Supplementary material

1 The Description and code of deep learning-based super-resolution reconstruction

In recent years, deep learning-based super-resolution (SR) reconstruction technology has achieved significant development in enhancing the spatial resolution of medical images. The Onekey platform (http://medai.icu) offered a 3D SR reconstruction technology for medical imaging that utilizes a Generative Adversarial Network (GAN) as its fundamental architecture. GAN was composed of two parts: a generator network, which is responsible for converting low-resolution images into high-resolution ones, and a discriminator network, tasked with distinguishing between real and generated images. These networks were trained through an adversarial process, where the generator learns to produce increasingly realistic images to deceive the discriminator, while the discriminator improves its ability to identify real and fake images. The dataset used for training the 3D super-resolution reconstruction technology was made up of a large collection of medical images. Before training, these images were meticulously preprocessed to remove noise and artifacts, and their intensity values were normalized. Then, the images were paired to form low-resolution and high-resolution image pairs, with the low-resolution images obtained by down-sampling the high-resolution counterparts. These paired image datasets were employed to train the GAN model. In the GAN model, the loss function employed is composed of three parts: gradient loss, L1 loss, and perceptual loss. The gradient loss ensures that the gradients of the generated images are consistent with those of the high-resolution images. The L1 loss measures the pixel-wise differences between the generated images and the high-resolution images. The perceptual loss evaluates the discrepancies between the feature representations of the generated images and those of the high-resolution images obtained from a pre-trained deep learning model. The combination of these three loss functions helps to ensure that the generated

images are visually highly similar to the high-resolution images. The 3D super-resolution reconstruction technology provided by the Onekey platform has shown remarkable effectiveness in enhancing the spatial resolution of medical images. After being assessed across a variety of medical imaging modalities, such as Computed Tomography, Magnetic Resonance Imaging, and ultrasound imaging, it has demonstrated significant advantages in improving both image quality and spatial resolution.

The relevant code for this paper has been made publicly available at http://medai.icu. For access to the source code or for further technical details of this study, please contact the corresponding author of this paper.

2 The Description of Automatic Tumor Segmentation.

Firstly, image preprocessing was carried out mainly including the following aspects. (1) Resampling. Due to the possibility of varying spacing in images from different patients, resampling was employed to ensure segmentation accuracy by adjusting the varying spacing to a consistent spacing. (2) N4 Bias Field Correction. MR images often suffer from non-uniform brightness, known as bias field, which may arise from magnetic field inhomogeneities or various factors during the scanning process. N4 Bias Field Correction could assess and correct for this bias, thereby facilitating more accurate identification of structures and performing quantitative analysis. (3) Intensity range clipping. Due to the significant differences in image values across various tissues and lesions, it is necessary to truncate the range of image intensity to reduce the interference of external information on segmentation results. (4) Normalization. The specific formula is as follows:

$$x' = \frac{x - \mu}{\sigma} \qquad [1]$$

x represents the gray value of each voxel in the current MRI image, μ represents the mean gray value of all voxel in the current image, σ represents the standard deviation of the gray values in the current image,

and x' represents the result after the standardization.

Subsequently, data augmentation was performed. To mitigate the issue of insufficient data, and to ensure that the model has robustness and better generalization performance with limited data, the algorithm primarily employed two major types of data augmentation methods: spatial transformations and intensity transformations. Spatial transformations included the following aspects: random rotation, scaling, translation, and mirror flipping. Intensity transformations mainly included Gaussian noise and contrast enhancement. Training of the automatic tumor segmentation model was then initiated based on 3D U-Net. This network was composed of an encoder, a decoder, dense skip pathways, and deep supervision. A sliding window was used to slice the MRI data to preserve the spatial information between different layers. The process of model training consisted of 500 epochs with a batch size of 1, and each epoch included 1000 batches. The optimizer selected for model training was Stochastic Gradient Descent. The initial learning rate is set to 1e-03 for tumor. The dataset was randomly divided into a training set and a validation set in an 8:2 ratio for training. The test set was a separate dataset sourced differently from the training and validation sets. Testing on a dataset from a different source could more realistically reflect the model's generalizability and robustness. Each epoch included two processes: training and validation. At the end of each epoch, the trained model is tested on the validation dataset to obtain the Dice metric for that epoch's model on the validation set. If the model's metric on the validation set is the best to date, the model is updated; otherwise, training continues to the next epoch. Ultimately, the model that was trained based on the training and validation sets was applied to our study to calculate the model's Dice metric (95.12%). For access to the source code or for further technical details of this study, please contact the corresponding author of this paper.

