Texture analysis of CT images based on machine learning to identify whether adrenal cortical adenoma is functional or not

Guohua Yan
Department of Radiology, The Affiliated Hospital of Southwest Medical University, Taiping Street, Luzhou 646000, Sichuan, China.

Yu Zhang
Department of Radiology, The Affiliated Hospital of Southwest Medical University, Taiping Street, Luzhou 646000, Sichuan, China.

Limei Wang
Department of Radiology, The First Affiliated Hospital of the Army Medical University (Southwest Hospital), Chongqing 400038, China.

Xiaopeng Yao
Southwest Medical University, Zhongshan Road, Luzhou 646000, Sichuan, China.

Fugang Han
Department of Radiology, The Affiliated Hospital of Southwest Medical University, Taiping Street, Luzhou 646000, Sichuan, China.

Jian Shu
Department of Radiology, The Affiliated Hospital of Southwest Medical University, Taiping Street, Luzhou 646000, Sichuan, China.

Jiao Bai (baijiao77@163.com)
Department of Radiology, The Affiliated Hospital of Southwest Medical University, Taiping Street, Luzhou 646000, Sichuan, China.

Research Article

Keywords: non-functional, CT scanning texture analysis, machine learning, adrenal cortical adenoma

Posted Date: January 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1224615/v1

Abstract

Background: The detection rates and surgical resection rates of adrenal incidentaloma are increasing, but malignant tumors only account for 10%-15%. Taking adrenal cortical adenomas (ACA) as an example, this paper aims to explore if texture analysis of preoperative computed tomography (CT) images based on machine learning can reliably identify if ACA are functional.

Methods: The clinical and imaging data were collected retrospectively for 75 patients with adrenal cortical adenoma confirmed by surgery and pathology in our hospital from November 2018 to November 2020. MaZda image analysis software was used to segment image data and extract features. Support vector machine (SVM) was used to create an image omics model to predict the functionality of the ACA. Lastly, the area under the receiver operating characteristic (ROC) curve (AUC) was used to evaluate the performance of the radiohistology model.

Results: Four histological models were developed (models 1-4, which represent plain scan, arterial enhancement, venous enhancement and delayed scan, respectively). All models 1-4 showed a good ability to distinguish between functional and non-functional ACA in the training sample, with an average AUC of 0.96, 0.91, 0.91 and 0.88, respectively.

Conclusion: Texture analysis of CT images using machine learning can effectively identify whether adrenal cortical adenoma is functional.

Introduction

Adrenal Cortical Adenoma

Detection and surgical resection rates of adrenal incidentaloma are increasing, while malignant tumors only account for 10%-15% [1]. Adrenal cortical adenoma (ACA) is the most common adrenal incidentaloma and the most common benign adrenal tumor, accounting for 50-80% of all adrenal tumors [2, 3]. Clinically ACA is divided into non-functional and functional subgroups [3]. Functional ACA produces aldosterone, cortisol and sex hormones and causes corresponding clinical symptoms [4]. There are differences in treatment methods between functional ACA and non-functional ACA. Since non-functional ACA with a diameter <4 cm has a good prognosis, it can be followed regularly without surgical resection [5]. However, for functional ACA or ACA with a diameter ≥4cm, surgical treatment is recommended [6]. Diagnosis of functional ACA depends on clinical presentation, as well as results from laboratory, imaging and pathological examinations [4]. Conventional imaging methods alone are not sufficient to distinguish functional adenoma from non-functional adenoma.

Texture Analysis

Texture analysis is a computational quantitative technique, which provides a method for measuring the inhomogeneity of lesions according to the local change of image brightness to distinguish different

pathological areas [7]. Image texture analysis based on machine learning, also known as radiohistology, is a new field of radiology. It uses various calculation methods to obtain quantitative parameters from computed tomography (CT) and magnetic resonance imaging (MRI); it is even possible to automatically analyze heterogeneous tumors [8]. To date, it has been successfully investigated as an auxiliary means for diagnosis, differential diagnosis, metastasis and prognosis of glioma [9], lung cancer [10], breast cancer [11] and rectal cancer [12].

The purpose of this study is to explore whether texture analysis of CT images based on machine learning can be used as a suitable method to identify the function of adrenal cortical adenoma.

Materials And Methods

1.1 General information

The research was approved by the ethics committee of the affiliated hospital of Southwest Medical University (KY2020063). All patient information was completely anonymous before analysis. According to Helsinki Declaration (2000 edition), this study is a retrospective study, the institutional review committee (IRB) of our hospital waived informed consent.

