



# Article Optimized Resolution-Oriented Many-to-One Intensity Standardization Method for Magnetic Resonance Images

# Yuan Gao <sup>1</sup>, Yuanyuan Wang <sup>1,2,\*</sup> and Jinhua Yu <sup>1,2</sup>

- <sup>1</sup> Department of Electronic Engineering, Fudan University, Shanghai 200433, China; 15110720016@fudan.edu.cn (Y.G.); jhyu@fudan.edu.cn (J.Y.)
- <sup>2</sup> Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention (MICCAI) of Shanghai, Shanghai 200433, China
- \* Correspondence: yywang@fudan.edu.cn

Received: 30 October 2019; Accepted: 12 December 2019; Published: 16 December 2019



Abstract: With the development of big data, Radiomics and deep-learning methods based on magnetic resonance (MR) images, it is necessary to conduct large databases containing MR images from multiple centers. Having huge intensity distribution differences among images reduced or even eliminated, robust computer-aided diagnosis models could be established. Therefore, an optimized intensity standardization model is proposed. The network structure, loss function, and data input strategy were optimized to better avoid the image resolution loss during transformation. The experimental dataset was obtained from five MR scanners located in four hospitals and was divided into nine groups based on the imaging parameters, during which 9152 MR images from 499 participants were collected. Experiments show the superiority of the proposed method to the previously proposed unified model in resolution metrics including the peak signal-to-noise ratio, structural similarity, visual information fidelity, universal quality index, and image fidelity criterion. Another experiment further shows the advantage of the proposed method in increasing the effectiveness of following computer-aided diagnosis models by better preservation of MR image details. Moreover, the advantage over conventional standardization methods are also shown. Thus, MR images from different centers can be standardized using the proposed method, which will facilitate numerous data-driven medical imaging studies.

**Keywords:** cycle generative adversarial network; intensity standardization; magnetic resonance imaging; resolution oriented; advanced weak-pair strategy

# 1. Introduction

As the most commonly used imaging method in the diagnosis of brain diseases, magnetic resonance imaging (MRI) is one of the research hotspots of computer-aided brain diagnosis in recent years [1]. Many magnetic resonance (MR) image based studies, such as computer aided diagnosis [2–4], differential diagnosis [5], treatment options selection [6], and prognosis estimation [7], have made great progress, which also put forward higher requirements for not only the quantity, but also the quality of the image data.

To conduct a larger unified training set contains MR images obtained from different MR scanners, the scale and intensity distribution difference of such images should be suppressed. There are already some mature methods for solving scale inconsistency, including image slice resampling [8] and scale-adaptive feature extraction [9]. To deal with different intensity distribution, preventing the increasingly complex diagnostic models from being over-fitted and unstable as well as making the

model more generalizable, there is a great need for methods to eliminate the inter-group difference by intensity standardization, to ensure that the intensities of the same tissue type are the same in different images. In this case, many standardized data from multiple sources can be used for model training, and the models can be used in a wide range. To this end, many methods have been proposed, which can be divided into three main categories: the global histogram-matching methods, the joint histogram registration methods, and deep-learning-based methods.

The simplest method category among the three main categories is the histogram matching methods. These methods generate a series of corresponding intensity landmarks based on the features of the target images and the reference images, and then conduct a piecewise linear function based on these intensity landmarks, which is treated as the transformation equation. Nyúl et al. applied both overall and "foreground" percentile markers as intensity landmarks [10]. Collewet et al. utilized and compared three landmark series, which contain the maximum intensity of MR images, the mean intensity of MR images, and three sigma intensity points [11]. Madabhushi et al. employed the intensity features extracted from the largest fuzzily connected homogenous region to avoid the influence of the diseased tissue [12]. Sun et al. brought forward the maximum and the minimum decile along with the mean intensity [13]. Nunzio et al. firstly formed linear transforming functions for different brain tissues, and then built the final standardization function by joining the different transforming functions with spline smoothness [14]. However, the robustness of such methods is not sufficient. If there is a difference in tissue intensity relationship between the target images and the reference images, for example, if tissues with the same intensity in the target images differ in intensity in the reference ones, the piecewise linear function may not correctly describe the transforming function due to the absence of the multi-modality information as well as neighborhood area information. Meanwhile, such global transformation method is also unable to deal with spatial intensity nonuniformities due to the radio-frequency field inhomogeneities.

The joint histogram registration methods, on the other hand, use intensities of the same spatial location in multiple modalities among one scan as coordinates of a intensity vector and form a multi-dimensional point cloud (i.e., the joint histogram) based on the intensity vectors with different spatial location. Then, the elastic registration function of two point-clouds can be used as the intensity vector mapping function between the corresponding scans. Jager et al. directly applied the multimodality histograms to form the joint probability density functions (PDFs) [15,16]. Dzyubachyk et al. applied Jager's method into body MRI scan and further came up with regional equalization of the transforming function in different body parts [17]. Robitaille et al. firstly generated the joint histogram, and then extracted a set of characteristic points for histogram matching method [18]. Our group formed the target and reference point cloud with weighted sub-region intensity distribution instead of the joint PDFs [19,20]. These methods can better describe the transforming function but require that the target dataset and the reference one used for training the elastic registration function have the same multi-modality MR image data obtained from the same patient/volunteer and that these data should be accurately registered, which is quite difficult to achieve.