Supplementary Table 1 Scanning parameters of each magnetic resonance imaging

Manufacturer	Model	Field Strength	Sequence	TR (ms)	TE (ms)	Filp angle (°)	Slice thickness (mm)	Acquisition Type	Matrix	Pixel Spacing (mm)
Philips	Multiva	1.5T	T2WI	1600	70	90	6	2D	268×53	0.73
			DWI	5200	66	90	5	2D	128×125	1.7
			DCE T1WI	5.9	0	15	5	3D	190×190	0.84
United Imaging	uMR560	1.5T	T2WI	4000	80	150	7	2D	320×167	0.67
			DWI	3800	76	90	7	2D	112×70	1.68
			DCE T1WI	3.2	1.5	10	2.5	3D	256×133	0.84
United Imaging	uMR570	1.5T	T2WI	4000	80	150	7	2D	320×142	0.75
			DWI	3000	100	90	7	2D	176×83	2.71
			DCE T1WI	3.2	1.5	10	2.5	3D	256×114	0.94

United Imaging	uMR770	3.0T	T2WI	2400	80	165	7	2D	288×154	0.97
			DWI	2500	77	90	7	2D	128×98	1.64
			DCE T1WI	2.8	1.2	10	2.5	3D	256×144	1.09
GE	MR750	3.0T	T2WI	6000	88	110	5	2D	320×320	0.82
			DWI	5000	55	90	5	2D	128×160	1.64
			DCE T1WI	3.3	1.4	12	4	3D	256×192	0.82
SIEMENS	Skyra	3.0T	T2WI	5200	75	72	5	2D	320×320	1.19
			DWI	5600	60	90	5	2D	128×104	1.48
			DCE T1WI	4.1	1.3	12	3.5	3D	320×192	1.19

Supplementary Table 2 Delong Test and NRI for Pairwise Comparison Among Different Models

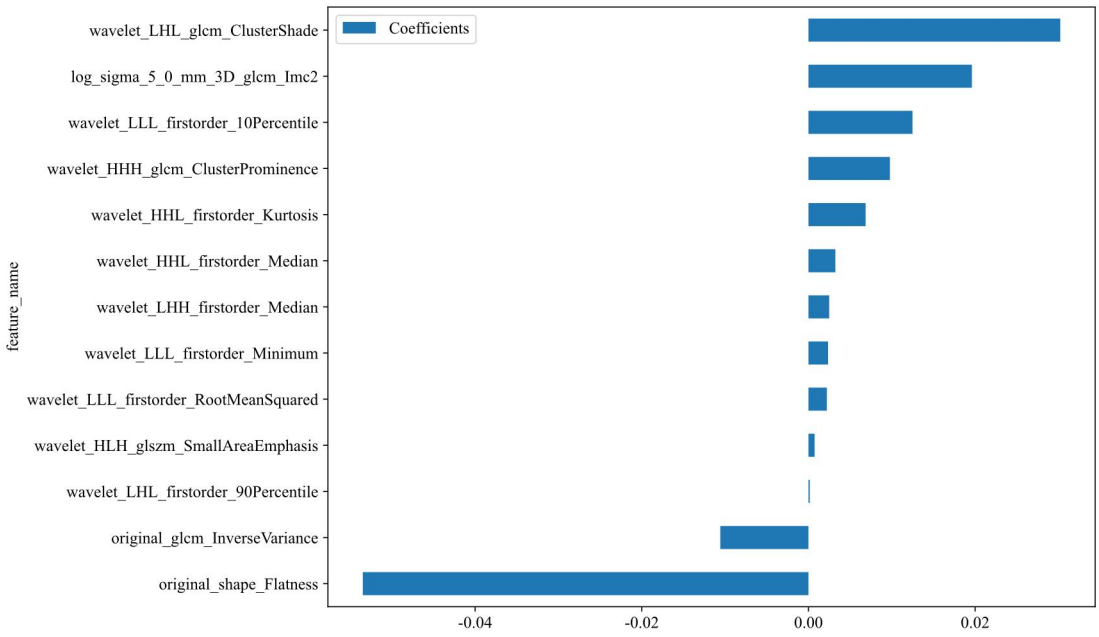
Original NR	SR MRI	Cohort	Delong	NRI

MRI				test*	
T2WI	*vs.*	T2WI	Training	0.171	0.081
			Validation	0.778	0.039
			Test	0.451	0.081
DWI	*vs.*	DWI	Training	0.424	0.02
			Validation	0.753	-0.016
			Test	0.993	0.022
PVP	*vs.*	PVP	Training	0.397	0.14
			Validation	0.615	0.024
			Test	0.41	0.122
All-sequence	*vs.*	All-sequence	Training	0.701	0.018
			Validation	0.473	0.053

