

Machine learning-based multiparametric traditional multislice computed tomography radiomics for improving the discrimination of parotid neoplasms

ZHIFENG XU^{1*}, YABIN JIN^{2*}, WENXIU WU¹, JINMIAN WU³, BING LUO¹, CHENGLONG ZENG¹, XIUQIN GUO⁴, MINGCONG GAO⁵, SHIQIN GUO⁶ and AIZHEN PAN¹

¹Department of Radiology; ²Clinical Research Institute, The First People's Hospital of Foshan, Foshan, Guangdong 528000;

³Department of Radiology, The Sanshui People's Hospital, Foshan, Guangdong 528100; ⁴Department of Radiology, The Foshan Chancheng District Central Hospital; ⁵Department of Radiology, The Foshan Medicine Hospital; ⁶Department of Radiology, The Second People's Hospital of Nanhai District, Foshan, Guangdong 528000, P.R. China

Received February 19, 2021; Accepted July 28, 2021

DOI: 10.3892/mco.2021.2407

Abstract. Characterization of parotid tumors is important for treatment planning and prognosis, and parotid tumor discrimination has recently been developed at the molecular level. The aim of the present study was to establish a machine learning (ML) predictive model based on multiparametric traditional multislice CT (MSCT) radiomic and clinical data analysis to improve the accuracy of differentiation among pleomorphic adenoma (PA), Warthin tumor (WT) and parotid carcinoma (PCa). A total of 345 patients (200 with WT, 91 with PA and 54 with PCa) with pathologically confirmed parotid tumors were retrospectively enrolled from five independent institutions between January 2010 and May 2019. A total of 273 patients recruited from institutions 1, 2 and 3 were randomly assigned to the training model; the independent validation set consisted of 72 patients treated at institutions 1, 4 and 5. Data were investigated using a linear discriminant analysis-based ML classifier. Feature selection and dimension reduction were conducted using reproducibility testing and a wrapper method. The diagnostic accuracy of the predictive model was compared with histopathological findings as reference results. This classifier achieved a satisfactory performance for the discrimination of PA, WT and PCa, with a total accuracy of 82.1% in the training cohort and 80.5% in the validation cohort. In conclusion, ML-based

multiparametric traditional MSCT radiomics can improve the accuracy of differentiation among PA, WT and PCa. The findings of the present study should be validated by multicenter prospective studies using completely independent external data.

Introduction

Characterization of parotid tumors (PTs) is important for treatment planning and prognosis, and PT discrimination has recently been developed at the molecular level. Diffusion-weighted imaging (DWI)- and dynamic contrast-enhanced (DCE)-MRI are the most important methods described to date, with a single or multiple technology/parameter combination (1-3), largely expanding predictive power. Additionally, molecular imaging, such as diffusion tensor imaging, has been used to differentiate malignant from benign tumors, as well as Warthin tumors (WTs) from pleomorphic adenomas (PAs), both with an accuracy of 86% (4). Unfortunately, such measurements based on molecular biomarkers are expensive, time-consuming, involve complicated analyses and are available at few facilities. Therefore, it is of paramount importance to explore simple and cost-effective approaches to improve the accuracy of differentiation among different PTs. Compared with MRI, multislice computed tomography (MSCT) has the characteristics of wide application and simple generalization of the results (5). The effectiveness of machine learning (ML) methods for high-throughput extraction of quantitative features from clinical images has recently been demonstrated in multiple disciplinary fields, including tumor identification (6) and clinical outcome prediction (7). However, there are few published reports using ML-based multiparametric CT radiomics to study PTs, which represents a knowledge gap. The aim of the present study was to establish a model through analysis of clinical data and dual contrast-enhanced CT coupled with ML-based algorithms to improve the accuracy of PT differential diagnosis and to test its efficacy upon biopsy completion to enable timely decisions for treatment.

