УДК 616.136.7:616.61-073.75-047.58
https://doi.org/10.20538/1682-0363-2026-2-21-31

Mathematical Modeling of the Intrarenal Arterial Bed Structure Using Regression Models and a Graph Neural Network

Zenin O.K.[1], Gorbachenko V.I.[1], Gribkov D.N.[1], Sergienko A.A.[1], Ilias Miltiadis[2]

[1] *Penza State University (PSU)*
40 Krasnaya St., 440026 Penza, Russian Federation

[2] *The University of Palermo*
61 Piazza Marina St., 90133 Palermo, Italy

ABSTRACT

The aim was to develop models for numerical modeling of the diameters and lengths of bifurcation segments, structural units of the intrarenal arterial bed (IAB).

Materials and methods. The study is based on previously obtained morphometric data on the structure of 9,642 arterial bifurcations (AB) of 60 corrosion casts of the IAB. AB is a section of the IAB consisting of a proximal arterial segment (AS) (D), two distal AS (with a larger (d_{max}) and smaller (d_{min}) internal diameter) and a branching point. Regression models were used for numerical modeling of the AS diameters, while a graph attention network (GAT) was used for the lengths of AS.

Results. The developed models demonstrate high prediction accuracy: for AS diameters with a larger value (d_{max}), the training set reached an R^2 value of 0.89, for AS with a smaller value (d_{min}) – R^2 = 0.8. For the length of the AS with a larger diameter, the training set reached an R^2 value of 0.75, and the test set reached an R^2 value of 0.74; for the length of the AS with a smaller diameter, the training set reached an R^2 value of 0.77, and the test set reached an R^2 value of 0.73.

Conclusion. The proposed models can be used for numerical modeling of the IAB structure and assessment of the adequacy of renal blood supply.

Keywords: intrarenal arterial bed, arterial bifurcation, numerical modeling, neural networks

Conflict of interest. The authors declare the absence of obvious or potential conflicts of interest related to the publication of this article.

Source of financing. The authors state that they received no funding for the research.

For citation: Zenin O.K., Gorbachenko V.I., Gribkov D.N., Sergienko A.A., Ilias Miltiadis. Mathematical Modeling of the Intrarenal Arterial Bed Structure Using Regression Models and a Graph Neural Network. *Bulletin of Siberian Medicine.* 2026;25(2):21–31. https://doi.org/10.20538/1682-0363-2026-2-21-31.

✉ *Zenin Oleg K.,* zen.olegz@gmail.com

Математическое моделирование структуры внутрипочечного артериального русла с помощью регрессионных моделей и графовой нейронной сети

Зенин О.К.[1], Горбаченко В.И.[1], Грибков Д.Н.[1], Сергиенко А.А.[1], Ilias Miltiadis[2]

[1] *Пензенский государственный университет (ПГУ)*
Россия, 440026, г. Пенза, ул. Красная, 40

[2] *Университет Палермо*
Италия, 90133, г. Палермо, ул. Пиацца Марина, 61

РЕЗЮМЕ

Цель – разработать модели для численного моделирования диаметров и длин сегментов бифуркаций – структурных единиц внутрипочечного артериального русла (ВАР).

Материалы и методы. Исследование основано на ранее полученных морфометрических данных о структуре 9 642 артериальных бифуркаций (АБ) 60 коррозионных препаратов ВАР. Артериальная бифуркация – участок ВАР, состоящий из проксимального артериального сегмента (АС) (D), двух дистальных АС (с большей (d_{max}) и меньшей (d_{min}) величиной внутреннего диаметра) и точки ветвления. Для численного моделирования диаметров (АС) использовали регрессионные модели, для моделирования длин АС – графовую нейронную сеть внимания (GAT).

Результаты. Разработаны модели, которые демонстрируют высокую точность прогнозирования: для диаметров АС с d_{max} обучающий набор достиг значения $R^2 = 0{,}89$, для АС с d_{min} – $R^2 = 0{,}8$. Для длины АС с d_{max} обучающий набор достигал значения $R^2 = 0{,}75$, а тестовый набор – $R^2 = 0{,}74$; для длины АС с d_{min} обучающий набор достиг значения $R^2 = 0{,}77$, а тестовый набор – $R^2 = 0{,}73$.

Заключение. Предложенные модели могут быть использованы для численного моделирования структуры ВАР и оценки адекватности почечного кровоснабжения.

Ключевые слова: внутрипочечное артериальное русло, артериальная бифуркация, численное моделирование, нейронные сети

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Источник финансирования. Авторы заявляют об отсутствии источника финансирования при проведении исследования.

Для цитирования: Зенин О.К., Горбаченко В.И., Грибков Д.Н., Сергиенко А.А., Ilias Miltiadis. Математическое моделирование структуры внутрипочечного артериального русла с помощью регрессионных моделей и графовой нейронной сети. *Бюллетень сибирской медицины.* 2026;25(2):21–31. https://doi.org/10.20538/1682-0363-2026-2-21-31.

INTRODUCTION

The prevalence of kidney disease is steadily increasing every year [1–3]. Recently, the P4 medicine – personalized, predictive, preventive, and participatory – has become increasingly important in modern medicine, including the fight against kidney disease [4]. Modern digital *in vivo* imaging technologies offer new opportunities for diagnosis [5, 6], predicting the results of angioplasty [7], and solving educational problems related to the anatomy and pathology of the intrarenal arterial bed (IAB) [8, 9]. Despite many studies, these problems remain unresolved, which, in our opinion, is largely due to the lack of a digital model adequately describing the structure of the IAB.

