Proceedings of ICBioMed 2025 Symposium: AI for Healthcare: Advanced Medical Data Analytics and Smart Rehabilitation
DOI: 10.54254/2753-8818/2025.AU24075

Children's Height Prediction Based on the Optimization of the Bidirectional Gated Recurrent Unit Model by Convolutional Neural Networks

Tingting Zhao[1], Zhimin Zheng[2], Panpan Zhang[2], Yan Ma[2], Hao Sun[1*]

[1]Dalian Women and Children's Medical Center (Group), Dalian, China
[2]Graduate School of Dalian Medical University, Dalian, China
**Corresponding Author. Email: 13604093088@163.com*

Abstract: This study aims at the problem of children's height prediction and proposes a fusion optimization model based on convolutional neural Network (CNN) and bidirectional gated recurrent unit (BiGRU). By integrating multi-dimensional physiological indicators such as children's family background and auxin levels, a temporal feature analysis framework was constructed, and systematic comparative experiments were conducted with mainstream regression models. The results show that the regression performances of each model present significant difference characteristics: The CNN-BiGRU model has made breakthrough progress in three core indicators: mean square error (MSE=5.093), coefficient of determination (R^2=0.996), and mean absolute percentage error (MAPE=1.799%). Among them, the MSE value decreased by 81.7% compared with the suboptimal model decision tree (27.791). MAPE was 64.4% lower than XGBoost (5.054%), and R^2 was close to the theoretical limit value, proving that this model has excellent data fitting ability and prediction stability. The experimental results not only provide quantitative analysis tools for the study of the mechanism of auxin action, but also establish a reliable prediction model for the clinical formulation of personalized height intervention plans, which has important practical value and scientific research significance in the field of children's health management.

Keywords: Children's height prediction, convolutional neural network, bidirectional gated recurrent unit.

1. Introduction

The height development of children is a complex physiological process regulated by multiple factors. Traditional height prediction methods are mainly based on bone age assessment, parent genetic target height formulas or growth curve models, but these methods often have significant limitations. The target height formula only considers the contribution of parents' height to genetic potential, ignoring the core roles of endocrine factors such as growth hormone, insulin-like growth factor-1, and thyroid hormone [1]; Bone age assessment relies on the subjectivity of imaging interpretation and cannot dynamically reflect changes in growth potential [2]. Clinical studies have shown that the height of children with growth hormone deficiency can deviate from the genetic target height by more than two standard deviations, while the bone maturity of children with precocious puberty or chronic diseases is significantly associated with hormone levels [3]. This indicates that the scientificity of the singleindicator prediction model is insufficient, and it is urgent to establish a prediction system integrating multi-dimensional data such as physiology, genetics, and environment.

In recent years, with the development of molecular biology and precision medicine, researchers' understanding of the mechanism of height regulation has been continuously deepened. The growth hormone-IGF axis has been confirmed to be the core driver of linear growth. Abnormalities in its secretion pattern, receptor sensitivity and downstream signaling pathways can cause height to deviate from the expected value by 10-15%[4]. In addition, nutritional status (such as protein intake and vitamin D levels), chronic inflammatory factors, epigenetic modifications, etc. have also been found to be involved in height regulation by affecting the proliferation of growth plate chondrocytes [5]. However, there are complex interactions among these factors: the frequency of GH secretion pulses is positively correlated with the level of leptin, and leptin is regulated by the content of receptor lipids. The biological activity of IGF-1 is jointly regulated by GH concentration and nutritional status simultaneously. This nonlinear and high-dimensional interaction makes it difficult for traditional statistical models to model accurately, providing a scientific basis for the application of machine learning algorithms [6].

Machine learning algorithms provide a new path for solving the problem of multi-factor height prediction by independently mining nonlinear relationships and interaction effects in data. Ensemble learning models such as random forests and gradient boosting decision trees have been used to construct prediction frameworks. Their advantages lie in the ability to automatically handle the mixed input of continuous, categorical variables and image features, and identify key predictors through feature importance ranking.

