

ORIGINAL RESEARCH

A random forest-neural network coupled model for predicting the recurrence location of uterine cancer

Fengchun Liu^{1,2,3,4,5}, Xiangdong Huang^{2,3,4,5,*}, Jian Wang^{2,5}, Liya Wang^{1,2,3,4,5}, Jingguo Qu^{1,3,4,6}, Dianbo Hua⁷

¹Hebei Engineering Research Center for the Intelligentization of Iron Ore Optimization and Ironmaking Raw Materials Preparation Processes, North China University of Science and Technology, 063509 Tangshan, Hebei, China

²Hebei Key Laboratory of Data Science and Application, North China University of Science and Technology, 063509 Tangshan, Hebei, China

³The Key Laboratory of Engineering Computing in Tangshan City, North China University of Science and Technology, 063509 Tangshan, Hebei, China

⁴Tangshan Intelligent Industry and Image Processing Technology Innovation Center, North China University of Science and Technology, 063509 Tangshan, Hebei, China

⁵College of Science, North China University of Science and Technology, 063509 Tangshan, Hebei, China

⁶College of Qian An, North China University of Science and Technology, 064400 Tangshan, Hebei, China

⁷Beijing Sitairui Cancer Data Analysis Joint Laboratory, 101149 Beijing, China

*Correspondence

huangxd@stu.ncst.edu.cn
(Xiangdong Huang)

1. Introduction

Cancer is an extremely important disease that poses serious threats to human health and is an important research topic in medicine, especially due to the associated high mortality rate. The cancer types vary according to the patient's disease location. The World Health Organization statistics indicated that more than 60% of the world's deaths are due to cancer, diabetes, cardiovascular disease and chronic respiratory disease. Uterine cancer is an extremely common type of female cancer. Based on comparative analyses of statistical data from China and the United States [1], it is reported that uterine cancer-related deaths continue to increase, making it one of the ten leading cancers in women. The age group of uterine cancer patients also covers most of the female population, including young women in their 20s and 30s.

Rapid developments and great progress in artificial intelligence have accelerated research on uterine cancer. Some researchers have used the Hibert-Schmidt independent criterion

Abstract

With increasing developments and progress, the status and influence of women in society have significantly improved, with more attention paid to women's health. According to relevant statistics, uterine cancer is a highly-incident malignant tumor in females and urges more research to improve the survival rate of uterine cancer patients. In this study, we established a prediction model to determine the location of uterine cancer recurrence by combining random forest and neural network algorithms. Data of uterine cancer patients were collected from major hospitals, and professional doctors evaluated and graded the patients' physical fitness indicators based on their experience, which were then used to construct the model and obtain the prediction results. Compared to traditional method, the proposed method of this paper showed that the model was more effective and accurate in predicting the location of uterine cancer recurrence, with a prediction accuracy rate of up to 88.63%.

Keywords

Uterine cancer; Random forest; Neural network; Coupled model

of the diversity of artificial fish to perform feature selection and used recurrent neural networks to establish a model to predict survival in patients with recurrent uterine cancer [2]. Similarly, some researchers have used machine learning algorithms to analyze the information data of 5112 women and compared their model analysis results with other algorithms of the same type, which showed excellent reliability of their proposed model [3]. To analyze and evaluate the risk status and annual recurrence risk of early uterine cancer patients, Cibula *et al.* [4] established a numerical model of annual risk of uterine cancer recurrence based on the proportional hazards model developed by the International Uterine Cancer Surveillance Consortium. Identifying potential biomarkers for uterine cancer treatment and prognosis prediction is beneficial in improving the outcomes of the patients [5]. Variation in Apparent diffusion coefficient (ADC) data of primary tumor will be a potential predictor of uterine cancer recurrence [6].