Correspondence to: Dr Zhifeng Xu or Dr Aizhen Pan, Department of Radiology, The First People's Hospital of Foshan, 81 Lingnan Avenue, Foshan, Guangdong 528000, P.R. China
E-mail: xuzf83@126.com
E-mail: pazhen2121@126.com

*Contributed equally

Key words: parotid tumor, machine learning, linear discriminant analysis, computed tomography, radiomics

Materials and methods

Patients. A total of 1,100 patients (427 women and 683 men; mean age, 54.5 years; age range, 19-78 years) with pathologically confirmed PTs who were treated by curative resection between January 2010 and May 2019 at five independent institutions [i) The First People's Hospital of Foshan; ii) The Sanshui People's Hospital; iii) The Foshan Chancheng District Central Hospital; iv) The Foshan Medicine Hospital; and v) The Second People's Hospital of Nanhai District; all in Foshan, China] were reviewed. Of the subjects reviewed, 345 were included in the final analysis according to the study criteria shown in Fig. 1. A total of 273 patients recruited from institutions 1, 2 and 3 were randomly assigned to the training model. The independent validation set consisted of 72 patients treated at institutions 1, 4 and 5.

Imaging feature extraction. This retrospective study was approved by the Institutional Review Board (Ethics Committee of The First People's Hospital of Foshan; approval no. FSYYY-EC-SOP-008-02.0-A09). Informed consent was obtained from all participants included in the study. Two CT radiologists with 10 years (reader 1) and 15 years (reader 2) of experience in PT imaging who were blinded to all clinical data independently reviewed baseline CT images to evaluate the following characteristics: Location, number, boundary, calcification, morphology, vascular marginalization (VM; small, newly formed blood vessels around tumors), enlarged lymph nodes (ELN) around tumors (diameter ≥ 5 mm), density (Hounsfield Units; HU) in plain phase (PP), arterial phase (AP) and venous phase (VP), perfusion rate (PR) in AP [$PR = (HU_{AP} - HU_{PP}) / HU_{PP}$] and clearance [$(HU_{VP} - HU_{AP}) / HU_{AP}$] (Figs. S1-S2). The tumor was deemed to have a clear boundary if it was well-demarcated along its entire circumference and to have an unclear boundary otherwise. When the assessment was inconsistent, consensus was reached through discussion. To assess tumor attenuation, a circular region of interest (ROI) that excluded obvious cystic and necrotic areas was identified, and the two averages of the ROI were taken as the final value. Moreover, patient age and sex were considered. A total of 16 indicators were examined in the statistical analysis.

ML models and statistical analysis. The Kruskal-Wallis H test and χ^2 test were used to compare parameters among WT, PA and parotid carcinoma (PCa). Parameters with significant differences among the three groups were used to construct a classifier model, and the other parameters were filtered. The R package MASS (<https://CRAN.R-project.org/package=MASS>) was utilized to construct a classifier model by linear discriminant analysis (LDA) using the training data. The model was analyzed as a classifier for the three tumors in the training cohort (institutions 1, 2 and 3) and validated in a combination cohort from three hospitals (institutions 1, 4 and 5). The effect of the classifier model was based on accuracy. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

Characteristics of different PTs. Compared with patients with PA and PCa, the majority of patients with WT had

Table I. Diagnostic performance of machine learning-based CT features and clinical data classifiers for discriminating WT, PA and PCa in the training and validation cohorts.

Training group		Pathological diagnosis	
Predictive	PCa	PA	WT
PCa	27	12	4
PA	6	42	5
WTs	6	14	147
Accuracy	82.1%		
Validation group			
Predictive	PCa	PA	WT
PCa	12	1	1
PA	1	13	2
WTs	2	9	41
Accuracy	80.5%		
PA, pleomorphic adenomas; WT, Warthin tumor; PCa, parotid carcinoma.			

multiple lesions in the bilateral parotid glands, and the difference was statistically significant ($P < 0.01$). A total of 79.5 and 74.4% of lesions in PCa patients displayed an unclear boundary and morphological irregularity, respectively, which was higher compared with the respective percentages in patients with WT and PA ($P < 0.01$). Additionally, ELN signs were observed around the lesions in ~64.7% of WT and in 64.1% of PCa cases, which was significantly higher compared with the percentage for PA (30.8%; $P < 0.01$). The majority of WTs (~66.7%) displayed VM signs ($P < 0.01$). Regarding quantitative CT characteristics, the Hu and arterial perfusion rate of WTs were both significantly higher compared with those of other tumors in plain and dual-enhanced scans ($P < 0.01$), and the majority of patients with WT exhibited characteristics of fast forward and fast retreat, with a high clearance rate (Figs. S1-S2). These features were selected for predictive ML model establishment.

