MRIQA: Subjective Method and Objective Model for Magnetic Resonance Image Quality Assessment

Qi Chen[1*], Fang Liu[2*], Huiyu Duan[1], Yao Wang[2], Xiongkuo Min[1], Yan Zhou[2†], Guangtao Zhai[1†]
[1] Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai, China
[2] Department of Radiology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China
{chen-qi1997, huiyuduan, minxiongkuo, zhaiguangtao}@sjtu.edu.cn,
{15221732204, wangyao852204526}@163.com, {clare1475}@hotmail.com

2022 IEEE International Conference on Visual Communications and Image Processing (VCIP) | 978-1-6654-7592-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/VCIP56404.2022.10008885

Abstract—**Magnetic Resonance Imaging (MRI) is widely used for medical diagnosis, staging and follow-up of disease. However, MRI images may have artifacts due to various reasons such as patient movement or machine distortion, which may be unintentionally introduced during the procedure of medical image acquisition, processing, *etc*. These artifacts may affect the effectiveness of diagnosis or even cause false diagnosis. To solve this problem, we propose a general medical image quality assessment (MIQA) methodology, including subjective MIQA procedures and objective MIQA algorithms. We further apply this methodology to MRI images in this paper due to its widespread use in practical applications. We first establish a magnetic resonance imaging quality assessment (MRIQA) database, which contains 3809 MRI images. Then a subjective image quality assessment experiment is conducted by expert doctors according to the diagnostic value of these images, which split all MRI images into 1285 low quality images and 2524 high quality images. We then conduct a baseline deep learning experiment, and propose an attention based MIQANet model to automatically separate MRI images into high quality and low quality based on their diagnosis value. Our proposed method achieves a great quality assessment accuracy of 96.59%. The constructed MRIQA database and proposed MIQA model will be public available to further promote medical IQA research.**

Index Terms—**Medical Image, Quality Assessment, Database, Deep Learning**

I. Introduction

During the procedure of medical image acquisition, processing, *etc.*, due to various reasons such as machine noise, electromagnetic interference, human interference, *etc.*, medical images may have artifacts such as motion blur or ghosting artifact [1], [2]. These artifacts result in poor image quality and seriously affect doctors' diagnostic accuracy and confidence [3]. The current quality assessment of medical images is mainly carried out by image quality inspectors to screen low-quality images. These inspectors (Human Observers (HumO)) are now regarded as the gold standard for medical image quality assessment for different tasks. Every day each hospital may generate tens of thousands of MRIs, CT, ultrasound images, *etc*. Each image may have hundreds of slice sequences. In order to detect distorted images, inspectors often need to manually check a large amount of image data, which causes a lot of work load. Therefore, there is an urgent need to construct an objective quality assessment system for different medical images, and automatically leave out or re-photograph low quality medical images for better diagnosis purpose [4]. Different with traditional image quality assessment for natural images, medical images are unnatural and have less aesthetic clue. The quality assessment of medical images should serve the diagnosis clues [5].

In this work, we mainly focus on exploring the subjective and objective methodologies for magnetic resonance imaging, which is widely used for medical diagnosis. Specifically, we first establish a magnetic resonance image quality assessment (MRIQA) database, and invite 8 professional doctors to give ground-truth label based on the diagnostic value of these images. As a result, the database is split into 1285 low quality images and 2524 high quality images. Based on the MRIQA database, we further propose our attention based medical image quality assessment net (MIQANet) for better assessing performance. Compared to other benchmark methods, our model achieves better results and higher robustness. Our MRIQA database, MIQANet and benchmark method will be public available to facilitate related medical IQA research.

In the following of this paper, we first overview some related papers to this work. Then we introduce our proposed magnetic resonance image quality assessment database (MRIQA) and its construction methodology. Finally we introduce our proposed attention based objective quality assessment algorithm and analyze the quality assessment result of our algorithm. We further compare out method with some traditional natural image IQA methods and baseline deep learning classification networks to validate the effectiveness of our model.

II. Related Work

In the past few decades, various image quality assessment algorithms designed to evaluate the quality of natural images have emerged in the scientific research community. Many researchers have also tied some of them to conduct quality assessment of medical images. However, when these algorithms are migrated to the quality assessment of medical images, the effect of them is limited.

Medical image quality assessment methods are roughly divided into two directions: complete reference image quality assessment and non-reference image quality assessment.

- Full reference image quality assessment (FR-IQA): FR-IQA provides both the original reference image and the

* Equal Contribution.
† Corresponding author.