To sum up, there are many studies on the prevention and treatment of uterine cancer. A risk assessment calculation

model has been formed in the risk assessment of early uterine cancer patients, and machine learning algorithms are better than traditional methods in predicting the survival rate and recurrence time of uterine cancer patients. Early uterine cancer patients have no obvious symptoms, and the age of onset presents a double-peak shape, which spreads to young women. Uterine cancer patients after cure still have a certain probability of recurrence, and once recurrence is difficult to be cured again, timely detection of recurrence symptoms, possible recurrence location and timely treatment can effectively improve the survival rate of patients. In order to help doctors diagnose uterine cancer patients, complete the scientific prediction of the recurrence position of postoperative uterine cancer patients. In this paper, using the basic principle of algorithm coupling, combining random forest algorithm and one-dimensional convolutional neural network, a new prediction model of postoperative recurrence position of uterine cancer patients was established. The six different index data in the body indicators are used as important reference indicators. The model established in this paper can achieve better results in predicting the recurrence position of uterine cancer patients, and the prediction accuracy is higher than that of the same type of algorithm, the prediction accuracy rate reached 88.63%.

2. Experimental model

2.1 Basic introduction to the model

Combining the relevant algorithms in the field of artificial intelligence can effectively help the medical field to complete the data sorting and analysis process for the prevention and treatment of uterine cancer patients, which can further improve the application effect of machine learning algorithms in the field of uterine cancer treatment. Experiments show that the random forest algorithm can better detect the influence between different parameter indicators, and has good model generalization ability. Neural networks have better feature extraction capabilities, but have more stringent requirements on the quality and size of datasets, but the random forest algorithm does not require high numbers of datasets. Therefore, this paper uses the basic principle of the coupling algorithm to establish a new prediction model for the postoperative recurrence position of uterine cancer patients by combining the random forest algorithm and the convolutional neural network algorithm. The model combines the random forest algorithm and neural network. The advantages of each network can further improve the accuracy of the algorithm model in predicting the location of postoperative recurrence in patients with uterine cancer. The basic process of predicting the postoperative recurrence position of uterine cancer patients through the constructed model is shown in Fig. 1. First, the original data of uterine cancer patients in major hospitals are collected and organized into the original data set of uterine cancer patients. Then professional doctors annotated and graded the patient's specific physical indicators according to their rich experience, and then organized into an annotated and graded uterine cancer patient dataset. Finally, these datasets are quantified and input into the model to get the final classification prediction results.

In the experiments carried out in this paper, the original

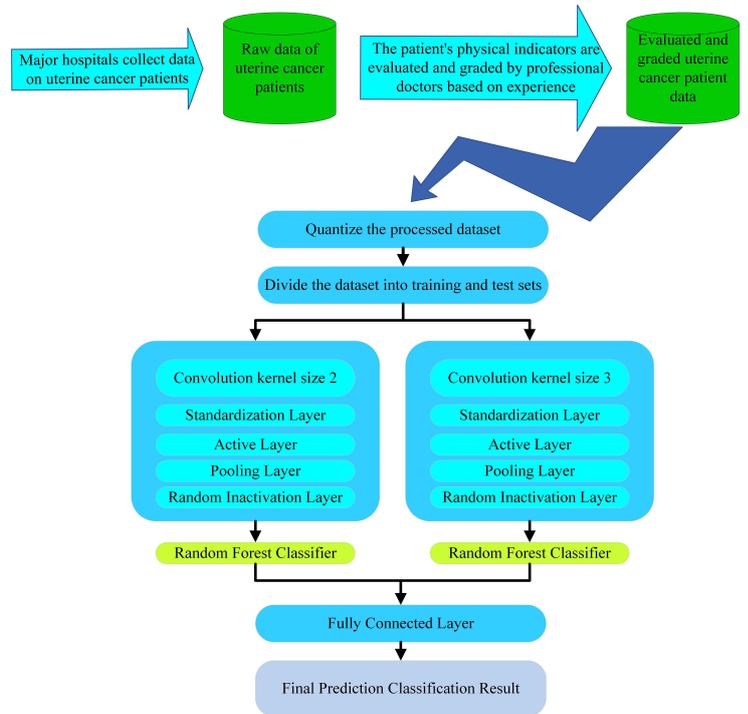


FIGURE 1. Random-neural network model for prediction of patient relapse location.

data of uterine cancer patients were first collected and sorted through major hospitals, and then professional doctors evaluated and analyzed some of the patient's physical characteristics data according to their experience, and divided them into reasonable levels. In order to make the data set suitable for the model established in this paper, the patient data marked by doctors is quantified and divided into training set and test set according to a certain proportion, and then the data set is put into the model as input data. After feature extraction through convolutional networks with convolution kernels of different sizes, the random forest classifier constructed by the random forest algorithm further improves the classification ability of the model. Finally, the calculation of multiple convolutional neural networks with different convolution kernel sizes. The results are merged into the fully connected layer, and finally the final classification result of the model is output by the fully connected layer.

