World Journal of Gastrointestinal Oncology

Submit a Manuscript: https://www.f6publishing.com

World J Gastrointest Oncol 2024 December 15; 16(12): 4663-4674

DOI: 10.4251/wjgo.v16.i12.4663

ISSN 1948-5204 (online)

ORIGINAL ARTICLE

Retrospective Study

Deep learning model combined with computed tomography features to preoperatively predicting the risk stratification of gastrointestinal stromal tumors

Yi Li, Yan-Bei Liu, Xu-Bin Li, Xiao-Nan Cui, Dong-Hua Meng, Cong-Cong Yuan, Zhao-Xiang Ye

Specialty type: Oncology

Provenance and peer review: Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's classification
Scientific Quality: Grade C
Novelty: Grade B
Creativity or Innovation: Grade A
Scientific Significance: Grade B

P-Reviewer: Zheng S

Received: June 25, 2024
Revised: October 2, 2024
Accepted: October 22, 2024
Published online: December 15, 2024
Processing time: 140 Days and 4.9 Hours

Yi Li, Xu-Bin Li, Xiao-Nan Cui, Dong-Hua Meng, Zhao-Xiang Ye, Department of Radiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Tianjin Key Laboratory of Digestive Cancer, Tianjin 300060, China

Yan-Bei Liu, School of Life Sciences, Tiangong University, Tianjin 300387, China

Cong-Cong Yuan, Department of Radiology, Tianjin First Central Hospital, Tianjin 300190, China

Co-corresponding authors: Zhao-Xiang Ye and Xu-Bin Li.

Corresponding author: Zhao-Xiang Ye, PhD, Chief, Chief Doctor, Professor, Department of radiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Tianjin Key Laboratory of Digestive Cancer, Huanhuxi Road, Hexi District, Tianjin 300060, China.
yezhaoxiang@163.com

Abstract

BACKGROUND

Gastrointestinal stromal tumors (GIST) are prevalent neoplasm originating from the gastrointestinal mesenchyme. Approximately 50% of GIST patients experience tumor recurrence within 5 years. Thus, there is a pressing need to accurately evaluate risk stratification preoperatively.

AIM

To assess the application of a deep learning model (DLM) combined with computed tomography features for predicting risk stratification of GISTs.

METHODS

Preoperative contrast-enhanced computed tomography (CECT) images of 551 GIST patients were retrospectively analyzed. All image features were independently analyzed by two radiologists. Quantitative parameters were statistically analyzed to identify significant predictors of high-risk malignancy. Patients were

randomly assigned to the training (n = 386) and validation cohorts (n = 165). A DLM and a combined DLM were established for predicting the GIST risk stratification using convolutional neural network and subsequently evaluated in the validation cohort.

RESULTS

Among the analyzed CECT image features, tumor size, ulceration, and enlarged feeding vessels were identified as significant risk predictors (P < 0.05). In DLM, the overall area under the receiver operating characteristic curve (AUROC) was 0.88, with the accuracy (ACC) and AUROCs for each stratification being 87% and 0.96 for low-risk, 79% and 0.74 for intermediate-risk, and 84% and 0.90 for high-risk, respectively. The overall ACC and AUROC were 84% and 0.94 in the combined model. The ACC and AUROCs for each risk stratification were 92% and 0.97 for low-risk, 87% and 0.83 for intermediate-risk, and 90% and 0.96 for high-risk, respectively. Differences in AUROCs for each risk stratification between the two models were significant (P < 0.05).

CONCLUSION

A combined DLM with satisfactory performance for preoperatively predicting GIST stratifications was developed using routine computed tomography data, demonstrating superiority compared to DLM.

Key Words: Gastrointestinal stromal tumors; Deep learning; Risk stratification; Tomography, X-ray computed; Prognosis

Core Tip: The deep learning model (DLM) was validated to accurately predict the risk classification of gastrointestinal stromal tumors. The combined DLM outperformed DLM in predicting risk stratification. The combined model has potential to guide and facilitate clinical decision-making.

Citation: Li Y, Liu YB, Li XB, Cui XN, Meng DH, Yuan CC, Ye ZX. Deep learning model combined with computed tomography features to preoperatively predicting the risk stratification of gastrointestinal stromal tumors. *World J Gastrointest Oncol* 2024; 16(12): 4663-4674

URL: https://www.wjgnet.com/1948-5204/full/v16/i12/4663.htm

DOI: https://dx.doi.org/10.4251/wjgo.v16.i12.4663

INTRODUCTION

As is well documented, gastrointestinal stromal tumors (GIST) are prevalent tumors originating from the gastrointestinal mesenchyme. Approximately 50% of GIST patients experience tumor recurrence within 5 years, and surgical intervention alone is inadequate for achieving an optimal prognosis[1,2]. Earlier studies established that 85% of GISTs are associated with receptor tyrosine kinase (c-kit) mutations, and 3%–5% are linked to platelet derived growth factor receptor alpha gene mutation[1,3-5]. The mutation type determines response to imatinib[6-8]. Considering the lack of effective conventional chemotherapy drugs, and the activity of Imatinib, adjuvant Imatinib may be a potential treatment option for patients with GIST[1,5,9]. A study assessing the efficacy of long-term imatinib treatment in advanced GIST patients documented a median overall survival exceeding four years[10]. Moreover, considering the recurrence and metastasis rate of intermediate- and high-risk GISTs after surgery, targeted drug therapy can improve the prognosis of GISTs[11,12]. Therefore, there is an urgent need to accurately evaluate risk stratification prior to surgery to obtain valuable information for evaluating the necessity of surgery and adjuvant treatment.

At present, the risk stratification of GISTs is based on histologic (mitotic index) and imaging characteristics (including tumor size and site) of the lesion as outlined in the National Institutes of Health (NIH) consensus classification system [13]. However, it is challenging to determine the mitotic index without histological examination. Indeed, the malignancy risk of most tumors is confirmed histologically after surgery. Although, endoscopic biopsy has also been widely used preoperatively, its utility may be limited if the tumor sample contains large areas of necrosis or hemorrhage, yielding inconclusive results[14-16].

Recently, the deep learning model (DLM), composed of multi-types of self-learning units, has emerged as a promising technique for analyzing medical imaging data[17-20]. Notably, DLM has demonstrated efficacy in clinical applications such as the assessment of differentiation grades in meningioma and renal cell carcinoma, as well as in predicting the molecular subtypes and grades of glioma[17-19,21]. Overall, deep learning transforms medical images into high-dimensional mineable data, offering rapid insights with high repeatability and providing a novel approach for GIST risk assessment[22-24]. This study aimed to establish and validate DLMs for predicting preoperative GIST risk stratification based on routine post-contrast computed tomography (CT) and clinical data.

MATERIALS AND METHODS

Characteristics of patients

This study was approved by our institutional ethics committee. All the patients signed the informed consent form before examinations. From January 2012 to December 2022, 606 patients with GISTs were initially enrolled in this retrospective study. A total of 55 patients were subsequently excluded for the following reasons: (1) Lack of preoperative contrast-enhanced CT images (n = 19); (2) Suboptimal CT image quality (n = 9); (3) Preoperative therapy or experienced metastasis to other sites (n = 25); and (4) Absence of required pathologic data (n = 2). Finally, a total of 551 patients were included in this study (256 men and 295 women; mean age 60.3 ± 9.8 years). The study population flow chart is illustrated in Figure 1.