The clinical, pathology and imaging data were collected for 75 patients with ACA confirmed by surgery and pathology at our hospital from November 2018 to November 2020. There were 47 functional adenomas and 28 nonfunctional adenomas, including 36 males and 39 females.

The inclusion criteria are as follows: (1) adrenal cortical adenoma confirmed by clinical diagnosis and postoperative pathological examination; (2) multi-slice spiral CT used for imaging with complete imaging data and clear images; and (3) complete clinical data reviewed by an endocrinologist. The exclusion criteria were as follows: (1) partial clinical and imaging data; (2) a lesion diameter ≥ 4 cm; and (3) image is unable to be processed by imaging software.

1.2 Methods

1.2.1 Classification standard

This study used clinical symptoms and endocrine function evaluation results to distinguish functional ACA from non-functional ACA. Non-functional ACA patients were identified by extensive diagnosis. These standards include, but are not limited to: (1) no full moon face, buffalo back, hair, purple lines and ecchymosis, no sudden severe hypertension, paroxysmal flaccid paralysis, nocturia, hypokalemia and other symptoms. (2) ACTH-F rhythm is normal, 1mg DST cortisol level at midnight is less than 50nmol/L, and 24h urinary free cortisol (24h UFC) is normal, urinary catecholamine, norepinephrine (ne), epinephrine (E) and dopamine (DA) are within normal limits for at least two 24 hour intervals; spironolactone/mineralocorticoid antagonist discontinued for at least 4 weeks; diuretics, angiotensin-converting enzyme inhibitors, angiotensin II receptor antagonists and beta receptor blockers discontinued

for at least 2 weeks; urine aldosterone is within normal limits for 24 hours; and the ratio of aldosterone (pmol/L)/ renin activity in resting plasma [μg/(L.h)] (ARR) < 30. [5]

1.2.2 Grouping method

According to the clinical symptoms and endocrine evaluation results, patients were divided into functional (F) and non-functional (N) groups. The image data for each group were extracted and analyzed according to four categories: plain scan, artery enhancement, vein enhancement and delayed scan.

1.2.3 Image segmentation and feature extraction

We used MaZda image analysis software (Version 4.6, http//www.eletel.p.londz.pl/program/Mazda/) to segment images and extract features. A radiologist with experience using this software sketched a region of interest (ROI) of the collected image data. All patients' image data was stored in the radiation work system.

The texture analysis process began with the radiologist selecting the image of the largest slice of the lesion (Figure 1a-d), saving it in a bitmap (BMP) format, importing the image into the MaZda software and sketching the ROI. We ensured that the distance between the sketching line and the edge of the lesion was approximately 1-2 mm to avoid including free fat outside the lesion and reduce the error caused by edge samples (Figure 2). After sketching the ROI, the MaZda software automatically analyzed the texture feature parameters of the selected area. Lastly, between 295 and 334 texture features were extracted from each image (Figure 3).

To evaluate the reproducibility of this approach, we randomly selected 20 samples from which the first radiologist and a second radiologist also with software experience simultaneously completed the same process described above. Each worked independently and was blinded to the order and grouping of sample data. By comparing the texture features extracted by two radiologists from the ROI description, the consistency between observers was evaluated. An intraclass correlation coefficient (ICC) was used to evaluate the consistency within and between observers. An ICC>0.75 indicates good reliability between observers.

1.2.4 Data feature processing

Using principal component analysis (PCA) dimensionality reduction, 300 effective texture features were selected from all the extracted texture features. First, 300 features of the four categories (plain scan, artery enhancement, vein enhancement and delayed scan) were normalized to eliminate the adverse image effects caused by outlier data. For this study, we defined normalization as (each data-minimum value)/ (maximum value-minimum value). A value of 0 was assigned for negative samples (non-functionality) and a value of 1 was assigned for positive samples (functionality). For the plain scan period, the number of negative samples (19 cases), was approximately equal to the number of positive samples (22 cases). Thus, there was no need to balance samples. However, for the arterial phase,

negative samples (28 cases) and positive samples (43 cases) were imbalanced. To reduce the imbalance of samples, we applied the Synthetic Minority Over-sampling Technique (SMOTE), resulting in 84 negative samples and 86 positive samples.