2.2 Random Forest Part

The random forest algorithm belongs to a form of decision tree algorithm and also represents a kind of ensemble algorithm. It was developed by Leo Breiman Cutler, and because the algorithm has a certain anti-noise ability, its use in the industrial field is more advantageous compared with other algorithms of the same type [7–9]. The random forest algorithm is often a combination of multiple different decision tree classifiers, enabling random forests to handle nonlinear data. At the same time, it can also process high-dimensional data. These advantages make the random forest algorithm have advantages that other algorithms do not have when dealing with medical data with high data dimensions, various types and large amounts of data. These advantages also make the

random forest algorithm have many practical applications in the medical field [10, 11]. The Gini coefficient is an important parameter of each classifier in the random forest algorithm, and the calculation formula is shown as follows:

$$\text{Gini}(T) = 1 - \sum_i^c [p(i | T)]^2 \quad (1)$$

Here, c represents the number of categories in the dataset, i represents the i -th category, and T refers to the dataset. A larger Gini coefficient represents a greater uncertainty, and vice versa, and a smaller uncertainty represents a more thorough data segmentation. If the dataset T is divided into n subsets of T_a by the feature A , the Gini index of the subsets divided by the attribute A after the splitting is:

$$\text{Gini}_A(T) = \sum_{a=1}^n \frac{|T_a|}{T} \text{Gini}(T_a) \quad (2)$$

And its gain index calculation formula is shown in formula (3):

$$\Delta \text{Gini}(A) = \text{Gini}(T) - \text{Gini}_A(T) \quad (3)$$

In the prediction model built in this paper, adding the random forest algorithm structure is beneficial to further improve the model's ability to extract features from the dataset and resist overfitting, thereby improving the model's accuracy.

2.3 Convolutional Neural Network Part

The convolutional neural network algorithm is one of the most widely used artificial intelligence algorithms, and it is also a relatively mature algorithm among many machine learning algorithms. It has practical applications in many fields and has achieved excellent results, especially in Outstanding results have been achieved in areas such as image recognition, image processing, character recognition, and speech recognition. In the field of medical research, neural network algorithms play more roles in image data processing such as medical image segmentation and medical image recognition [12–15], while in the processing and analysis of other text-based datasets in the medical field, such as Neural network algorithms and their improved algorithms are also used in classification and regression tasks on some medical datasets [16, 17]. The convolutional neural network contains three different network layer structures, and the calculation method of the j -th feature image of the first layer in the convolutional layer can be expressed as:

$$X_j^l = g \left(\sum_{x_i^{j-1} \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \right) \quad (4)$$

Here, g represents a nonlinear activation function and the feature map dataset connecting the j -th feature map in the 1-1 layer and the first layer is marked as M_j , which is the input

feature image collection. b_j^l refers to the offset value, K_{ij}^l refers to the convolution kernel connecting the i -th feature map in layers 1-1 and the j -th map in layer l . In the pooling layer network structure, for a pooling layer denoted as layer l , the j -th feature map x_j^l of the first layer can be expressed as:

$$x_j^l = g (w_j^l \text{pool}(X_j^{l-1}) + b_j^l) \quad (5)$$

Here, w_j^l refers to the weight coefficient, b_j^l refers to the bias parameter value, $\text{pool}()$ refers to the pooling function in the pooling layer. In the fully connected layer part, the calculation formula of the final output vector x^l of the fully connected layer is:

$$x^l = g \left((\beta^l)^T v^{l-1} + b^l \right) \quad (6)$$

Here, v^{l-1} refers to the feature map vector data obtained by the pooling layer of the 1-1 layer or the vector representation of the feature map output of the convolution layer. b^l refers to the bias parameter value and β^l refers to the weight coefficient matrix.

The convolutional neural network used in the model established in this paper is a one-dimensional neural network used for processing text datasets. At the same time, the dimensions of three different network layer structures are adjusted for application to text datasets.