Construction of the ML classifier model. An ML classifier model was successfully constructed based on the selected clinical and CT characteristics. These characteristics differed significantly among WT, PA and PCa. The results regarding the ability of ML-based CT characteristic and clinical data analysis to discriminate the three subtypes of PTs in the training and validation cohorts are summarized in Table I and Fig. 2. In general, this classifier achieved satisfactory performance, with a total accuracy of 82.1% in the training cohort and 80.5% in the validation cohort. The cases of 3 representative patients with WT, PA and PCa for whom the proposed ML model correctly predicted pathology are presented in Fig. 3.

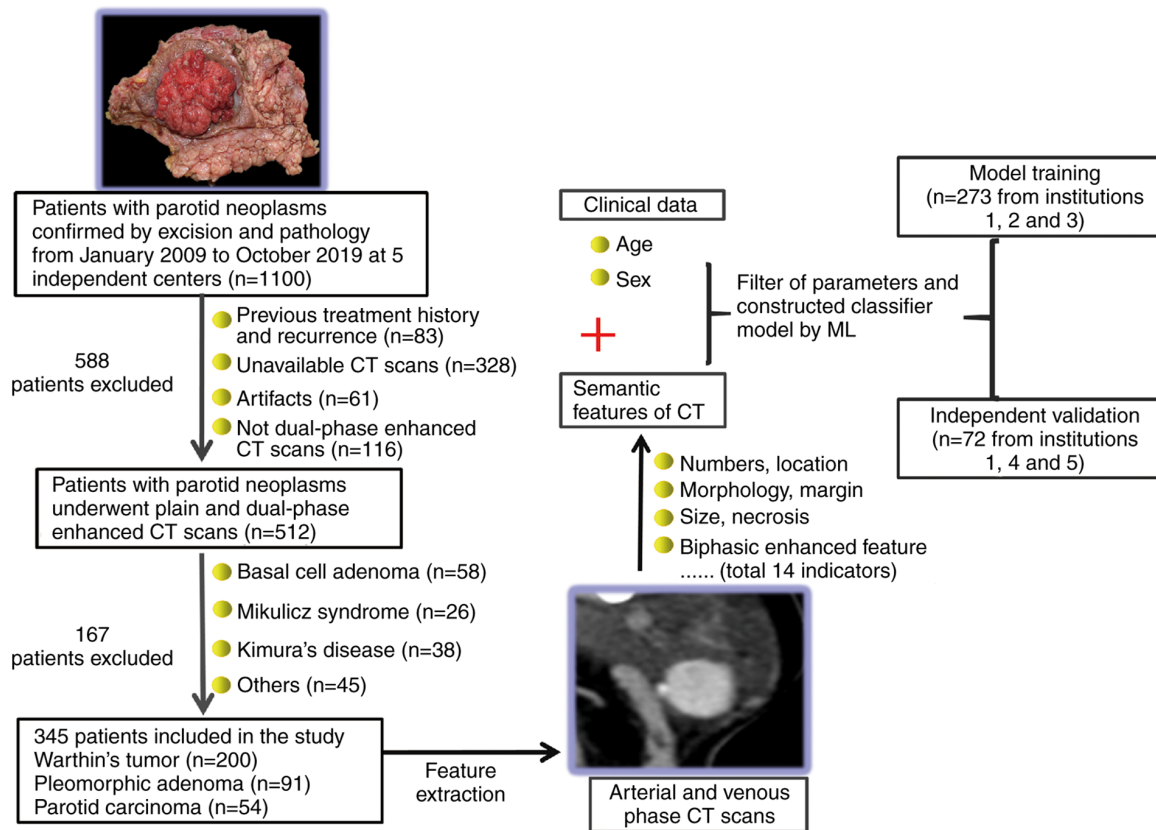


Figure 1. Flowchart of the study design. ML, machine learning.

