IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 5, 2017, accepted January 17, 2018, date of publication January 23, 2018, date of current version March 16, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2796632

Can Signal-to-Noise Ratio Perform as a Baseline Indicator for Medical Image Quality Assessment

ZHICHENG ZHANG[1,2], GUANGZHE DAI[3], XIAOKUN LIANG[2], SHAODE YU[1,2], LEIDA LI[4], (Member, IEEE), AND YAOQIN XIE[1]

[1]Shenzhen Institutes of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China
[2]Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China
[3]Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang 110819, China
[4]School of Information and Control Engineering, Chinese University of Mining and Technology, Xuzhou 221116, China

Corresponding author: Yaoqin Xie (yq.xie@siat.ac.cn)

This work is supported in part by the National Natural Science Foundation of China under Grant 61471349, Grant 61771473, and Grant 61379143, in part by the Leading Talent of Special Support Project in Guangdong under Grant Y77504, in part by the Shenzhen Key Technical Research Project under Grant JSGG20160229203812944, in part by the National Science Foundation of Guangdong under Grant 2014A030312006, in part by the Beijing Center for Mathematics and Information Interdisciplinary Sciences, in part by the Six Talent Peaks High-level Talents in Jiangsu Province under Grant XYDXX-063, in part by the Qing Lan Project, and in part by the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology.

ABSTRACT Natural image quality assessment (NIQA) wins increasing attention, while NIQA models are rarely used in the medical community. A couple of studies employ the NIQA methodologies for medical image quality assessment (MIQA), but building the benchmark data sets necessitates considerable time and professional skills. In particular, the characteristics of synthesized distortions are different from those of clinical distortions, which make the results not so convincing. In clinic, signal-to-noise ratio (SNR) is widely used, which is defined as the quotient of the mean signal intensity measured in a tissue region of interest (ROI) and the standard deviation of the signal intensity in an air region outside the imaged object, and both regions are outlined by specialists. We take advantage of the knowledge that SNR is routinely used and concern whether SNR measure can perform as a baseline metric for the development of MIQA algorithms. To address the issue, the inter-observer reliability of SNR measure is investigated regarding to different tissue ROIs [white matter (WM); cerebral spinal fluid (CSF)] in magnetic resonance (MR) images. A total of 192 T_2^*, 88 T_1, 76 T_2 and 55 contrast-enhanced T_1 (T_1C) weighted images are analyzed. Statistical analysis indicates that SNR values show consistency between different observers to the same ROI in each modality (Wilcoxon rank sum test, $p_w \geq 0.11$; and paired sample t-test, $p_p \geq 0.28$). Moreover, whether off-the-shelf NIQA models can predict MR image quality is considered by using SNR values as the reference scores. Four NIQA models (BIQI, BLIINDS-II, BRISQUE, and NIQE) are evaluated, and the correlation between SNR values and NIQA results is evaluated. Pearson correlation coefficient (r_p) shows that WM-based SNR values correlates well with BIQI, BLIINDS-II and BRISQUE in T_2^* images ($r_p \geq 0.77$), BRISQUE and NIQE in T_1 images ($r_p \geq 0.75$), BLIINDS-II in T_2 images ($r_p \geq 0.67$), and BRISQUE and NIQE in T_1C images ($r_p \geq 0.58$), while CSF-based SNR values correlates well with BLIINDS-II in T_2^* images ($r_p \geq 0.64$) and T_2 images ($r_p \geq 0.60$), and all $p_p < 10^{-4}$. The prediction performance analysis further proves the result from the correlation analysis. Conclusively, SNR measure is reliable to different observations and can perform as a baseline indicator for the development of MIQA algorithms. In general, BRISQUE and BLIINDS-II are full of potential to be conditionally used as objective MIQA models toward human brain MR images. This paper presents the first attempt of using SNR measure to bridge the gap between NIQA and MIQA, and large-scale experiments should be further conducted to confirm the conclusion in this paper.

INDEX TERMS Signal-to-noise ratio, natural image quality assessment, medical image quality assessment, magnetic resonance imaging.

I. INTRODUCTION

Natural image quality assessment (NIQA) wins increasing attention [1]–[8] and a variety of NIQA models are available [2], [9]–[14], such as BIQI [15], BLIINDS-II [16], BRISQUE [17] and NIQE [18]. Furthermore, a large number of novel algorithms, improvement of existing methods and applications of NIQA models to other fields are presented in each year [19]–[23]. However, the general NIQA models are

rarely used in the medical community [24]. One reason comes from the fact that various artifacts in medical images are not seen in natural images. These artifacts, hardware-related and human-related, hamper the direct application of the state-of-the-art NIQA models for medical image quality assessment (MIQA). Another cause is from imaging characteristics, and for accurate decision making, images are acquired from different modalities. Last but not the least, high-quality images support clinical diagnosis, while the quantification of medical image quality is not a reliable indicator of diagnostic accuracy.

The quality of medical images is closely related to image interpretation, disease diagnosis, surgical planing and treatment delivery. In each day, various types of imaging modalities are used, such as ultrasound, computerized tomography (CT) and magnetic resonance (MR) imaging, not to speak of these devices under clinical trial [25]–[27]. MIQA has been used in quality assurance [28]–[30], comparison of image restoration methods [31]–[35] and clinical diagnosis [36]–[39]. Generally, MIQA can be divided into subjective and objective assessment [40]–[43]. In routine work, subjective assessment is common, such as hand-hold ultrasound imaging for diagnosis and X-ray imaging for lung cancer screening. Objective assessment can be further categorized into full- and no-reference estimation. The former require high-quality reference images, while the reference is hard or unavailable to access in medical community. To tackle this problem, researchers make use of the images from advanced imaging devices as the reference to validate the designed algorithms with images from common devices. However, this kind of approaches causes new problems, such as uncontrollable motion and different imaging characteristics [44]–[46]. Relatively, no-reference MIQA is more useful but challenging, because no information regarding the reference is available for the assessment.

To the best of our knowledge, a couple of MIQA studies utilize NIQA methodologies. The first study [47] builds a benchmark data set which includes 6 types of synthesized distortions based on 25 MR images. The distortions include Rician noise, Gaussian white noise, Gaussian blur, Discrete Cosine Transform, JPEG and JPEG2000 compression. It involves 28 subjects and a total of 21,700 human evaluation to quantify the MR image quality. The study investigates the correlation between the subjective scores and 13 full-reference NIQA models. The second one [48] is based on the benchmark database built in the first study. It modifies a general NIQA model of BRISQUE [17] to predict the quality of MR images. However, the drawbacks of both studies come from the high cost, large disparities between the simulated distortions and the real-world medical imaging. In addition, limited MR images and two MR imaging sequences are concerned.

