

Detailed elaboration of the deep learning and image processing pipeline

To ensure methodological transparency for readers interested in the computational background of our computerized tongue image analysis (CTIA), the specific computational architectures and algorithms deployed in our automated pipeline are elaborated below. All semantic segmentation, image processing, and quantitative feature extraction procedures were implemented and executed in Python (version 3.10) using the PyTorch (version 2.1.2) deep learning framework, with OpenCV (version 4.12.0) used for computer vision operations.

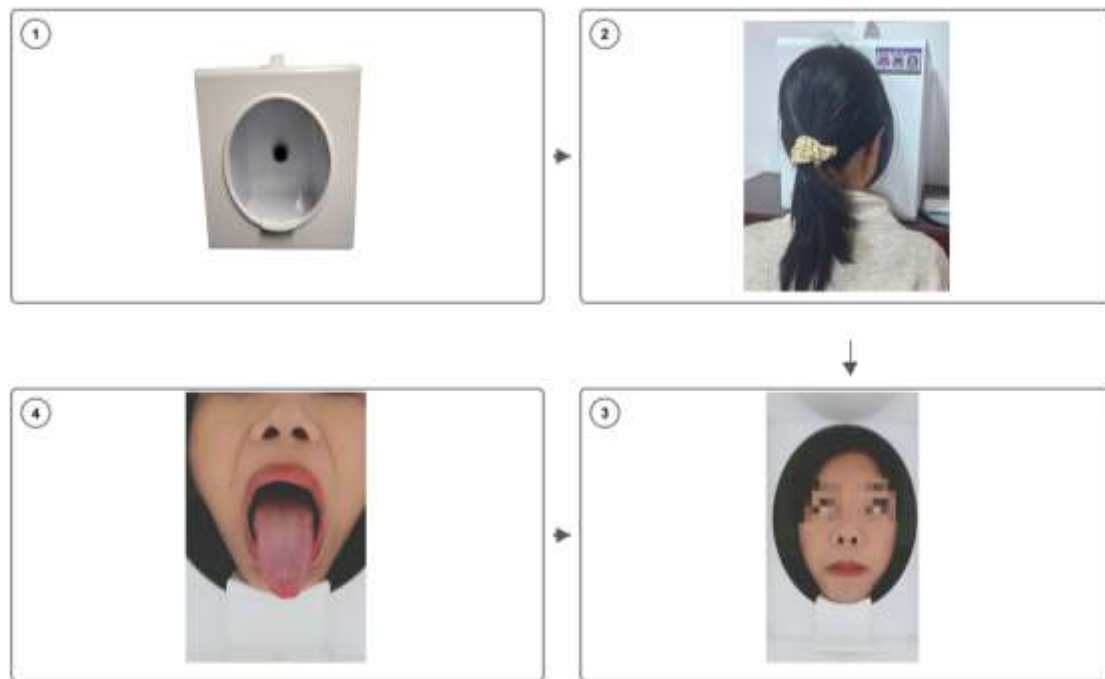
Automated quality control and target localization: To isolate the diagnostic target from high-noise backgrounds, a YOLO-series (You Only Look Once) object detection network was deployed. This module dynamically cropped the region of interest (ROI) and effectively eliminated peripheral visual interference (e.g., teeth, lips, and mucosal shadows), preventing false-positive identification caused by spectral similarities between lips and tongue mucosa.

Semantic segmentation of the pure tongue body: To extract precise morphological parameters and provide a clean substrate for color extraction, we employed an ICNet-based (Image Cascade Network) semantic segmentation model optimized for high-resolution images. By processing the image across multiple spatial resolutions, the network preserved raw high-frequency edge details, generating a highly accurate, binarized pure tongue mask.

