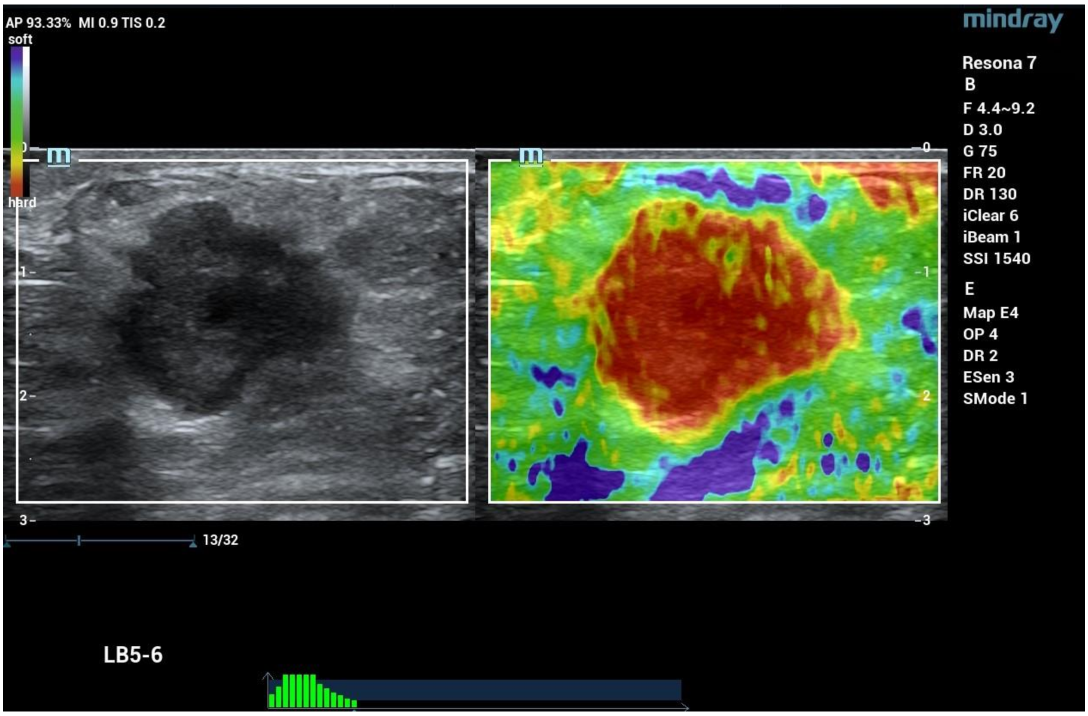

Fig S1. An example of US image collected from high-end US device. In this figure, BUS and EUS are shown in the white boxes on the left and right sides respectively.

1

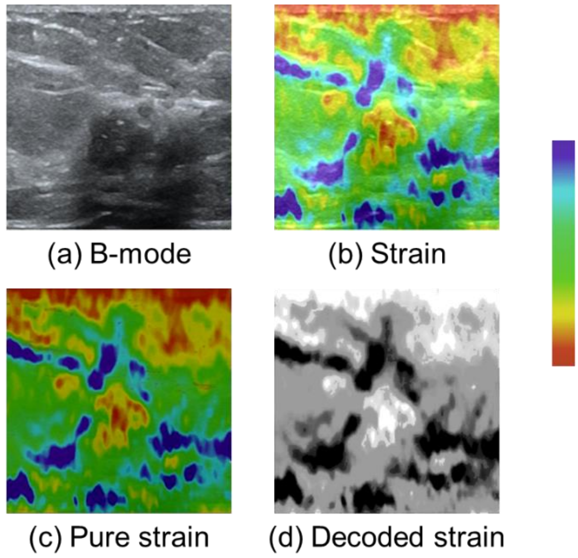

Fig S2. Pseudo color image decoding process. (a) the B-mode US image; (b) the original strain overlay on the B-mode US image by pseudo-color. (c) pure strain image obtained by subtracting the B-mode US image from the strain image. (d) gray-scale strain image reconstructed by decoding pure strain image.

