

## RESEARCH ARTICLE

# Improved Classification of Different Brain Tumors in MRI Scans Using Patterned-GridMask

Ji-Hyeon Lee<sup>1</sup>, Jung-Woo Chae<sup>1,2</sup>, and Hyun-Chong Cho<sup>1,2</sup>, (Member, IEEE)

<sup>1</sup>Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, Chuncheon-si 24341, Republic of Korea

<sup>2</sup>Department of Electronics Engineering, Kangwon National University, Chuncheon-si 24341, Republic of Korea

Corresponding authors: Jung-Woo Chae (cowjddn94@kangwon.ac.kr) and Hyun-Chong Cho (hyuncho@kangwon.ac.kr)

This work was supported in part by the “Regional Innovation Strategy (RIS)” through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) under Grant 2022RIS-005; in part by the Basic Science Research Program through NRF funded by the MOE under Grant 2022R111A3053872; and in part by Kangwon National University, funded by the National University Development Project, in 2023.


This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committees of Nanfang Hospital and General Hospital, Tianjin Medical University.

**ABSTRACT** Continuous advancements in deep learning are affecting various research areas, especially research on applications in the medical sector. A computer-aided diagnosis system that utilizes deep learning is used for classifying and detecting brain tumors in magnetic resonance imaging. Regarding brain tumors, the main diagnostic indicators are patient symptoms and outcomes of magnetic resonance imaging. Frequent changes in the symptoms of these tumors have raised serious concerns about potential misdiagnoses. Implementing computer-aided diagnosis systems can support diagnostic methods that rely on the visual assessments of physicians, potentially reducing misdiagnosis rates. In this study, we propose an enhanced computer-aided diagnosis algorithm that is optimized for brain tumor classification. We removed noise from the magnetic resonance imaging results by applying Gaussian filters, and we employed GridMask to improve the generalization performances of the deep learning models. Then, we applied Patterned-GridMask, which is a method we proposed to reduce the issue of brain tumors being obscured by standard GridMask. Under the application of Patterned-GridMask, a performance improvement of up to 6% was demonstrated across the four deep learning models used in the experiments: ViT-B/16, MaxViT-B, TresNet-M, and EfficientNetV2-M, with the highest performance being represented by an accuracy and F1-score of 97.74% and 97.75%, respectively. Using the proposed computer-aided diagnosis system, improved diagnosis results can be obtained, thereby resulting in more accurate rates of early detection, better patient outcomes, and more appropriate treatment selection.

**INDEX TERMS** Brain tumor, classification, computer-aided diagnosis system, deep learning, magnetic resonance imaging, Patterned-GridMask.

## I. INTRODUCTION

The brain is a vital organ that controls the central nervous system. It consists of six structures: the cerebrum, cerebellum, spinal cord, medulla oblongata, midbrain, and diencephalon. Each part of the brain regulates body movement, behavior, and stability. Additionally, these parts are responsible for

The associate editor coordinating the review of this manuscript and approving it for publication was R. K. Tripathy .

various functions such as cognitive reasoning, emotional regulation, memory, and learning [1].

Brain tumors are tumors that arise within the cranial cavity. Unlike other types of cancers, brain tumors are not classified by stage because they rarely metastasize to other organs. Since these tumors grow within the confined spaces of skulls, an increase in tumor volume, which leads to elevated pressure, is considered the most critical factor contributing to worsening symptoms. Such elevated intracranial pressure

initially manifests as headaches and vomiting. Further pressure build-up that compresses the surrounding nerves can result in paralysis and impaired physical activity [2], clearly emphasizing the life-threatening nature of brain tumors and the importance of early diagnosis and prompt treatment.

Progress in medical technologies has enabled the extension of the survival durations for various types of cancer through various therapeutic interventions. However, in cases where cancer persists over a prolonged period without being resolved over a prolonged period, it frequently metastasizes to the brain.

Additionally, as people age, they face an increasing risk of exposure to brain tumors, which tend to occur at older ages. According to the 2020 Cancer Registration Statistics from the Korean Statistical Office, the incidence of brain tumors has sharply increased from 45–49, with the highest incidence observed in the age group of 60–64 [3]. Fig. 1 depicts the number of brain tumor cases by age. Furthermore, cancer statistics from the National Brain Tumor Foundation indicate that by 2023, approximately 18,990 deaths will be attributable to brain tumors. Significantly, brain tumor-induced deaths are projected to be the 10th leading cause of cancer-related deaths in all age groups [4].

