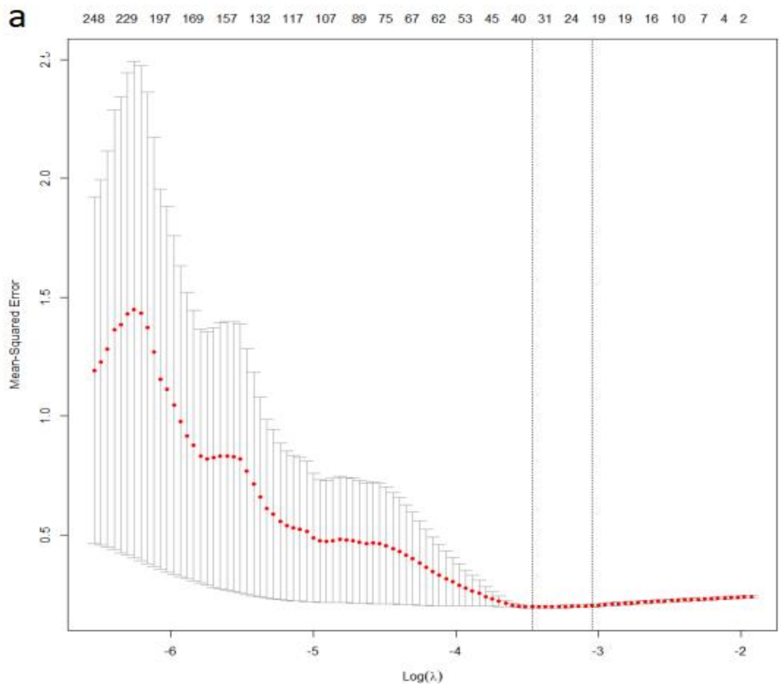

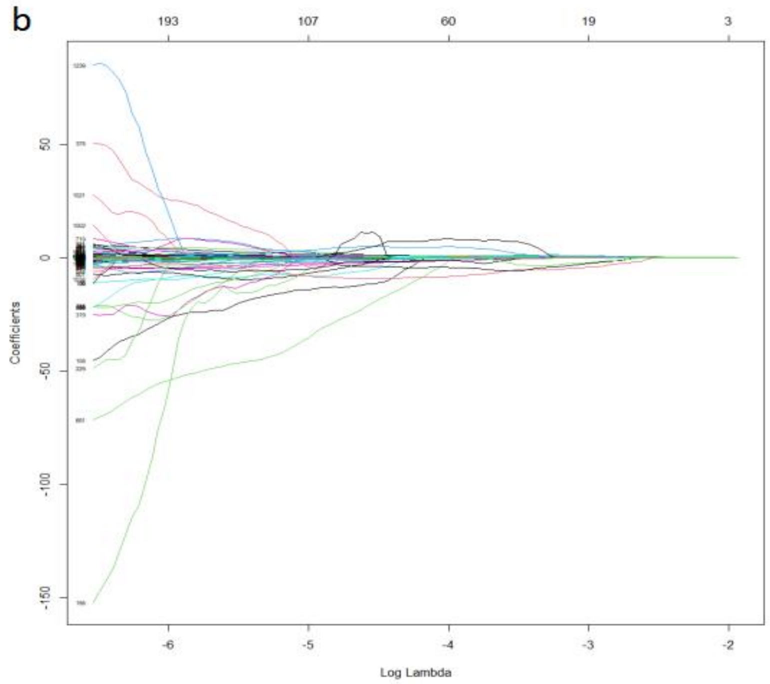

Fig. S1 LASSO coefficients of radiomics features. (a) A coefficient profile plot was produced against the log (λ) sequence. (b) feature selection was performed by using the least absolute shrinkage and selection operator (LASSO) logistic regression model