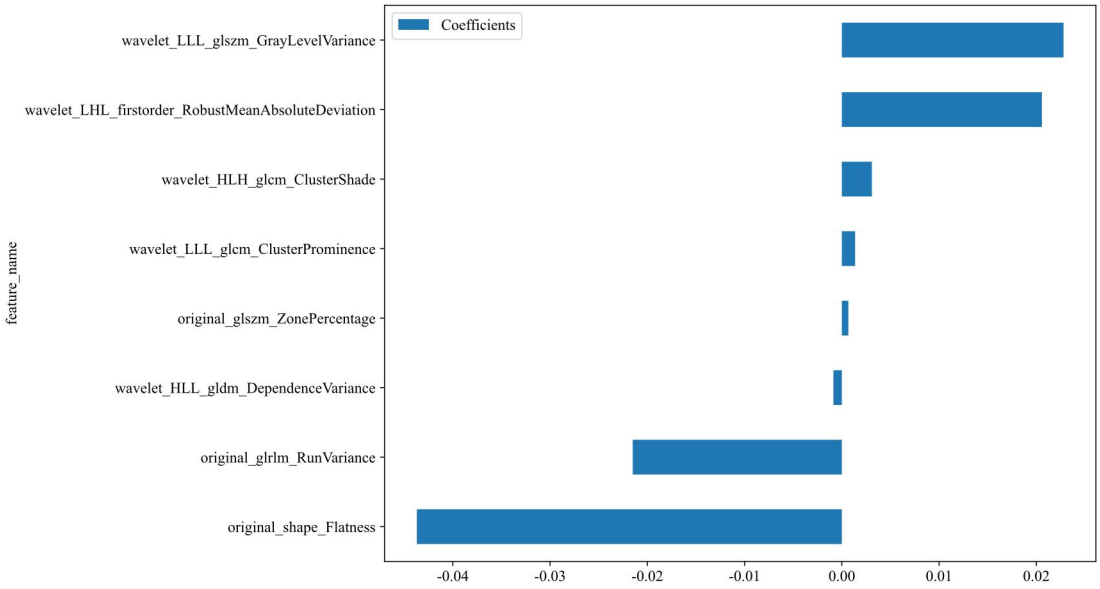
	Test	0.695	0.021

NR: Normal-resolution; SR: Super-resolution; MRI: Magnetic resonance imaging; *: p-value; IDI: Integrated discrimination index; NRI: Net reclassification index; T2WI: T2-weighted imaging; DWI: Diffusion-weighted imaging; PVP: Contrast-enhanced T1WI (portal venous phase); All-sequence: Including three sequences (T2WI, DWI, PVP).

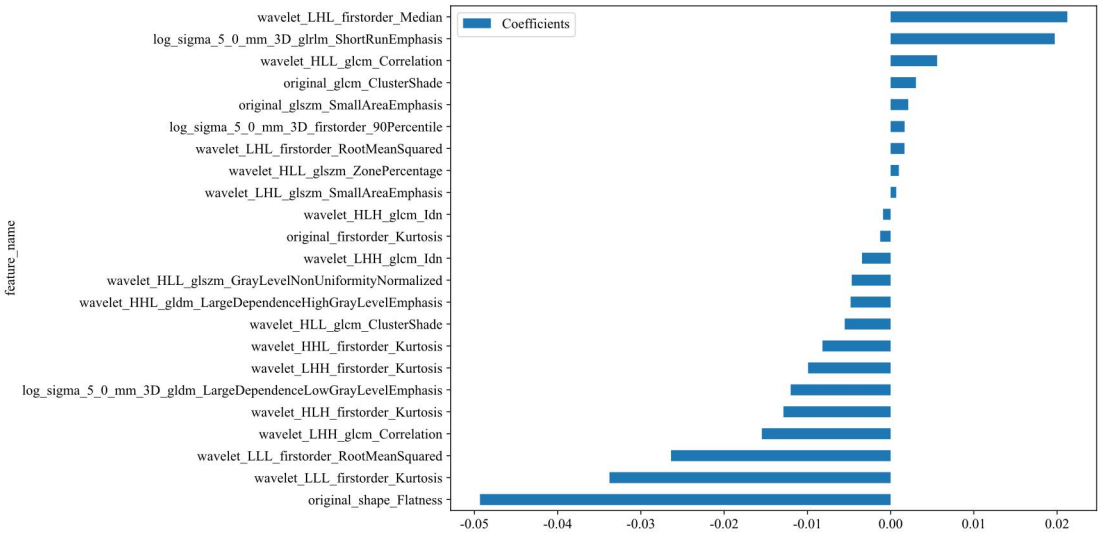
A


Coefficients
wavelet_LHL_glcm_ClusterShade
log_sigma_5_0_mm_3D_glcm_Imc2
wavelet_LLL_firstorder_10Percentile
wavelet_HHH_glcm_ClusterProminence
wavelet_HHL_firstorder_Kurtosis
wavelet_HHL_firstorder_Median
wavelet_LHH_firstorder_Median
wavelet_LLL_firstorder_Minimum
wavelet_LLL_firstorder_RootMeanSquared
wavelet_HLH_glszm_SmallAreaEmphasis
wavelet_LHL_firstorder_90Percentile
original_glcm_InverseVariance
original_shape_Flatness
feature_name
-0.04
-0.02
0.00
0.02