2

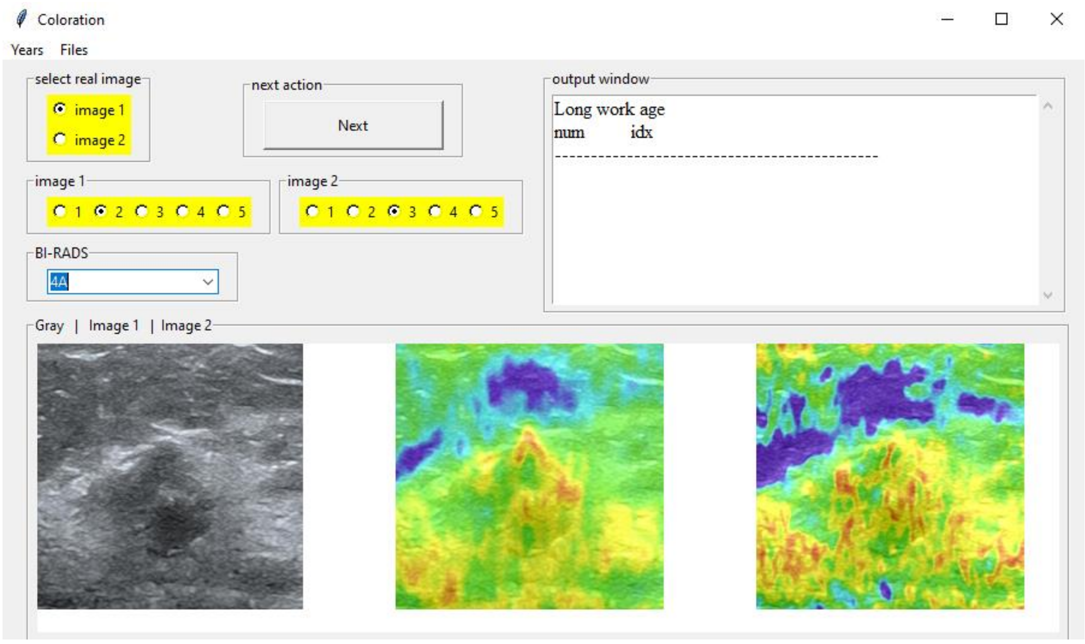

Fig S3. Software interface for blind evaluation.

3

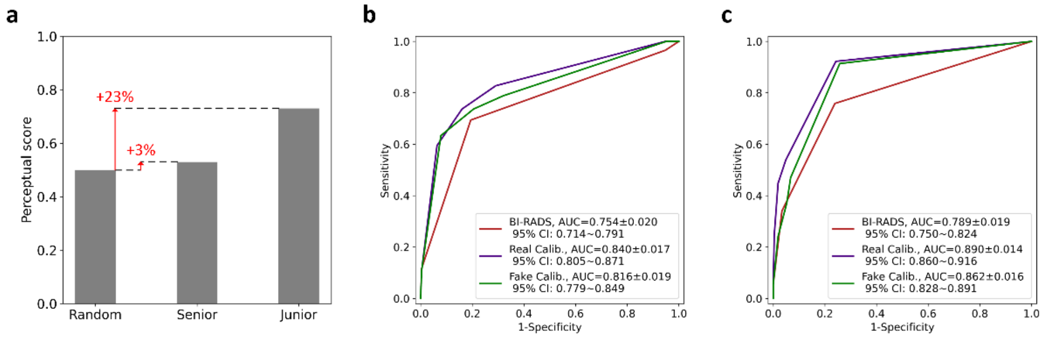

Fig. S4 Blind evaluation results of AR-EUS and real EUS. a. Perceptual score comparison of blind evaluation results of the junior radiologist and senior radiologist with random group. b. ROCs comparison of blind evaluation results of the junior radiologist in diagnosing breast cancer using BI-RADS, real EUS combined with BI-RADS, AR-EUS combined with BI-RADS. c. ROCs comparison of blind evaluation results of the senior radiologist in diagnosing breast cancer using BI-RADS, real EUS combined with BI-RADS, AR-EUS combined with BI-RADS.

