IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 28, 2020, accepted August 25, 2020, date of publication August 31, 2020, date of current version September 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3020369

HS–GS: A Method for Multicenter MR Image Standardization

GUOHUA ZHAO[1,2,3], JIE BAI[2], PEI PEI WANG[2], GUAN YANG[4], LEI SHI[5], YONGCAI TAO[1], YUSONG LIN[3,5,6], (Member, IEEE), AND JINGLIANG CHENG[2]

[1]School of Information Engineering, Zhengzhou University, Zhengzhou 450000, China
[2]Department of Magnetic Resonance Imaging, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450004, China
[3]Collaborative Innovation Center for Internet Healthcare, Zhengzhou University, Zhengzhou 450052, China
[4]School of Computer Science, Zhongyuan University of Technology, Zhengzhou 450007, China
[5]School of Software, Zhengzhou University, Zhengzhou 450002, China
[6]Hanwei IoT Institute, Zhengzhou University, Zhengzhou 450002, China

Corresponding authors: Jingliang Cheng (fccchengjl@zzu.edu.cn) and Yusong Lin (yslin@ha.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 81772009, and in part by the Scientific and Technological Research Project of Henan Province under Grant 182102310162.

ABSTRACT The access to and sharing of medical image data is essential to accelerate the research progress of complex diseases and sudden disease outbreaks. Multicenter image data is collected from different medical institutions, and the contrast and brightness of the images are significantly different, making it difficult to use the images directly. Herein, we introduce a standardized method based on magnetic resonance imaging, referred to as Histogram specification–grid search (HS–GS), which is mainly used to eliminate differences in image contrast and brightness. A Gaussian probability density function with adjustable parameters is used to generate the cumulative distribution function, and the transfer function required for the HS mapping is constructed to obtain standardized image sets based on the controllable parameters. The image sets are used to perform the GS task of radiomics classification to find the optimal controllable parameter combination and classification results, and then obtain the optimal standardized image sets. We used HS–GS to test and verify the predictive ability of the standardized mixed image sets for glioma grading, and compared it with existing methods. The experiments indicate that the standardized image sets generated by the HS–GS algorithm retain excellent stability after mixing and also show excellent classification performance. This novel image set standardization technique has proven to be a promising solution for integration into medical expert systems.

INDEX TERMS Data standardization, histogram specification, grid search, MRI, radiomics.

I. INTRODUCTION

Although pathological examination is considered as the gold standard for diagnosis, it is often obtained too late [1]. Medical imaging systems can acquire the status of patients' lesions earlier, which plays a key role in clinical practice [2]. In addition, the imaging system can accurately describe anatomical and physiological characteristics in a nearly non-invasive manner, thereby allowing clinicians to better understand complex or rare diseases. Different medical imaging technologies, such as magnetic resonance imaging (MRI) and computer tomography (CT), can obtain different types of medical images. The images typically contain a significant amount of information; however, the information obtained by relying on the clinician's observation is minimal [3]. In recent years, the analysis of big data in medicine using artificial intelligence (AI) has received increasing attention, particularly with research on medical images becoming the focus of attention [4]–[7]. Compared with other modalities of medical data, medical images contain rich information and are widely used. Furthermore, as the medical community has found that using machine learning methods (radiomics [8]–[10]) to process medical image data shows great potential, computer-aided analysis has expanded to cover diseases of various organs throughout the body, and some studies have achieved significant results. However, this area is constrained by the scale of data. By relying only on image data from a single source for analysis, the understanding of complex

The associate editor coordinating the review of this manuscript and approving it for publication was Michele Nappi.

diseases becomes highly limited. Therefore, it is crucial to merge data from different medical institutions to form a larger data volume.

MRI is the most important method for obtaining soft tissue imaging, and also the main method for early detection of most complex diseases [11]. Compared with CT, high-field intensity MRI (T = 3.0) can provide higher resolution and contrast scan images and performs better in tumor detection and diagnosis, especially in determining the degree of tumor invasion. MRI is the main source of big medical data [12]. In fact, different medical institutions usually use magnetic resonance equipment provided by different manufacturers, and the parameters of the equipment configuration are also different; hence, almost all the collected magnetic resonance data are unique [13]. Even if data sharing is achieved, the merged image data will have large internal differences, and these data can still be regarded as individual ''data islands.'' Moreover, these differences will not only affect the clinician's subjective judgment of the image but also affect the parameter measurement during data post-processing. The most serious result is misdiagnosis. Therefore, it is crucial to eliminate the internal differences of data and realize data standardization [14], [15]. The advantages are not limited to revealing certain features between medical images but include providing doctors with richer and more valuable information, helping them identify, diagnose and treat abnormalities more effectively. Moreover, it can be used to build a unified standard large-scale medical image database in order to gain a deeper understanding of the imaging and pathological features of complex and rare diseases [16].

Image enhancement technology is generally used to improve the visual effect of images and improve the image quality [17]. Compared with the original image, the enhanced image can describe more image details and obtain previously hidden or ignored features. In the field of medical imaging, most enhancement techniques are used for image preprocessing, providing clinicians with clearer images, and also providing high-quality original images for computer-aided analysis systems. HS is a classic image enhancement algorithm [18]. Through the transfer function, it can transform the input original image into an image that better conforms to people's subjective expectations. HS can also be used for image standardization. First, the clinician manually selects a set of representative images from the data set and constructs a transfer function based on this set of representative images [19]. All data sets are subjected to histogram matching according to the transfer function, and the images have a highly similar histogram distribution. Although this method is simple and effective, the manual selection of representative images by clinicians involves significant subjectivity, which can cause some uncertainty. In addition, there is no guarantee that the selected representative image is ''standard,'' or that the transformed image is optimal. Introducing relevant evaluation indexes to evaluate the quality of reconstructed images is the most effective solution in this regard [20]. However, there is currently no unified theory of image quality measurement to define and evaluate the effect of image standardization.

In this paper we propose a standardized method for magnetic resonance images, referred to as HS–GS, which is mainly used for multi-center magnetic resonance data. We introduce a Gaussian function with adjustable parameters as the representative image in the HS. On this basis, the cumulative distribution function (CDF) is calculated to construct the transfer function in the HS, thus guiding the histogram mapping. In the next step, HS–GS is applied to clinical problems involving MRI, namely, MRI-based diagnosis of glioma grades. Since the verification mode of result feedback often requires significant human–computer interaction, we develop a dedicated GS script. The grading results predicted by radiomics are used as indicators to evaluate the standardization results of images, and the parameters in the HS can be adjusted in order to find the best grading results and best standardized images.

The main contributions of HS–GS can be briefly summarized as follows: (1) In the HS, the representative image used to construct the CDF is replaced by the Gaussian function, thus replacing the clinician's selection. We have effectively avoided errors caused by doctors' subjective judgments, and, simultaneously, added parameters that can be flexibly adjusted according to specific clinical application scenarios and data to achieve a quantitative description of image standardization. (2) Matching GS strategies to reduce unnecessary human–computer interaction and realize search automation are provided.