In medical community, one daily used indicator of image quality is signal-to-noise ratio (SNR) [49]–[54]. It has been applied in the estimation of the development of new hardware and image processing algorithms [37], [47]. How to calculate SNR is based on the signal intensities of two separate regions from a single image. Thus, SNR measure is also known as a ''two-region'' approach. One region is outlined to determine the tissue signal and the other is to measure the background noise [49]–[51]. And accordingly, fair comparison of medical image quality with SNR is burdensome across studies [40]. Above all, SNR values change in accordance with the outlined regions. Specially, different tissue regions are outlined for specific purposes; and even for the same purpose, the regions delineated by two observers or delineated twice by one observer are not the same. Besides, a number of factors influence the quality of acquired images. Taking MR imaging for instance, these factors are not limited to device venders, magnetic field strengths, pulse sequences, field of view, matrix size, slice thickness and reconstruction methods. In addition, image acquisition is prone to noise and artifacts which are embodied in diverse imaging characteristics across modalities.

In this study, we take advantage of the knowledge that SNR is routinely used and concern whether SNR can perform as a baseline metric for the development of MIQA algorithms. To address the issue, the inter-observer reliability of SNR measure is first evaluated. In total, 192 T_2^*, 88 T_1, 76 T_2 and 55 contrast-enhanced T_1 (T_1C) weighted MR images are analyzed. White matter (WM) and cerebral spinal fluid (CSF) are outlined as the tissue region of interest (ROI) for SNR measure. After the consistency of SNR measure is verified, we further consider the potentiality of using SNR measure as a baseline indicator for the development of MIQA models. Consequently, the performance of four general NIQA (BIQI, BLIINDS-II, BRISQUE and NIQE) models are studied. Based on the 411 *in vivo* human brain images and 4 MR sequences, this study not only has verified the reliability of SNR measure, but also has shed light on developing automatic, objective and no-reference MIQA algorithms by using SNR as the baseline indicator.

II. MATERIALS AND METHODS

A. DATA COLLECTION

On a 3.0 Tesla MR scanner (SIEMENS, Germany) by using an 8-channel brain phased-array coil, a total of 411 images are collected among which 192 are T_2^* weighted images of normal brain, while 88 T_1, 76 T_2 and 55 contrast-enhanced T_1 (T_1C) weighted images of brain with tumors.

Specifically, T_2^* images are acquired by using gradient-echo pulse sequence (flip angle: 15°; field of view: 220×220 mm^2; image matrix: 384×384; slice thickness: 3.0 mm; repetition time (TR): 200 ms; and echo time (TE) ranging from 2.61 ms to 38.91 ms with an equal interval 3.3 ms) [32]. On the other hand, T_1, T_2 and T_1C images are scanned by using spin echo sequence with different TR (535 ms, 3500 ms and 650 ms) and TE (8 ms, 105 ms and 9 ms) pairs. Besides, flip angle is 15°, field of view 220×220 mm^2 and slice thickness 1 or 2 mm. In particular,

the matrix size of T_1 and T_1C images ranges from 512×432 to 668×512 and the matrix size of T_2 images is from 384×324 to 640×640.

B. EXPERIMENT DESIGN

The whole experiment can be divided into three phases. At first, the reliability of SNR measure is evaluated between different observations regarding the same tissue ROI in each imaging sequence. Wilcoxon rank sum test [55] and paired sample t-test [56] are used to analyze the inter-observer difference. The significant level is $p = 0.05$.

Secondly, the correlation between SNR values and NIQA results is analyzed by using Pearson correlation coefficient (r_p) [57]. Note that R (http://www.R-project.org) is used to perform the statistical analysis in the first and the second phase.

At last, when SNR_{wm} or SNR_{csf} performs as the surrogate of the reference MIQA indicator, the prediction performance of NIQA models (BIQI, BLIINDS-II, BRISQUE and NIQE) is evaluated with two criteria. One criterion is Pearson linear correlation coefficient (PLCC) which quantifies the prediction accuracy. The other is Spearman rank-order correlation coefficient (SROCC) which measures the prediction monotonicity. The values of PLCC and SROCC range in [0, 1] and a higher value indicates better rating prediction. In the end, Kendall rank-order correlation coefficient (KROCC) is also estimated. Note that before PLCC, SROCC and KROCC estimation, a nonlinear regression is applied to map the predicted scores to the reference score scope by using a five-parameter logistic function as follows,

$$Q(s) = q_1\left(\frac{1}{2} - \frac{1}{1 + e^{q_2(s-q_3)}}\right) + q_4 s + q_5, \tag{1}$$

where s and $Q(s)$ are the input score and the mapped score, and q_i (i=1,2,3,4,5) are optimized during the nonlinear fitting.

C. SNR MEASURE

There are two approaches to measure SNR. One method is to delineate two separate regions from a single image [49], [50]. By taking the signal (S) to be the mean pixel intensity in a tissue ROI (μ_{ROI}) and the noise (δ) to be the standard deviation of pixel intensities values in a background air region (σ_{AIR}), SNR value to the image is quantified as below,

$$SNR_{ROI} = \frac{S}{\delta} = 0.655 \times \frac{\mu_{ROI}}{\sigma_{AIR}}. \tag{2}$$

The factor of 0.655 is due to the Rician distribution of the background noise in a magnitude MR image. It arises because noise variations can be negative and positive. In addition, the delineated background region should be free of anatomical structures and ghosting artifacts.

The other one takes image homogeneity into consideration. If the image is not with so good homogeneity, the SNR may be derived from the following method [58]. First, two images are acquired by consecutive scans with identical receiver and transmitter settings. Then, subtract the images one from the other to generate a difference image. Provided the image has not suffered from ghosting artifacts or any other instability, the only difference between the two original images should be due to noise. Using either of the original images, the signal (S) is defined as the mean pixel intensity value in an *o*riginal tissue ROI (μ_{oROI}), and the noise (δ) is the standard deviation in the same region on the *s*ubtracted image (σ_{sROI}), then SNR is estimated as below,

$$SNR_{ROI} = \frac{S}{\delta} = \sqrt{2} \times \frac{\mu_{oROI}}{\sigma_{sROI}}. \tag{3}$$

The factor $\sqrt{2}$ arises due to the fact that the standard deviation is derived from the subtraction image and not from the original image.

In this study, we use the first approach (2) for SNR measure of MR image quality, since the second method is targeted to estimate the quality of MR images with inhomogeneity.

D. IMAGE PRE-PROCESSING

To each image, pixel intensities are linearly scaled to 8-bit storage ([0, 255]). Two ROIs (WM and CSF) and two AIR regions are outlined. A non-physician (observer A, OA) and a physician (observer B, OB) with more than 15-year experience are asked to localize each region with six points independently. They agree on that the size of outlined tissue regions should be as large as possible. As to T_1, T_2 and T_1C images, they further agree on that ROIs should be homogeneous and keep away from tumor regions. Then, outlined regions are refined by using a close-form curve-fitting method [59] which takes the six points to each outlined region as the control points and utilizes Hermite cubic curve [60] for smooth interpolation between the successive points in a clockwise direction. At last, the refined regions are as input to the in-house built code with MATLAB to quantify the values of WM-based SNR (SNR_{wm}) and CSF-based SNR (SNR_{csf}).

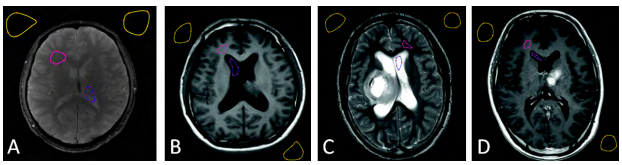


FIGURE 1. Tissue and air regions for SNR measure. (A), (B), (C) and (D) respectively shows one T_2^*, T_1, T_2 and T_1C weighted MR images. Note that red sparkles are primarily points manually localized by observers. Outlined WM, CSF and AIR regions are in closed curves with different colors, pink, blue and yellow, respectively. Note that images have been cropped for display purpose.