978-1-6654-7592-1/22/$31.00 ©2022 IEEE

distorted image during the objective quality assessment experiment. Such an objective quality assessment experiment is the easiest to carry out.

- No reference image quality assessment (NR-IQA): NR-IQA method does not have any information of the reference image as a reference in the process of calculating the image quality. Compared with FR-IQA, NR-IQA has a wider range of application scenarios, especially in the medical field. Because of the particularity of medical image content and image structure, many poor-quality medical images are not obtained by adding noise or blurring to good-quality reference images like natural images. The distortion of medical images is often caused by the acquisition equipment. Distortion and distortion caused by patient movement, etc. The special nature of the distortion inducement of medical images leads to the fact that most of the time, medical images have only one distorted image and have no corresponding reference image. Therefore, NR-IQA is the most used quality assessment method for medical image quality assessment.

Due to the particularity of medical images, it is difficult for researchers to find a perfect original image as a reference [6] when performing objective quality assessment experiments. Therefore, the most commonly used image quality assessment method in the field of medical images is the No Reference Quality Assessment (NR-IQA) [7], there is still great research potential in this field.

A. Nature Image NR-IQA

NR-IQA can also perform non-reference quality assessment based on natural scene statistics (NSS). The underlying mechanism of the non-reference image quality assessment model based on NSS is that natural image scenes usually follow a certain specific statistical distribution, such as Gaussian distribution And so on, and the existence of distortion may make the image violate these statistical distributions. Therefore, the quality of the distorted image can be effectively judged by the method of scene statistics. The most common NSS-based algorithm is BRISQUE [8], which is a non-reference image quality assessment algorithm in space and within. Kim *et al.* [9] proposed an NR-IQA framework based on convolutional neural network DIQA. This method divides the training of NR-IQA into two stages: the objective distortion part and the HVS-related part. In the first stage, the convolutional neural network mainly learns to predict the objective error map, and then the model learns to predict the subjective quality score of the image in the second stage. Ma *et al.* [10] proposed a multi-task end-to-end optimized deep neural network (MEON) for non-reference image quality assessment. Pan *et al.* [11] proposed a simple and efficient no-reference quality assessment model. The model is a novel network result composed of a fully convolutional neural network (FCNN) and a pooling layer. Zhang *et al.* [12] proposed a deep bilinear model, which is suitable for synthetic and real-distorted images.

B. Medical Image Quality Assessment

Current medical image quality assessment is mainly for magnetic resonance imaging (MRI), computer tomography (CT) and ultrasound imaging images. So far, since medical images have many different image features and contents in various imaging modalities, it is difficult to carry out quality assessment experiments based on the texture and other characteristics of natural images. Therefore, in designing image quality assessment databases and automated quality suitable for medical images There are many difficulties when evaluating algorithms. Therefore, there is no gold standard in the field of medical image quality assessment.

Mortamet *et al.* [13] proposed a fully automatic method for measuring the quality of 3D sMRI images. Woodard *et al.* [14] used variance analysis to determine the two most effective non-reference quality assessment quality metrics: one is based on natural scene statistics, and the other is originally developed to measure distortion caused by image compression. Kalayeh *et al.* [15] proposed two new machine learning-based numerical observers (NO) for medical image quality assessment, and developed a kernel-based regression model. Nakhaie *et al.* [16] proposed an NR-IQA method based on spread spectrum technology and discrete wavelet transform processing the region of interest. Dutta *et al.* [17] reviewed traditional statistical analysis techniques and evaluated the quality of medical images by calculating two key indicators: resolution (determined by local impulse response) and covariance. Brendan *et al.* [18] developed a computational model observer that can reliably learn the detectability of human observers in CT images based on experience and basic knowledge in iterative reconstruction and filtered back-projection reconstruction. Kustner *et al.* [19] proposed a reference-free MRI image quality assessment framework based on machine learning. Lei *et al.* [20] proposed a framework with a multi-task convolutional neural network model, which uses calibration labels for training and supports the two most common artifacts in MRI: noise artifacts and motion blur. Liu *et al.* [21] introduced a multi-site incremental IQA (MSI-IQA) method for sMRI.

III. MRIQA Database

A. MRIQA Data Content

MRI is a commonly used medical image imaging technology based on radiology principles. It is another major advancement in electronic computed tomography (CT) technology and medical imaging, and has become one of the most commonly used medical imaging formats in clinical practice. The nuclear magnetic resonance instrument can obtain the information of the internal tissues of the human body by placing the human body in a special magnetic field, using the phenomenon of nuclear magnetic resonance, processing by the receiver and the corresponding electronic computer.