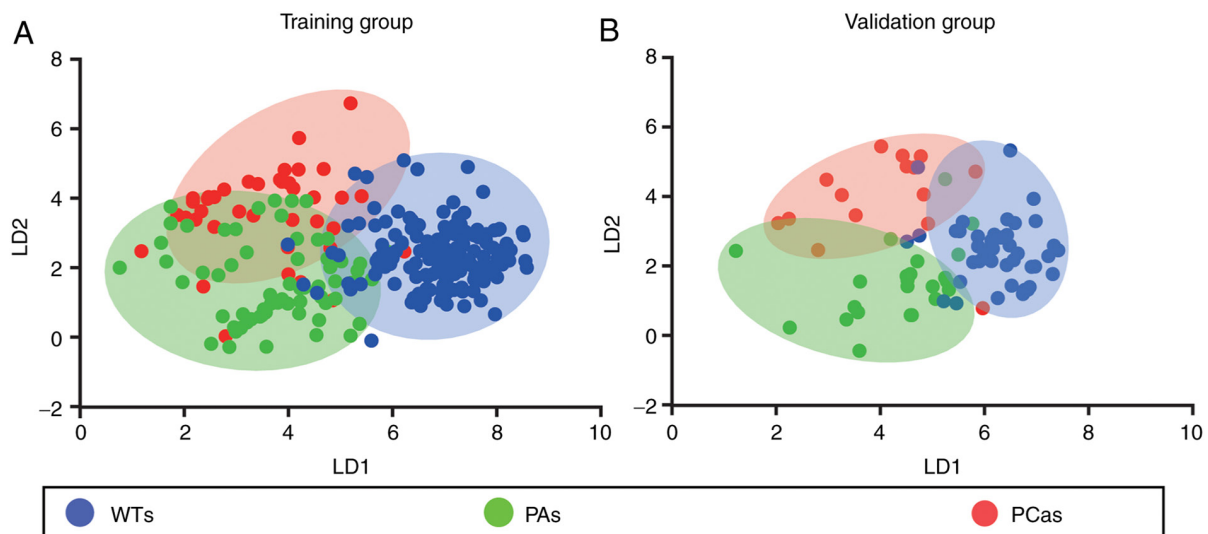


Figure 2. Results of machine learning-based CT feature and clinical data analysis for parotid tumor discrimination. The scatter plot shows that the model was able to distinguish the three types of tumors in the (A) training and (B) validation cohort by linear discriminant analysis. PA, pleomorphic adenoma; WT, Warthin tumor; PCa, parotid carcinoma; LD, linear discriminant.

Discussion

The characterization of PTs is important for preoperative treatment planning, as well as assessment of therapeutic responsiveness and prognosis (1). In the present study, LDA-based ML with selected clinical data and CT radiomics were applied to discriminate between benign (PA and WT) and malignant (PCa) PTs preoperatively, with satisfactory

results. The results demonstrated that the ML approach is a promising non-invasive method that is feasible and reliable for the evaluation of PTs.

In general, the clinical symptoms of PTs are non-specific, and preoperative imaging is crucial for diagnosis. CT and MRI are promising examination techniques for PTs based on the analysis of various morphological parameters and molecular imaging. Over the last 5 years, CT multiphase enhancement (8),