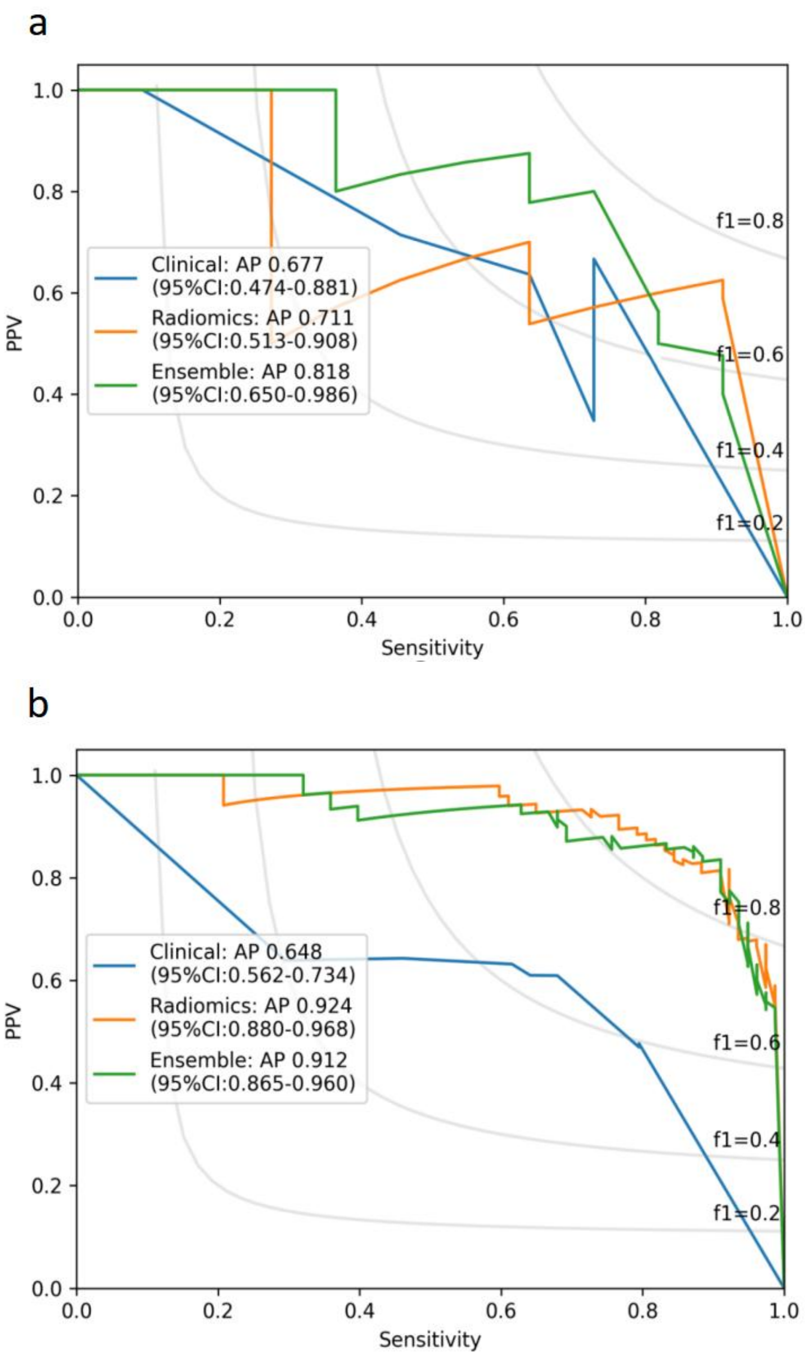

Fig. S2 The comparisons of precision-recall curves in this study. (a): In the internal validation set: AP = 0.677 for the clinical model, 0.711 for the radiomics model, and 0.818 for the ensemble model; (b): In the external testing set: AP = 0.648 for the clinical model, 0.924 for the radiomics model, and 0.912 for the ensemble model. AP: average precision