3. Experiment Description

3.1 Experimental environment and design of comparative experiments

The basic experimental environment to establish the prediction model of predicting the location of postoperative uterine cancer recurrence used in this paper were: Python v3.7 and the third-party libraries that relied on neural networks, including Keras and TensorFlow. The operating system of the hardware platform was Window 10, processor: AMD Ryzen 5 4600H with Radeon Graphics 3.00 GHz, with an onboard RAM of 16.0 GB.

The comparative experiment used the following four algorithms: the basic random forest algorithm, the nearest neighbor algorithm, the one-dimensional convolutional neural network algorithm, and the Bayesian algorithm. The same datasets were tested using the experimental models established in this paper. The results of the algorithms were evaluated experimentally based on the evaluation indicators designed in this experiment to compare their pros and cons in predicting the location of postoperative uterine cancer recurrence.

3.2 The dataset used in the experiments

The data set used in this article is derived from the original data of uterine cancer patients that have been collated from several hospitals for a long time. When obtaining some data information of the patient, the patient will be notified in advance, and the patient will be notified after the consent of the patient. Data

collection and organization. When actually collecting data on uterine cancer patients, professional doctors carefully measure and evaluate some of the patient's physical parameters. According to the existing classification and grading standards, the obtained patient data information will be evaluated and graded. In this study, a data set containing 1652 uterine cancer patients collected from multiple hospitals was used, and 1324 patient data sets were finally used after preliminary data screening. During the experiment, this part of the data set was divided according to the ratio of training set: test set to 8:2. At the same time, in order to further test the training effect of the model, the training set is subdivided into the training set and the validation set according to the ratio of 8:2.

After analyzing a large amount of patient data information and consulting with professional doctors, the collected original data of uterine cancer patients were further deleted and processed. It is found that the recurrence of uterine cancer patients is closely related to 6 different indications in the data information of uterine cancer patients, which are immune indications, tumor indications, microenvironment indications, psychological indications, and nutritional indications. and aerobic exercise and advanced occupational index. After obtaining the original data of uterine cancer patients, professional doctors will evaluate some of the patient's physical data indicators according to their experience according to the numbers 1–9. The larger the number, the more obvious the indicator is. The six different indications are calculated and evaluated by doctors based on their experience. For example, the immunization index is the first test of the patient's data on the immune index by the hospital, and their weight relationship is shown in Table 1:

TABLE 1. Immune indication assessment form.

Immune Index	Weights
CD3 + CD4 + CD8 +/CD45 +	1.0
CD3 + CD4 +/CD45 +	0.8
CD4 +/CD8 +	0.8
CD3 + CD16 + CD56 +/CD45 +	0.5
Exercise ECG	1.5
Sports Galvanic	1.5

CD-x in Table 1 refers to six different indicators including immune index, tumor index, microenvironment index, psychological index, nutritional index, aerobic exercise and advanced work. ECG in Table 1 refers to Electrocardiogram, Exercise ECG refers to ECG during exercise. The calculation formula of the immunity pointer in Table 1 was based on the following formula:

$$I = \frac{\sum_{i=1}^6 x_i w_i}{\sum_{i=1}^6 w_i} \quad (7)$$

Here, x_j refers to the value of the i -th immune index, and refers to the weight of the i -th immune index. The scores ranged from 5 to 37 and were divided into 1–9 different grades. Similarly, the evaluation of microenvironmental indicators was performed (Table 2), which were also rated by professional doctors on a scale of 1–9 based on their experience.

TABLE 2. Microenvironmental indication assessment form.

Microenvironmental Indication	Weights
O ₂	0.5
pH	0.2
Interstitial Pressure	0.2
Inflammatory Response	1.0
Vascular Permeability	0.3
CTCs (circulating tumor cells)	3.0
Protein Profiling	2.0

The CTCs in Table 2 are called circulating tumor cell assays, which are used to detect tumor cells circulating in the peripheral blood. In the experimental data in this paper, after evaluation by professional doctors, the final data indicators used are shown in Table 3.