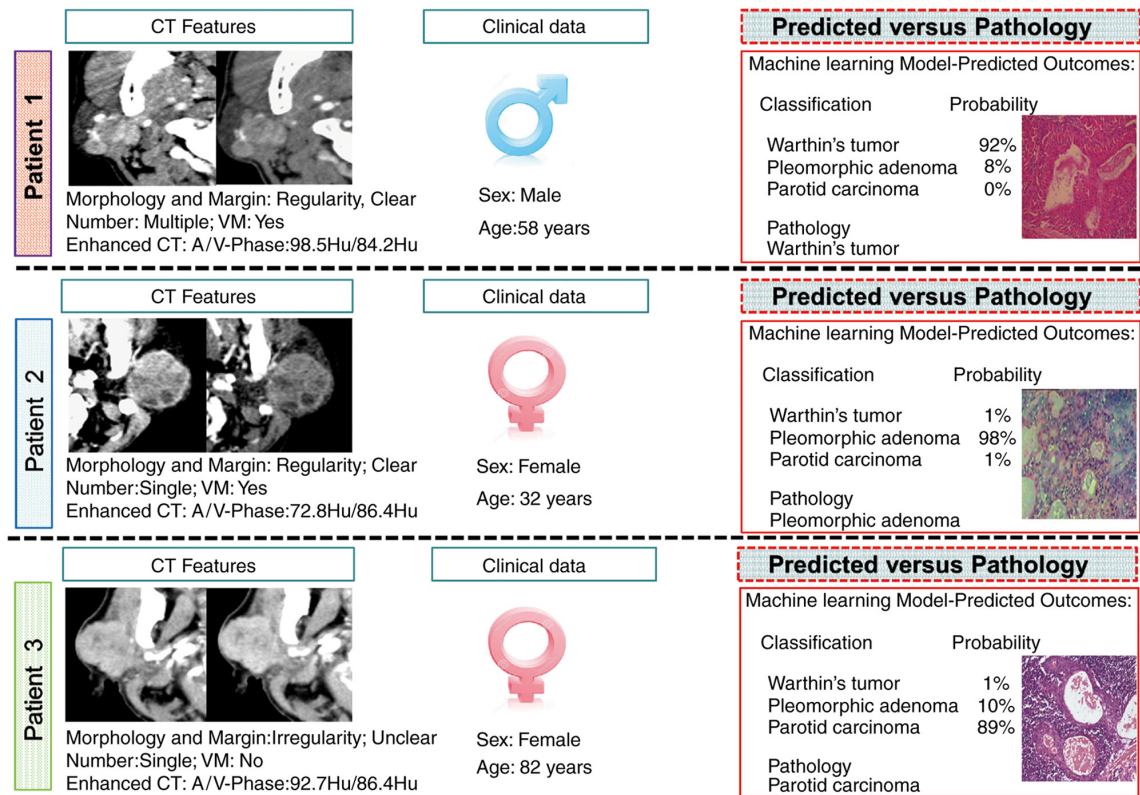


Figure 3. Three representative cases showing the clinical translation of the CT feature-derived model. Patient 1: A 58-year-old male with multiple tumors in the right parotid gland had a high predicted probability of Warthin tumor, and this result was consistent with the pathology. Patient 2: A 32-year-old female with one tumor in the left parotid gland had a high predicted probability of pleomorphic adenoma, and this result was consistent with the pathology. Patient 3: An 82-year-old female with one tumor in the right parotid gland had a high predicted probability of parotid carcinoma, and this result was consistent with the pathology.

CT perfusion (9), DCE-MRI (1,3,10) and apparent diffusion coefficients derived from DWI (1-3,10) have been applied to improve the diagnostic efficiency of PTs, with diagnostic accuracies ranging from 70 to 97% (3,8-12). However, the large influence of subjective factors, the small case sample size and a single evaluation index are the main limitations of these studies, and results among studies have not been replicated. In ML, the predominant task is predictive modeling, namely, the creation of models to characterize new examples. In terms of the clinical demand discussed above, the histopathology of PT must be predicted accurately and preoperatively, for which there is an opportunity to apply ML technology. As our sample number was lower compared with that in common industrial cases, 2 clinical indices and 14 CT features were selected; 16 indicators were included in the statistical analysis according to prior clinical knowledge for prediction. ML has been shown to have a great advantage in tumor diagnosis as well as prognosis and recurrence prediction, with an increasing number of reports involving hepatic carcinoma (13), thyroid nodules (14), renal tumors (15) and colon cancer (16), among others. Previous studies have demonstrated that ML-based texture analysis has diagnostic accuracy ranging from 98.3 to 100% for neoplastic lesions of the abdomen. To the best of our knowledge, this is the first ML study for PT discrimination, revealing powerful diagnostic performance.

The key to building a powerful predictive model is to select efficient indicators and appropriate modeling methods. In the present study, radiomic features were extracted from plain scans and dual-phase enhancement of CT examinations. The analysis