Previous attempts at numeric modeling and creating a digital model of the IAB described its structure as a fractal system. This system consists of arterial bifurcations (ABs), which include a parent (D) and two daughter arterial segments (ASs), with larger (d_{max}) and smaller ((d_{min}) internal

diameters. This issue was first addressed in the doctoral dissertation of the German anatomist and embryologist Wilhelm Roux [10]. On the basis of his observations, he concluded that the structure of AB resembles the shape of a jet of liquid flowing out of a tube opening. He was the first to establish a connection between the AB branching angle and the diameters of the lumen of the parent and daughter ASs [10]. This issue was subsequently studied by C.D. Murray [11], H.B. Uylings [12], M. Kretowski et al. [13], F. Cassot et al. [14], P.J. Blanco et al. [15], T. Fredrich et al. [16], F.M. Kashkooli et al. [17], and P. Li et al. [18].

Nevertheless, an adequate quantitative (numerical) description of IAB remains an unsolved problem for two main reasons. Firstly, the channel covers a wide range of AS diameters – from several micrometres to several millimeters [19], which is undeservedly ignored by most authors. Secondly, the complexity of visualizing and obtaining reliable information about AS lengths, which is extremely necessary for their numerical modeling.

The aim was to develop models for numerical modeling of the diameters and lengths of bifurcation segments – structural units of IAB.

MATERIALS AND METHODS

The material for the study was previously obtained from the morphometry of 60 corrosion casts of IAB [20]. This method allows for the measurement of ASs with an inner diameter of 0.1 mm with accuracy of 0.01 mm. IABs are considered to be fractal systems consisting of interconnected ABs (Fig. 1, *a, b*).

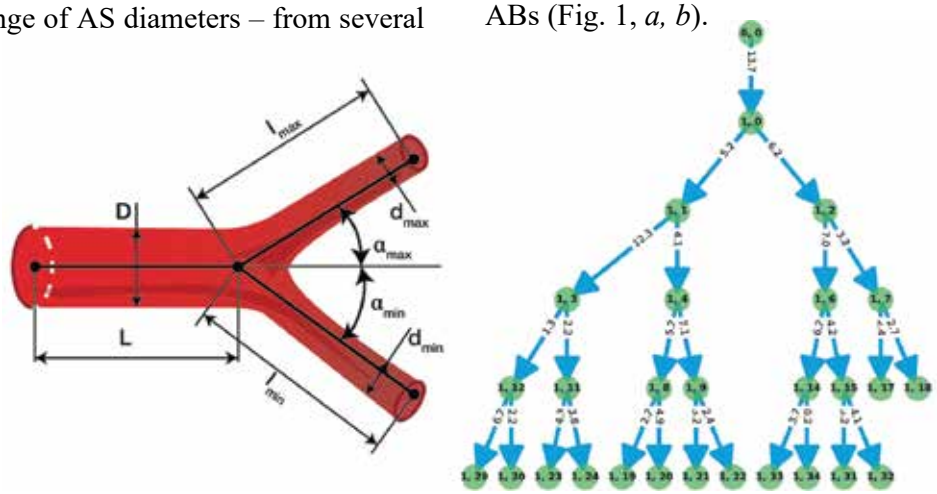

Fig. 1. Diagram of IAB as a fractal system: *a* – AB IAB diagram; *b* – graphical representation of the protocol; 1 – unique numbers are indicated at the graph nodes, and the edges indicate the lengths of AS. *D* – diameter of the proximal segment; d_{max} – diameter of the distal segment with a larger value; d_{min} – diameter of the distal segment with a smaller value; *L* – length of the proximal segment; l_{max} – length of the distal segment with a larger diameter; l_{min} – length of the distal segment with a smaller diameter; α_{max} – branching angle between the distal segment with a larger diameter and the proximal segment; α_{min} – branching angle between the distal segment with a smaller diameter and the proximal segment

The study included 9,642 ABs located at 22 division levels (*i*) and 60 corrosive IAB preparations. The ABs were divided into the following groups: two sex-based groups (male or female ABs); two age groups: middle-aged (men aged 36–60 years, women aged 36–55 years) and elderly individuals (men aged 61–75 years, women aged 56–75 years) (Classification of the USSR Academy of Medical Sciences (1965)); as well as ABs belonging to the corrosion preparations of the left and right kidneys. Previous studies have revealed morphological differences in the arterial bed of the kidneys in individuals of different sexes [21], as well as in the left and right kidneys [22].

The following indicators were used to describe this structure numerically: *i* – serial number of the division level – newly formed AS series; *D* – diameter of the parent proximal AS located at the initial division level; d_{max} – diameter of the daughter distal AS with a larger value; d_{min} – diameter of the daughter distal AS with a smaller value; *L* – length of the parent proximal AS located at the initial division level, distance between the nearest branching points; l_{max} – length of the daughter distal AS with a larger diameter; l_{min} – length of the daughter distal AS with a smaller diameter.

For the diameters of the AB IAB segments, calculations and comparisons of several regression model options were performed. A formal description of the model in statistics assumes the presence of observations, a set of all measured pairs of features – responses, expressed as $\{(\mathbf{x}_i, y_i)\}_{(i=1)}^{n}$, where the observation vector x_i is one element of the sample. In our case, this is one row of the available database [20], where *i* is the observation index, x_i is the value of the predictor (feature) for the i-th observation, and y_i is the response value for this observation.

When a simple linear regression model is applied, the observation vector contains only one element x_i. From a mathematical point of view, simple linear regression has the form represented as $y_i = \alpha + \beta x_i + \varepsilon_i$, where α is the intercept (free term) common to all observations, which does not depend on *x*; β is the slope coefficient, which describes how *y* changes on average when *x* changes by one unit, assuming that the relationship is linear (straight); ε_i is a random error (noise), i.e., the part of the variation in *y* that is not explained by the linear relationship.