The data shown in Table 1 was processed as the input data of the model, which was then quantified using the convolutional neural network algorithm. There are several methods for data quantization, such as one-hot encoding, distributed representation, ASCII encoding representation, *etc.* The method of data quantization used in this paper was the distributed representation method, which can effectively display all the features of text data and is beneficial to improving the effects of the neural network during the feature extraction process.

3.3 Evaluation Metrics of the Experiments

During the evaluation process, the model parameters used included precision rate, recall rate, F1 parameter index, false positive rate and true positive rate, which can reflect the actual classification effects of the model. The calculation method was performed using the formula below:

$$A_{avg} = \frac{A_{pos} + A_{neg}}{2} \quad (8)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (9)$$

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (10)$$

TABLE 3. Some raw data of uterine cancer patients used in the experiments.

Patient serial number	immune Indication	tumor indication	Microenvironmental Indications	psychological Indications	nutritional Indications	Aerobic Exercise and Advanced Occupational Indication	In-situ I/Ex-in-situ 2
1	2	5	1	6	3	8	1
2	1	9	1	9	4	8	1
3	1	8	2	8	2	5	2
4	9	3	6	6	7	3	1
...

$$FPR = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FP} + N_{TN} + N_{FN}} \quad (11)$$

$$I(P_{pos}, P_{neg}) = \begin{cases} 1, & P_{pos} > P_{neg} \\ 0.5, & P_{pos} = P_{neg} \\ 0, & P_{pos} < P_{neg} \end{cases} \quad (12)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (13)$$

Here, A_{avg} refers to the average accuracy, A_{pos} refers to the accuracy of positive sample classification, A_{neg} refers to the accuracy of negative sample classification, R refers to the recall rate, N_{TP} refers to the number of false positive samples, P_{pos} refers to the probability of predicting a positive sample, P_{neg} refers to the probability of predicting a negative sample, M refers to the number of positive samples, N refers to the number of negative samples, FPR refers to the false positive rate and can reflect the feasibility of the model results, and lastly, $F1-Score$ refers to the F1 score. A larger value represents a better model performance.

3.4 Experimental Results and Analysis

To verify the actual predicting effects of the postoperative uterine cancer recurrence location model, the pre-set uterine cancer patient dataset was used as the model input. Data preprocessing was performed before the dataset was input into the model to facilitate the experiment and testing process of the convolutional neural network algorithm. The preprocessing work included using a distributed model to quantify the domain name dataset and scrambling the dataset to avoid affecting the training effect of the model due to the uneven distribution of the dataset types.

Next, the random forest algorithm structure was added on the basis of the basic convolutional neural network to improve the original algorithm model. Although the random forest algorithm could achieve certain results in the dataset of uterine cancer patients, it still had some prediction accuracy limitations, such as being prone to overfitting, but can be improved

by adjusting the number of decision trees. However, due to the insufficient number of data samples, the experimental prediction effect of a simple one-dimensional convolutional neural network was not good, and it was difficult for the neural network model to be trained to have a reliable and acceptable accuracy. The various indicators of the experimental results of the random forest algorithm and the one-dimensional convolutional neural network algorithm when processing the dataset in this paper are shown in Fig. 2 and Fig. 3.

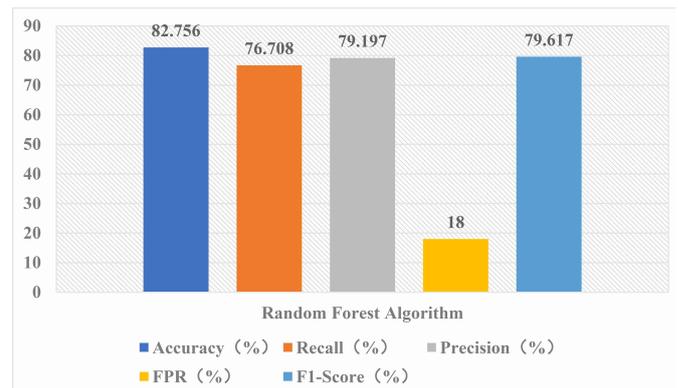


FIGURE 2. Experiment results of random forest algorithm and various index data.

