SUPPLEMENTARY MATERIAL

Data pre-processing

In order to prepare good-quality colonoscopy images for the CAD system, the following pre-processing were performed, which helped to clean and standardize the data, allowing for more accurate and reliable results from the CAD system: (1) Image resizing: All of the images were resized to a uniform size of 1280*1280 since the input to the system is required to be the same dimensions; (2) Data normalization: The pixel values in all images were normalized to a range between 0 and 1 to improve the stability and performance of the model; and (3) Data augmentation: To overcome the size limitations of training data and improve model generalization capability, data augmentation techniques were applied to the images in the train data, such as random flipping, rotation, and scaling of the images.

Polyp recognition and classification

In this study, we used the YOLOv7 model for polyp **recognition** and classification, which is an improvement over the previous YOLO models with increased accuracy and faster processing times. The YOLOv7 architecture consists of multiple scales of object detection, from small to large objects, and each scale has its own set of anchors to determine the size of predicted bounding boxes. For each detected object, besides the bounding box which contains the object, the model also can predict its category.

In order to simultaneously detect and classify objects, the detection loss (L_{IOU}), the classification loss (L_{CE}), and the coordinate loss (L_{CD}) were adopted during the training of the YOLOv7 model. Subsequently, the total loss L_{total} of the model can be computed as Eq. 1.

$$L_{total} = L_{IOU} + L_{CE} + L_{CD} \quad (1)$$

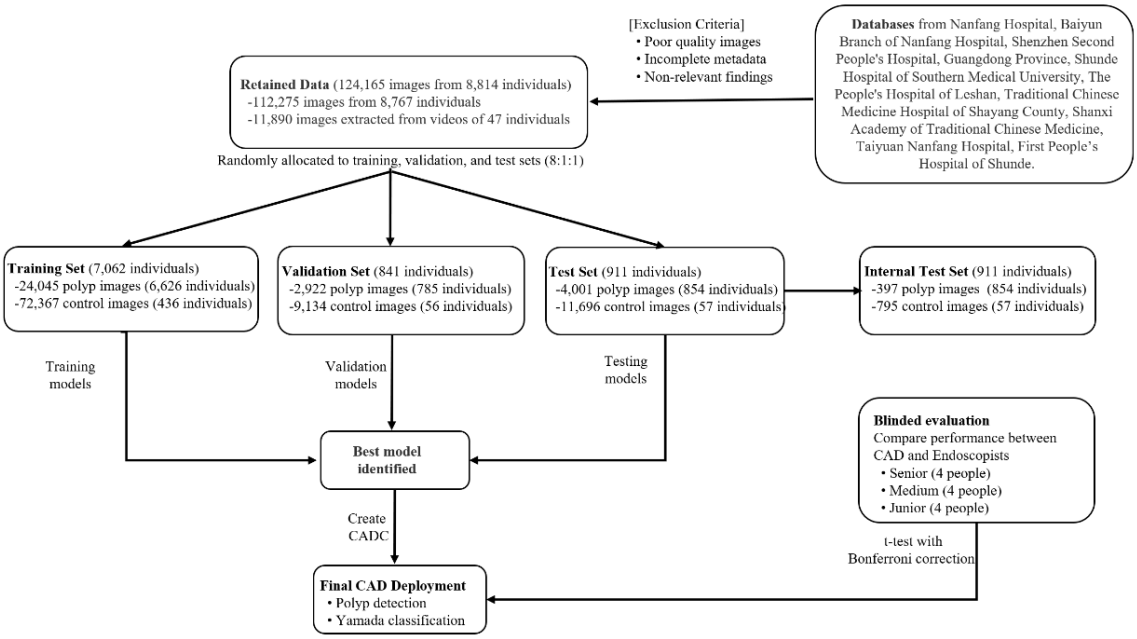
The detection loss is responsible for ensuring that the model is able to accurately predict the presence of objects in an image. It is calculated based on the Intersection over Union (IoU) between the predicted bounding boxes and

the ground-truth bounding boxes. As defined in Eq.2, the IoU measures the overlap between the two bounding boxes, normalized by the area of their union, where a value of 1 indicates a perfect overlap and a value of 0 indicates no overlap.

$$\mathrm{IoU}(A, B) = \frac{A \cap B}{A \cup B} \quad (2)$$

where A is the ground-truth bounding boxes and B is the predicted bounding box.

