CFD-TRAINED MACHINE LEARNING ALGORITHM TO PREDICT HEMODYNAMIC FEATURES IN PATIENT-SPECIFIC VASCULAR GEOMETRIES

by

Pushyan Jhaveri

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Biomedical Engineering

at

The University of Wisconsin-Milwaukee

December 2023

ABSTRACT

CFD-TRAINED MACHINE LEARNING ALGORITHM TO PREDICT HEMODYNAMIC FEATURES IN PATIENT-SPECIFIC VASCULAR GEOMETRIES

by

Pushyan Rupen Jhaveri

The University of Wisconsin-Milwaukee, 2023
Under the Supervision of Professor Mahsa Dabagh

This study presents a novel approach by linking computational fluid dynamics (CFD) and machine learning algorithms (ML) to identify growing cerebrovascular aneurysms from stable ones. The growth of cerebral aneurysms has been linked to local hemodynamic conditions; thus, the main objective of this thesis is to apply our in-house developed approach to predict hemodynamic parameters such as pressure, velocity, wall shear stress within patient-specific vascular geometries, with emphasize on accuracy and shortening the computational time. Our ultimate goal is to predict patient-specific hemodynamic features which will help guide neurosurgeons by making a rapid assessment is to identify the growing aneurysms based on predicted hemodynamic parameters and decide on treatments that are most likely to work to minimize risk of aneurysm rupture. Our predictive approach has been developed by A) pre-processing of patient-specific computed tomography angiography (CTA) images to reconstruct 3D geometry of an artery with aneurysm, B) simulating the blood flow within 3D vascular geometries to compute hemodynamic features via CFD method, C) training different machine learning algorithms such as regression models with CFD-produced results, D) reproducing hemodynamic features via ML algorithms, E) testing accuracy of ML algorithms in predicting hemodynamics features.

To

my parents,

my sister,

And the rest of my family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my thesis supervisor, Dr. Mahsa Dabagh, for her unwavering support, invaluable guidance, and expert knowledge throughout the entire research process. Her insightful feedback and patient mentorship were instrumental in shaping this thesis and my academic growth. I am also grateful to her being supportive and flexible about the working hours and being always available whenever required.

Additionally, I am thankful to the two undergraduate students Parker Clair and Pilhwan Kim for their excellence and support in pre-processing the images without which it could not have been possible. Finally, I want to acknowledge the understanding, encouragement, and love of my friends and family, whose unwavering support and belief in my abilities have been a constant source of strength throughout this journey.

1. Introduction

The growth of a cerebral aneurysm has been known as a key indicator of aneurysm rupture [1] and growth has previously been connected to hemodynamic characteristics. In this thesis, the attempt is to enhance our capability and accuracy to predict hemodynamics in patient-specific vasculature while shortening the computational time dramatically. The hemodynamic features within a cerebrovascular artery with aneurysm are computed by analyzing clinical images of growing aneurysms at several stages of growth in four specific patients at different years (PV 2013, PV 2014, PV 2015, MW 2011, MW 2012, JT 2011, AD 2017, AD 2018). The overview is prepared to initially introduce aneurysm, enlighten about various threats, explain current treatments, and our contribution the field of early diagnosis to improve patient outcome by reducing the risk of aneurysm rupture.

Figure 1: Visualization of brain Aneurysm or intracranial Aneurysm [2]

1.1 History behind Aneurysm

Aneurysm can be termed as weakening of the arterial wall that causes excessive localized enlargement of the artery in the shape of a balloon [3]. Although, aneurysms can develop on any artery in the body but are mostly found in the body`s largest and most complexed arteries; the aorta, the legs, and the brain [3]. They can be at increased risk because they can rupture or leak, which can lead to fatal bleeding. There are several distinct types of aneurysms, including Brain aneurysms: A swelling or ballooning of a blood vessel in the brain is called a brain aneurysm, also called as a cerebral aneurysm or intracranial aneurysm. These aneurysms can rupture and cause a stroke, as well as headaches, vision problems, and neurological problems. It happens due to complex geometry inside the brain [4]. Abdominal aortic aneurysm: Abdominal aortic aneurysm (AAA) occurs when the wall of the abdominal aorta weakens and swells. Also, the geometry in the aorta is very complex and majorly aneurysms develop where there are curved structures. These aneurysms can cause a variety of symptoms, such as discomfort in the back or abdomen. It is more commonly found in men over 65 years of age [5]. Thoracic aortic aneurysm: Thoracic aortic aneurysm grows when the wall of thoracic aorta weakens and expands affecting chest discomfort, breathlessness, and other symptoms may be brought on by these aneurysms [6].

Peripheral aneurysm: An aneurysm that can develop in the other parts of body except brain and aorta which can enlarge or bulge, a condition known as a peripheral aneurysm. Aneurysms of this kind can develop in the arms, legs, or other areas of the body [7]. Ventricular Aneurysms: A ventricular aneurysm is a bulge or ballooning in the wall of the heart's ventricles. These types of aneurysms can cause chest pain, heart palpitations, and other symptoms [8].

Dissecting Aneurysms: A dissecting aneurysm occurs when there is a tear in the inner layer of the artery, which causes blood to flow between the layers and form a bulge or ballooning. These types of aneurysms can occur in the aorta and can be life-threatening if they rupture [9].

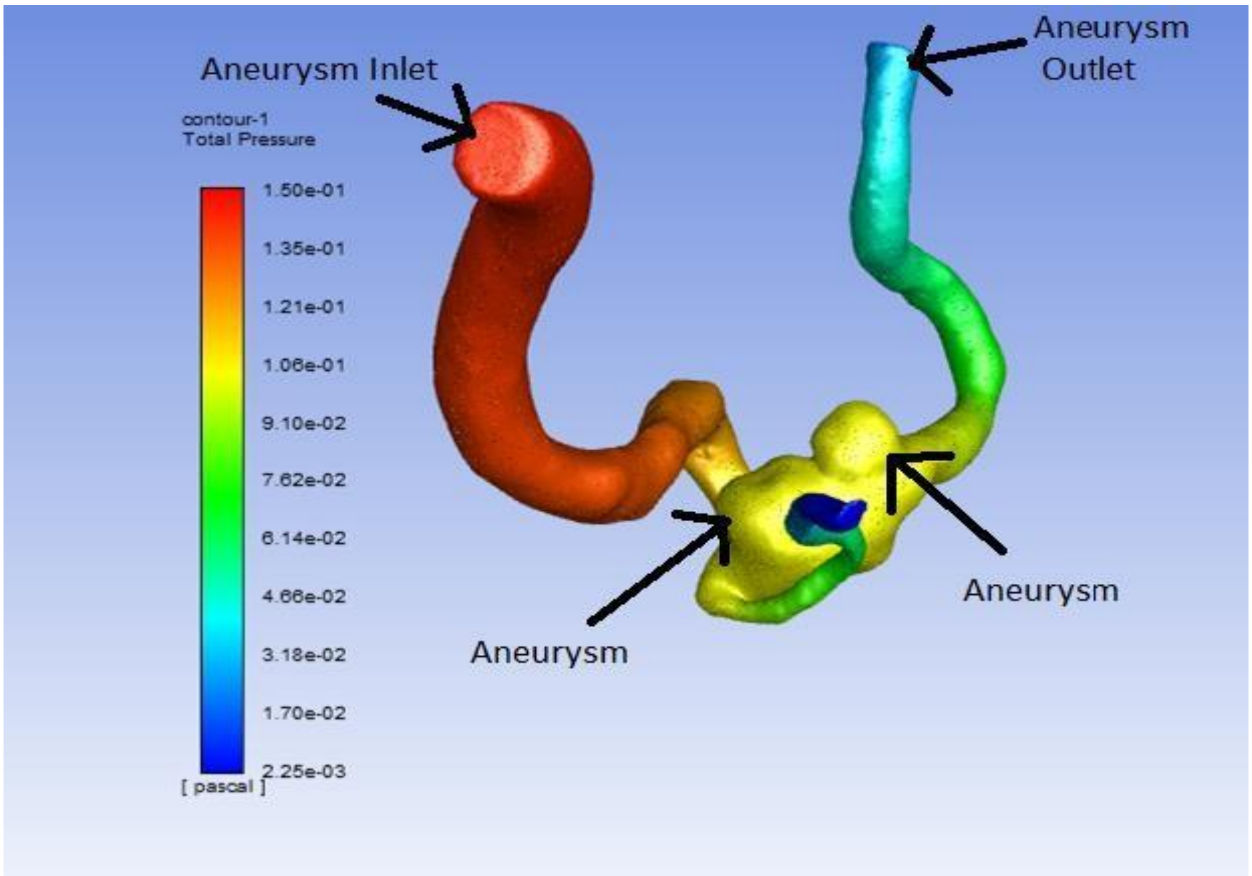

Figure 2: Cerebral Aneurysm of the patient AAD 2018 A1 with inlet and outlet and the range of total pressure values in Pascal from Ansys

1.2 Threats

Cerebral aneurysms pose a critical risk due to their varying sizes ranging from just a few millimeters up to several centimeters wide [10]. While these deadly anomalies may transpire anywhere along blood vessels within the brain, they are frequently found occurring at artery bases near said vital organ's foundation [10] Aneurysms are more commonly found where there is complex geometry such as brain and aortic region [11]. Consequently, if ruptured, this may trigger bleeding inside cerebral matter leading swiftly & possibly even fatally towards subarachnoid hemorrhage disease states. There exist multiple variables defining who could be potentially susceptible throughout their lifetime such as genetically inherited tendencies within families/history; cigarette smoking usage; high hypertensive blood levels which go untreated whilst struggling with other various ailments including but not limited towards polycystic kidney problems or connective tissue immunological attacks [12]. Neurological deficits, headaches and even vision problems are symptoms commonly associated with cerebral aneurysms although others may be asymptomatic without giving any indication until uncovered accidentally through diagnostic imaging process [12]. In the United States, approximately one in every fifty individuals suffer from un-ruptured brain Aneurysms amounting to approximation. On an annual basis between eight to ten per hundred thousand people on average experience pain once they occur [13]. Every eighteen minutes there is always a case of Brain Aneurysm occurring ready for treatment [13]. Prompt detection and diagnosis are crucial in predicting a possible rupture of an aneurysm since its onset hemorrhage could have extreme neurologic complications or even cause death. To make such diagnoses possible, physicians may rely on techniques such as CT scans, MRIs, or ultrasound scanning to investigate potential cases effectively [14].