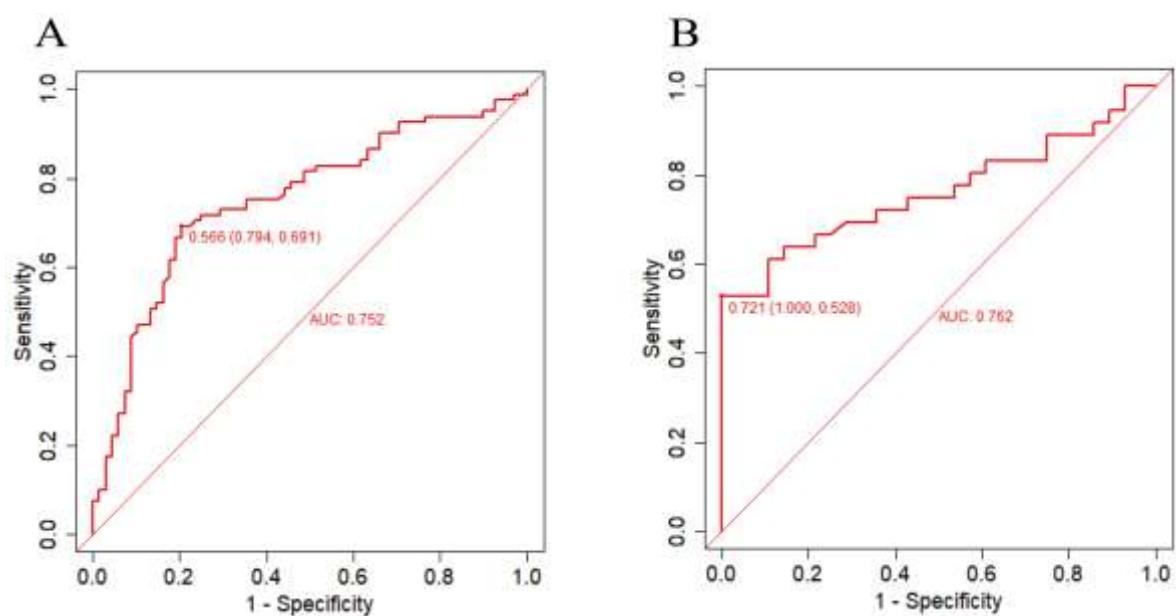
Unsupervised clustering for physical decoupling: Tongue coating is often unevenly distributed over the tongue body. To eliminate subjective truncation errors caused by preset hard color thresholds, an unsupervised k -means clustering algorithm was introduced. By mapping valid pixels into advanced