The classification loss is used to make the model accurately classify the detected objects in the image. This loss is calculated based on the cross-entropy between the probabilities of the predicted categories and the ground-truth categories, which measures the difference between the predicted and actual class distributions. The coordinate loss is used to help the model accurately predict the location of the objects in the image. This loss is calculated based on the differences between the predicted and ground-truth bounding box coordinates.



Supplementary Figure 1 The study design of this study.

Supplementary Table 1 Clinical characteristics of patients in training, validation and test set

Clinical characteristics	Training set	Validation set	Test set
Gender			
women	2675 (38.1%)	299 (35.8%)	352 (38.9%)
men	4350 (61.9%)	537 (64.2%)	554 (61.1%)
Age			
0-10	2 (0.0%)	9 (0.0%)	0 (0.0%)
11-20	16 (0.2%)	1 (0.1%)	4 (0.4%)
21-30	168 (2.4%)	26 (3.1%)	18 (2.0%)
31-40	574 (8.2%)	85 (10.2%)	86 (9.5%)
41-50	1520 (21.6%)	179 (21.4%)	195 (21.5%)
51-60	1980 (28.2%)	216 (25.8%)	260 (28.7%)
61-70	1842 (26.2%)	237 (28.3%)	255 (28.1%)
71-80	790 (11.2%)	77 (9.2%)	79 (8.7%)
81-90	129 (1.8%)	15 (1.8%)	9 (1.0%)
91-100	4 (0.1%)	0 (0.0%)	0 (0.0%)
Classes			
Normal (Patient/Image)	399/4329	51/584	52/608
Class I (Patient/Image)	3208/49527	375/5886	471/8746
Class II (Patient/Image)	4505/60291	536/7398	633/11229
Class III (Patient/Image)	1805/24569	208/3145	258/4619
Class IV (Patient/Image)	1279/17441	135/1853	171/3370

Supplementary Table 2 The performance of computer-aided diagnosis and endoscopists on polyp detection (image-based), %

Group	PPV	Sensitivity	NPV	Specificity	Accuracy
Experts	96.0 ± 3.0	96.4 ± 1.5[b]	98.2 ± 0.7[b]	97.9 ± 1.6	97.4 ± 0.7[a]
Seniors	94.9 ± 2.7[b]	97.3 ± 2.5	98.7 ± 1.2	97.3 ± 1.5[b]	97.3 ± 0.6[a]
Juniors	93.3 ± 3.5[b]	96.1 ± 4.5	98.1 ± 2.2	96.5 ± 1.9[b]	96.4 ± 0.3[a]
Average	94.8 ± 3.0[a]	96.6 ± 2.9[b]	98.3 ± 1.4[b]	97.2 ± 1.7[a]	97.0 ± 0.7[a]
AI	99.0	98.5	99.2	99.5	99.2

[a]$P < 0.01$.

[b]$P < 0.05$.

Supplementary Table 3 The performance of CAD and endoscopists on Yamada classification (image-based), %

Group	PPV	Sensitivity	NPV	Specificity	Accuracy
Experts	78.2 ± 2.0[a]	73.9 ± 4.4	97.6 ± 0.2[a]	97.7 ± 0.5[b]	95.9 ± 0.4[a]
Seniors	71.4 ± 2.9[a]	73.4 ± 1.7[a]	97.4 ± 0.3[a]	97.2 ± 0.8[b]	95.2 ± 0.9[a]
Juniors	63.8 ± 4.2[a]	69.7 ± 7.5[b]	96.6 ± 0.8[a]	96.4 ± 0.6[a]	93.7 ± 1.1[a]
Average	71.1 ± 6.8[a]	72.3 ± 5.0[a]	97.2 ± 0.7[a]	97.1 ± 0.8[a]	94.9 ± 1.2[a]
AI	83.0	79.2	98.4	98.4	97.2

[a]$P < 0.01$.