Figure 1 shows four cases of outlined regions. Red sparkles are primarily points localized by observers, and refined WM, CSF and AIR regions are in closed curves with pink, blue and yellow lines, respectively.

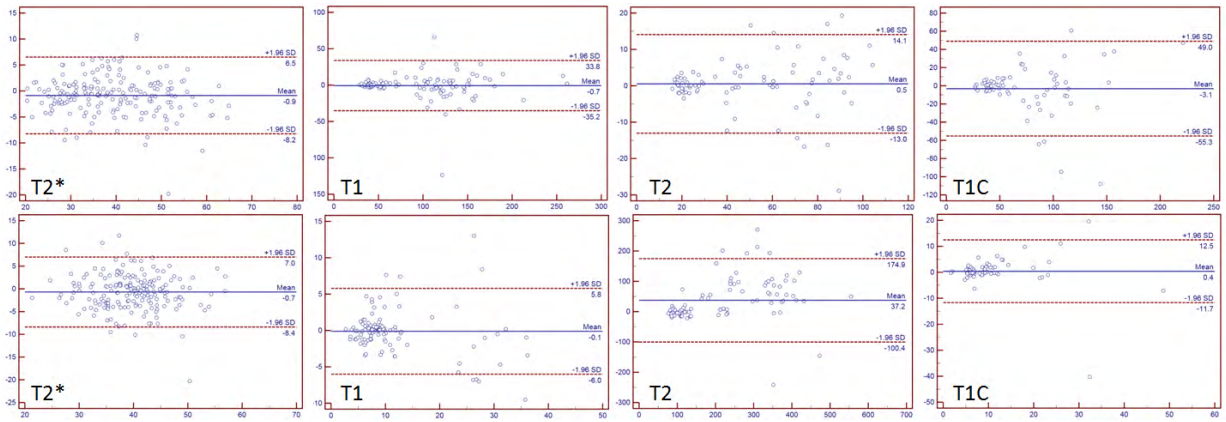


FIGURE 2. Quantitative analysis of the inter-observer reliability of SNR measure using Bland & Altman plots. The solid lines indicate the mean values (Mean) of two SNR observations and the dashed lines indicate the 95% confident interval of the difference between the observations.

E. NO-REFERENCE NIQA ALGORITHMS

This study tests 4 no-reference NIQA models primarily used in computer vision field. These algorithms utilize natural scene statistics (NSS) model to estimate the general quality of natural images. Specially, BIQI [15] needs no prior knowledge of the distorting process after trained and it can be extended to any kinds of distortions. BLIINDS-II [16] needs minimal training and adopts a simple probabilistic model for score prediction. BRISQUE [17] uses the scene statistics of locally normalized luminance coefficients to quantify possible losses of ''naturalness'' in the image due to the presence of distortions, thereby leading to a holistic measure of quality. NIQE [18] constructs a ''quality-aware'' collection of statistical features based on a simple and successful space domain NSS model.

All four NIQA algorithms are implemented with MATLAB and the codes provided by the authors are evaluated without any modifications. For full details of these algorithms, please refer to corresponding literature [15]–[18].

III. RESULTS

A. INTER-OBSERVER RELIABILITY

Figure 2 shows the inter-observer reliability of SNR values by using Bland & Altman plot which illustrates the difference against the average of two observations [61], [62]. In each plot, the horizontal and the vertical axis respectively denotes the average and the difference of two SNR observations. Furthermore, the blue solid line is drawn at the mean difference (Mean) between SNR measures, and the brown dashed ones are drawn at the limits of agreement which are defined as the Mean plus and minus 1.96 times of the standard deviation (SD) of the difference of SNR measure. Fig. 2 indicates that major points (>89%) are localized between the limits of agreement.

Table 1 shows the inter-observer reliability of SNR measure analyzed with Wilcoxon rank sum test (p_w) and paired t-test (p_p). It is observed that the minimal p_w is 0.11 and the minimal p_p is 0.28, both of which is larger than 0.05.

TABLE 1. Inter-observer reliability of SNR measure in each modality.

	T_2^*		T_1		T_2		T_1C	
	WM	CSF	WM	CSF	WM	CSF	WM	CSF
p_w	0.54	0.39	0.88	0.74	0.99	**0.11**	0.69	0.56
p_p	0.41	0.30	0.98	0.59	0.94	**0.28**	0.77	0.46

Meanwhile, it is found that the p-value from SNR_{wm} is larger than that from SNR_{csf} in each imaging sequence.

B. CORRELATION BETWEEN SNR VALUES AND NIQA RESULTS

Table 2 shows Pearson correlation coefficients (r_p) between SNR_{ROI} values and NIQA results, in which the left to the right lists is the names of NIQA models, the r_p between NIQA results and SNR_{wm} and SNR_{csf} values, respectively. The bold-faced values denote $r_p \geq 0.60$. When SNR_{wm} values play as the surrogate of the reference MIQA indicator, the result from BIQI, BLIINDS-II and BRISQUE correlates strongly on T_2^* ($r_p \geq 0.77$), the result from BRISQUE and NIQE correlates strongly on T_1 ($r_p \geq 0.76$), the result from BLIINDS-II correlates strongly on T_2 ($r_p \geq 0.67$), the result from BRISQUE and NIQE correlates moderately on T_1C ($r_p \geq 0.61$) images; while when SNR_{csf} values perform as the MIQA indicator, the result from BLIINDS-II shows moderate correlation on T_2^* ($r_p \geq 0.64$) and T_2 ($r_p \geq 0.61$) images.

C. PERFORMANCE OF NIQA MODELS ON MEDICAL IMAGES

Table 3 shows the prediction accuracy of NIQA models (the left) on the medical MR images when using SNR_{wm} (the middle) and SNR_{csf} (the right) as the reference scores of image quality. The bold-faced values denote the PLCC values larger than 0.60. It is observed that, when SNR_{wm} performs the reference score, BRISQUE demonstrates the notable performance that show good prediction on T_2^*, T_1 and T_1C weighted images, followed by BLIINDS-II

TABLE 2. Quantitative analysis of the correlation between ROI-based SNR values and NIQA results. The left is the NIQA models, the middle and the right are the correlation between NIQA results and the SNR_{wm} and the SNR_{csf} values, respectively. Note that the SNR_{wm} and the SNR_{csf} values respectively perform as the surrogate of the reference MIQA indicator to evaluate the prediction correlation of four no-reference NIQA models primarily used in computer vision field.