The structure of this manuscript is as follows: Section II outlines application scenarios along with related literature and existing methods; Section III describes the principles of HS–GS data normalization and the automated GS process for data normalization. The experimental data and design are described in Section IV. Section V presents the experimental results along with the discussion. Finally, Section VI reports the concluding observations.

II. RELATED WORK

It is challenging for a single medical institution to obtain large-scale data [21]. When using a computer to analyze large data for complex or rare diseases, it is often impossible to find laws or obtain general conclusions. Therefore, the most effective solution is to integrate data from multiple medical institutions to form a larger multi-center data set [22]. However, the directly merged multi-center data set has many challenges. Owing to the differences in image acquisition equipment, even under the same acquisition protocol, it is not guaranteed to obtain consistent image characteristics; therefore, it is difficult to analyze how the image acquisition parameters affect the generated image [23]. In addition, inconsistency is also a serious problem when using multi-center data. Specifically, the data format is not uniform, the image size is inconsistent, and the contrast and brightness of different images vary significantly. At present, most

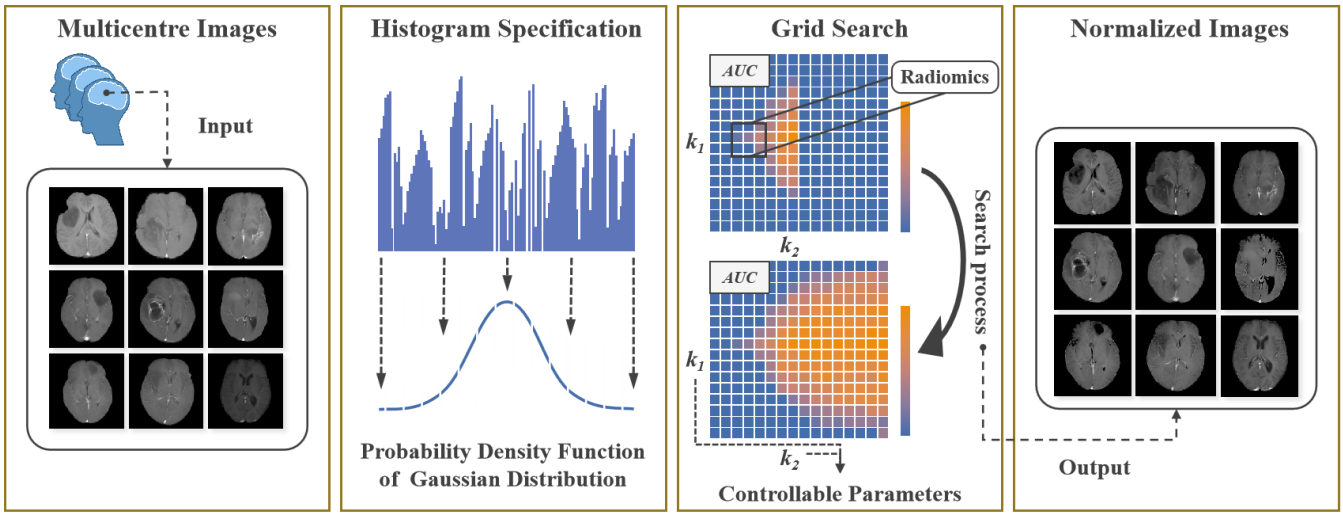


FIGURE 1. HS-BC workflow.

medical image analyses and research have focused on multi-center issues, emphasizing that the data used should have a uniform data format and image size, while the differences in image contrast and brightness are largely ignored. It has been proven that multi-center MRI research has more statistical significance than a single medical institution. However, the huge difference in contrast and brightness in many images is likely to reduce or completely offset the statistical advantages of multi-center data.

The adjustment of image brightness and contrast belongs to the problem of image enhancement. As a typical spatial domain calculation method and image enhancement technology, HS can effectively solve the problem of image contrast and brightness difference [17], [24], [25]. Histogram technology is a relatively old technology. Following continuous improvement, researchers have developed a series of methods for solving different problems. From the conventional to the dynamic HS, the features of the input image can be preserved intact [18], [26]. On this basis, the precise HS is further proposed. This method can produce almost strict ordering of the pixels of the input image and assign it to the desired grayscale [27]. In addition, the local histogram is normalized, and the global histogram is applied to the segmented local image blocks to perform image enhancement [28], [29]. However, the local HS is an operation that relies on sliding windows, and the additional generation of the checkerboard effect is still an urgent problem to be solved [30]. In recent years, deep learning techniques have also been used to solve image enhancement problems. For example, images under weak lights usually contain more noise and have extremely low contrast. The LL-RefineNet network can extract deep-level global features, thereby improving the original image contrast and optimizing training loss to reduce mixed noise [31]. The methods listed above focus on the field of image enhancement. Of course, the standardization of medical images has gradually been noticed. Being influenced by the operation of clinical technicians and ultrasonic testing equipment, the intensity of ultrasound images often changes, which affects the quantitative measurement and calculation of images. A histogram specification mechanism based on constructed CDF is used, which can retain the distribution of echo texture in ultrasound images, and does not damage the reflectivity of the organ [32]. Data standardization has additional requirements compared to image enhancement, as the standardization of medical images is different from the specific enhancement of natural images. The standardization operation emphasizes the batch processing of the entire data set. It is necessary to ensure that the image has a consistent gray space and approximate gray distribution. In addition, the normalized image needs to be competent for each data analysis task. The main novelty of HS–GS is that it is a method that can flexibly adjust the degree of data standardization. It uses a Gaussian function with adjustable parameters to indirectly control the degree of data standardization, and uses a dedicated GS to find the target solution distribution and thereafter the desired standardized image.

III. METHODOLOGY

HS–GS is a global enhancement technology which can perform uniform transmission and mapping of the contrast and brightness of all images in the image data set. This batch operation can ensure that the contrast and brightness of all images maintain an approximate dynamic range, thereby achieving image standardization. Generally, a complete magnetic resonance image is composed of 2D slices, and if the pixel-level operation is performed, the desired result cannot be obtained. Therefore, HS–GS performs batch processing in units of 2D slices (i.e., performs a mapping operation on slices of the same layer of magnetic resonance). Figure 1 describes the HS–GS workflow. HS–GS introduces a Gaussian probability density function with two controllable parameters, constructs a transfer function through the CDF obtained by calculation,

and guides the histogram to specify the map by the transfer function, thereby obtaining a standardized image. Since the parameters are discrete values that lack constraints, in theory, we may generate a large number of normalized images where most of the images do not meet our requirements. Therefore, we use the generated image to perform the task of radiomics classification, use the area under the receiver operating characteristic curve (AUC) as the evaluation index for classification, filter the combination of controllable parameters with high AUC scores, and obtain the optimal labeled image.

A. HISTOGRAM SPECIFICATION BASED ON CONTROLLABLE CONTRAST AND BRIGHTNESS

In this module, HS–GS is used to generate standardized medical images. Suppose the original image set is X', $\mathrm{X}' = \{X_1, X_2, \ldots, \mathrm{X}_n\}$. Each image in the dataset is used as the input. Let the input image be defined as $\mathrm{X} = \{x\,(i,j)\,|1 \le i \le M, 1 \le j \le N\}, \mathrm{X} \in \mathrm{X}'$, and be assumed to have a dynamic range $[x_{min}, x_{max}]$ (i.e. $x\,(i,j) \in [x_{min}, x_{max}]$). In addition, the generated image is defined as $Y = \{Y\,(i,j)\,|1 \le i \le M, 1 \le j \le N\}$. The dataset of generated images is $Y' = \{Y_1, Y_2, \ldots, Y_n\}, Y \in Y'$.