All patients underwent complete surgical resection. GIST risk classification was based on National Comprehensive Cancer Network (NCCN) guidelines[15]. According to risk categories, patients in this study were classified into the high-risk (high risk), intermediate-risk (intermediate risk), and low-risk (very low and low risk) groups.

CT image acquisition

All patients underwent abdominal contrast-enhanced CT examination covering the whole tumor. After a non-contrast CT scan (Scanner: Philips Iqon, GE Healthcare Discovery CT750 HD or SIEMENS 64-MDCT) with a thickness of 1.0-1.5 mm, three-phase contrast-enhanced scans were performed, with 90 to 100 mL iodine contrast medium (Ultravist 370, Bayer Schering Pharma, Germany) intravenously injected at a rate of 2.5 to 3.0 mL/s.

Clinical and CT image feature analysis

All CT images were independently analyzed by two radiologists with 3 and 13 years of experience in abdominal radiology. In cases of disagreement, the radiologist with 13 years of experience reviewed the images to reach a consensus. Clinical information, image features and pathologic characteristics, including gender, age, tumor location, growth pattern (exophytic, endoluminal, mixed), tumor size (measured as the maximum diameter of the largest tumor section), tumor morphology (round or oval shape was considered regular, and lobulated or other irregular shapes were categorized as irregular), necrosis, ulceration, internal hemorrhagic foci, calcification, lymph node status, presence of enlarged feeding vessels, tumor boundary (clear or blurred), the pattern and degree of tumor enhancement during the venous phase and the range of tumor enhancement across the three phases, were derived from CT images and medical records. For CT value measurements of each tumor, regions of interest (ROIs) were delineated to cover tumor parenchyma while avoiding areas with evidence of cystic, necrotic or hemorrhagic changes at the level of the largest solid tumor regions and their adjacent upper and lower levels during the plain phase, arterial phase, venous phase and delay phase, respectively. The ROIs for CT value measurements were consistently sized using the copy and paste function across the different phases of images. Next, the average of three measurements was calculated. According to the difference between the venous phase and plain CT, the enhancement degree was defined as mild (CT value difference ≤ 20 HU), moderate (CT value difference between 20 HU and 40 HU), and obvious enhancement (CT value difference > 40 Hu). According to differences between the CT values of the venous and arterial phases, the enhancement pattern was defined as continuous (CT value difference ≥ 0) and attenuation (CT value difference < 0). The enhancement tumor range was calculated during the arterial phase (ER1 = arterial phase CT value-precontrast CT value), venous phase (ER2 = venous phase CT value-precontrast phase CT value), and delay phase (ER3 = delay phase CT value - precontrast scan CT value).

Image segmentation

All the CT images were exported in Joint Photographic Experts Group format. Then, two radiologists with extensive experience in abdominal imaging diagnosis (3 years and 13 years, respectively) participated in the segmentation of the entire tumor. One radiologist manually delineated the ROIs of the entire tumor layer by layer on venous phase CT images. The segmented images were subsequently confirmed by the other radiologist. Both radiologists were blinded to the pathological reports for risk stratification. Based on recommendations from previous literature[21], ImageJ (NIH, Bethesda, MD) was employed to apply an adaptive contrast filter to images. Besides, CT- segmented images were randomly selected from 20 patients, and the Dice similarity coefficient (DSC) was used to evaluate the inter-reader consistency in image segmentation. Detailed information of image preprocessing can be found in the Supplementary material 1.

The DLM construction

The DLM was constructed in two steps: Tumor features and tumor classification were initially extracted from CT images to generate the DLM, followed by the establishment of the combined model for tumor classification by integrating subject clinical-imaging features after statistical analysis. Figure 2 displays the detailed framework of this process. In the current study, a stratified random split was utilized at the patient level to randomly divide all patients into a training cohort and a validation cohort in a 7:3 ratio.

The 3D residual network (ResNet) was used to train our image dataset and establish the DLM. The 3D-ResNet is a three-dimensional convolutional neural network model based on the ResNet architecture. It is an extension of ResNet in two-dimensional image classification tasks used to process three-dimensional volume data. Besides, it accepts 3D volumetric data as input, positioning it as a powerful model for learning volume datasets. Furthermore, it can be adjusted according to the task complexity and dataset characteristics with variable depth. It can enhance the network depth by increasing the number of stacked layers of residual blocks, thereby optimizing the model's expressive ability. During the training phase, 3D-Resnet was used to extract and learn deep tumor features related to GIST risk stratification from each patient's CT images.

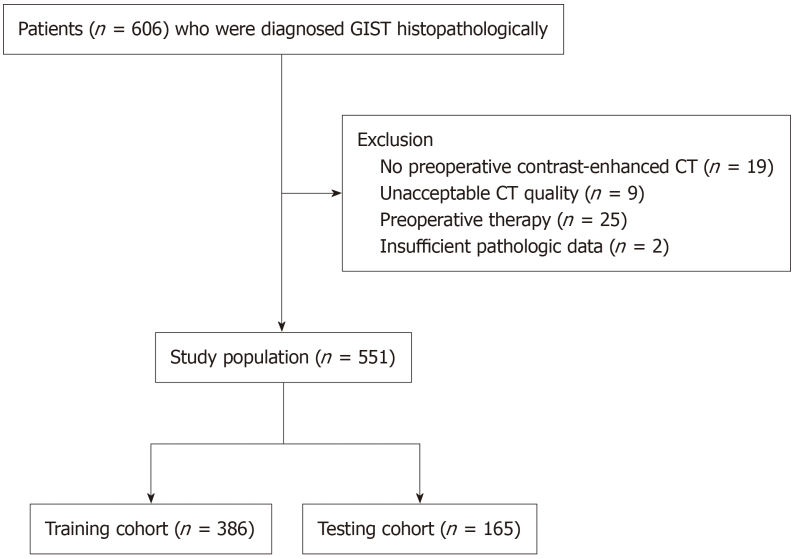

Figure 1 Study population flowchart. CT: Computed tomography; GIST: Gastrointestinal stromal tumors.

