IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 18 December 2024, accepted 31 December 2024, date of publication 7 January 2025, date of current version 16 January 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3526684

RESEARCH ARTICLE

Combined Oriented Data Augmentation Method for Brain MRI Images

AHMEED SULIMAN FARHAN[1,2], (Member, IEEE), MUHAMMAD KHALID[1], AND UMAR MANZOOR[3]

[1]School of Computer Sciences, University of Hull, HU6 7RX Hull, U.K.
[2]Computer Center, University of Anbar, Ramadi 31001, Iraq
[3]School of Engineering, Computing and Mathematical Sciences, University of Wolverhampton, WV1 1LY Wolverhampton, U.K.

Corresponding author: Ahmeed Suliman Farhan (ahmeedsuliman@uoanbar.edu.iq)

ABSTRACT In recent years, deep learning's use in medical imaging has grown exponentially. However, one of the biggest problems with training deep learning models is the unavailability of large amounts of data, which leads to overfitting. Collecting large quantities of labelled medical images is expensive, time-consuming, and depends on specialists' availability. In this paper, we proposed a novel method namely Oriented Combination MRI (OCMRI) for augmenting brain MRI dataset. The proposed method helps CNN models overcome overfitting and address class imbalance issues by combining Brain MRI images to generate new images. The image fusion is performed by selecting two images of the same tumor class if the Mean Squared Error (MSE) between these two images is greater than threshold 1 and lower than threshold 2. Both thresholds are adjustable, initially set by the user and automatically fine-tuned by the algorithm to control the number of images produced for each class, thus helping to address the data imbalance problem. The proposed approach was evaluated by training and testing the PRCnet model on four publicly available datasets before and after applying the proposed method to the datasets. Where the classification accuracy without data augmentation was 85.19% for dataset A, 90.12% for dataset B, 94.77% for dataset C, and 90% for dataset D respectively. After adding the synthetic data; the accuracy improved to 92.7% for dataset A, 95.37% for dataset B, 96.51% for dataset C and 98% for dataset D respectively.

INDEX TERMS Data augmentation, brain tumor, medical imaging, deep learning, MRI, brain tumor classification, convolutional neural network.

I. INTRODUCTION

In recent years, notable advancements have been witnessed in disease diagnosis and treatment through the application of computer models [1], [2]. Particularly, deep neural networks have emerged as pivotal tools in the realm of medical image processing, such as the segmentation of tumor or classification of tumor type [3], [4], [5], [6]. Brain tumors are directly responsible for cancer-related deaths, and are the most common type of cancer in the world [7], [8]. Tumors are formed due to abnormal or excessive growth of cells and brain tumors are of two forms namely benign and malignant [9]. Benign tumors develop relatively slowly whereas malignant tumors grow more rapidly and aggressively invade the surrounding cells, spread to other brain cells, and significantly affect the neurological system [10], [11], [12], [13]. MRI is considered one of the best and most accurate tests for detecting and identifying brain tumors because it produces high-quality, multi-dimensional images [14], [15], [16]. MRI detects changes in the structure of the brain Which enhances the ability to detect tumor. The brain MRI can be imaged in three planes axial, coronal, and sagittal as shown in Figure 1 [17], [18].

Early tumour detection increases the patient's chance of survival and plays an essential role in assessing the condition and planning treatment. However, due to the heterogeneous nature of cancer cells, manual tumor detection is complex, time-consuming, prone to errors and depends on radiologist's expertise. Therefore, developing a reliable and accurate system is crucial to assist

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

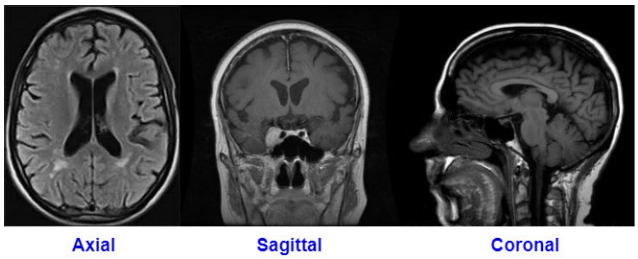

FIGURE 1. MRI image planes.

radiologists [10], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28].

CNN performed exceptionally well on medical image classification and disease detection tasks and, in some cases, even outperformed human experts [29]. Recent studies have reinforced the significant role of convolutional neural networks (CNNs) in the analysis of brain tumors using MRI data. Aamir et al. demonstrated the effectiveness of utilizing segmentation and clustering techniques to identify high-quality tumor regions, significantly improving classification performance [30]. Guan et al. tackled the challenge of low-quality MRI images by improving image quality through contrast enhancement and non-linear stretching techniques for tumour localization and refinement. EfficientNet is used for feature extraction and data augmentation is incorporated to address the issue of overfitting [31]. Despite significant advancements, deep learning faces considerable challenges when applied to medical imaging, primarily due to the limited availability of training data. Acquiring labeled medical data is particularly difficult because of patient privacy concerns, as well as the substantial costs and time required for data collection and annotation. Furthermore, convolutional neural network (CNN) models contain numerous parameters that must be trained on large datasets to mitigate the risk of overfitting [32], [33]. To tackle these challenges, data augmentation techniques are frequently employed to generate additional training data by applying modifications to the existing data while maintaining their original labels. This approach enhances the model's generalizability and ultimately improves its performance [34], [35], [36], [37].

Imbalanced medical image datasets pose a significant challenge that can adversely impact the performance of deep learning models. This issue arises when the distribution of instances across various classes is uneven, a condition referred to as ''imbalanced data.'' In such cases, certain classes, often associated with rare diseases, have substantially fewer examples compared to others, making it difficult to collect a balanced dataset. This imbalance can lead machine learning models to favor the majority class, introducing bias that results in suboptimal performance, particularly in accurately identifying and classifying instances from minority classes. Although the model may achieve high overall accuracy by predominantly predicting the majority class, its performance in detecting minority classes – where examples are scarce – tends to be inadequate. This disparity underscores the importance of addressing data imbalance in the development of robust and reliable deep learning systems for medical image analysis [38], [39].

Although data augmentation techniques generally enhance the performance of deep learning models, not all methods are suitable or effective for medical datasets. The appropriate augmentation techniques must aligns with the unique characteristics of the given dataset [40]. In this paper, we proposed the state-of-the-art method called OCMRI for brain MRI image data augmentation. The proposed method helps to address the lack of data and data imbalance in brain MRI datasets. The proposed method when evaluated on the state-of-the-art brain tumour classification model namely PRCnet [41] leads to overall improvement in the performance on different dataset.

The main contributions of the OCMRI method are:

- A novel method called OCMRI for brain MRI image augmentation.
- Contributes to addressing the lack of data and data imbalance in brain MRI datasets.
- OCMRI customisable threshold values allow the user to control the number of images produced for each class which help in solving the imbalanced data problem.
- OCMRI is tested on four different databases, and the results desmostrate the effectiveness its effect on the efficiency of training the PRCnet model was tested.

II. LITERATURE REVIEW

Data augmentation is a technique used to expand training datasets by synthetically generating image-label pairs from existing data. This approach has gained significant attention in deep learning applications, especially with the emergence of deep convolutional neural networks (CNNs) [35]. In this section, we review established data augmentation strategies originally developed for general computer vision tasks, as well as recent advancements specifically tailored for the analysis of medical images using deep learning models.

A. BASIC AUGMENTATION TECHNIQUES

Basic augmentation techniques encompass a range of methods, including applying various transformations to input images, performing oversampling, or adjusting image intensity values, to generate augmented versions of the original images. Table 1 shows common techniques that are used for basic data augmentation [35], [36], [42], [43], [44].

B. DEEP LEARNING AUGMENTATION METHOD

The methods of augmenting data based on deep learning models to create realistic images can help address the problem of overfitting during training. The process of artificially producing images is referred to as image synthesis [36]. The architecture of most deep learning networks for data augmentation is based on the generative adversarial network (GAN). Generative Adversarial Networks (GANs) have garnered significant attention for their exceptional

TABLE 1. The common basic augmentation techniques.

Techniques	Summary
Flipping	Flipping is one of the most common data augmentations techniques which is widely adopted in training deep learning models. The flipping is either on horizontal axes or vertical axes [35], [43].
Image cropping	Image cropping is the process of cropping a random section of the image and then resizing the section to the size of the original image [35], [43], [44].
Rotation	The image is rotated right or left on an axis between 1° and 359° to perform rotation augmentations [42]. Rotating images by a large degree may not be helpful and may lead to a change in the data. For example, in the images of numbers, when a number is rotated 9 180°, it becomes 6. Therefore, the degree of rotation must be considered according to the data.
Noise injection	Noise injection is an important method for data augmentation, it helps avoid overfitting in the deep learning models [43], [45].
Color space	Changing the color distribution of images or changing random factors of color saturation such as brightness and contrast may contribute to training the model and overcoming overfitting. In addition, applying color jittering in the medical field is primarily driven by the desire to simulate observations made under varying lighting conditions and variances brought on by various imaging and scanning hardware [35].
Merging images	In this technique, the new image is created by merging two or more original images [36]. For example, Takahashi R. et al. suggested a new data augmentation called RICAP, which randomly crops four regions from four images and then merges these regions to create a new image. In addition, the proposed method combines the four image labels to produce the advantage of soft labels [46].
Random erasing	Random erasing is based on randomly selecting a rectangle from the image and erasing the pixels in the specified area [47].

performance, particularly in the context of data augmentation. These networks have the capacity to produce fresh training data, which can enhance the performance of classification models [48].