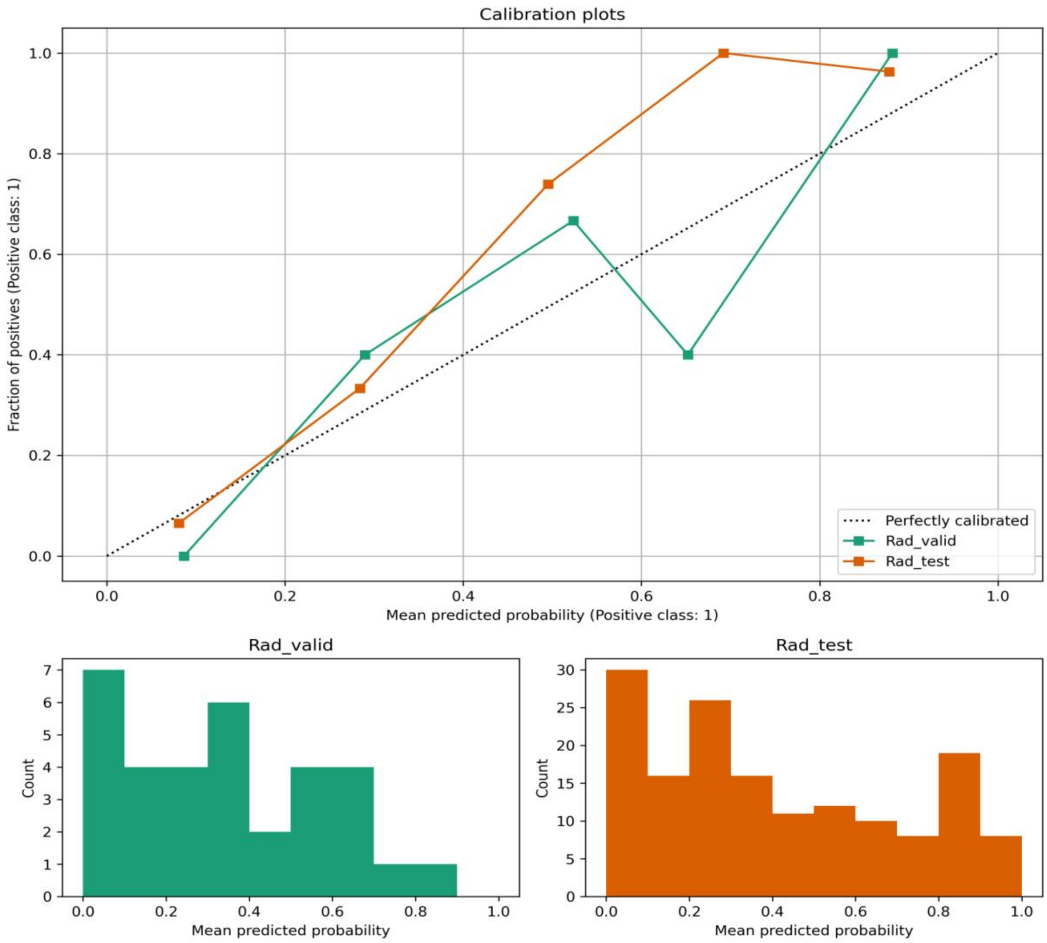

Fig. S3 Calibration curves of the radiomics model in this study. (a) Internal validation set (the green solid line and histogram); (b) External testing set (the orange solid line and histogram). P>0.05 in the Hosmer–Lemeshow test for the internal validation and external testing sets, indicating an appropriate agreement between the predicted perineural invasion status and actual observed perineural invasion status. The y-axis represents the actual perineural invasion probability. The x-axis represents the predicted perineural invasion probability. The black dotted line represents the best model with the perfect agreement. The green and orange solid lines have a closer fit to the black dotted line represents the radiomics model has a better calibration.

Table S1 The parameters of CT scanning and contrast agents used by the two centers

Parameter	Center 1	Center 2
Machine model	Philips Brilliance iCT	GE Optima CT680
Tube current	300 mA	350mA
Tube voltage	120 kV	120 kV
Slice thickness	5 mm	1.5mm
Rotation time	0.75 s	0.5s
Pitch	0.993	1.375
Detector collimation	128×0.625 mm	128×0.625 mm
Matrix	512×512	512×512
FOV	350×350 mm	400×400 mm
Contrast agent	Ioversol	Ioversol
Contrast dose	75ml	75ml
Injection rate	3.5ml/s	3.5ml/s

Table S2 The detailed search space of nine candidate classifiers.

classifier	Parameter search space
Support Vector Machine	Kernal = linear, rbf
	C = -4 – 4
	Gamma = (1, 2, 3, 'auto')
	Shrinking = True, False
Logistic Regression	Penalty = L1, L2
	C =-4 – 4
	Solver = liblinear, lbfgs
Naive Bayes	Var_smoothing = 1e-9
Multi-layer Perceptron	Alpha=10 ** -arange(1, 10)
	Activation = tanh, relu
	Solver = adam
	Learning_rate = adaptive
	Tol=1e-5
	Layers = [input, input * 10, output=2]
K-nearest Neighbors	Weights = uniform, distance
	Algorithm = auto, ball_tree, kd_tree, brute
	Leaf_size = 1, 2, 3
	N_neighbors = 2, 8
Decision Tree	Max_depth = 20-200
	Min_samples_leaf = 1,5,10,20,50,100

	Max_features = auto, sqrt
	Min_samples_split = 2, 5, 10
Random Forest	N_estimator = 200 – 1000
	Max_features = auto, sqrt
	Max_depth = 10 – 110
	Min_samples_split = 2, 5, 10
	Min_samples_leaf = 10, 20, 50, 100
	Boostrap = True, False
ADABoost	N_estimators = 200 – 1000
XGBoost	Objective = binary:logistic
	N_estimators = 200 – 1000
	Max_depth = 10