1.2.5 Model establishment and testing

Based on the SVM, a single group of the four categories (models 1-4, which represent plain scan, artery enhancement, vein enhancement and delayed scan, respectively) was created by combining the texture features to predict the performance of four omics models. We randomly selected 80% of the total samples as the training set and designated the remaining 20% as the test set. We used this training set to train the SVM model, tested with the trained model and repeated the testing 100 times. We calculated the AUC by drawing the ROC curve of the 100 test results.

The discrimination of training model performance mainly depends on the AUC of the ROC. We also calculated the average accuracy, maximum accuracy, minimum accuracy, average negative predictive value (NPV), average positive predictive value (PPV), average sensitivity and average specificity of the 100 tests. Thus, the performance of the established model was evaluated according to all indicators of the validation sample.

Results

The 300 radiohistological features achieved satisfactory consistency between observers. Therefore, no radiohistological features needed to be excluded. The average ICC of the internal observer protocol is 0.92 (range 0.774 to 1, $p < 0.001$), and that of inter-observer protocol is 0.9 (range 0.631 to 1, $p < 0.001$).

The ROC curves of the omics models 1-4 to predict ACA functionality are shown in Figure 4. In the training sample, Model 1 had better performance (average AUC = 0.96; maximum AUC = 1.00, minimum AUC = 0.75; and average accuracy = 92.34%). The performance of models 2 and 3 are similar, with average AUC of 0.91, maximum AUC of 0.99 and 1.00, minimum AUC of 0.75 and 0.76, and average accuracy of 86.21% and 84.27%, respectively. The performance of model 4 is the worst, with average AUC of 0.88, maximum AUC of 0.99, minimum AUC of 0.75 and average accuracy of 80.95% (Table 1).

Table 1 The calculation results

	Model 1	Model 2	Model 3	Model 4
Average AUC	0.96	0.91	0.91	0.88
Maximum AUC	1	0.99	1	0.99
Minimum AUC	0.75	0.75	0.76	0.75
Average accuracy (%)	92.34	86.21	84.27	10.95
Maximum accuracy (%)	100	100	100	100
Minimum accuracy (%)	71.42	50	63.63	50
Average negative predictive value (NVP) (%)	88.03	76.50	79.53	76.12
Average positive predictive value (PVP) (%)	92.34	86.21	84.27	80.95
Average sensitivity (%)	86.95	78.66	79.48	78.30
Average specificity (%)	93.22	85.54	84.08	79.57

Discussion

Background

Adrenal incidentaloma is an asymptomatic adrenal mass. Yet, previous research found that suspected adrenal diseases were not examined on imaging [13]. As machine learning improves, its use is becoming an important aspect of medical care and treatment [5]. The new urology guidelines recommend surgery for adrenal masses with indicators including obvious malignant tendency, tumor diameter ≥4 cm, excessive hormone secretion or obvious clinical symptoms [13]. For masses without obvious surgical indicators, clinical and imaging follow-up is recommended. For masses with a diameter <1 cm, no follow-up is required [6]. General adrenal masses with obvious surgical indicators are easily identified in the clinic and imaging. However, for the masses with no obvious clinical symptoms, no excessive hormone secretion and a tumor diameter between 1-4 cm, the clinical surgical decision-making criteria are rather vague [5, 14]. Therefore, many adrenal incidentalomas are excessively diagnostically resected, which could be followed up by clinical and imaging studies [15].

Adrenal cortical adenoma (ACA), the most common adrenal incidentaloma, is mainly benign and nonfunctional [3]. It is representative of adrenal masses recommended for follow-up. At present, there is a single method to evaluate whether ACA is functional, that is, by measuring clinical symptoms and hormone secretion levels [5]. Endocrine assessment (determination of hormone secretion level) is usually used in patients with adrenal incidentaloma [16]. However, non-functional adenoma is often over diagnostically resected because it is not easy to diagnose and patients lack knowledge of the disease [15].

Innovations and research results.

The difference between treatments for functional ACA and non-functional ACA emphasizes the importance of functional evaluation for adrenal tumors [13]. However, the more complicated endocrine assessment cannot meet the needs of surgeons. Diagnostic resection brings unnecessary pain and expense to patients [14]. Artificial intelligence is widely used in medicine, and most of the differential studies of primary tumors focus on benign and malignant tumors [17–20]. Because of the endocrine particularity of adrenal tumors and the differences in recommended treatment methods, we chose a new angle, that is, first to identify the tumor's functionality, and then to make suggestions for the treatment of these tumors.