Bowles et al. [49] studied the feasibility of augmenting training data using GAN in two brain segmentation tasks and this leads to improvements between 1 and 5 percentage in the Dice Similarity Coefficient (DSC).

Mukherkjee et al. [9] proposed AGGrGAN to generate synthetic brain tumors MRI images by combining three base GAN models namely Wasserstein GAN (WGAN) [50] and two variants of Deep Convolutional Generative Adversarial Network (DCGAN) [51]. AGGrGAN was evaluated on two datasets namely the Multimodal Brain Tumor Segmentation Challenge (BraTS) 2020 and the brain tumor and experimental results demonstrates that the proposed model can produce fine-quality images with maximum Structural Similarity Index Measure (SSIM) scores of 0.57 and 0.83 respectively.

Li et al. [52] proposed a novel model called TumorGAN to generate images for brain tumor segmentation pairs based on unpaired adversarial training and used BraTS 2017 dataset for evaluation. The experimental results show that the synthetic data pairs produced by TumorGAN improve the performance of tumor segmentation.

Han et al. [53] utilized GANs to produce a comprehensive and diverse dataset of brain Magnetic Resonance (MR) images, aiming to address the challenges associated with training Convolutional Neural Networks (CNNs) for brain tumor detection. Specifically, Progressive Growing GANs (PGGANs) were employed as an advanced generative method to create full-scale 256 × 256 MR images, effectively mitigating the instability typically encountered with traditional GANs. The findings demonstrate that data augmentation using PGGANs leads to significant improvements in tumor detection performance.

Frid-Adar et al. [54] employed GANs to produce synthetic medical images of liver lesions. The dataset utilized in this study comprised of 182 computed tomography (CT) images of liver lesions. Before incorporating synthetic data augmentation, the model achieved a sensitivity of 78.6% and a specificity of 88.4%. With the addition of synthetic data augmentation, these metrics improved to 85.7% sensitivity and 92.4% respectively.

Konidaris et al. [55] introduced an innovative data augmentation approach leveraging GANs to create realistic synthetic images, specifically targeting the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset for the classification of Alzheimer's Disease and normal control subjects. The integration of GAN-generated images notably enhances model performance, achieving an 11.68% improvement in accuracy compared to conventional methods.

Sharma and Sampath [56] proposed a data augmentation strategy to improve brain tumor segmentation using the MONAI framework, a specialized deep learning toolkit for medical imaging. Leveraging MONAI's comprehensive data augmentation capabilities, including geometric transformations and intensity variations, the authors enriched MRI brain tumor datasets to boost segmentation performance. Experimental results demonstrate that this approach significantly enhances model robustness by increasing the diversity of the training dataset, thereby improving the accuracy of tumor segmentation.

C. LIMITATIONS OF EXISTING AUGMENTATION TECHNIQUES

Data augmentation techniques, including geometric transformations (e.g. rotation, flipping, cropping, and scaling), intensity-based modifications, and synthetic data generation, are extensively utilized in medical imaging. While these methods have demonstrated significant potential in

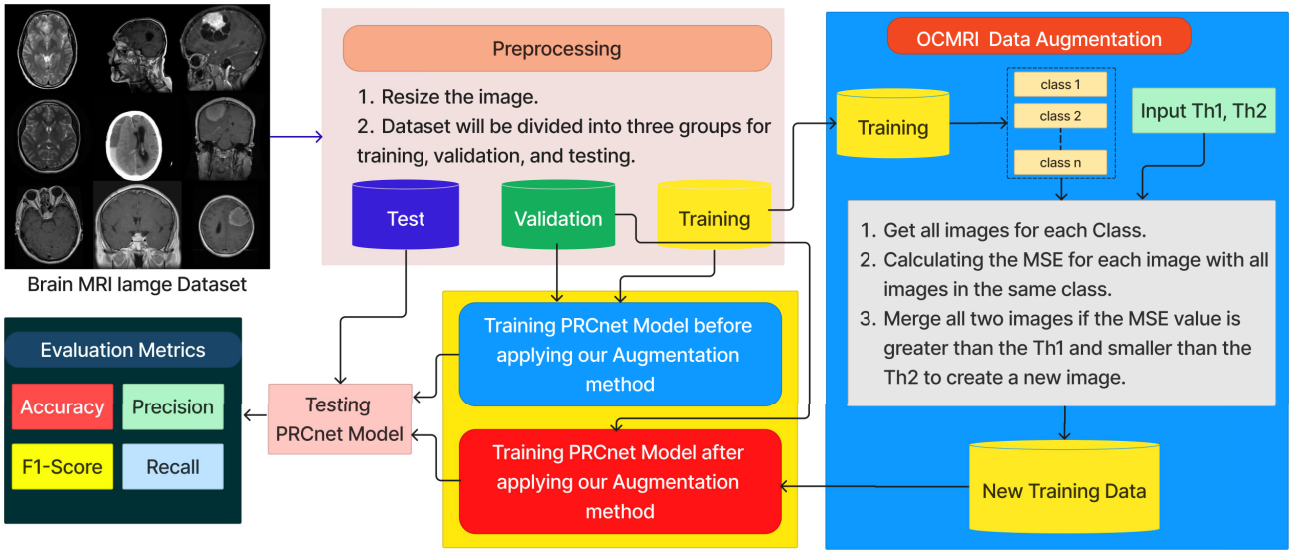

FIGURE 2. Flow of various steps including pre-processing, a new augmentation method, and evaluation matrix.

enhancing the performance of deep learning models, they exhibit limitations in addressing key challenges, such as:

1- Data Insufficiency and Diversity: Traditional augmentation techniques struggle to produce the diversity needed for deep learning models to generalize effectively. They often fail to replicate realistic variations within the data distribution.

2- Class Imbalance: Standard augmentation approaches apply transformations uniformly across all classes, which can exacerbate class imbalance issues in medical datasets.

3- Context Preservation: Basic transformations may introduce artifacts or distort critical elements of medical images, compromising diagnostic accuracy and clinical reliability.

The OCMRI Solution: Addressing Limitations with Context-Aware Augmentation The proposed Optimized Contextual Medical Image Augmentation (OCMRI) technique has been designed to address these limitations through a novel, context-dependent approach. OCMRI introduces several key innovations:

1- Automatic Augmentation Control: OCMRI employs threshold values (Th1 and Th2) to automatically regulate the number of augmented images generated for each class, ensuring balanced representation across the dataset.

2- Realistic Data Synthesis: By integrating complementary features across multiple samples, OCMRI produces augmented images that maintain anatomical realism while significantly enhancing data diversity.

3- Focus on Minority Classes: The technique prioritizes the augmentation of underrepresented classes, such as rare tumor types, reducing class imbalance and improving the model's ability to accurately represent and diagnose less common medical conditions.

Through these advancements, OCMRI offers a robust solution to traditional augmentation challenges, enabling the development of more accurate and generalizable deep learning models in medical imaging.

III. PROPOSED METHODOLOGY

The proposed data augmentation technique, namely Optimized Contextual Medical Image Augmentation (OCMRI) is a novel technique for brain MRI augmentation. Figure 2 illustrates the five key stages of the proposed augmentation approach. The initial stage involves data acquisition, whereby four publicly available datasets are selected to assess the efficacy of the proposed augmentation method. Subsequently, the second stage encompasses data pre-processing, whereby all images undergo resizing to align with the input dimensions of the Convolutional Neural Network (CNN) model. Additionally, each dataset is partitioned into three groups–training, test, and validation groups, as outlined in Table 2. Moving to the third stage, the proposed augmentation method is applied to the training group of each dataset for the generation of augmented samples. The fourth stage involves the training and testing of the PRCnet model on targeted datasets before applying the proposed augmentation method, followed by re-training and testing after applying the proposed augmentation method. Finally, the fifth stage culminates in evaluating the PRCnet model's performance on the testing data of each dataset, both before and after the application of the novel augmentation method.

These five stages collectively form a comprehensive framework, wherein the proposed augmentation method plays a pivotal role in enhancing the training efficiency of the PRCnet model for brain MRI image classification. The ensuing sections delve into the specific details of each stage,

providing a deeper understanding of the intricacies involved in the proposed methodology and its impact on the overall performance of the CNN model.

A. DATASETS ACQUISITION

In this research, four public datasets were used (Brain MRI [57], brain-tumor-dataset [58], Brain tumor dataset [59], Brain MRI Images for Brain Tumor Detection [60]). We referred to these four datasets as Dataset A, Dataset B, Dataset C, and Dataset D, respectively.

1) DATASET A

This dataset includes 3264 labelled images of four types: 901 MRI images of pituitary tumors, 500 MRI images of no tumor, 937 MRI images of meningioma, and 926 MRI images of glioma [57].

2) DATASET B

This dataset consists of 4,292 slices grouped into four categories: no tumor (681 slices), glioma (1038 slices), meningioma (1318 slices), and pituitary (1,255 slices) [58].