To construct a generalized MR image standardization model that could transform MR images acquired from different MR scanners and/or using different parameters, while multi-modality MR images are not needed for training and spatial intensity nonuniformities might be eliminated with the fusion of regional and global information, we propose a universal intensity standardization method based on cycle generative adversarial network [21]. This method applies a many-to-one framework with jump connections in the generators and weak-pair data augmentation strategy. This method is able to produce proper intensity standardization results. However, the original CycleGAN is mainly used in image segmentation as well as low-resolution image generation, while the cycle consistency is an indirect structural similarity indicator which is easily affected by other factors and unable to remove image blurring introduced by convolutional neural networks. Meanwhile, the weak-pair method used in the original method does not guarantee a consistent balance of randomness and structural

similarity when the axial density of the dataset varies. These all result in a reduced resolution of the standardized images.

To achieve better preservation of MR image spatial resolution and generate precision standardization results, in this paper, we propose a resolution-oriented MR image intensity standardization method to generate MR image datasets with strong intensity and spatial uniformity. This method is based on the cycle generative adversarial network (CycleGAN) [22] under the extended many-to-one framework. Each generator in this framework, in particular, applies cascading residual blocks to enhance super resolution performance. The generators also apply normalized mutual information (NMI) as a part of the loss function, which is able to reduce and even eliminate image blurring by directly measuring structural errors. Moreover, an advanced weak-pair data augmentation method is applied to adapt to the varying MR image axial density.

The paper is organized as follows. Section 2 expatiates on the proposed method. The dataset and data preprocessing steps are shown in Section 3. Section 4 describes the experiments and the results. The discussion is presented in Section 5 and the conclusion is made in Section 6.

#### 2. Methods

#### 2.1. Network Architecture

Figure 1 illustrates the network architecture under the extended many-to-one framework. The aim of the intensity standardization generator  $G_{forward}$  is to transform every T2-FLAIR MR image slice  $x_{n,m}$ , where  $n \in \{1, 2, \dots, N\}$  represents the group which the slice is from and  $m \in \{1, 2, \dots, M_n\}$  is the slice number in the  $n_{th}$  group, ensuring  $G_{forward}(x_{n,m})$  "seems to be" from the  $N_{th}$  group, which is treated as the reference, while keeping its own structure/tissue properties. N reverse transformations  $G_{backward n}$  are established, ensuring that  $G_{backward n}\left(G_{forward}(x_{n,m})\right) \approx x_{n,m}$ , which could indicate the preservation of image-specific structure/tissue properties. A forward discriminator  $D_{forward}$  and N backward discriminators  $D_{backward n}$  are also applied to form the GAN structures. Each G is optimized in turn with the corresponding D during the training process to form the adversarial scheme so that the intensity distribution of the generated images is gradually close to that of the corresponding reference ones.



**Figure 1.** The network architecture of the proposed method. Different groups of MR images acquired from different machines or using different acquisition parameters are standardized with a same forward generator  $G_{forward}$  while the standardized images are transformed back using  $G_{backward n}$  to evaluate the information consistency.

Each generator shown in Figure 2, in particular, consists of a convolutional concentration path, a low-dimensional feature extraction residual convolution path, a convolutional dispersion path, and jump connection paths. The convolutional concentration path contains a preprocessing convolution layer and two strided convolution layers. The dispersion path contains two strided transposed convolution layers and a synthetical transposed convolution layer. Two jump connections feed the feature maps before and during the concentration process into the corresponding position in the dispersion path to provide more details of input images [23].



**Figure 2.** The structure of the proposed generative networks. The orange arrow indicates the cascading block in the dotted box at the bottom right of the figure. The plum arrow within the cascading block indicates the residual block in the dotted box at the bottom left of the figure. All submodules in the same parent module share a same set of parameters.

The low-dimensional feature extraction residual convolution path employs cascading residual network (CARN) blocks in order to provide more "fast propagate" jump connections and acquire multi-level representations by integrating features between layers [24]. This may significantly increase the information richness of the generated images, thus be conducive to image resolution. Such residual

convolution path consists of three cascading blocks. The  $1 \times 1$  convolutional layers are instantiated after each cascading block to synthesize feature maps with fixed output channel number from all previous cascading blocks along with the original result of the concentration path. The cascading block consists of three residual blocks. Inside the cascading block,  $1 \times 1$  convolutional layers are also instantiated after each residual block to build a fixed-channel intermediate feature map. A residual block contains two  $3 \times 3$  convolutional layers, and the filter number is fixed to 256.

The discriminators are classical PatchGAN with three successive  $4 \times 4$  strided convolutions and two synthetical convolutions. Each convolution process applies leaky rectified linear unit (LeakyReLU) as the activation function. The discriminative feature map is  $30 \times 30$  with a single layer, which is then used to create the final discriminative result.

# 2.2. Adversarial Loss and Cycle Consistency Loss

The adversarial loss is used to judge the distribution similarity between the transformed images and the reference images. Applying least squares generative adversarial networks (LSGAN), the forward adversarial losses is defined as

$$\mathcal{L}_{LSGAN}\left(G_{forward}, D_{forward}, X_n, X_N\right) = \mathbb{E}_{\{x_N\}}\left[\left(D_{forward}\left(x_N\right) - 1\right)^2\right] + \mathbb{E}_{\{x_n\}}\left[D_{forward}\left(G_{forward}\left(x_n\right)\right)^2\right].$$
(1)