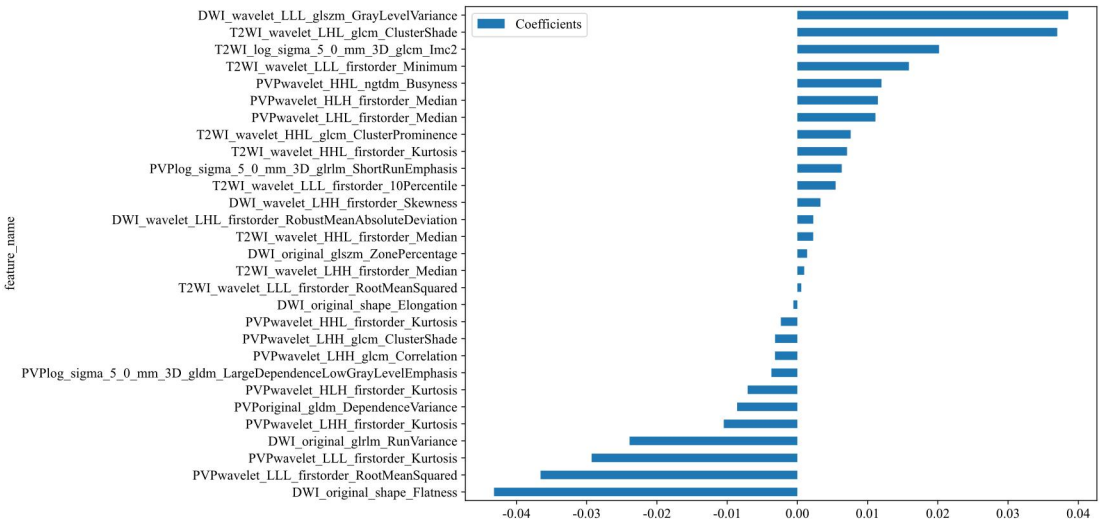
B

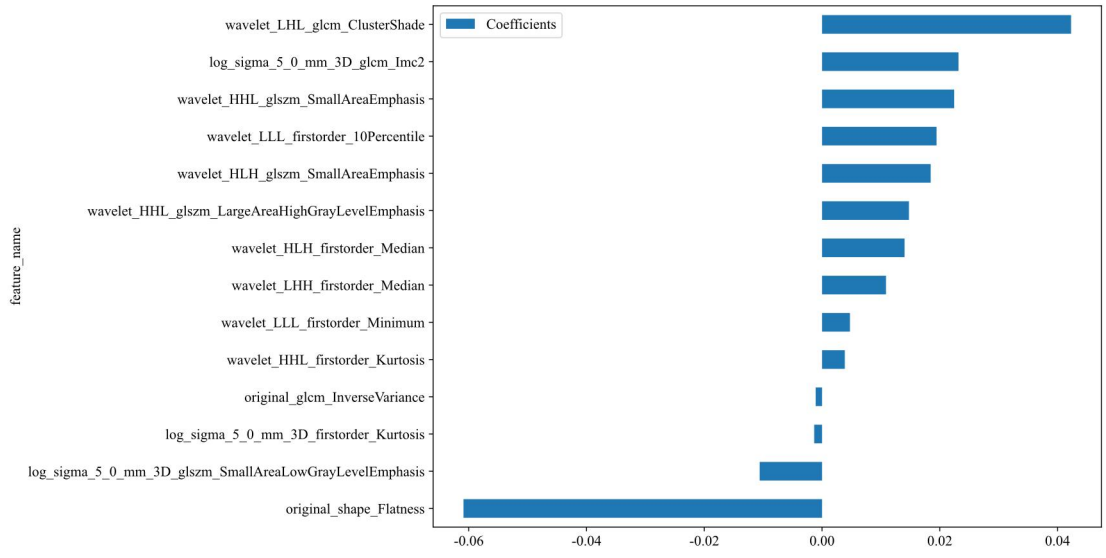

Coefficients
wavelet_LLL_glszm_GrayLevelVariance
wavelet_LHL_firstorder_RobustMeanAbsoluteDeviation
wavelet_HLH_glcm_ClusterShade
wavelet_LLL_glcm_ClusterProminence
original_glszm_ZonePercentage
wavelet_HLL_gldm_DependenceVariance
original_glrlm_RunVariance
original_shape_Flatness
feature_name
-0.04
-0.03
-0.02
-0.01
0.00
0.01
0.02

