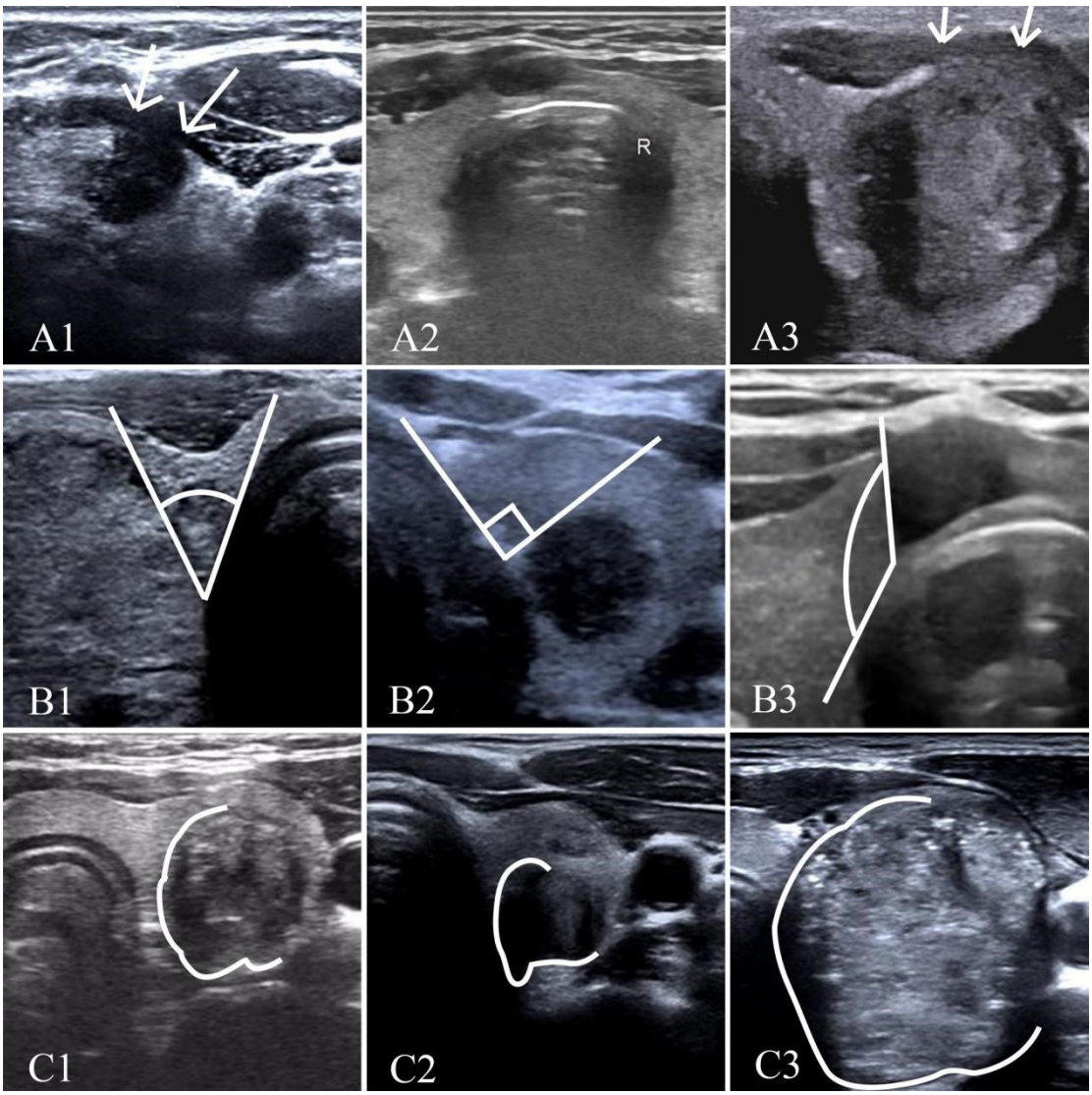

Figure S1 Ultrasound morphology of thyroid nodules: capsular, tracheal, and tracheoesophageal groove relationships. A1–A3: capsular status. A1: capsular disruption (arrow, lost capsule-nodule contact); A2: capsular bulging (capsule displaced by nodule); A3: replacement of strap muscle (indistinct muscle margins, arrow). B1–B3: tumor-trachea angle. B1: acute angle; B2: right angle; B3: obtuse angle. C1–C3: tracheoesophageal groove interaction. C1: no contact with tracheoesophageal groove; C2: contact with tracheoesophageal groove (intact capsule); C3: invasion into tracheoesophageal groove.