In GANs, the objective of generator optimization is to "fool" the corresponding discriminator by reducing the discriminator result gap between generated images and reference images while that of discriminator optimization is to better distinguish the two kinds of images. Therefore, the objective of the forward adversarial loss is min  $\max_{G_{forward}} \mathcal{L}_{LSGAN} (G_{forward}, D_{forward}, X_n, X_N)$ , and that of the backward one is min  $\max_{G_{backward n}} \mathcal{L}_{LSGAN} (G_{backward n}, D_{backward n}, X_N, X_n)$ . The forward generator and the backward one is combined to form an end-to-end structure

The forward generator and the backward one is combined to form an end-to-end structure similar to the self-encoder. Therefore, the  $G_{forward}(x_{n,m})$  is treated as the "code". To ensure the preservation of image-specific feature in the "code", cycle consistency loss is applied to compare  $x_{n,m}$  with  $G_{backward n}(G_{forward}(x_{n,m}))$  and compare  $x_{N,m}$  with  $G_{forward}(G_{backward n}(x_{N,m}))$ . Utilizing the  $L_1$  norm, the *n*th cycle consistency loss function is

$$\mathcal{L}_{cycle}\left(G_{forward}, G_{backward n}\right) = \mathbb{E}_{\{x_n\}}\left[\left\|G_{backward n}\left(G_{forward}\left(x_n\right)\right) - x_n\right\|_1\right] + \mathbb{E}_{\{x_N\}}\left[\left\|G_{forward}\left(G_{backward n}\left(x_n\right)\right) - x_n\right\|_1\right].$$
(2)

#### 2.3. Normalized Mutual Information Loss

To ensure the consistency of the morphological features and tissue characteristics between two MR image slices, information consistency losses are applied. In particular, we apply the NMI loss as part of the entire loss function [25] so as to ensure that the transformed image is consistent with the target image in structure, that is, only the intensities corresponding to the tissues are about to change. One of the NMI losses is expressed as

$$\mathcal{L}_{NMI}\left(G_{forward}\right) = \mathbb{E}_{\{x_n, x_N\}}\left[\frac{H\left(G_{forward}\left(x_n\right)\right) + H\left(x_N\right)}{H\left(G_{forward}\left(x_n\right), x_N\right)}\right]$$
(3)

where marginal entropies  $H(x) = -\sum Hist(x) \log [Hist(x)]$ , and the joint entropy  $H(x, y) = -\sum JointHist(x, y) \log [JointHist(x, y)]$ . Therefore, the NMI loss and the cycle consistency loss are applied together to make both direct and indirect consistency measurement.

#### 2.4. Entire Loss Function

The entire loss function is defined as

$$\mathcal{L}\left(G_{forward}, G_{backward n}, D_{forward}, D_{backward n}\right)$$

$$= \mathcal{L}_{LSGAN}\left(G_{forward}, D_{forward}, X_n, X_N\right) + \mathcal{L}_{LSGAN}\left(G_{backward n}, D_{backward n}, X_N, X_n\right)$$
(4)
$$+ \mathcal{L}_{cycle}\left(G_{forward}, G_{backward n}\right) + \mathcal{L}_{NMI}\left(G_{forward}\right) + \mathcal{L}_{NMI}\left(G_{backward n}\right).$$

Therefore, the objective of the optimization is to find the optimal  $G_{forward}$  and the N optimal  $G_{backward n}$  by

$$G_{forward}^{*}, G_{backward n}^{*} = \arg \min_{G_{forward}, G_{backward n}} \max_{D_{forward}, D_{backward n}} \mathcal{L}\left(G_{forward}, G_{backward n}, D_{forward}, D_{backward n}\right).$$
<sup>(5)</sup>

Finally, the  $G_{forward}$  is used to standardize MR image slices by transforming any MR image slices into the reference domain.

#### 3. Data and Preprocessing

#### 3.1. Dataset

This study was approved by the Ethics Committee of all four participating hospitals while informed consent was obtained from every patient and volunteer.

In total, 8192 MR image slices were obtained from 489 patients and divided into nine groups according to the different acquisition parameter used. These image slices were used to train the proposed model. The details of the patient images are shown in Table 1. Image Groups 1–5 were from Department of Neurosurgery, Huashan Hospital, Fudan University, the first three groups of which were acquired with a Siemens Magnetom Verio 3.0T MRI scanner and the other two groups were acquired with a GE Discovery MR750 3.0T MRI scanner. Image Groups 6 and 7 were from Department of Neurosurgery/Neuro-oncology, Sun Yat-sen University Cancer Center using a Siemens Magnetom Trio Tim 3.0T MRI scanner. Image Group 8 was obtained from Department of Neurosurgery, Huadong Hospital, Fudan University with a Siemens Magnetom Verio 3.0T MRI scanner. Image Group 9 was obtained from Department of Neurosurgery, Huadong Hospital, Fudan University with a Siemens Magnetom Verio 3.0T MRI scanner. Image Group 9 was obtained from Department of Neurosurgery, Huadong Hospital, Fudan University with a Siemens Magnetom Verio 3.0T MRI scanner. Image Group 9 was obtained from Department of Neurosurgery, Shandong Provincial Hospital using a Siemens Magnetom Skyra 3.0T MRI scanner.

Meanwhile, two groups of T2-FLAIR MR images were acquired successively with the Siemens Magnetom Verio 3.0T MRI scanner (Group 3) and GE Discovery MR750 3.0T MRI scanner (Group 6) from ten volunteers. MR images of another two modalities were acquired at the same time to meet the requirements of the joint histogram registration method (Groups 1 and 2 with the Siemens scanner and Groups 4 and 5 with the GE scanner). Therefore, 60 image sets from the ten volunteers were treated as the paired gold standard to perform all paired comparisons. The details of the volunteer images are shown in Table 2.