3) DATASET C

This dataset comprises a total of 3064 brain MRI images gathered from 233 patients. It contains three classes based on the type of brain tumor. It includes 1426 MRI images of glioma tumors, 708 MRI images of meningioma tumors, and 930 MRI images of pituitary tumors [59], [61].

4) DATASET D

This dataset consists of 253 MRI images classified into two parts, 155 MRI images with a tumor and 98 MRI images without tumor [60].

All images in all datasets were resized to 224 × 224 to fit the PRCnet Model input layers.

TABLE 2. Distribution of the datasets.

Dataset	Classes	total	training	Validation	Testing
Dataset A	glioma	926	648 (70%)	138 (15%)	140 (15%)
	meningioma	937	655 (70%)	140 (15%)	142 (15%)
	no tumor	500	350 (70%)	75 (15%)	75 (15%)
	pituitary	901	630 (70%)	135 (15%)	136 (15%)
Dataset B	glioma	1038	726 (70%)	155 (15%)	157 (15%)
	meningioma	1318	922 (70%)	197 (15%)	199 (15%)
	no tumor	681	476 (70%)	102 (15%)	103 (15%)
	pituitary	1255	878 (70%)	188 (15%)	189 (15%)
Dataset C	glioma	1426	998 (70%)	214 (15%)	214 (15%)
	meningioma	708	496 (70%)	106 (15%)	106 (15%)
	pituitary	930	652 (70%)	139 (15%)	139 (15%)
Dataset D	no tumor	98	68 (70%)	11 (10%)	19 (20%)
	with a tumor	155	105 (70%)	19 (10%)	31 (20%)

B. THE PROPOSED AUGMENTATION METHOD

This paper introduces an innovative augmentation method designed to address the challenges posed by limited data in deep learning models and mitigate issues related to class imbalance. The proposed methodology centers around the fusion of brain MRI images to generate new, diverse images, providing a solution to data scarcity. Figure 3 illustrates the augmentation method in detail, with the process involving the merging of two selected brain MRI images. The merging operation is executed by extracting odd columns from the first image and even columns from the second image as demonstrated in Figure 4. The selection of the two images involved in the merging process is not arbitrary; rather, it is determined through a calculated Mean Squared Error (MSE) between the two images. Specifically, the choice is made if the MSE value is larger than the first threshold (Th1) and less than the second threshold (Th2). Equation 1 shows how the MSE is calculated [62]. The value of the Th1 and Th2 is variable and is entered by the user. Changing the values of the two thresholds affects the number of images that are produced. We initialize Th1 and Th2 based on the dataset structure and the desired diversity of augmented images.

1) TH1 (LOW THRESHOLD)

Regulates the minimum similarity between two images, enabling a meaningful fusion process. Reducing the Th1 value too much leads to redundancy. A higher Th1 value increases diversity but may reduce the structural integrity of the merged images.

2) TH2 (UPPER THRESHOLD)

Determines the maximum difference allowed. Too much Th2 value can produce noisy or blurry images.

The user sets the acceptable class imbalance ratio before the augmentation method starts. During the augmentation method process, the method dynamically adjusts the thresholds based on the data distribution. The algorithm will cease tuning Th2 once the number of images produced for all classes has a balanced distribution and an acceptable error margin.

$$MSE = \frac{1}{WH}\sum_{i=0}^{W}\sum_{j=0}^{H}[im1(i,j) - im2(i,j)]^2 \qquad (1)$$

We use the MSE between two images to determine which images to merge in our method due to the following reasons:

3) ENSURING IMAGE SIMILARITY

Merging two medical images generates a new synthetic image that preserves key features from both images. MSE measures the pixel-wise differences between two images. By setting upper and lower thresholds on the MSE values, we ensure that the selected images are similar enough in structure but not so identical that the new synthetic image lacks diversity. This helps us create new images that still represent their class but add enough variation to improve the model's generalization.

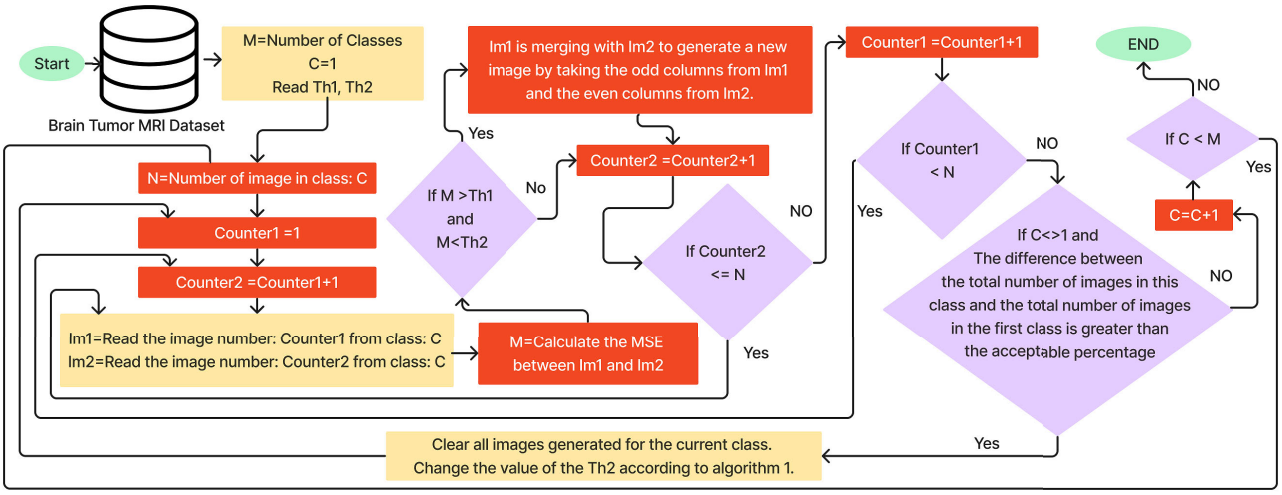

FIGURE 3. Flowchart for the new augmentation method.

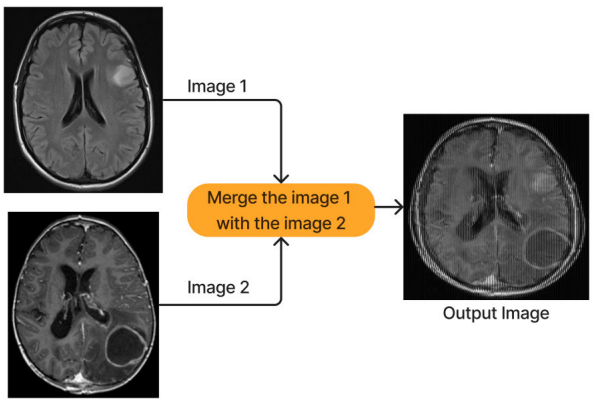

FIGURE 4. The result of merge image 1 with image 2.

4) AVOIDING REDUNDANT DATA

By using MSE thresholds, we can avoid selecting two images that are too similar or different. Merging remarkably similar images would result in synthetic images that provide little new information to the model. On the other hand, merging highly dissimilar images could produce unrealistic or noisy data that confuses the model. The controlled selection process based on MSE ensures we generate useful and varied synthetic data.

5) FLEXIBILITY OF MSE THRESHOLDS

Adjustable MSE thresholds provide control over the augmentation process. By changing the threshold limits, we can fine-tune the diversity of the generated images and ensure that the process generates a balanced dataset, which is particularly helpful in dealing with class imbalance. The proposed augmentation method automates the control of the Th1 and Th2. The user initiates the process by providing values for these thresholds, and the method subsequently generates images for the first class based on the user input thresholds. Following this, the total number of images for the first class is computed by combining the original images with the newly generated ones. The process then advances to the second class, where image generation ensues, and the total number for the second class is calculated. This number is then compared with the total images in the first class. In the event of an unacceptable difference in the number of images between classes, the value of the Th2 is dynamically adjusted, either increased or decreased, until the difference becomes acceptable. This iterative adjustment ensures a balanced distribution of images across different classes, thereby alleviating data imbalance concerns. The process continues iteratively for subsequent classes, ensuring optimal adjustment of Th2 to achieve balanced representation. This adaptability is a key advantage over other methods that do not allow for dynamic control over the selection process. Algorithm 1 elucidates the intricacies of this adaptive threshold control process, providing a detailed step-by-step of the automated mechanism implemented by the proposed augmentation method.

Note the acceptable difference percentage is determined by the user.

Advantages of our method compared with other methods: Many other data augmentation methods (e.g., random transformations like flipping, rotation, and cropping) are helpful. However, it does not consider the structure of medical images. Such random augmentations may distort diagnostic features, which is especially problematic in medical imaging, where even subtle differences matter. In contrast, our method ensures that the merged images preserve information related to brain tumors. This results in more meaningful data augmentation that enhances model performance.

Algorithm 1 Steps to Change the Second Threshold Value

Input: Th2, min, max, accept error.

```
/* Th2: represents the second threshold
   value.                                  */
/* min, max: They are initially assigned
   the value of the first and second
   threshold.                              */
```

diff = (Total number of images in the current class) - (total number of images in the first class)

if | *Diff* | < *accept error* **then**
    Save the images generated from the current category.
    go to the next category or stop if there are no other categories.
else
    if *diff* >*0* **then**
        max ← ((Th2 - min) * 2) + min
        min ← Th2
        Th2 ← max
    else
        Th2 ← ((Th2 - min) / 2) + min
        max ← Th2
    end
end

IV. RESULT

In this section, we will present the evaluation metrics, and the results of the proposed augmentation method on the four datasets and the effect of the proposed augmentation method on the accuracy of the PRCnet model.