The regression problem was solved via the method of least squares. The accuracy of each predicted value is measured by the square of its residual (the vertical distance between the data point and the constructed line), and the goal of the method is to minimize the sum of these squares of deviations. In the case of simple linear regression, the method of least squares leads to a system of two linear equations, the solutions of which are the coefficients α and β.

We also employed a multiple linear regression model that uses a vector of features $x_i = [x_{i1}\, x_{i2} \ldots x_{ip}]^{\mathrm{T}}$ for *p* features, in contrast to simple linear regression, which relies on a single variable. Then, the regression equation takes the form $y_i = \sum_{j=0}^{p} a_j x_{ji}$, where $x_{0i} = 1$. In matrix form, the regression equation takes the form **Xa** = y, where $y = [y_1\, y_2 \ldots y_n]^{\mathrm{T}}$ is the vector of dependent variable values, T is the transpose symbol, and $a = [a_0\, a_1, \ldots, a_k]^{\mathrm{T}}$ is the vector of regression coefficients:

$$X = \begin{bmatrix} 1\, x_{11} \ldots x_{1k} \\ 1\, x_{21} \ldots x_{2k} \\ \vdots\ \vdots \quad\ \vdots \\ 1\, x_{n1} \ldots x_{nk} \end{bmatrix}$$

is the matrix of independent variables (design matrix).

In modern libraries, the regression equation is solved via the singular value decomposition method rather than the method of least squares.

In addition, we explored the use of polynomial regression of degrees 2 and 3. It is a linear model in the space of extended basis features obtained by raising the initial variables to power and multiplying them together. From a mathematical point of view, the general form of polynomial regression is represented as $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \ldots + \beta_n x^n + \varepsilon$, where *y* is the dependent variable response, β_n are unknown terms, ε is noise, and *n* is the degree of the polynomial. In general, the matrix form is $\vec{y} = X\vec{\beta} + \vec{\varepsilon}$, and the solution is $\hat{\vec{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\vec{y}$.

For regression modeling of diameters, the dataset was divided into a training set (7,710 samples, 80 %) and a test set (1,930 samples, 20%). In a simple linear regression approach, the diameter of the parent segment was used as the only predictor. For all other models, a complete set of features was used, including the diameter of the parent vessel, sex, age, and division level.

The following metrics were selected to verify the quality of the results. R^2 is the coefficient of determination. An R^2 close to 1 indicates that the model almost completely explains the variation. An R^2 of 0 means that the model does not explain the variance any better than the mean, while a negative R^2 suggests the model's predictions are worse than simply using the mean as a forecast. The coefficient is calculated by subtracting from one the ratio of the residual sum of squares to the total sum of squares: $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$, where y_i are the actual values, $\hat{y}_i$ are the predicted values, and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ is the average value of the observed data.

The RMSE and MAE metrics were used to evaluate the accuracy of the forecasts. The RMSE is the root mean squared error, measured in the same units as the target variable. The formula is as follows: $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$, where n is the number of observations and $\hat{y}_i$ is the predicted value. The MAE is the mean absolute error , which is less sensitive to outliers than the RMSE: $MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$.

The models were implemented in Python version 3.11.11 in the Ubuntu 24.04 software environment via the scikit-learn library version 1.7.2. The forecasts were visualized via the matplotlib library version 3.10.6. The same linear regression models were used for forecasting $l_{max}$and l_{min} as for d_{max} and d_{min}. The only change for simple linear regression was that the dependence of the daughter length on a single predictor L – vessel length – was investigated, whereas for multiple regression and polynomial regression, the set of features was expanded with the parameters for d_{max} and d_{min}.

RESULTS

Mathematical modeling of AB as a structural element of the fractal system IAB is a difficult task. To construct an accurate digital model of ABs, four tasks were solved simultaneously. Specifically, the following were determined: 1. Diameters d_{max} of distal ASs with large values; 2. Diameters d_{min} of distal ASs with smaller values; 3. Lengths l_{max} of distal ASs with large diameters; 4. Lengths l_{min} of distal ASs with smaller diameters.

The results of the calculations for d_{max} and d_{min} are presented in Table 1 and Figure 2, *a*:

Table 1

Model Quality Indicators for Predicting d_{max} and d_{min}

Model/metrics	Prediction			Prediction		
		RMSE, mm	MAE, mm		RMSE, mm	MAE, mm
Linear regression	0.87	0.21	0.15	0.78	0.22	0.15
Multiple linear regression	0.87	0.21	0.15	0.78	0.21	0.14
Second-degree polynomial regression	0.88	0.20	0.15	0.78	0.21	0.14
Third-degree polynomial regression	0.83	0.24	0.17	0.75	0.23	0.16

Note. Here and in Tables 2 and 3: d_{max} – diameter of the distal AS with a larger value; d_{min} – diameter of the distal AS with a smaller value; R^2 – coefficient of determination; RMSE – root mean squared error; MAE – mean absolute error.