Moreover, to prove that the proposed method as a pre-processing method not only improves the visual perception of the resolution and related indicators, but also is conducive to the performance of various resolution-sensitive post-processing algorithms, the effectiveness of the proposed intensity standardization method for aiding the computer-aided diagnostic algorithms was evaluated. Two sets of brain glioma MR images from two MR scanners were used to build a radiomics model for differentiating high grade glioma (HGG) from lower grade glioma (LGG). The clinical image data and radiomics model are briefly described as follows.

		(a)			
Image Group	1	2	3	4	5
Image number	1904	1472	464	1196	656
Patient number	119	92	29	80	41
Slice thickness (mm)	6	6	6	6	4
Pulse repetition time/Echo time/ Inverse time (ms/ms/ms)	8000/102/2370	9000/1022500	8500/1022439	8800/152/2100	8525/142/2100
Imaging frequency (MHz)	123.1678	123.2407	123.2463	127.7692	127.7706
Pixel bandwidth (Hz/pixel)	287	287	287	195.312	195.312
Pixel spacing (mm)	0.4492	0.4492	0.4492	0.4688	0.4688
Province and I	12-channel	12-channel	12-channel	8-channel	8-channel
Receive coll	head coil	head coil	head coil	head coil	head coil
		(b)			
Image Group		6	7	8	9
Image number		924	548	644	384
Patient number		46	30	32	20
Slice thickness (mm	ι)	5	5	5.5	5
Pulse repetition time/Echo time/ Inverse time (ms/ms/ms)		8500/90/2439	8500/91/2439	9000/83/2500	9000/128/2500
Imaging frequency (MHz)		123.2587	123.2622	123.2001	123.2292
Pixel bandwidth (Hz/pixel)		289	287	201	285
Pixel spacing (mm)		0.6875	0.4297	0.8984	0.8984
Receive coil		12-channel head coil	12-channel head coil	12-channel head coil	32-channel head coil

Table 1. The details of the patient images.

Table 2. The details of the volunteer images.

Image Group	1	2	3	4	5	6
Pulse sequence	Axial T2W BLADE	T1W FLAIR	T2W FLAIR	Axial T2 PRO- PELLER	T1 FLAIR	T2 FLAIR
Slice thickness (mm)	6	6	6	6	6	6
Pulse repetition time/Echo time/ Inverse time (ms/ms/ms)	3500/95/NA	2000/17/857	8000/102/2370	4582/96/NA	1872/22/720	8525/146/2100
Imaging frequency (MHz)	123.1678	123.1678	123.1678	127.7705	127.7705	127.7705
Pixel bandwidth (Hz/pixel)	287	287	287	195.312	195.312	195.312
Pixel spacing (mm)	0.8984	0.8984	0.4492	0.4688	0.4688	0.4688

NA: not applicable.

The MR images of glioma patients from the Department of Neurosurgery, Huashan Hospital affiliated to Fudan University were used, which were collected from a SIEMENS MAGNETOM<sup>(R)</sup> Verio 3.0T MRI scanner and a GE Discovery<sup>TM</sup> MR750 3.0T MRI scanner, respectively. Among them, the images obtained from the SIEMENS scanner were treated as a subset of the reference images in the proposed standardization method, and the images obtained from the GE scanner comprised a subset of one target image group. HGG was considered as positive sample, and LGG was considered as negative sample. Therefore, we used the MR images obtained from the SIEMENS scanner (34 HGG and 32 LGG) to train the model and tested the model with the original images from the GE scanner and the standardized GE images (17 HGG and 11 LGG). The detail acquisition protocols are shown in Table 3.

Group	Train	Test
Instrument	SIEMENS MAGNETOM	GE Discovery MR750
instrument	Verio 3.0T MRI scanner	3.0T MRI scanner
Patient with HGG	34	17
Patient with LGG	32	11
Slice thickness (mm)	6	4
Pulse repetition time/Echo time/ Inverse time (ms/ms/ms)	8000/102/2370	8525/141.9/2100
Imaging frequency (MHz)	123.1678	127.7706
Pixel bandwidth (Hz/pixel)	287	195.312
Transmit coil name	Body	8HRBRAIN
Pixel spacing (mm)	0.4492	0.4688

**Table 3.** The details of patient imaging acquisition protocols for the images used for glioma grade classification.

#### 3.2. Advanced Weak-Pair Data Input Strategy

For the case where the axial position of the input MR image slice is close to the boundary, since the axial density of the reference MR images is lower there, the structural similarity of the weak-pair selection of the reference image may be lowered if the same random criterion is applied. Therefore, to make the selected reference MR image structurally similar to the input MR image and guarantee a certain randomness, an advanced weak-pair data input strategy is proposed. Firstly, the axial location of each brain MR image slice in Montreal Neurological Institute (MNI) space is extracted with SPM12 tool. Then, in every epoch, 30 MR images in the reference dataset axially closest to  $x_{n,m}$  is used to form a set  $Xsel1_{n,m}$ . Meanwhile, MR images in the reference dataset within the range of  $\pm 3$  mm from  $x_{n,m}$  is used to form another set  $Xsel2_{n,m}$ . One image  $x_{N,\tilde{m}}$  in the larger set above is randomly selected to be paired for  $x_{n,m}$  in the training process. After that, a random number  $rand_{n,m}$  is generated according to a Gaussian distribution with  $\mu = 256$  and  $\sigma = N_{epoch \ left} + 64$ . The augmented image size is

$$size_{n,m} = round \left(256 + abs \left(256 - rand_{n,m}\right)\right). \tag{6}$$

Therefore, the  $x_{n,m}$  and corresponding  $x_{N,\tilde{m}}$  are scaled to  $size_{n,m} \times size_{n,m}$ , randomly cropped to 256 × 256, and then randomly flipped. The above random procedures are all synchronized within an image pair to eliminate the structure disparity.