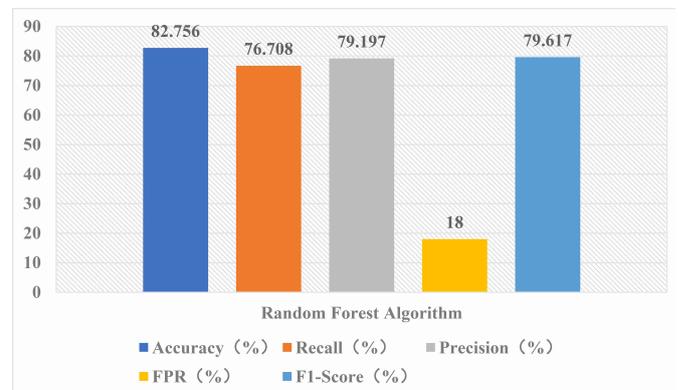


FIGURE 3. The experimental results of the one-dimensional convolutional neural network algorithm.

The results from Fig. 2 and Fig. 3 indicate that the prediction

accuracy of the random forest algorithm was 82.756%, which was higher than 75.947% of the one-dimensional convolutional neural network algorithm, but the recall of the one-dimensional convolutional neural network algorithm. The rate was higher than the random forest algorithm. The comparison of the experimental results of the above two algorithms shows that both have different advantages. The advantage of the random forest algorithm is that its overall performance is relatively stable and is not easily affected by some discrete data values in the dataset. In theory, the random forest algorithm should have a strong anti-overfitting ability. The advantage of the convolutional neural network algorithm is that it can extract the text dataset features better than random forests, but the algorithm effects depend on the support of a large number of data samples. In the model proposed in this paper, the two algorithms were combined, and the random forest algorithm was added to the convolutional neural network to improve the predictive ability of the model. Fig. 4 shows a comparison chart of five different algorithms to further compare the pros and cons of the proposed model.

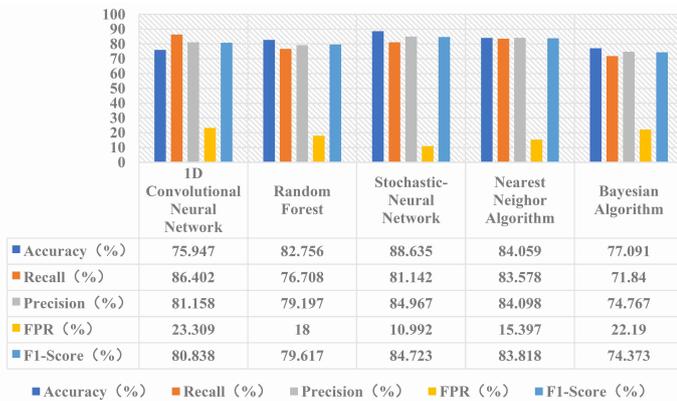


FIGURE 4. The experimental results of the five models and the index data.

The comparative experiments of the five different models in Fig. 4 show that the model constructed in this paper combined the random forest algorithm and the one-dimensional convolutional neural network algorithm with accuracy, precision, and F1-Score. All parameters were improved. Compared with the nearest neighbor algorithm and the Bayesian algorithm, the proposed model demonstrated better effects on various indices, indicating that the random-neural network model could be used in the original. Based on some algorithms, the prediction effect of the model was improved and could better predict the recurrence location of uterine cancer after surgery.

Considering that the small dataset used in this study might lead to deviations in the experimental results, to further improve its performance and prediction effects, a 5-fold cross-validation method was used (Table 4):

1D-CNN in Table 4 refers to one-dimensional convolutional neural network, and Random Forest + CNN refers to the coupled model of random forest and convolutional neural network proposed in this paper. To detect some comprehensive performances of the false alarm rate and recall rate of the model, the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) was determined, with a larger AUC value

referring to better comprehensive performance of the model. To further compare the actual predicting effects of the proposed model, the ROC curves of these five algorithms were compared and shown in Fig. 5.