We used MaZda image analysis software to sketch ROI and extract 300 tumor-related texture features for machine learning. The training results of four groups of models were obtained, with the average AUC of 0.96, 0.91, 0.91 and 0.88, representing the prediction results of plain scan, artery enhancement, vein enhancement and delayed scan, respectively. The results showed that CT texture analysis based on machine learning has a strong performance in the functional prediction of ACA. However, in this study, the image evaluation of CT contrast delayed scanning is not as ideal as the other three groups, which may be related to the rapid clearance of adenoma enhanced scanning [13].

Limitations and solutions

Of course, radio-omics still faces major challenges, such as replicability of research, standardization of images and data, and ethical and regulatory considerations. In our research, we used ICC to evaluate the reproducibility within and between observers, and used MaZda image analysis software to standardize images. Although we adopted SMOTE to solve the problem of sample imbalance in our research, this technique can only reduce the biases to a certain extent. This may be a shortcoming of our research. Nevertheless, this study confirms that radiohistology still has a strong predictive power.

Conclusion

The training results of models 1-4 suggest well-validated models, with the average AUC of 0.96, 0.91, 0.91 and 0.88, respectively. Our results suggest that texture analysis of preoperative CT images based on machine learning can be used to effectively predict the functionality of ACA.

Declarations

Ethics approval and consent to participate

This thesis is based on the Helsinki Declaration and has been approved by the Ethics Review Committee of the Affiliated Hospital of Southwest Medical University. In our research, all procedures followed the regulations of the Affiliated Hospital of Southwest Medical University, and obtained the ethical approval of the hospital ethics committee (KY2020063). The Review Committee of the Affiliated Hospital of Southwest Medical University approved this retrospective study, and exempted the need to obtain informed consent from patients.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available dueto patient privacy and ethical requirements but are available from the corresponding author on reasonable request.

Competing interests

The authors declares that they have no competing interests.

Funding

This study has no financial support.

Authors' contributions

GHY, YZ and LMW collected all the images needed in this paper, and finished the post-processing of the images. GHY wrote the paper and edited the manuscript. XPY analyzed all the data. All authors have read and approved the final manuscript.

Acknowledgements

None declared

Author details

[1] Department of Radiology, The Affiliated Hospital of Southwest Medical

University, Taiping Street, Luzhou 646000, Sichuan, China. [2] Department of Radiology, The First Affiliated Hospital of the Army Medical University (Southwest Hospital), Chongqing 400038, China. [3] Southwest Medical University, Zhongshan Road, Luzhou 646000, Sichuan, China.

References

1. Musella M, Conzo G, Milone M, Corcione F, Belli G, De Palma M, Tricarico A, Santini L, Palazzo A, Bianco P, Biondi B, Pivonello R, Colao A. Preoperative workup in the assessment of adrenal incidentalomas: outcome from 282 consecutive laparoscopic adrenalectomies. BMC Surg. 2013 Nov 27;13:57.
2. Low G, Sahi K. Clinical and imaging overview of functional adrenal neoplasms. Int J Urol. 2012 Aug;19(8):697–708.