© AME Publishing Company.
https://dx.doi.org/10.21037/qims-2025-1-2654

Table S1 C-TIRADS classification for malignancy risk stratification of thyroid nodules

C-TIRADS category	Score	Probability of malignancy (%)	Characteristic features
Nodule-free	Unscored	0	Nodule-free
Benign nodules	–1	0	–
Benign nodules might	0	<2	–
Suspicious for malignancy nodules			
Low suspicious	1	2–10	–
Intermediate suspicious	2	10–50	–
High suspicious	3–4	50–90	–
Highly suggestive of malignancy nodules	5	>90	–
Biopsy-proven malignancy	–	–	Biopsy-proven malignancy
Positive indicators (+1 point each)	Taller-than-wide; solid composition; markedly hypoechoic; punctate echogenic foci (suspicious for microcalcifications); irregular/ill-defined margins or extrathyroidal extension		
Negative indicators (–1 point each)	hyperechoic foci with a comet-tail artifact		

C-TIRADS, Chinese thyroid imaging reporting and data system (1).

References

1. Zhou J, Yin L, Wei X, Zhang S, Song Y, Luo B, et al. 2020 Chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules: the C-TIRADS. Endocrine 2020;70:256-79.

Table S2 Performance of the combined model across 10 repeated training-validation splits

Split No.	Training cohort				Validation cohort			
	AUC (95%CI)	ACC	SEN	SPE	AUC (95%CI)	ACC	SEN	SPE
1	0.873 (0.835–0.912)	0.793	0.774	0.826	0.831 (0.763–0.899)	0.733	0.727	0.741
2	0.868 (0.829–0.907)	0.786	0.767	0.822	0.846 (0.781–0.910)	0.748	0.747	0.750
3	0.865 (0.826–0.905)	0.783	0.761	0.822	0.851 (0.788–0.915)	0.756	0.760	0.750
4	0.885 (0.826–0.904)	0.780	0.756	0.825	0.857 (0.795–0.920)	0.763	0.775	0.750
5	0.869 (0.830–0.908)	0.786	0.764	0.832	0.852 (0.787–0.917)	0.748	0.754	0.742
6	0.883 (0.846–0.919)	0.803	0.792	0.824	0.799 (0.725–0.873)	0.710	0.671	0.754
7	0.879 (0.842–0.916)	0.796	0.790	0.807	0.810 (0.788–0.883)	0.725	0.688	0.779
8	0.875 (0.837–0.913)	0.796	0.787	0.810	0.828 (0.760–0.896)	0.725	0.702	0.766
9	0.878 (0.849–0.915)	0.790	0.789	0.790	0.825 (0.757–0.896)	0.740	0.707	0.821
10	0.855 (0.814–0.896)	0.773	0.739	0.842	0.893 (0.841–0.945)	0.779	0.826	0.726
Mean ± SD	0.873±0.009 (0.833±0.011–0.910±0.007)	0.789±0.009	0.772±0.018	0.820±0.014	0.839±0.027 (0.779±0.030–0.893±0.044)	0.743±0.020	0.736±0.046	0.758±0.026

ACC, accuracy; AUC, area under the curve; CI, confidence interval; SD, standard deviation; SEN, sensitivity; SPE, specificity.

© AME Publishing Company.
https://dx.doi.org/10.21037/qims-2025-1-2654

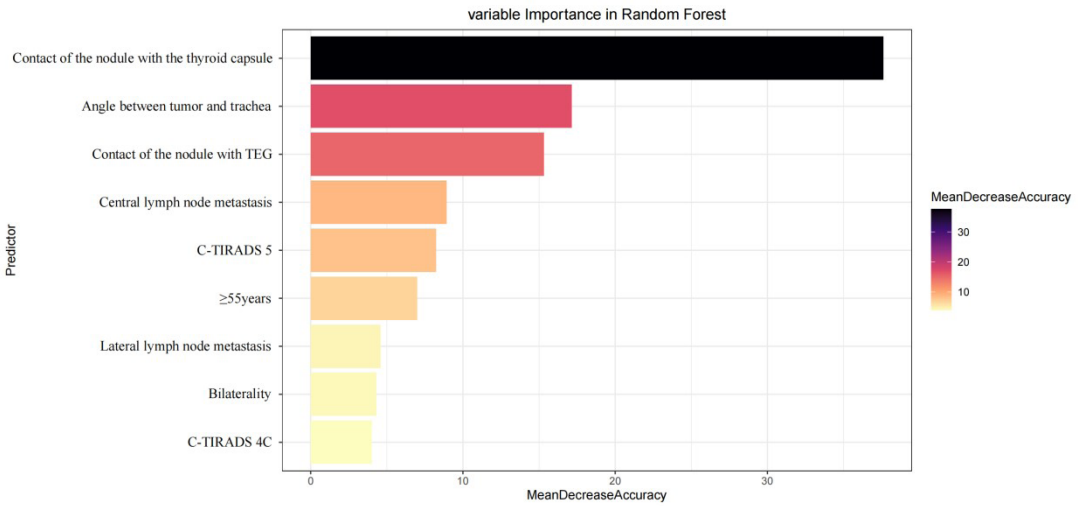

Figure S2 Random forest variable importance (Mean Decrease Accuracy) for predicting extrathyroidal extension (ETE). The Mean Decrease Accuracy metric quantifies the reduction in model prediction accuracy when a feature is removed. A higher value indicates greater importance of the feature in predicting ETE. The top three most impactful features were Contact of the nodule with the thyroid capsule, Angle between tumor and trachea, and Contact of the nodule with TEG (tracheoesophageal groove).