Let $\chi = \{x_1, x_2, \ldots, x_k\}$ be the ordering of k different gray levels of the input image X, and satisfy $x_1 < x_2 < \ldots < x_k, x_1 = x_{min}, x_k = x_{max}$.

HS–BC defines histograms to perform the prescribed operations:

$$H_x = \{h_x\,(m)\,|m = 1, \ldots, k\} \tag{1}$$

where $h_x\,(m) \in \Re^+$, and

$$h_x\,(m) = \frac{scr\,(m)}{\sum_{i=1}^{k} scr\,(x_k)} \tag{2}$$

$scr\,(m)$ represents the number of pixels with gray level m in the image.

As the image histograms are samples of the probability distribution function, the cumulative distribution can be written as

$$P_x = \{P_x\,(m)\,|m = 1, \ldots, k\} \tag{3}$$

$$P_x\,(m) = \sum_{i=1}^{m} h_x\,(i) \tag{4}$$

In addition, based on the image histogram, the average gray value of the image can be defined as

$$a = \sum_{i=1}^{k} x_i h_x\,(x_i) \tag{5}$$

Concurrently, the standard deviation of the gray values of the image can be defined as:

$$u = \left[\sum_{i=1}^{k} (x_i - a)^2\, h_x\,(x_i)\right]^{1/2} \tag{6}$$

Unlike the conventional histogram specification operation, we no longer use the representative image manually selected by the clinician to construct the transfer function. We use the probability density function of the Gaussian distribution based on the discrete value as the initial function. The CDF calculated by the initial function is used for the construction transfer function, and the mean value and standard deviation of the gray value of the image are introduced. Let $\gamma = \{y_1, y_2, \ldots, y_l\}$ be the order of l different gray levels of the output image Y, and satisfy $y_1 < y_2 < \ldots < y_l, y_1 = y_{min}$, $y_l = y_{max}$. The probability density function of the Gaussian distribution can be written as

$$H_t = h_t\,(m') = \frac{1}{\sqrt{2\pi}\,k_2 u} e^{-\frac{(m' - k_1 a)^2}{2(k_2 u)^2}} \quad |m' = 1, \ldots, l \tag{7}$$

Here, $k_1 a$ and $k_2 u$ represent the mean and standard deviation of this Gaussian distribution, respectively, which can in turn be determined by the mean a and standard deviation u of the gray values of the input histogram and the corresponding positive real numbers k_1 and k_2. Similarly, the CDF can be written as

$$P_t = \{P_t\,(m')\,|m' = 1, \ldots, l\} \tag{8}$$

$$P_t\,(m') = \sum_{i=1}^{m'} h_t\,(i) \tag{9}$$

After obtaining the CDFs of the input and the output images using (4) and (9), respectively, we construct the transfer function required for histogram mapping. The group mapping rule (GML) in the histogram specification maps the input grayscale x_m to the output gray scale $y_{m'}$. Let an integer function $I\,(x), x = 1, \ldots, l$, satisfy $0 \le I\,(1) \le \ldots \le I\,(l) \le x_{max}$; thus, $I\,(x)$ is determined by minimizing (10).

$$argmin\,|P_x\,[I\,(m') - P_t\,(m')]| \tag{10}$$

By (10), each gray level of the input image is mapped to the corresponding gray level in the output image. The corresponding algorithm is provided in Algorithm 1.

Algorithm 1 Algorithm 1 HS–BC

Input: input image X, parameters k_1, k_2,
Output: output image Y
1: Initialize expectation a, standard deviation u, image histogram h, Gaussian distribution H, enhanced image Y, pixel intensity m.
2: Compute the image histogram h by (2)
3: Compute expectation a by (5)
4: Compute standard deviation u by (6)
5: Compute the Gaussian distribution H by (7)
6: Compute the cumulative distribution function $\mathrm{P_x}\,(\mathrm{m})$ of histogram of input image by (4)
7: Compute the cumulative distribution function $\mathrm{P_t}\,(m')$ of the Gaussian distribution by (9)
8: Compute and determine the integer function $I\,(x)$ by (10)
9: Perform histogram mapping according to GML

TABLE 1. Image acquisition equipment and main parameter statistics.

Num	Manufacturer	Magnetic Field	Model	Slice Thickness	Slice Interval	Resolution	Voxel Size
1	SISMENS	3 T	Verio	5 mm	1.5 mm	320×224×20	0.6875/0.9213
2	SISMENS	3 T	Skyra	5 mm	1.75 mm	288×216×20	0.7986/1.0648
3	SISMENS	3 T	TrioTim	5 mm	1.5 mm	320×224×22	0.6875/0.8594
4	SISMENS	3 T	Prisma	5 mm	1.75 mm	288×216×20	0.7986/1.0648
5	SISMENS	1.5 T	Avanto	5 mm	1.5 mm	320×256×20	0.7188/0.7304
6	GE	3 T	Signa HDxt	5 mm	2 mm	320×224×18	0.7500/0.8036
7	GE	3 T	Discovery MR750	5 mm	1.5 mm	256×256×20	0.7986/1.0648
8	PHILIPS	3 T	Ingenia	6 mm	1 mm	320×208×18	0.7187/0.8750

B. GS STRATEGY

We have developed a dedicated GS script for screening optimal parameters and correspondingly generated standardized image datasets. The GS is based on the clinical application of standardized images. Specifically, we can define search tasks and select evaluation indicators according to clinical needs. In this paper, we define the classification of the imaging group as the search task. Since multicenter datasets typically have skewed data, AUC is used instead of Accuracy (ACC) as the evaluation index. If there are no constraints, the GS traverses all nodes in the search domain. In fact, the search task is an optimization problem. Unconstrained traversal only adds additional computing tasks and has extremely low efficiency. Therefore, we set the search threshold T, which terminates the process of AUC lower than T during the search process. The search process can be described as: (1) Divide the search grid with k_1 as the ordinate and k_2 as the abscissa, initialize k_1, k_2, and step size. First, perform a bidirectional search along the positive and negative directions of k_1. When the AUC is lower than the threshold T, the search process in this direction is terminated. (2) When the processes on k_1 are stopped, the search is performed along the positive and negative directions of k_2. Similarly, the search process stops when the AUC is below the threshold T.