4

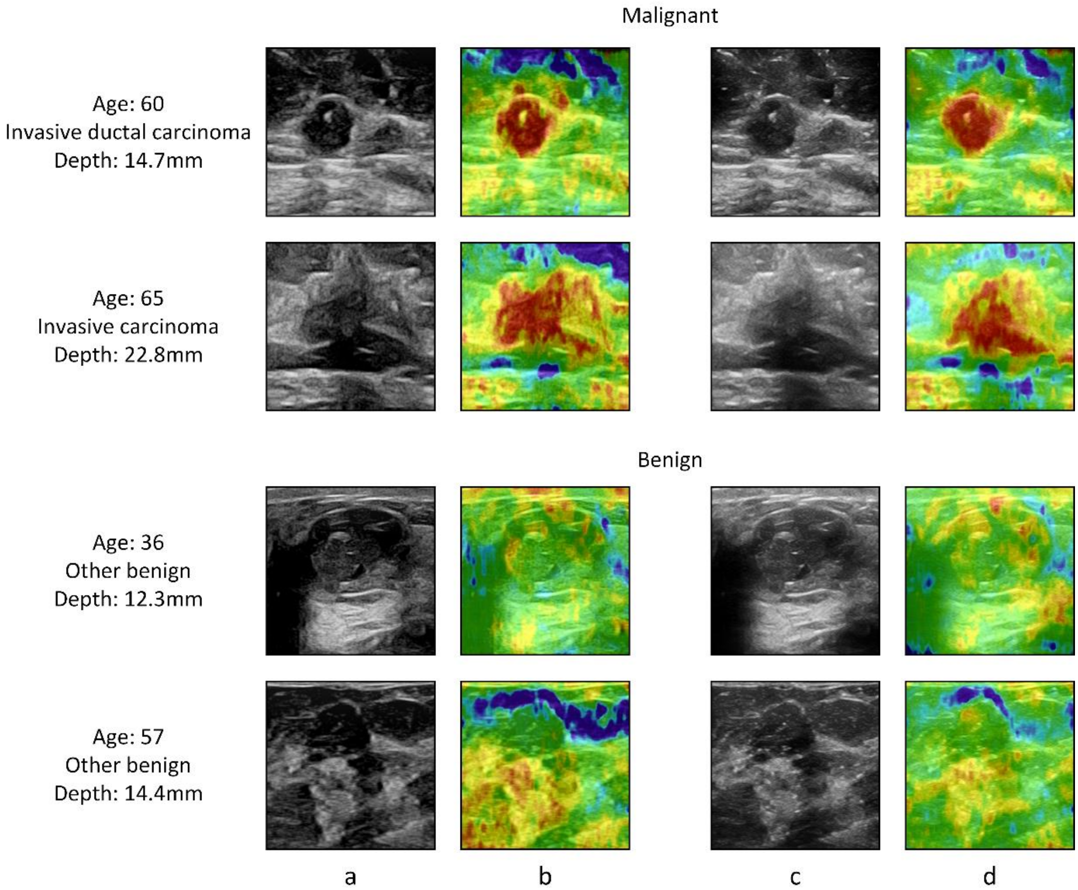

Fig. S5 Comparison of results before and after standardization of low-quality US. a. low-quality US. b. AR-EUS generated from low-quality US. c. Standardized low-quality US. d. AR-EUS generated from standardized low-quality US.

