Supplementary Appendix 1. CT morphological characteristics

Max-diameter: The sizes of the tumors were measured by determining the maximal cross-sectional diameter.

Number: We observed the lesions of parotid tumor patients on the picture archiving and communication systems (PACS) of our hospital. If there was only one lesion, it was considered single, and if there were two or more lesions, it was multiple.

Symptoms (with or without): We assessed the symptoms based on the clinical record, including pain/tenderness or facial nerve palsy.

Location: superficial or deep lobe, defined by a dashed line delineated from the lateral edge of the mandible to the lateral border of the digastric muscle's posterior belly and retromandibular vein.

Density: The homogeneous or heterogeneous density of the lesion was assessed on the non-contrast CT.

Calcification: Calcification was defined as the CT value of the foci within the tumor is higher than 100Hu.

Cystic areas: cystic area was defined as having a CT scan attenuation of 20 HU or less.

Enhanced-peak phase: We measured CT values (in HU) on non-enhanced, arterial and venous CT scans by placing the largest possible circular region of interest within the solid portion of the lesion with caution to avoid the cystic area. The phase of the highest CT values was defined as enhanced-peak phase.

Enhancement degree: Obvious enhancement was defined as the CT value of tumor enhancement on postcontrast CT is 40Hu higher than it on non-enhanced CT scan. Slight enhancement was defined as the CT value of tumor enhancement on postcontrast CT below 20 Hu on the basis of non-enhanced CT value. Moderate enhanced CT values fell somewhere in between.

Enlarged lymph nodes (with or without): We evaluated ipsilateral lymph node metastases based on imaging features and intraoperative records. The maximal axial dimension criteria for metastatic lymph nodes on imaging were >15 mm for level I and II nodes, 8 mm for retropharyngeal nodes and 10 mm for all other node levels.

Supplementary Appendix 2. Principal Component Analysis (PCA) method

PCA is a statistical and data science technique designed to simplify the complexity inherent in high-dimensional data while retaining essential trends and patterns. The main idea behind PCA is to identify the directions in which the data varies the most. The first principal component accounts for the most variance in the data, the second principal component (uncorrelated with the first) accounts for the second most, and so on.