#### 3.3. Optimized Training Strategy with Synchronized Batch Normalization

Batch normalization regards the statistics of neuron outputs with mini-batch input as the estimation of the whole dataset and normalize the outputs according to the estimation in order to prevent internal covariate shift and thus improve model stability. When multiple GPUs are used in training process, the mini-batch is defined as the per-GPU input data, which works well in low resolution scenario.

However, when higher input image resolution and a deeper network are applied, the graphics memory within a GPU can handle several or only one input image. In this case, the statistics of the mini-batch can no longer represent the entire dataset thus the batch normalization degenerates into the instance normalization. Therefore, synchronized batch normalization is applied, which extracts the statistics among all input samples distributed on multiple GPUs. Thus, the estimate based on such samples are more accurate, so that the model stability is guaranteed.

#### 4. Experiments and Results

We firstly evaluated the resolution preservation ability enhancement brought by the proposed method over the original unified method previously proposed by our group. As two unified models, they should be able to convert any MR image within a certain range into an image of a reference

group. Therefore, we first use a larger number of images with the same imaging parameters regarded as the reference image group as the generator input, comparing the resolution loss of the output by measuring consistency of the input and the output. We then used the paired volunteer dataset to measure resolution parameters of the standardized images in the presence of intensity distribution variation between the target and reference datasets.

Then, we compared the proposed method in general properties with the original unified method and also the representative methods of the two major types of intensity standardization methods, the histogram matching method proposed by Sun et al. and the joint histogram registration method previously proposed by our group. To meet the dataset requirements of all above methods, the volunteer dataset was applied.

To conduct a numerical comparison between these methods, the resolution-oriented metrics including the peak signal-to-noise ratio (PSNR), the structural similarity index (SSIM), the visual information fidelity (VIF), the universal quality index (UQI), and the image fidelity criterion (IFC) were utilized. The widely-used general properties containing the gradient magnitude similarity deviation (GMSD) and the mean square error (MSE) were applied. The histogram correlation (HC) and the average disparity (AD) we proposed before were also employed. All numerical results were obtained by comparing the transformed (generated) three-dimensional (3D) image and the reference 3D image and then averaging between different participants of the corresponding experiment. The HC is defined as

$$HC = 1 - \sqrt{1 - \frac{\sum \sqrt{Cnt_{standard} \cdot Cnt_{ref}}}{\sqrt{\sum Cnt_{standard} \times \sum Cnt_{ref}}},$$
(7)

where the  $Cnt_{standard}$  records the values of histogram bins in the standardized image and  $Cnt_{ref}$  records the values of histogram bins in the reference image. The AD is defined as

Average Disparity = 
$$100 \times mean\left(\frac{\left|I_{standard} - I_{ref}\right|}{I_{ref}}\right)$$
, (8)

where  $I_{standard}$  and  $I_{ref}$  are the standardized and reference image, respectively. Better standardization result is denoted with larger PSNR, SSIM, IFC, UQI, VIF, and HC as well as smaller GMSD, MSE, and AD.

#### 4.1. The Resolution Preservation Ability Enhancement Brought by the Optimizations on Methodology

To prove the proposed method's improvement over the resolution preservation of the standardized images, the resolution-related metrics of the methods was firstly explored. For the sake of achieving stable test results on a large test set, the first experiment was conducted by obtaining 320 images with the same image acquisition conditions as Group 1, which is the reference group, in the training set. Then, the standardized images were obtained with both the original unified method and the proposed method. In theory, the standardized images should be the same as the original images, thus we used the original images as the gold standard to calculate the resolution-related features of the two sets of standardized images. PSNR, SSIM, VIF, UQI and IFC were applied.

Table 4 illustrates such a comparison. In this experiment, the proposed method was superior to the original unified method in most resolution-related metrics while the fidelity-related indicators were significantly improved. These demonstrate the improved effectiveness of the proposed method in maintaining MR image resolution.

<b>Evaluation Criteria</b>	PSNR	SSIM	IFC	UQI	VIF
The proposed method	37.31	0.9663	4.164	0.6169	0.1588
Original universal method	37.21	0.9690	3.212	0.6138	0.1233

**Table 4.** The resolution-related metrics comparison between the proposed method and the original universal method over the images with the same conditions as the reference group.

To better reflect the resolution preservation ability enhancement of the proposed method in the presence of intensity distribution variation between input MR images and output ones, we conducted another experiment by comparing the standardization result of volunteer MR images obtained from GE scanner with the paired MR images obtained from SIEMENS scanner. Table 5 illustrates the comparison of PSNR and SSIM. The proposed method not only successfully eliminated the huge difference between the original target image and the reference image, but also had a significant performance improvement compared with the original unified method, due to the absence of resolution-oriented optimizations.

**Table 5.** The resolution-related metrics comparison between the proposed method and the original universal method over the volunteer images.