IV. EXPERIMENTS

A. DATASET

The data set used in the experiment contains multi-center MRI images of gliomas. The data set consists of two parts. The first part is the public dataset BraTS2017 [33], [34], which consists of MRI images of 285 patients from 19 medical institutions, including 210 high-grade (III, IV) and 75 low-grade (I, II) glioma patients. The other part of the data set is from the Department of Magnetic Resonance of the First Affiliated Hospital of Zhengzhou University, called GI2019. GI2019 is a retrospective of 408 cases diagnosed by the pathology department, including 302 high-grade (III, IV) and 106 low-grade (I, II) glioma patients. It should be emphasized that GI2019 is collected using different models of different manufacturers, including five SIEMENS devices (one 1.5T device, four 3T devices), two GE devices (two 3T devices), and one PHILIPS piece of equipment (one 3T device). Table 1 lists the MRI equipment used for data acquisition and the main information of the equipment, including manufacturer, model, magnetic field strength, slice thickness, and slice interval. The two-part data set includes a total of 693 patients. BraTS2017 does not contain any privacy information, and GI2019 also deletes all privacy information during retrospective data collection. We used axial images of contrast-enhanced T1 weighted imaging (CET1) for processing and analysis.

B. PREPROCESSING

Prior to data processing, the necessary preprocessing steps were performed to eliminate differences in the data formats: (1) FSL (http://www.itksnap.org) was used to perform skull dissection on the GI2019 patient images. (2) The region of interest (ROI) in the GI2019 images was defined as the largest abnormal signal area in the tumor, which was manually delineated by two experienced radiologists, using the ITK-SNAP (http://www.itksnap.org) for image segmentation [35]. (3) The shortest distance interpolation resampling technique was used to perform proportional resampling on each data sample. In addition, each data sample had a uniform size of 240 × 240 × 155 voxels, with a uniform layer thickness of 1 mm.

We define the classification of radiomics as the task of GS. Here we list the radiomics process and the parameters of the key steps. Fig. 2 shows the process of radiomics built into GS. The standardized glioma images were used as inputs to predict the grade of glioma. The ratio of training set to test set is set to 4:1. In the feature extraction process, pyradiomics [36] was used to extract 558 features, including first-order statistical features, spatial geometric features, texture features, and wavelet features. The selected feature dimensionality reduction methods include principal component analysis (PCA), kernel principal component analysis (KPCA). In the feature selection method, the radiological features are selected by the minimum redundancy maximum

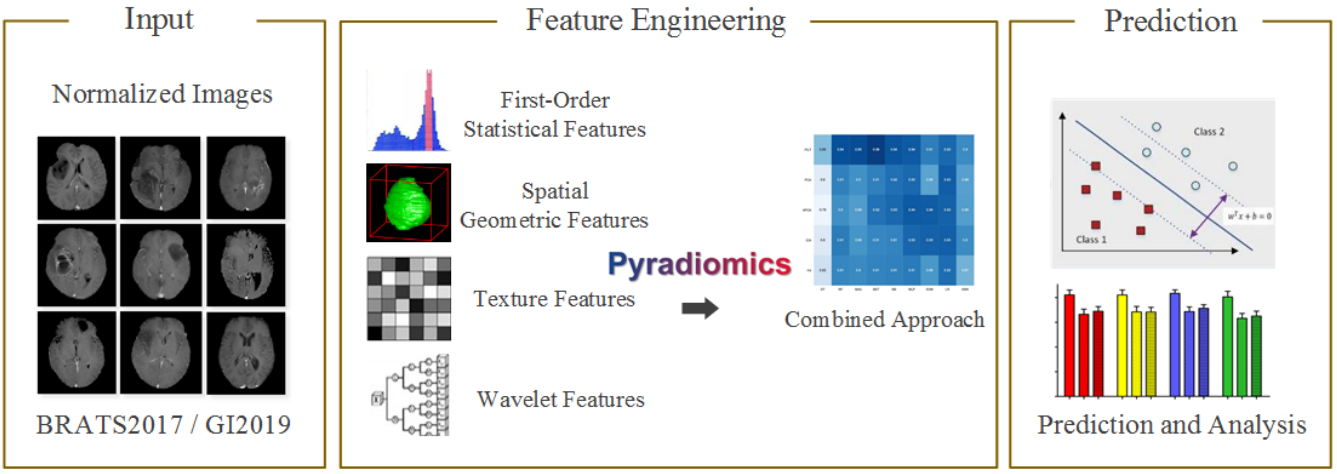


FIGURE 2. Flow chart of glioma grading by our method.

correlation (MRMR) algorithm, which ranks the features according to importance based on mutual information. Next, a recursive feature elimination (RFE) algorithm is used to select the features by recursively considering fewer and fewer feature sets. The following classifiers were trained, namely the decision tree (DT), random forest (RF), bagging (BAG), binary search tree (BST), naive Bayes (NB), multi-layer perception (MLP), support vector machine (SVM), logistic regression (LR), and k-nearest neighbor (KNN) algorithms. We combined different methods of feature dimensionality reduction, feature selection and classifier and obtained the optimal classification results.

C. EXPERIMENTAL STEPS

Multiple experiments were set up to test the effect of HS–BC on data normalization and the scalability of the proposed algorithm.

Experiment (1): After GI2019 and BraTS2017 were processed using HS–BC to generate the normalized images, these images were divided into the training and test sets for radiomics classification tasks. Datasets that were not processed using HS–BC were used for comparing the results.

Experiment (2): After GI2019 and BraTS2017 were processed using HS–BC to generate the normalized images, these images were used as training test sets for radiomics classification tasks. Datasets that were not processed using HS–BC were used for comparing the results.

Experiment (3): After the GI2019 and BraTS2017 datasets were processed using HS–BC to generate the normalized images, these images were all mixed and then divided into the training and test sets for radiomics classification tasks. The mixed dataset without HS–BC processing was used to compare the results.

Experiment (4): We add the comparison method of experiment (3). Radiologist selection and use the classic histogram specification [19], and we also compare the performance of deep learning for glioma grading [37], [38].

D. EVALUATION METRICS

In addition to the AUC, we also counted ACC, sensitivity (Sens), and specificity (Spec) to evaluate the experimental results. ACC refers to the proximity of the measured value to the actual value. In medical statistics, Sens and Spec are common objective evaluation indicators. Sens describes the ratio of the positive samples identified in all the positive samples, whereas Spec indicates the proportion of the negative samples identified in all the negative samples. The calculation formulas for the above three evaluation indicators are as follows:

$$ACC = TP + TN \big/ TP + FN + FP + TN \tag{11}$$

$$Sens = TP \big/ TP + FN \tag{12}$$

$$Spec = TN \big/ FP + TN \tag{13}$$

TP, FP, TN, and FN stand for true positive, false positive, true negative, and false negative, respectively. True positive indicates that high-grade glioma cases are correctly judged as high-grade gliomas. False positive indicates that high-grade glioma cases are wrongly judged as low-grade gliomas. True negative indicates that low-grade glioma cases are correctly judged as low-grade gliomas. False negatives indicates that low-grade glioma cases are wrongly judged as high-grade gliomas.

V. RESULTS AND DISCUSSION

In order to more intuitively compare the effect of the image before and after standardization, Fig. 3 shows the original image after only preprocessing and the normalized image after HS–GS processing. The difference in contrast and brightness of different original images can be clearly observed in the figure. When processed by the HS–GS algorithm, these differences are hardly observed at the image level, and the grayscale of different standardized images can be kept in a similar dynamic range.