	T_2^*		T_1		T_2		T_1C		T_2^*		T_1		T_2		T_1C	
	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB
BIQI	**0.80**	**0.79**	0.18	0.13	0.15	0.20	0.35	0.37	0.57	0.54	0.14	0.10	0.08	0.29	0.07	0.13
BLIINDS-II	**0.77**	**0.78**	0.22	0.21	**0.67**	**0.69**	0.32	0.39	**0.70**	**0.64**	0.01	0.05	**0.69**	**0.61**	0.12	0.17
BRISQUE	**0.80**	**0.80**	**0.76**	**0.80**	0.45	0.35	**0.64**	**0.71**	0.56	0.52	0.18	0.21	0.53	0.26	0.34	0.31
NIQE	0.23	0.24	**0.80**	**0.80**	0.53	0.45	**0.61**	**0.70**	0.33	0.00	0.23	0.26	0.54	0.34	0.31	0.28

TABLE 3. PLCC values of NIQA models when SNR measure performs as the baseline indicator of medical image quality. The left lists the name of NIQA models, the middle and the right are the PLCC value of NIQA models when the SNR_{wm} and the SNR_{csf} measure respectively plays as the surrogate of the reference scores.

	T_2^*		T_1		T_2		T_1C		T_2^*		T_1		T_2		T_1C	
	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB
BIQI	**0.86**	**0.79**	0.18	0.13	0.15	0.24	0.35	0.41	0.59	0.57	0.14	0.10	0.31	0.29	0.37	0.19
BLINDS-II	**0.79**	**0.80**	0.23	0.23	**0.71**	**0.74**	0.50	0.39	**0.70**	**0.69**	**0.65**	**0.62**	**0.73**	**0.66**	0.14	0.36
BRISQUE	**0.81**	**0.83**	**0.76**	**0.90**	0.45	0.44	**0.62**	**0.77**	0.57	0.53	0.41	0.39	0.53	0.26	0.45	0.42
NIQE	0.31	0.30	**0.80**	**0.83**	0.53	0.47	**0.64**	**0.70**	0.33	0.43	0.38	0.35	0.54	0.34	0.46	0.28

TABLE 4. SROCC values of NIQA models when SNR measure performs as the baseline indicator of medical image quality. The left lists the name of NIQA models, the middle and the right are the SROCC value of NIQA models when the SNR_{wm} and the SNR_{csf} measure respectively plays as the surrogate of the reference scores.

	T_2^*		T_1		T_2		T_1C		T_2^*		T_1		T_2		T_1C	
	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB
BIQI	**0.73**	**0.66**	0.19	0.16	0.07	0.26	0.41	0.39	0.49	0.47	0.25	0.17	0.15	0.20	0.23	0.29
BLINDS-II	**0.79**	**0.82**	0.35	0.34	**0.72**	**0.72**	0.46	0.33	**0.71**	**0.69**	0.25	0.55	**0.75**	**0.70**	0.02	0.14
BRISQUE	**0.82**	**0.84**	**0.73**	**0.81**	0.54	0.52	**0.78**	**0.78**	0.53	0.50	0.37	0.36	0.59	0.46	0.36	0.48
NIQE	0.32	0.35	**0.80**	**0.78**	0.53	0.51	**0.79**	**0.76**	0.35	0.44	0.27	0.30	0.54	0.45	0.46	0.51

TABLE 5. KROCC values of NIQA models when SNR measure performs as the baseline indicator of medical image quality. The left lists the name of NIQA models, the middle and the right are the KROCC value of NIQA models when the SNR_{wm} and the SNR_{csf} measure respectively plays as the surrogate of the reference scores.

	T_2^*		T_1		T_2		T_1C		T_2^*		T_1		T_2		T_1C	
	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB	OA	OB
BIQI	0.49	0.44	0.13	0.13	0.01	0.02	0.11	0.24	0.34	0.31	0.17	0.11	0.02	0.08	0.14	0.03
BLINDS-II	**0.58**	**0.62**	0.28	0.29	0.40	0.42	0.37	0.48	**0.52**	**0.50**	0.33	0.36	0.41	0.43	0.09	0.03
BRISQUE	**0.61**	**0.62**	**0.52**	**0.54**	0.42	0.38	**0.56**	**0.54**	0.37	0.35	0.23	0.24	0.37	0.41	0.30	0.36
NIQE	0.21	0.23	**0.56**	**0.57**	0.38	0.32	**0.53**	**0.51**	0.23	0.29	0.19	0.22	0.38	0.32	0.34	0.38

(T_2^* and T_2), NIQE (T_1 and T_1C) and BIQI (T_2^*); while when using SNR_{csf} as the reference score, only BLIINDS-II shows the remarkable prediction over T_2^*, T_1 and T_2 weighted images.

Table 4 demonstrates the SROCC values of NIQA models (the left) on MR images when using SNR_{wm} (the middle) and SNR_{csf} (the right) as the reference. The bold-faced values in red and blue denote the SROCC values larger than 0.60. When using SNR_{wm} as the reference, BRISQUE again achieves good prediction on T_2^*, T_1 and T_1C weighted images, followed by BLIINDS-II (T_2^* and T_2), NIQE (T_1 and T_1C) and BIQI (T_2^*); while when using SNR_{csf} as the reference, only BLIINDS-II shows notable results over T_2^* and T_2 weighted images.

Table 5 summerizes the KROCC values of NIQA models (the left) when using SNR_{wm} (the middle) and SNR_{csf}

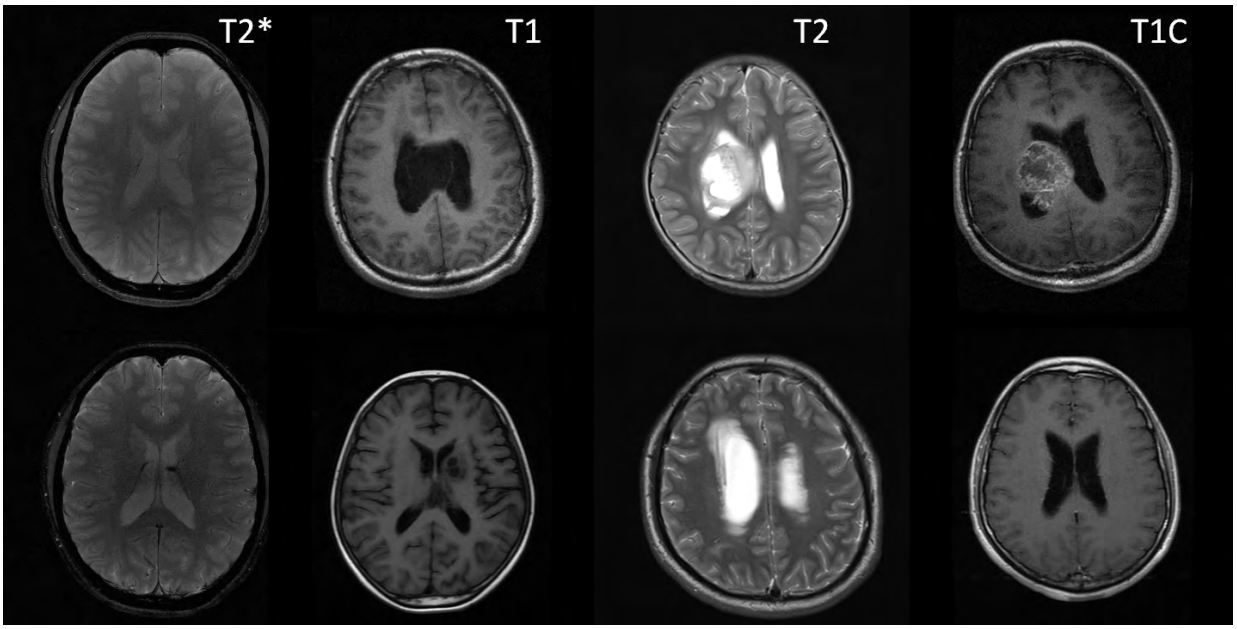


FIGURE 3. Perceived medical image quality. To each imaging sequences (T_2^*, T_1, T_2 and T_1C), two MR images are illustrated. Note that images have been cropped for display.