Evaluation Criteria	PSNR	SSIM	IFC	UQI	VIF
The proposed method	65.51	0.9992	4.318	0.7568	0.2413
Original universal method	65.13	0.9989	4.294	0.7538	0.2431

Figure 3 shows the local details of some of the MR images within the experiment above. Since the proposed method uses a network structure that is more focused on improving image resolution, a loss function with normalized mutual information, an advanced weak-pair data augmentation method, and a multi-GPU training strategy, the output images of the proposed method have a higher local accuracy and resolution. This allows a series of image texture features to be extracted more accurately on the standardized image generated by the proposed method.



**Figure 3.** The local detail of the standardization result of two example MR images with the original and proposed standardization method. The aim is to transform target images into the reference group: (**a1,b1**) the reference MR images, the red boxes are zoomed in (**a2,b2**), respectively; (**a3,b3**) the details of the target MR images; (**a4,b4**) the details of the results produced by the original many-to-one method; and (**a5,b5**) the details of the results produced by the proposed method.

Moreover, the performances of differential diagnosis with the original dataset and the datasets standardized with different methods were evaluated to reflect the advantage of the resolution preservation ability enhancement brought by the proposed method. Classical and highly interpretable radiomics model was applied to better illustrate the superiority of the proposed method. First, the obtained images from glioma patients and their standardized results were segmented using the GrowCut method to obtain a glioma region [26]. Then, for the segmented regions, 555 features including intensity features (21), shape features (15), texture features (39), and wavelet features (480) were extracted using a self-adaptive feature extraction method [9]. The minimum-redundancy-maximum-relevance (mRMR) based genetic-algorithm (GA) was applied for feature selection. Finally, the selected features were used to train the SVM classifier.

Table 6 show the classification result of the original and standardized test set. The accuracy of differential diagnosis between HGG and LGG was noticeably increased. Therefore, we believe that the proposed method can effectively improve the accuracy of differential diagnosis methods by increasing the preservation of image details and texture features as well as wavelet features, which highly rely on image details.

	10010 0. 11		
Test Group	Original Images	Standardized Images with the Original Method	Standardized Images with the Proposed Method
Accuracy	0.7143	0.8214	0.8571
Sensitivity	0.6471	0.7647	0.8824
Specificity	0.8182	0.9091	0.8182

Table 6. The result of HGG/LGG classification

#### 4.2. Other Visual and Numerical Comparison with the Previous Methods

Figure 4 shows the standardized results of the proposed method, the original unified many-to-one method, the histogram matching method, and the joint histogram registration method on the volunteer MR image dataset while Figure 5 shows the logarithmic absolute error image between the standardized image and corresponding reference image. The results of the histogram matching method are deviated in terms of overall image intensity and contrast of gray matter and white matter. The joint histogram registration method has good overall results, but has some image intensity errors in the skull, muscle, and skin areas. Both the original unified method and the proposed method could produce the state-of-the-art image intensity standardization results. In Figure 5, the mean intensity value of the error image corresponding to the proposed method is relatively small, but the spatial distribution of the errors is more extensive. This may be mainly because the standardization results of the proposed method are highly similar to the original target MR image structure, while the original target image and the reference "gold standard" image still have difference due to scanning interval and registration accuracy. Therefore, such difference might be reflected in the error images.

Table 7 shows the numerical comparison for these different methods. The proposed method outperformed the original unified many-to-one method based on CycleGAN in the overwhelming majority of metrics. Meanwhile, the proposed method had a significant effect improvement over the two conventional standardization methods.

In terms of runtime, the histogram matching method is mainly piecewise linear functions, which takes only 0.5867 s to standardize an image slice. The joint histogram registration method requires b-spline fitting, so the amount of calculation is large, and the time taken for processing a slice reaches 2.1651 s. For the original many-to-one universal method and the proposed method, since the convolution is very suitable for parallel computing using the graphics processing units (GPUs), with the help of two NVIDIA TITAN Xp GPUs, the original many-to-one universal method only needs 0.2868 s to transform an MR image slice. Because the parameters are shared between layers in the cascading block and between the cascading blocks, the proposed method only needs 0.1978 s, which is fast enough for a preprocessing step.



**Figure 4.** The tests of four standardization methods on five example slices from volunteers: (a) T2 FLAIR MR images from GE scanner treated as the target; (b) T2 FLAIR MR images from SIEMENS scanner treated as the reference; (c) standardization results with the proposed method; (d) standardization results with the original many-to-one method; (e) standardization results with the histogram matching method; and (f) standardization results with the joint histogram registration method.



**Figure 5.** The logarithmic absolute error image of four standardization methods in Figure 4: (a) standardization results with the proposed method; (b) standardization results with the original many-to-one method; (c) standardization results with the histogram matching method; and (d) standardization results with the joint histogram registration method.

Table 7. The comparison results of the different methods.

Evaluation Criteria	Original Images	The Proposed Method	Original Universal Method	Histogram Matching Method	Joint Histogram Registration Method
PSNR	53.75	65.51	65.13	61.39	64.27
SSIM	0.9829	0.9992	0.9989	0.9968	0.9988
HC	0.2724	0.9273	0.8865	0.8049	0.8825
GMSD	0.1328	0.1026	0.1027	0.1291	0.1245
MSE	$18.50  imes 10^3$	$1.222 \times 10^3$	$1.353  imes 10^3$	$1.652  imes 10^3$	$1.559  imes 10^3$
AD	30.8	4.505	1.093	13.54	8.467

# 5. Discussion

# 5.1. Analysis of the Model Stability of the Anomalies and Lesions

To support various post-processing computer-aided diagnosis models, the dataset we used in training the model includes MR images from patients with various brain diseases. Therefore, the resized and cropped MR images used in the training may be either completely normal or contain various types of lesion/tumor areas. Therefore, the diversity in data can guarantee the robustness of the transformation model in one aspect. Meanwhile, on the other hand, since we use a method based on GANs, all regions of the image, including lesions and tumors, will tend to match the intensity distribution corresponding to the same tissue type of the reference image group during training to minimize the adversarial loss. Furthermore, the consistency-based loss functions including both the mutual information loss and the cycle consistency loss ensure the stability of the structure (including the type of tissue) before and after the transformation (standardization). These loss functions ensure that, although the transformation model has undergone regional intensity changes, such changes should also be stable for abnormal tissues.