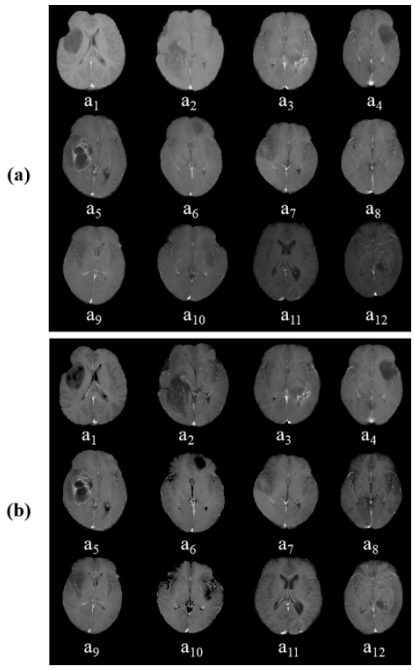


FIGURE 3. (a) Original image that has only been preprocessed. (b) Image generated after HS-BC processing. Here, a_1 and b_1 represent the same slice of the same sample, and 1, 2, . . . , 12 represent different samples.

Fig. 4 depicts the GS process of the image collections of GI2019 and BraTS2017 used in experiment (3). These images have been processed by the HS–GS algorithm, and the two data sets have all been mixed together (refer to TABLE 3-3.2). The search results show the AUC obtained by performing the radiomics classification task, and only display the AUC above the search threshold. Each rectangle represents a set of k_1 and k_2 values, and the color of the gradient indicates the level of AUC. We initialize T = 0.080, $k_1 = 0.5$, and $k_2 = 0.5$. The AUC is divided into columns, and the maximum AUC obtained by the current column is displayed in the corresponding rectangle. We can easily obtain the highest AUC and the corresponding k_1 and k_2, which means that we can also get the best standardized image. In addition, it can be seen in Fig. 4 that the k_1 and k_2 parameters are more sensitive to the GS with the task of imaging histology classification, and the smaller step size results in a significant change in AUC. At the same time, we also found that the maximum value of AUC (AUC_{max}) is also related to the step size. Setting different steps results in different maximum values and corresponding k_1 and k_2. Fig. 5 shows the distribution of AUC_{max} under several groups of steps. Furthermore, Table 2 lists the values of k_1 and k_2 corresponding to the step size and AUC_{max}. According to the results in Fig. 5 and Table 2, we found that when the set step size is small, we obtain a higher AUC_{max}, because k_1 and k_2 can obtain more accurate values at small step sizes. Conversely, when the set step size is larger, AUC_{max} decreases accordingly. In particular, when the step size is set to 1, AUC_{max} is less than the threshold and is terminated. Therefore, the best option is to set a smaller step size to search. We changed the search step size to obtain the fuzzy positioning of AUC_{max}. It can be predicted that AUC_{max} is located in a rectangle restricted by k_1, $k_2 \in [0.4,0.8]$.

TABLE 2. The K value corresponding to step sizes.

Step Size	AUC_{max}	k_1	k_2	Step Size	AUC_{max}	k_1	k_2
0.05	**0.956**	**0.60**	**0.65**	0.55	0.933	0.55	0.55
0.10	0.954	0.60	0.70	0.60	0.952	0.60	0.60
0.15	0.953	0.60	0.75	0.65	0.948	0.65	0.65
0.20	0.953	0.60	0.60	0.70	0.951	0.70	0.70
0.25	0.941	0.50	0.75	0.75	0.928	0.75	0.75
0.30	0.952	0.60	0.60	0.80	0.887	0.80	0.80
0.35	0.935	0.70	0.70	0.85	0.869	0.85	0.85
0.40	0.938	0.40	0.80	0.90	0.825	0.90	0.90
0.45	0.891	0.45	0.90	0.95	0.810	0.95	0.95
0.50	0.922	0.50	0.50	1.00	--	1.00	1.00

Table 3 lists the results of experiments (1) –(3) and the comparison of each group. "□ " indicates that HS–GS has not been implemented, and "√" indicates that the HS–GS method has been implemented. The results of experiments (1.1-4) show that the data set processed by the HS–GS method can obtain better classification prediction ability. For example, in experiment (1.4), the AUC of BraTS2017 increased from 0.847 to 0.979, while other indicators also improved significantly. In the two groups of experiments (1.1 and 1.3) and experiments (2.1 and 2.3), it can be confirmed that there are obvious differences between different data sets, and these differences produce different graded prediction results. When different data sets are combined, these differences can have a significant impact on the results of the graded prediction. Experiments (2.1–2.2) and (2.3–2.4) show that when the HS–GS method is used to normalize the data, the differences between different data sets are reduced and the classification effect is significantly improved. When BraTS2017 is used as the training set and GI2019 is used as the test set, the AUC may reach 0.948. In MRI, the imaging features of the tumor are mainly reflected in the change of the specific grayscale of the image, and the change of grayscale is mainly reflected in the image contrast and brightness. Generally, before data set A is normalized, the features extracted by feature engineering cannot fully describe data set B. When the HS–GS processed data set is used as the training set and in turn as the test set, there is no significant difference in classification effect. It is thus verified that the HS–GS method processing the data set has almost no variation. Experiments (3.1–2) show that simple preprocessing of the mixed datasets does not produce ideal results. Therefore, it is necessary to standardize the image processing after