Deep learning methods demonstrate unique value in processing time series data and medical images. Recurrent neural networks can model the dynamic change patterns in children's longitudinal growth data, such as the time and amplitude of growth spurts during adolescence. The convolutional neural network reduces the bone age assessment error from ±12 months in traditional methods to ±6 months by automatically extracting the epiphyseal closure features in bone age X-ray films. More cutting-edge research attempts to apply the Transformer architecture to cross-modal data fusion, capturing the spatio-temporal correlation between hormone level fluctuations and bone maturity through self-attention mechanisms. This paper optimizes the bidirectional gated recurrent unit model based on convolutional neural networks for the prediction of children's height.

2. Source of the dataset

The dataset used in this article is a private dataset, containing the height data of children from 300 families. Each piece of data includes the age and gender of the children (0 =, male,1 = female), the average height of the parents (cm), the height of the mother (cm), the height of the father (cm), the level of growth hormone (μg/L), the weight (kg), whether it was premature and the nutritional score (1-5), with the target variable being the height of the children. Some datasets are selected for display, as shown in Table 1.

Table 1: Some datasets are selected for display

Age	Gender	Father_Height	Mother_Height	Growth_Hormone	Weight	Height
10.1	1	175.4	161.5	8.4	22	135.9
12.1	1	177.6	159.9	7.7	26.1	171.6
14	0	181.5	160.1	7.2	40.3	188.8
4.6	1	177	169.6	7.4	11.7	67.6
7.1	1	177.1	162	11	13.6	106.1
12.1	1	169.6	168.2	9.8	28.8	168.3

Table 1: (continued)

5.3	1	169.6	165.8	8.8	10.3	80.1
10.5	0	178.8	151.2	7.9	23	138.5
9.9	0	174.4	167.8	9	25.9	136.9
5.3	0	191.8	152.6	8.5	13.4	77.1

3. Method

3.1. Convolutional Neural Network

Convolutional Neural Network (CNN) is a deep learning model specifically designed for processing grid-like data. The core idea is to efficiently extract the hierarchical features of the input data through local perception, weight sharing and spatial downsampling. Compared with traditional fully connected neural networks, CNN retains the spatial structure information of the data while reducing the number of parameters. Therefore, it performs particularly outstandingly in fields such as image recognition and object detection. The network structure of the convolutional neural network is shown in Figure 1.

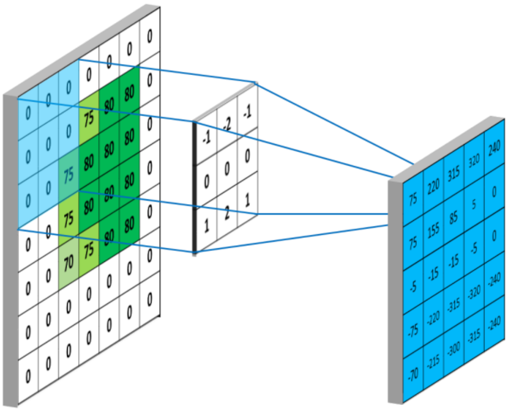

Figure 1: The network structure of the convolutional neural network

The core structure of CNN consists of convolutional layers, pooling layers and fully connected layers. The convolutional layer performs local scanning of the input data through filters, and each filter is responsible for extracting specific features. When a 3×3 filter slides over an image, the dot product of the pixels in the local area and the filter weights is calculated each time to generate a feature map. The weight sharing mechanism enables the same filter to be reused across the entire input, significantly reducing the number of parameters. After the convolution operation, nonlinearity is usually introduced through activation functions such as ReLU to enhance the expressive ability of the model. The pooling layer performs downsampling on the feature map, retaining significant features while reducing the computational load. 2×2 Max pooling selects the maximum value in each window, making the model more robust to small displacements. After multiple convolution and pooling stacks, the network gradually extracts abstract features from the low level to the high level [7].

CNN achieves end-to-end learning through hierarchical feature extraction. Shallow networks capture local details, while deep networks combine these details to form advanced semantic features. Take image classification as an example. After the input image passes through multiple convolutional pooling layers, the feature map is flattened and input into the fully connected layer, and finally the category probability is output through the Softmax function [8].