## 5.2. The Fusion of Various Losses during the Training Process

The loss functions we use in the proposed method include adversarial loss, cycle consistency loss, and NMI loss. Among them, the adversarial loss is the basis of the generative adversarial network and measures the difference in intensity distribution between the generated image and the reference image. Consistency-based losses including cycle consistency loss and NMI loss measure the feature preservation, especially structure retention of the original target after the intensity distribution transformations of the MR images. All the loss functions are essential for the generator to make high-resolution precise transformed MR images with standardized intensity distribution. However, in the case of multiple losses acting simultaneously, the weight of the loss largely determines the final performance of the model. In our study, the weights of these three losses have been manually fine-tuned to achieve a more consistent decay curve, which has resulted in better convergence results. Moreover, recent studies have pointed out that based on the noise or systematic uncertainty of each loss corresponding to the task, an optimal loss weight may be derived [27,28]. Therefore, we may try to give each loss element a learnable weight while giving a constraint to the overall loss, and automatically obtaining the optimal loss weights based on the systematic uncertainty estimates during the iteration process in the future.

# 6. Conclusions

In this paper, a resolution-oriented universal MR image standardization method is proposed in order to standardize MR images from different MR scanner with different acquisition parameter. The low-dimensional feature extraction residual convolution path in the generator, the loss function, the data augmentation strategy and training strategy are all optimized for achieving better transformation resolution. The experiments show the superiority of the proposed method over the original unified many-to-one method as well as the conventional methods. This method has far-reaching implications for the establishment of large-scale, homogeneous MR image datasets. Based on such datasets, various identification and prediction models can achieve higher robustness.

**Author Contributions:** writing–original draft preparation, Y.G.; writing–review and editing, J.Y. and Y.W. **Funding:** This research was funded by the National Basic Research Program of China under Grant 2015CB755500. **Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

MRI	Magnetic Resonance Imaging
MR	Magnetic Resonance
CycleGAN	Cycle Generative Adversarial Network
PDFs	Probability Density Functions
NMI	Normalized Mutual Information
CARN	Cascading Residual Network

LeakyReLU	Leaky Rectified Linear Unit
LSGAN	Least Squares Generative Adversarial Networks
HGG	High Grade Glioma
LGG	Low Grade Glioma
MNI	Montreal Neurological Institute
PSNR	Peak Signal-to-noise Ratio
SSIM	Structural Similarity
VIF	Visual Information Fidelity
UQI	Universal Quality Index
IFC	Image Fidelity Criterion
GMSD	Gradient Magnitude Similarity Deviation
MSE	Mean Square Error
HC	Histogram Correlation
AD	Average Disparity
mRMR	Minimum-redundancy-maximum-relevance
GA	Genetic-algorithm
GPUs	Graphics Processing Units

# References

- 1. Bradley, J.E.; Panagiotis, K.; Zeynettin, A.; Timothy, L.K. Machine learning for medical imaging. *RadioGraphics* **2017**, *37*, 505–515.
- Yu, J.; Shi, Z.; Lian, Y.; Li, Z.; Liu, T.; Gao, Y.; Wang, Y.; Chen, L.; Mao, Y. Noninvasive IDH1 mutation estimation based on a quantitative radiomics approach for grade II glioma. *Eur. Radiol.* 2017, 27, 3509–3522. [CrossRef] [PubMed]
- Feis, R.A.; Bouts, M.J.R.J.; Panman, J.L.; Jiskoot, L.C.; Dopper, E.G.P.; Schouten, T.M.; Vos, F.D.; Grond, J.V.D.; Swieten, J.C.V.; Rombouts, S.A.R.B. Single-subject classification of presymptomatic frontotemporal dementia mutation carriers using multimodal MRI. *NeuroImage Clin.* 2019, 22, 101718. [CrossRef] [PubMed]
- 4. Narayanan, B.N.; Hardie, R.C.; Kebede, T.M.; Sprague, M.J. Optimized feature selection-based clustering approach for computer-aided detection of lung nodules in different modalities. *Pattern Anal. Appl.* **2019**, *22*, 559–571. [CrossRef]
- Raj, A.; Kuceyeski, A.; Weiner, M. A network diffusion model of disease progression in dementia. *Neuron* 2012, 73, 1204–1215. [CrossRef] [PubMed]
- 6. Shiradkar, R.; Podder, T.K.; Algohary, A.; Viswanath, S.; Ellis, R.J.; Madabhushi, A. Radiomics based targeted radiotherapy planning (Rad-TRaP): A computational framework for prostate cancer treatment planning with MRI. *Radiat. Oncol.* **2016**, *11*, 148. [CrossRef] [PubMed]
- Ingrisch, M.; Schneider, M.J.; Norenberg, D.; Figueiredo, G.N.D.; Maier-Hein, K.; Suchorska, B.; Schuller, U.; Albert, N.; Bruckmann, H.; Reiser, M.; et al. Radiomic analysis reveals prognostic information in T1-weighted baseline magnetic resonance imaging in patients with glioblastoma. *Investig. Radiol.* 2017, 52, 360–366. [CrossRef] [PubMed]
- 8. Narayanan, B.N.; Hardie, R.C.; Kebede, T.M. Performance analysis of a computer-aided detection system for lung nodules in CT at different slice thicknesses. *J. Med. Imaging* **2018**, *1*, 014504. [CrossRef] [PubMed]
- Gao, Y.; Shi, Z.; Wang, Y.; Yu, J.; Chen, L.; Guo, Y.; Zhang, Q.; Mao, Y. Histological grade and type classification of glioma using Magnetic Resonance Imaging. In Proceedings of the 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, Datong, China, 15–17 October 2016; pp. 1808–1813.
- Nyul, L.G.; Udupa, J.K.; Zhang, X. New variants of a method of MRI scale standardization. *IEEE Trans. Med. Imaging* 2000, 19, 143–150. [CrossRef] [PubMed]
- 11. Collewet, G.; Strzelecki, M.; Mariette, F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn. Reson. Imaging* **2004**, *22*, 81–91. [CrossRef] [PubMed]
- 12. Madabhushi, A.; Udupa, J.K. New methods of MR image intensity standardization via generalized scale. *Med. Phys.* **2006**, *33*, 3426–3434. [CrossRef] [PubMed]