(the right) as the reference. The bold-faced values in red and blue denote the values larger than 0.50. When using SNR_{wm} as the reference, BRISQUE obtains fair prediction on T_2^*, T_1 and T_1C weighted images, followed by NIQE (T_1 and T_1C) and BLIINDS-II (T_2^*); while when SNR_{csf} values perform as the reference scores, only BLIINDS-II shows fair results over T_2^* weighted images. It should be mentioned that only KROCC values from BRISQUE results larger than 0.60 in terms of SNR_{wm} scores on T_2^* weighted images.

D. PERCEIVED IMAGE QUALITY

Figure 3 demonstrates eight MR images for medical image quality comparison. In each sequence, two images are selected. Based on two SNR measures, to the T_2^*, T_1, T_1C and T_2 weighted images in the top row, the average value of SNR_{wm} and SNR_{csf} is 42.10 and 42.05, 31.97 and 5.51, 27.03 and 5.80, and 104.28 and 396.14, respectively. Similarly, the average value of SNR_{wm} and SNR_{csf} is 30.82 and 40.28, 63.00 and 12.55, 98.56 and 11.95, and 43.43 and 309.67 to the MR images in the bottom row. Visually, the CSF region is as dark as the background air region in T_1 and T_1C MR images, and the intensity of CSF region changes dramatically from T_1, T_1C to T_2^* and T_2 imaging.

To each image, the predicted score of NIQA models after non-linear fitting and its baseline score of SNR_{wm} are shown in pair in Figure 4, where the horizontal axis indicates SNR measures and the vertical axis indicates nonlinearly fitted NIQA scores. For differentiation, the points in red, green, blue and pink colors stand for the results from BIQI, BLIINDS-II, BRISQUE and NIQE, respectively. Moreover, the markers of circle and square correspond to the first (OA) and the second (OB) observation of SNR measure. The line

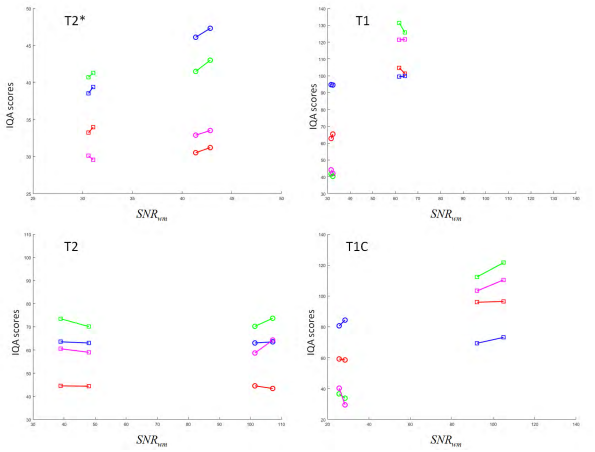


FIGURE 4. Comparison of the predicted score of NIQA models after non-linear fitting and the baseline score of SNR_{wm}. The points in red, green, blue and pink colors stand for the results from BIQI, BLIINDS-II, BRISQUE and NIQE, respectively. Moreover, the markers of circle and square correspond to the first (OA) and the second (OB) observation of SNR measure. The line connecting two points further quantifies the difference of two baseline SNR measures (absolute horizontal distance) and two predicted score after fitting with different SNR observations as the reference (absolute vertical distance).

connecting two points further quantifies the difference of two baseline SNR measures (absolute horizontal distance) and two non-linearly predicted scores with different SNR measure as the reference (absolute vertical distance). As shown in Figure 4, the absolute difference of two SNR measures are less than 15 in each plot. Meanwhile, the nonlinearly fitted scores of each NIQA model are close to each other when using SNR_{wm} values from OA and OB as the surrogate indicator of medical image quality.

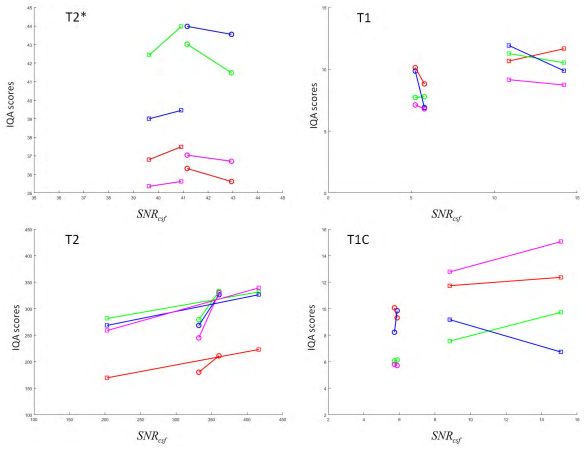


FIGURE 5. Comparison of the predicted score of NIQA models after non-linear fitting and the baseline score of SNR_{csf}. The points in red, green, blue and pink colors stand for the results from BIQI, BLIINDS-II, BRISQUE and NIQE, respectively. Moreover, the markers of circle and square correspond to the first (OA) and the second (OB) observation of SNR measure. The line connecting two points further quantifies the difference of two baseline SNR measures (absolute horizontal distance) and two predicted score after fitting with different SNR observations as the reference (absolute vertical distance).

Figure 5 compares the non-linearly fitted scores of NIQA models and the baseline scores of SNR_{csf}. As aforementioned, the axis, the markers, the colors and the line are interpreted in the same way. It suggests that the absolute difference of two SNR measures are less than 10, except the difference of 200 is found in the second case of T_2 weighted MR image. Meanwhile, the SNR_{csf} is less than 16 in T_1 and T_1C weighted MR images, while it increases to 45 in T_2^* images and then reaches 430 in T_2 weighted MR images.

IV. DISCUSSION

This study has verified the inter-observer reliability of SNR measure regarding different tissue ROIs in each of four MR imaging sequences. The correlation analysis and the prediction performance of four NIQA models by using SNR values as the baseline indicator have further verified the feasibility of two models, BLIINDS-II and BRISQUE, for automated, objective and no-reference MIQA applications. Accordingly, SNR measure can conditionally perform as a baseline indicator for the future development of MIQA algorithms.

The inter-observer reliability of SNR measure is validated with Bland & Altman plot (Figure 2) and Pearson correlation analysis (Table 1). And thereby, a non-physician can independently perform the SNR measure as well as an experienced physician does. Moreover, both p_w and p_p from SNR_{wm} are respectively larger than those from SNR_{csf} in each imaging protocol. This result indicates that WM is a relatively better choice than CSF as the tissue ROI in the quality estimation of MR brain images. That is mainly because the intensity of CSF regions in different sequences changes more obviously than those of WM regions (Figure 3), which influences the quantification of SNR_{csf} values (Figure 5).