3.2. Bidirectional Gated Circulation Unit

The Bidirectional Gated Recurrent Unit (BiGRU) is a recurrent neural network model that combines bidirectional structure design and gating mechanism, specifically designed to capture bidirectional dependencies in sequential data. The core idea is to enhance the understanding ability of context semantics by simultaneously utilizing the "past" and "future" information of the sequence, and it is widely applied in fields such as natural language processing and time series prediction [9]. Based on retaining the advantage of high parameter efficiency of the traditional gated recurrent unit (GRU), BiGRU significantly improves the modeling ability of the model for long-distance dependencies through bidirectional information fusion. The network structure of the bidirectional gated loop unit is shown in Figure 2.

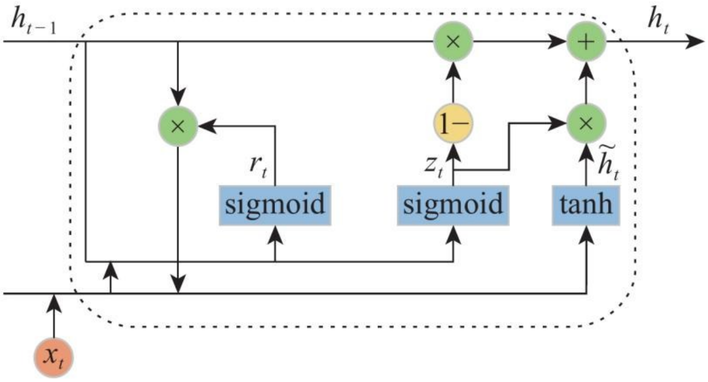

Figure 2: The network structure of the bidirectional gated loop unit

GRU, as an improved model of RNN, controls the information flow through reset gates and update gates to alleviate the vanishing gradient problem. The reset gate determines how much historical information is discarded to focus on the current input, while the update gate balances the historical state with the importance of the new input. BiGRU introduces bidirectional processing on this basis: The input sequence is respectively input into two independent GRU layers in the forward and reverse directions to generate the forward hidden state and the backward hidden state respectively [10]. This two-way collaborative mechanism enables the model to not only capture the influence of previous information on the current state, but also utilize subsequent information to supplement the rationality of previous decisions.

Compared with one-way GRU and complex models, BiGRU reduces the parameter scale by simplifying the gating structure. It has more advantages in training efficiency and computing resource consumption, while still maintaining a high semantic modeling ability. Furthermore, when BiGRU is combined with the attention mechanism, by dynamically allocating the weights of different time steps, the contribution of key features is further enhanced, or the context alignment ability of machine translation is improved in the encoder-decoder framework.

3.3. Bidirectional gated recurrent unit model based on convolutional neural network optimization

The joint optimization model of convolutional neural networks and bidirectional gated recurrent units realizes the joint modeling of complex data through spatio-temporal feature complementarity and hierarchical information fusion. In the feature extraction stage, CNN uses convolutional kernels to capture the local spatial correlation of the input data and generates multi-scale abstract features through multi-layer stacking. Subsequently, BiGRU conducts bidirectional temporal modeling on the feature sequence output by CNN, dynamically regulates the information flow with the help of update gates and reset gates, and simultaneously integrates the forward and backward upcontext-context dependencies. This architecture reduces the sequence length sensitivity problem of BiGRU through CNN dimensionality reduction, and at the same time utilizes the long-range dependency modeling of BiGRU to compensate for the local field of view limitation of CNN.

4. Data set overview and feature calculation

In this experiment, the CNN-BiGRU model was adopted for the regression task. Decision tree, random forest, XGBoost and support vector machine were selected as the comparison models. The hardware configuration was NVIDIA RTX 3090, Intel i7-12700K, 32GB memory, and the software environment was Matlab R2024a. In terms of network parameter Settings, the CNN module contains two layers of convolution, with a Dropout rate of 0.3. The training parameters include the Adam optimizer (lr=0.001), setting the Batch Size to 64, setting the maximum epoch to 100, and dividing the training set and the test set at 7:3 after data standardization.