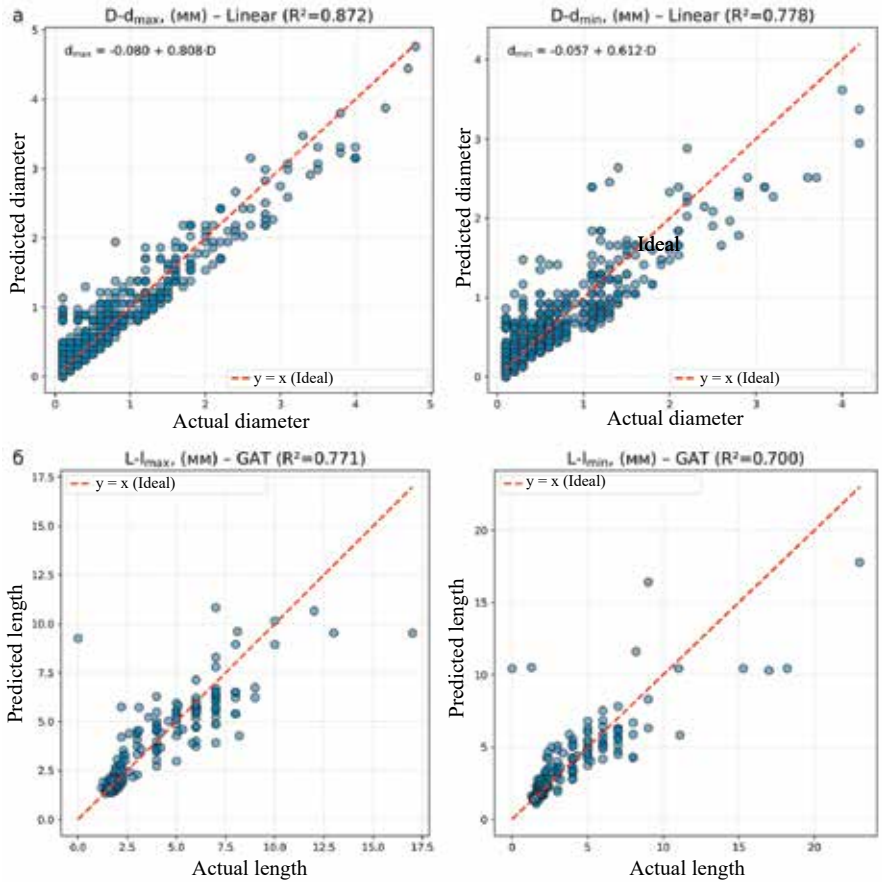

Fig. 2. Predicted/actual graphs for: *a* – simple linear regression for d_{max} and d_{min} of the relationship equations: $d_{max} = –0.080 + 0.808D$ $(R^2 = 0.87)$ and $d_{min} = –0.057 + 0.612D$ $(R^2 = 0.77)$; b – GAT for l_{max} and l_{min}: the coefficient of determination for l_{max} is $R^2 = 0.77$ and l_{min} is $R^2 = 0.70$

As shown above (Table 1 and Fig. 2, *a*), all the models yield good results ($R^2 \geq 0.70$), with both linear regression and second-degree polynomial regression showing the best results. The third-degree polynomial regression models are overfit: R^2 decreases to 0.83 (for d_{max}) and 0.75 (for d_{min}), and the RMSE increases by approximately 10–15% (Table 1).

The results of the calculations for l_{max} and l_{min} are presented in Table 2.

Table 2

Model Quality Indicators for Predicting l_{max} and l_{min}

Model/metrics	Prediction			Prediction		
		RMSE, mm	MAE, mm		RMSE, mm	MAE, mm
Linear regression	0.55	1.43	0.93	0.42	1.48	0.88
Multiple linear regression	0.57	1.40	0.90	0.51	1.35	0.86
Second-degree polynomial regression	0.58	1.38	0.85	0.55	1.31	0.79
Third-degree polynomial regression	0.56	1.41	0.85	0.56	1.30	0.77

The results of regression analysis show that for simple linear regression, the dependence is too complex to detect. As the complexity of the regression model increases, the results improve, but even for second- and third-degree polynomial regressions, the results do not exceed 58%. In this context, the use of neural networks, which can approximate arbitrarily complex functions [23] and capture hidden dependencies between variables that are not accessible to standard regressors [24], has been proposed.

This was the prerequisite for using a neural network approach to solve these problems. The first architecture used was the basis and progenitor of all other networks – a multilayer perceptron [24]. The advantages of a multilayer perceptron are its simplicity of implementation, fast training, and easy integration into the workflow pipeline (as well as into regressors). The main limitation is the inability to account for the structure of the input data (for example, temporal or graph data).

The main idea of a multilayer perceptron is that each neuron of layer l is connected to all the neurons of the previous layer $l-1$. This connection defines a linear transformation, which is usually followed by nonlinear activation. From a mathematical point of view, this can be represented as $\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}$, $\mathbf{a}^{(l)} = \mathbf{f}(\mathbf{z}^{(l)})$, , where $\mathbf{W}^{(l)}$ is the weight matrix of layer l, $\mathbf{b}^{(l)}$ is the bias vector of layer l, $\mathbf{f}(\mathbf{z}^{(l)})$ is a nonlinear activation function (ReLU, tanh, sigmoid, etc.), $\mathbf{a}^{(l-1)}$ is the activation vector of the previous layer, and $\mathbf{a}^{(l)}$ is the activation vector of the current layer. MLP training is implemented via the backpropagation algorithm, which calculates the gradients of the loss function across the network parameters. Parameter optimization is performed via stochastic gradient descent methods or their modern adaptive variants, such as Adam. The choice of loss function depends on the task at hand; MSE is often used for regression.

The next architecture chosen was a recurrent neural network LSTM [25] (long short-term memory network). The choice was made because segments of the vascular system are represented as sequences, the characteristics of which can potentially be predicted by LSTM. The main advantage of LSTM is the preservation of long-term dependencies in memory and resistance to vanishing gradients (unlike conventional recurrent networks). When applied to the vascular structure, each individual AB segment can be represented as a time series, where a daughter vessel grows from a larger vessel, and so on. In vascular bed modeling, the level of vascular bed division acts as an analog of the time series step.

The disadvantages of LSTM include the high computational cost of training and limitations in parallelization. The main advantage of LSTM is its ability to store information effectively over many steps, solving the problem of gradient decay. This allows the network to identify long-term dependencies in the data, which is impossible for conventional recurrent networks. This is achieved by introducing a memory cell that transmits information in steps and three control gates that regulate the flow of information. The gates are fully connected layers

with sigmoid activation functions that output values close to 0 ("closed") and 1 ("open").