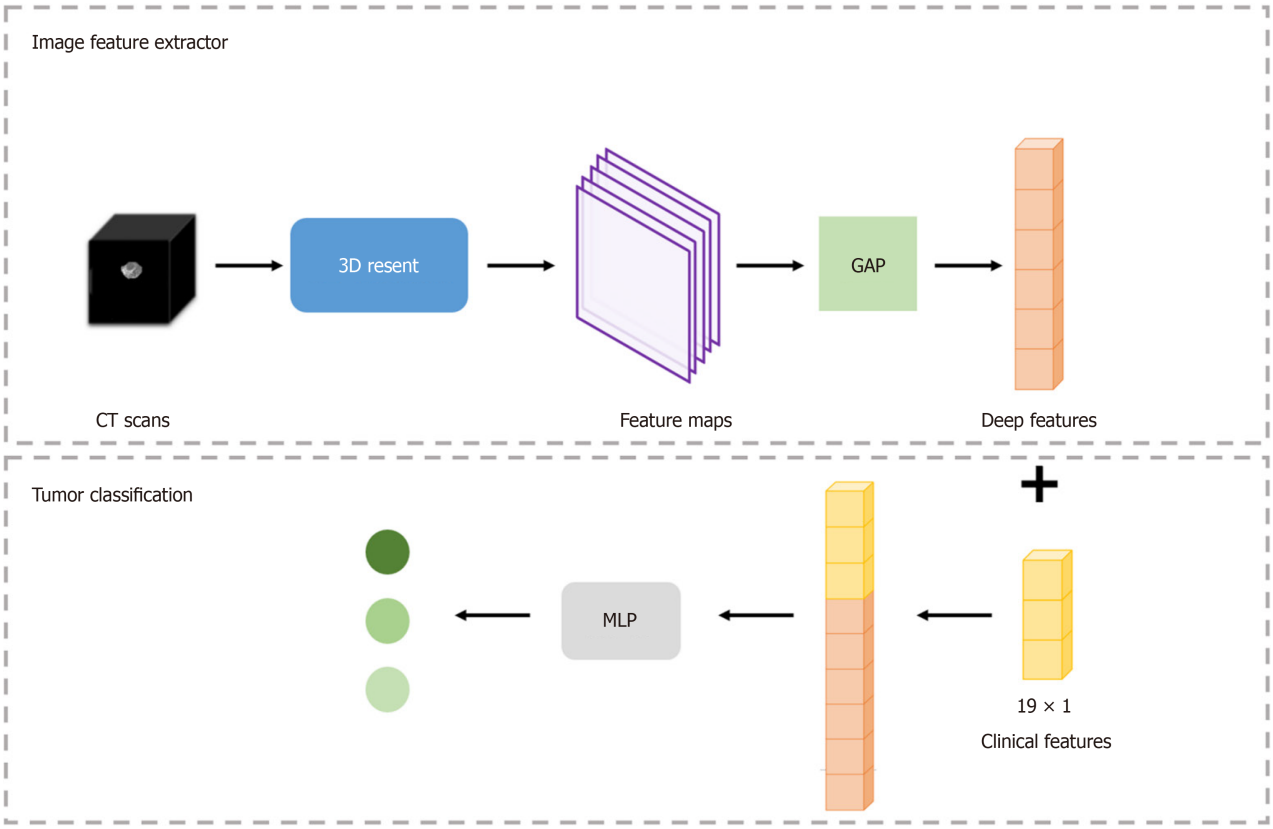

Figure 2 Overall gastrointestinal stromal tumor risk stratification framework. CT: Computed tomography; MLP: Multilayer perceptron; GAP: Global average pooling.

Multilayer perceptron (MLP) was employed for tumor risk stratification by combining imaging features and clinical data. It is a universal function approximator based on feedforward artificial neural networks that can learn and represent nonlinear relationships through multiple fully connected hidden layers and an output layer, making it suitable for various machine learning tasks. In the current model, a batch normalization (BN) layer was introduced following each linear layer in MLP to accelerate the convergence rate of the neural network, reduce the dependence of the model on the initial parameters, and improve the robustness of the model. The BN layer normalized each mini-batch data to stabilize the input of the neural network, thus improving the convergence rate and generalizability of the model. Figure 3 shows the feature extraction process. In the present study, the subject clinical-image data features of patients were concatenated with imaging features extracted by the 3D ResNet, and their feature vectors were inputted into the MLP to establish the

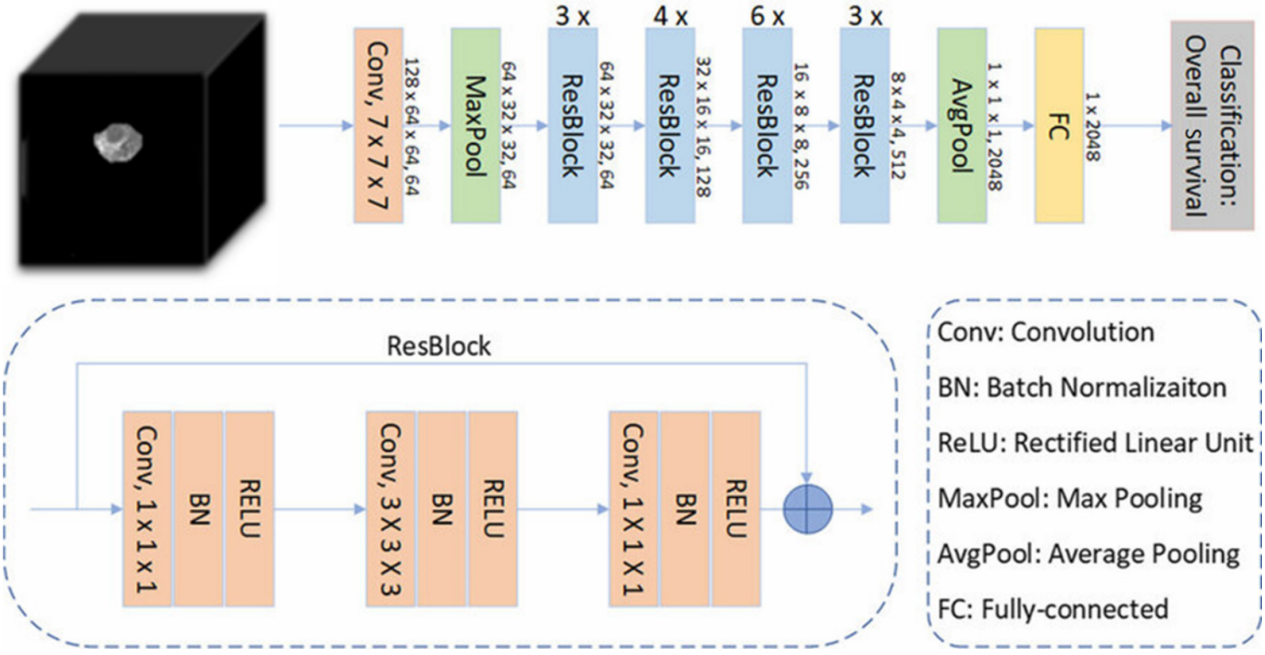

Figure 3 Schematic diagram of feature extraction process. BN: Batch normalization; RELU: Rectified linear unit.

combined DLM for tumor risk stratification.

Training details

In this study, the network architecture was implemented using the Pytorch framework and trained on NVIDIA GPU to accelerate training speed. Additionally, the transfer learning method was adopted, leveraging pre-training weights to improve model performance. Detailed information of model training process can be found in the Supplementary materials 2.

Statistical analysis

Statistical analysis was performed using the software SPSS version 22.0 (SPSS Inc., Chicago, IL, United States) and MedCalc version 22.002 (MedCalc Software Ltd, Ostend, Belgium). A P value of < 0.05 was considered statistically significant.

The interobserver agreement of measurements between the two radiologists was evaluated using the interclass correlation coefficient (ICC). The χ^2 test, independent two-sample t-test, one-way analysis of variance and Bonferroni tests were used to evaluate the significance of correlations between various clinical- imaging features and pathological GIST risk classifications from surgical resection specimens. Ordinary logistic regression was performed to identify significant predictive factors for relapse[25,26].

To assess the performance of DLM and combined DLM, five different indicators, namely area under the receiver operating characteristic curve (AUROC), F1 score (F1), accuracy (ACC), sensitivity (SEN), and specificity (SPE), were used. AUROC was calculated along with its 95%CI.

RESULTS

Clinical-imaging characteristics

ICC analysis showed a good concordance of measurements between the two radiologists (tumor size, ICC = 0.985; CT value in the plain phase, ICC = 0.812; CT value in the arterial phase, ICC = 0.906; CT value in the venous phase, ICC = 0.921; CT value in the delay phase, ICC = 0.848). Among the analyzed CECT image features, tumor size, tumor morphology, tumor location, growth pattern, necrosis, ulceration, calcification, lymph node status, presence of enlarged feeding vessels, tumor enhancement pattern during the venous phase, and the range of tumor enhancement across the three phases were found to be significantly associated with GIST risk categories ($P < 0.05$). The distribution of these features in the risk categories and the results of the χ^2 test are listed in Table 1. Meanwhile, ordinary logistic regression analysis identified tumor size, ulceration, and the presence of enlarged feeding vessels as statistically significant predictors ($P < 0.05$, Table 2).