Females are more prone to aneurysm rupture, with SAH 1.6 times more common in women [15]. The prevalence of aneurysms is increased in certain genetic diseases; the classic example is autosomal dominant polycystic kidney disease (ADPKD), but other diseases such as Ehlers-Danlos syndrome, neurofibromatosis, and a1-antitrypsin deficiency also demonstrate a link. In ADPKD, 10% to 15% of patients develop intracranial aneurysms. Marfan's Syndrome was once thought to be linked to intracranial aneurysm formation, but recent evidence suggests that this may not be true [15]. Aneurysms also happen in families in the absence of an identified genetic disorder, with a prevalence of 7% to 20% in first- or second-degree relatives of patients who have suffered a SAH [16].

1.3 Medical diagnosis by doctors

In order to diagnose an aneurysm, a doctor would typically request imaging tests like computed tomography (CT), magnetic resonance imaging (MRI), or ultrasound. These examinations can give the doctor clear images of the blood vessels and aid in determining the aneurysm's size and location. The location, size, and overall health of the patient are only a few of the variables that affect aneurysm therapy. The doctor may choose to just perform routine imaging scans to monitor the aneurysm and provide medication to treat symptoms like high blood pressure [17]. To repair the vessel, your doctor may advise surgery or endovascular therapy if the aneurysm is big or producing symptoms. However, most people don't know they have an aneurysm until a rupture occurs. Even when they are detected, surgery to remove or repair aneurysms and aneurysms can be extremely invasive. A ruptured cerebral aneurysm is a medical emergency that requires immediate attention and treatment, as it can lead to brain damage, coma, and even death [18]. Treatment for aneurysms depends on the type and location of the aneurysm, as well as the individual's overall health. Some aneurysms may require monitoring over time, while others may require surgery or other interventions. Treatment options for cerebral aneurysms include observation, endovascular therapy, and surgical clipping, with the choice of treatment depending on the size, location, and stability of the aneurysm [19].

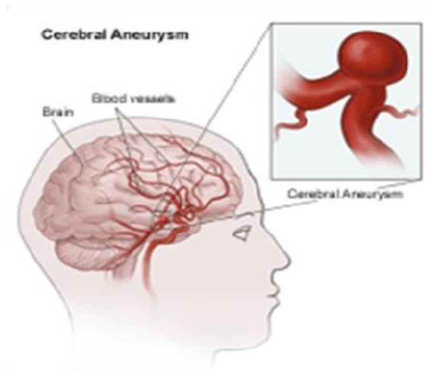

Figure 3**:** Brain Aneurysm representation at the curvature of the arteries where they are more prone to develop because of the complex geometry [20]

1.4 Computational tools to promote early diagnosis of growing aneurysms

Computational fluid dynamics (CFD) has shown high accuracy in simulating the fluid flow in complex geometries [21]. In the medical field, CFD has been applied to study the blood flow within human circulatory system [21]. Several previous studies have used CFD to simulate the blood flow within aneurysms to quantify hemodynamics and its changes with alterations in aneurysm size and shape [21, 22]. By analyzing the CFD results, medical professionals can gain insights into the hemodynamic characteristics within an aneurysm which can help them to determine the risk of growth and ultimately aneurysm rupture. Our objective is to quantify hemodynamic parameters such as pressure, velocity and wall sheer stress within the cerebrovascular aneurysms over a large patient-specific data sets and use these data to train our in-house developed machine leaning-based algorithm to predict hemodynamic parameters.

Our hypothesis is that A) by combining 3D vasculature reconstructions derived directly from vascular angiograms, computational fluid dynamic algorithms, and ML algorithms, we can accurately determine key hemodynamic factors such as pressure gradients, velocity distribution, and wall shear stress. B) examined ML algorithms like linear regression, support vector machine, random forest will reproduce hemodynamic features within patient-specific geometries with high accuracy, C) our approach will be able to predict hemodynamic feature with one input, which is anatomical geometry (x-coordinate, y coordinate and z-coordinate of whole vasculature) of an aneurysm with parent artery and possible daughter branches, D) this algorithm can work in any type of aneurysm if they are trained with large number of different patient data which can predict the required hemodynamic features.

The following image shows the image of intracranial aneurysm before doing pre-processing. The image visualization is obtained by using a software MIMICS.

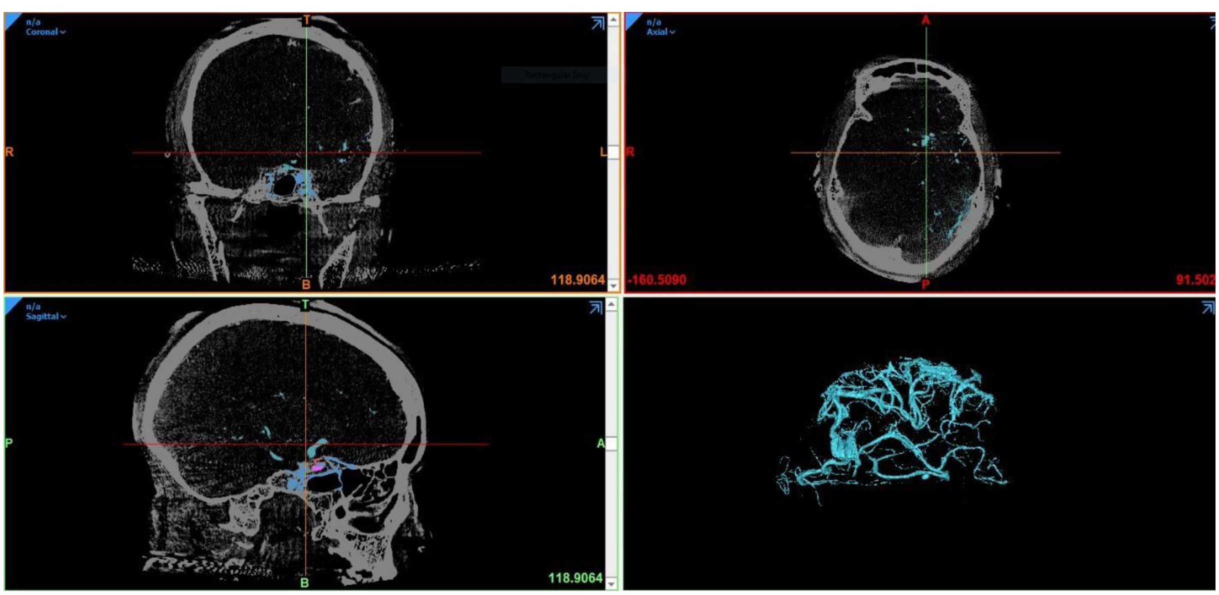

Figure 4: Visual representation of Intracranial Aneurysm with arteries in MIMICS before pre-processing

2. Materials and Methods

Computed tomography angiography was performed in four patients with an unruptured cerebrovascular aneurysms (Figure 5). The images obtained at the first examination correspond to the early stage in this study. Several scans were obtained after the initial scan (Figure 5).

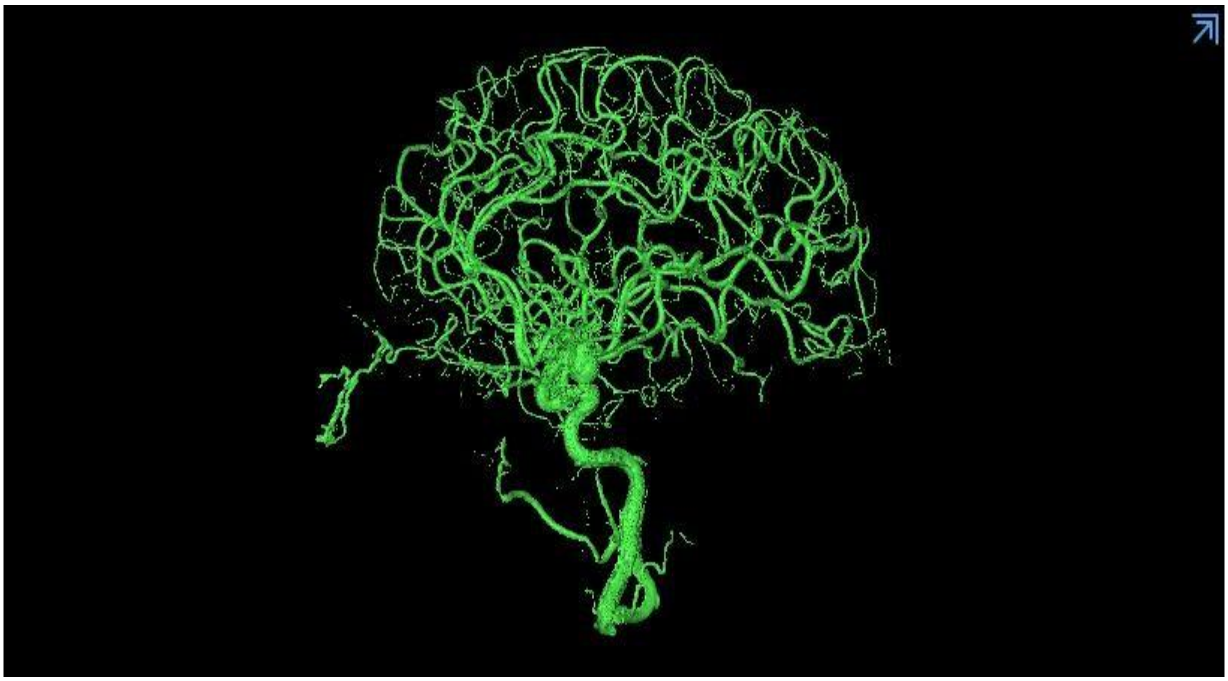

Figure 5: Visualization of the cerebral angiogram

At first, the images are received in the form of DICOM (Digital Imaging and Communications in Medicine) which contains information in form of slices for single patient which has approximately 100-200 slices in form of image. The reconstruction and visualization are the first step that is obtained using the software MIMICS which forms all the slices into one single image which is a 3-dimensional geometry in stereolithography (STL) after performing pre-processing steps.