The Pearson correlation analysis between SNR values and NIQA results suggests a good correlation between BLIINDS-II and SNR_{csf} on T_2^* and T_2 images (Table 2). Since in T_2^* and T_2 weighted MR images, CSF region presents relatively higher intensity over other tissue regions that enhances the reliability of SNR_{csf} measure (Figure 3). In comparison to CSF region, WM is once again proved to be a relatively better choice in the quality estimation of brain MR images, since more NIQA models correlate well with SNR_{wm}. This phenomenon is verified in the perceived image quality analysis (Figure 3, 4 and 5). Moreover, the prediction performance of NIQA models on medical images has been investigated (Table 3, 4 and 5). Notably, BRISQUE works well on T_2^*, T_1 and T_1C sequences, while BLIINDS-II shows superiority on T_2^* and T_2^* images independent of the selection of tissue regions. Therefore, with necessary modifications, it is possible to transfer NIQA models primarily designed in computer vision field for medical image quality estimation [48].

SNR is an important quantity and two preconditions should be considered for its measure based on the "two-region" approach. First, it requires a spatially homogeneous distribution of noise over the whole image. Second, the statistical intensity distribution of the noise should be known, so that the noise properties measured in a background area can be used to deduce the noise distribution overlaying the anatomic structures in the foreground. Fortunately, these preconditions can be laid aside because of the development of advanced MR imaging sequences [52], [54], [63]–[67].

This study involves 411 MR images from four imaging sequences. We take advantage of the knowledge that SNR is daily used in clinic and concern the inter-observer reliability of SNR measure. After the inter-observer reliability is verified, we further consider the feasibility of using SNR as the baseline indicator for the development of MIQA methods. Four off-the-shelf NIQA models are evaluated and the result is promising. First, NIQA models show good correlation to SNR_{wm} measure on one to three protocols. That means, with considerably modifications, NIQA models can be delicately applied to specific MR imaging modalities for quality estimation. Second, the images in each modality are not acquired with the same imaging sequence, such as different TE in T_2^*, slice thickness and matric size in T_1, T_2 and T_1C. Accordingly, SNR measure has certain ability of generalization that places great demands on the development of MIQA models.

Further improvement can be made from several points. This study measures local image quality and the SNR values correspond to the delineated regions in the brain MR image. In fact, SNR can be formulated from global signal intensity by replacing the signal from a tissue region with the signal from the whole object region [40]. More general approaches include utilizing Shannon's theory to represent the image content and to model the spatial spectral power density of the image [41] or analyzing the air background of structural brain magnitude images [68]. Notably, some studies explore to connect SNR values and the diagnostic

accuracy or detectability [30], [36], [69], since the ultimate purpose of medical imaging is for detectability or diagnosis of a certain disease.

V. CONCLUSION

This study has verified the inter-observer reliability of ''two-region'' SNR measure regarding different tissue regions of interest in the brain MR images. The correlation analysis of SNR values and NIQA results indicates that SNR can perform as a baseline indicator to assess MR image quality. Furthermore, by using SNR values as the surrogate of image quality scores, the prediction performance of NIQA models suggests that both BRISQUE and BLIINDS-II can be conditionally used for automated, objective and no-reference MIQA. This study presents the first attempt of using SNR to bridge the gap between NIQA and MIQA. A large-scale experiment should be further conducted to confirm the conclusion in this work.

ACKNOWLEDGEMENT

Zhicheng Zhang and Shaode Yu contribute equally in this work. The authors would like to thank the editor and reviewers for their valuable advices that have helped to improve the paper quality.

REFERENCES

[1] Q. Wu *et al.*, ''Blind image quality assessment based on multichannel feature fusion and label transfer,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 3, pp. 425–440, Mar. 2016.

[2] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot, ''No-reference image blur assessment based on discrete orthogonal moments,'' *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 39–50, Jan. 2016.

[3] Y. Zhang, Y. Fang, W. Lin, X. Zhang, and L. Li, ''Backward registration-based aspect ratio similarity for image retargeting quality assessment,'' *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4286–4297, Sep. 2016.

[4] K. Ma *et al.*, ''Waterloo exploration database: New challenges for image quality assessment models,'' *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 1004–1016, Feb. 2017.

[5] T.-J. Liu, K.-H. Liu, J. Y. Lin, W. Lin, and C.-C. J. Kuo, ''A paraboost method to image quality assessment,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 1, pp. 107–121, Jan. 2017.

[6] S. Yu, S. Wu, L. Wang, F. Jiang, Y. Xie, and L. Li, ''A shallow convolutional neural network for blind image sharpness assessment,'' *PLoS ONE*, vol. 12, no. 5, p. e0176632, 2017.

[7] K. Gu, L. Li, H. Lu, X. Min, and W. Lin, ''A fast reliable image quality predictor by fusing micro- and macro-structures,'' *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3903–3912, May 2017.

[8] K. Ma, W. Liu, T. Liu, Z. Wang, and D. Tao, ''dipIQ: Blind image quality assessment by learning-to-rank discriminable image pairs,'' *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3951–3964, Aug. 2017.

[9] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, ''Image quality assessment: From error visibility to structural similarity,'' *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[10] L. Zhang, L. Zhang, X. Mou, and D. Zhang, ''FSIM: A feature similarity index for image quality assessment,'' *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[11] L. Li, D. Wu, J. Wu, H. Li, W. Lin, and A. C. Kot, ''Image sharpness assessment by sparse representation,'' *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1085–1097, Jun. 2016.

[12] L. Li, W. Xia, W. Lin, Y. Fang, and S. Wang, ''No-reference and robust image sharpness evaluation based on multiscale spatial and spectral features,'' *IEEE Trans. Multimedia*, vol. 19, no. 5, pp. 1030–1040, May 2017.

[13] G. Dai, Z. Wang, Y. Li, Q. Chen, S. Yu, and Y. Xie, ''Evaluation of no-reference models to assess image sharpness,'' in *Proc. Int. Conf. Inf. Autom. (ICIA)*, Jul. 2017, vol. 19. no. 5, pp. 683–687.

[14] R. A. Manap and L. Shao, ''Non-distortion-specific no-reference image quality assessment: A survey,'' *Inf. Sci.*, vol. 301, pp. 141–160, Apr. 2015.

[15] A. K. Moorthy and A. C. Bovik, ''A two-step framework for constructing blind image quality indices,'' *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.

[16] M. A. Saad, A. C. Bovik, and C. Charrier, ''DCT statistics model-based blind image quality assessment,'' in *Proc. IEEE ICIP*, vol. 1. Sep. 2011, pp. 3093–3096.

[17] A. Mittal, A. K. Moorthy, and A. C. Bovik, ''No-reference image quality assessment in the spatial domain,'' *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[18] A. Mittal, R. Soundararajan, and A. C. Bovik, ''Making a 'completely blind' image quality analyzer,'' *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[19] D. M. Chandler *et al.*, ''Seven challenges in image quality assessment: Past, present, and future research,'' *ISRN Signal Process.*, vol. 2013, Nov. 2013, Art. no. 905685.