Note: the training procedure, hyperparameters, and model selection criteria for the PRCnet model were implemented as described in the original PRCnet paper [41].

A. EVALUATION METRICS

Performance measures are essential metrics used to evaluate the effectiveness and accuracy of deep learning models across various tasks [63]. The choice of multiple metrics is pivotal, as the results of any deep learning model may be good against specific metrics and not be good or bad against other metrics. Metrics contribute to a nuanced understanding of a model's strengths and weaknesses in addressing specific classification challenges [64]. Common performance measures include accuracy, precision, recall, and F1-score, each serving a distinct role in capturing different aspects of classification performance [65], [66].

1) **Accuracy**: Accuracy is a metric used to evaluate the overall performance of deep learning models at scale. It represents the number of instances of correct predictions divided by the total number of instances.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{2}$$

2) **Recall**: Recall or sensitivity focuses on a model's ability to identify positive cases among all true positives correctly.

$$Recall = \frac{TP}{(TP + FN)} \tag{3}$$

3) **Precision**: The proportion of correctly predicted positives among all predicted positives. It is an important metric in medical applications where false positives can have serious consequences.

$$Precision = \frac{TP}{(TP + FP)} \tag{4}$$

4) **F1-Score**: An F1 score is the harmonic mean of precision and recall, providing a balanced metric that considers false positives and false negatives. This metric is particularly valuable in scenarios where there is a need to balance precision and recall, as it provides an overall measure of the model's effectiveness.

$$F1Score = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \tag{5}$$

5) **Specificity**:

$$Specificity = \frac{TN}{(TN + FP)} \tag{6}$$

where,TN = True Negative, FN = False Negative, TP = True Positive, and FP = False Positive.

The combination of accuracy, precision, recall, and F1-score allows us to comprehensively evaluate our proposed Oriented Combination MRI (OCMRI) method. While accuracy gives a broad overview of the model's correctness, precision and recall provide insights into specific aspects of performance related to false positives and false negatives. The F1-score, being a harmonic mean, emphasizes the balance between precision and recall, making it particularly useful in medical image classification tasks where both aspects are critical.

B. EXPERIMENT RESULTS

The proposed augmentation method was verified using four different datasets. After the datasets were identified, each dataset was divided into training, validation, and testing sets, as shown in Table 2. Then, the OCMRI augmentation method was applied to the training set for each dataset. The initial value for Th1 chosen for all datasets was 500, while the initial value for th2 was 1300 for datasets A and B and was 2000 and 1150 for datasets D and C, respectively. The initial value for th2 was determined according to the number of images that were produced. Table 3 shows the number of images of the training set for each class before and after applying the OCMRI augmentation method. It also shows the th2 values for each class, where the algorithm changes them depending on the number of images produced. Changing the th2 values automatically for each class according to the number of images produced contributes to solving the problem of data imbalance in training. The Table 4 shows the details of applying the OCMRI augmentation method to dataset A and how the algorithm repeats the steps and changes

the value of the second threshold to solve the data imbalance problem. Table 5 summarizes the training group sizes for each dataset before and after applying the OCMRI augmentation method.

TABLE 3. Number of training images for each class before and after applying the OCMRI augmentation method. B: Before applying augmentation. A: After applying augmentation. R: The percentage increase in images after augmentation.

Dataset	Classes	Number of training images			Th1	Th2
		B	A	R		
DatasetA	glioma	648	4901	656.32%	500	825
	meningioma	655	4740	623.66%	500	1300
	no tumor	350	5283	1409.42%	500	2000
	pituitary	630	5165	719.84%	500	1300
DatasetB	glioma	726	5084	600.27%	500	825
	meningioma	922	5362	481.56%	500	1300
	no tumor	476	5170	986.13%	500	1600
	pituitary	878	4978	466.97%	500	1150
DatasetC	glioma	998	2758	176.35%	500	580
	meningioma	496	2503	404.63%	500	1150
	pituitary	652	2524	287.11%	500	1108
DatasetD	no tumor	68	188	176.47%	500	1719
	with tumor	105	201	91.42%	500	2000

TABLE 4. Details of iteration to produce images to Dataset A. N: Number of image before augmentation. R: Number of image after augmentation.

Classes	N	Th1	Th2	R	State
Meningioma	655	500	1300	4740	Done
Glioma	648	500	1300	36974	Change Th2
	648	500	900	7479	Change Th2
	648	500	700	2208	Change Th2
	648	500	800	3566	Change Th2
	648	500	850	5658	Change Th2
	648	500	825	4901	Done
Pituitary	630	500	1300	5165	Done
No tumor	350	500	1300	1263	Change Th2
	350	500	2100	6275	Change Th2
	350	500	1700	3045	Change Th2
	350	500	1900	4437	Change Th2
	350	500	2000	5283	Done

TABLE 5. Training group sizes before and after applying the OCMRI augmentation method. B: Before applying augmentation. A: After applying augmentation. R: The percentage increase in images after augmentation.

Dataset	Number of training images		
	B	A	R
Dataset A	2283	20089	779.7%
Dataset B	3002	20594	586.1%
Dataset C	2146	7785	262.8%
Dataset D	173	389	124.3%

To verify the effectiveness of the OCMRI augmentation method, we trained and tested the PRCnet model on four datasets before and after applying the OCMRI augmentation method. Table 6 shows the results achieved by the PRCnet model before applying the OCMRI augmentation method, while Table 7 shows the results after applying the OCMRI augmentation method. We note that the results achieved by the PRCnet model after applying the augmentation method were better. For example, the accuracy on dataset A achieved by the PRCnet model before applying the proposed augmentation method was 85.19, and it became 92.7 after applying the proposed augmentation method. The significant improvement in the results is due to a large increase in the training data and balancing the data between the classes.

Figure 5 demonstrates the difference between the accuracy obtained by of the PRCnet model before and after applying OCMRI data augmentation for all datasets.

Figure 6 shows the confusion matrix for all datasets during the testing process before applying the OCMRI data augmentation method. Contrastingly, figure 7 shows the results after applying the proposed augmentation method. The discernible shift in the confusion matrices indicates a

TABLE 6. Results for all Datasets before applying the OCMRI augmentation method.

Dataset	Accuracy	Precision	Recall or sensitivity	F1-Score	Specificity
Dataset A	85.19	85.56	85.19	85.26	95.00
Dataset B	90.12	90.44	90.12	90.08	96.44
Dataset C	94.77	94.73	94.77	94.74	97.01
Dataset D	90.00	89.97	90.00	89.94	87.76

TABLE 7. Results for all Datasets after applying the OCMRI augmentation method.

Dataset	Accuracy	Precision	Recall or sensitivity	F1-Score	Specificity
Dataset A	92.70	92.75	92.70	92.66	97.51
Dataset B	95.37	95.40	95.37	95.38	98.49
Dataset C	96.51	96.50	96.51	96.49	98.06
Dataset D	98.00	98.10	98.00	98.01	98.77

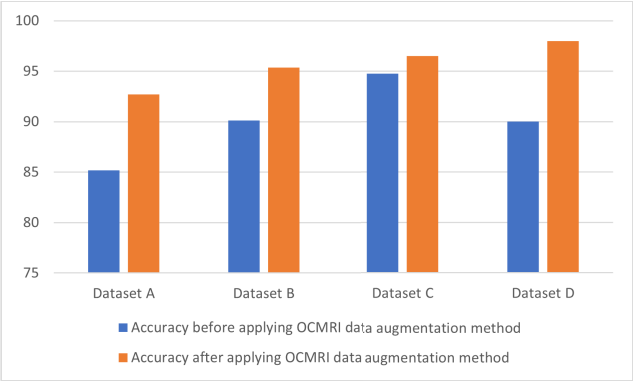

FIGURE 5. The accuracy obtained before and after applying OCMRI data augmentation for all datasets.

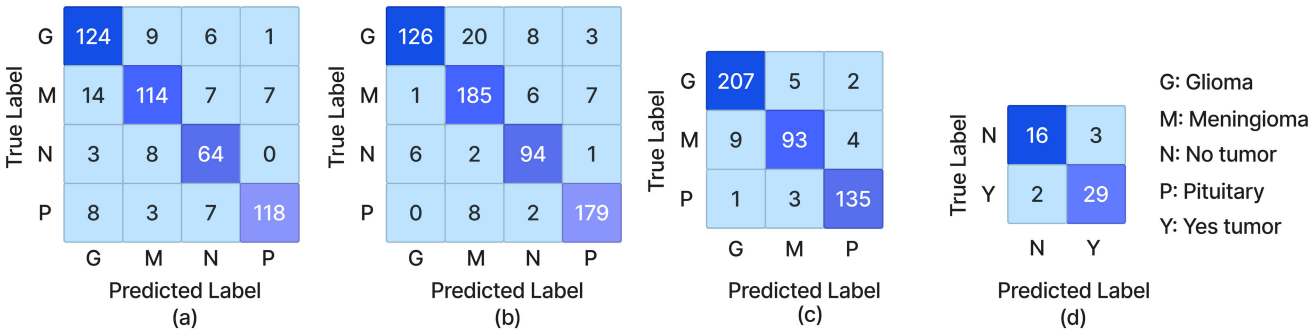

FIGURE 6. Confusion matrix before applying our augmentation method. Confusion matrix for (a) Dataset A, (b) Dataset B, (c) Dataset C, and (d) Dataset D.