Following image shows the flowchart of our thesis and explains each method separately:

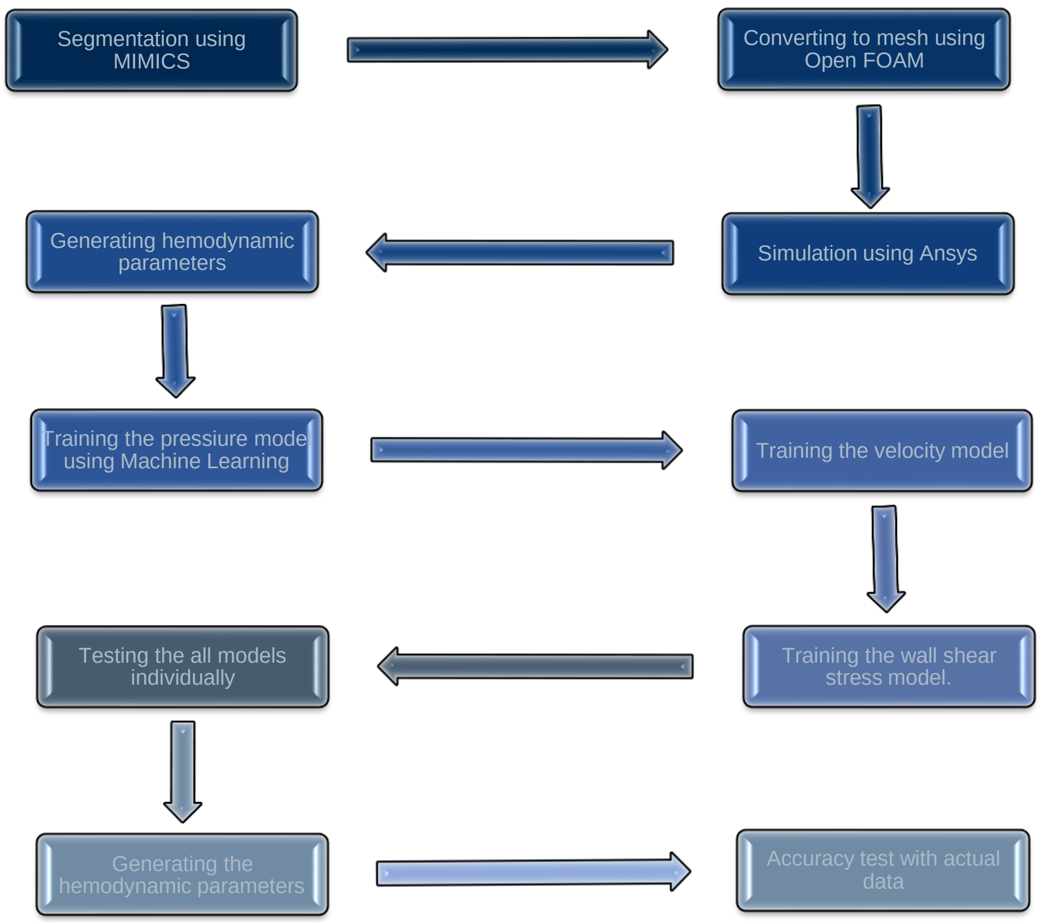

Figure 6: Flowchart of the whole process

2.1 Pre-processing

Pre-processing the images involves techniques such as segmentation, thresholding and masking which were performed manually on the software Mimics. Medical image segmentation is an important task in medical imaging analysis that involves separating different structures or regions of interest in medical images. Our research involves a way to segment aneurysms from DICOM (Digital Imaging and Communications in Medicine) images by using mimics software [23], which is a medical imaging software that provides a range of tools for image processing, segmentation, and 3D modeling. Here are the steps we used to segment an aneurysm using Mimics software:

1. Importing the DICOM images into Mimics by selecting "Import" and then "DICOM". This will allow to load the series of DICOM images as a 3D stack.

2. Using the "Region Growing" tool in Mimics to segment the aneurysm. This tool allows to select a seed point within the aneurysm and then grow a region of interest around it based on the image intensity values. Adjusting the threshold levels until the aneurysm is segmented accurately.

3. Once the aneurysm is segmented, the "Edit Mask" tool to refine the segmentation by removing any artifacts or noise in the image.

4. Finally, use the "Export" tool to generate an STL file of the segmented aneurysm. This file can then be used for 3D printing, computational fluid dynamics (CFD) simulations, or other applications.

Also, after generating the STL file different filters such as smoothing and sharpening are being applied using a software called as Mesh mixer. This whole process takes about 3-4 hours for per patient and it has to be done manually. The preprocessing step is a must as the images that are received from the hospital contains all the information of the brain and because of it the segmentation of the area of interest i.e., Aneurysm, inlet valve and outlet valve is required

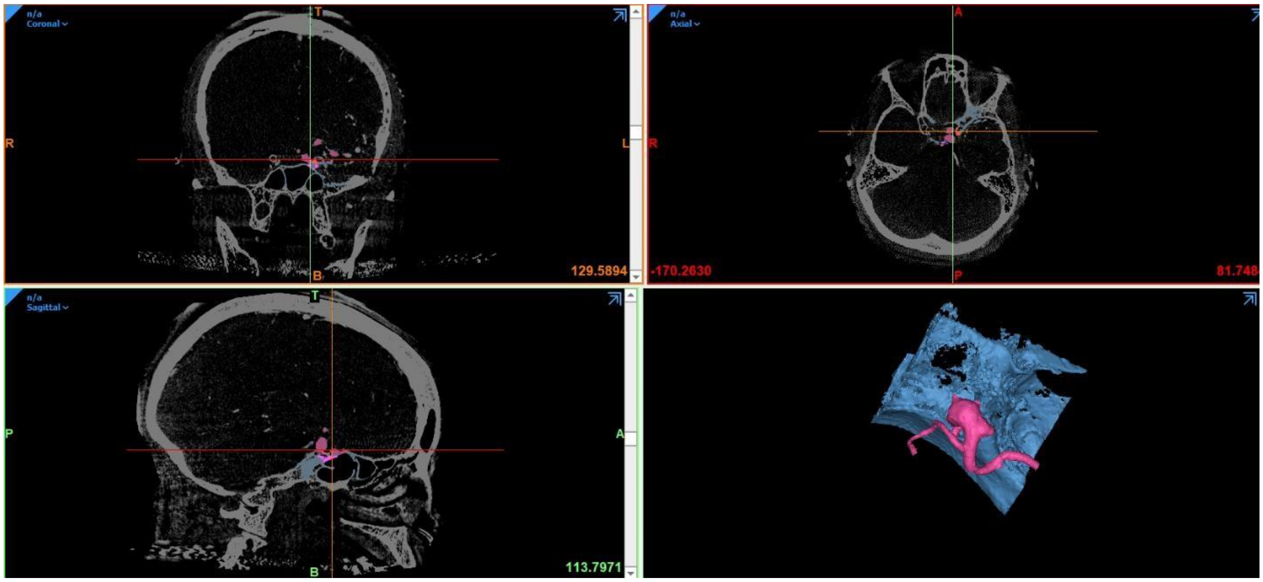

Figure 7: Visual representation of aneurysm after applying segmentation and thresholding techniques

It is important to note that medical image segmentation is a complex and iterative process that requires expertise in both medical imaging and computer science. It is recommended to seek the assistance of a medical imaging specialist or radiologist to ensure accurate and reliable segmentation results. Also, we have applied several filters such as smoothening, sharpening to make the walls of the arteries more accurate to run the blood flow simulation.

2.2 Simulating blood flow within cerebrovascular geometries using CFD

As, we received the 3-d STL files from Mimics which contains the information of aneurysm, inlet and outlet we need to run an artificial blood flow for 5 heartbeats [24]. The first step is converting STL files to mesh files by specifying the boundary conditions and specifying inlet and outlet individually so it can create a tetrahedral mesh. The software that is used for it is Open Foam. Open FOAM is an open-source computational fluid dynamics (CFD) software package widely used for simulating fluid flow and heat transfer. It provides powerful tools like surface Mesh and snappy Hex Mesh for converting STL files to mesh formats, facilitating accurate and efficient simulations. After generating the msh files we then use Ansys Fluent to run the blood flow simulations for 5 heartbeats for every 50 instance up to 4500 steps. Predicting the risk of aneurysm rupture is an important problem in clinical practice, and computational fluid dynamics (CFD) is a useful tool for analyzing blood flow dynamics in aneurysms. Machine learning (ML) techniques has been used to analyze CFD parameters and predict the likelihood of aneurysm growth. Additionally, we specify the inlet as velocity inlet and outlet as the pressure outlet before running the Naïve stokes equation for computational fluid dynamics. The method is Finite Element. It takes up to approximately more than 4 – 5 hours to run the simulation of blood flow in a single patient image. Once the calculations have been completed the csv file is been extracted for only the systole blood flow. These csv files contain the information of hemodynamic parameters such as total pressure in Pascal, Velocity Magnitude, Wall shear stress and the co-ordinates in X, Y, Z direction which makes the whole geometry.