[20] G. Yue, C. Hou, K. Gu, S. Mao, and W. Zhang, ''Biologically inspired blind quality assessment of tone-mapped images,'' *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2525–2536, Mar. 2018.

[21] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. Lin, ''A patch-structure representation method for quality assessment of contrast changed images,'' *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2387–2390, Dec. 2015.

[22] K. Gu, G. T. Zhai, and M. Lin, ''The analysis of image contrast: From quality assessment to automatic enhancement,'' *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 284–297, Jan. 2015.

[23] F. Li, F. Shao, Q. Jiang, R. Fu, G. Jiang, and M. Yu, ''Local and global sparse representation for no-reference quality assessment of stereoscopic images,'' *Inf. Sci.*, vol. 422, pp. 110–121, Jan. 2018.

[24] L. S. Chow and R. Paramesran, ''Review of medical image quality assessment,'' *Biomed. Signal Process. Control*, vol. 27, no. 1, pp. 145–154, 2016.

[25] G. Y. Sandhu, C. Li, O. Roy, S. Schmidt, and N. Duric, ''Frequency domain ultrasound waveform tomography: Breast imaging using a ring transducer,'' *Phys. Med. Biol.*, vol. 60, no. 14, p. 5381, 2015.

[26] M. Ahmad, M. Bazalova-Carter, R. Fahrig, and L. Xing, ''Optimized detector angular configuration increases the sensitivity of X-ray fluorescence computed tomography (XFCT),'' *IEEE Trans. Med. Imag.*, vol. 34, no. 5, pp. 1140–1147, May 2015.

[27] Z. Zhang, S. Yu, X. Liang, Y. Zhu, and Y. Xie, ''A novel design of ultrafast micro-CT system based on carbon nanotube: A feasibility study in phantom,'' *Phys. Medica*, vol. 32, no. 10, pp. 1302–1307, 2016.

[28] J. I. Peltonen, T. Mäkelä, A. Sofiev, and E. Salli, ''An automatic image processing workflow for daily magnetic resonance imaging quality assurance,'' *J. Digit. Imag.*, vol. 30, no. 2, pp. 163–171, 2017.

[29] C. H. Jenkins, L. Xing, and B. P. Fahimian, ''Automating position and timing quality assurance for high dose rate brachytherapy using radioluminescent phosphors and optical imaging,'' *Brachytherapy*, vol. 15, no. 1, p. S28, 2016.

[30] B. L. Eck *et al.*, ''Computational and human observer image quality evaluation of low dose, knowledge-based CT iterative reconstruction,'' *Med. Phys.*, vol. 42, no. 10, pp. 6098–6111, 2015.

[31] M. Razaak, M. G. Martini, and K. Savino, ''A study on quality assessment for medical ultrasound video compressed via HEVC,'' *IEEE J. Biomed. Health Inform.*, vol. 18, no. 5, pp. 1552–1559, Sep. 2014.

[32] S. Yu *et al.*, ''Linear-fitting-based similarity coefficient map for tissue dissimilarity analysis in T_2^*-w magnetic resonance imaging,'' *Chin. Phys. B*, vol. 24, no. 12, p. 128711, 2015.

[33] R. Zhang, W. Zhou, Y. Li, S. Yu, and Y. Xie, ''Nonrigid registration of lung CT images based on tissue features,'' *Comput. Math. Methods Med.*, vol. 1, no. 1, p. 834192, 2013.

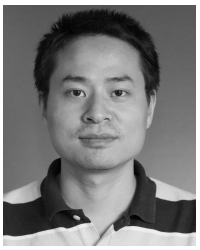
[34] S. Yu, R. Zhang, S. Wu, J. Hu, and Y. Xie, ''An edge-directed interpolation method for fetal spine MR images,'' *Biomed. Eng. Online*, vol. 12, no. 1, p. 102, 2013.

[35] X. Li, W. Huang, and W. D. Rooney, ''Signal-to-noise ratio, contrast-to-noise ratio and pharmacokinetic modeling considerations in dynamic contrast-enhanced magnetic resonance imaging,'' *Magn. Reson. Imag.*, vol. 30, no. 9, pp. 1313–1322, 2012.

[36] P. C. Cosman, R. M. Gray, and R. A. Olshen, ''Evaluating quality of compressed medical images: SNR, subjective rating, and diagnostic accuracy,'' *Proc. IEEE*, vol. 82, no. 6, pp. 919–932, Jun. 1994.

[37] Z. Cao, J. Park, Z.-H. Cho, and C. M. Collins, ''Numerical evaluation of image homogeneity, signal-to-noise ratio, and specific absorption rate for human brain imaging at 1.5, 3, 7, 10.5, and 14T in an 8-channel transmit/receive array,'' *J. Magn. Reson. Imag.*, vol. 41, no. 5, pp. 1432–1439, 2015.