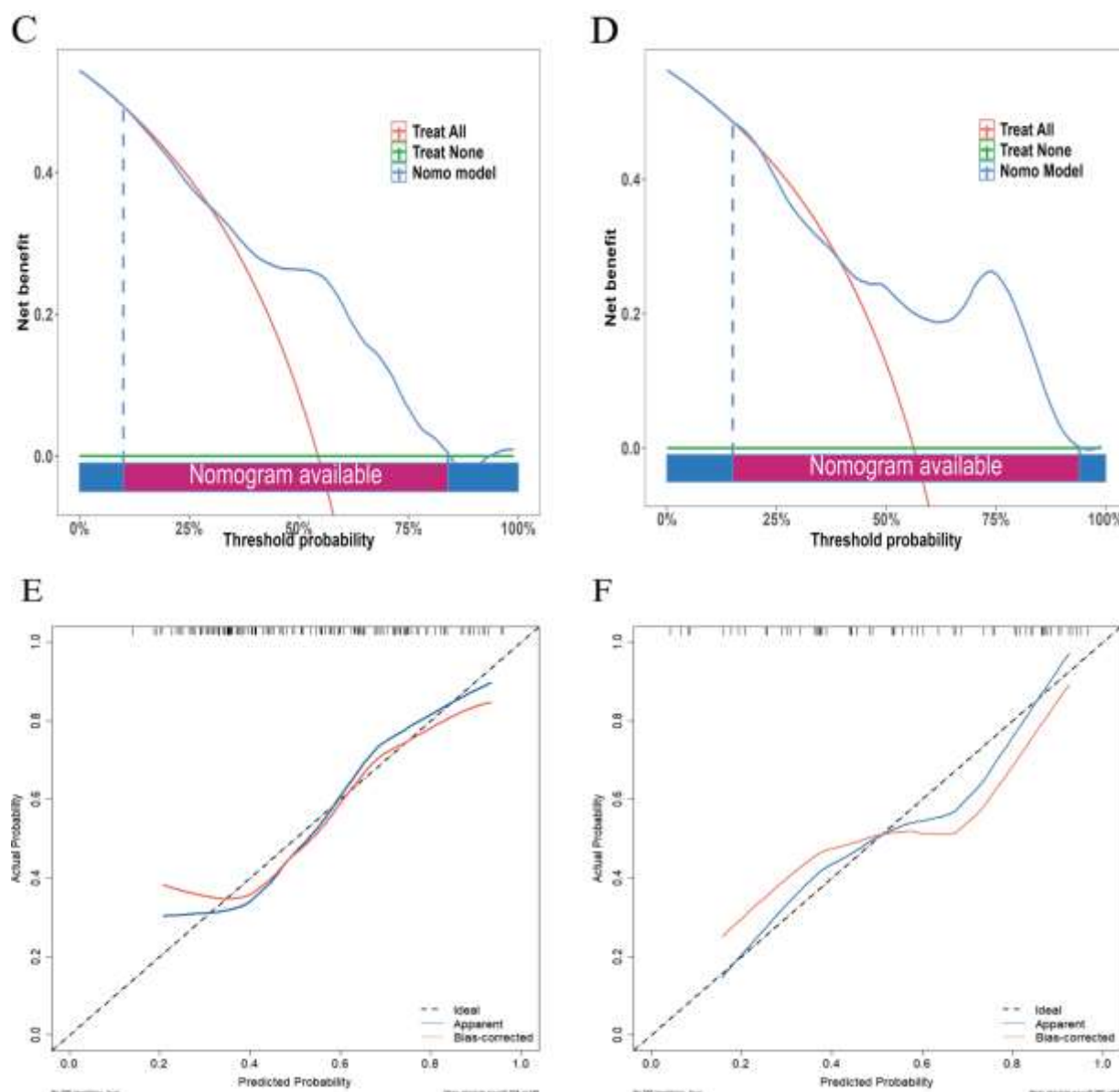
multidimensional color spaces, the algorithm adaptively decoupled the pixels into mutually independent "tongue body" and "tongue coating" masks.

Feature extraction and representation: The segmented regions were partitioned into five subregions (anterior, center, posterior, and bilateral margins). Key tongue-surface characteristics were identified using a multitask learning network with MobileNetV3 as the backbone. Furthermore, a capsule-network framework with residual blocks was utilized to enhance feature representation by integrating both shallow and deep spatial information. This architecture effectively handled the complex, concurrent manifestations of tongue phenotypes, ultimately outputting the 157 structured quantitative features used for diagnostic modeling.



Supplementary Figure 1 Visual demonstration of the standardized tongue image acquisition process. The portable standardized tongue-imaging device (MQ-SXZN-ZA1) includes a controlled illumination chamber to minimize ambient light interference. The accompanying photographs show participant positioning and standardized image acquisition in the clinical setting. The participant's eyes were obscured to protect privacy.





Supplementary Figure 2 Performance evaluation of the tongue image-clinical nomogram in the APCS score ≥ 4 high-risk subgroup. A: ROC curve in the high-risk training cohort; B: ROC curve in the high-risk validation cohort; C: DCA in the high-risk training cohort; D: DCA in the high-risk validation cohort; E: Bootstrap-corrected calibration curve (500 resamples) in the high-risk training cohort; F: Bootstrap-corrected calibration curve (500 resamples) in the high-risk validation cohort. APCS: Asia-Pacific Colorectal Screening; ROC: Receiver operating characteristic; AUC: Area under the curve; DCA: Decision curve analysis.