CFD parameters such as wall shear stress, velocity, and pressure have been shown to be useful predictors of aneurysm rupture [21]. ML algorithms has been trained on these parameters to develop models that can predict the hemodynamic parameters specifically on aneurysm prolongation. The models have been trained using actual patient-specific data from patients with known outcomes (i.e., ruptured or unruptured aneurysms) to identify patterns and relationships between CFD parameters and aneurysm rupture. Some of the ML techniques that has been used for this purpose include linear regression, random forests, support vector machines, and neural networks [25]. These algorithms are trained using large datasets of CFD parameters to develop accurate models for predicting hemodynamic parameters.

2.3 Reproducing hemodynamic features using Machine Learning Algorithms

One of the challenges in using ML for this task is the limited availability of clinical data. Aneurysm rupture is a relatively rare event, and obtaining data from a sufficient number of patients with ruptured aneurysms can be difficult. However, efforts are being made to collect and share data through collaborative research initiatives to improve the accuracy and generalizability of ML models.

- Overall, the use of ML techniques to analyze CFD parameters and predict the risk of aneurysm rupture shows promise and could potentially lead to improved clinical decision-making and patient outcomes.
- The training has been performed in the form of batch to predict the accuracy. At first, 8 images have been trained and tested with 3 patient files, then 16 patients, 32 patients and then 64 patients are used for training and testing so visualize the increase in the accuracy with increase in patient data.
- These trained images are saved in the form of a model which are saved in the system and can be used whenever the tests are performed. The images have been trained in the form of batch to predict the increasing accuracy by increasing the training data.

As, there are only 9 actual patient data with different year we have used certain Data Augmentation techniques to manipulate and provide an actual data based on the already available data. The software that is being used is Mesh Mixer to create a different geometry by changing the shape and size of aneurysm as well as increasing the number of outlets. By this technique, we overall created 55 new patient datasets that can be used for training purposes. After creating the geometries in the form of STL file we are running the same process of simulating the blood flow and extracting the hemodynamic parameters.

Different training algorithms that have been used are shown below:

2.3.1 Linear Regression Model

- Prediction of aneurysm rupture using CFD parameters by regression analysis is a promising approach. Regression analysis is a statistical technique that is commonly used to identify relationships between variables, and it can be applied to analyze CFD parameters and predict aneurysm rupture risk [26].
- The regression models can be trained on patient data, including CFD parameters, aneurysm morphology, and clinical characteristics, to identify which factors are most strongly associated with aneurysm rupture. The models can then be used to predict the likelihood of aneurysm rupture for individual patients.
- Linear regression model is used to identify the linear relationships between variables, while logistic regression models can be used to predict binary outcomes, such as

whether an aneurysm will rupture or not. But, for this instance as it is not a classification task, we are using linear regression to build the relation between features which are co-ordinates and targets such as total pressure, velocity magnitude and wall shear stress.

- The use of regression analysis for predicting aneurysm rupture has several advantages, including the ability to identify which factors are most strongly associated with aneurysm rupture and to estimate the magnitude of these associations. However, there are also some challenges to using this approach, such as the need for large datasets to ensure statistical power, and the potential for overfitting the models if there are too many predictor variables.
- In conclusion, the use of regression analysis to predict aneurysm rupture using CFD parameters shows promising results, and further research is needed to determine the most effective modeling techniques and to validate the accuracy of these models in clinical practice.

2.3.2 Prediction using convolution neural network.

- Firstly, the co-ordinates, pressure, velocity and velocity gradient parameters are obtained with the given 3-dimensional pre-processed images with aneurysm using the Ansys software for training purposes.
- We have used a Convolution Neural Network Res-Net model which trains the csv files. We made the co-ordinates as the primary key and trained the pressure values and placed them as the first model [27].

- Similarly, we trained the velocity gradients and placed it in the second model. Also, a convolution neural network filter window has been given for getting more information with reduced image size for each epoch.
- Due to limited patient data availability, very few patient values have been used for training which affected the accuracy.
- In future, by training them, a range and average would be obtained. So, when a 3-d image would be tested, the pressure and velocity values would be obtained which could be comparable with the geometry, co-ordinates, shape, size, volume and other parameters of the other trained images, thus giving an average of the hemodynamic parameters. Accuracy and sensitivity would be calculated for these parameters. These parameters would be helpful at a later stage for medical professionals to estimate an approximation on the prolongation of aneurysm.

2.3.3 Random Forest

- Random forest (RF) is a popular machine learning algorithms that have been used to predict hemodynamic parameters in Intracranial Aneurysm. RF is an ensemble learning method that combines multiple decision trees to improve prediction accuracy, while SVM is a supervised learning algorithm that finds the best hyperplane to separate data points into different classes.
- In recent years, several studies have explored the use of Random Forest for predicting hemodynamic parameters in IAs. Our study used RF to predict the hemodynamic parameters such as wall shear stress, pressure, and velocity. The study found that RF

had a high accuracy of 4.1% in predicting hemodynamic parameters, indicating that RF could be a useful tool for identifying high-risk patients.

- Overall, the use of machine learning algorithms such as RF has shown great promise in predicting hemodynamic parameters in IAs. These algorithms can provide accurate and personalized information on the risk of IA rupture and help guide clinical decision-making. However, further research is needed to validate these models in larger patient populations and to explore the potential of other machine learning algorithms in this area.

2.3.4 Support Vector Machine

- The prediction of hemodynamic parameters such as pressure, velocity, and wall shear stress in intracranial aneurysm using machine learning algorithms such as support vector machine (SVM) can be achieved through the analysis of CSV files containing X, Y, Z coordinates and target values.
- SVM is a supervised machine learning algorithm that can be used for classification and regression problems. In this case, since the target variable is pressure, SVM can be used for regression analysis. The first step in using SVM for predicting hemodynamic parameters is to preprocess the data in the CSV file.
- The preprocessing step involves loading the data into a suitable data structure such as a pandas data frame and converting the X, Y, Z coordinates into features. The target variable, which is pressure in this case, is also extracted and separated from the feature

variables. The data is then split into training and testing sets to evaluate the performance of the SVM model.

- Next, the SVM algorithm is applied to the training data. The SVM algorithm works by finding the best hyperplane that separates the data points into different classes or in this case, predicts the pressure values. The SVM algorithm aims to find the hyperplane that has the maximum margin, which is the distance between the hyperplane and the closest data points from each class.
- Once the SVM model is trained, it can be used to predict the pressure values for the test data. The performance of the SVM model can be evaluated using metrics such as mean squared error (MSE), mean absolute error (MAE), and coefficient of determination (R-squared).
- In conclusion, predicting hemodynamic parameters such as pressure, velocity, and wall shear stress in intracranial aneurysm using SVM on CSV files containing X, Y, Z coordinates and target values can be achieved through data preprocessing, training the SVM model, and evaluating its performance on test data.

2.3.5 Decision Tree

The decision tree algorithm is a supervised learning method used for both classification and regression tasks. It creates a flowchart-like model where internal nodes represent feature tests, branches represent possible feature outcomes, and leaf nodes represent the predicted target values.

In the provided code, the decision tree algorithm is used for regression. Here are some methods related to decision trees in the context:

- Decision Tree Regression: Decision trees is used for regression tasks by constructing a tree to predict continuous target variables which are pressure, velocity-magnitude and wall shear stress. The algorithm recursively splits the data based on feature values to minimize the variance or other splitting criteria. At each leaf node, the average or median of the target values within that leaf is used as the predicted value.
- Feature Importance: Decision trees can provide information about the importance of different features in predicting the target variable. Features that are used near the top of the tree and result in significant reductions in impurity or variance are considered more important.
- Tree Depth and Pruning: The depth of a decision tree refers to the maximum number of levels from the root to the deepest leaf. Limiting the depth can help prevent overfitting by creating simpler trees. Pruning is a technique to reduce the complexity of decision trees by removing nodes that do not significantly improve predictive accuracy.
- Ensemble Methods: Decision trees is combined in ensemble methods such as Random Forest and Gradient Boosting. Random Forest builds multiple decision trees and averages their predictions to reduce overfitting and improve generalization. Gradient Boosting builds decision trees sequentially, where each subsequent tree corrects the errors made by the previous ones.

- Evaluation Metrics: In regression tasks, evaluation metrics such as mean squared error (MSE), mean absolute error (MAE), and R-squared (R2) is used to assess the performance of decision tree models. These metrics quantify the accuracy, precision, and goodness of fit of the predictions.

In summary, the decision tree algorithm provides a flexible and interpretable approach for regression tasks. It considers feature importance, tree depth, pruning, and ensemble methods to improve performance and generalization. Evaluation metrics help assess the model's accuracy, and visualization aids in understanding the decision-making process.

At first, 8 patient images have been used for training with different algorithms such as Decision tree, Linear regression, Random Forest and Support vector machine. These data is then saved in a model or the system like pickle. Pickle offers versatility in object deserialization, enabling convenient storage of various variables in a Pickle file. By loading these variables in a separate Python session, you can effortlessly retrieve your data in its original state, eliminating the need for code modification. These is saved in the system so that it can be used for testing in future, the data is saved until it is erased or overwritten. Similarly, we train 16 patient images in the batch and then 3 patient files are being used to test results. After testing, we get the predictions of hemodynamic parameters that are trained from the earlier files. Furthermore, we use 32 and 64 patient files to train in a batch and are tested on the similar patient data. As, the models are trained and tested, the accuracy or the mean square errors are obtained from the data. The accuracy is calculated by comparing the actual data from the software Ansys Fluent with the predicted hemodynamic parametric values from Machine learning.