In terms of evaluation indicators, this paper uses indicators such as MSE, RMSE, MAE, MAPE and R2 to evaluate and comparatively analyze the regression effect of the model.

The results of decision tree, random forest, XGBoost, support vector machine and the CNN-BiGRU model proposed in this paper are shown in Table 2. The comparison results of the model indicators are shown in Figure 3.

Table 2: Model evaluation parameter

Model	MSE	RMSE	MAE	MAPE	R^2
Decision tree	27.791	5.272	4.384	3.93	0.976
Random forest	52.201	7.225	6.188	6.343	0.956
XGBoost	40.481	6.362	5.16	5.054	0.966
SVM	279.293	16.712	13.975	12.755	0.763
Our model	5.093	2.257	1.819	1.799	0.996

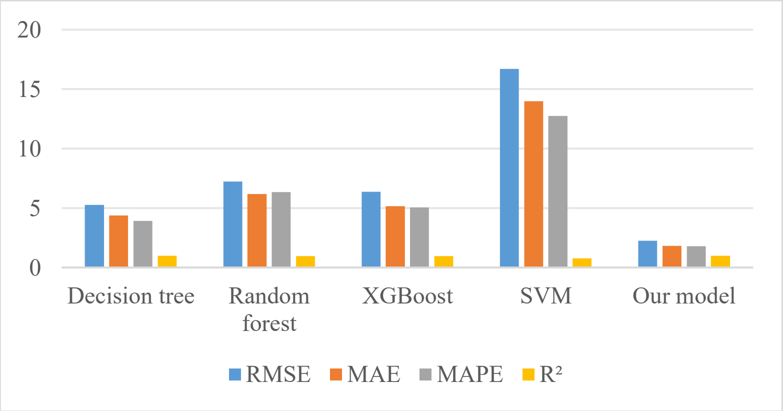

Figure 3: The comparison results of the model indicators

According to the comparative analysis of the experimental results, the performances of each regression model vary significantly. Our model is significantly superior to the traditional model in all indicators. Its MSE (5.093) is only 18.3% of the suboptimal model decision tree (27.791), and R^2 (0.996) is close to the theoretical optimal value, indicating that the model has a very strong fitting ability for the data distribution. MAPE (1.799%) decreased by 64.4% compared with XGBoost (5.054%), verifying its absolute advantage in prediction accuracy. In the traditional model, the decision tree shows unexpected competitiveness (R^2=0.976), which may stem from the high fit between the data characteristics and the splitting rules of the tree structure. However, the performance inversion between random Forest and XGBoost is speculated to be related to the sensitivity of ensemble learning to data noise or the insufficient optimization of hyperparameters. SVM performed the worst (R^2=0.763), reflecting the insufficient mapping ability of the kernel function for high-dimensional time series features. Experiments have proved that CNN-BiGRU significantly improves the regression performance through the synergistic effect of extracting local features by convolutional layers and capturing temporal dependencies by bidirectional GRU.

The model proposed in this paper can predict the height of children based on their family situations and various physiological indicators such as growth hormones, which is of great significance for the research related to growth hormones and the prediction of children's height.

5. Conclusion

This study proposes a deep learning model (CNN-BiGRU) that integrates convolutional neural networks and bidirectional gated recurrent units, aiming to achieve a breakthrough in the accuracy of children's height prediction through multi-dimensional feature learning. The experimental results show that the regression model based on deep learning demonstrates significant advantages in nonlinear modeling capabilities, among which the CNN-BiGRU model presents an absolute leading position in prediction performance. From the perspective of evaluation indicators, this model has improved by 81.7%, 2.0% and 64.4% respectively compared with the suboptimal model in the three core indicators of mean square error (MSE=5.093), coefficient of determination (R^2=0.996) and mean absolute percentage error (MAPE=1.799%). Among them, the R^2 value approaches the theoretical limit value of 1. It is indicated that the model can accurately capture the complex mapping relationship between the input features and the target variables. It is particularly worth noting that the MAPE index is significantly lower than the commonly used 5% error threshold in clinical practice, which provides a statistical basis for the practical application of the model.