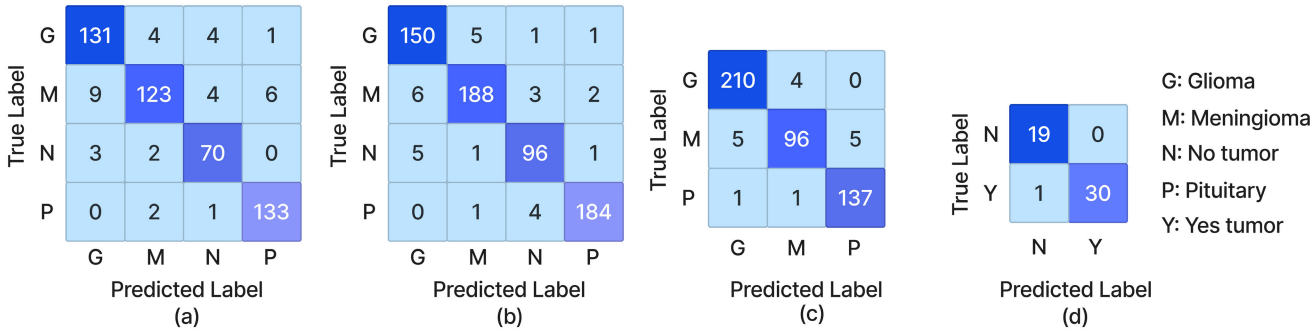

FIGURE 7. Confusion matrix after applying our augmentation method. Confusion matrix for (a) Dataset A, (b) Dataset B, (c) Dataset C, and (d) Dataset D.

substantial improvement in the performance of the PRCnet model after applying the OCMRI augmentation to the datasets. This improvement can be seen by comparing the results of Figure 6 with the results of Figure 7. For example, in Figure 6 (d), the number of cases that the model predicted as no tumor was 16 out of 19, and it predicted 29 cases as yes tumor out of 31 cases. After applying the OCMRI data augmentation approach, the number of cases predicted by the model as no tumor became 19 out of 19, and 30 cases predicted as yes tumor out of 31, as shown in Figure 7 (d).

Figure 8 shows losses and accuracy for the PRCnet model during the training and validation process on all datasets before and after applying the OCMRI augmentation method. We notice that there is a noticeable improvement and better stability after applying the OCMRI augmentation method.

In Figure 9, we showcase six illustrative examples of brain imaging images generated by applying the OCMRI augmentation method. The Figure 9 provides a visual representation of the input images selected by the algorithm, alongside the resulting merged image. As elucidated, the merged image encapsulates essential information and details from the two input images and this helps improve the training of deep learning models.

C. IMPACT OF TH1 AND TH2 ON METHOD PERFORMANCE

We did an ablation study where we changed the starting values of Th1 and Th2 and watched how that affected classification accuracy, precision, recall, F1 score, and specificity. This helped us learn more about how thresholds affect model performance. The goal was to determine the sensitivity of the OCMRI augmentation method to different threshold configurations.

Th1 variations: We tested three different values of Th1: 100, 500, and 1000. A lower Th1 value leads to merging more similar images, which can lead to redundancy. Whereas a higher Th1 value increases diversity but may reduce the structural integrity of the merged images.

Th2 variations: We also tested three different values of Th2: 1500, 2000, and 2500. A higher Th2 value allows for merging more dissimilar images, which increases diversity but leads to greater variance and produces unrealistic images.

Table 8 shows the results of the ablation study on dataset D. The ablation study examined the impact of varying starting values of the thresholds Th1 and Th2 on the performance of the OCMRI augmentation method. When Th1 was set to a lower value (e.g., 100), the resulting synthetic images were too similar, causing redundancy and limiting the model's learning ability, which led to lower accuracy (90.00%) and F1-score (90.11%). Conversely, a high Th1 value (e.g., 1000) increased diversity but often resulted in unrealistic images, reducing performance. The best balance was found with a Th1 value of 500, which led to a significant improvement with an accuracy of 98.00% and an F1-score of 98.01%.

Similarly, Th2 controls how dissimilar the merged images can be. A low Th2 value (1500) produced less diverse images, limiting the benefit of augmentation and obtaining an accuracy of 92.00%. In contrast, a higher Th2 (2500)

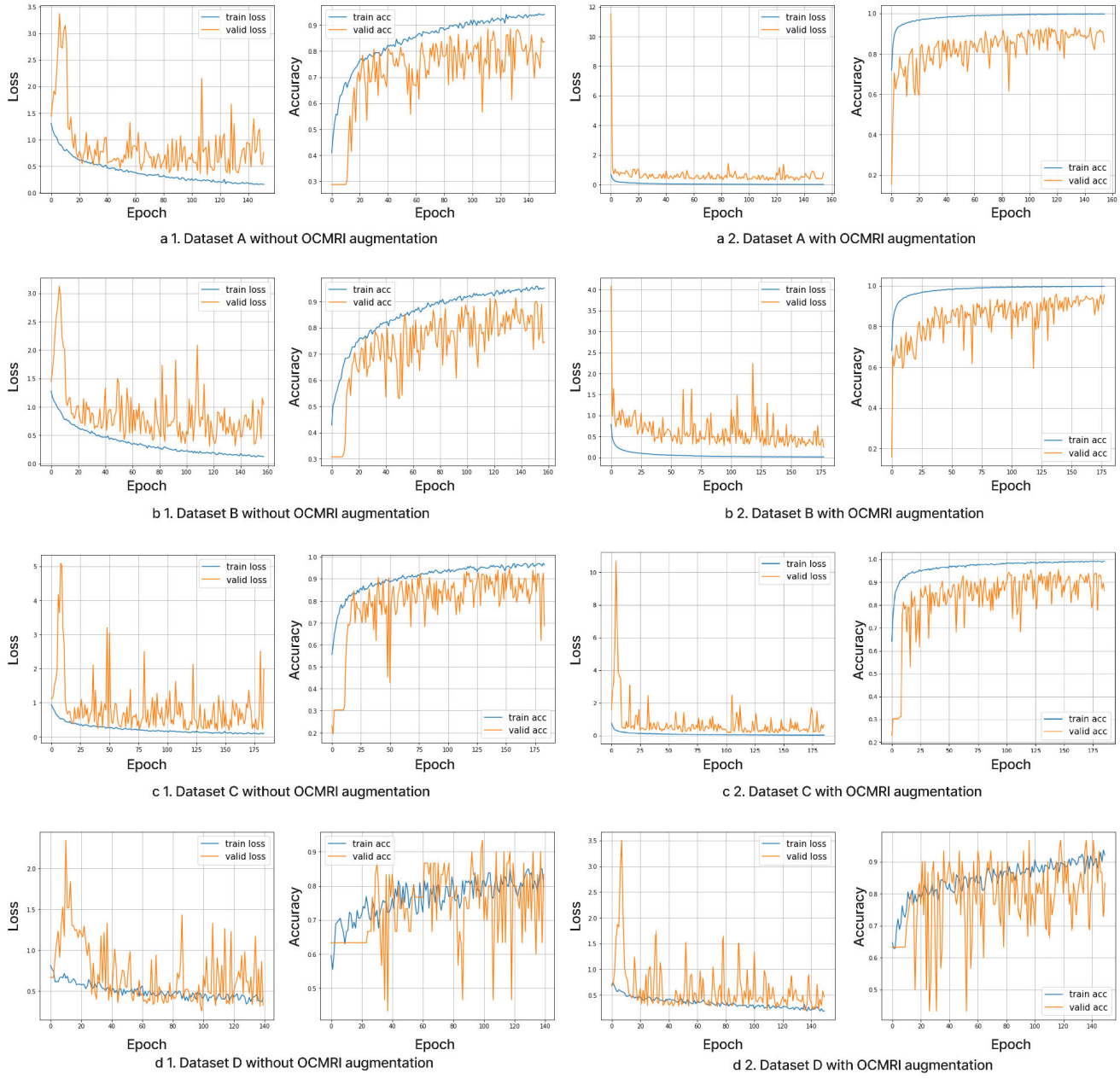

FIGURE 8. Losses and accuracy for the PRCnet model during the training and validation process on dataset A, dataset B, dataset C, and dataset D before and after applying the OCMRI augmentation method.

introduced more variation but produced noisier images, slightly decreasing accuracy (94.00%). The optimal Th2 value of 2000 generated the most effective balance between diversity and structural coherence, leading to the highest model performance, with an accuracy of 98.00%. Overall, the ablation study highlights that selecting moderate values for Th1 and Th2 (500 and 2000, respectively) leads to the most effective augmentation, significantly improving model performance. However, it is essential to note that the optimal values of Th1 and Th2 are not fixed across all datasets. The best-performing values can vary depending on the characteristics of the dataset in use, such as the number of classes, the complexity of the tumor features, and the overall distribution of the images. Therefore, while Th1 = 500 and Th2 = 2000 work best for Dataset D, they may need to be adjusted for other datasets to achieve optimal results.