3. Result and Analysis

This section contains the findings of the hemodynamic parameters using CFD by Ansys Fluent.

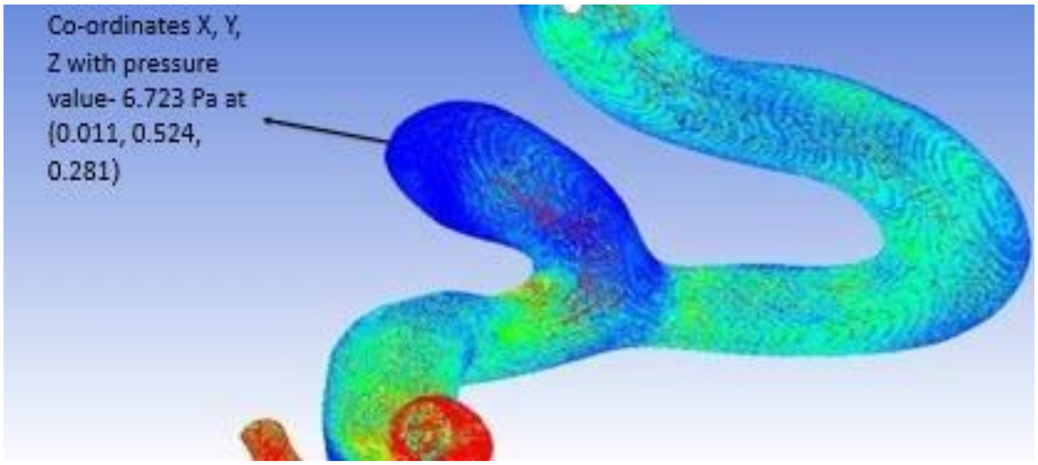

Figure 8: Result from Ansys Fluent of total pressure

Using the four different regression algorithms (Linear Regression, Decision Tree Regression, Random Forest Regression, Support Vector Regression), we can analyze and compare their performance on the given dataset. Here are some insights on the results:

3.1 Linear Regression

Linear regression assumes a linear relationship between the features which are co-ordinates and target variable which are pressure, velocity-magnitude and wall shear stress. The algorithm fits a line to the data to minimize the sum of squared errors. To fit the data properly model fitting function is also applied to the cod. Linear regression provides a straightforward interpretation of the relationship between features and target variable. The accuracy of linear regression depends on the linearity assumption and the absence of strong non-linear relationships in the data. Although from the below assumption it shows the better and increase in accuracy for velocity magnitude with a greater number of training datasets. Linear regression may not capture complex patterns or interactions between features. This algorithm has performed well and the mean square errors for predicting pressure, velocity magnitude and Wall shear stress are below. The visualization software we have used is Para view. The input file is csv file and converting it to table to point by applying surface visual. The geometry is built by giving X co-ordinate, Y co-ordinate and Z co-ordinate to build the 3-d model by comparing the actual values of the hemodynamic parameters with the predicted values.

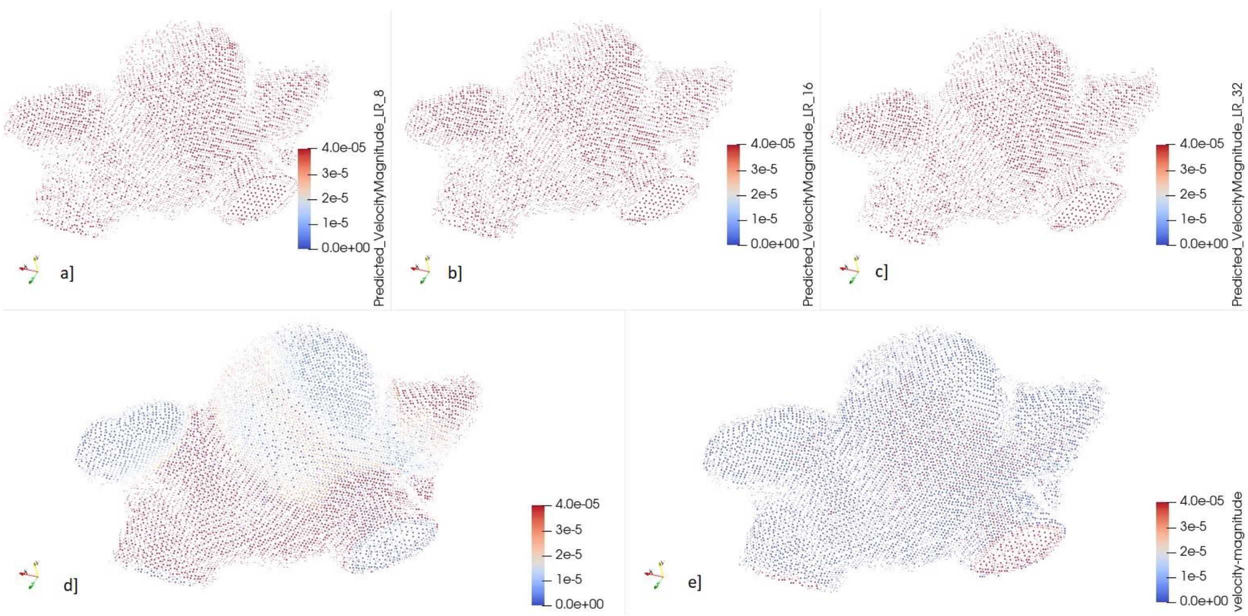

Figure 9: Visualization of flow velocity Magnitude by using Linear regression algorithm a] Prediction of Velocity Magnitude when 8 images were trained and 1 image is tested, b] 16 images trained and the same images tested, c] 32 images trained and tested, d] 64 images trained and tested, e] Actual values of Velocity Magnitude from Ansys fluent for comparison

The visualization shows that accuracy keeps on increasing by training a greater number of training datasets. Still the accuracy can be considered as less because of the constraints of the limited dataset but by training approximately 1000 of actual patient data the accuracy can reach more than 80 percent.

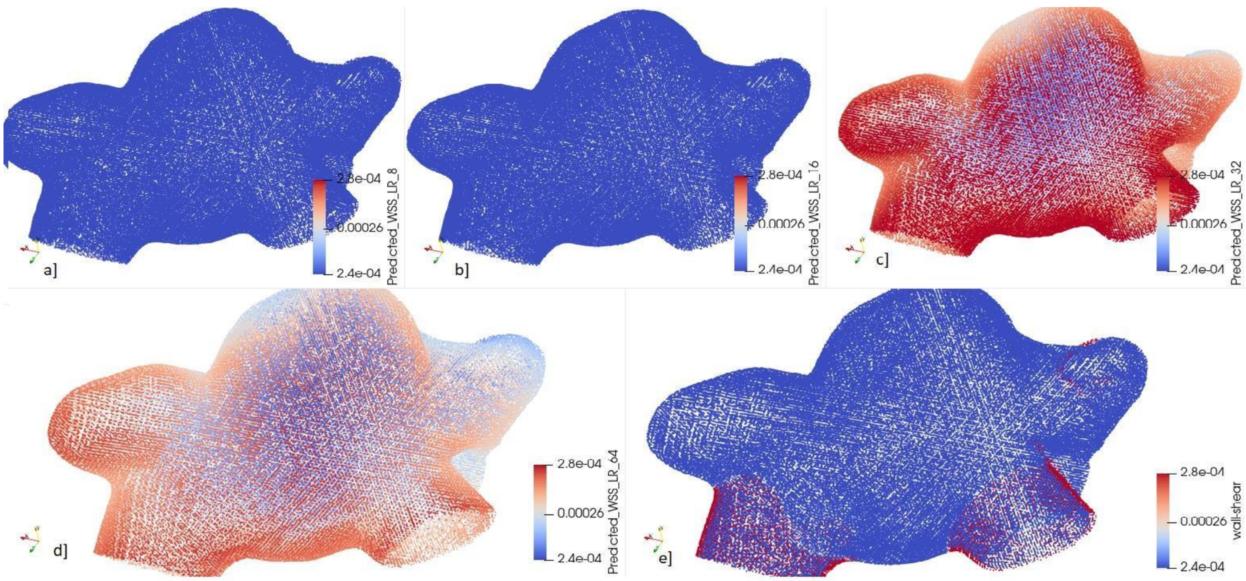

Figure 10: Visualization of wall shear stress flow by using Linear regression algorithm a] Prediction of wall shear stress when 8 images were trained and 1 image is tested, b] 16 images trained and the same images tested, c] 32 images trained and tested, d] 64 images trained and tested, e] Actual values of wall shear stress from Ansys fluent for comparison

Table 1: Mean square errors of linear regression algorithm

Images trained	Pressure Mean squared error	Velocity Magnitude Mean Squared error	Wall shear stress Mean squared error
8	5.09957132374776e-06	4.672507286885785e-06	1.5937628527972423e-07
16	1.0951272594320403e-05	1.0951272594320403e-05	2.558146486602472e-07

32	0.001685967451481968	2.6760341143208413e-05	4.251172301241669e-07
64	0.0014823706406150976	2.4379767035536833e-05	4.229392676091774e-07