At step t, the LSTM components work as follows. The forget gate (f_t) determines what proportion of information from the previous state of the cell (c_{t-1}) needs to be "forgotten." The forget gate is described by the expression $\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f)$, where $[\mathbf{h}_{t-1}, \mathbf{x}_t]$ is the concatenation of the previous hidden state ($\mathbf{h}_{t-1}$) and the current input ($\mathbf{x}_t$), and $\mathbf{b}_f$ is the bias vector. The input gate ($\mathrm{i}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i)$)) determines which part of the new information to update, and the candidate value creates new candidate values that can be added to the cell ($\hat{\mathbf{c}}_t = \tanh(\mathbf{W}_{\hat{c}}[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\hat{c}}$). The cell state update is performed by combining the forgetting of old information and the addition of a new $\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \hat{\mathbf{c}}_t$, where the symbol $\odot$ denotes element-wise multiplication. Old information is multiplied by the "forgetting" coefficient, and new information is multiplied by the "input" coefficient. The output gate determines which part of the updated cell state ($\mathbf{c}_t$) will be used to form the output signal $\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o$. The new hidden state is formed according to the formula

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t).$$

LSTM training, like other recurrent networks, is performed via the backpropagation through time (BPTT) algorithm. The key mechanism of LSTM, additive state updates through gates, allows gradients to propagate without attenuation. This enables the network to remember information over many steps, as the model itself learns to decide when to "open" or "close" the gates.

For comparison with the multilayer perceptron, we selected the Kolmogorov–Arnold Network (KAN), a novel neural network architecture [26]. The KAN architecture is based on the Kolmogorov–Arnold theorem, which states that any continuous function of many variables can be represented as a finite superposition of continuous functions of a single variable. For the input vector $\mathrm{x} = (x_1, x_2, \ldots, x_n)$ KAN, the regression problem for predicting vessel diameters is described by the expression

$$y = \sum_{j=1}^{2n+1} \Phi_j \left(\sum_{i=1}^{n} \varphi_{ij}(x_i) \right),$$

where φ_{ij} are B-splines (approximating basis functions), and Φ_j are nonlinear functions of the output layer (B-splines). The network has two layers with activations on the edges rather than on the nodes. B-splines with trainable coefficients are used as activation functions. The main advantage of Kolmogorov–Arnold networks is their high accuracy with fewer parameters.

Since the AS components of AB in the IAB system can be represented as a graph, the use of graph neural networks [27] to predict vessel diameters has been proposed. The best results were obtained via graph attention networks (GAT) [28]. In this work, PyTorch Geometric (PyG), a specialized PyTorch extension for working with graph neural networks, was used to implement the graph attention network [29].

Let us consider the implementation of a GAT, taking into account the features of PyG. The graph at the input of the graph neural network was represented by two matrices: a vertex matrix and an adjacency matrix. The vertex matrix is a concatenation (combination) of vectors (embeddings) describing the vertices. Considering the bifurcations of the daughter vessel length , they were represented in the node embeddings, which made it possible to exclude the edge matrix. The adjacency matrix, as is customary in graph theory, is a binary matrix in which the element (i, j) is equal to one if the vertex i is connected to the vertex j. Since the adjacency matrix is sparse, PyG is represented as an array of edges containing the indices of the graph edges.

A multi-headed graph attention network is a GAT architecture in which each node of the graph is trained to dynamically weigh the importance of its neighbors simultaneously in several independent projections of the feature space («attention heads»). Multi-headed attention uses K independent attention mechanisms («heads») in parallel to improve stability and expressiveness. Each head k has its own trainable weight matrix $\mathbf{W}^k$ of size $f \times f$, where f is the dimension of the node embeddings at the layer input, and a trainable attention matrix $\mathbf{A}_{\text{att}}^k$ of size $n \times n$, where n is the number of graph nodes. The attention matrix shares structure with the graph adjacency matrix, but the elements of the adjacency matrix are constant and equal to 0 or 1, whereas the elements of the attention matrix are trainable and capture the varying degrees of influence of neighboring nodes. Each attention head performs the following operations (in PyG, the operations are vectorized and performed in parallel).

Linear transformation $\mathbf{H}^k = \mathbf{X}\mathbf{W}^k$ for $k = 1, \ldots K$, where $\mathbf{X}$ is the matrix of input embeddings of the layer with size $n \times f$.

Calculation of attention matrices $\mathbf{A}_{\text{att}}^k$.

Aggregation $\mathbf{Z}^k = A_{\text{att}}^k \mathbf{H}^k$ for $k = 1, \ldots, K$.

The results of all heads are combined. For the hidden layers, concatenation is used $\mathbf{X}'$ = Concat $[\sigma(\mathbf{Z}^1), \sigma(\mathbf{Z}^2), .., \sigma(\mathbf{Z}^K)]$, where σ is the activation function, usually ReLU. The output dimension of the network layer increases $n \times (K \cdot f)$. In the input layer, averaging

$$\mathbf{X}' = \frac{1}{K}\sum_{k=1}^{K}\sigma(\mathbf{Z}^k)$$

is used, and the output dimension of the network becomes equal to the dimension of the input embedding array. Since the regression problem is being solved, a linear activation function is used in the output layer.

The embeddings of nodes in a single layer of a multi-headed attention graph network consider the properties of neighboring nodes. The perception area (how far the influence spreads) is controlled by the number of layers in the network. A graph network with m layers will aggregate information from nodes at a distance of up to m transitions.

To implement neural networks, the cuda library version 12.8 is added to the above information, together with the PyTorch library for implementing neural networks in Python version 2.7.0. Pycan library version 0.0.2 was used to implement KAN. Owing to the limitations of this library, KAN training was also performed on the CPU via the PyTorch library. PyTorch-Geometric library version 2.6.1 [29] was used to implement the GAT graph model.

