prediction by considering other vessel segments with longer vessel lengths than those used in our study. For example, Wright et al. analyzed PCAs, ACAs, and MCAs, including their distal branches, in order to associate their morphometric measurements with age. They demonstrated a Pearson correlation coefficient of 0.30 between age and bifurcating branch length, a Pearson correlation coefficient of 0.18 between age and fractal dimension, and a Pearson correlation coefficient of 0.44 between age and tortuosity [19]. Bullitt et al. indicated that the number of unbranched vessels was lower for an older age group. They interpreted their results by stating that as a result of the thickening of vessel wall due to aging, the lumen of the small vessel became too thin to be imaged for the old age group [17]. Chen et al. [23] suggested a detailed analysis by tracing more distal arteries and labeling arteries in more precise categories than the method used in the study of Wright et al. [19].

A 3D CNN model trained using MRA image data was investigated for the purpose of predicting age [20]. Unlike our approach, their method was an end-to-end approach that took a 3D TOF MRA image as an input and the predicted age as the output. In contrast to the deep CNN model development and validation, our approach was more direct in explaining what features contributed the most to age prediction. The deep CNN model automatically finds features from the data itself, but one of the major concerns is its black-box nature, which lacks in the model's interpretability of the prediction results. Our study evaluated the rankings of feature importance in four tree-based ML models and indicated that the mean vessel diameter information in the right ICA segment was the highest-contributing feature for age prediction.

There was a similar study that investigated the performance of age prediction using ML regression models. Simfukwe and Youn applied ML regression models to features of brain region volumes in order to predict brain age [36]. In their study, the Bayesian ridge regression model showed the highest R^2 value of 0.3 compared to other regression models, including the gradient boosting regression, support vector regression, and linear regression models. In our study, the extracted tortuosity and diameter features could be used for other prediction purposes. One potential area of application is risk prediction for intracranial aneurysm, as demonstrated by Nouri et al. [37].

The proposed method involved image segmentation, feature extraction, and ML-based prediction. Automatic segmentation based on deep learning may help improve the object or lesion identification performance [38,39]. Radiomic feature extraction could increase the dimensions of the features for precision medicine [40]. We focused on age prediction, but this framework could be applied to other purposes such as diagnosis [41],

clinical decision-making [42], and outcome prediction [43] in acute stroke neuroimaging and tumor imaging.

SHapley Additive exPlanations (SHAP) [44] was not applied to our study, though it has the potential to interpret a machine learning model by finding a distribution of the contribution of each feature value in the model's outcome. The feature with the widest distribution has the highest impact, and thus, it is regarded as the most important feature. SHAP has numerous medical applications in interpreting a predictive machine learning model. For example, its interpretability has been demonstrated in a prognostic risk stratification model for acute ischemic stroke [45], a feature selection task for the diagnosis of Parkinson's disease [46], and a classification task for the diagnosis of glaucoma [47].

Our study had several limitations. First, the data used in our study did not have clinical information such as history of stroke, smoking, hypertension, dyslipidemia, etc. The predictions were made solely using angiographic image features. Second, the sample size was small, and hence, the data did not encompass substantial biological and anatomical variations among the subjects. Third, the data used in our study had static measurements and did not reveal temporal changes that occurred over an individual's lifetime. Fourth, our image analysis approaches were sophisticated such that it was difficult to obtain large-scale data from multiple institutions. They involved multiple steps for image preprocessing such that any error in any processing step would result in invalid data for the training and validation of a machine learning prediction model.

5. Conclusions

This study showed the feasibility of age prediction using tortuosity and vessel diameter features extracted from individual vessel segments in circle of Willis arteries in magnetic resonance angiography data. Six machine learning regression models were developed and validated for age prediction. The linear regression model showed the highest correlation between real age and predicted age. A drawback of the proposed approach was the need for manual annotations of individual vessel segments as these are time consuming. The proposed method based on an ML regression of arterial segments' vessel features had the advantage of explicitly determining the importance of features, while deep CNN prediction on an MRA image lacks interpretability. The proposed semi-automatic technique for quantifying segmental blood vessels' morphometric features on TOF-MRA images is viable not only for age prediction, but also for other research areas such as the evaluation of treatment effects in atherosclerotic vessels and the assessment of risk of aneurysm in the bifurcations of vessels.