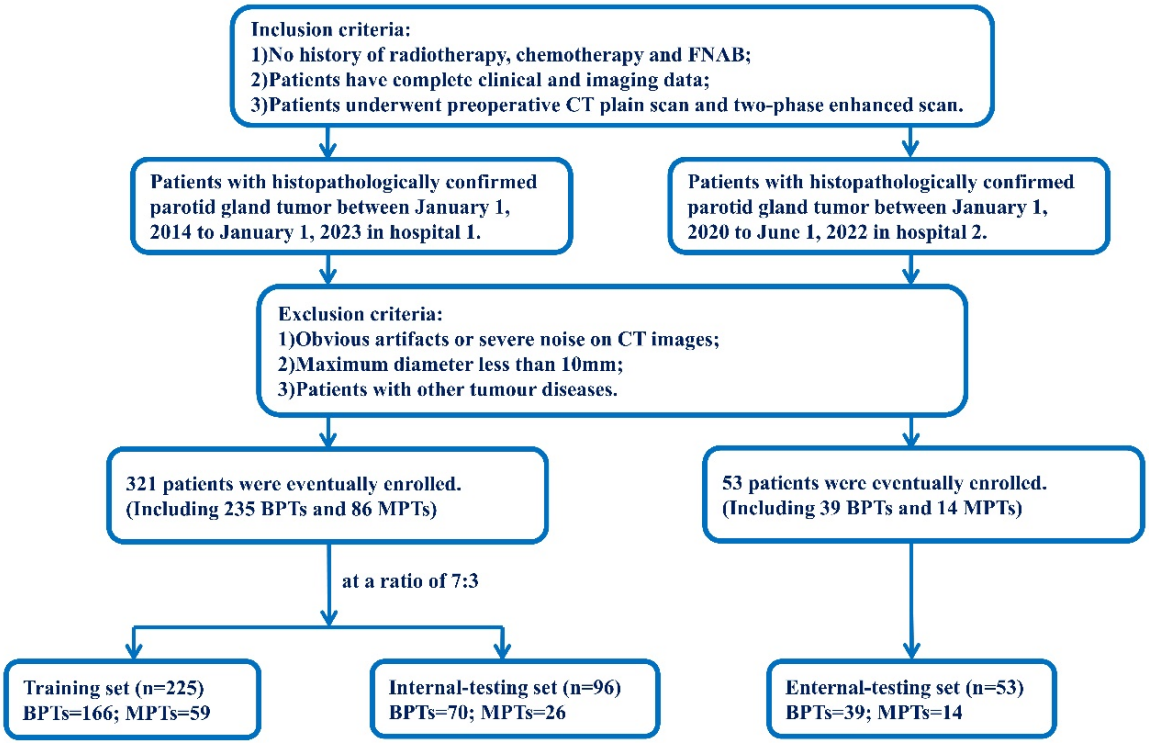

Figure S1. Flowchart for selecting the study population.

Table S1. Histopathological types and numbers of parotid tumors

BPT	Number (Center 1/Center 2)	MPT	Number (Center 1/Center 2)
Pleomorphic adenoma	128 (113/15)	Mucoepidermoid carcinoma	38 (30/8)
Warthin tumor	100 (82/18)	Adenoid cystic carcinoma	6 (4/2)
Basal cell adenoma	33 (28/5)	Acinic cell carcinoma	18 (15/3)
Myoepithelial tumor	3 (3/0)	Squamous cell carcinoma	5 (5/0)
Oncocytoma	4 (3/1)	Lymphoepithelial carcinoma	2 (2/0)
Ductal papillomas	2 (2/0)	Basal cell adenocarcinoma	2(2/0)
Lipoma	4 (4/0)	Myoepithelial carcinoma	7 (7/0)
		Salivary ductal carcinoma	9 (9/0)
		Lymphoma	6 (5/1)
		Secretory carcinoma	1 (1/0)
		Carcinoma in pleomorphic adenoma	1 (1/0)
		Undifferentiated carcinoma	3 (3/0)
		Eosinophilic cell carcinoma	2 (2/0)

Center 1, The First Affiliated Hospital of Chongqing Medical University; Center 2, The Affiliated Hospital of Southwest Medical University.

Table S2. CT protocols of the two centers

Parameters	Center 1			Center 2	
CT scanners	Discovery CT750 HD	SOMATOM Definition Flash	SOMATOM Definition Force	Philips iCT 256	United imaging uCT 550
Tube voltage	100-120 kV	100-120 kV	100 kV	100-120 kV	120 kV
Tube current	Automatic tube-current	Automatic tube-current	Automatic tube-current	Automatic tube-current	Automatic tube-current
Gantry rotation time	0.6 s	0.5 s	0.28 s	0.5 s	0.8 s
Detector collimation	64×0.625 mm	128×0.6 mm	128×0.6 mm	128×0.625 mm	64×0.6 mm
Section thickness	5 mm	5 mm	5 mm	5mm	5 mm
Section interval	5 mm	5 mm	5 mm	5 mm	5 mm
Image matrix	512×512	512×512	512×512	512×512	512×512
Contrast agent type	Omnipaque	Ioversol	Ioversol	Ioversol	Ioversol
Contrast agent concentration	300 mgI/mL	320 mgI/mL	320 mgI/mL	350 mgI/mL	350 mgI/mL
Contrast agent dosage	1.5 mL/kg	1.2 mL/kg	1.2 mL/kg	1.2 mL/kg	1.2 mL/kg
Contrast agent infused rate	3.0-4.0 mL/s	3.0-4.0 mL/s	3.0-4.0 mL/s	2.0-3.0 mL/s	2.0-3.0 mL/s
Arterial phase scan	30 s after the contrast injection	25 s after the contrast injection	25 s after the contrast injection	28 s after the contrast injection	25 s after the contrast injection
Venous phase scan	65 s after the contrast injection	60 s after the contrast injection	60 s after the contrast injection	65 s after the contrast injection	60 s after the contrast injection

Center 1, The First Affiliated Hospital of Chongqing Medical University; Center 2, The Affiliated Hospital of Southwest Medical University. CT: computed tomography.