We first collect 3809 MRI images to construct the database. In real applications, the doctors only need MRI images with good quality, and just leave out low quality MRI images. Therefore, after thorough analysis and discussion with professional doctors, we decide to classify the quality of MRI

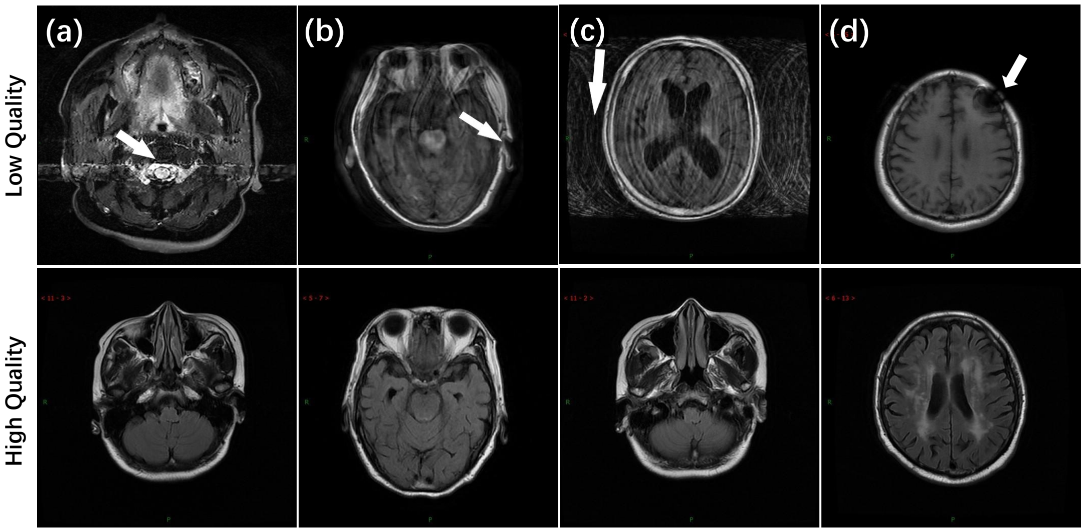

Fig. 1. The first line are MRI images with low diagnostic value in the MRIQA database: (a) chemical shift artifact, (b&c) metal artifact, (d) motion blur, *etc*.; The second line are high diagnostic value images [7]

images into two levels, i.e., low quality images and high quality images. In MRI images, image distortion may be caused by Bofield inhomogeneity, RF noise or irregularities, chemical shifts, ghosting, electromagnetic interference, *etc* [22]. We define the low quality MRI images as the images with ghosting artifacts, chemical shift artifacts, metal artifacts, and motion blur artifacts, *etc.*, since these artifacts are frequently encountered during clinical process, and may seriously affect doctors to make the correct diagnosis. Figure 1 shows some examples of the high quality images and low quality images in our constructed MRIQA database. The low quality images will affect the doctors to make accurate diagnosis and need to be re-photographed.

B. Subjective Quality Assessment Labeling

Images for medical diagnosis usage usually need rigorous data labeling procedure [23]–[26]. Therefore, in the subjective quality assessment and labeling procedure of constructing our MRIQA database, we invite 8 expert doctors to conduct a subjective quality assessment experiment. Medical image quality assessment is very different from traditional natural image quality assessment in subjective scoring, since it is difficult to give multiple quality levels (*e.g.*, 5 levels or 10 levels) for them from aesthetic views. After discussing subjective grading and the diagnostic value of MRI images with professional doctors, we decided to divide the data in the MRIQA database into two categories: low diagnostic quality and high diagnostic quality. Low diagnostic quality means that this MRI image may have one or more of the above-mentioned artifacts, which makes the MRI image unable to assist the doctor in completing an accurate diagnosis, while high diagnostic quality means that the quality of the image is relatively standard, which is very important for doctors in clinical diagnosis. In the MRIQA database, we use the majority vote of 8 expert doctors to determine the quality label of each image, and we finally classify the total 3809 images into 1285 low quality images and 2524 high quality images.