Each of the four neural networks was trained on an identical dataset. All the networks used a single set of features: parent diameter, diameter of the larger and smaller daughter ASs, parent AS length, sex, age, and division level. The target data were the lengths of the larger and smaller daughter ASs. For each of the trained networks, the Adam optimizer and MSE validation for early stopping were used to prevent overfitting on the training set. The MLP is represented by a single hidden layer with a dimension of 128 with a ReLU activation function and a learning rate of 0.0005. The LSTM is represented by four layers, with 1,024 hidden neurons, at the output of a linear layer for compressing the output data into one neuron, with a learning rate of 0.005. The KAN consists of two hidden layers with two neurons in each, the grid size (number of "nodes" in each function) is 5, the grid range is from –5 to 5, and a learning rate of 0.01. The GAT is represented by two layers with 1,024 neurons, one attention head, a learning rate of 0.001, and weight decay (L2 regularization) of 0.00001.

The graph models (GATs) demonstrate the best prediction quality for both the maximum length of IAB AS (l_{max}) and the minimum length of IAB AS (l_{min}). This confirms the idea that the relationships between individual ASs play a key role in assessing the morphology of the vascular system: at l_{max}, it outperforms all other models on all the metrics; for l_{min}, the GAT retains the best explanatory power (R^2 = 0.70) and the lowest MAE, but its RMSE (1.46 mm) is slightly greater than that of the MLP. This may indicate that when predicting shorter lengths (l_{min}), the model is more sensitive to "outliers" in the data, whereas the MLP is more stable on average; the poor performance of the LSTM indicates that the current dataset does not have a clearly defined sequential structure that can be used by a recurrent architecture.

Neural networks use nonlinear transformations that cannot be written as fixed analytical formulas without fully disclosing all the weights and biases. Therefore, the article presents a structural description of the models, their hyperparameters, and results, which fully meets the requirements of a scientific publication and allows the experiment to be reproduced. In neural networks, prediction is a sequence of matrix multiplication and nonlinear activation functions with many parameters (weights and biases). It is meaningless to write them down as a few formulas; instead, the architecture, hyperparameters, and metric results are published, allowing other researchers to reproduce the model.

The results are presented in Table 3 and Fig.2, *b*.

Table 3

Neural Network Training Results l_{max} and l_{min}

Model/metrics	Prediction			Prediction		
		RMSE, mm	MAE, mm		RMSE, mm	MAE, mm
MLP neural network	0.61	1.32	0.79	0.61	1.21	0.73
LSTM neural network	0.43	1.59	0.95	0.34	1.51	0.93
KAN	0.63	1.28	0.78	0.55	1.29	0.77
GAT neural network	0.77	1.11	0.56	0.70	1.45	0.66

DISCUSSION

Two models were trained to find the diameters of daughter ASs with larger and smaller values. For practical implementation, simple linear regression equations can be used: $d_{max} = -0.080 + 0.808\ D$ ($R^2 = 0.87$) and $d_{min} = -0.057 + 0.612\ D$ ($R^2 = 0.77$) (Fig. 2, *a*). Two models were trained to find the length of AS with larger and smaller diameters, both of which solve the regression problem (Fig. 2, *b*). The training used features that were explicitly known or could be accurately calculated, including the following: current AS diameter; AS diameters with larger values; AS diameters with smaller values; division level (*i*); sex; age group; and current AS length.

Studies have shown that using the organ position (i.e., right or left) impairs the predictive power of the models. Using AS lengths with larger and smaller diameter values increases the accuracy of the model's predictions, but these characteristics cannot be used because of their uncertainty in predicting distal AS lengths.

As a result of the experiments, the best solution was a graph network consisting of two layers of GAT and two attention heads, implemented in the PyTorch Geometric library [29]. The best results were obtained via the Adam optimizer [32], with a learning rate of 0.001 and weight_decay of 0.00001. Retraining was controlled by monitoring the values of R^2 for the test sample: if there was no improvement in the indicator by a delta of 0.01 within 100 epochs, training was stopped.

For accurate modeling of the geometry of the IAB, it is necessary to combine the developed neural network models to determine the distal ASs with maximum and minimum diameters. The results of testing the program were presented in the form of model training quality metrics. The coefficient of determination R^2 was used to evaluate quality. For the length of the distal segment with a large diameter, the training set reached a value of $R^2 = 0.75$, and the test set reached a value of $R^2 = 0.74$. The mean square error function was used as the loss function.

Similar metrics and loss functions were used for the length of the distal segment with a calculation model of a smaller diameter. The training set achieves a value of $R^2 = 0.77$, and the test set achieves a value of $R^2 = 0.73$. The results of our study confirm that the use of a GNN for modeling AB IAB is a promising approach. In particular, the GAT demonstrated high prediction accuracy, which is consistent with other studies in the field of vascular structure modeling [33, 34].

The results obtained coincide with the results of previous studies, such as those by P.J. Blanco et al., which showed that digital models and computer simulations can accurately describe vascular beds. However, unlike these studies, the present study focuses on neural networks, which allows for the consideration of more complex relationships between the AS parameters that make up the AB. Moreover, the neural network approach made it possible for the first time to quantitatively predict the model of the IAB structure – to determine the values of the AS length for each AB, which makes it possible to conduct a more in-depth study and perform numerical modeling of hemodynamics within the IAB.

The use of the GAT offers several advantages over traditional modeling methods. First, graph networks consider the topology of the IAB, which is crucial for modeling complex bifurcation structures [35]. Second, the use of GAT mechanisms allows for more accurate consideration of the importance of AS (edges) and AB (nodes) connections in the graph, as GAT calculates trainable attention coefficients for each AS that determine how important a neighboring AB is to the current one, thereby improving the quality of prediction [28].