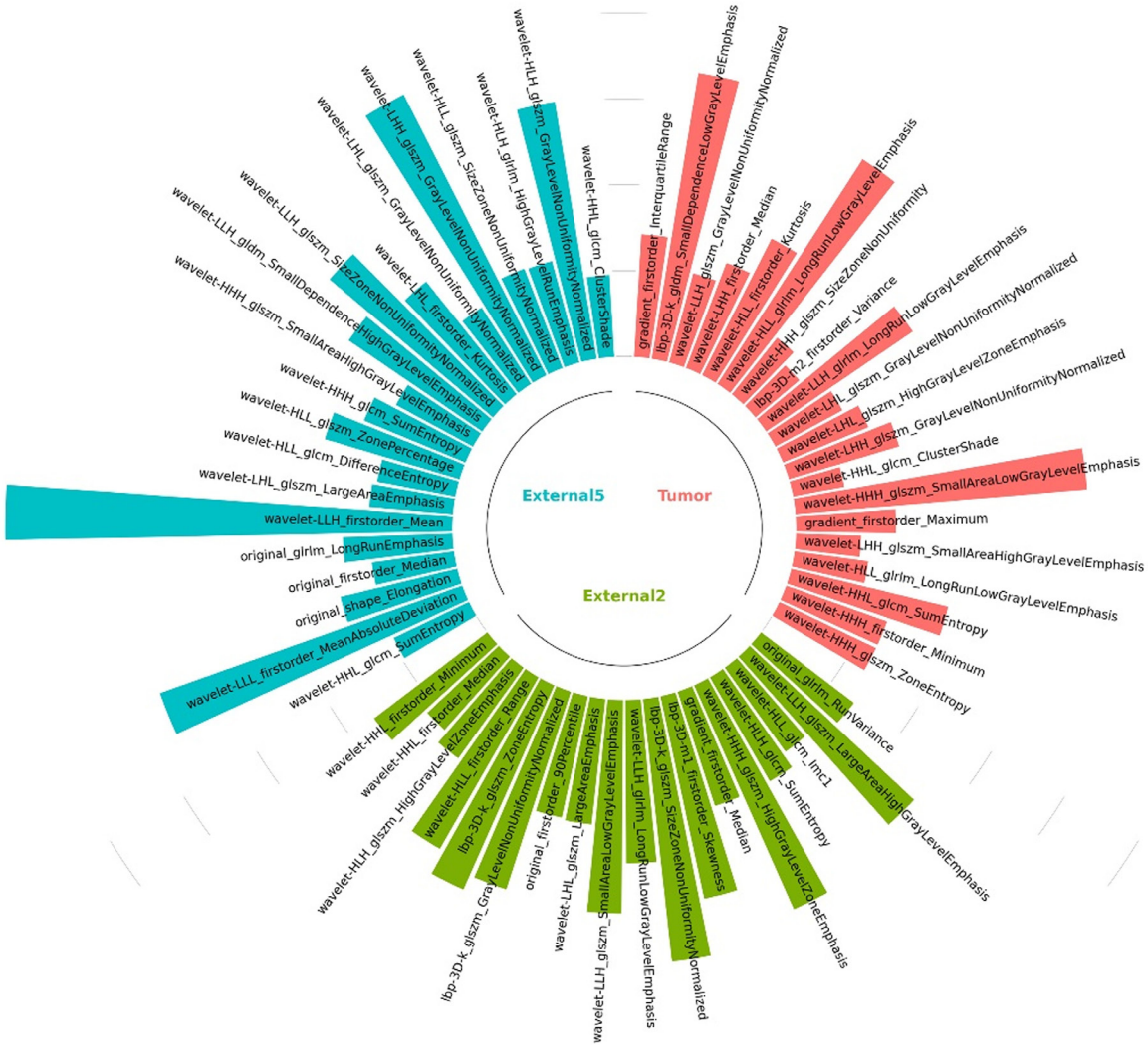

Figure S2. Radiomic feature selection results of Tumor, External2, External5, respectively.