D. COMPARISON

In this section, we compare the performance of the proposed OCMRI method with two other augmentation approaches:

TABLE 8. Impact started value of Th1 and Th2 on method performance. B: Number of training images for each class Before applying augmentation. A: Number of training images for each class After applying augmentation. R: The percentage increase in images after augmentation.

Th1	Th2	Classes	Number of training images			Accuracy	Precision	Recall or sensitivity	F1-Score	Specificity
			B	A	R					
100	2000	no tumor	68	210	208.82%	90.00	90.88	90.00	90.11	91.83
		with tumor	105	203	93.33%					
500	2000	no tumor	68	188	176.47%	98.00	98.10	98.00	98.01	98.77
		with tumor	105	201	91.42%					
1000	2000	no tumor	68	188	176.47%	90.00	89.97	90.00	89.94	87.76
		with tumor	105	201	91.42%					
500	1500	no tumor	68	115	69.11%	92.00	92.43	92.00	92.07	93.06
		with tumor	105	114	8.57%					
500	2500	no tumor	68	423	522.05%	94.00	94.13	94.00	94.03	94.29
		with tumor	105	435	314.28%					

the Random Image Cropping and Patching (RICAP) method and the Deep Convolutional Generative Adversarial Network (DC-GAN) model.

The RICAP introduced by Takahashi et al. [46]. Diversify training data and enhances the performance of CNN models. RICAP involves randomly cropping and patching four images to generate a new training image. Random cropping involves extracting random regions from original images, while patching combines random patches to create new synthetic images. The goal is to improve the generalization and robustness of deep CNN models by diversifying the training data [46]. Similarly, DC-GAN models are commonly used to generate synthetic images. It has been widely adopted in medical imaging tasks, particularly for generating synthetic data to augment small datasets, such as MRI or CT scans. Brain MRI image analysis, for example, can generate additional synthetic images of tumors, helping to address the common problem of data scarcity [67], [68].

Table 9 and figure 10 present a comparative analysis of the accuracy, precision, recall, F1-score, and specificity for the PRCnet model when augmented with OCMRI, RICAP, and DC-GAN. The PRCnet model, enhanced with OCMRI, consistently outperformed both RICAP and DC-GAN across all datasets. Specifically, the OCMRI method achieved accuracy improvements of 92.7% on Dataset A, 95.37% on Dataset B, 96.51% on Dataset C, and 98% on Dataset D. In contrast, RICAP achieved accuracy of 86.61%, 91.98%, 94.99%, and 90%. In comparison, DC-GAN achieved 83.37%, 85.49%, 91.94%, and 96% for Datasets A, B, C, and D, respectively. The observed difference in results can be attributed to the RICAP method's reliance on cutting and rearranging four images to create a new image. In the context of medical images, this approach may generate new images lacking tumour-related information. Similarly, DC-GAN struggled with preserving fine details in tumour regions, leading to lower classification accuracy. Conversely, the OCMRI method creates new images by merging two images, ensuring that the resulting image retains information from both images, thus contributing to its superior performance.

TABLE 9. The results of the OCMRI method comparison with the RICAP approach.

Dataset	Method	Accuracy	Precision	Recall or sensitivity	F1-Score	Specificity
Dataset A	RICAP	86.61	87.47	86.61	86.68	95.18
	OCMRI	**92.70**	**92.75**	**92.70**	**92.66**	**97.51**
	DC-GAN	83.37	85.87	83.37	83.31	93.70
Dataset B	RICAP	91.98	92.15	91.98	92.01	97.30
	OCMRI	**95.37**	**95.40**	**95.37**	**95.38**	**98.49**
	DC-GAN	85.49	85.92	85.49	85.39	94.88
Dataset C	RICAP	94.99	94.95	94.99	94.96	97.39
	OCMRI	**96.51**	**96.50**	**96.51**	**96.49**	**98.06**
	DC-GAN	91.94	91.87	91.94	91.86	96.01
Dataset D	RICAP	90.00	90.88	90.00	90.11	91.83
	OCMRI	**98.00**	**98.10**	**98.00**	**98.01**	**98.77**
	DC-GAN	96.00	96.00	96.00	96.00	95.51

V. DISCUSSION

In this paper, we proposed a new method for augmenting brain MRI images called OCMRI. The proposed method combines MRI images of the brain to create new images. Every two MR images of the same class are selected by taking the odd columns from the first image and the even columns from the second image to produce new pictures. The mean squared error between the two images must also be greater than the first threshold value and smaller than the second threshold value. Controlling the value of the first and second thresholds led to controlling the number of images to be produced, as well as solving the problem of class imbalance.

Table 4 details applying the OCMRI augmentation method to dataset A. We notice that as the value of the Th2 increases, the number of images produced by the algorithm increases due to an acceptable rise in the ratio of MSE between image 1 and image 2. We also note that repeatedly and auto-adjusting the value of the Th2 to achieve a balance between the classes contributed to solving the data imbalance problem.

We conducted experiments to evaluate the effectiveness of the proposed Oriented Combination MRI (OCMRI) augmentation method on four publicly available datasets. The

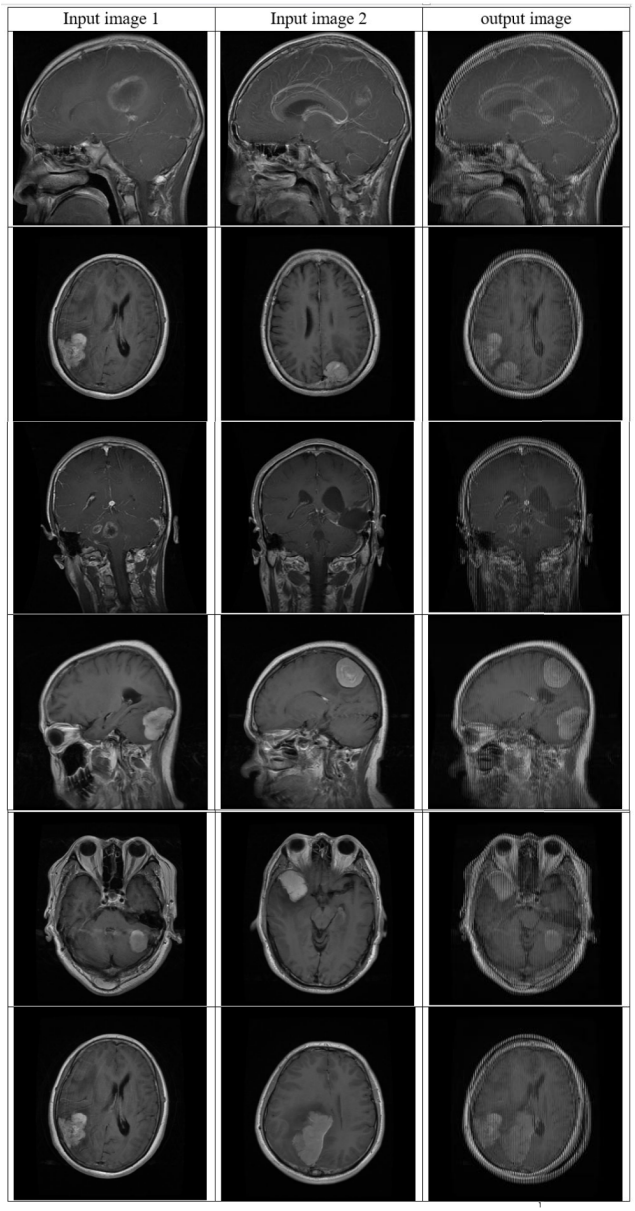

FIGURE 9. Examples of brain imaging images produced by OCMRI augmentation.

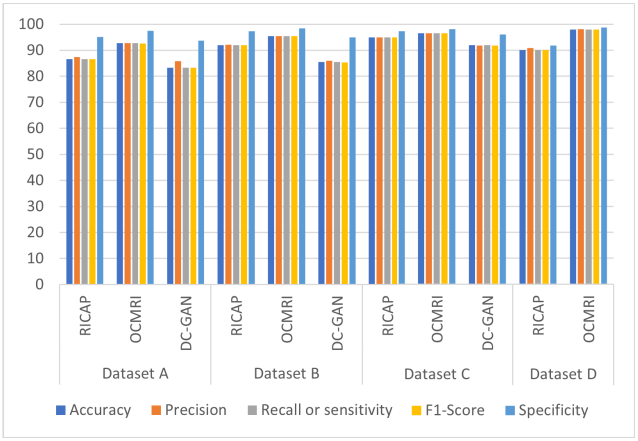

FIGURE 10. The results of the OCMRI method comparison with the RICAP approach.

experimental results demonstrated a substantial improvement in efficiency when the PRCnet model was trained and tested before and after applying the OCMRI augmentation method. The significant enhancement in results can be attributed to the considerable expansion of the training data and the successful mitigation of class imbalance issues, which, in turn,

Furthermore, the fusion of images through the OCMRI method helped the model recognize and distinguish better. The generated images, containing two tumors with distinct locations and shapes, allowed the model to better comprehend and differentiate between intricate scenarios. A meticulous analysis of the data pre- and post-application of the OCMRI augmentation method revealed that all correctly predicted images remained correctly predicted after augmentation. Moreover, specific images that were initially mispredicted were subsequently correctly classified following the OCMRI augmentation.