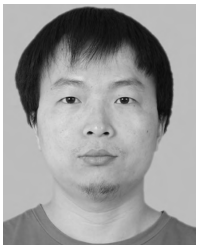
[38] J. Hausleiter *et al.*, "Image quality and radiation exposure with prospectively ECG-triggered axial scanning for coronary CT angiography: The multicenter, multivendor, randomized PROTECTION-III study," *JACC, Cardiovascular Imag.*, vol. 5, no. 5, pp. 484–493, 2012.
[39] B. Bischoff *et al.*, "Comparison of sequential and helical scanning for radiation dose and image quality: Results of the prospective multicenter study on radiation dose estimates of cardiac CT angiography (protection) I study," *Amer. J. Roentgenol.*, vol. 194, no. 6, pp. 1495–1499, 2010.
[40] M. Welvaert and Y. Rosseel, "On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data," *PLoS ONE*, vol. 8, no. 11, p. e77089, 2013.
[41] M. Fuderer, "The information content of MR images," *IEEE Trans. Med. Imag.*, vol. MI-7, no. 4, pp. 368–380, Dec. 1988.
[42] A. Geissler, A. Gartus, T. Foki, A. R. Tahamtan, R. Beisteiner, and M. Barth, "Contrast-to-noise ratio (CNR) as a quality parameter in fMRI," *J. Magn. Reson. Imag.*, vol. 25, no. 6, pp. 1263–1270, 2007.
[43] J. M. P. Dias, C. M. Oliveira, and L. A. da Silva Cruz, "Retinal image quality assessment using generic image quality indicators," *Inf. Fusion*, vol. 19, pp. 73–90, Sep. 2014.
[44] T. Niu and L. Zhu, "Scatter correction for full-fan volumetric CT using a stationary beam blocker in a single full scan," *Med. Phys.*, vol. 38, no. 11, pp. 6027–6038, 2011.
[45] T. Niu, M. Sun, J. Star-Lack, H. Gao, Q. Fan, and L. Zhu, "Shading correction for on-board cone-beam CT in radiation therapy using planning MDCT images," *Med. Phys.*, vol. 37, no. 10, pp. 5395–5406, 2010.
[46] X. Liang *et al.*, "Iterative image-domain ring artifact removal in cone-beam CT," *Phys. Med. Biol.*, vol. 62, no. 1, pp. 5276–5292, 2017.
[47] L. S. Chow, H. Rajagopal, R. Paramesran, and Alzheimer's Disease Neuroimaging Initiative, "Correlation between subjective and objective assessment of magnetic resonance (MR) images," *Magn. Reson. Imag.*, vol. 34, no. 6, pp. 820–831, 2016.
[48] L. S. Chow and H. Rajagopal, "Modified-BRISQUE as no reference image quality assessment for structural MR images," *Magn. Reson. Imag.*, vol. 43, pp. 74–87, Nov. 2017.
[49] R. M. Henkelman, "Erratum: Measurement of signal intensities in the presence of noise in MR images [Med. Phys. 12, 232 (1985)]," *Med. Phys.*, vol. 12, no. 2, pp. 232–233, 1985.
[50] L. Kaufman, D. M. Kramer, L. E. Crooks, and D. A. Ortendahl, "Measuring signal-to-noise ratios in MR imaging," *Radiology*, vol. 173, no. 1, pp. 265–267, 1989.
[51] B. W. Murphy, P. L. Carson, J. H. Ellis, Y. T. Zhang, R. J. Hyde, and T. L. Chenevert, "Signal-to-noise measures for magnetic resonance imagers," *Magn. Reson. Imag.*, vol. 11, no. 3, pp. 425–428, 1993.
[52] C. D. Constantinides, E. Atalar, and E. R. McVeigh, "Signal-to-noise measurements in magnitude images from NMR phased arrays," *Magn. Reson. Med.*, vol. 38, no. 1, pp. 852–857, 1997.
[53] P. Shokrollahi, J. M. Drake, and A. A. Goldenberg, "Signal-to-noise ratio evaluation of magnetic resonance images in the presence of an ultrasonic motor," *Biomed. Eng. Online*, vol. 16, no. 1, p. 45, 2017.
[54] S. B. Reeder *et al.*, "Practical approaches to the evaluation of signal-to-noise ratio performance with parallel imaging: Application with cardiac imaging and a 32-channel cardiac coil," *Magn. Reson. Med.*, vol. 54, no. 3, pp. 748–754, 2005.
[55] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, vol. 1, no. 6, pp. 80–83, 1945.
[56] D. W. Zimmerman, "Teacher's Corner: A note on interpretation of the paired-samples *t* test," *J. Edu. Behav. Stat.*, vol. 22, no. 3, pp. 349–360, 1997.
[57] F. Galton, "Regression towards mediocrity in hereditary stature," *J. Anthropol. Inst. Great Britain Ireland*, vol. 15, no. 1, pp. 246–263, 1886.
[58] M. J. Firbank, A. Coulthard, R. M. Harrison, and E. D. Williams, "A comparison of two methods for measuring the signal to noise ratio on MR images," *Phys. Med. Biol.*, vol. 44, no. 12, p. N261, 1999.
[59] W. Zhou and Y. Xie, "Interactive contour delineation and refinement in treatment planning of image-guided radiation therapy," *J. Appl. Clin. Med. Phys.*, vol. 15, no. 1, pp. 141–166, 2014.
[60] L. Lu, "A note on curvature variation minimizing cubic Hermite interpolants," *Appl. Math. Comput.*, vol. 259, no. 1, pp. 596–599, 2015.
[61] D. G. Altman and J. M. Bland, "Measurement in medicine: The analysis of method comparison studies," *Statistician*, vol. 32, no. 1, pp. 307–317, 1983.
[62] D. Giavarina, "Understanding bland altman analysis," *Biochem. Medica*, vol. 25, no. 2, pp. 141–151, 2015.
[63] D. K. Sodickson and W. J. Manning, "Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays," *Magn. Reson. Med.*, vol. 38, no. 4, pp. 591–603, 1997.
[64] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, 1999.
[65] W. A. Willinek, J. Gieseke, M. von Falkenhausen, B. Neuen, H. H. Schild, and C. K. Kuhl, "Sensitivity encoding for fast MR imaging of the brain in patients with stroke," *Radiology*, vol. 228, no. 3, pp. 669–675, 2003.
[66] Q. Chen *et al.*, "On improving temporal and spatial resolution of 3D contrast-enhanced body MR angiography with parallel imaging," *Radiology*, vol. 231, no. 3, pp. 893–899, 2004.
[67] P. Hunold, S. Maderwald, M. E. Ladd, V. Jellus, and J. Barkhausen, "Parallel acquisition techniques in cardiac cine magnetic resonance imaging using TrueFISP sequences: Comparison of image quality and artifacts," *J. Magn. Reson. Imag.*, vol. 20, no. 3, pp. 506–511, 2004.
[68] B. Mortamet *et al.*, "Automatic quality assessment in structural brain magnetic resonance imaging," *Magn. Reson. Med.*, vol. 62, no. 2, pp. 365–372, 2009.
[69] J. Galbally, S. Marcel, and J. Fierrez, "Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 710–724, Feb. 2014.



ZHICHENG ZHANG received the B.S. degree from Sun Yat-sen University, Guangzhou, China in 2010. He is currently pursuing the Ph.D. degree with the Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China. He received the Best Oral Award in AOCMP 2015. From 2017 to 2018, he has been a Visiting Ph.D. student with the Virginia Tech-Wake Forest University, School of Biomedical Engineering and Sciences, Virginia Polytechnic Institute and State University, USA. His research interests are medical image analysis, computer vision and deep learning.

GUANGZHE DAI is currently a Senior Student with the Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang, China. Since 2016, she has been Visiting Student with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. Her research interests mainly include image quality assessment and machine learning.

XIAOKUN LIANG received the B.S. degree from Southern Medical University, Guangzhou, China, in 2013, and the M.S. degree from Guangdong Medical University, Dongguan, China, in 2016. He is currently pursuing the Ph.D. degree with the Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China. His research interests focus on medical image analysis, medical physics and image guided radiotherapy.



SHAODE YU received the B.S. and M.S. degrees from Beijing Normal University, Beijing, China, in 2008 and 2011, respectively. He is currently pursuing the Ph.D. degree with the Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China. In 2014, he was a Visiting Scholar with the School of Medicine, Wayne State University, USA. He is currently an Assistant Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interests include machine learning and biomedical engineering.

LEIDA LI (M'14) received the B.S. and Ph.D. degrees from Xidian University in 2004 and 2009, respectively. In 2008, he was a visiting Ph.D. student at the Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Taiwan. From 2014 to 2015, he was a Visiting Research Fellow with the Rapid-Rich Object Search (ROSE) Laboratory, School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore. From 2016 to 2017, he was a Senior Research Fellow at the ROSE Laboratory, NTU, Singapore. He is currently a Full Professor and the Ph.D. supervisor with the School of Information and Control Engineering, China University of Mining and Technology, China. His research interests include multimedia quality assessment, information hiding, and image forensics.

YAOQIN XIE received the B.S., M.S., and Ph.D. degrees from Tsinghua University, Beijing, China, in 1995, 1998 and 2002, respectively. From 2002 to 2010, he was a Lecturer at the School of Physics, Peiking University, China. From 2006 to 2008, he was a Research Fellow with the Department of Radiation Oncology, Stanford Medical School, USA. Since 2010, he has been a Full Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interests include medical image analysis, medical physics, and image-guided radiotherapy.