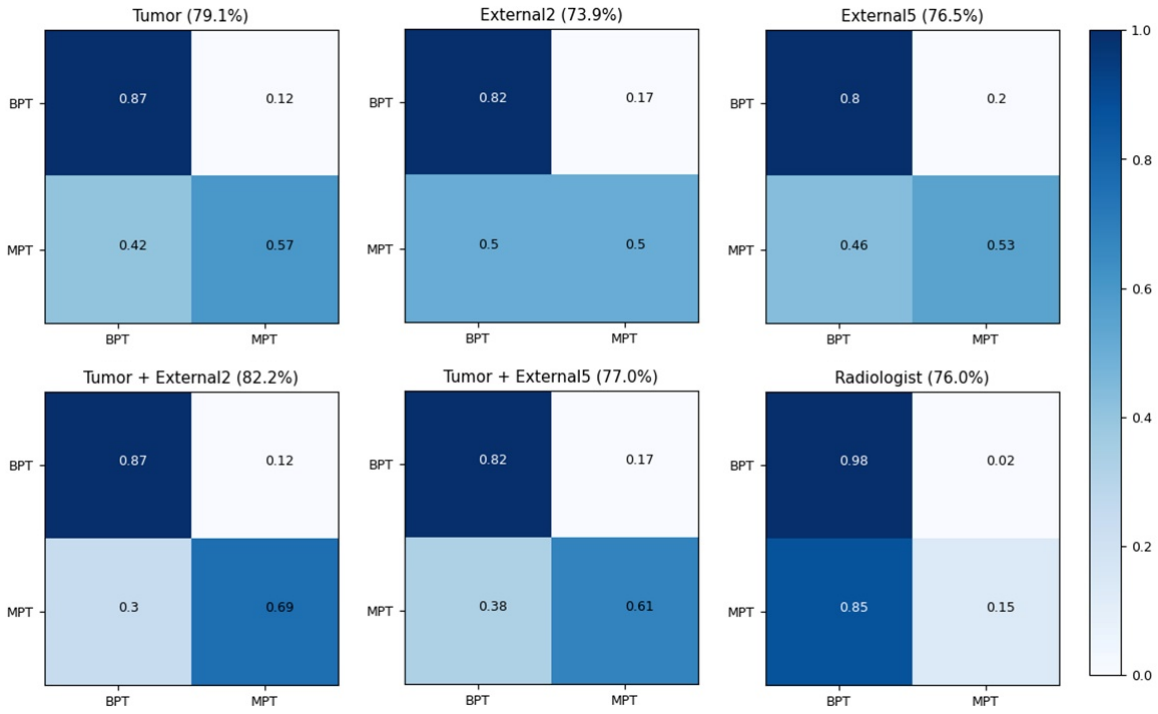

Figure S3. The 2×2 diagnostic confusion matrix analysis for different radiomic models, and the radiologist.

Table S3. The performance of Tumor + External2 radiomics using six machine learning methods

Model	AUC [95% CI]	Accuracy	Sensitivity	Specificity
Training set				
SVM	0.971 (0.962-0.981)	0.954	0.927	0.981
RF	0.965 (0.950-0.980)	0.930	0.921	0.939
LR	0.959 (0.952-0.966)	0.906	0.903	0.909
XGboost	0.988 (0.977-0.996)	1.000	1.000	1.000
DT	0.869 (0.826-0.912)	0.885	0.927	0.843
KNN	0.947 (0.927-0.967)	0.864	0.740	0.987
Internal-testing set				
SVM	0.827 (0.799-0.855)	0.822	0.871	0.692
RF	0.781 (0.747-0.815)	0.739	0.728	0.769
LR	0.801 (0.781-0.822)	0.822	0.857	0.730
XGboost	0.780 (0.748-0.813)	0.760	0.771	0.576
DT	0.728 (0.659-0.797)	0.750	0.771	0.692
KNN	0.761 (0.712-0.810)	0.677	0.642	0.769
External-testing set				
SVM	0.745 (0.701-0.785)	0.773	0.794	0.714
RF	0.721 (0.661-0.782)	0.622	0.666	0.500
LR	0.673 (0.613-0.732)	0.679	0.769	0.428
XGboost	0.684 (0.620-0.747)	0.716	0.794	0.5
DT	0.608 (0.467-0.750)	0.660	0.743	0.428
KNN	0.644 (0.554-0.733)	0.547	0.512	0.642