IV. Objective Quality Assessment Algorithm

Our model is implemented based on PyTorch 1.4.0 and all experiments are conducted on a server with two NVIDIA GTX 1080 graphics cards. Our proposed MIQANet realize a deep image quality assessment network based on attention mechanism. Since the image content in the MRIQA database constructed in this paper is not very complicated, and the total amount of data is not large, we choose a relatively shallow deep neural network as our backbone. In our experiments, we find that using Resnet-34 [27] as the backbone network can achieve the best performance. In order to further improve the performance of the model, we introduce the squeeze and excitation (SE) module [28] in our proposed network, and learn the global information in the image through the SE module. The SE module contains two steps: squeeze and excitation: the squeeze step considers the channel dependence, it can compress the global spatial information into the channel descriptor through the global average pooling layer, then the excitation step can be used to completely capture the channel dependency from the information and aggregate in the extrusion step. To learn the nonlinear interaction between channels, the excitation module uses a fully connected neural network (FC) with two hidden layers [28]. The SE module can improve the sensitivity of the network to information characteristics, which can be used in subsequent conversions.

A. Network and Methods

In the general workflow, the preprocessed image, whose size is normalized to 512×512, is fed into our proposed medical image quality assessment network MIQANet (Figure 2), and gradually filtered to extract higher semantic features. The final features are sent to the fully connected layer activated by

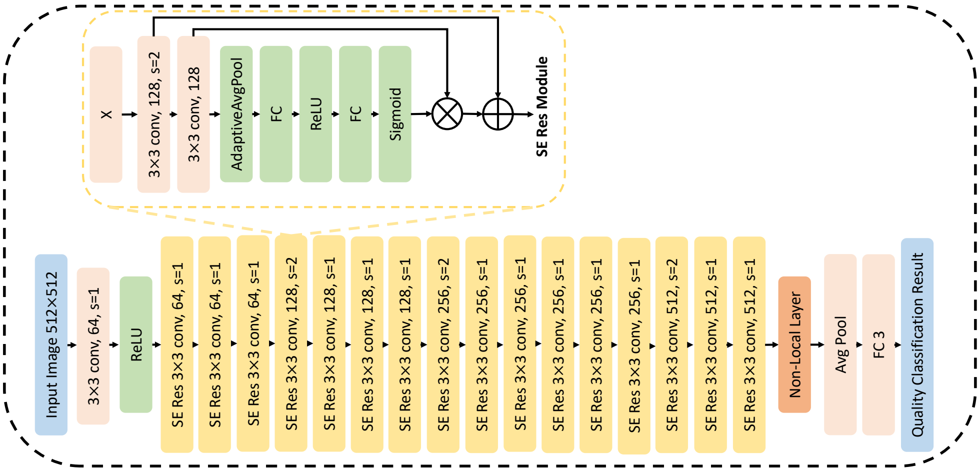

Fig. 2. Proposed Objective Quality Assessment Network **MIQANet**

TABLE I
THE QUALITY ASSESSMENT ACCURACY ANALYSIS OF MIQANET PROPOSED IN THIS EXPERIMENT ON PROPOSED MRIQA DATABASES

Method	Accuracy
BLIINDS-II [29]	56.96%
CORNIA [30]	59.19%
BRISQUE [8]	77.69%
DIIVINE [31]	71.13%
VGG16 [32]	87.14%
SqueezeNet [33]	87.66%
Resnet34 [27]	92.39%
Proposed MIQANet	**96.59%**

softmax to obtain the final classification output. We use cross entropy as our loss function, and the batch size is set to 10 during training. We randomly split the dataset into 10 folds and take 7 folds for training, 1 fold for validation and 2 folds for testing. As a result, our model achieves the best classification performance of **96.59%** on our proposed MRIQA database. It has high robustness and can be effectively applied in practical clinical applications to help reduce the workload of doctors from manually inspect image quality.

B. Experimental Validation and Results

In order to verify the effectiveness of our proposed network, we also compare it with some traditional NR-IQA methods designed for natural image, and some common classification models. We first validate the performance of several traditional non-deep learning NR-IQA methods on our MRIQA database, including BLIINDS-II [29], BRISQUE [8], DIIVINE [31], and CORNIA [30]. These methods can give a quality score for each image, and then we determine a threshold to divide the images into two categories. We also validate the performance of some baseline networks designed for classification tasks in deep learning, such as VGG [32], Resnet [27], Squeezenet [33] and so on. The classification results of our proposed MRIQA database are shown in the Table I. We can conclude from the result that our proposed MIQANet can have better quality assessment performance than traditional methods.