- Sun, X.; Shi, L.; Luo, Y.; Yang, W.; Li, H.; Liang, P.; Li, K.; Mok, V.C.; Chu, W.C.; Wang, D. Histogram-based normalization technique on human brain magnetic resonance images from different acquisitions. *Biomed. Eng. Online* 2015, 14, 73. [CrossRef] [PubMed]
- 14. Nunzio, G.D.; Cataldo, R.; Carla, A. Robust intensity standardization in brain magnetic resonance images. *J. Digit. Imaging* **2015**, *28*, 727–737. [CrossRef] [PubMed]
- Jager, F.; Deuerling-Zheng, Y.; Frericks, B.; Wacker, F.; Hornegger, J. A new method for MRI intensity standardization with application to lesion detection in the brain. In Proceedings of the 11th International Fall Workshop Vision, Modeling, and Visualization, Aachen, Germany, 22–24 November 2006; Volume 1010, pp. 269–276.
- 16. Jager, F.; Hornegger, J. Nonrigid registration of joint histograms for intensity standardization in magnetic resonance imaging. *IEEE Trans. Med. Imaging* **2009**, *28*, 137–150. [CrossRef] [PubMed]
- 17. Dzyubachyk, O.; Staring, M.; Reijnierse, M.; Lelieveldt, B.P.; Geest, R.J.V.D. Inter-station intensity standardization for whole-body MR data. *Magn. Reson. Med.* **2017**, *77*, 422–433. [CrossRef] [PubMed]
- Robitaille, N.; Mouiha, A.; Crepeault, B.; Valdivia, F.; Duchesne, S. Tissue-based MRI intensity standardization: application to multicentric datasets. *Int. J. Biomed. Imaging* 2012, 2012, 347120. [CrossRef] [PubMed]
- Gao, Y.; Pan, J.; Guo, Y.; Yu, J.; Zhang, J.; Geng, D.; Wang, Y. N-D point cloud registration for intensity normalization on magnetic resonance images. In Proceedings of the VipIMAGE 2017, Porto, Portugal, 18–20 October 2017; pp. 121–130.
- Gao, Y.; Pan, J.; Guo, Y.; Yu, J.; Zhang, J.; Geng, D.; Wang, Y. Optimised MRI intensity standardisation based on multi-dimensional sub-regional point cloud registration. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* 2019, 7, 1–10. [CrossRef]
- Gao, Y.; Liu, Y.; Wang, Y.; Shi, Z.; Yu, J. A universal intensity standardization method based on a many-to-one weak-paired cycle generative adversarial network for magnetic resonance images. *IEEE Trans. Med. Imaging* 2019, *38*, 2059–2069. [CrossRef] [PubMed]
- Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2242–2251.
- Ronneberger, O.; Fischer, P.; Brox, T. U-net: convolutional networks for biomedical image segmentation. In Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
- 24. Ahn, N.; Kang, B.; Sohn, K.A. Fast, accurate, and lightweight super-resolution with cascading residual network. In Proceedings of the 2018 European Conference onComputer Vision, Munich, Gemany, 8–14 September 2018.
- 25. Wang, C.; Macnaught, G.; Papanastasiou, G.; MacGillivray, T.; Newby, D. Unsupervised learning for cross-domain medical image synthesis uing deformation invariant cycle consistency networks. In Proceedings of the 2018 International Workshop on Simulation and Synthesis in Medical Imaging, Granada, Spain, 16 September 2018; pp. 52–60.
- Ji, C.; Yu, J.; Wang, Y.; Chen, L.; Shi, Z.; Mao, Y. Brain tumor segmentation in MR slices using improved GrowCut algorithm. In Proceedings of the 7th International Conference on Graphic and Image Processing, Singapore, 23–25 October 2015.
- Kendall, A.; Gal, Y.; Cipolla, R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 7482–7491.
- BenTaieb, A.; Hamarneh, G. Uncertainty driven multi-loss fully convolutional networks for histopathology. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention Workshop on Large-scale Annotation of Biomedical data and Expert Label Synthesis (MICCAI LABELS), Quebec City, QC, Canada, 10–14 September 2017; pp. 155–163.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).