Supplementary Table 1 Variable dictionary for tongue-image features used in model development and figures

Raw code	Manuscript label	Figure label (short form)
Texture_Thickness	Tongue-coating thickness	Coating thickness
Tongue_Center_a	Central tongue redness (CIELAB a* channel)	Central tongue redness (a*)
Tongue_Tip_V	Anterior tongue brightness (HSV V channel)	Anterior tongue brightness (V)
TB.5.Divisions_Center_V	Central tongue-body brightness (HSV V channel)	Central TB brightness (V)
TB.5.Divisions_Root_H	Posterior tongue-body hue (HSV H channel)	Posterior TB hue (H)
TC.5.Divisions_Root_b	Posterior tongue-coating yellowness (CIELAB b* channel)	Posterior TC yellowness (b*)
TC.5.Divisions_Tip_a	Anterior tongue-coating redness (CIELAB a* channel)	Anterior TC redness (a*)

TB: Tongue body; TC: Tongue coating; HSV, Hue saturation value; CIELAB: CIE L*a*b* color space.

Supplementary Table 2 Tongue-image features significantly associated with advanced adenoma in Univariate analysis (training cohort)

Characteristics	B	SE	OR	CI	Z	P	P_raw
Tooth. Marks	0.435	0.19436	1.546	1.546(1.056-2.262)	2.241	<0.05	0.025
Cracks_num	-0.796	0.27604	0.451	0.451(0.263-0.775)	-2.885	<0.01	0.004
Texture_Thickness	1.306	0.18917	3.691	3.691(2.548-5.348)	6.904	<0.001	0

Tongue. Body_a	-0.078	0.02527	0.925	0.925(0.88-0.972)	-3.103	<0.01	0.002
Tongue. Body_R	-0.038	0.01142	0.963	0.963(0.942-0.985)	-3.316	<0.001	0.001
Tongue. Body_S	-3.391	1.57437	0.034	0.034(0.002-0.737)	-2.154	<0.05	0.031
Tongue. Body_V	-5.192	2.13175	0.006	0.006(0-0.363)	-2.436	<0.05	0.015
Tongue. Coating_a	-0.148	0.03031	0.862	0.862(0.813-0.915)	-4.89	<0.001	0
Tongue. Coating_G	0.016	0.00795	1.016	1.016(1-1.032)	2.022	<0.05	0.043
Tongue. Coating_S	-5.142	1.58349	0.006	0.006(0-0.13)	-3.248	<0.01	0.001
Tongue_Center_a	-0.14	0.02655	0.869	0.869(0.825-0.915)	-5.291	<0.001	0
Tongue_Center_R	-0.024	0.01015	0.976	0.976(0.957-0.996)	-2.375	<0.05	0.018
Tongue_Center_H	-0.002	0.00056	0.998	0.998(0.997-0.999)	-3.231	<0.01	0.001
Tongue_Center_S	-4.828	1.4967	0.008	0.008(0-0.15)	-3.226	<0.01	0.001
Tongue_Left_a	-0.092	0.02767	0.912	0.912(0.864-0.963)	-3.319	<0.001	0.001
Tongue_Left_S	-3.157	1.58553	0.043	0.043(0.002-0.952)	-1.991	<0.05	0.046
Tongue_Root._a	-0.055	0.02545	0.946	0.946(0.9-0.995)	-2.171	<0.05	0.03
Tongue_Root._b	0.155	0.041	1.168	1.168(1.078-1.266)	3.792	<0.001	0
Tongue_Root._H	-0.001	0.00055	0.999	0.999(0.998-1)	-1.984	<0.05	0.047