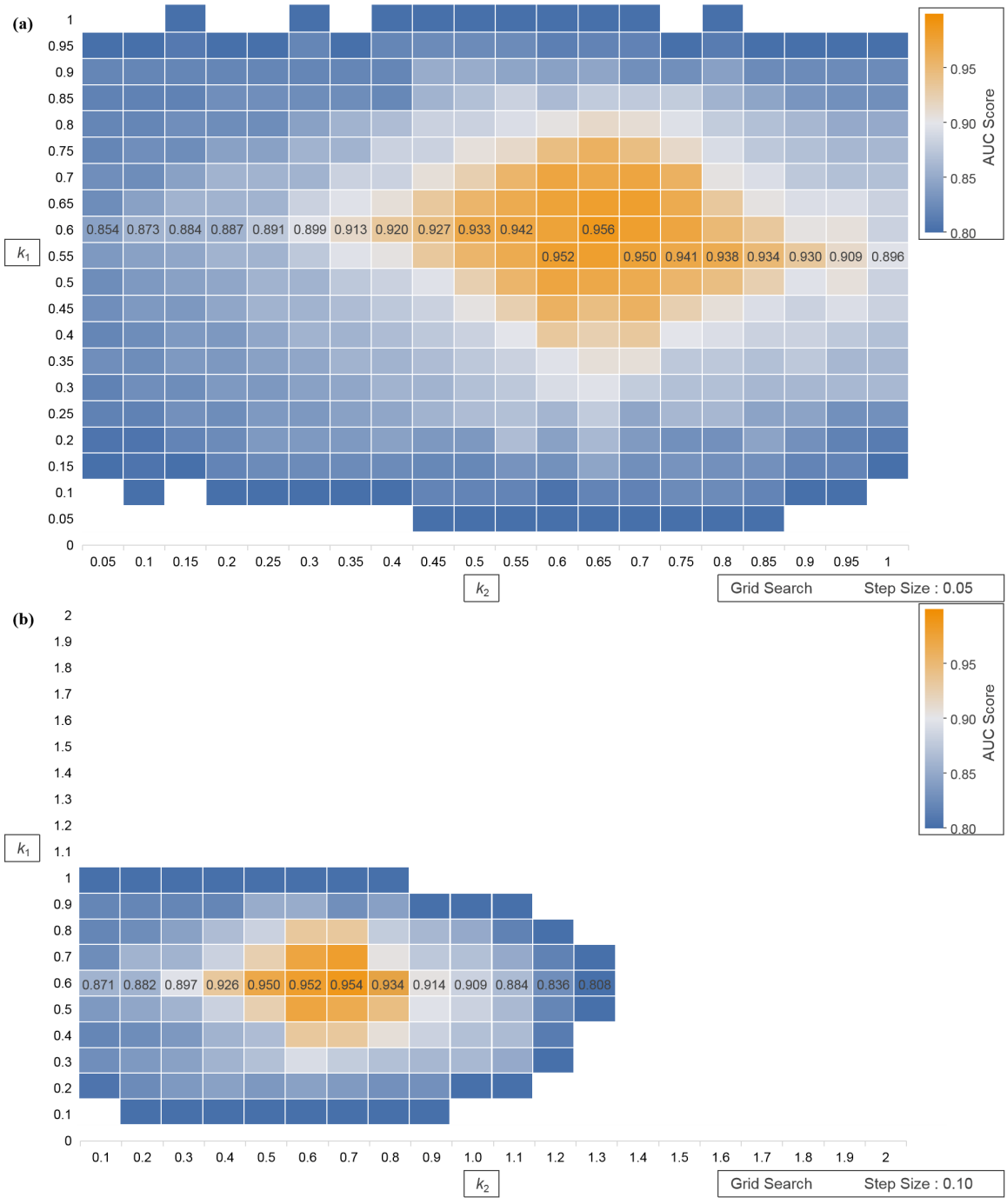


FIGURE 4. The grid search process under asynchronous length. The step sizes were set to (a) 0.05 and (b) 0.10.

mixing multi-center data. The AUC of the predicted classification of the mixed data after HS–GS processing is 0.956, which is 26.96% higher than the result of not performing standardization processing.

The experiments in Table 4 (4.1–2) compare the classification prediction effect of the traditional method and the HS–GS method proposed in this paper after the image is standardized. In the traditional method, the radiologist manually selects the representative images and normalizes them with histogram specification. Compared with traditional methods, the HS–GS method can clearly obtain better prediction results. Although the grading prediction results of "Radiologist Selection + HS" are still acceptable, this method depends largely on the subjective manual selection of the doctor, and it is impossible to judge whether the result of image standardization is the best. Experiments (4.3–4) list

TABLE 3. Comparisons between the results of experiments (1)–(3).

Exp	Num	Train Set	Test Set	Status	AUC (Std)	ACC (Std)	Sens (Std)	Spec (Std)
1.	1	GI2019	GI2019	□	0.821(0.060)	0.597(0.041)	0.832(0.085)	0.650(0.155)
	2	GI2019	GI2019	√	**0.964(0.032)**	**0.929(0.017)**	**0.966(0.026)**	**0.856(0.064)**
	3	BraTS2017	BraTS2017	□	0.847(0.075)	0.796(0.048)	0.801(0.060)	0.790(0.066)
	4	BraTS2017	BraTS2017	√	**0.979(0.012)**	**0.946(0.019)**	**0.969(0.020)**	**0.906(0.049)**
2.	1	GI2019	BraTS2017	□	0.664(0.093)	0.567(0.082)	0.516(0.125)	0.676(0.119)
	2	GI2019	BraTS2017	√	**0.910(0.051)**	**0.869(0.037)**	**0.897(0.056)**	**0.814(0.102)**
	3	BraTS2017	GI2019	□	0.698(0.040)	0.731(0.043)	0.805(0.075)	0.590(0.087)
	4	BraTS2017	GI2019	√	**0.948(0.024)**	**0.900(0.030)**	**0.933(0.043)**	**0.836(0.049)**
3.	1	GI2019+ BraTS2017		□	0.753(0.070)	0.742(0.048)	0.805(0.081)	0.623(0.118)
	2	GI2019+ BraTS2017		√	**0.956(0.036)**	**0.913(0.035)**	**0.931(0.033)**	**0.873(0.050)**

TABLE 4. Diagnostic performance of the proposed and compared methods.

Exp	Num	Train Set	Test Set	Method	AUC (Std)	ACC (Std)	Sens (Std)	Spec (Std)
4.	1	GI2019+	BraTS2017	Expert Selection + HS	0.888(0.034)	0.835(0.055)	0.882(0.064)	0.743(0.110)
	2	GI2019+	BraTS2017	HS-BC	**0.956(0.036)**	**0.913(0.035)**	**0.931(0.033)**	**0.873(0.050)**
	3	GI2019+	BraTS2017	VGG-19 DNN	0.921(0.043)	0.902(0.032)	0.943(0.036)	0.813(0.700)
	4	GI2019+	BraTS2017	GoogLeNet	0.939(0.028)	0.896(0.036)	0.930(0.043)	0.820(0.069)

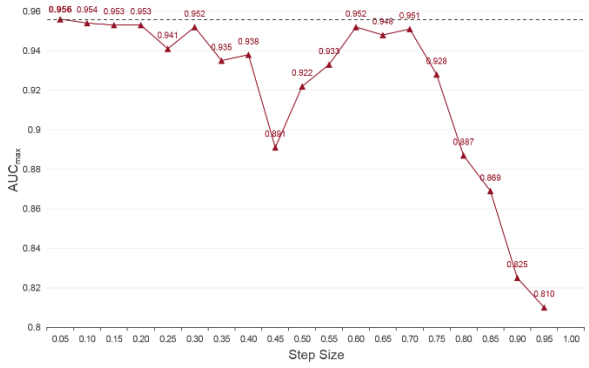


FIGURE 5. AUC_{max} statistics under different step sizes.

the results of deep learning models VGG-19 DNN [37] and GoogLeNet [38] for glioma grading. The data used is a mixed data set processed by the HS–GS method, and the necessary preprocessing of the data meets the input requirements of the deep learning model. In order to control the additional variables, we did not pre-train the network. The current results show that the standardized images of our proposed HS–GS method are also suitable for deep learning networks, and the results of prediction are also acceptable.

Similar to current image enhancement technologies, histogram specification improves the visual effect of images by optimizing image contrast [39]. GS is usually used for parameter selection, testing all possible choices to obtain the best parameters and results. In the field of medical image standardization, HS–GS provides a new way of thinking, transforming "optimizing contrast and brightness" into "controlling contrast and brightness." Our data standardization system can reduce inefficient interactive decision-making tasks for doctors.

It should be emphasized that, similar to the public dataset BraTS2017, the internal dataset GI2019 is also a multi-center dataset. We verified the performance of HS–GS on different data sets. The most important thing is whether the image standardization operation is still effective after mixing different data sets, because multi-center data is often mixed. The results of the experiment were approved by an experienced radiologist. In fact, HS–GS can be extended to other magnetic resonance-based medical image scenarios, such as brain metastases and pituitary tumors, and can be used for sequences other than the CET1 sequence.