mainly included the location, number, boundary, morphology, VM, ELN, density and enhancement pattern, among others, which should represent the underlying tumor biology. Previous studies have reported that feature optimization can enhance the predictive value of radiomic features (6,7,17). Therefore, statistical analysis was first used to eliminate indicators without significant differences, and among the tumor categories, no significant differences were observed regarding characteristics such as tumor location or nodule-in-nodule sign. Finally, 14 CT characteristics that were valuable in the differential diagnosis of PA, WT and PCa were selected, which was consistent with the results of our previous studies (3,5,9). In addition, given the importance of sex and age in the identification of PTs found in previous studies, these characteristics were included as modeling indicators to improve the success of the model. It was previously demonstrated that LDA-based models had slightly better diagnostic performance compared with support vector machine-based models (6,18). LDA is a representative linear classifier that uses a straight line (a vector) to separate three classes (19), namely PA, WT and PCa in the present study. In general, this classifier achieved satisfactory performance, with a total accuracy of 82.1% in the training cohort and 80.5% in the validation cohort. Our previous study (3) established a simple model for PT discrimination based on multiparametric analysis derived from DCE-MR and DWI, and LDA analysis was used. Our results revealed the biomarker that appeared to be the best indicator (extracellular extravascular space volume fraction + time-intensity curve), with a prediction accuracy of 75%, while the predictive model in the present study exhibited

a higher accuracy at 80.5%. Therefore, we believe that the model established in the present study is more valuable. First, dual-phase enhanced CT scanning has a wide range of applications and is relatively simple and practical. Although CT involves exposure to radiation, the progress of low-dose scanning and image reconstruction algorithms has greatly reduced the radiation dose (20). Second, the modeling indicators combined with clinical and CT characteristics in the present study were efficient and simple. Additionally, the number of cases in this study was large (345 cases) and, as the data from the training and prediction models were both derived from multiple centers, the model is considered as reliable and accurate.

There were three primary limitations to the present study. First, as this was a retrospective study, selection bias could not be fully avoided, and the number of patients with each subtype of PT was not balanced in the validation set, which may have influenced the performance metrics to a certain extent. However, it is not clear whether this had a significant impact on the modeling. Second, a manual ROI method was used for the determination of tumor enhancement indicators. Although multiple point and multiple site measurements were adopted to reduce errors, the enhancement characteristics of the tumors could not be fully and accurately reflected. We are confident that, in the future, semiautomated software will be able to recognize PTs with CT images. Finally, only three types of tumors (WTs, PA and PCa) were analyzed in the present study. Other types of tumors, such as basal cell adenoma, eosinophiloma and oncocytic adenoma, are rare tumors with a small number of cases. Therefore, these pathological types of PTs were not included in the diagnostic scope in this study. The current diagnostic model still needs to be further improved.

In summary, we herein established an ML classifier based on LDA analysis with selected CT characteristics derived from dual-enhanced parotid CT examination and clinical characteristics (sex and age) that achieved 80.5% diagnostic accuracy for differentiating among PA, WT and PCa. To the best of our knowledge, this is the first successful ML classifier established for PT discrimination. This predictive model has high diagnostic accuracy and can be quickly translated into clinical applications and popularized in primary hospitals. Therefore, this classifier will benefit patients with PTs and help in providing appropriate treatment.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Science Innovative Project of Foshan (grant no. FSOAA-KJ218-1301-0021), the Foshan Ascending Peak Plan Project (grant no. 2020B003), and the Medical Scientific Research Foundation of Guangdong Province of China (grant no. A2021493).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

ZX performed the majority of experiments and wrote the manuscript; YJ wrote the manuscript and prepared the graphs; WW, JW, BL, CZ, XG, MG and SG collected the data and analyzed images; AP designed the study and was involved in editing the manuscript. ZX and AP have seen and can confirm the authenticity of the raw data. All the authors have read and approved the final manuscript.