5

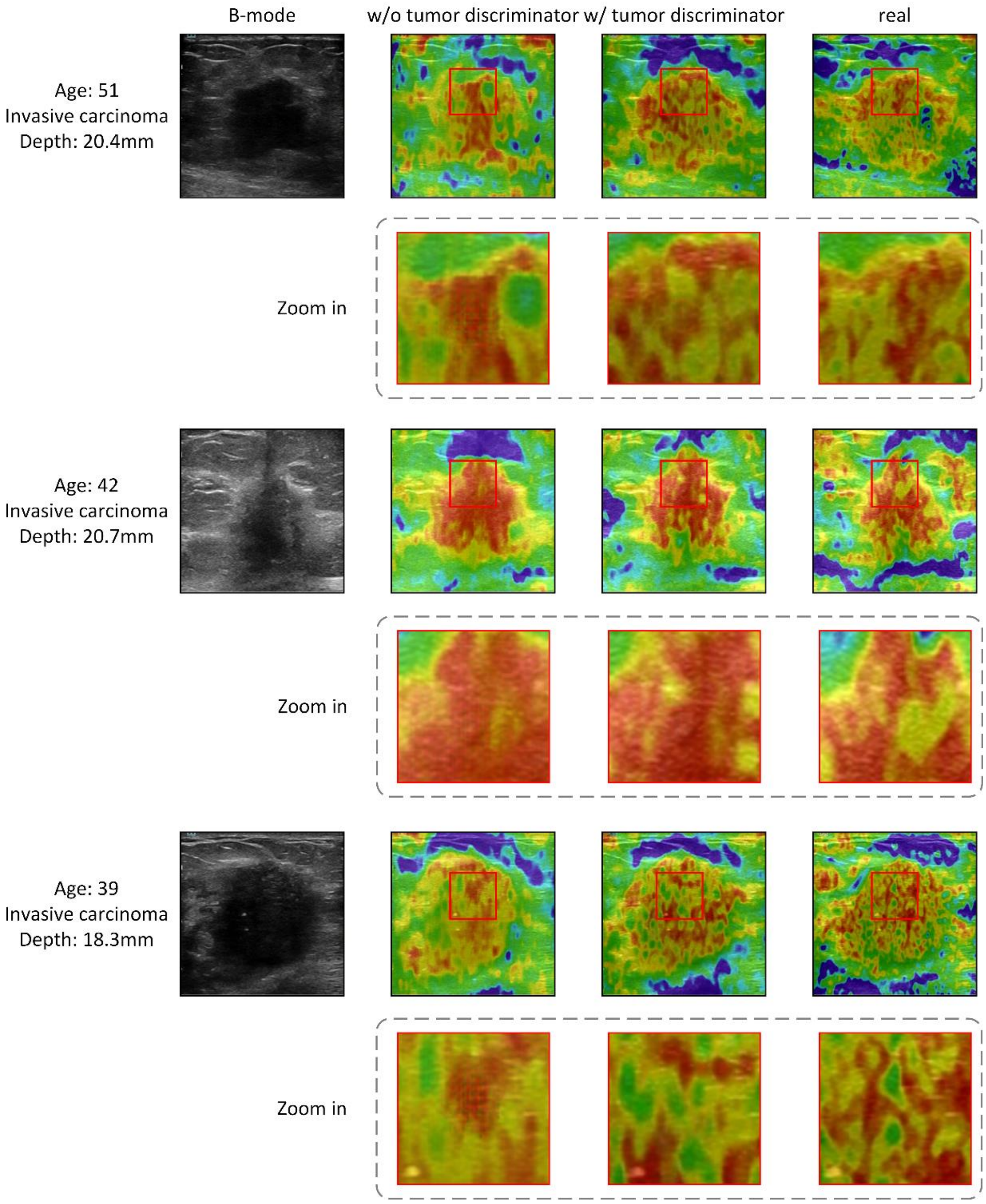

Fig S6. Examples of AR-EUS generated by model with or without tumor discriminator. For each example, we show BUS, real EUS and two AR-EUS images generated by model with or without tumor discriminator. The zoomed area (marked by red box) shows that the AR-EUS generated by the model with tumor discriminator can effectively avoid the ‘checkerboard’ artifacts.