However, The computational complexity of the proposed methodology represents a notable limitation arising from the iterative computation of the Mean Squared Error (MSE) between each image and all other images within the dataset. This iterative process involves exhaustive calculations of MSE for each image in comparison with every other image present in the dataset. Consequently, the method may incur increased computational demands, potentially leading to longer processing times and heightened resource utilization.

VI. CONCLUSION

The brain's available MRI images are insufficient to train deep learning models. For this reason, deep learning models that need large amounts of data for classification or segmentation may suffer from the problem of overfitting. To solve this problem, data augmentation methods are used to produce additional medical images to the original data. In this work, we proposed a new method for augmenting brain MRI data which we named Oriented Combination MRI (OCMRI). The proposed method combines the two MRI images by taking the odd columns from the first image and the even columns from the second image to produce new images. The two images are within the same class, and the mean squared error between the two images is also required to be greater than the value of the first threshold and smaller than the value of the second threshold. The value of the first and second thresholds is variable and is specified by the user at the beginning. Adjusting the thresholds controls the number of generated images. Therefore, the threshold value is controlled in a way that controls the number of images produced for each class, which contributes to solving the data imbalance problem. Experiments were carried out on four datasets. The proposed method proved its efficiency by improving the performance of the PRCnet model. The accuracy of the PRCnet model was 85.19% for dataset A, 90.12% for dataset B, 94.77% for

dataset C, and 90% for dataset D. After adding synthetic data, The accuracy increased to 92.7% for dataset A, 95.37% for dataset B, 96.51% for dataset C, and 98% for dataset D.

One limitation of the proposed methodology lies in its computational complexity, stemming from the iterative process of computing the Mean Squared Error (MSE) between each image and all other images within the dataset. To address this issue in future research, efforts could be directed towards optimizing the computational demands by adopting selective pixel sampling strategies. For instance, an approach involving the selection of pixels exclusively from either even or odd column rows within the image could be explored to alleviate the computational burden. Additionally, to enhance the method's overall performance, an avenue for improvement involves incorporating user-defined parameters. Specifically, allowing users to specify the desired number of generated images would empower the algorithm to dynamically adjust the first and second threshold values, thereby optimizing the augmentation process based on user input.

REFERENCES

[1] A. C. Kara and F. Hardalaç, "Detection and classification of knee injuries from MR images using the MRNet dataset with progressively operating deep learning methods," *Mach. Learn. Knowl. Extraction*, vol. 3, no. 4, pp. 1009–1029, Dec. 2021.

[2] Y. Chen, X.-H. Yang, Z. Wei, A. A. Heidari, N. Zheng, Z. Li, H. Chen, H. Hu, Q. Zhou, and Q. Guan, "Generative adversarial networks in medical image augmentation: A review," *Comput. Biol. Med.*, vol. 144, Mar. 2022, Art. no. 105382.

[3] N. F. Aurna, M. A. Yousuf, K. A. Taher, A. K. M. Azad, and M. A. Moni, "A classification of MRI brain tumor based on two stage feature level ensemble of deep CNN models," *Comput. Biol. Med.*, vol. 146, Jul. 2022, Art. no. 105539.

[4] R. A. Zeineldin, M. E. Karar, Z. Elshaer, J. Coburger, C. R. Wirtz, O. Burgert, and F. Mathis-Ullrich, "Explainability of deep neural networks for MRI analysis of brain tumors," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 17, no. 9, pp. 1673–1683, Apr. 2022.

[5] J. Nodirov, A. B. Abdusalomov, and T. K. Whangbo, "Attention 3D U-Net with multiple skip connections for segmentation of brain tumor images," *Sensors*, vol. 22, no. 17, p. 6501, Aug. 2022.

[6] D. Kaur, S. Singh, W. Mansoor, Y. Kumar, S. Verma, S. Dash, and A. Koul, "Computational intelligence and metaheuristic techniques for brain tumor detection through IoMT-enabled MRI devices," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–20, Jan. 2022.

[7] S. Das, G. K. Nayak, L. Saba, M. Kalra, J. S. Suri, and S. Saxena, "An artificial intelligence framework and its bias for brain tumor segmentation: A narrative review," *Comput. Biol. Med.*, vol. 143, Apr. 2022, Art. no. 105273.

[8] E. Dandıl and S. Karaca, "Detection of pseudo brain tumors via stacked LSTM neural networks using MR spectroscopy signals," *Biocybern. Biomed. Eng.*, vol. 41, no. 1, pp. 173–195, Jan. 2021.

[9] D. Mukherkjee, P. Saha, D. Kaplun, A. Sinitca, and R. Sarkar, "Brain tumor image generation using an aggregation of GAN models with style transfer," *Sci. Rep.*, vol. 12, no. 1, pp. 1–16, Jun. 2022.

[10] T. Hossain, F. S. Shishir, M. Ashraf, M. A. A. Nasim, and F. M. Shah, "Brain tumor detection using convolutional neural network," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–6.

[11] J. Walsh, A. Othmani, M. Jain, and S. Dev, "Using U-Net network for efficient brain tumor segmentation in MRI images," *Healthcare Anal.*, vol. 2, Nov. 2022, Art. no. 100098.

[12] A. K. Anaraki, M. Ayati, and F. Kazemi, "Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 63–74, Jan. 2019.

[13] B. Ahmad, J. Sun, Q. You, V. Palade, and Z. Mao, "Brain tumor classification using a combination of variational autoencoders and generative adversarial networks," *Biomedicines*, vol. 10, no. 2, p. 223, Jan. 2022.

[14] E. U. Haq, H. Jianjun, K. Li, H. U. Haq, and T. Zhang, "An MRI-based deep learning approach for efficient classification of brain tumors," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 6, pp. 6697–6718, Jun. 2023.

[15] A. S. Farhan, M. Khalid, and U. Manzoor, "XAI-MRI: An ensemble dual-modality approach for 3D brain tumor segmentation using magnetic resonance imaging," *bioRxiv*, vol. 2024, p. 11, Nov. 2024.

[16] D. S. Wankhede and R. Selvarani, "Dynamic architecture based deep learning approach for glioblastoma brain tumor survival prediction," *Neurosci. Informat.*, vol. 2, no. 4, Dec. 2022, Art. no. 100062.

[17] A. Shukla, R. Tiwari, and S. Tiwari, "Alzheimer's disease detection from fused PET and MRI modalities using an ensemble classifier," *Mach. Learn. Knowl. Extraction*, vol. 5, no. 2, pp. 512–538, May 2023.

[18] L. Fang and X. Wang, "Brain tumor segmentation based on the dual-path network of multi-modal MRI images," *Pattern Recognit.*, vol. 124, Apr. 2022, Art. no. 108434.

[19] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69215–69225, 2019.

[20] N. Dey, V. Rajinikanth, F. Shi, J. M. R. S. Tavares, L. Moraru, K. A. Karthik, H. Lin, K. Kamalanand, and C. Emmanuel, "Social-group-optimization based tumor evaluation tool for clinical brain MRI of flair/diffusion-weighted modality," *Biocybern. Biomed. Eng.*, vol. 39, no. 3, pp. 843–856, Jul. 2019.

[21] T. Yang, J. Song, and L. Li, "A deep learning model integrating SK-TPCNN and random forests for brain tumor segmentation in MRI," *Biocybern. Biomed. Eng.*, vol. 39, no. 3, pp. 613–623, Jul. 2019.

[22] P. Sriramakrishnan, T. Kalaiselvi, and R. Rajeswaran, "Modified local ternary patterns technique for brain tumour segmentation and volume estimation from MRI multi-sequence scans with GPU CUDA machine," *Biocybern. Biomed. Eng.*, vol. 39, no. 2, pp. 470–487, Apr. 2019.

[23] A. Veeramuthu, S. Meenakshi, G. Mathivanan, K. Kotecha, J. R. Saini, V. Vijayakumar, and V. Subramaniyaswamy, "MRI brain tumor image classification using a combined feature and image-based classifier," *Frontiers Psychol.*, vol. 13, Mar. 2022, Art. no. 848784.

[24] A. Stadlbauer, F. Marhold, S. Oberndorfer, G. Heinz, M. Buchfelder, T. M. Kinfe, and A. Meyer-Bäse, "Radiophysiomics: Brain tumors classification by machine learning and physiological MRI data," *Cancers*, vol. 14, no. 10, p. 2363, May 2022.

[25] P. Li, W. Wu, L. Liu, F. M. Serry, J. Wang, and H. Han, "Automatic brain tumor segmentation from multiparametric MRI based on cascaded 3D U-Net and 3D U-Net++," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, Art. no. 103979.

[26] S. A. Y. Al-Galal, I. F. T. Alshaikhli, and M. M. Abdulrazzaq, "MRI brain tumor medical images analysis using deep learning techniques: A systematic review," *Health Technol.*, vol. 11, no. 2, pp. 267–282, Mar. 2021.