Diagnostic performance of the DLM

All patients were randomly associated into two independent cohorts, namely a training cohort (386 patients: 176 males, mean age, 60.2 ± 9.8 years; 210 females, mean age, 59.9 ± 10.2 years) and a validation cohort (165 patients: 80 males, mean

Table 1 Distribution of the analyzed clinical-imaging features across pathologic risk categories, mean ± SD

CECT features		Pathologic risk categories			*P* value
		High (*n* = 213)	Moderated (*n* = 143)	Low (*n* = 195)	
Gender	Male	107	60	89	0.296
	Female	106	83	106	
Location	Gastric	139	136	171	0.000
	Non-gastric	74	7	24	
Morphology	Regular	54	94	170	0.000
	Irregular	159	49	25	
Growth pattern	Endoluminal	35	54	93	0.000
	Mixed	51	22	28	
	Exophytic	127	67	74	
Degree of contrast enhancement in the venous phase	Mild (≤ 20 HU)	30	20	20	0.424
	Moderate (20-40 HU)	101	74	90	
	Obvious (≥ 40 HU)	82	49	85	
Contrast enhancement pattern during the venous phase	Continuous	198	142	193	0.000
	Attenuation	15	1	2	
Calcification	Present	22	16	21	0.967
	Absent	191	127	174	
Necrosis	Present	155	65	33	0.000
	Absent	58	78	162	
Ulceration	Present	53	23	8	0.000
	Absent	160	120	187	
Enlarged feeding vessels	Present	183	60	14	0.000
	Absent	30	83	181	
Lymph nodes	Present	19	1	1	0.000
	Absent	194	142	194	
Age		59.44 ± 10.41	61.39 ± 9.83	60.37 ± 9.13	0.183
Size		9.03 ± 4.42	4.90 ± 1.89	2.77 ± 1.20	0.000
Range of tumor enhancement during the arterial phase		18.83 ± 17.81	14.24 ± 11.53	18.45 ± 17.47	0.004
Range of tumor enhancement during the venous phase		41.59 ± 25.73	38.97 ± 18.96	45.22 ± 25.70	0.038
Range of tumor enhancement during the delay phase		42.73 ± 18.89	43.72 ± 17.97	47.97 ± 20.27	0.016

CECT: Contrast-enhanced computed tomography.

age, 59.8 ± 9.9 years; 85 females, mean age, 60.0 ± 10.1 years). There were 136 (35.2%) cases of low-risk GISTs, 101 (26.2%) cases of intermediate-risk GISTs, and 149 (38.6%) cases of high-risk GISTs in the training cohort. In contrast, the validation cohort comprised 59 (35.8%) cases of low-risk GISTs, 42 (25.5%) cases of intermediate-risk GISTs, and 64 (38.8%) cases of high-risk GISTs.

The DSC value showed a good concordance of image segmentation between the two radiologists (DSC = 99.96%). The results for the different algorithms are detailed in Table 3. The model constructed using 3D-ResNet with 34 Layers demonstrated the optimal performance, with an ACC of 75%, a SEN of 72%, a SPE of 87%, and a F1 score of 72%. The overall AUROC for DLM was 0.88 (0.83, 0.93). The ROCs are depicted in Figure 4A. In DLM, the ACC and AUROCs for each stratification were 87% (144/165) and 0.96 (0.94, 0.98) for low-risk GISTs, 79% (131/165) and 0.74 (0.67, 0.81) for intermediate-risk GISTs, and 84% (138/165) and 0.90 (0.85, 0.95) for high-risk GISTs, respectively. The results for the validation cohorts were visualized as confusion matrices to compare the GIST risk stratification predicted by DLM against the pathological risk stratification (Figure 5A).

Table 2 Logistic regression analysis of risk classification based on clinical-imaging feature

	P value	95%CI	
		Lower bound	Upper bound
Size	0.000	-0.763	-0.473
Range of tumor enhancement during the arterial phase	0.131	-0.035	0.005
Range of tumor enhancement during the venous phase	0.220	-0.007	0.032
Range of tumor enhancement during the delay phase	0.858	-0.021	0.017
Morphology	0.602	-0.387	0.608
Location	0.074	-0.063	1.386
Ulceration	0.004	-1.622	-0.300
Enlarged feeding vessels	0.000	-2.134	-1.094
Growth pattern	0.224	-0.833	0.328
Contrast enhancement during the venous phase	0.428	-2.266	1.384
Necrosis	0.236	-0.195	0,793
Lymph nodes	0.890	-1.934	1.678

Table 3 Different algorithms for predicting gastrointestinal stromal tumor risk classification

Different method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 score (%)
3DCNN	60	51	68	52
3DResnet_50	66	58	71	59
3DResnet_18	71	66	76	67
3DResnet_34	75	72	87	72
Combined model (3DResnet + MLP)	84	83	92	83

3DCNN: Three-dimensional convolutional neural network; 3DResnet: Three-dimensional residual network; MLP: Multilayer perceptron.

The combined DLM achieved satisfactory performance in assessing GIST risk stratification. The overall ACC and AUROC were 84% (139/165) and 0.94 (0.93, 0.95) for the combined model. The ROCs are delineated in Figure 4B. The ACC and AUROCs for each tumor risk stratification were 92% (152/165) and 0.97 (0.96, 0.98) for low-risk GISTs, 87% (143/165) and 0.83 (0.78, 0.88) for intermediate-risk GISTs, and 90% (148/165) and 0.96 (0.94, 0.98) for high-risk GISTs, respectively. The results for the validation cohorts were visualized as confusion matrices to compare the GIST risk stratification predicted by the combined model against pathological risk stratification (Figure 5B).

The ACC, SEN, SPE, F1 score and AUROCs for each tumor risk stratification across different models are summarized in Table 4. Importantly, differences in AUROCs between DLM and the combined model were significant ($P < 0.001$).

DISCUSSION

In this retrospective research, a DLM and a combined DLM were constructed. Notably, the latter (AUROC = 0.94) outperformed the former (AUROC = 0.88) in assessing GIST grading.

According to the modified NIH criteria and NCCN guidelines, the need of adjuvant treatment for GIST patients and the duration of treatment are associated with the risk stratification of GISTs[8,15,27,28]. Combining adjuvant treatment such as Imatinib before and after surgery may extend the recurrence free survival and overall survival of intermediate and high-risk GIST patients[1,5,29]. Therefore, an accurate preoperative categorization of risk classification, especially in high-risk GISTs, can provide valuable information for evaluating the necessity of surgical resection and adjuvant treatments[30-32]. In this study, two models were developed to predict preoperative GIST risk stratification: DLM and combined DLM. To the best of our knowledge, studies that combined clinical-imaging features and convolutional neural

Table 4 Accuracy, sensitivity, specificity, F1 score and areas under the receiver operating characteristic curves for each tumor risk stratification, *n* (%)/95%CI

		Accuracy (*n* = 165)	Sensitivity	Specificity	F1 score (%)	AUROC
DLM	High	138 (84); (78-90)	81 (52/64); (76-86)	85 (86/101); (78-91)	79	0.90 (85-95)
	Moderate	131 (79); (71-87)	50 (21/42); (25-74)	89 (110/123); (82-96)	55	0.74 (67-81)
	Low	144 (87); (81-93)	86 (51/59); (81-91)	88 (93/106); (83-94)	83	0.96 (94-98)
	Overall	75	72	87	72	0.88 (83-93)
Combined model	High	148 (90); (86-94)	88 (56/64); (83-93)	91 (92/101); (83-98)	87	0.96 (94-98)
	Moderate	143 (87); (83-92)	69 (29/42); (66-71)	93 (114/123); (86-98)	72	0.83 (78-88)
	Low	152 (92); (85-96)	92 (54/59); (89-93)	92 (98/106); (86-97)	89	0.97 (96-98)
	Overall	84	83	92	83	0.94 (93-95)

DLM: Deep learning model; AUROC: Areas under the receiver operating characteristic curve.