Author Contributions: Conceptualization, Y.-C.K.; methodology, H.-S.Y., J.O. and Y.-C.K.; software, H.-S.Y., J.O. and Y.-C.K.; validation, Y.-C.K.; formal analysis, Y.-C.K.; investigation, Y.-C.K.; resources, Y.-C.K.; data curation, H.-S.Y. and Y.-C.K.; writing—original draft preparation, Y.-C.K.; writing—review and editing, H.-S.Y. and Y.-C.K.; visualization, H.-S.Y. and Y.-C.K.; supervision, Y.-C.K.; project administration, Y.-C.K.; funding acquisition, Y.-C.K. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the use of publicly available data.

Informed Consent Statement: Patient consent was waived due to the use of publicly available data.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guillén, M.F. *2030: How Today's Biggest Trends Will Collide and Reshape the Future of Everything*; St. Martin's Press: New York, NY, USA, 2020.
2. Baecker, L.; Garcia-Dias, R.; Vieira, S.; Scarpazza, C.; Mechelli, A. Machine learning for brain age prediction: Introduction to methods and clinical applications. *eBioMedicine* **2021**, *72*, 103600. [CrossRef]
3. MacDonald, M.E.; Pike, G.B. MRI of healthy brain aging: A review. *NMR Biomed.* **2021**, *34*, e4564. [CrossRef]
4. Hepp, T.; Blum, D.; Armanious, K.; Scholkopf, B.; Stern, D.; Yang, B.; Gatidis, S. Uncertainty estimation and explainability in deep learning-based age estimation of the human brain: Results from the German National Cohort MRI study. *Comput. Med. Imaging Graph.* **2021**, *92*, 101967. [CrossRef]
5. Sajedi, H.; Pardakhti, N. Age Prediction Based on Brain MRI Image: A Survey. *J. Med. Syst.* **2019**, *43*, 279. [CrossRef]
6. He, S.; Pereira, D.; David Perez, J.; Gollub, R.L.; Murphy, S.N.; Prabhu, S.; Pienaar, R.; Robertson, R.L.; Ellen Grant, P.; Ou, Y. Multi-channel attention-fusion neural network for brain age estimation: Accuracy, generality, and interpretation with 16,705 healthy MRIs across lifespan. *Med. Image Anal.* **2021**, *72*, 102091. [CrossRef]
7. De Lange, A.G.; Anaturk, M.; Suri, S.; Kaufmann, T.; Cole, J.H.; Griffanti, L.; Zsoldos, E.; Jensen, D.E.A.; Filippini, N.; Singh-Manoux, A.; et al. Multimodal brain-age prediction and cardiovascular risk: The Whitehall II MRI sub-study. *NeuroImage* **2020**, *222*, 117292. [CrossRef]
8. Peng, H.; Gong, W.; Beckmann, C.F.; Vedaldi, A.; Smith, S.M. Accurate brain age prediction with lightweight deep neural networks. *Med. Image Anal.* **2021**, *68*, 101871. [CrossRef]
9. Hussein, M.K. Comparison of various machine learning regression models based on Human age prediction. *Wasit J. Eng. Sci.* **2022**, *10*, 1–16. [CrossRef]
10. Bermudez, C.; Plassard, A.J.; Chaganti, S.; Huo, Y.; Aboud, K.S.; Cutting, L.E.; Resnick, S.M.; Landman, B.A. Anatomical context improves deep learning on the brain age estimation task. *Magn. Reson. Imaging* **2019**, *62*, 70–77. [CrossRef]
11. Cherubini, A.; Caligiuri, M.E.; Peran, P.; Sabatini, U.; Cosentino, C.; Amato, F. Importance of Multimodal MRI in Characterizing Brain Tissue and Its Potential Application for Individual Age Prediction. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 1232–1239. [CrossRef]
12. Hofmann, S.M.; Beyer, F.; Lapuschkin, S.; Goltermann, O.; Loeffler, M.; Muller, K.R.; Villringer, A.; Samek, W.; Witte, A.V. Towards the interpretability of deep learning models for multi-modal neuroimaging: Finding structural changes of the ageing brain. *NeuroImage* **2022**, *261*, 119504. [CrossRef] [PubMed]
13. Jiang, H.; Lu, N.; Chen, K.; Yao, L.; Li, K.; Zhang, J.; Guo, X. Predicting Brain Age of Healthy Adults Based on Structural MRI Parcellation Using Convolutional Neural Networks. *Front. Neurol.* **2019**, *10*, 1346. [CrossRef] [PubMed]