3. Lam AK. Update on Adrenal Tumours in 2017 World Health Organization (WHO) of Endocrine Tumours. Endocr Pathol. 2017 Sep;28(3):213–227.
4. Zheng Y, Liu X, Zhong Y, Lv F, Yang H. A Preliminary Study for Distinguish Hormone-Secreting Functional Adrenocortical Adenoma Subtypes Using Multiparametric CT Radiomics-Based Machine Learning Model and Nomogram. Front Oncol. 2020 Sep 29;10:570502.
5. Bernini GP, Moretti A, Oriandini C, Bardini M, Taurino C, Salvetti A. Long-term morphological and hormonal follow-up in a single unit on 115 patients with adrenal incidentalomas. Br J Cancer. 2005 Mar 28;92(6):1104–9.
6. Bhat HS, Tiyadath BN. Management of Adrenal Masses. Indian J Surg Oncol. 2017 Mar;8(1):67–73.
7. Ho LM, Samei E, Mazurowski MA, Zheng Y, Allen BC, Nelson RC, Marin D. Can Texture Analysis Be Used to Distinguish Benign From Malignant Adrenal Nodules on Unenhanced CT, Contrast-Enhanced CT, or In-Phase and Opposed-Phase MRI? AJR Am J Roentgenol. 2019 Mar;212(3):554–561.
8. Daye D, Staziaki PV, Furtado VF, Tabari A, Fintelmann FJ, Frenk NE, Shyn P, Tuncali K, Silverman S, Arellano R, Gee MS, Uppot RN. CT Texture Analysis and Machine Learning Improve Post-ablation Prognostication in Patients with Adrenal Metastases: A Proof of Concept. Cardiovasc Intervent Radiol. 2019 Dec;42(12):1771–1776.
9. Lu CF, Hsu FT, Hsieh KL, Kao YJ, Cheng SJ, Hsu JB, Tsai PH, Chen RJ, Huang CC, Yen Y, Chen CY. Machine Learning-Based Radiomics for Molecular Subtyping of Gliomas. Clin Cancer Res. 2018 Sep 15;24(18):4429–4436.
10. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, Ye Z, Makrigiorgos M, Fennessy F, Mak RH, Gillies R, Quackenbush J, Aerts HJWL. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. Cancer Res. 2017 Jul 15;77(14):3922–3930.
11. Li H, Zhu Y, Burnside ES, Huang E, Drukker K, Hoadley KA, Fan C, Conzen SD, Zuley M, Net JM, Sutton E, Whitman GJ, Morris E, Perou CM, Ji Y, Giger ML. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. NPJ Breast Cancer. 2016;2:16012–.
12. Bibault JE, Giraud P, Housset M, Durdux C, Taieb J, Berger A, Coriat R, Chaussade S, Dousset B, Nordlinger B, Burgun A. Deep Learning and Radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. Sci Rep. 2018 Aug 22;8(1):12611.
13. Fassnacht M, Arlt W, Bancos I, Dralle H, Newell-Price J, Sahdev A, Tabarin A, Terzolo M, Tsagarakis S, Dekkers OM. Management of adrenal incidentalomas: European Society of Endocrinology Clinical Practice Guideline in collaboration with the European Network for the Study of Adrenal Tumors. Eur J Endocrinol. 2016 Aug;175(2):G1-G34.
14. Mayo-Smith WW, Song JH, Boland GL, Francis IR, Israel GM, Mazzaglia PJ, Berland LL, Pandharipande PV. Management of Incidental Adrenal Masses: A White Paper of the ACR Incidental Findings Committee. J Am Coll Radiol. 2017 Aug;14(8):1038–1044.
15. Akbulut S, Erten O, Kahramangil B, Gokceimam M, Kim YS, Li P, Remer EM, Berber E. A Critical Analysis of Computed Tomography Washout in Lipid-Poor Adrenal Incidentalomas. Ann Surg Oncol.

2021 May;28(5):2756–2762.

16. Glazer DI, Mayo-Smith WW. Management of incidental adrenal masses: an update. Abdom Radiol (NY). 2020 Apr;45(4):892–900.
17. Torresan F, Crimì F, Ceccato F, Zavan F, Barbot M, Lacognata C, Motta R, Armellin C, Scaroni C, Quaia E, Campi C, Iacobone M. Radiomics: a new tool to differentiate adrenocortical adenoma from carcinoma. BJS Open. 2021 Jan 8;5(1):zraa061.
18. Elmohr MM, Fuentes D, Habra MA, Bhosale PR, Qayyum AA, Gates E, Morshid AI, Hazle JD, Elsayes KM. Machine learning-based texture analysis for differentiation of large adrenal cortical tumours on CT. Clin Radiol. 2019 Oct;74(10):818.e1-818.e7.
19. Yi X, Guan X, Chen C, Zhang Y, Zhang Z, Li M, Liu P, Yu A, Long X, Liu L, Chen BT, Zee C. Adrenal incidentaloma: machine learning-based quantitative texture analysis of unenhanced CT can effectively differentiate sPHEO from lipid-poor adrenal adenoma. J Cancer. 2018 Sep 8;9(19):3577-3582.
20. Nakajo M, Jinguji M, Nakajo M, Shinaji T, Nakabeppu Y, Fukukura Y, Yoshiura T. Texture analysis of FDG PET/CT for differentiating between FDG-avid benign and metastatic adrenal tumors: efficacy of combining SUV and texture parameters. Abdom Radiol (NY). 2017 Dec;42(12):2882–2889.