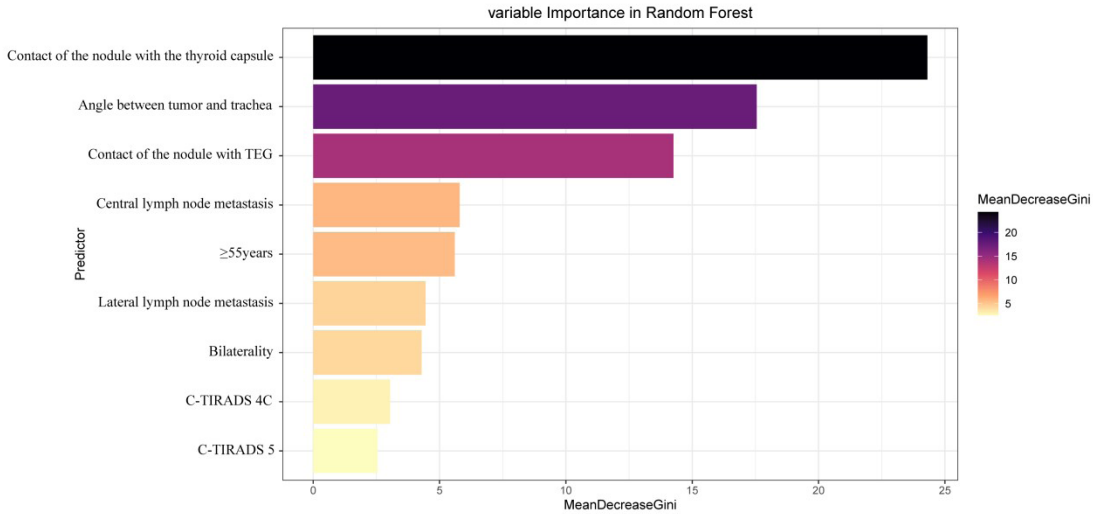

Figure S3 Random forest variable importance (Mean Decrease Gini) for predicting extrathyroidal extension (ETE). Mean Decrease Gini measures the reduction in node impurity when a feature is removed, reflecting its contribution to splitting the decision tree. Higher values indicate greater feature importance in predicting ETE, with contact of the nodule with the thyroid capsule being the most impactful predictor.

© AME Publishing Company.
https://dx.doi.org/10.21037/qims-2025-1-2654

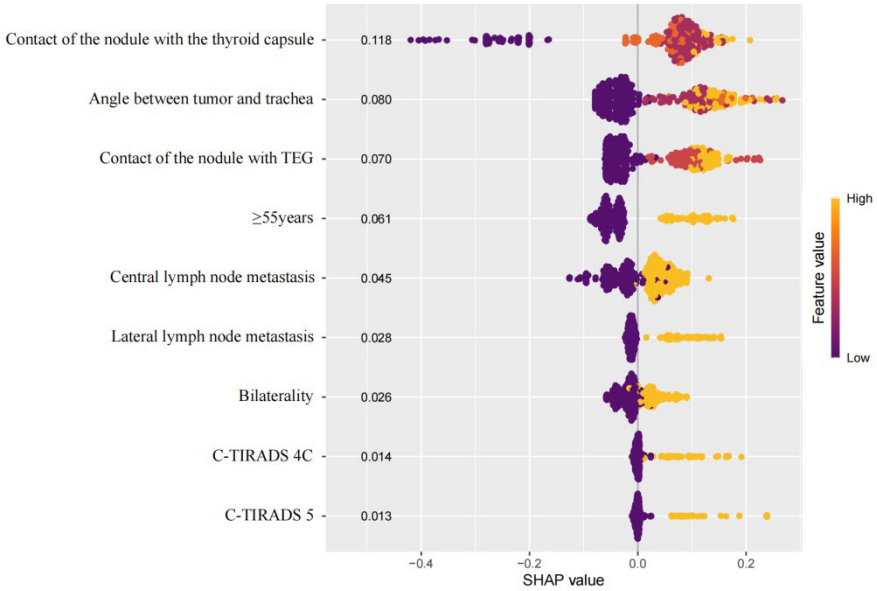

Figure S4 SHAP summary plot of feature contributions to extrathyroidal extension (ETE) Prediction. The SHAP summary plot illustrates each feature's contribution to ETE prediction. The horizontal axis shows the SHAP value (positive values increase ETE risk, negative values decrease it), and the color gradient represents the feature's magnitude (yellow = high value, purple = low value). Contact of the nodule with the thyroid capsule shows the strongest positive association with ETE.

© AME Publishing Company.
https://dx.doi.org/10.21037/qims-2025-1-2654