[b]$P < 0.05$.

Supplementary Table 4 The Performance of CAD and Endoscopists on Yamada Class I (Image-based)

Group	PP	Sensitivity	NPV	Specificity	Accuracy
Experts	77.0 ± 7.1[a]	67.9 ± 10.8	96.8 ± 1.0	97.8 ± 1.3	95.1 ± 0.7[a]
Seniors	75.1 ± 8.8	68.4 ± 8.1[b]	96.9 ± 0.8[b]	97.6 ± 1.2	94.9 ± 1.1[b]
Juniors	64.8 ± 3.6[a]	65.7 ± 5.4[a]	96.5 ± 0.5[a]	96.3 ± 0.8[a]	93.5 ± 0.4[a]
Average	72.3 ± 8.4[a]	67.3 ± 7.7[a]	96.7 ± 0.7[a]	97.2 ± 1.2[a]	94.5 ± 1.0[a]
AI	83.7	79.1	97.9	98.4	96.6

[a]$P < 0.01$.

[b]$P < 0.05$.

Supplementary Table 5 The Performance of CAD and Endoscopists on Yamada Class II (Image-based), %

Group	PPV	Sensitivity	NPV	Specificity	Accuracy
Experts	79.4 ± 8.5	80.4 ± 10.7	96.1 ± 2.0	95.4 ± 2.8	92.8 ± 1.1[a]
Seniors	78.3 ± 4.5[b]	75.9 ± 8.3[b]	95.2 ± 1.6[a]	95.6 ± 1.3	92.3 ± 1.3[a]
Juniors	77.5 ± 5.7[b]	59.7 ± 13.8[a]	92.2 ± 2.4[a]	96.4 ± 1.3	90.2 ± 1.9[a]
Average	78.4 ± 5.9[a]	72.0 ± 13.7[a]	94.5 ± 2.5[a]	95.8 ± 1.8	91.8 ± 1.8[a]
AI	84.7	90.6	98.1	96.7	95.6

[a]$P < 0.01$.

[b]$P < 0.05$.

Supplementary Table 6 The Performance of CAD and Endoscopists on Yamada Class III (Image-based), %

Group	PPV	Sensitivity	NPV	Specificity	Accuracy
Experts	64.3 ± 6.4[b]	74.5 ± 12.7	98.8 ± 0.6	97.9 ± 0.8[b]	96.9 ± 0.5
Seniors	52.9 ± 12.2[a]	71.8 ± 16.3	98.6 ± 0.8	96.4 ± 2.2	95.3 ± 1.4[b]
Juniors	38.5 ± 9.1[a]	74.1 ± 13.9	98.7 ± 0.7	94.0 ± 2.2[a]	93.1 ± 2.2[a]
Average	51.9 ± 14.0[a]	73.5 ± 13.1[b]	98.7 ± 0.6[b]	96.1 ± 2.4[a]	95.1 ± 2.2[a]
AI	76.1	63.6	98.3	99.0	97.4

[a]$P < 0.01$.

[b]$P < 0.05$.

Supplementary Table 7 The Performance of CAD and Endoscopists on Yamada Class IV (Image-based), %

Group	PPV	Sensitivity	NPV	Specificity	Accuracy
Experts	92.1 ± 8.6	72.6 ± 5.0[a]	99.0 ± 0.2[a]	99.7 ± 0.3	98.8 ± 0.2[b]
Seniors	79.1 ± 14.8	77.4 ± 17.9	99.2 ± 0.6	99.1 ± 0.8	98.3 ± 0.1[a]
Juniors	74.4 ± 15.2	79.2 ± 11.6	99.2 ± 0.4	98.8 ± 0.7	98.2 ± 0.5[b]
Average	81.9 ± 14.3	76.4 ± 11.8	99.1 ± 0.4[b]	99.2 ± 0.7	98.4 ± 0.4[a]
AI	87.5	83.3	99.4	99.6	99.0

[a]$P < 0.01$.

[b]$P < 0.05$.