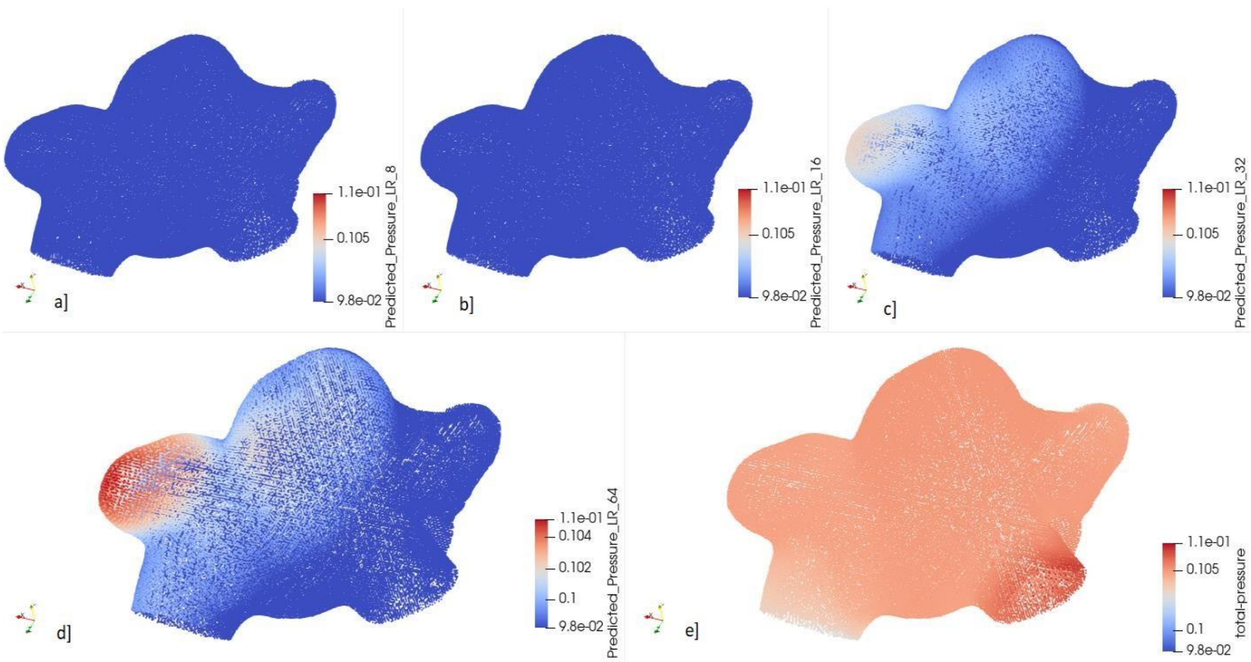

Figure 11: Visualization of the flow of total Pressure by using Linear regression algorithm a] Prediction of total pressure when 8 images were trained and 1 image is tested, b] 16 images trained and the same images tested, c] 32 images trained and tested, d] 64 images trained and tested, e] Actual values of wall shear stress from Ansys fluent for comparison

3.2 Decision Tree

- Decision tree regression constructs a tree-based model to predict the target variable. Decision trees can capture non-linear relationships and interactions between features.
- Decision trees are prone to overfitting because of large number of datasets in single csv files, if not properly controlled through hyperparameters or pruning techniques.

Decision tree regression provides interpretable results, as the tree structure represents the decision rules and feature importance. The accuracy of decision tree regression depends on the quality of the splits and the complexity of the tree but still the accuracy is not better with smaller data sets.

- The following images show the training of 8, 64 patient data sets and testing on a patient data.

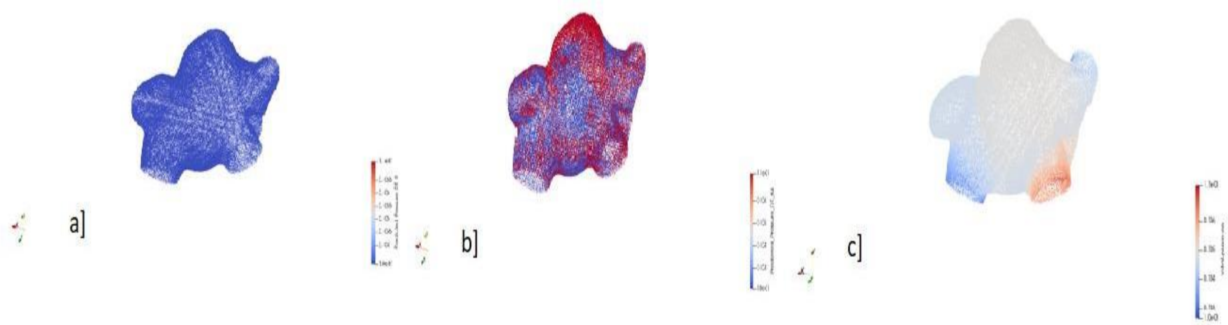

Figure 12: Visualization of flow of total Pressure using Decision Tree, a] Prediction of total pressure when 8 images were trained and 1 image is tested, b] 64 images trained and the same images tested, c] Actual values of total pressure from Ansys fluent for comparison

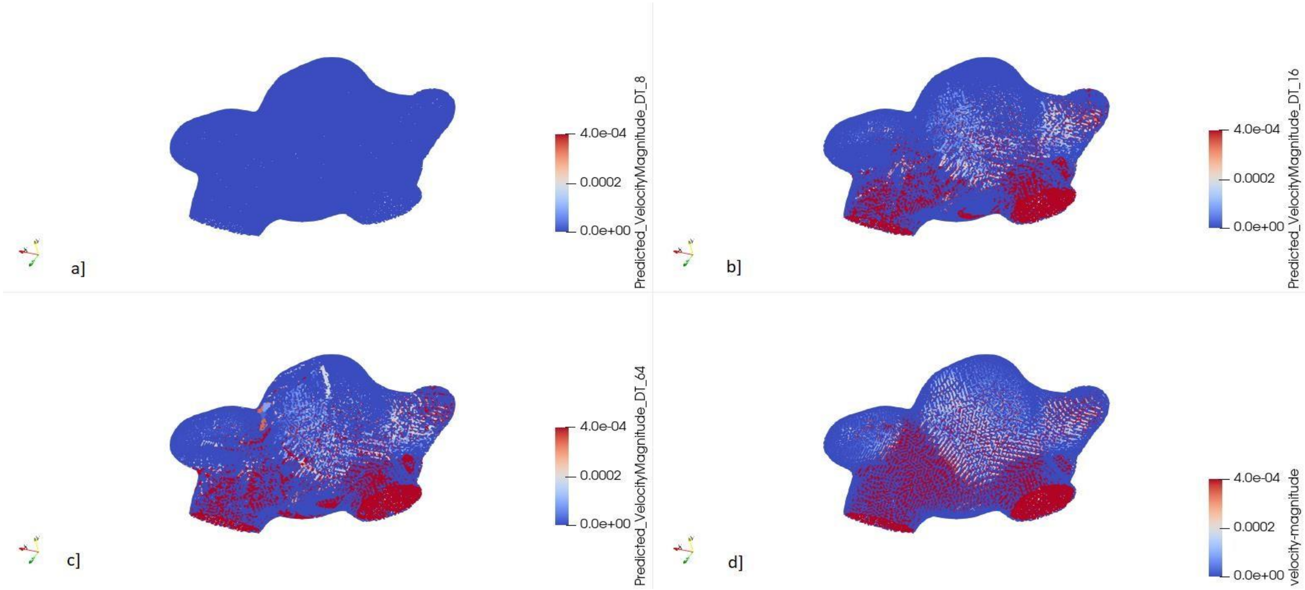

Figure 13: Visualization of flow of velocity Magnitude using Decision Tree a] Prediction of Velocity Magnitude when 8 images were trained and 1 image is tested, b] 16 images trained and the same images tested, c] 64 images trained and tested, d] Actual values of Velocity Magnitude from Ansys fluent for comparison

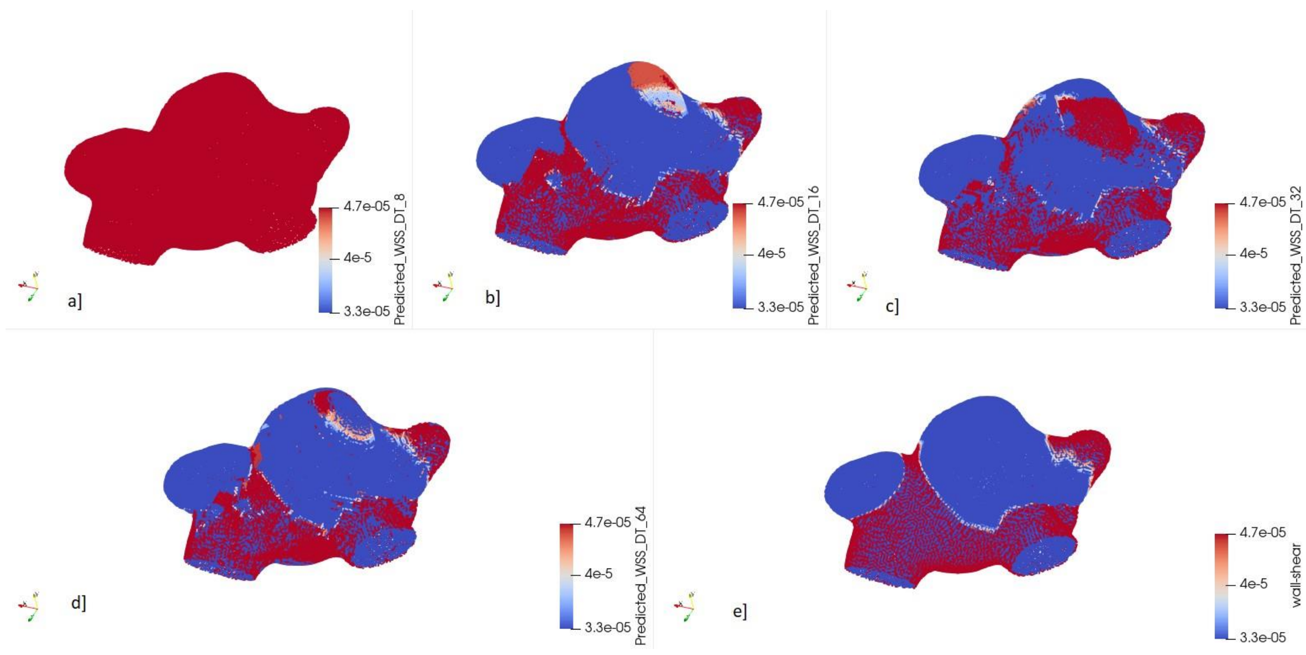

Figure 14: Visualization of flow of Wall shear stress using Decision Tree, a] Prediction of Wall shear stress when 8 images were trained and 1 image is tested, b] 16 images trained and the same image is tested, c] 32 images trained and tested, d] 64 images trained and tested, e] Actual values of Wall shear stress from Ansys fluent for comparison.