C

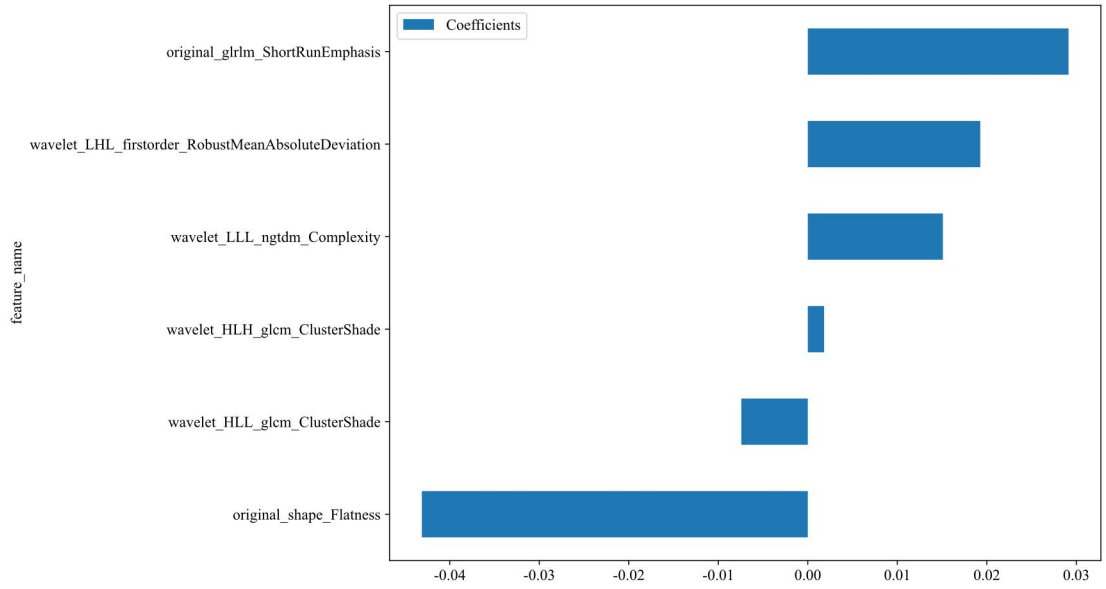

Coefficients
feature_name
wavelet_LHL_firstorder_Median
log_sigma_5_0_mm_3D_glrlm_ShortRunEmphasis
wavelet_HLL_glcm_Correlation
original_glcm_ClusterShade
original_glszm_SmallAreaEmphasis
log_sigma_5_0_mm_3D_firstorder_90Percentile
wavelet_LHL_firstorder_RootMeanSquared
wavelet_HLL_glszm_ZonePercentage
wavelet_LHL_glszm_SmallAreaEmphasis
wavelet_HLH_glcm_Idn
original_firstorder_Kurtosis
wavelet_LHH_glcm_Idn
wavelet_HLL_glszm_GrayLevelNonUniformityNormalized
wavelet_HHL_gldm_LargeDependenceHighGrayLevelEmphasis
wavelet_HLL_glcm_ClusterShade
wavelet_HHL_firstorder_Kurtosis
wavelet_LHH_firstorder_Kurtosis
log_sigma_5_0_mm_3D_gldm_LargeDependenceLowGrayLevelEmphasis
wavelet_HLH_firstorder_Kurtosis
wavelet_LHH_glcm_Correlation
wavelet_LLL_firstorder_RootMeanSquared
wavelet_LLL_firstorder_Kurtosis
original_shape_Flatness
-0.05
-0.04
-0.03
-0.02
-0.01
0.00
0.01
0.02