Figures

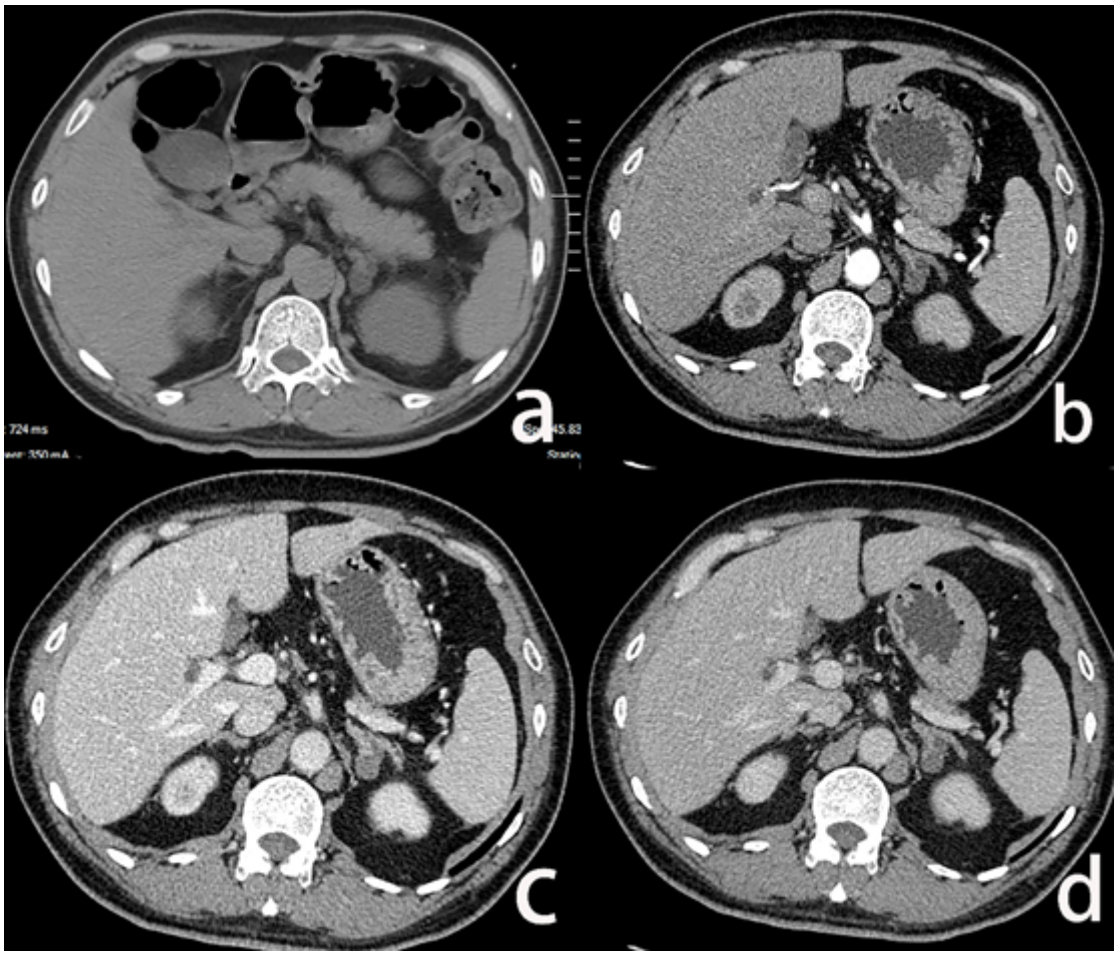

Figure 1

Selected the largest level of tumor. a) plain scan; b) arterial phase; c) venous phase; and d) delayed phase.