VI. CONCLUSION

A method of image data standardization is proposed, which is specially used for multicenter medical maps based on MRI, called HS–GS. The HS–GS algorithm uses a Gaussian probability density function with adjustable parameters as a representative image, and the CDF obtained by calculation is used to construct a histogram specification transfer function

to generate a standardized image. In order to find the optimal standardized image, the classification of radiomics is designated as the GS task, the optimal parameter combination is found through a dedicated GS, and the optimal standardized image set is also obtained. The HS–GS algorithm abandons the method of manually selecting representative images by clinicians, avoids the uncertainty caused by subjective selection, and reduces the human–computer interaction in the main link. We tested and verified the performance of the HS–GS algorithm in solving clinical practical problems (prediction of glioma grading), and compared traditional methods with the latest methods. Overall, it performs well in terms of algorithm robustness and image quality. This novel image standardization technique has proven to be a promising solution for medical expert systems.

As a future extension of this work, we will consider adding more GS tasks and evaluation indicators for the abundant search tasks. In addition, standardization of other data types, such as CT images, will also be considered. Generally, as CT images have a larger image size, for their standardization the HS–GS algorithm should pay more attention to reducing the time cost.

COMPETING INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] J. D. Rudie, A. M. Rauschecker, R. N. Bryan, C. Davatzikos, and S. Mohan, "Emerging applications of artificial intelligence in neuro-oncology," *Radiology*, vol. 290, no. 3, pp. 607–618, Mar. 2019.

[2] L. Rundo, A. Tangherloni, M. S. Nobile, C. Militello, D. Besozzi, G. Mauri, and P. Cazzaniga, "MedGA: A novel evolutionary method for image enhancement in medical imaging systems," *Expert Syst. Appl.*, vol. 119, pp. 387–399, Apr. 2019.

[3] C. Parmar, J. D. Barry, A. Hosny, J. Quackenbush, and H. J. W. L. Aerts, "Data analysis strategies in medical imaging," *Clin. Cancer Res.*, vol. 24, no. 15, pp. 3492–3499, Aug. 2018.

[4] B. Jie, M. Liu, and D. Shen, "Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease," *Med. Image Anal.*, vol. 47, pp. 81–94, Jul. 2018.

[5] J. Zhang, Y. Xia, Y. Xie, M. Fulham, and D. D. Feng, "Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 5, pp. 1521–1530, Sep. 2018.

[6] C.-S. Lin, Y.-C. Huang, S.-H. Chen, Y.-L. Hsu, and Y.-C. Lin, "The application of deep learning and image processing technology in laser positioning," *Appl. Sci.*, vol. 8, no. 9, p. 1542, Sep. 2018.

[7] J. Liu, F. Chen, C. Pan, M. Zhu, X. Zhang, L. Zhang, and H. Liao, "A cascaded deep convolutional neural network for joint segmentation and genotype prediction of brainstem gliomas," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 1943–1952, Sep. 2018.

[8] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, Feb. 2016.

[9] T. Upadhaya, M. Vallieres, A. Chatterjee, F. Lucia, P. A. Bonaffini, I. Masson, A. Mervoyer, C. Reinhold, U. Schick, J. Seuntjens, C. C. L. Rest, D. Visvikis, and M. Hatt, "Comparison of radiomics models built through machine learning in a multicentric context with independent testing: Identical data, similar algorithms, different methodologies," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 192–200, Mar. 2019.

[10] Y. Q. Huang, "Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer," *J. Clin. Oncol.*, vol. 34, no. 18, pp. 2157–2164, Jun. 20 2016.

[11] S. P. Sourbron and D. L. Buckley, "Classic models for dynamic contrast-enhanced MRI," *NMR Biomed.*, vol. 26, no. 8, pp. 1004–1027, Aug. 2013.

[12] A. M. Rauschecker, "Artificial intelligence system approaching neuroradiologist-level differential diagnosis accuracy at brain MRI," *Radiology*, vol. 295, no. 3, Apr. 2020, Art. no. 190283.

[13] L. K. Young, S. J. Gandy, L. Priba, S. Z. Matthew, and J. Graeme Houston, "The impact of different magnetic resonance imaging equipment and scanning parameters on signal intensity ratio measurements in phantoms and healthy volunteers: Implications for interpreting gadolinium signal changes within the brain," *Investigative Radiol.*, vol. 54, no. 3, pp. 169–176, Mar. 2019.

[14] U. Bağci, J. K. Udupa, and L. Bai, "The role of intensity standardization in medical image registration," *Pattern Recognit. Lett.*, vol. 31, no. 4, pp. 315–323, Mar. 2010.

[15] M. Drozdzal, G. Chartrand, E. Vorontsov, M. Shakeri, L. Di Jorio, A. Tang, A. Romero, Y. Bengio, C. Pal, and S. Kadoury, "Learning normalized inputs for iterative estimation in medical image segmentation," *Med. Image Anal.*, vol. 44, pp. 1–13, Feb. 2018.

[16] H. L. Zhang, Y. Zhao, C. Pang, and J. He, "Splitting large medical data sets based on normal distribution in cloud environment," *IEEE Trans. Cloud Comput.*, vol. 8, no. 2, pp. 518–531, Apr. 2020.

[17] X. Sun, Q. Xu, and L. Zhu, "An effective Gaussian fitting approach for image contrast enhancement," *IEEE Access*, vol. 7, pp. 31946–31958, 2019.

[18] R. C. Conzalez and R. E. Woods, *Digial Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[19] X. Chen, "Automatic histogram specification for glioma grading using multicenter data," *J. Healthcare Eng.*, vol. 2019, Dec. 2019, Art. no. 9414937.

[20] A. Draa and A. Bouaziz, "An artificial bee colony algorithm for image contrast enhancement," *Swarm Evol. Comput.*, vol. 16, pp. 69–84, Jun. 2014.

[21] X.-H. Jiang, S.-B. Wang, Q. Tian, C. Zhong, G.-L. Zhang, Y.-J. Li, P. Lin, Y. You, R. Guo, Y.-H. Cui, and Y.-Q. Xing, "Right-to-left shunt and subclinical ischemic brain lesions in chinese migraineurs: A multicentre MRI study," *BMC Neurol.*, vol. 18, no. 1, p. 18, Feb. 2018.

[22] A. E. Fetit, J. Novak, D. Rodriguez, D. P. Auer, C. A. Clark, R. G. Grundy, A. C. Peet, and T. N. Arvanitis, "Radiomics in paediatric neuro-oncology: A multicentre study on MRI texture analysis," *NMR Biomed.*, vol. 31, no. 1, p. e3781, Jan. 2018.

[23] R. Berenguer, M. D. R. Pastor-Juan, J. Canales-Vázquez, M. Castro-García, M. V. Villas, F. Mansilla Legorburo, and S. Sabater, "Radiomics of CT features may be nonreproducible and redundant: Influence of CT acquisition parameters," *Radiology*, vol. 288, no. 2, pp. 407–415, Aug. 2018.