TABLE II
PERFORMANCE OF OUR PROPOSED *MIQANet*

Diagnosis Quality	Precision/%	Recall/%	$F1score/\%$
Low	95.26	94.51	94.88
High	97.24	97.62	97.43

V. CONCLUSION

In addition to the accuracy on the test set, we also introduce other three important indicators of the basic classification task: recall (R), precision (P) and F1 score to effectively evaluate the performance of the quality assessment algorithm in this experiment. The other three classification performance results of our proposed network are shown in Table II.

In this paper, we study the subjective and objective medical image quality assessment method, especially for MRI images. We first construct a MRIQA database, which includes 3809 MRI images. Then a subjective assessment labeling experiment is conducted by 8 professional doctors, which further split all MRI images into 1285 low quality images and 2524 high quality images. Then we build a general medical image quality assessment network MIQANet by integrating the powerful feature extraction capabilities of deep neural networks and combining the attention mechanism. Our proposed network achieves excellent results on MRIQA database. The accuracy of quality classification is of great clinical significance for assisting doctors in automatic image quality assessment, which can effectively reduce the workload of the doctors and improve the effectiveness of their diagnosis.

VI. ACKNOWLEDGEMENT

This work was supported by National Key R&D Program of China 2021YFE0206700, NSFC 61831015 and 61901260, National Natural Science Foundation of China (Grant No. 81901693, 82001457, and 82171885), Shanghai Science and Technology Committee Project (Natural Science Funding: Grant No. 20ZR1433200), Shanghai Science and Technology Committee Project (Explorer Project Funding: Grant No. 21TS1400700), Shanghai "Rising Stars of Medical Talent" Youth Development Program.

REFERENCES

[1] Jeffrey J Ma, Ukash Nakarmi, Cedric Yue Sik Kin, Christopher M Sandino, Joseph Y Cheng, Ali B Syed, Peter Wei, John M Pauly, and Shreyas S Vasanawala, "Diagnostic image quality assessment and classification in medical imaging: Opportunities and challenges," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2020, pp. 337–340.

[2] Qi Chen, Xiongkuo Min, Huiyu Duan, Yucheng Zhu, and Guangtao Zhai, "Muiqa: Image quality assessment database and algorithm for medical ultrasound images," in *2021 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2021, pp. 2958–2962.

[3] Elizabeth A Krupinski, "Current perspectives in medical image perception," *Attention, Perception, & Psychophysics*, vol. 72, no. 5, pp. 1205–1217, 2010.

[4] Lucie Lévêque, Meriem Outtas, Hantao Liu, and Lu Zhang, "Comparative study of the methodologies used for subjective medical image quality assessment," *Physics in Medicine & Biology*, vol. 66, no. 15, pp. 15TR02, 2021.

[5] Meriem Outtas, Lu Zhang, Olivier Deforges, W Hammidouche, Amina Serir, and Christine Cavaro-Ménard, "A study on the usability of opinion-unaware no-reference natural image quality metrics in the context of medical images," in *2016 International Symposium on Signal, Image, Video and Communications (ISIVC)*. IEEE, 2016, pp. 308–313.

[6] Huiyu Duan, Xiongkuo Min, Yucheng Zhu, Guangtao Zhai, Xiaokang Yang, and Patrick Le Callet, "Confusing image quality assessment: Towards better augmented reality experience," *IEEE Transactions on Image Processing (TIP)*, 2022.

[7] Li Sze Chow and Raveendran Paramesran, "Review of medical image quality assessment," *Biomedical signal processing and control*, vol. 27, pp. 145–154, 2016.

[8] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on image processing*, vol. 21, no. 12, pp. 4695–4708, 2012.

[9] Jongyoo Kim, Anh-Duc Nguyen, and Sanghoon Lee, "Deep cnn-based blind image quality predictor," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 1, pp. 11–24, 2018.

[10] Kede Ma, Wentao Liu, Kai Zhang, Zhengfang Duanmu, Zhou Wang, and Wangmeng Zuo, "End-to-end blind image quality assessment using deep neural networks," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1202–1213, 2017.

[11] Da Pan, Ping Shi, Ming Hou, Zefeng Ying, Sizhe Fu, and Yuan Zhang, "Blind predicting similar quality map for image quality assessment," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6373–6382.

[12] Weixia Zhang, Kede Ma, Jia Yan, Dexiang Deng, and Zhou Wang, "Blind image quality assessment using a deep bilinear convolutional neural network," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 1, pp. 36–47, 2018.