D

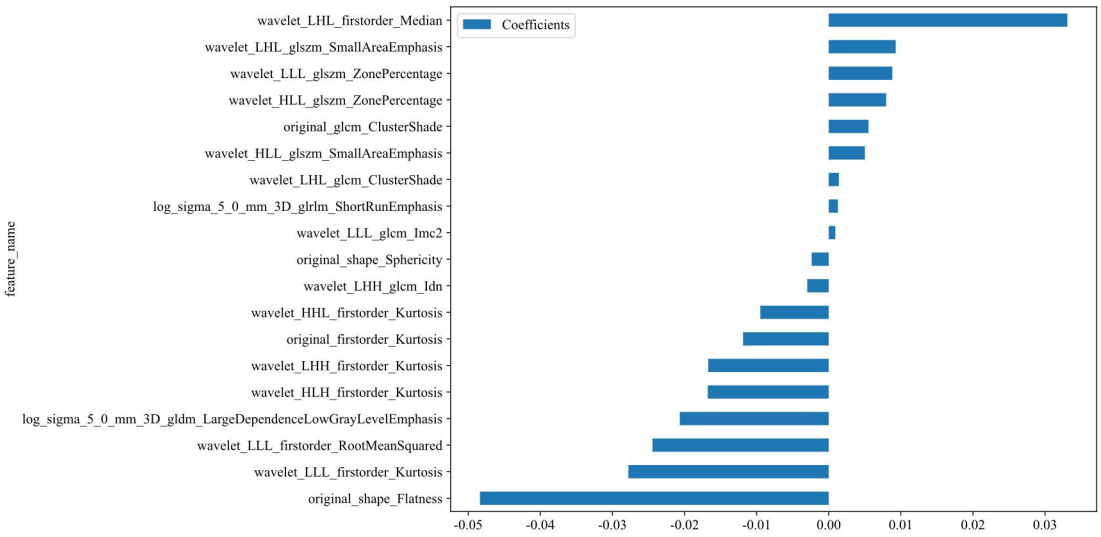

Coefficients
feature_name
DWI_wavelet_LLL_glszm_GrayLevelVariance
T2WI_wavelet_LHL_glcm_ClusterShade
T2WI_log_sigma_5_0_mm_3D_glcm_Imc2
T2WI_wavelet_LLL_firstorder_Minimum
PVPwavelet_HHL_ngtdm_Busyness
PVPwavelet_HLH_firstorder_Median
PVPwavelet_LHL_firstorder_Median
T2WI_wavelet_HHL_glcm_ClusterProminence
T2WI_wavelet_HHL_firstorder_Kurtosis
PVPlog_sigma_5_0_mm_3D_glrlm_ShortRunEmphasis
T2WI_wavelet_LLL_firstorder_10Percentile
DWI_wavelet_LHH_firstorder_Skewness
DWI_wavelet_LHL_firstorder_RobustMeanAbsoluteDeviation
T2WI_wavelet_HHL_firstorder_Median
DWI_original_glszm_ZonePercentage
T2WI_wavelet_LHH_firstorder_Median
T2WI_wavelet_LLL_firstorder_RootMeanSquared
DWI_original_shape_Elongation
PVPwavelet_HHL_firstorder_Kurtosis
PVPwavelet_LHH_glcm_ClusterShade
PVPwavelet_LHH_glcm_Correlation
PVPlog_sigma_5_0_mm_3D_gldm_LargeDependenceLowGrayLevelEmphasis
PVPwavelet_HLH_firstorder_Kurtosis
PVPoriginal_gldm_DependenceVariance
PVPwavelet_LHH_firstorder_Kurtosis
DWI_original_glrlm_RunVariance
PVPwavelet_LLL_firstorder_Kurtosis
PVPwavelet_LLL_firstorder_RootMeanSquared
DWI_original_shape_Flatness
-0.04
-0.03
-0.02
-0.01
0.00
0.01
0.02
0.03
0.04


Coefficients
wavelet_LHL_glcm_ClusterShade
log_sigma_5_0_mm_3D_glcm_Imc2
wavelet_HHL_glszm_SmallAreaEmphasis
wavelet_LLL_firstorder_10Percentile
wavelet_HLH_glszm_SmallAreaEmphasis
wavelet_HHL_glszm_LargeAreaHighGrayLevelEmphasis
wavelet_HLH_firstorder_Median
wavelet_LHH_firstorder_Median
wavelet_LLL_firstorder_Minimum
wavelet_HHL_firstorder_Kurtosis
original_glcm_InverseVariance
log_sigma_5_0_mm_3D_firstorder_Kurtosis
log_sigma_5_0_mm_3D_glszm_SmallAreaLowGrayLevelEmphasis
original_shape_Flatness
feature_name
-0.06
-0.04
-0.02
0.00
0.02
0.04

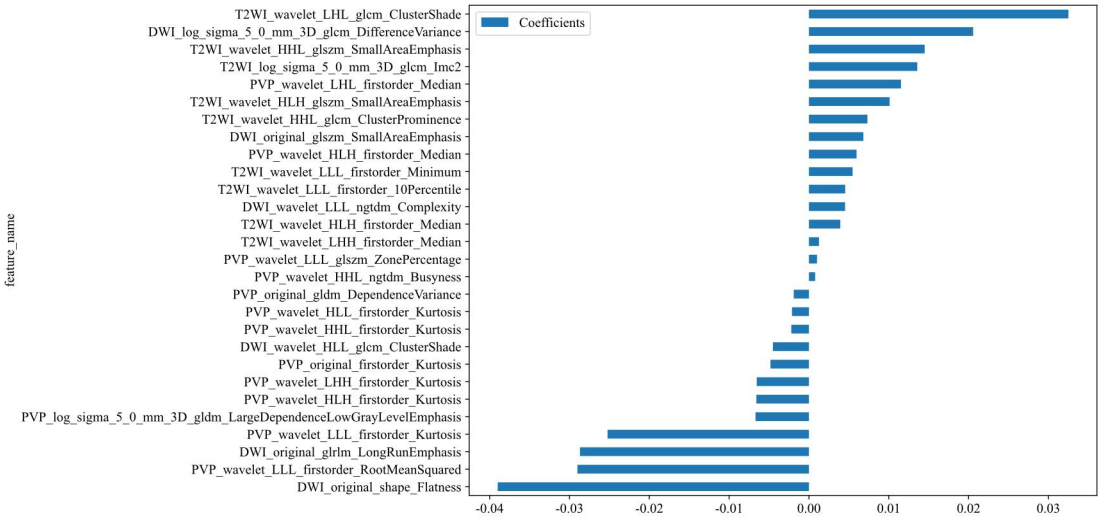
E


Coefficients
original_glrlm_ShortRunEmphasis
wavelet_LHL_firstorder_RobustMeanAbsoluteDeviation
wavelet_LLL_ngtdm_Complexity
wavelet_HLH_glcm_ClusterShade
wavelet_HLL_glcm_ClusterShade
original_shape_Flatness
feature_name
-0.04
-0.03
-0.02
-0.01
0.00
0.01
0.02
0.03