[24] T. Celik, "Two-dimensional histogram equalization and contrast enhancement," *Pattern Recognit.*, vol. 45, no. 10, pp. 3810–3824, Oct. 2012.

[25] D. Coltuc, P. Bolon, and J.-M. Chassery, "Exact histogram specification," *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1143–1152, May 2006.

[26] C.-C. Sun, S.-J. Ruan, M.-C. Shie, and T.-W. Pai, "Dynamic contrast enhancement based on histogram specification," *IEEE Trans. Consum. Electron.*, vol. 51, no. 4, pp. 1300–1305, Nov. 2005.

[27] M. Nikolova and G. Steidl, "Fast ordering algorithm for exact histogram specification," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5274–5283, Dec. 2014.

[28] C. Wang, J. Peng, and Z. Ye, "Flattest histogram specification with accurate brightness preservation," *IET Image Process.*, vol. 2, no. 5, pp. 249–262, Oct. 2008.

[29] C. Liu, X. Sui, X. Kuang, Y. Liu, G. Gu, and Q. Chen, "Optimized contrast enhancement for infrared images based on global and local histogram specification," *Remote Sens.*, vol. 11, no. 7, p. 849, Apr. 2019.

[30] B. Xiao, H. Tang, Y. Jiang, W. Li, and G. Wang, "Brightness and contrast controllable image enhancement based on histogram specification," *Neurocomputing*, vol. 275, pp. 2798–2809, Jan. 2018.

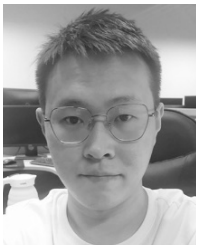
[31] L. Jiang, Y. Jing, S. Hu, B. Ge, and W. Xiao, "Deep refinement network for natural low-light image enhancement in symmetric pathways," *Symmetry*, vol. 10, no. 10, p. 491, Oct. 2018.

[32] R. Roy, S. Ghosh, and A. Ghosh, "Clinical ultrasound image standardization using histogram specification," *Comput. Biol. Med.*, vol. 120, May 2020, Art. no. 103746.

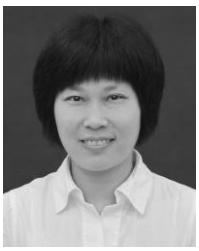
[33] B. H. Menze, A. Jakab, and S. Bauer, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.

[34] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, no. 1, Sep. 2017, Art. no. 170117.

[35] L. Li, K. Wang, X. Ma, Z. Liu, S. Wang, J. Du, K. Tian, X. Zhou, W. Wei, K. Sun, Y. Lin, Z. Wu, and J. Tian, "Radiomic analysis of multiparametric magnetic resonance imaging for differentiating skull base chordoma and chondrosarcoma," *Eur. J. Radiol.*, vol. 118, pp. 81–87, Sep. 2019.
[36] J. J. M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. H. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, and H. J. W. L. Aerts, "Computational radiomics system to decode the radiographic phenotype," *Cancer Res.*, vol. 77, no. 21, pp. e104–e107, Nov. 2017.
[37] A. Muneer K. V., V. R. Rajendran, and P. J. K., "Glioma tumor grade identification using artificial intelligent techniques," *J. Med. Syst.*, vol. 43, no. 5, p. 113, Mar. 2019.
[38] Y. Yang, L.-F. Yan, X. Zhang, Y. Han, H.-Y. Nan, Y.-C. Hu, B. Hu, S.-L. Yan, J. Zhang, D.-L. Cheng, X.-W. Ge, G.-B. Cui, D. Zhao, and W. Wang, "Glioma grading on conventional MR images: A deep learning study with transfer learning," *Frontiers Neurosci.*, vol. 12, p. 804, Nov. 2018.
[39] J.-J. Fernandez, T. E. Torres, E. Martin-Solana, G. F. Goya, and M.-R. Fernandez-Fernandez, "PolishEM: Image enhancement in FIB–SEM," *Bioinformatics*, vol. 36, no. 12, pp. 3947–3948, Mar. 2020.

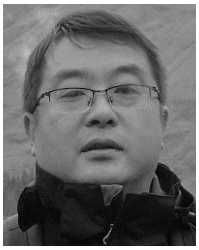


GUOHUA ZHAO received the bachelor's and master's degrees in software engineering from Zhengzhou University. He is currently pursuing the Ph.D. degree with the Department of Information Engineering, Zhengzhou University, and The First Affiliated Hospital of Zhengzhou University. His research interests include medical imaging processing and machine learning.



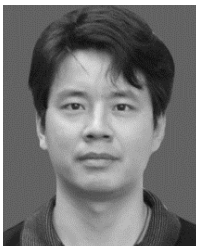
JIE BAI received the bachelor's degree in medical imaging and nuclear medicine and the master's degree from Zhengzhou University, in 2001 and 2008, respectively. She is currently working with the Department of Radiology, The First Affiliated Hospital of Zhengzhou University. Her major is medical imaging, which is mainly good at magnetic resonance diagnosis of central nervous system diseases.

PEI PEI WANG received the bachelor's and master's degrees in biomedical engineering from Southern Medical University, in 2015 and 2018, respectively. He is currently working as an Assistant Engineer with the Department of Radiology, The First Affiliated Hospital of Zhengzhou University. His research interests include medical image processing and pattern recognition.



GUAN YANG received the B.S. degree in probability theory and mathematical statistics from Northwest University, Xi'an, China, in 1997, and the M.S. degree in applied mathematics and the Ph.D. degree in mathematics from Sun Yat-sen University, Guangzhou, China, in 2005 and 2011, respectively. He is currently an Associate Professor with the School of Computer Science, Zhongyuan University of Technology. His research interests include computer visioning, machine learning, and medical image processing.

LEI SHI received the M.Sc. degree from Nanjing University, China, in 1992, and the Ph.D. degree from the Beijing Institute of Technology, China, in 2006. He is a Professor with the School of Software, Zhengzhou University. His research interests include cloud computing, distributed systems, big data, and artificial intelligence.



YONGCAI TAO received the Ph.D. degree in computer engineering from the Huazhong University of Science and Technology, China, in 2009. He is currently a Researcher with the High Performance Computing Laboratory, Zhengzhou University, China. His research interests include cluster and grid computing, fault-tolerance, Web services, and cloud computing.

YUSONG LIN (Member, IEEE) received the B.Sc. degree from Nanjing University, in 1993, the M.Sc. degree from Zhengzhou University, in 1998, and the Ph.D. degree from Information Engineering University, in 2005. He is currently a Professor with Zhengzhou University. He is also an Associate Director of the Henan Provincial Education and Research Network Center. His research interests include medical image processing, artificial intelligence, and Internet protocols.

JINGLIANG CHENG is the Director of the Department of Magnetic Resonance, The First affiliated Hospital of Zhengzhou University, the Director of the Medical Imaging Center, and a Doctoral Tutor. He is also the Director of the Key Laboratory of Magnetic Resonance Imaging and Molecular Imaging of Henan Province. He has published more than 300 research articles and edited 30 monographs on medical imaging. He has been involved in medicine, teaching, and research work in radiodiagnosis.