Tongue_Right_a	-0.064	0.02658	0.938	0.938(0.89-0.988)	-2.417	<0.05	0.016
Tongue_Tip_a	-0.076	0.02142	0.927	0.927(0.889-0.967)	-3.544	<0.001	0
Tongue_Tip_R	-0.042	0.01061	0.959	0.959(0.939-0.979)	-3.939	<0.001	0
Tongue_Tip_V	-11.738	2.91467	0	0(0-0.002)	-4.027	<0.001	0
TB.5.Divisions_Center_a	-0.088	0.02395	0.915	0.915(0.873-0.959)	-3.694	<0.001	0
TB.5.Divisions_Center_R	-0.035	0.01043	0.966	0.966(0.946-0.986)	-3.315	<0.001	0.001
TB.5.Divisions_Center_V	-6.867	1.8233	0.001	0.001(0-0.037)	-3.766	<0.001	0
TB.5.Divisions_Left_a	-0.076	0.02665	0.927	0.927(0.88-0.976)	-2.853	<0.01	0.004
TB.5.Divisions_Left_R	-0.024	0.00941	0.976	0.976(0.959-0.995)	-2.544	<0.05	0.011
TB.5.Divisions_Left_V	-3.227	1.60108	0.04	0.04(0.002-0.915)	-2.015	<0.05	0.044
TB.5.Divisions_Root_H	-0.003	0.00082	0.997	0.997(0.995-0.999)	-3.531	<0.001	0
TB.5.Divisions_Right_a	-0.065	0.02599	0.937	0.937(0.89-0.986)	-2.513	<0.05	0.012
TB.5.Divisions_Right_R	-0.023	0.00909	0.977	0.977(0.96-0.995)	-2.559	<0.05	0.011
TB.5.Divisions_Tip_a	-0.079	0.02207	0.924	0.924(0.885-0.965)	-3.584	<0.001	0
TB.5.Divisions_Tip_R	-0.038	0.01049	0.962	0.962(0.943-0.982)	-3.668	<0.001	0
TB.5.Divisions_Tip_S	-2.895	1.47414	0.055	0.055(0.003-0.994)	-1.964	<0.05	0.05
TB.5.Divisions_Tip_V	-10.88	2.88146	0	0(0-0.005)	-3.776	<0.001	0

TC.5.Divisions_Center_a	-0.143	0.02716	0.866	0.866(0.822-0.914)	-5.277	<0.001	0
TC.5.Divisions_Center_b	0.074	0.03145	1.077	1.077(1.013-1.146)	2.365	<0.05	0.018
TC.5.Divisions_Center_R	-0.019	0.0096	0.981	0.981(0.962-0.999)	-2.027	<0.05	0.043
TC.5.Divisions_Center_H	-0.002	0.00051	0.998	0.998(0.997-0.999)	-3.729	<0.001	0
TC.5.Divisions_Center_S	-4.057	1.55397	0.017	0.017(0.001-0.364)	-2.611	<0.01	0.009
TC.5.Divisions_Left_a	-0.126	0.02748	0.882	0.882(0.835-0.93)	-4.585	<0.001	0
TC.5.Divisions_Left_S	-4.655	1.50729	0.01	0.01(0-0.182)	-3.089	<0.01	0.002
TC.5.Divisions_Root_a	-0.064	0.02578	0.938	0.938(0.892-0.986)	-2.492	<0.05	0.013
TC.5.Divisions_Root_b	0.156	0.03971	1.169	1.169(1.082-1.264)	3.94	<0.001	0
TC.5.Divisions_Root_H	-0.001	0.00055	0.999	0.999(0.998-1)	-2.432	<0.05	0.015
TC.5.Divisions_Right_a	-0.125	0.02784	0.882	0.882(0.835-0.932)	-4.503	<0.001	0
TC.5.Divisions_Right_S	-5.475	1.47975	0.004	0.004(0-0.076)	-3.7	<0.001	0
TC.5.Divisions_Tip_a	-0.124	0.02655	0.883	0.883(0.838-0.93)	-4.675	<0.001	0
TC.5.Divisions_Tip_R	-0.026	0.00854	0.974	0.974(0.958-0.99)	-3.097	<0.01	0.002
TC.5.Divisions_Tip_S	-3.922	1.47274	0.02	0.02(0.001-0.355)	-2.663	<0.01	0.008
TC.5.Divisions_Tip_V	-3.901	1.88596	0.02	0.02(0.001-0.815)	-2.069	<0.05	0.039