[13] Bénédicte Mortamet, Matt A Bernstein, Clifford R Jack Jr, Jeffrey L Gunter, Chadwick Ward, Paula J Britson, Reto Meuli, Jean-Philippe Thiran, and Gunnar Krueger, "Automatic quality assessment in structural brain magnetic resonance imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 62, no. 2, pp. 365–372, 2009.

[14] Jeffrey P Woodard and Monica P Carley-Spencer, "No-reference image quality metrics for structural mri," *Neuroinformatics*, vol. 4, no. 3, pp. 243–262, 2006.

[15] Mahdi M Kalayeh, Thibault Marin, and Jovan G Brankov, "Generalization evaluation of machine learning numerical observers for image quality assessment," *IEEE Transactions on Nuclear Science*, vol. 60, no. 3, pp. 1609–1618, 2013.

[16] Arash Ashtari Nakhaie and Shahriar B Shokouhi, "No reference medical image quality measurement based on spread spectrum and discrete wavelet transform using roi processing," in *2011 24th Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 2011, pp. 000121–000125.

[17] Joyita Dutta, Sangtae Ahn, and Quanzheng Li, "Quantitative statistical methods for image quality assessment," *Theranostics*, vol. 3, no. 10, pp. 741, 2013.

[18] Brendan L Eck, Rachid Fahmi, Kevin M Brown, Stanislav Zabic, Nilgoun Raihani, Jun Miao, and David L Wilson, "Computational and human observer image quality evaluation of low dose, knowledge-based ct iterative reconstruction," *Medical physics*, vol. 42, no. 10, pp. 6098–6111, 2015.

[19] Thomas Küstner, Sergios Gatidis, Annika Liebgott, Martin Schwartz, Lukas Mauch, Petros Martirosian, Holger Schmidt, Nina F Schwenzer, Konstantin Nikolaou, Fabian Bamberg, et al., "A machine-learning framework for automatic reference-free quality assessment in mri," *Magnetic resonance imaging*, vol. 53, pp. 134–147, 2018.

[20] Ke Lei, John M Pauly, and Shreyas S Vasanawala, "Artifact-and content-specific quality assessment for mri with image rulers," *arXiv preprint arXiv:2111.03780*, 2021.

[21] Siyuan Liu, Kim-Han Thung, Weili Lin, and Pew-Thian Yap, "Multi-site incremental image quality assessment of structural mri via consensus adversarial representation adaptation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 381–389.

[22] Steren Chabert, Juan Sebastian Castro, Leonardo Muñoz, Pablo Cox, Rodrigo Riveros, Juan Vielma, Gamaliel Huerta, Marvin Querales, Carolina Saavedra, Alejandro Veloz, et al., "Image quality assessment to emulate experts' perception in lumbar mri using machine learning," *Applied Sciences*, vol. 11, no. 14, pp. 6616, 2021.

[23] Huiyu Duan, Guangtao Zhai, Xiongkuo Min, Yi Fang, Zhaohui Che, Xiaokang Yang, Cheng Zhi, Hua Yang, and Ning Liu, "Learning to predict where the children with asd look," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 704–708.

[24] Huiyu Duan, Guangtao Zhai, Xiongkuo Min, Zhaohui Che, Yi Fang, Xiaokang Yang, Jesús Gutiérrez, and Patrick Le Callet, "A dataset of eye movements for the children with autism spectrum disorder," in *Proceedings of the ACM Multimedia Systems Conference*, 2019, pp. 255–260.

[25] Huiyu Duan, Xiongkuo Min, Yi Fang, Lei Fan, Xiaokang Yang, and Guangtao Zhai, "Visual attention analysis and prediction on human faces for children with autism spectrum disorder," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 15, no. 3s, pp. 1–23, 2019.

[26] Huiyu Duan, Guangtao Zhai, Xiongkuo Min, Yucheng Zhu, Wei Sun, and Xiaokang Yang, "Assessment of visually induced motion sickness in immersive videos," in *Pacific Rim Conference on Multimedia*. Springer, 2017, pp. 662–672.

[27] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[28] Jie Hu, Li Shen, and Gang Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.

[29] Michele A Saad, Alan C Bovik, and Christophe Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339–3352, 2012.

[30] Peng Ye, Jayant Kumar, Le Kang, and David Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012, pp. 1098–1105.

[31] Anush Krishna Moorthy and Alan Conrad Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE transactions on Image Processing*, vol. 20, no. 12, pp. 3350–3364, 2011.

[32] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[33] Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and $<$ 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.