Ethics approval and consent to participate

This retrospective study was approved by the Institutional Review Board (Ethics Committee of The First People's Hospital of Foshan; approval no. FSYYY-EC-SOP-008-02. 0-A09). Written consent was obtained from all participants included in the study.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Eissa L, Seif SA, Desooky SE, Eid M and Koraitim T: Accuracy assessment of combined diffusion weighed and dynamic gadolinium MR sequences in characterization of salivary gland tumors. *Egypt J Radiol Nucl Med* 47: 127-139, 2016.
2. Razek AAK, Samir S and Ashmalla GA: Characterization of parotid tumors with dynamic susceptibility contrast perfusion weighted magnetic resonance imaging and diffusion-weighted MR imaging. *J Comput Assist Tomogr* 41: 131-136, 2017.
3. Xu Z, Zheng S, Pan A, Cheng X and Gao M: A multiparametric analysis based on DCE-MRI to improve the accuracy of parotid tumor discrimination. *Eur J Nucl Med Mol Imaging* 46: 2228-22344, 2019.
4. Razek AA: Characterization of salivary gland tumours with diffusion tensor imaging. *Dentomaxillofac Radiol* 47: 20170343, 2018.
5. Xu ZF, Yong F, Yu T, Chen YY, Gao MY, Zhou T, Pan AZ and Wu RH: Different histological subtypes of parotid gland tumors: CT findings and diagnostic strategy. *World J Radiol* 5: 313-320, 2013.
6. Fan Y, Chen C, Zhao F, Tian Z, Wang J, Ma X and Xu J: Radiomics-based machine learning technology enables better differentiation between glioblastoma and anaplastic oligodendroglioma. *Front Oncol* 9: 1164, 2019.
7. Hua X, Huang X, Huang ZZ, Song CG, Deng JP, Long ZQ, Lin HX and Yuan ZY: Establishment of prognostic nomograms based on skeletal muscle index and serum biomarker in breast cancer patients receiving radiotherapy. *Clin Transl Med* 10: e115, 2020.
8. Jin GQ, Su DK, Xie D, Zhao W, Liu LD and Zhu XN: Distinguishing benign from malignant parotid gland tumours: Low-dose multiphasic CT protocol with 5-minute delay. *Eur Radiol* 21: 1692-1698, 2011.
9. Xu Z, Rong F, Yu T, Chen YY, Gao Q, Zhou T and Pan AZ: Pleomorphic adenoma versus Warthin tumor of the parotid gland: Diagnostic value of CT perfusion imaging and its pathologic explanation:CT Perfusion of pleomorphic adenoma versus Warthin. *J Tumor* 4: 419-425, 2016.
10. Yuan Y, Tang W and Tao X: Parotid gland lesions: Separate and combined diagnostic value of conventional MRI, diffusion-weighted imaging, and dynamic contrast-enhanced MRI. *Br J Radiol* 89: 20150912, 2016.
11. Matsusue E, Fujihara Y, Matsuda E, Tokuyasu Y, Nakamoto S, Nakamura K and Ogawa T: Vanishing parotid tumors on MR imaging. *Yonago Acta Med* 61: 33-39, 2018.

12. Kato H, Kanematsu M, Watanabe H, Mizuta K and Aoki M: Salivary gland tumors of the parotid gland: CT and MR imaging findings with emphasis on intratumoral cystic components. *Neuroradiology* 56: 789-795, 2014.
13. Ji GW, Zhu FP, Xu Q, Wang K, Wu MY, Tang WW, Li XC and Wang XH: Machine-learning analysis of contrast-enhanced CT radiomics predicts recurrence of hepatocellular carcinoma after resection: A multi-institutional study. *EBioMedicine* 50: 156-165, 2019.
14. Colakoglu B, Alis D and Yergin M: Diagnostic value of machine learning-based quantitative texture analysis in differentiating benign and malignant thyroid nodules. *J Oncol* 31: 6328329, 2019.
15. Ding J, Xing Z, Jiang Z, Chen J, Pan L, Qiu J and Xing W: CT-based radiomic model predicts high grade of clear cell renal cell carcinoma. *Eur J Radiol* 103: 51-56, 2018.
16. Gessert N, Bengs M, Wittig L, Drömann D, Keck T, Schlaefer A and Ellebrecht DB: Deep transfer learning methods for colon cancer classification in confocal laser microscopy images *Int J Comput Assist Radiol Surg* 14: 1837-1845, 2019.
17. van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RG, Fillion-Robin JC, Pieper S and Aerts HJ: Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 77: e104-e107, 2017.
18. Bellingegni AD, Gruppioni E, Colazzo G, Davalli A, Sacchetti R, Guglielmelli E and Zollo L: NLR, MLP, SVM, and LDA: A comparative analysis on EMG data from people with trans-radial amputation. *J Neuroeng Rehabil* 15: 82, 2017.
19. Xiao qing D and Hong cai C: Feature extraction using a fusion method based on sub-pattern row-column two-dimensional linear discriminant analysis. *Journal of Computer Applications* 2014; 12: 3593-3598.
20. Zhu Y and Ding Y: Auto-optimized paralleled sinogram noise reduction method based on relative quality assessment for low-dose X-ray computed tomography. *J Med Imaging Health Informatics* 7: 278-282, 2017.