6

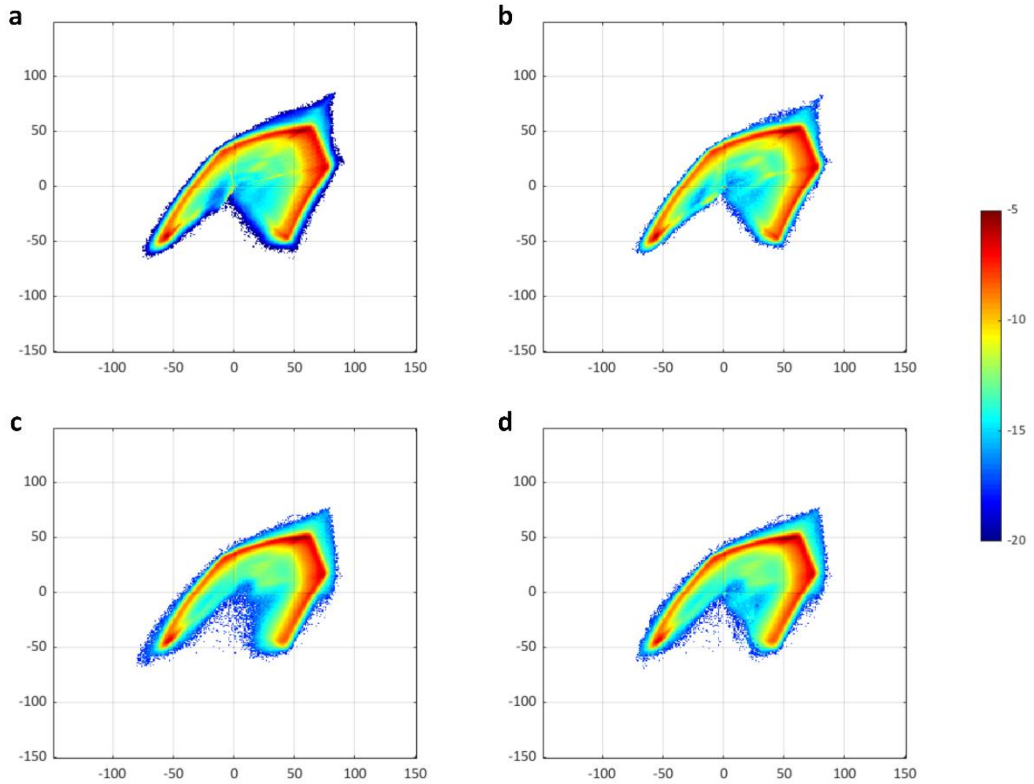

Fig S7. Comparison of the empirical distribution of *ab* values, shown in log scale. a. The empirical distribution of *ab* values in train set. b. The empirical distribution of *ab* values of real EUS in test set. c. The empirical distribution of *ab* values of AR-EUS in test set without using prior boost module. d. The empirical distribution of *ab* values of AR-EUS in test set using prior boost module.

7

Table S1. ablation experiment results

	Tumor enhance	Prior boost	SSIM	PSNR	MAE
Exp. 1	-	-	0.896±0.062	22.088±3.068	6.282±3.713
Exp. 2	√	-	0.897±0.062	22.364±3.100	6.058±3.604
Exp. 3	-	√	0.887±0.055	22.167±3.239	5.980±3.530
Exp. 4	√	√	**0.903±0.059**	**22.464±3.090**	**5.908±3.495**

A. The calculation methods of SSIM, MAPE, CHC

Structure similarity index measurement (SSIM), mean absolute percentage error (MAPE), and color histogram correlation (CHC) are metrics that commonly used to evaluate the similarity between two images. The calculation formulas are listed as below:

1. The SSIM is a metric to measure the similarity between two images, as defined in (1):

$$\mathrm{SSIM} = \frac{(2\mu_{real}\mu_{virtual} + C_1)(2\sigma_{real,virtual} + C_2)}{(\mu_{real}^2 + \mu_{virtual}^2 + C_1)(\sigma_{real}^2 + \sigma_{virtual}^2 + C_2)} \tag{1}$$

Where μ_{real} and $\mu_{virtual}$ are the average of real EUS and AR-EUS, respectively. σ_{real} and $\sigma_{virtual}$ are the variance of real EUS and AR-EUS, respectively. $\sigma_{real,virtual}$ is the covariance of real EUS and AR-EUS. C_1 and C_2 are constants.

2. The MAPE denotes percentage of the mean absolute error between strain ratios (SR) of real EUS and AR-EUS, as defined in (2):

$$\mathrm{MAPE} = \frac{1}{m}\sum_{i=1}^{m}\frac{\left|p_i^{real} - p_i^{virtual}\right|}{p_i^{real}} \tag{2}$$

Where p_i^{real} and $p_i^{virtual}$ represents the strain scores of real EUS and AR-EUS, respectively.

3. The CHC is a metrics to measure the histogram correlation (HC) between two images in hue and saturation color space. The HC is defined in (3):

$$\mathrm{HC} = 1 - \sqrt{1 - \frac{\sum \sqrt{Cnt_{real} \cdot Cnt_{virtual}}}{\sqrt{\sum Cnt_{real} \cdot \sum Cnt_{virtual}}}} \quad (3)$$

Where the Cnt_{real} and $Cnt_{virtual}$ are vectors containing the count of every bin in the histogram of real EUS and AR-EUS respectively. Therefore, CHC is the mean of HC in hue and saturation color space.

B. Decoding pseudo color of EUS via color bar and strain ratio calculation

A typical US image collected from high-end US instrument (Resona 7 ultrasound system, Mindray Medical International, Shenzhen, China) is shown in Fig S1. The pseudo color which depicts tissue elasticity is overlaid on BUS to form EUS, as shown in the right of Fig S1.

As shown by the color bar in the upper left of Fig S1, the red in EUS indicates that the tissue is hard, and the blue indicates that the tissue is soft. In order to qualify the difference in tissue elasticity represented by different colors in EUS, we use the color bar to decode the colors in EUS. Specifically, we first subtract BUS from EUS to obtain the pure color. Then, we decode the pure color image by using the color bar[1]. The elasticity modulus value is encoded with 256 pseudo-color levels from blue to red. Each pixel in pure color image is compared with all the color in color bar and the modules value with the smallest distance is taken as the elasticity value at this pixel. The distance d between the pixel and the color is defined as:

$$\mathrm{d(c)} = \sqrt{(R_{pixel} - R_c)^2 + (G_{pixel} - G_c)^2 + (B_{pixel} - B_c)^2}$$

where R, G, B is three color channel and the subscripts $pixel$ and c represent the pixel in pure strain image and pixel in color bar. In our experiment, the color bar is divided in to 256 color classes from blue to red, so the range of c is 1 to 256. This process is shown in Fig S2.