Despite these promising results, the study has several limitations. Firstly, the accuracy of predictions depends on the quality of the input data and model parameters. Secondly, the use of neural networks requires significant computational resources, which may limit their application in clinical practice. Further improvements will focus on using more accurate data and optimizing training algorithms.

CONCLUSION

Thus, models were developed for numerical simulation of the diameters and lengths of AB segments, i.e., structural units of IAB. The first module consisting of two regression equations allows the diameters of two daughter (distal) ASs to be calculated on the basis of the diameter of the parent (proximal) AS. The second module, a graph neural attention network, allows for the calculation of the lengths of two daughter (distal) ASs on the

basis of the length of the parent (proximal) AS to determine the lengths of the AB segments of the IAB. The use of a graph attention network has made it possible to achieve high prediction accuracy, opening new opportunities for clinical applications, including preoperative modeling and assessment of whether renal blood supply is normal.

The developed two-module model can serve as a basis for creating methods for the numerical prediction of the volume and area of the kidney region supplied by a given artery. This will help in planning and performing surgical interventions, as well as evaluating their results. Further improvements will focus on using more accurate data and optimizing training algorithms.

REFERENCES

1. Borg R., Carlson N., Søndergaard J., Persson F. The growing challenge of chronic kidney disease: an overview of current knowledge. *Int. J. Nephrol*. 2023:9609266. DOI: 10.1155/2023/9609266.
2. Eckardt K.U., Delgado C., Heerspink H.J.L., Pecoits-Filho R., Ricardo A.C., Stengel B. et al. Trends and perspectives for improving quality of chronic kidney disease care: conclusions from a kidney disease: Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney Int.* 2023;104(5):888–903. DOI: 10.1016/j.kint.2023.05.013.
3. Eckardt K.U., Coresh J., Devuyst O., Johnson R.J., Köttgen A., Levey A.S. et al. Evolving importance of kidney disease: from subspecialty to global health burden. *Lancet*. 2013;382(9887):158–169. DOI: 10.1016/S0140-6736(13)60439-0.
4. Jenkins S.L., Ma'ayan A. Systems pharmacology meets predictive, preventive, personalized and participatory medicine. *Pharmacogenomics*. 2013;14(2):119–122. DOI: 10.2217/pgs.12.186.
5. Tsai S.H., Chu S.J., Chen S.J., Fan Y.M., Chang W.C., Wu C.P. et al. Acute renal infarction: a 10-year experience. *Int. J. Clin. Pract*. 2007;61(1):62–67. DOI: 10.1111/j.1742-1241.2006.01136.x
6. Siegel R.L., Miller K.D., Jemal A. Cancer statistics, 2018. *CA Cancer J. Clin.* 2018;68(1):7–30. DOI: 10.3322/caac.21442.
7. Heuer R., Gill I.S., Guazzoni G., Kirkali Z., Marberger M., Richie J.P. et al. A critical analysis of the actual role of minimally invasive surgery and active surveillance for kidney cancer. *Eur. Urol*. 2010;57(2):223–232. DOI: 10.1016/j.eururo.2009.10.023.
8. Pather N., Blyth P., Chapman J.A., Dayal M.R., Flack N.A.M.S., Fogg Q.A. et al. Forced disruption of anatomy education in Australia and New Zealand: an acute response to the Covid-19 pandemic. *Anat. Sci. Educ*. 2020;13(3):284–300. DOI: 10.1002/ase.1968.
9. Baptiste Y.M. Digital feast and physical famine: the altered ecosystem of anatomy education due to the Covid-19 pandemic. *Anat. Sci. Educ*. 2021;14(4):399–407. DOI: 10.1002/ase.2098.
10. Roux W. Ueber die verzweigungen der blutgefsse. *Eine morphologische Studie*. *Z Naturwiss.* 1878;12:205–266.
11. Murray C.D. The physiological principle of minimum work applied to the angle of branching of arteries. *J. Gen. Physiol*. 1926;9(6):835–841. DOI: 10.1085/jgp.9.6.835.
12. Uylings H.B. Optimization of diameters and bifurcation angles in lung and vascular tree structures. *Bull. Math. Biol*. 1977;39(5):509–520. DOI: 10.1007/BF02461198.
13. Kretowski M., Rolland Y., Bézy-Wendling J., Coatrieux J.L. Physiologically based modeling of 3-D vascular networks and CT scan angiography. *IEEE Trans. Med. Imaging*. 2003;22(2):248–257. DOI: 10.1109/TMI.2002.808357.
14. Cassot F., Lauwers F., Fouard C., Prohaska S., Lauwers-Cances V. A novel three-dimensional computer-assisted method for a quantitative study of microvascular networks of the human cerebral cortex. *Microcirculation*. 2006;13(1):1–18. DOI: 10.1080/10739680500383407.
15. Blanco P.J., de Queiroz R.A., Feijóo R.A. A computational approach to generate concurrent arterial networks in vascular territories. *Int. J. Numer. Method Biomed. Eng*. 2013;29(5):601–614. DOI: 10.1002/cnm.2547.
16. Fredrich T., Welter M., Rieger H. Tumorcode: A framework to simulate vascularized tumors. *Eur. Phys. J. E. Soft Matter*. 2018;41(4):55. DOI: 10.1140/epje/i2018-11659-x.
17. Kashkooli F.M., Abazari M.A., Soltani M., Ghazani M.A., Rahmim A. A spatiotemporal multi-scale computational model for FDG PET imaging at different stages of tumor growth and angiogenesis. *Sci. Rep*. 2022;12(1):10062. DOI: 10.1038/s41598-022-13345-4.
18. Li P., Pan Q., Jiang S., Yan M., Yan J., Ning G. Development of novel fractal method for characterizing the distribution of blood flow in multi-scale vascular tree. *Front. Physiol*. 2021;12:711247. DOI: 10.3389/fphys.2021.711247.
19. Montino Pelagi G., Regazzoni F., Huyghe J.M., Baggiano A., Alì M., Bertoluzza S. et al. Modeling cardiac microcirculation for the simulation of coronary flow and 3D myocardial perfusion. *Biomech. Model. Mechanobiol*. 2024;23(6):1863–1888. DOI: 10.1007/s10237-024-01873-z.
20 Database Registration Certificate No. 2021620982 Russian Federation. Quantitative Anatomy of the Intrarenal Arterial Bed in the Human Kidney : No. 2021620806 : applied on April 27, 2021 : published on May 18, 2021 / Zenin O.K., Dmitriev A.V., Miltykh I.S., Yurchenko O.O.; applicant: Federal State Budgetary Educational Institution of Higher Education Penza State University. (In Russ.) EDN JDPUFN.
22. Kolsanov A.V., Nazaryan A.K., Voronin A.S., Gelashvili O.A. Variant Anatomy of the Renal Arteries: a Computed Tomography Angiography Study. *Journal of Anatomy and Histopathology*. 2020;9(1):43–48. (In Russ.). DOI: 10.18499/2225-7357-2020-9-1-43-48
23. Liu C., Tan Z., He M. Overview of Artificial Intelligence in Medicine. In: Raz M., Nguyen T.C., Loh E. Artificial intelligence in medicine: applications, limitations and future directions. Singapore: Springer, 2022: 23–34. DOI: 10.1007/978-981-19-1223-8_2.
24. Aggarwal C.C. Neural networks and deep learning: a textbook, 2nd ed. Switzerland: Springer International Publishing AG, 2023:553. DOI: 10.1007/978-3-031-29642-0_13.