F



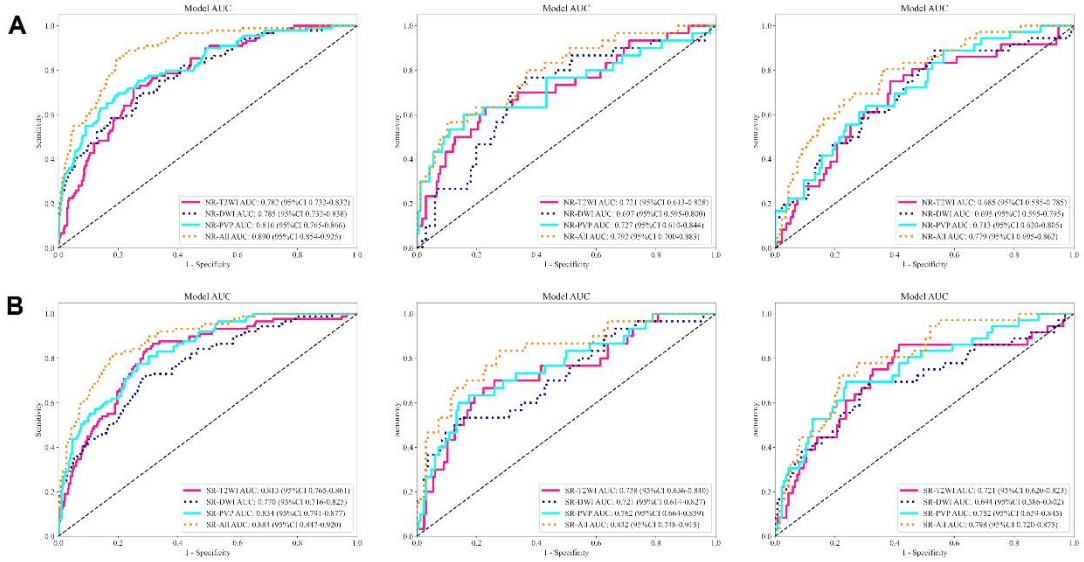
G



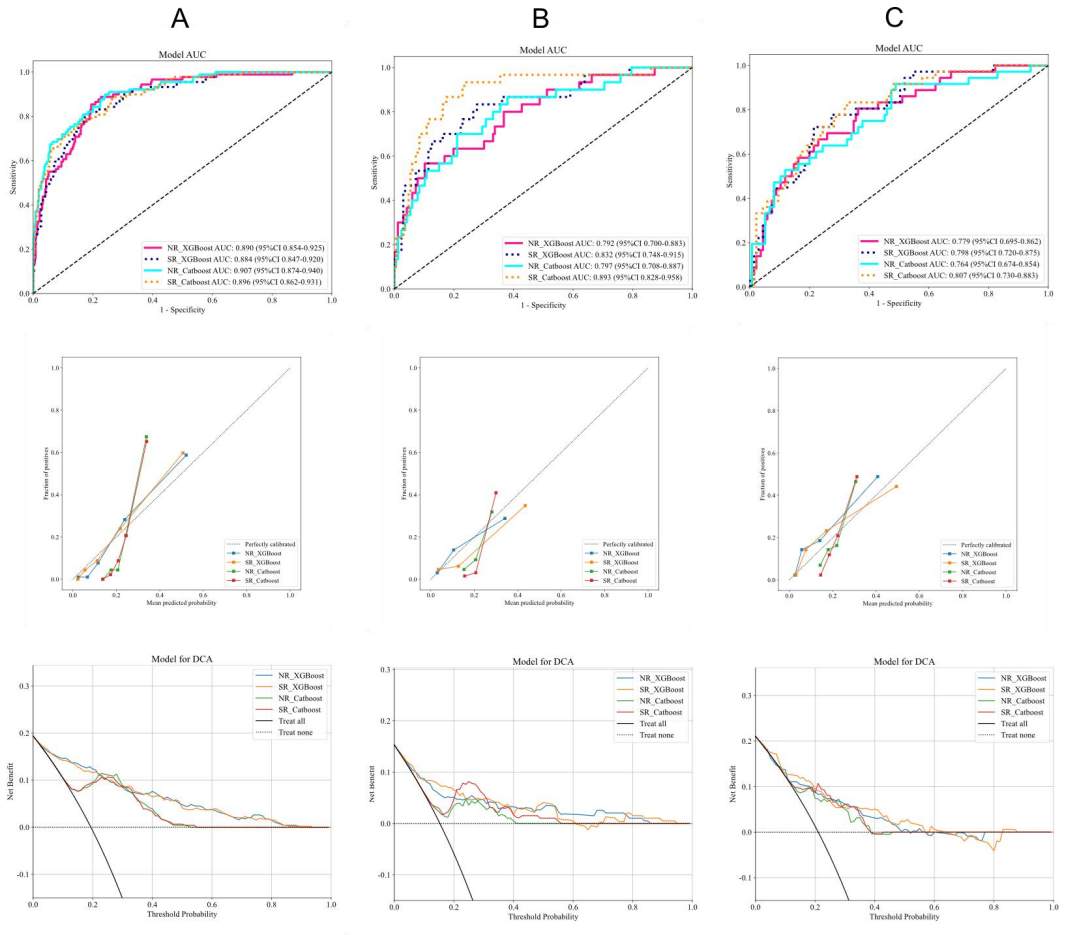
H

Supplementary Figure 1 The weights of radiomics features selected from different radiomics models. A: Details of radiomic features extracted from T2WI based on NR MRI; B: Details of radiomic features extracted from DWI based on NR MRI; C: Details of radiomic features extracted from PVP based on NR MRI; D: Details of radiomic features extracted from All-sequence based on NR MRI; E: Details of radiomic features extracted from T2WI based on SR MRI; F: Details of radiomic features extracted from DWI based on SR MRI;G:

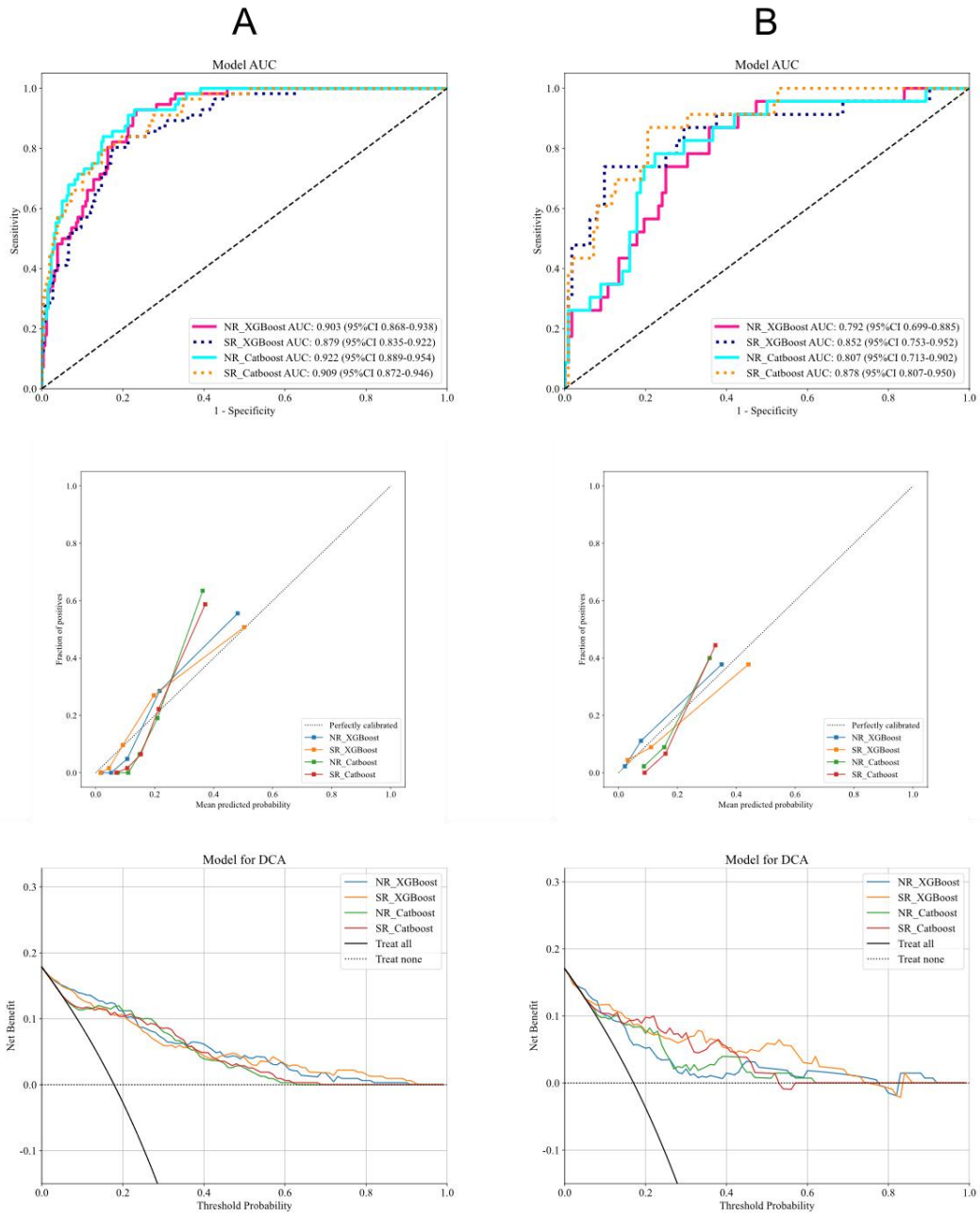
Details of radiomic features extracted from PVP based on SR MRI; H: Details of radiomic features extracted from All-sequence based on SR MRI.



Supplementary Figure 2 ROC curves. A and B: Various radiomics models based on NR (A) and SR (B) MRI in training, validation, and test cohorts, respectively.



Supplementary Figure 3 The receiver operating characteristic curves, calibration curves, and decision curve analysis based on Catboost. A: Training cohort; B: Validation cohort; C: Test cohort.



Supplementary Figure 4 The predictive efficacy of subgroup analyses based on United imaging 560 and 570. A: Training cohort; B: Test cohort.