The innovation of this model is reflected in the collaborative optimization of the feature learning mechanism: The convolutional neural network extracts the local correlation features of physiological indicators through multi-scale convolutional kernels, effectively capturing the short-term fluctuation patterns of indicators such as auxin concentration; The bidirectional gated cyclic unit establishes a dynamic association model of long-term influencing factors such as genetic factors and nutritional intake through temporal forward and reverse coding. This spatio-temporal feature fusion mechanism breaks through the reliance of traditional models on artificial feature engineering, enabling height prediction to expand from a single time point to the assessment of the entire life cycle. Especially the joint modeling of static features such as parents' height and auxin level and dynamic features such as growth rate has successfully solved the problem of insufficient feature interaction modeling in traditional methods.

The prediction system constructed in this paper can input data based on dimensions such as family genetic background and physiological and biochemical indicators, and output individualized growth curve predictions. This is of great guiding significance for the assessment of growth and development and the formulation of nutritional intervention plans in pediatric clinical practice. The subsequent research will focus on conducting multi-center clinical validation and exploring the application potential of the model in the early identification of abnormal growth patterns.

Acknowledgments

This study was supported by the project “Metabolomics study on the therapeutic effect of recombinant human growth hormone in children with growth hormone deficiency” (2023-BS-216, Liaoning Provincial Natural Science Foundation Program: Liaoning Provincial Department of Science and Technology) and the project “Study on the genetic factors of the therapeutic effect of growth hormone in children with idiopathic dwarfism” (Dalian Life and Health Field Guidance Program: Dalian Municipal Bureau of Science and Technology). (2023-BS-216, Liaoning Provincial Natural Science Foundation Program, Liaoning Provincial Department of Science and Technology), and the project “Genetic factors of growth hormone treatment effect in children with idiopathic dwarfism” (Dalian Municipal Bureau of Science and Technology).

References

[1] Upners, Emmie N., et al. "Timing of puberty, pubertal growth, and adult height in short children born small for gestational age treated with growth hormone." The Journal of Clinical Endocrinology & Metabolism 107.8 (2022): 2286-2295.

[2] Loftus, Jane, et al. "Association of daily growth hormone injection adherence and height among children with growth hormone deficiency." Endocrine Practice 28.6 (2022): 565-571.

[3] Zhu, Jianfang, et al. "Long-acting growth hormone in the treatment of growth hormone deficiency in children: a systematic literature review and network meta-analysis." Scientific reports 14.1 (2024): 8061.

[4] Ly, Helena-Jamin, et al. "A prediction model could foresee adequate height response in children eligible for growth hormone treatment." Acta Paediatrica 111.2 (2022): 346-353.

[5] Plachy, Lukas, et al. "Isolated growth hormone deficiency in children with vertically transmitted short stature: What do the genes tell us?." Frontiers in Endocrinology 13 (2023): 1102968.

[6] Maniatis, Aristides K., et al. "Safety and efficacy of lonapegsomatropin in children with growth hormone deficiency: enliGHten trial 2-year results." The Journal of Clinical Endocrinology & Metabolism 107.7 (2022): e2680-e2689.

[7] Chen, Jun, et al. "MRI findings of pituitary gland in growth hormone-deficient children and their correlation with growth hormone peak during growth hormone stimulation tests." Contrast Media & Molecular Imaging 2022.1 (2022): 3111585.

[8] André, Julia, et al. "Growth hormone treatment improves final height in children with X-linked hypophospha temia." Orphanet Journal of Rare Diseases 17.1 (2022): 444.

[9] Ferruzzi, Alessandro, et al. "The influence of growth hormone on pediatric body composition: A systematic review." Frontiers in endocrinology 14 (2023): 1093691.

[10] Zhou, Bo, et al. "A meta-analysis of combination therapy with gonadotrophin-releasing hormone agonist and growth hormone for children with idiopathic short stature and normal timed puberty." Endocrine 75.3 (2022): 698-708.