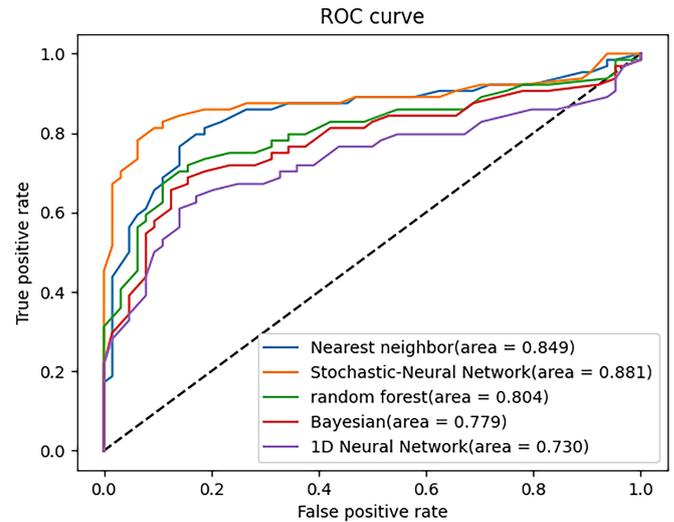


FIGURE 5. ROC curve comparison results of five models.

The ROC in Fig. 5 refers to the receiver operating characteristic curve, which enables a more in-depth comparison of the performance of the algorithmic models. The closer the ROC curve is to the upper left corner, the higher the accuracy of the trial. It can be seen from the ROC curve results in Fig. 5 that the area under the ROC curve AUC of the random-neural network model constructed in this paper is 0.881, which is the best comprehensive performance compared to the other four algorithms. Therefore, the comparison of the ROC curve results of different algorithms further shows that the model established in this paper can better predict the postoperative recurrence position of uterine cancer patients.

4. Conclusion

This paper proposes a prediction model based on combining the random forest algorithm and one-dimensional convolutional neural network algorithm for predicting the location of postoperative uterine cancer recurrence. This proposed model used the feature extraction ability of multiple convolution kernels of the convolutional neural network to extract the features of six types of physical state index data in the quantified uterine cancer dataset and also introduced the random forest algorithm into the convolutional neural network. While improving the classification effect of the model, this also increased the anti-overfitting ability of the model. The experimental results showed that the prediction model for the postoperative recurrence position of uterine cancer patients established in this paper could effectively improve the prediction accuracy of the model based on the original algorithm; however, due to the complexity of the actual treatments and heterogeneity of uterine cancer, more factors that affect the recurrence of uterine cancer need to be further considered.

TABLE 4. Five-fold cross-validation results.

Algorithm Model	Accuracy				
1D CNN	0.7467	0.7632	0.7535	0.7712	0.7588
Random Forest	0.7923	0.8145	0.8312	0.8275	0.8320
Nearest Neighbor	0.8512	0.8411	0.8356	0.8413	0.8405
Bayesian	0.7776	0.7832	0.7709	0.7645	0.7689
Random Forest + CNN	0.8912	0.8863	0.8790	0.8802	0.8689

AUTHOR CONTRIBUTIONS

In this research, the first author, FCL—responsible for organizing the paper and revising the overall framework. The second author, XDH—responsible for the idea of the paper, the overall paper writing, algorithm fusion research, and experimental analysis. The third author, JW—responsible for the operation of the experimental part of the paper, the fourth author and the fifth author, JGQ, DBH—responsible for data analysis and sorting.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was approved by the Shanxi Provincial People's Hospital, Taiyuan, China. All patients provided written informed consent. The study was approved by The Fourth Affiliated Hospital of Hebei Medical University, Shijiazhuang, China. All patients provided written informed consent. The study was approved by the Affiliated Hospital of North China University of Science and Technology, Tangshan, China. All patients provided written informed consent.

The patient dataset used in this study contains patient data collected and collated from the aforementioned hospitals.

The data pre-processing process of this research was carried out in Beijing Stairui Health Technology Co., Ltd. (TIES.IO). The construction of the research model and the realization of the algorithm were completed by the Tangshan Key Laboratory of Engineering Computing (North China University of Science and Technology).

ACKNOWLEDGMENT

This research supported by North China University of Science and Technology Provincial Universities Basic Scientific Research Business Expenses Project (No. JST2022001).