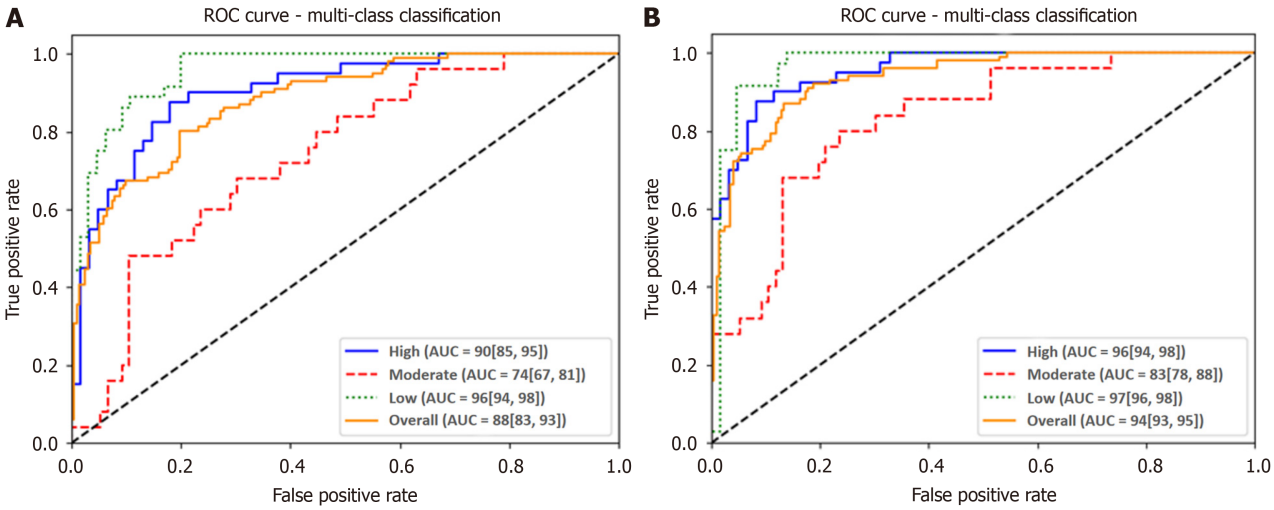

Figure 4 Receiver operating characteristic curves for the deep learning model and combined model. A: Receiver operating characteristic curve (ROC) for the deep learning model in the validation cohort; B: ROC curve for the combined model in the validation cohort. ROC: Receiver operating characteristic; AUC: Area under the curve.

network models to establish a combined model for predicting GIST risk stratification and distinguishing between different categorizations of risk classification (high-risk, intermediate-risk and low-risk GISTs) are scarce.

Analysis of clinical-imaging features revealed that tumor size, morphology, location, growth pattern, the presence of necrosis, ulceration, calcification, lymph nodes, and enlarged feeding vessels, as well as the tumor enhancement pattern during the venous phase and the range of tumor enhancement, were significantly associated with pathologic GIST risk categories. Logistic regression analysis subsequently identified tumor size, the presence of ulcers, and enlarged feeding vessels as predictors of pathologic risk categories, consistent with the results of previous studies[25,26,33]. Zhou *et al*[34] reported that tumors with large sizes (> 10 cm) and enlarged feeding vessels were more likely to be a high -risk GISTs [34]. Moreover, mucosal destruction promotes the formation of ulcers due to the influence of gastric acid[35]. The NCCN guideline recommend patients with GISTs larger than 2 cm to undergo surgical resection[28], while according to Nishida's[36] report, small GIST tumors may also be invasive and linked to a poor prognosis. Therefore, evaluating the risk stratification of GISTs exclusively based on tumor size could be insufficient. Other imaging features were also subjectively assessed and heavily relied on the experience of observers. While the degree of contrast enhancement is typically considered a characteristic of tumor biological activity, it showed no significant association with pathologic risk stratification as a predictive factor in our study, in line with the results of previous articles[26,33].

The results of our study unveiled that the DLM could accurately predict GIST risk classifications, with an AUROC of 0.88 in the validation cohort. However, the performance of the combined DLM was relatively higher (AUROC = 0.94), attributable to the combination of DLM with clinical-imaging features increasing the ability to assess the GIST risk classification. Overall, our study offers a novel method for optimizing the preoperative assessment of GIST risk stratification based on CT images, moving beyond dependence on postoperative specimens. Zhou *et al*[34] documented that the AUROC of the multinomial logistic regression model with subjective CT image features for GIST risk stratification was 0.806. At the same time, Wang *et al*[22] divided patients with GISTs into the high malignant potential group (intermediate

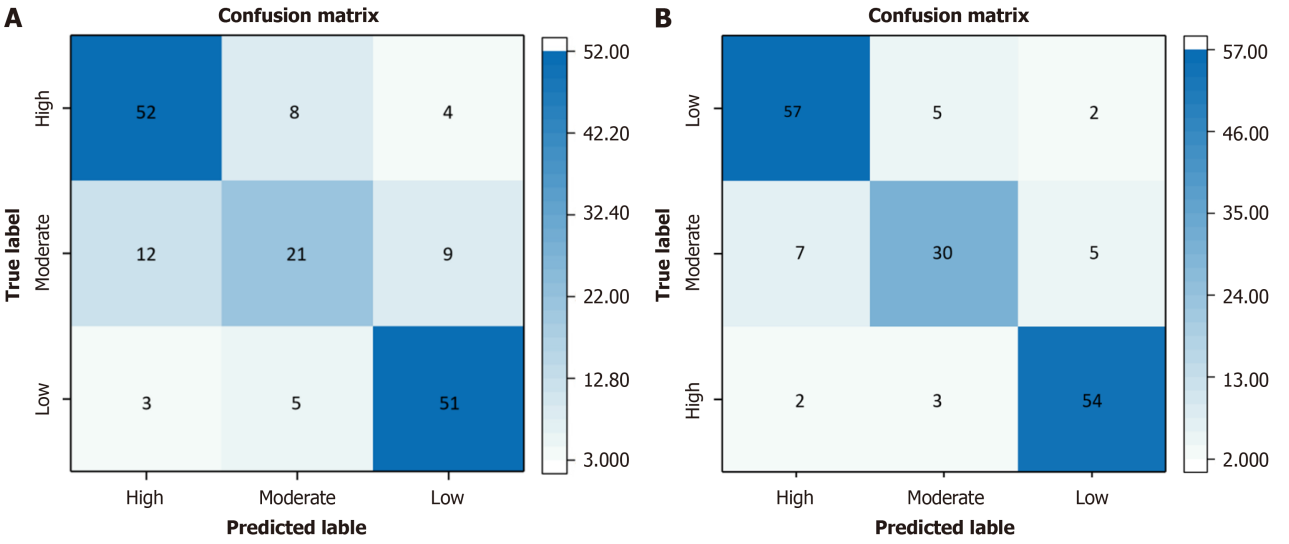

Figure 5 Confusion matrix for the deep learning model and combined model for gastrointestinal stromal tumors risk stratification. A: Confusion matrix for the deep learning model; B: Confusion matrix for the combined model.