[27] X. Guan, G. Yang, J. Ye, W. Yang, X. Xu, W. Jiang, and X. Lai, "3D AGSE-VNet: An automatic brain tumor MRI data segmentation framework," *BMC Med. Imag.*, vol. 22, no. 1, pp. 1–18, Dec. 2022.

[28] S. Devi, M. N. Sahoo, and S. Bakshi, "A novel privacy-supporting 2-class classification technique for brain MRI images," *Biocybern. Biomed. Eng.*, vol. 40, no. 3, pp. 1022–1035, Jul. 2020.

[29] K. Seddiki, P. Saudemont, F. Precioso, N. Ogrinc, M. Wisztorski, M. Salzet, I. Fournier, and A. Droit, "Cumulative learning enables convolutional neural network representations for small mass spectrometry data classification," *Nature Commun.*, vol. 11, no. 1, pp. 1–11, Nov. 2020.

[30] M. Aamir, Z. Rahman, W. A. Abro, U. A. Bhatti, Z. A. Dayo, and M. Ishfaq, "Brain tumor classification utilizing deep features derived from high-quality regions in MRI images," *Biomed. Signal Process. Control*, vol. 85, Aug. 2023, Art. no. 104988.

[31] Y. Guan, M. Aamir, Z. Rahman, A. Ali, W. A. Abro, Z. A. Dayo, M. S. Bhutta, and Z. Hu, "A framework for efficient brain tumor classification using MRI images," *Math. Biosci. Eng.*, vol. 18, no. 5, pp. 5790–5815, 2021.

[32] J. Zhang, L. Zhao, J. Zeng, P. Qin, Y. Wang, and X. Yu, "Deep MRI glioma segmentation via multiple guidances and hybrid enhanced-gradient cross-entropy loss," *Expert Syst. Appl.*, vol. 196, Jun. 2022, Art. no. 116608.

[33] M. Sun, W. Zou, N. Hu, J. Wang, and Z. Chi, "Iterative brain tumor retrieval for MR images based on user's intention model," *Pattern Recognit.*, vol. 127, Jul. 2022, Art. no. 108650.
[34] Z. Hussain, F. J. G. Fuentes-Guerra, D. Yi, and D. L. Rubin, "Differential data augmentation techniques for medical imaging classification tasks," in *AMIA Annu. Symp. Proc.*, vol. 2017, Jan. 2017, pp. 979–984.
[35] B. Abdollahi, N. Tomita, and S. Hassanpour, "Data augmentation in training deep learning models for medical image analysis," in *Deep Learners and Deep Learner Descriptors for Medical Applications*. Cham, Switzerland: Springer, 2020, pp. 167–180.
[36] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, and A. Haworth, "A review of medical image data augmentation techniques for deep learning applications," *J. Med. Imag. Radiat. Oncol.*, vol. 65, no. 5, pp. 545–563, Aug. 2021.
[37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
[38] A. Galdrán, G. Carneiro, and M. A. Ballester, "Balanced-MixUp for highly imbalanced medical image classification," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*. Cham, Switzerland: Springer, Jan. 2021, pp. 323–333.
[39] L. Gao, L. Zhang, C. Liu, and S. Wu, "Handling imbalanced medical image data: A deep-learning-based one-class classification approach," *Artif. Intell. Med.*, vol. 108, Aug. 2020, Art. no. 101935.
[40] T. Agustin, E. Utami, and H. A. Fatta, "Implementation of data augmentation to improve performance CNN method for detecting diabetic retinopathy," in *Proc. 3rd Int. Conf. Inf. Commun. Technol. (ICOIACT)*, Nov. 2020, pp. 83–88.
[41] A. S. Farhan, M. Khalid, and U. Manzoor, "PRCNet: An efficient model for automatic detection of brain tumor in MRI images," *bioRxiv*, vol. 9, pp. 1–33, Oct. 2023.
[42] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019.
[43] C. Khosla and B. S. Saini, "Enhancing performance of deep learning models with different data augmentation techniques: A survey," in *Proc. Int. Conf. Intell. Eng. Manage. (ICIEM)*, Jun. 2020, pp. 79–85.
[44] J. Shijie, W. Ping, J. Peiyi, and H. Siping, "Research on data augmentation for image classification based on convolution neural networks," in *Proc. Chin. Autom. Congr. (CAC)*, Oct. 2017, pp. 4165–4170.
[45] F. J. Moreno-Barea, F. Strazzera, J. M. Jerez, D. Urda, and L. Franco, "Forward noise adjustment scheme for data augmentation," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Nov. 2018, pp. 728–734.
[46] R. Takahashi, T. Matsubara, and K. Uehara, "Data augmentation using random image cropping and patching for deep CNNs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 9, pp. 2917–2931, Sep. 2020.
[47] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random erasing data augmentation," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, Apr. 2020, pp. 13001–13008.
[48] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1–15.
[49] C. Bowles, L. Chen, R. Guerrero, P. Bentley, R. Gunn, A. Hammers, D. A. Dickie, M. Valdés Hernández, J. Wardlaw, and D. Rueckert, "GAN augmentation: Augmenting training data using generative adversarial networks," 2018, *arXiv:1810.10863*.
[50] J. Wu, Z. Huang, J. Thoma, U. D. Acharya, and L. Van Gool, "Wasserstein divergence for GANs," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Jan. 2018, pp. 673–688.
[51] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.
[52] Q. Li, Z. Yu, Y. Wang, and H. Zheng, "TumorGAN: A multi-modal data augmentation framework for brain tumor segmentation," *Sensors*, vol. 20, no. 15, p. 4203, Jul. 2020.
[53] C. Han, L. Rundo, R. Araki, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi, "Infinite brain MR images: PGGAN-based data augmentation for tumor detection," in *Neural Approaches to Dynamics of Signal Exchanges*. Cham, Switzerland: Springer, 2020, pp. 291–303.
[54] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based data augmentation for improved liver Lesion classification," Tech. Rep., 2018.
[55] F. Konidaris, T. Tagaris, M. Sdraka, and A. Stafylopatis, "Generative adversarial networks as an advanced data augmentation technique for MRI data," in *Proc. 14th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2019, pp. 48–59.
[56] S. P. Sharma and N. Sampath, "Data augmentation for brain tumor segmentation using MONAI framework," in *Proc. 2nd Int. Conf. Intell. Technol. (CONIT)*, Jun. 2022, pp. 1–8.
[57] A. Panda, "Brain MRI," Tech. Rep., 2021.
[58] M. M. Sherif, "Brain-tumor-dataset," Tech. Rep., 2020.
[59] J. Cheng, "Brain tumor dataset, figshare dataset," 2017, doi: 10.6084/m9.figshare.1512427.v5.
[60] N. Chakrabarty, "Brain MRI images for brain tumor detection," Tech. Rep., 2019.
[61] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0140381.
[62] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study," *J. Comput. Commun.*, vol. 7, no. 3, pp. 8–18, 2019.
[63] M. Gong, "A novel performance measure for machine learning classification," *Int. J. Manag. Inf. Technol.*, vol. 13, no. 1, pp. 11–19, Feb. 2021.
[64] A. Al-Kubaisi and N. N. Khamiss, "A transfer learning approach for lumbar spine disc state classification," *Electronics*, vol. 11, no. 1, p. 85, Dec. 2021.
[65] L. Alzubaidi, M. Al-Amidie, A. Al-Asadi, A. J. Humaidi, O. Al-Shamma, M. A. Fadhel, J. Zhang, J. Santamaría, and Y. Duan, "Novel transfer learning approach for medical imaging with limited labeled data," *Cancers*, vol. 13, no. 7, p. 1590, Mar. 2021.
[66] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *Int. J. Data Mining Knowl. Manage. Process*, vol. 5, no. 2, pp. 1–11, Mar. 2015.
[67] D. Mourad and K. O. Oseni, "Synthetic brain images: Bridging the gap in brain mapping with generative adversarial model," 2024, *arXiv:2404.08703*.
[68] S. S. Alrumiah, N. Alrebdi, and D. M. Ibrahim, "Augmenting healthy brain magnetic resonance images using generative adversarial networks," *PeerJ Comput. Sci.*, vol. 9, p. e1318, Apr. 2023.

AHMEED SULIMAN FARHAN (Member, IEEE) received the M.S. degree in computer science from the Computer Science and Information Technology, University of Anbar, Iraq. He is currently pursuing the Ph.D. degree with the University of Hull, U.K. His research interests include image processing and artificial intelligence.

MUHAMMAD KHALID received the Ph.D. degree in computer science from Northumbria University, Newcastle upon Tyne, U.K. He is currently a Lecturer with the Department of Computer Science and Technology, University of Hull, U.K. Earlier, he was a Research Fellow with the University of Lincoln, U.K. His research interests include machine learning, autonomous systems, safety in robotics, EV charging and scheduling, the Internet of Things, wireless sensor networks, and autonomous valet parking.

UMAR MANZOOR has more than 12 years of experience in the field of artificial intelligence (AI) and machine learning (ML). He has developed a range of state-of-the-art AI/ML algorithms and applied them across diverse fields from healthcare to robotics. He is currently a full-time Faculty Member with the Department of Computer Science, Aston University.