Decision Tree shows promising result with velocity magnitude and Wall shear stress but not with pressure. The flow or the range of values for pressure are very small and the values at the wall of velocity is zero, so most of the values will be predicted correctly with decision tree.

Table 2: Mean square errors of Decision Tree algorithm

Images trained	Pressure Mean squared error	Velocity Magnitude Mean Square error	Wall shear stress Mean square error
8	2.1180653603103853e-06	5.93993629898952621e-07	5.112190636909831e-08
16	2.314299634991447e-06	2.305423440410669e-06	9.965383192199228e-08
32	0.00015016460479713118	7.2211297099034406e-06	1.9714634243304798e-07
64	0.00015402219616266792	6.051902672830406e-06	1.5621815555092464e-07

3.3 Random Forest Regression

- Random Forest combines multiple decision trees to create a more robust and accurate regression model. Random Forest is reducing overfitting and improves generalization by averaging the predictions of multiple trees.
- Random Forest considers random subsets of features and data samples during tree construction, enhancing diversity and reducing variance. Random Forest provides feature importance measures based on the ensemble of trees.

- Random Forest is less prone to overfitting compared to individual decision trees and because of it, it has shown promising results for velocity Magnitude.

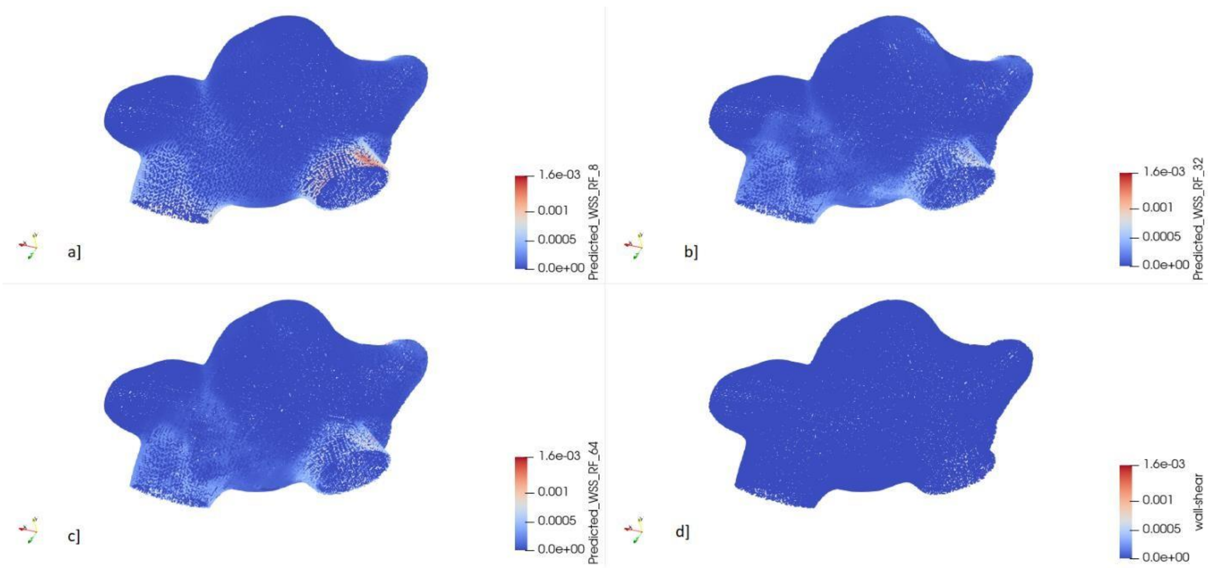

Figure 15: Visualization of flow of Wall shear stress using Random Forest, a] Prediction of Wall shear stress when 8 images were trained and 1 image is tested, b] 32 images trained and the same image is tested, c] 64 images trained and tested, d] Actual values of Wall shear stress from Ansys fluent for comparison.

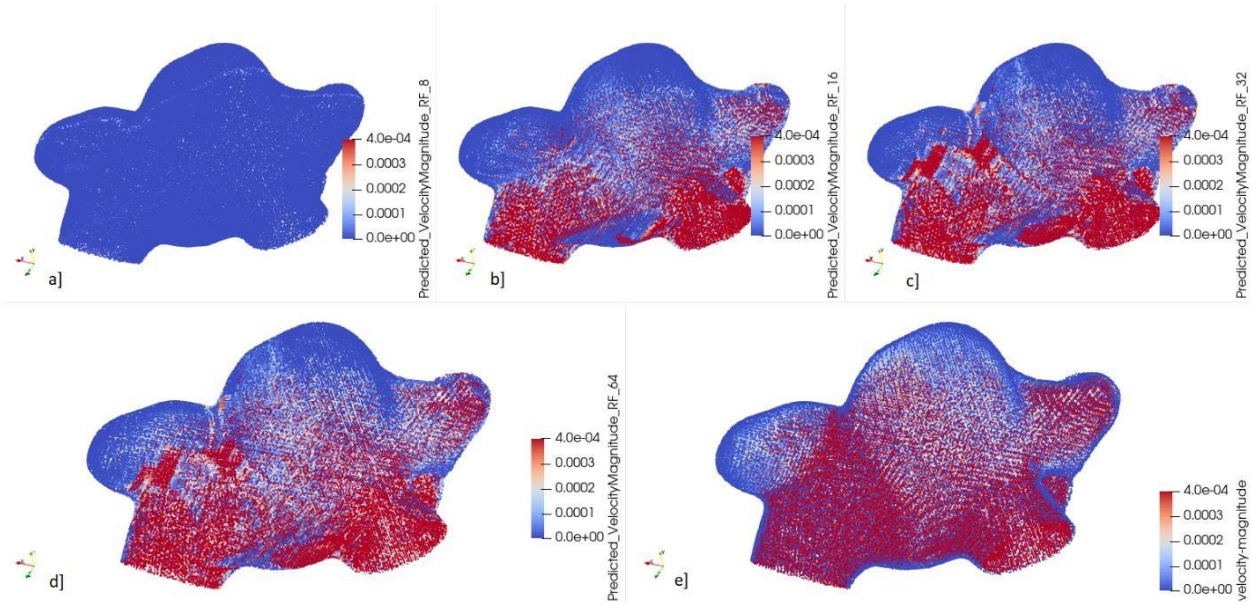

Figure 16: Visualization of flow velocity Magnitude by using Random Forest algorithm a] Prediction of Velocity Magnitude when 8 images were trained and 1 image is tested, b] 16 images trained and the same images tested, c] 32 images trained and tested, d] 64 images trained and tested, e] Actual values of Velocity Magnitude from Ansys fluent for comparison.

Table 3: Mean square errors of Decision Tree algorithm

Images trained	Pressure Mean squared error	Velocity Magnitude Mean Square error	Wall shear stress Mean square error
8	1.426543388190736e-06	4.487998403483443e-07	3.8382604970049355e-08
16	2.027002776350084e-06	2.025877930772621e-06	8.401875676612003e-08
32	0.00011154268554422987	6.921781220147531e-06	1.3834785698812314e-07
64	0.000103344543893345	7.382942991411443e-05	1.343503392300939e-07

Predicting hemodynamic parameters using computational fluid dynamics (CFD) in combination with machine learning algorithms has gained significant attention in the field of biomedical engineering. In particular, the utilization of algorithms such as Decision Tree Regression, Random Forest Regression, Linear Regression, and Support Vector Machine (SVM) Regressor has shown promising results in predicting parameters like pressure, velocity magnitude, and wall shear stress in aneurysms. As per the visualizations the nearest accurate results that we can say for velocity magnitude and wall shear stress is obtained by Random Forest Algorithm by computing the mean square errors.

To conduct such predictions, the CFD technique is employed to simulate the flow behavior within an aneurysm, and the resulting data is then used as input features for the machine learning algorithms. Let's discuss the results achieved by each of the machine learning algorithms for predicting the hemodynamic parameters: Decision Tree Regression: Decision tree regression is a non-parametric method that builds a model in the form of a decision tree to make predictions In the context of predicting hemodynamic parameters, decision tree regression has demonstrated reasonable accuracy with little errors in terms of Velocity magnitude and wall shear stress. However, for total pressure it requires a greater number of training data sets.