risk and high risk) and the low malignancy potential group (very low risk and low risk) and described that the area under the curve (AUC) of the combined model (clinical features and the radiomics) was significantly higher in the validation group (0.913 *vs* 0.792, P = 0.019) compared to the clinical model. Importantly, Kang *et al*[37] concluded that the DLM (AUROCs; testing, 0.89; external validation, 0.85) outperformed the radiomics model in terms of GIST risk classification. In this study, the DLM and the combined DLM had an AUROC of 0.88 and 0.94 for distinguishing between the three types of GIST risk classifications (high-risk, intermediate-risk and low-risk) in the validation cohort. Indeed, the combined model outperformed DLM, with ACCs and AUROCs of 92% and 0.97 for low-risk GISTs, 87% and 0.83 for intermediate-risk GISTs, and 90% and 0.96 for high-risk GISTs, respectively. Notably, the AUROCs for different risk stratifications in the combined model were significantly superior to those of the DLM. These results collectively indicated that the combined DLM had superior predictive capabilities, especially for low- and high-risk GISTs.

However, the results for the intermediate-risk class were relatively unsatisfactory, with an accuracy of 87% and an AUROC of 0.83. This may be ascribed to imbalanced sample sizes. Specifically, there were only 143 cases of intermediate-risk GISTs in this study, which was lower than those of non-intermediate-risk GISTs (408 cases). This may result in relatively fewer features being extracted from intermediate-risk GISTs compared to machine learning algorithms, thereby introducing model bias. Nevertheless, compared to DLM (AUC = 0.74), the combined model (AUC = 0.83) demonstrated advantages in predicting intermediate-risk GISTs.

Of note, deep learning is a subfield of artificial intelligence that performs tasks by analyzing relationships between existing data points[38-40]. In recent years, image analysis based on deep learning algorithm has been increasingly applied to tumor diagnosis, grading, staging, prediction, and treatment evaluation. Zhu *et al*[41] concluded that the DLM outperformed in assessing the risk of screening-detected breast cancer. Similarly, Doppalapudi *et al*[42] pointed out that the accuracy of lung cancer classification predicted by DLM was 71.18%, while that of traditional machine learning models was merely 61.12%, indicating that DLM displayed superior performance for predicting lung cancer subtypes. A study investigating glioma showed that DLM achieved high performance in predicting molecular subtypes and grades, with an isocitrate dehydrogenase-AUC of 0.90, an 1p/19q co-deletion AUC of 0.85, and a grade AUC of 0.81 (grade II/III/IV)[17]. Wang *et al*[43] developed the convolutional neural network models with varying layers, achieving AUROCs above 0.8 for differentiating high-risk gastric GISTs from intermediate-risk and very low/low-risk gastric GISTs in the validation dataset. In the present study, the DLM based on the 3D-ResNet method increased the network depth by increasing the number of stacked layers of residual blocks, thereby improving the model's expressive ability. In addition, the clinical data of patients were concatenated with the imaging features extracted by the 3D- ResNet and then incorporated their feature vectors into the MLP for risk classification. As anticipated, the results uncovered that the DLM based on the 3D-ResNet method combined with clinical-imaging features could accurately predict GIST risk classifications.

Nevertheless, this study has several limitations that cannot be overlooked. Firstly, this was a retrospective study based on a limited sample size, resulting in an imbalance in the data for risk stratification. Therefore, an ideal DLM should be constructed with a larger training set containing datasets from multiple-centers to balance the data for different risk stratifications. Further prospective studies with external validation cohorts are warranted to validate our results. Secondly, the DLM developed in our study required manual segmentation of tumors on CT images remains a semi-automatic model. Thirdly, radiomics features were used for the risk stratification of GISTs in previous studies[22,32,39,44,45]. Therefore, future studies can compare the performance of radiomics models with DLM.

CONCLUSION

In summary, a high-performance combined DLM for preoperative prediction of the GIST risk stratification was developed and validated in this study. Noteworthily, this model has the potential to guide and facilitate clinical decision-making for GIST patients.

ACKNOWLEDGEMENTS

We are grateful to Yan-Bei Liu and Cong-Cong Yuan for their supports. We would also like to thank Yan-Bei Liu for his scientific advice and intellectual discussions.

FOOTNOTES

Author contributions: Li Y completed the research design and data analysis, and was a major contributor in writing the manuscript; Liu YB completed data analysis and model establishment; Cui XN, Yuan CC, and Meng DH participated in data collection and preliminary reports; Li XB and Ye ZX provided guidance on the paper and objectively reviewed it; Li Y and Cui XN were responsible for the processing of inspection analysis and statistical data to ensure the quality of the data. All authors read and approved the final manuscript.

Supported by The Chinese National Key Research and Development Project, No. 2021YFC2500400 and No. 2021YFC2500402; and Tianjin Key Medical Discipline (Specialty) Construction Project, No. TJYXZDXK-009A.

Institutional review board statement: The study was reviewed and approved by Tianjin Medical University Cancer Institute and Hospital (EK20240015).

Informed consent statement: All study participants, or their legal guardian, provided informed written consent prior to study enrollment.

Conflict-of-interest statement: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data sharing statement: Technical appendix, statistical code, and dataset available from the corresponding author at yezhaoxiang@163.com.

Country of origin: China

ORCID number: Xu-Bin Li 0000-0002-9888-4267; Zhao-Xiang Ye 0000-0003-3157-8393.

S-Editor: Liu H
L-Editor: A
P-Editor: Xu ZH

REFERENCES

1 **Dematteo RP**, Ballman KV, Antonescu CR, Maki RG, Pisters PW, Demetri GD, Blackstein ME, Blanke CD, von Mehren M, Brennan MF, Patel S, McCarter MD, Polikoff JA, Tan BR, Owzar K; American College of Surgeons Oncology Group (ACOSOG) Intergroup Adjuvant GIST Study Team. Adjuvant imatinib mesylate after resection of localised, primary gastrointestinal stromal tumour: a randomised, double-blind, placebo-controlled trial. *Lancet* 2009; **373**: 1097-1104 [PMID: 19303137 DOI: 10.1016/S0140-6736(09)60500-6]

2 **DeMatteo RP**, Lewis JJ, Leung D, Mudan SS, Woodruff JM, Brennan MF. Two hundred gastrointestinal stromal tumors: recurrence patterns and prognostic factors for survival. *Ann Surg* 2000; **231**: 51-58 [PMID: 10636102 DOI: 10.1097/00000658-200001000-00008]

3 **Yang J**, Chen Z, Liu W, Wang X, Ma S, Jin F, Wang X. Development of a Malignancy Potential Binary Prediction Model Based on Deep Learning for the Mitotic Count of Local Primary Gastrointestinal Stromal Tumors. *Korean J Radiol* 2021; **22**: 344-353 [PMID: 33169545 DOI: 10.3348/kjr.2019.0851]

4 **Wang Y**, Call J. Mutational Testing in Gastrointestinal Stromal Tumor. *Curr Cancer Drug Targets* 2019; **19**: 688-697 [PMID: 30914028 DOI: 10.2174/1568009619666190326123945]