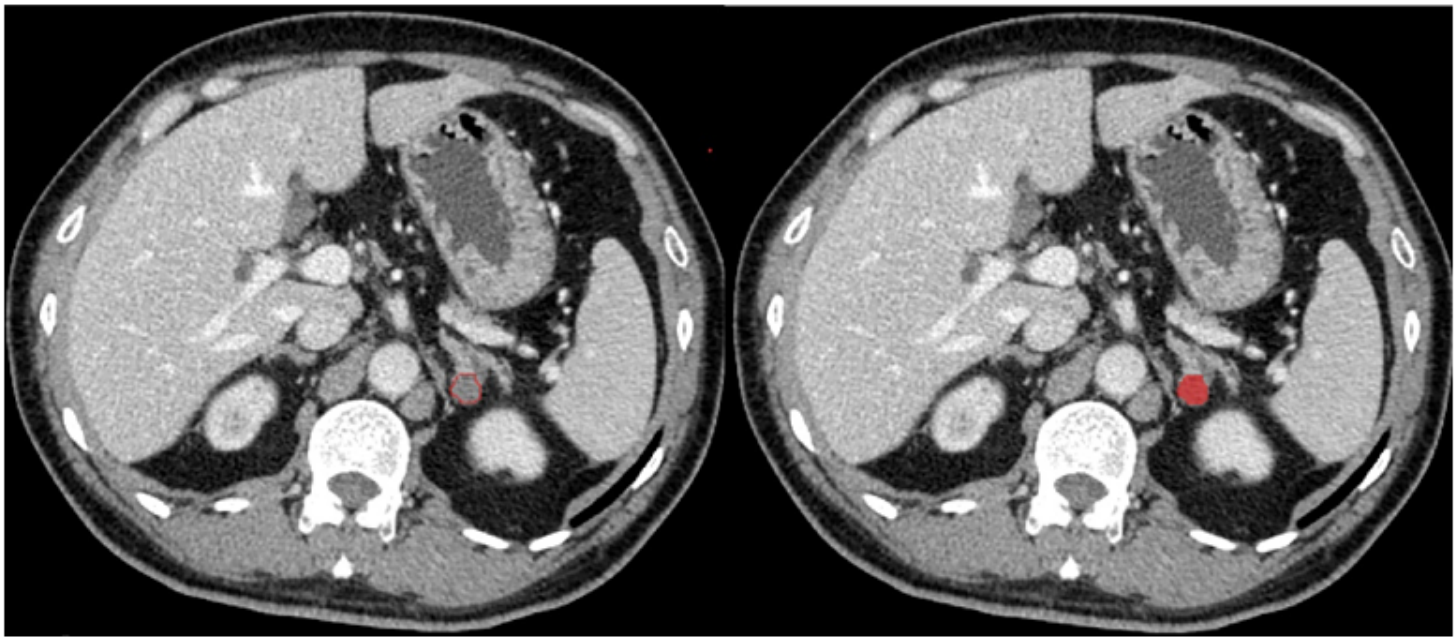

Figure 2

Manually sketched ROI. The distance between the sketching line and the edge of the lesion is about 1-2mm to avoid including free fat outside the lesion and reduce the error caused by edge samples.