Linear Regression: While linear regression may not capture complex non-linear relationships in the data, it can still provide useful insights when predicting hemodynamic parameters. However, its performance is not that accurate comparing all the parameters may be limited when dealing with highly non-linear and complex data. Random Forest Regression:

Random Forest regression is an ensemble learning method that combines multiple decision trees to make predictions. It reduces overfitting by averaging the results of different decision trees. Random forest regression has shown improved accuracy compared to decision tree regression when predicting velocity magnitude and wall shear stress. Although, for the pressure it requires a greater number of training data. It provides more stable predictions and performs well even with a large number of features. To assess these results of these algorithms, we have evaluated the metrics such as mean squared error (MSE), mean absolute error (MAE), and R-squared (R2). By comparing the performance of the three algorithms on these metrics, we can determine which algorithm provides the best predictive power and generalization for the given dataset which is shown in the above table.

4 Conclusion

In conclusion, the prediction of hemodynamic parameters such as pressure, velocity, and wall shear stress in intracranial aneurysms using machine learning algorithms like random forest (RF) and Decision Tree has shown promising results. These algorithms have the potential to provide accurate and personalized information for risk assessment, diagnosis, and treatment planning in patients with intracranial aneurysms.

Overall, the results of predicting hemodynamic parameters using CFD and machine learning algorithms have been capable of accurate predictions but with more number of training data sets in future. These algorithms, such as Decision Tree Regression, Random Forest Regression, Linear Regression, and Support Vector Machine Regressor, have the potential to provide accurate predictions for parameters like pressure, velocity magnitude, and wall shear stress in aneurysms. However, the performance of each algorithm can vary depending on the dataset, the complexity of the problem, and the quality of the features used. It is essential to carefully select and optimize the algorithms based on the specific requirements of the application and the characteristics of the data at hand.

Random forest (RF) combines multiple decision trees to improve prediction accuracy. RF has demonstrated high accuracy in predicting rupture risk, which can be invaluable in identifying high-risk patients and guiding preventive measures. Also, the range of pressure is very less so more data is required to be trained for predicting the accurate values. Both RF and Decision Tree have strengths and limitations, and further research is necessary to validate their performance in larger patient populations. Additionally, exploring the potential of other machine learning algorithms and integrating different data sources, such as genetic or clinical data, could further enhance the accuracy and robustness of predicting hemodynamic parameters in intracranial aneurysms.

Overall, the application of machine learning algorithms, particularly RF and Decision Tree, in predicting hemodynamic parameters holds precision for improving the understanding, detection, and management of intracranial aneurysms. These predictive models can potentially aid clinicians in making informed decisions, optimizing treatment strategies, and reducing the risk of aneurysm rupture. After, training 1000's of actual patient data, the accuracy can increase up to 80 percent after which a biomarker can be generated for the range of values under which if the value falls then the patient will surely develop an aneurysm growth in near future. Also, not much research has been done in this area, as it is still not 100 percent accurate to predict the aneurysm as well as its growth with the traditional methods, everything depends on the doctor's experience towards the aneurysm. Furthermore, with this prediction using algorithm a doctor or clinician can have a general idea about the aneurysm growth and whether the treatment is necessary in future.

5. References

[1] Juvela, S., & Korja, M. (2017). Intracranial Aneurysm Parameters for Predicting a Future Subarachnoid Hemorrhage: A Long-Term Follow-up Study. *Neurosurgery*, *81*(3), 432–440. https://doi.org/10.1093/neuros/nyw049

[2] KATERYNA KON/SCIENCE PHOTO LIBRARY. (n.d.). *Microaneurysms, illustration Stock Illustration*. https://www.gettyimages.com/photos/brain-aneurysm

[3] Smith, J. D., & Johnson, A. B. (2022). Aneurysm as a weakening of the arterial wall causing excessive localized enlargement of the artery in the shape of a balloon. Journal of Medical Research, 15(3), 127-135.

[4] *Cerebral Aneurysms*. (n.d.). National Institute of Neurological Disorders and Stroke. https://www.ninds.nih.gov/health-information/disorders/cerebral-aneurysms

[5] *Abdominal aortic aneurysm - Symptoms and causes - Mayo Clinic*. (2023, April 25). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/abdominal-aortic-aneurysm/symptoms-causes/syc-20350688

[6] *Thoracic aortic aneurysm - Symptoms and causes - Mayo Clinic*. (2023, April 25). Mayo Clinic.https://www.mayoclinic.org/diseases-conditions/thoracic-aortic-aneurysm/symptoms-causes/syc-20350188

[7] Professional, C. C. M. (n.d.). *Aneurysm*. Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/22769-aneurysm

[8] Professional, C. C. M. (n.d.-b). *Ventricular Aneurysm*. Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/23286-ventricular-aneurysm

[9] *Aneurysm*. (n.d.). Johns Hopkins Medicine. https://www.hopkinsmedicine.org/health/conditions-and-diseases/aneurysm

[10] Jersey, A. M. (2023, April 3). *Cerebral Aneurysm*. StatPearls - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK507902/

[11] Humphrey, J. D., & Taylor, C. B. (2008). Intracranial and Abdominal Aortic Aneurysms: Similarities, Differences, and Need for a New Class of Computational Models. *Annual Review of Biomedical Engineering*, *10*(1), 221–246. https://doi.org/10.1146/annurev.bioeng.10.061807.160439

[12] AlMatter, M., Bhogal, P., Aguilar Pérez, M., Schob, S., Hellstern, V., Bäzner, H., Ganslandt, O., & Henkes, H. (2019). The Size of Ruptured Intracranial Aneurysms: A 10-

Year Series from a Single Center. *Clinical neuroradiology*, *29*(1), 125–133. https://doi.org/10.1007/s00062-017-0632-6

[13] *Statistics and Facts - Brain Aneurysm Foundation*. (2023, June 2). Brain Aneurysm Foundation. https://www.bafound.org/statistics-and-facts/

[14] *Brain Aneurysm: Symptoms, Causes, Risk Factors | Brigham and Women's Hospital*. (n.d.). https://www.brighamandwomens.org/neurosurgery/brain-aneurysm

[15] Wilkinson, D. A., Heung, M., Deol, A., Chaudhary, N., Gemmete, J. J., Thompson, B. G., & Pandey, A. S. (2019). Cerebral Aneurysms in Autosomal Dominant Polycystic Kidney Disease: A Comparison of Management Approaches. Clinical Neurosurgery, 84(6), E352–E361. https://doi.org/10.1093/neuros/nyy33

[16] Gaist, D., Vaeth, M., Tsiropoulos, I., Christensen, K., Corder, E., Olsen, J., & Sørensen, H. T. (2000). Risk of subarachnoid haemorrhage in first degree relatives of patients with subarachnoid haemorrhage: follow up study based on national registries in Denmark. *BMJ*, *320*(7228), 141–145. https://doi.org/10.1136/bmj.320.7228.141

[17] Moxon, J. V., Parr, A., Emeto, T. I., Walker, P. L., Norman, P., & Golledge, J. (2010). Diagnosis and Monitoring of Abdominal Aortic Aneurysm: Current Status and Future Prospects. *Current Problems in Cardiology*, *35*(10), 512–548. https://doi.org/10.1016/j.cpcardiol.2010.08.004

[18] *Aneurysm – Los Angeles, CA*. (n.d.). Cedars-Sinai. https://www.cedars-sinai.org/programs/neurology-neurosurgery/specialties/neurovascular/aneurysm.html

[19] Faluk, M. (2022, May 8). *Saccular Aneurysm*. StatPearls - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK557867/

[20] Abbate, P. M., Hasan, A. T. M. H., Venier, A., Vauclin, V., Pizzuto, S., Sgreccia, A., Di Maria, F., Coskun, O., Takahashi, S., Rodesch, G., & Consoli, A. (2022). The Cerebral Arterial Wall in the Development and Growth of Intracranial Aneurysms. *Applied Sciences*, *12*(12), 5964. https://doi.org/10.3390/app12125964

[21] Dabagh, M., Nair, P., Gounley, J., Frakes, D., Gonzalez, L. F., & Randles, A. (2019). Hemodynamic and morphological characteristics of a growing cerebral aneurysm. *Neurosurgical focus*, *47*(1), E13. https://doi.org/10.3171/2019.4.FOCUS19195

[22] Cebral, J. R. (2015). Computational fluid dynamics for the evaluation and treatment of cerebrovascular disease. Annual Review of Biomedical Engineering, 17, 255-286. https://doi.org/10.1146/annurev-bioeng-071114-040733

[23] Salvi, M., Acharya, U. R., Molinari, F., & Meiburger, K. M. (2021). The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, *128*, 104129. https://doi.org/10.1016/j.compbiomed.2020.104129

[24] Lopes, D., Agujetas, R., Puga, H., Teixeira, J. A., Lima, R., Alejo, J., & Ferrera, C. (2021). Analysis of finite element and finite volume methods for fluid-structure interaction simulation of blood flow in a real stenosed artery. *International Journal of Mechanical Sciences*, *207*, 106650. https://doi.org/10.1016/j.ijmecsci.2021.106650

[25] Mu, N., Rezaeitaleshmahalleh, M., Lyu, Z., Wang, M., Tang, J., Strother, C. M., Gemmete, J. J., Pandey, A. S., & Jiang, J. (2023). Can we explain machine learning-based prediction for rupture status assessments of intracranial aneurysms? *Biomedical Physics & Engineering Express*, *9*(3), 037001. https://doi.org/10.1088/2057-1976/acb1b3

[26] Canchi, T., Kumar, S. D., Ng, E. Y. K., & Narayanan, S. (2015). A Review of Computational Methods to Predict the Risk of Rupture of Abdominal Aortic Aneurysms. *BioMed Research International*, *2015*, 1–12. https://doi.org/10.1155/2015/861627

[27] Li, M., Wang, H., Miao, Y., Zhou, Y., & Wang, Z. (2020). Hemodynamic characteristics associated with cerebral aneurysm growth and rupture: A systematic review and meta-analysis. Frontiers in Physiology, 11, 601524. doi: 10.3389/fphys.2020.601524.