14. Ciurica, S.; Lopez-Sublet, M.; Loeys, B.L.; Radhouani, I.; Natarajan, N.; Vikkula, M.; Maas, A.; Adlam, D.; Persu, A. Arterial Tortuosity. *Hypertension* **2019**, *73*, 951–960. [CrossRef] [PubMed]
15. Hyvarinen, S. Arteriographic findings of claudication patients. *Ann. Clin. Res.* **1984**, *16* (Suppl. 41), 1–45.
16. Del Corso, L.; Moruzzo, D.; Conte, B.; Agelli, M.; Romanelli, A.M.; Pastine, F.; Protti, M.; Pentimone, F.; Baggiani, G. Tortuosity, kinking, and coiling of the carotid artery: Expression of atherosclerosis or aging? *Angiology* **1998**, *49*, 361–371. [CrossRef]
17. Bullitt, E.; Zeng, D.; Mortamet, B.; Ghosh, A.; Aylward, S.R.; Lin, W.; Marks, B.L.; Smith, K. The effects of healthy aging on intracerebral blood vessels visualized by magnetic resonance angiography. *Neurobiol. Aging* **2010**, *31*, 290–300. [CrossRef] [PubMed]
18. Xu, X.; Wang, B.; Ren, C.; Hu, J.; Greenberg, D.A.; Chen, T.; Xie, L.; Jin, K. Age-related Impairment of Vascular Structure and Functions. *Aging Dis.* **2017**, *8*, 590–610. [CrossRef]
19. Wright, S.N.; Kochunov, P.; Mut, F.; Bergamino, M.; Brown, K.M.; Mazziotta, J.C.; Toga, A.W.; Cebral, J.R.; Ascoli, G.A. Digital reconstruction and morphometric analysis of human brain arterial vasculature from magnetic resonance angiography. *NeuroImage* **2013**, *82*, 170–181. [CrossRef]
20. Nam, Y.; Jang, J.; Lee, H.Y.; Choi, Y.; Shin, N.Y.; Ryu, K.H.; Kim, D.H.; Jung, S.L.; Ahn, K.J.; Kim, B.S. Estimating age-related changes in in vivo cerebral magnetic resonance angiography using convolutional neural network. *Neurobiol. Aging* **2020**, *87*, 125–131. [CrossRef]
21. Mouches, P.; Wilms, M.; Rajashekar, D.; Langner, S.; Forkert, N.D. Multimodal biological brain age prediction using magnetic resonance imaging and angiography with the identification of predictive regions. *Hum. Brain Mapp.* **2022**, *43*, 2554–2566. [CrossRef]
22. Diedrich, K.T. *Arterial Tortuosity Measurement System for Examining Correlations with Vascular Disease*; The University of Utah: Salt Lake City, UT, USA, 2011.
23. Chen, L.; Mossa-Basha, M.; Balu, N.; Canton, G.; Sun, J.; Pimentel, K.; Hatsukami, T.S.; Hwang, J.N.; Yuan, C. Development of a quantitative intracranial vascular features extraction tool on 3D MRA using semiautomated open-curve active contour vessel tracing. *Magn. Reson. Med.* **2018**, *79*, 3229–3238. [CrossRef] [PubMed]
24. Van der Walt, S.; Schönberger, J.L.; Nunez-Iglesias, J.; Boulogne, F.; Warner, J.D.; Yager, N.; Gouillart, E.; Yu, T. Scikit-image: Image processing in Python. *PeerJ* **2014**, *2*, e453. [CrossRef] [PubMed]
25. Kim, S.O.; Kim, Y.C. Effects of Path-Finding Algorithms on the Labeling of the Centerlines of Circle of Willis Arteries. *Tomography* **2023**, *9*, 1423–1433. [CrossRef] [PubMed]

26. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272; Correction to *Nat. Methods* **2020**, *17*, 352. [CrossRef] [PubMed]

27. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

28. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

29. Hastie, T.; Tibshirani, R.; Friedman, J.H.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2009; Volume 2.

30. Bishop, C.M.; Nasrabadi, N.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4.

31. Freund, Y. An adaptive version of the boost by majority algorithm. *Mach. Learn.* **2001**, *43*, 293–318. [CrossRef]

32. Chen, T.Q.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the KDD'16: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]

33. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.

34. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]

35. Bullitt, E.; Rahman, F.N.; Smith, J.K.; Kim, E.; Zeng, D.; Katz, L.M.; Marks, B.L. The effect of exercise on the cerebral vasculature of healthy aged subjects as visualized by MR angiography. *Am. J. Neuroradiol.* **2009**, *30*, 1857–1863. [CrossRef]

36. Simfukwe, C.; Youn, Y.C. Prediction of East Asian Brain Age using Machine Learning Algorithms Trained with Community-based Healthy Brain MRI. *Dement. Neurocogn. Disord.* **2022**, *21*, 138–146. [CrossRef] [PubMed]

37. Nouri, A.; Autrusseau, F.; Bourcier, R.; Gaignard, A.; L'Allinec, V.; Menguy, C.; Veziers, J.; Desal, H.; Loirand, G.; Redon, R. Characterization of 3D bifurcations in micro-scan and MRA-TOF images of cerebral vasculature for prediction of intra-cranial aneurysms. *Comput. Med. Imaging Graph.* **2020**, *84*, 101751. [CrossRef] [PubMed]

38. Kim, Y.C.; Lee, J.E.; Yu, I.; Song, H.N.; Baek, I.Y.; Seong, J.K.; Jeong, H.G.; Kim, B.J.; Nam, H.S.; Chung, J.W.; et al. Evaluation of Diffusion Lesion Volume Measurements in Acute Ischemic Stroke Using Encoder-Decoder Convolutional Network. *Stroke* **2019**, *50*, 1444–1451. [CrossRef] [PubMed]

39. Heydarheydari, S.; Birgani, M.J.T.; Rezaeijo, S.M. Auto-segmentation of head and neck tumors in positron emission tomography images using non-local means and morphological frameworks. *Pol. J. Radiol.* **2023**, *88*, e365–e370. [CrossRef] [PubMed]

40. Ardakani, A.A.; Bureau, N.J.; Ciaccio, E.J.; Acharya, U.R. Interpretation of radiomics features—A pictorial review. *Comput. Methods Programs Biomed.* **2022**, *215*, 106609. [CrossRef] [PubMed]

41. Choi, Y.S.; Bae, S.; Chang, J.H.; Kang, S.G.; Kim, S.H.; Kim, J.; Rim, T.H.; Choi, S.H.; Jain, R.; Lee, S.K. Fully automated hybrid approach to predict the IDH mutation status of gliomas via deep learning and radiomics. *Neuro Oncol.* **2021**, *23*, 304–313. [CrossRef]

42. Kim, Y.C.; Kim, H.J.; Chung, J.W.; Kim, I.G.; Seong, M.J.; Kim, K.H.; Jeon, P.; Nam, H.S.; Seo, W.K.; Kim, G.M.; et al. Novel Estimation of Penumbra Zone Based on Infarct Growth Using Machine Learning Techniques in Acute Ischemic Stroke. *J. Clin. Med.* **2020**, *9*, 1977. [CrossRef]

43. Salmanpour, M.R.; Hosseinzadeh, M.; Rezaeijo, S.M.; Rahmim, A. Fusion-based tensor radiomics using reproducible features: Application to survival prediction in head and neck cancer. *Comput. Methods Programs Biomed.* **2023**, *240*, 107714. [CrossRef]

44. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.

45. Wang, K.; Hong, T.; Liu, W.; Xu, C.; Yin, C.; Liu, H.; Wei, X.; Wu, S.N.; Li, W.; Rong, L. Development and validation of a machine learning-based prognostic risk stratification model for acute ischemic stroke. *Sci. Rep.* **2023**, *13*, 13782. [CrossRef]

46. Liu, Y.C.; Liu, Z.H.; Luo, X.; Zhao, H.J.T. Diagnosis of Parkinson's disease based on SHAP value feature selection. *Biocybern. Biomed. Eng.* **2022**, *42*, 856–869. [CrossRef]

47. Oh, S.; Park, Y.; Cho, K.J.; Kim, S.J. Explainable Machine Learning Model for Glaucoma Diagnosis and Its Interpretation. *Diagnostics* **2021**, *11*, 510. [CrossRef] [PubMed]