To calculate the strain ratio (SR), we also select a reference area with a size of 25*25 pixels in the pure strain image, decode it and average the modules value to obtain the elastic value of the reference region. According to the definition of SR, we divide the elastic value of tumor region and reference region to obtain the SR.

C. Blind evaluation

As shown in Fig S3, the blind evaluation software mainly composed of 3 modules, including real EUS and AR-EUS judgement, BUS BI-RADS grading and EUS 5-point scoring. The process of blind evaluation is as follows:

1. Participant chose their working years (senior or junior) in the menu, and then select the folder to be evaluated.

2. The US image pair being evaluated is displayed in the image display area at the bottom of the software interface. The display position of BUS is fixed and always displayed on the left side of the image display area. Real EUS and AR-EUS are randomly displayed on the right side of the BUS, and their positions are marked as image 1 and image 2 respectively. Participant is requested to give the BI-RADS score of BUS, the Tsukuba score of image 1 and image 2, and pick out the real EUS from image 1 and image 2.

3. The status window on the right side of the software interface outputs the current evaluation status. Click the 'Next' button to evaluate the next pair of US images.

D. The Tsukuba scoring system

The EUS is scored by sonographer based on 5-point scoring system[2]. The detailed scoring rules are as follows: 1=even strain over the entire lesion; 2=strain in most of the lesion; 3=strain at the periphery of the lesion with sparing in the center of a lesion; 4=no strain in the entire lesion; 5=no strain in the entire lesion or the surrounding area.

E. Blind evaluation results of AR-EUS

The blind evaluation results are shown in Fig. S4.

F. Results of AR-EUS after BUS standardization

The results of AR-EUS after BUS standardization are shown in Fig. S5.

G. Tumor discriminator in removing artifacts examples

Fig S6. shows the effect of tumor discriminator on removing artifacts from AR-EUS. We can see that when the tumor discriminator is not added, the AR-EUS will show obvious ‘checkerboard’ artifacts, which may be caused by the over-fitting of the discriminator. After applying the tumor discriminator, the model can effectively eliminate artifacts and make AR-EUS looks more visually authentic.

H. Ablation experiment results

The ablation experiment results are listed in Table S1.

I. Color rebalancing coefficient calculation

EUS have the characteristics of simple color distribution, mainly in blue and red. To accommodate this, we used the color prior coefficient to regularize the L1 loss. We first statistic the color prior of the training data to get the color distribution over *ab* space, as shown in Fig 5a. Then, in order to simplify the calculation, we quantize the *ab* color space with a grid size of 10 and kept the $\mathrm{Q} = 313$ values which were in-gamut[3]. In this way, we can have access to the prior probability of each bin in-gamut, and for each *ab* value falling into the bin, we can use formula (1) to obtain the corresponding boost coefficient:

$$\gamma_p = (\alpha\tilde{P} + (1-\alpha)/Q)^{-1} \tag{1}$$

Where $\tilde{P}$ is the empirical distribution of pixel p, $1/Q$ is a uniform distribution. α is the temperature to control the punishment of $\tilde{P}$. In our experiment, $\alpha = 0.8$ works well.

As shown in Fig S7, when the prior boost module is not added, the *ab* value distribution of AR-EUS (shown in Fig S7c) is quite different from that of real EUS (shown in Fig S7b). By adding the prior boost module to the model, the *ab* value distribution of AR-EUS (shown in Fig S7d) is consistent with that of real EUS.

Reference

1 Xiao, Y. *et al.* Computer-Aided Diagnosis Based on Quantitative Elastographic Features with Supersonic Shear Wave Imaging. *Ultrasound Med Biol* **40**, 275-286 (2014).

2 Itoh, A. *et al.* Breast disease: Clinical application of US elastography for diagnosis. *Radiology* **239**, 341-350 (2006).

3 Zhang, R., Isola, P. & Efros, A. A. Colorful Image Colorization. *Computer Vision - Eccv 2016, Pt Iii* **9907**, 649-666, doi:10.1007/978-3-319-46487-9_40 (2016).