25. Nielsen A. Practical time series analysis: prediction with statistics and machine learning. Sebastopol: O'Reilly Media, Inc., 2019:504.
26. Wu Y., Su T., Du B., Hu S., Xiong J., Pan D. Kolmogorov–Arnold network made learning physics laws simple. *J. Phys. Chem. Lett*. 2024;15(50):12393–12400. DOI: 10.1021/acs.jpclett.4c02589.
27. Wu L., Cui P., Pei J., Zhao L. Graph Neural networks: foundations, frontiers, and applications. Singapore: Springer Nature, 2022:725. DOI: 10.1007/978-981-16-6054-2.
28. Veličković P., Cucurull G., Casanova A., Romero A., Liò P., Bengio Y. Graph Attention Networks. Published as a conference paper at ICLR, 2018:12. arXiv:1710.10903v3 [stat.ML].
29. PyG documentation. URL: https://pytorch-geometric.readthedocs.io/en/latest/#.
30. Hagberg A.A., Schult D.A., Swart P.J. Exploring Network Structure, Dynamics, and Function using NetworkX. Proceedings of the 7th Python in Science Conference. Pasadena, CA USA. 2008:11–15. URL: https://aric.hagberg.org/papers/hagberg-2008-exploring.pdf
31. Harris C.R., Millman K.J., van der Walt S.J., Gommers R., Virtanen P., Cournapeau D. et al. Array programming with NumPy. *Nature*. 2020;585(7825):357–362. DOI: 10.1038/s41586-020-2649-2.
32. Kingma D.P., Ba J. Adam: A method for stochastic optimization. Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego. 2017. DOI: 10.48550/arXiv.1412.6980.
33. Zhang S., Yang G., Qian J., Xiaomei Zh., L J., Li Pu. et al. A novel 3D deep learning model to automatically demonstrate renal artery segmentation and its validation in nephron-sparing surgery. *Front. Oncol.* 2022;12:997911. DOI: 10.3389/fonc.2022.997911.
34. Xu P., Holstein-Rathlou N.H., Søgaard S.B., Gundlach C., Sørensen C.M., Erleben K. et al. A hybrid approach to full-scale reconstruction of renal arterial network. *Sci. Rep.* 2023;13(1):7569. DOI: 10.1038/s41598-023-34739-y.
35. Thomas S.R. Mathematical models for kidney function focusing on clinical interest. *Morphologie*. 2019;103(343):161–168. DOI: 10.1016/j.morpho.2019.10.043.

Author Contribution

Zenin O.K. — conception and design, final approval of the manuscript for publication. Gorbachenko V.I. — analysis and interpretation of the data. Gribkov D.N. — software development, analysis and interpretation of the data. Sergienko A.A. — critical revision for important intellectual content. Ilias Miltiadis — justification of the manuscript.

Author Information

Zenin Oleg K. – Dr. Sci. (Med.), Professor, Professor of Human Anatomy Department, Penza State University, Penza, zen.olegz@gmail.com, https://orcid.org/0000-0002-5447-1989

Gorbachenko Vladimir I. – Dr. Sci. (Tech.), Professor, Head of Computer Technologies Department, Penza State University, Penza, gorvi@mail.ru, https://orcid.org/0000-0002-1012-8855

Gribkov Dmitry N. – Postgraduate Student, Computer Technologies Department, Penza State University, Penza, rororol23@mail.ru, https://orcid.org/0000-0003-4728-5923

Sergienko Anna A. – Postgraduate Student, Human Anatomy Department, Penza State University, Penza, anya.gurieva00@mail.ru, https://orcid.org/0009-0005-1380-8767

Ilias Miltiadis – Graduate Student, University of Palermo, Palermo, Italy, contact@miltykh.com, https://orcid.org/0000-0002-9130-3255

(✉) **Zenin Oleg K.,** zen.olegz@gmail.com

Received on June 30, 2025;
approved after peer review on October 25, 2025;
accepted on October 30, 2025