FUNDING

This research received no external funding.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- [1] Xia C, Dong X, Li H, Cao M, Sun D, He S, *et al.* Cancer statistics in China and United States, 2022: profiles, trends, and determinants. *Chinese Medical Journal.* 2022; 135: 584–590.
- [2] Senthilkumar G, Ramakrishnan J, Frnda J, Ramachandran M, Gupta D, Tiwari P, *et al.* Incorporating artificial fish swarm in ensemble classification framework for recurrence prediction of uterine cancer. *IEEE Access.* 2021; 9: 83876–83886.
- [3] Guo C, Wang J, Wang Y, Qu X, Shi Z, Meng Y, *et al.* Novel artificial intelligence machine learning approaches to precisely predict survival and site-specific recurrence in uterine cancer: a multi-institutional study. *Translational Oncology.* 2021; 14: 101032.
- [4] Cibula D, Dostálek L, Jarkovsky J, Mom CH, Lopez A, Falconer H, *et al.* The annual recurrence risk model for tailored surveillance strategy in patients with uterine cancer. *European Journal of Cancer.* 2021; 158: 111–122.
- [5] Huo X, Zhou X, Peng P, Yu M, Zhang Y, Yang J, *et al.* Identification of a six-gene signature for predicting the overall survival of uterine cancer patients. *Oncotargets and Therapy.* 2021; 14: 809–822.
- [6] Okubo M, Itonaga T, Saito T, Shiraishi S, Yunaiyama D, Mikami R, *et al.* Predicting factors for primary uterine cancer recurrence after definitive radiation therapy. *BJR Open.* 2021; 3: 20210050.
- [7] Chai Z, Zhao C. Enhanced random forest with concurrent analysis of static and dynamic nodes for industrial fault classification. *IEEE Transactions on Industrial Informatics.* 2020; 16: 54–66.
- [8] Tohry A, Chehreh Chelgani S, Matin SS, Noormohammadi M. Power-draw prediction by random forest based on operating parameters for an industrial ball mill. *Advanced Powder Technology.* 2020; 31: 967–972.
- [9] Glock A. Explaining a random forest with the difference of two ARIMA models in an industrial fault detection scenario. *Procedia Computer Science.* 2021; 180: 476–481.
- [10] Mylona E, Kourou K, Manikis G, Kondylakis H, Marias K, Karademas E, *et al.* 'Prediction of poor mental health following breast cancer diagnosis using random forests'. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society. 31 October–04 November, 2021. IEEE. 2021.
- [11] Bayramli I, Castro V, Barak-Corren Y, Madsen EM, Nock MK, Smoller JW, *et al.* Temporally informed random forests for suicide risk prediction. *Journal of the American Medical Informatics Association.* 2021; 29: 62–71.
- [12] Yang R, Yu Y. Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. *Frontiers in Oncology.* 2021, 11: 573.
- [13] Ziller A, Usynin D, Braren R, Makowski M, Rueckert D, Kaissis G. Medical imaging deep learning with differential privacy. *Scientific Reports.* 2021; 11: 13524.
- [14] Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, *et al.* Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *Npj Digital Medicine.* 2021; 4: 65.
- [15] Alzubaidi L, Al-Amidie M, Al-Asadi A, Humaidi AJ, Al-Shamma O, Fadhel MA, *et al.* Novel transfer learning approach for medical imaging with limited labeled data. *Cancers.* 2021. 13: 1590.
- [16] Izonin I, Tkachenko R, Dronyuk I, Tkachenko P, Gregus M, Rashkevych M. Predictive modeling based on small data in clinical medicine: RBF-based additive input-doubling method. *Mathematical Biosciences and Engineering.* 2021; 18: 2599–2613.

- [17] Suzdaltseva M, Shamakhova A, Dobrenko N, Alekseeva O, Hammoud J, Gusarova N, *et al.* 'De-identification of medical information for forming multimodal datasets to train neural networks'. Proceedings of the 7th International Conference on Information and Communication Technologies for Ageing Well and E-Health. Science and Technology: Russia. 2021.

How to cite this article: Fengchun Liu, Xiangdong Huang, Jian Wang, Liya Wang, Jingguo Qu, Dianbo Hua. A random forest-neural network coupled model for predicting the recurrence location of uterine cancer. *European Journal of Gynaecological Oncology*. 2022; 43(6): 61-68. doi: 10.22514/ejgo.2022.059.