5 **Wozniak A**, Gebreyohannes YK, Debiec-Rychter M, Schöffski P. New targets and therapies for gastrointestinal stromal tumors. *Expert Rev Anticancer Ther* 2017; **17**: 1117-1129 [PMID: 29110548 DOI: 10.1080/14737140.2017.1400386]

6 **Gastrointestinal Stromal Tumor Meta-Analysis Group (MetaGIST)**. Comparison of two doses of imatinib for the treatment of unresectable

or metastatic gastrointestinal stromal tumors: a meta-analysis of 1,640 patients. *J Clin Oncol* 2010; **28**: 1247-1253 [PMID: 20124181 DOI: 10.1200/JCO.2009.24.2099]

7 **Patrikidou A**, Domont J, Chabaud S, Ray-Coquard I, Coindre JM, Bui-Nguyen B, Adenis A, Rios M, Bertucci F, Duffaud F, Chevreau C, Cupissol D, Pérol D, Emile JF, Blay JY, Le Cesne A; French Sarcoma Group. Long-term outcome of molecular subgroups of GIST patients treated with standard-dose imatinib in the BFR14 trial of the French Sarcoma Group. *Eur J Cancer* 2016; **52**: 173-180 [PMID: 26687836 DOI: 10.1016/j.ejca.2015.10.069]

8 **Tirumani SH**, Baheti AD, Tirumani H, O'Neill A, Jagannathan JP. Update on Gastrointestinal Stromal Tumors for Radiologists. *Korean J Radiol* 2017; **18**: 84-93 [PMID: 28096720 DOI: 10.3348/kjr.2017.18.1.84]

9 **Druker BJ**, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S, Sawyers CL. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med* 2001; **344**: 1031-1037 [PMID: 11287972 DOI: 10.1056/NEJM200104053441401]

10 **Li GZ**, Raut CP. Targeted therapy and personalized medicine in gastrointestinal stromal tumors: drug resistance, mechanisms, and treatment strategies. *Onco Targets Ther* 2019; **12**: 5123-5133 [PMID: 31308690 DOI: 10.2147/OTT.S180763]

11 **Park SH**, Lee HJ, Kim MC, Yook JH, Sohn TS, Hyung WJ, Ryu SW, Kurokawa Y, Kim YW, Han SU, Kim HH, Park DJ, Kim W, Lee SI, Cho H, Cho GS, Kim JJ, Kim KH, Yoo MW, Yang HK. Early experience of laparoscopic resection and comparison with open surgery for gastric gastrointestinal stromal tumor: a multicenter retrospective study. *Sci Rep* 2022; **12**: 2290 [PMID: 35145127 DOI: 10.1038/s41598-022-05044-x]

12 **Takahashi R**, Toyokawa T, Yoshii M, Tamura T, Tanaka H, Lee S, Muguruma K, Yashiro M, Ohira M. A Giant Gastric Gastrointestinal Stromal Tumor Successfully Resected Following Neoadjuvant Treatment With Imatinib Mesylate. *Anticancer Res* 2020; **40**: 1147-1152 [PMID: 32014967 DOI: 10.21873/anticanres.14056]

13 **Fletcher CD**, Berman JJ, Corless C, Gorstein F, Lasota J, Longley BJ, Miettinen M, O'Leary TJ, Remotti H, Rubin BP, Shmookler B, Sobin LH, Weiss SW. Diagnosis of gastrointestinal stromal tumors: a consensus approach. *Int J Surg Pathol* 2002; **10**: 81-89 [PMID: 12075401 DOI: 10.1177/106689690201000201]

14 **Eckardt AJ**, Adler A, Gomes EM, Jenssen C, Siebert C, Gottschalk U, Koch M, Röcken C, Rösch T. Endosonographic large-bore biopsy of gastric subepithelial tumors: a prospective multicenter study. *Eur J Gastroenterol Hepatol* 2012; **24**: 1135-1144 [PMID: 22797706 DOI: 10.1097/MEG.0b013e328356eae2]

15 **Demetri GD**, von Mehren M, Antonescu CR, DeMatteo RP, Ganjoo KN, Maki RG, Pisters PW, Raut CP, Riedel RF, Schuetze S, Sundar HM, Trent JC, Wayne JD. NCCN Task Force report: update on the management of patients with gastrointestinal stromal tumors. *J Natl Compr Canc Netw* 2010; **8** Suppl 2: S1-41; quiz S42 [PMID: 20457867 DOI: 10.6004/jnccn.2010.0116]

16 **Park CH**, Kim GH, Lee BE, Song GA, Park DY, Choi KU, Kim DH, Jeon TY. Two staging systems for gastrointestinal stromal tumors in the stomach: which is better? *BMC Gastroenterol* 2017; **17**: 141 [PMID: 29207963 DOI: 10.1186/s12876-017-0705-7]

17 **van der Voort SR**, Incekara F, Wijnenga MMJ, Kapsas G, Gahrmann R, Schouten JW, Nandoe Tewarie R, Lycklama GJ, De Witt Hamer PC, Eijgelaar RS, French PJ, Dubbink HJ, Vincent AJPE, Niessen WJ, van den Bent MJ, Smits M, Klein S. Combined molecular subtyping, grading, and segmentation of glioma using multi-task deep learning. *Neuro Oncol* 2023; **25**: 279-289 [PMID: 35788352 DOI: 10.1093/neuonc/noac166]

18 **Zhao Y**, Chang M, Wang R, Xi IL, Chang K, Huang RY, Vallières M, Habibollahi P, Dagli MS, Palmer M, Zhang PJ, Silva AC, Yang L, Soulen MC, Zhang Z, Bai HX, Stavropoulos SW. Deep Learning Based on MRI for Differentiation of Low- and High-Grade in Low-Stage Renal Cell Carcinoma. *J Magn Reson Imaging* 2020; **52**: 1542-1549 [PMID: 32222054 DOI: 10.1002/jmri.27153]

19 **Zhu Y**, Man C, Gong L, Dong D, Yu X, Wang S, Fang M, Wang S, Fang X, Chen X, Tian J. A deep learning radiomics model for preoperative grading in meningioma. *Eur J Radiol* 2019; **116**: 128-134 [PMID: 31153553 DOI: 10.1016/j.ejrad.2019.04.022]

20 **Jiang Y**, Yang M, Wang S, Li X, Sun Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Commun (Lond)* 2020; **40**: 154-166 [PMID: 32277744 DOI: 10.1002/cac2.12012]

21 **Banzato T**, Causin F, Della Puppa A, Cester G, Mazzai L, Zotti A. Accuracy of deep learning to differentiate the histopathological grading of meningiomas on MR images: A preliminary study. *J Magn Reson Imaging* 2019; **50**: 1152-1159 [PMID: 30896065 DOI: 10.1002/jmri.26723]

22 **Wang P**, Yan J, Qiu H, Huang J, Yang Z, Shi Q, Yan C. A radiomics-clinical combined nomogram-based on non-enhanced CT for discriminating the risk stratification in GISTs. *J Cancer Res Clin Oncol* 2023; **149**: 12993-13003 [PMID: 37464150 DOI: 10.1007/s00432-023-05170-7]

23 **Vernuccio F**, Cannella R, Comelli A, Salvaggio G, Lagalla R, Midiri M. [Radiomics and artificial intelligence: new frontiers in medicine.]. *Recenti Prog Med* 2020; **111**: 130-135 [PMID: 32157259 DOI: 10.1701/3315.32853]

24 **Esteva A**, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115-118 [PMID: 28117445 DOI: 10.1038/nature21056]