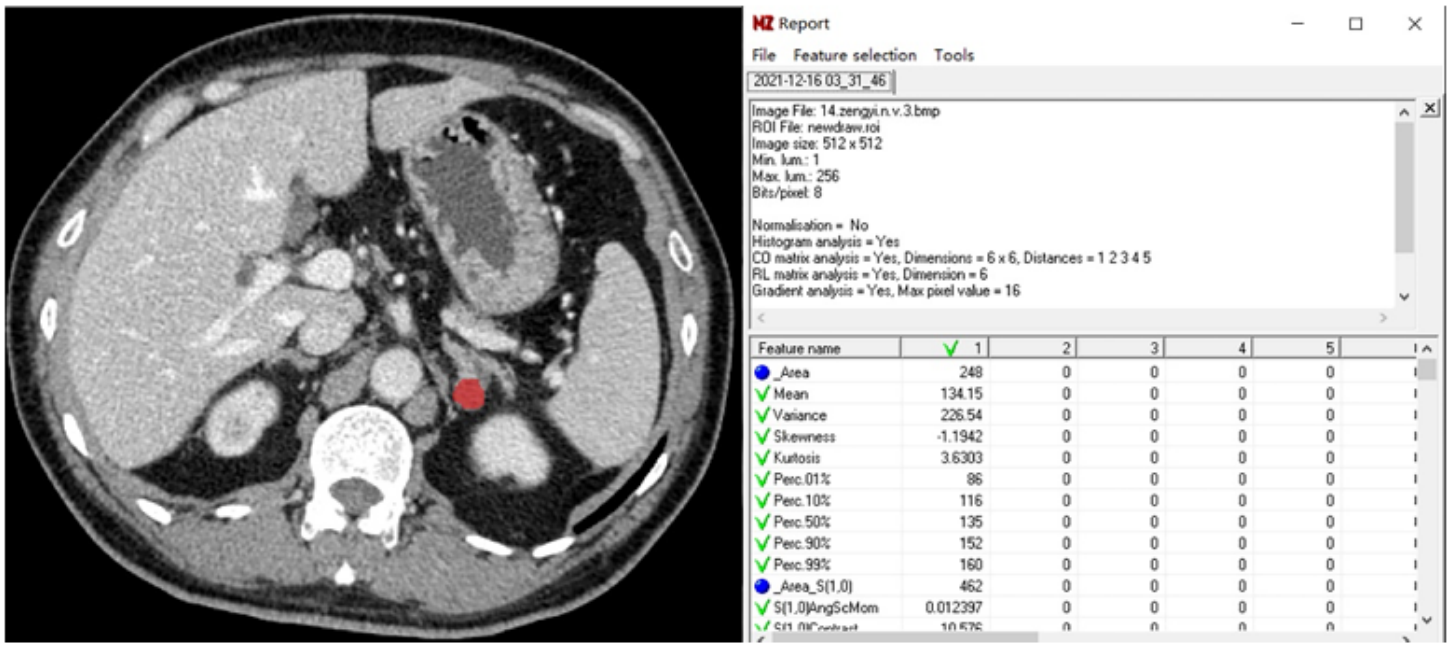

Figure 3

MaZda software used to automatically analyze the texture feature parameters of the selected area. Between 295-334 texture features are extracted from each image.

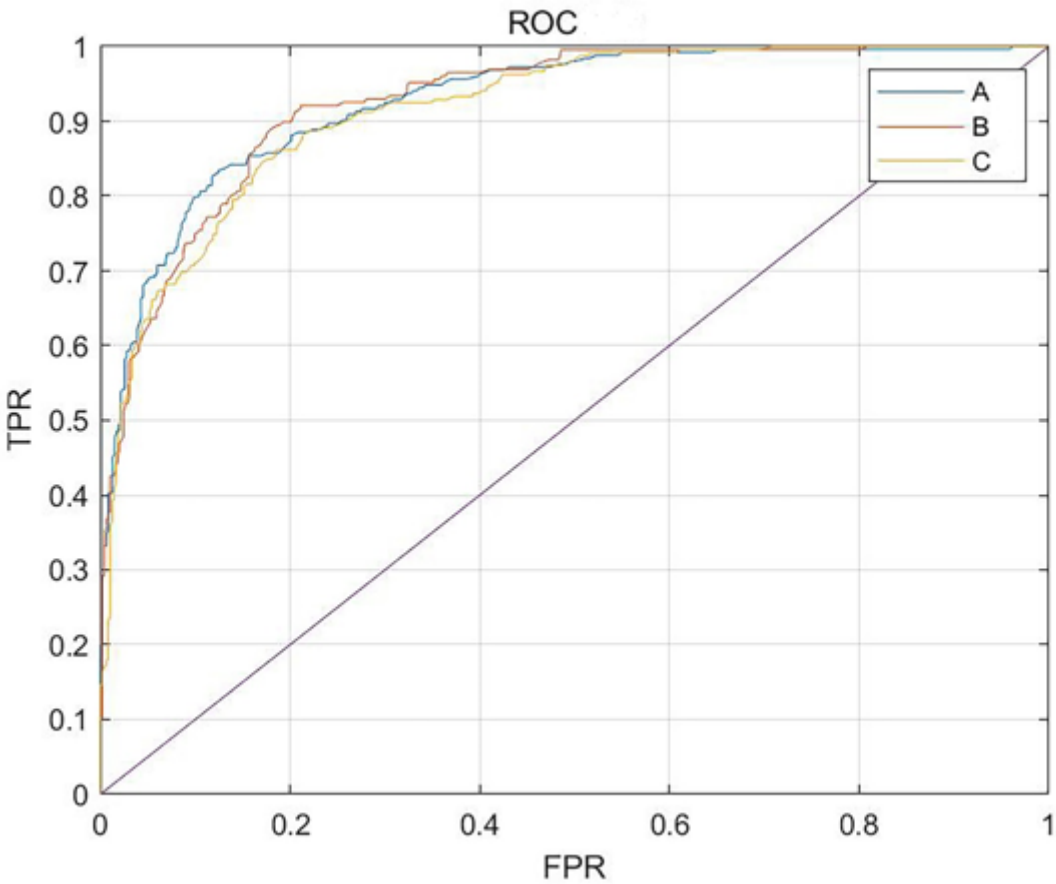

Figure 4

The receiver operating characteristic (ROC) curve for judging the performance of the training model is a three-class curve, in which grades A, B and C represent low, medium and high levels respectively.