SVM: Support Vector Machine; RF: Random Forest; LR: Logistic Regression; DT: Decision Tree; KNN: k-Nearest Neighbor.

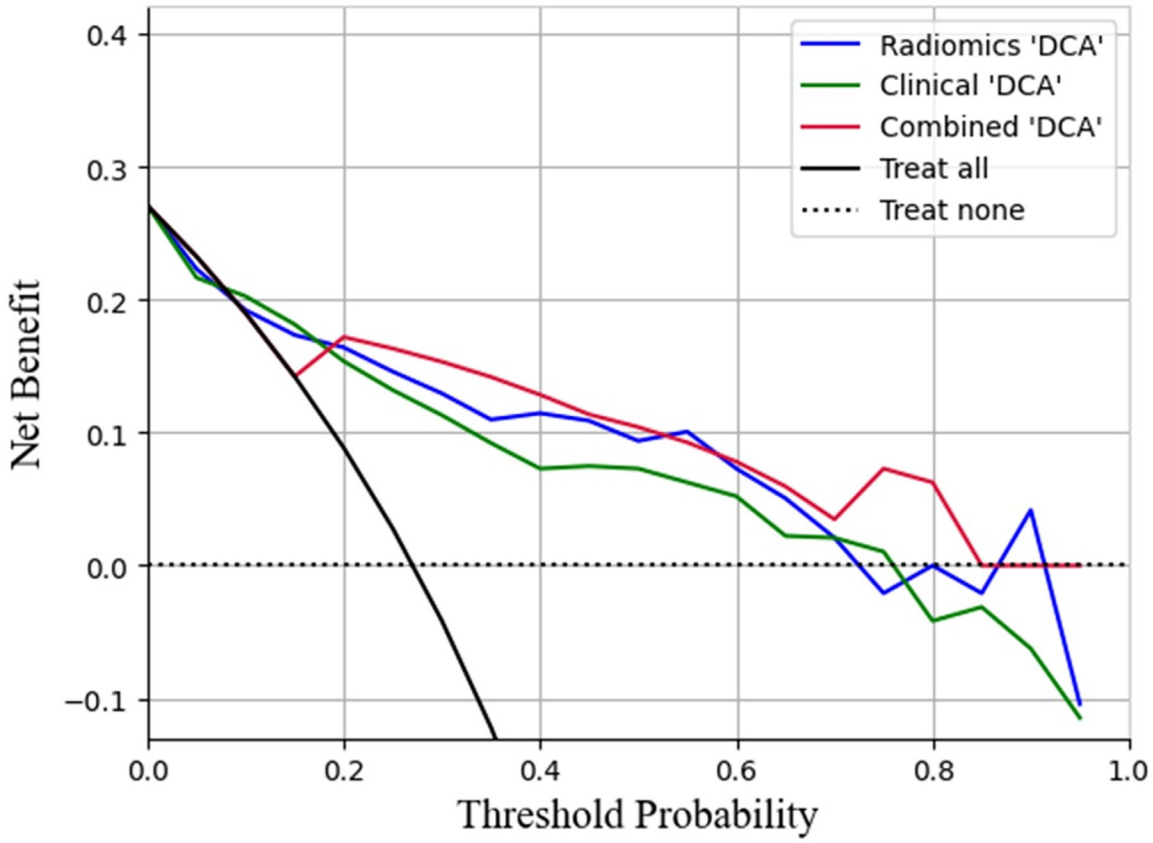

Figure S4. DCA to evaluate the clinical usefulness of the radiomic model, clinical model and combine model in classifying parotid gland tumors.