Among 157 tongue-image features, 52 features with $P < 0.05$ in Univariate analysis were retained as candidates for subsequent

LASSO selection. SE: Standard error; OR: Odds ratio; CI: Confidence interval; IQR: Interquartile range; LASSO: Least absolute shrinkage and selection operator

Supplementary Table 3 Baseline characteristics of participants with and without advanced adenoma in the APCS score ≥ 4 high-risk subgroup (training and validation cohorts)

Variable	Training cohort(<i>n</i> =149)			Validation cohort (<i>n</i> =64)			<i>P</i> ^a
	NAA (<i>n</i> =68)	AA (<i>n</i> =81)	<i>P</i>	NAA (<i>n</i> =28)	AA (<i>n</i> =36)	<i>P</i>	
Sex, <i>n</i> (%)			0.916			0.124	0.891
Female	8 (11.8%)	8 (9.9%)		6 (21.4%)	2 (5.6%)		
Male	60 (88.2%)	73 (90.1%)		22 (78.6%)	34 (94.4%)		
Age (yr), median (IQR)	55.0 (42.0–63.0)	60.0 (50.0–67.0)	0.243	60.5 (54.0–68.5)	61.0 (55.8–68.2)	0.855	0.125
BMI, kg/m ² , mean (SD)	24.3 (3.3)	24.0 (2.9)	0.607	24.2 (4.0)	23.8 (2.7)	0.669	0.670
Smoking status, <i>n</i> (%)			0.394			0.362	1.000

Never	22 (32.4%)	20 (24.7%)		10 (35.7%)	8 (22.2%)		
Ever (former/current)	46 (67.6%)	61 (75.3%)		18 (64.3%)	28 (77.8%)		
Family history of colorectal cancer, <i>n</i> (%)			0.978			0.362	0.661
No	51 (75.0%)	62 (76.5%)		18 (64.3%)	28 (77.8%)		
Yes	17 (25.0%)	19 (23.5%)		10 (35.7%)	8 (22.2%)		
APCS score, <i>n</i> (%)			0.652			1.000	0.951
4	54 (79.4%)	60 (74.1%)		21 (75.0%)	26 (72.2%)		
5	10 (14.7%)	14 (17.3%)		5 (17.9%)	7 (19.4%)		
6	3 (4.4%)	6 (7.4%)		2 (7.1%)	2 (5.6%)		
7	1 (1.5%)	1 (1.2%)		0 (0.0%)	1 (2.8%)		
M-APCS score, median (IQR)	4.0 [3.0;4.0]	4.0 [3.0;4.0]	0.119	4.0 [3.0;4.0]	4.0 [3.0;5.0]	0.329	0.521

Categorical variables are presented as *n* (%). Age and M-APCS score are presented as median (IQR), whereas BMI is presented as mean (SD). Within-cohort *P* values compare AA versus NAA in the training and validation cohorts. BMI was compared using Student's *t* test; other continuous variables were compared using the Mann-Whitney *U* test, and categorical variables were

compared using the χ^2 test or Fisher's exact test. P values compare baseline characteristics between the training and validation cohorts within the APCS score ≥ 4 high-risk subgroup. APCS: Asia-Pacific Colorectal Screening score; M-APCS: modified APCS score; AA: Advanced adenoma; NAA: Nonadvanced adenoma; IQR: Interquartile range; SD: Standard deviation; BMI: Body mass index.