25 **Yang TH**, Hwang JI, Yang MS, Hung SW, Chan SW, Wang J, Tyan YS. Gastrointestinal stromal tumors: computed tomographic features and prediction of malignant risk from computed tomographic imaging. *J Chin Med Assoc* 2007; **70**: 367-373 [PMID: 17908650 DOI: 10.1016/S1726-4901(08)70022-4]

26 **Grazzini G**, Guerri S, Cozzi D, Danti G, Gasperoni S, Pradella S, Miele V. Gastrointestinal stromal tumors: relationship between preoperative CT features and pathologic risk stratification. *Tumori* 2021; **107**: 556-563 [PMID: 33620027 DOI: 10.1177/0300891621996447]

27 **Joensuu H**, Hohenberger P, Corless CL. Gastrointestinal stromal tumour. *Lancet* 2013; **382**: 973-983 [PMID: 23623056 DOI: 10.1016/S0140-6736(13)60106-3]

28 **von Mehren M**, Randall RL, Benjamin RS, Boles S, Bui MM, Conrad EU 3rd, Ganjoo KN, George S, Gonzalez RJ, Heslin MJ, Kane JM 3rd, Koon H, Mayerson J, McCarter M, McGarry SV, Meyer C, O'Donnell RJ, Pappo AS, Paz IB, Petersen IA, Pfeifer JD, Riedel RF, Schuetze S, Schupak KD, Schwartz HS, Tap WD, Wayne JD, Bergman MA, Scavone J. Soft Tissue Sarcoma, Version 2.2016, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2016; **14**: 758-786 [PMID: 27283169 DOI: 10.6004/jnccn.2016.0078]

29 **Wang FH**, Zheng HL, Li JT, Li P, Zheng CH, Chen QY, Huang CM, Xie JW. Prediction of recurrence-free survival and adjuvant therapy benefit in patients with gastrointestinal stromal tumors based on radiomics features. *Radiol Med* 2022; **127**: 1085-1097 [PMID: 36057930 DOI: 10.1007/s11547-022-01549-7]

30 **Lin JX**, Chen QF, Zheng CH, Li P, Xie JW, Wang JB, Lu J, Chen QY, Cao LL, Lin M, Tu RH, Huang CM. Is 3-years duration of adjuvant imatinib mesylate treatment sufficient for patients with high-risk gastrointestinal stromal tumor? A study based on long-term follow-up. *J Cancer Res Clin Oncol* 2017; **143**: 727-734 [PMID: 28083710 DOI: 10.1007/s00432-016-2334-x]

31 **Liu S**, Pan X, Liu R, Zheng H, Chen L, Guan W, Wang H, Sun Y, Tang L, Guan Y, Ge Y, He J, Zhou Z. Texture analysis of CT images in

predicting malignancy risk of gastrointestinal stromal tumours. *Clin Radiol* 2018; **73**: 266-274 [PMID: 28969853 DOI: 10.1016/j.crad.2017.09.003]

32 **Choi IY**, Yeom SK, Cha J, Cha SH, Lee SH, Chung HH, Lee CM, Choi J. Feasibility of using computed tomography texture analysis parameters as imaging biomarkers for predicting risk grade of gastrointestinal stromal tumors: comparison with visual inspection. *Abdom Radiol (NY)* 2019; **44**: 2346-2356 [PMID: 30923842 DOI: 10.1007/s00261-019-01995-4]

33 **Iannicelli E**, Carbonetti F, Federici GF, Martini I, Caterino S, Pilozzi E, Panzuto F, Briani C, David V. Evaluation of the Relationships Between Computed Tomography Features, Pathological Findings, and Prognostic Risk Assessment in Gastrointestinal Stromal Tumors. *J Comput Assist Tomogr* 2017; **41**: 271-278 [PMID: 27753723 DOI: 10.1097/RCT.0000000000000499]

34 **Zhou C**, Duan X, Zhang X, Hu H, Wang D, Shen J. Predictive features of CT for risk stratifications in patients with primary gastrointestinal stromal tumour. *Eur Radiol* 2016; **26**: 3086-3093 [PMID: 26699371 DOI: 10.1007/s00330-015-4172-7]

35 **Meng X**, Liu J, Kang J, Wang M, Guan Z, Tian D, Chen X. Lamivudine protects mice from gastric ulcer by activating PGK1 to suppress ferroptosis. *Biochem Pharmacol* 2024; **227**: 116440 [PMID: 39029631 DOI: 10.1016/j.bcp.2024.116440]

36 **Nishida T**, Goto O, Raut CP, Yahagi N. Diagnostic and treatment strategy for small gastrointestinal stromal tumors. *Cancer* 2016; **122**: 3110-3118 [PMID: 27478963 DOI: 10.1002/cncr.30239]

37 **Kang B**, Yuan X, Wang H, Qin S, Song X, Yu X, Zhang S, Sun C, Zhou Q, Wei Y, Shi F, Yang S, Wang X. Preoperative CT-Based Deep Learning Model for Predicting Risk Stratification in Patients With Gastrointestinal Stromal Tumors. *Front Oncol* 2021; **11**: 750875 [PMID: 34631589 DOI: 10.3389/fonc.2021.750875]

38 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

39 **Ning Z**, Luo J, Li Y, Han S, Feng Q, Xu Y, Chen W, Chen T, Zhang Y. Pattern Classification for Gastrointestinal Stromal Tumors by Integration of Radiomics and Deep Convolutional Features. *IEEE J Biomed Health Inform* 2019; **23**: 1181-1191 [PMID: 29993591 DOI: 10.1109/JBHI.2018.2841992]

40 **Tran KA**, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med* 2021; **13**: 152 [PMID: 34579788 DOI: 10.1186/s13073-021-00968-x]

41 **Zhu X**, Wolfgruber TK, Leong L, Jensen M, Scott C, Winham S, Sadowski P, Vachon C, Kerlikowske K, Shepherd JA. Deep Learning Predicts Interval and Screening-detected Cancer from Screening Mammograms: A Case-Case-Control Study in 6369 Women. *Radiology* 2021; **301**: 550-558 [PMID: 34491131 DOI: 10.1148/radiol.2021203758]

42 **Doppalapudi S**, Qiu RG, Badr Y. Lung cancer survival period prediction and understanding: Deep learning approaches. *Int J Med Inform* 2021; **148**: 104371 [PMID: 33461009 DOI: 10.1016/j.ijmedinf.2020.104371]

43 **Wang J**, Shao M, Hu H, Xiao W, Cheng G, Yang G, Ji H, Yu S, Wan J, Xie Z, Xu M. Convolutional neural network applied to preoperative venous-phase CT images predicts risk category in patients with gastric gastrointestinal stromal tumors. *BMC Cancer* 2024; **24**: 280 [PMID: 38429653 DOI: 10.1186/s12885-024-11962-y]

44 **Chen T**, Ning Z, Xu L, Feng X, Han S, Roth HR, Xiong W, Zhao X, Hu Y, Liu H, Yu J, Zhang Y, Li Y, Xu Y, Mori K, Li G. Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumours preoperatively. *Eur Radiol* 2019; **29**: 1074-1082 [PMID: 30116959 DOI: 10.1007/s00330-018-5629-2]

45 **Ren C**, Wang S, Zhang S. Development and validation of a nomogram based on CT images and 3D texture analysis for preoperative prediction of the malignant potential in gastrointestinal stromal tumors. *Cancer Imaging* 2020; **20**: 5 [PMID: 31931874 DOI: 10.1186/s40644-019-0284-7]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: office@baishideng.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com