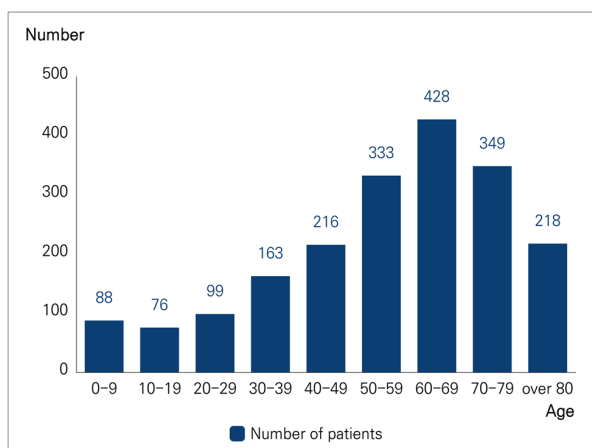


FIGURE 1. Incidence of brain tumor by age group in Korea.

Advancements in deep learning technology have led to significant increases in its application in medical imaging [5]. Computer-aided diagnosis (CADx) systems, which exhibit high accuracy and auxiliary capabilities, have greatly aided in medical diagnoses. Additionally, they have enabled more precise and early diagnoses of brain tumors, supporting preventions of tumor growth and metastasis, preservation of brain function, and avoidance of treatment-related complications.

Research on brain tumor classification using deep learning has been actively conducted. In their brain tumor classification study, Zahid et al. [6] proposed a CNN-based brain tumor classification model for classifying glioma, meningioma, pituitary tumors, or the absence of a tumor. They employed Contrast Limited Adaptive Histogram Equalization (CLAHE) in the preprocessing stage to enhance image

quality. While the reliability of the model was enhanced through the implementation of 5-fold cross-validation. Alsaif et al. [7] developed a brain tumor classification deep learning network with an improved ResNet50 architecture, incorporating data augmentation techniques such as flipping, rotation, and translation. Abdullah et al. [8] proposed an enhanced Fine-tuned ResNet50 and U-Net model based on CNN for tumor segmentation in brain tumor MRI images. Their model classifies between normal and abnormal, and for abnormal cases, segments the tumor, achieving comprehensive brain tumor detection based on high accuracy. Kumar et al. [9] propose an automatic glioma segmentation model based on a deep neural network (DNN) using the Double ConvNet (DCN) architecture in their brain tumor segmentation research.

Beyond brain tumors, Murtaza et al. [10] developed the Biopsy Microscopic Image Cancer Network model, which utilizes feature reduction schemes to classify breast cancer into eight different subtypes. Göreke et al. [11] proposed a computer-aided diagnostic system using a multi-layer deep neural network that classifies benign and malignant thyroid nodules by extracting the region of interest of the nodules from the radiographic images. Gopinath et al. [12] introduced a lung-cancer classification model using deep learning. They applied noise reduction through a filter application and performed contrast enhancement to enhance the ability of the deep learning model to detect abnormalities. Additionally, they combined a generative adversarial network and mask region-based convolutional neural network techniques. These examples demonstrated the effective utilization of deep learning based CADx systems in the medical field.

In this study, we propose a deep learning model-based CADx system for classifying the magnetic resonance imaging (MRI) results of brain tumors. To the best of our knowledge, this is the first study to propose an CADx system that has been specifically optimized for brain tumor classification. We classified gliomas, meningiomas, and normal and pituitary tissues. Gaussian filters were applied for noise reduction in MRI. Experiments were conducted with various combinations of kernel size ( $k$ ) and standard deviation ( $\sigma$ ) to adjust the blurring intensity. Then, the generalization performance and robustness of the deep learning model were enhanced using Patterned-GridMask, which is a method we proposed to overcome the limitations of the standard GridMask technique. Patterned-GridMask applies a mosaic-patterned mask to rectangular regions at 1-pixel intervals. Regarding these two enhancement techniques, experiments were conducted with various combinations of parameters, including  $d_1$ ,  $d_2$ , and ratio, which determined the size and distribution of the rectangular areas used for information removal. Focusing on these methods, we propose a brain tumor classification algorithm that employs this novel enhancement technique to improve the classification performances of different deep learning models.

The rest of the manuscript is organized as follows. Section II describes the dataset, performance enhancement

techniques, and deep learning models used in the proposed study. Then, section III presents the results of classifying different tumors in brain tumor MRI using the proposed technique. Finally, section IV provides the concluding remarks.

## II. MATERIALS AND METHODS

### A. DATASET CONFIGURATION

In this study, the Brain Tumor MRI Dataset provided by Nickparvar [13] was utilized. This dataset is a combination of three datasets: the “Brain Tumor” dataset provided by Cheng [14] that is publicly available on Figshare, the “Brain Tumor Classification (MRI)” dataset provided by Bhuvaji [15] on Kaggle, and the “Br35H: Brain Tumor Detection 2020” dataset provided by Hamada [16]. It includes glioma, meningioma, normal tissue, and pituitary tissue classes for use in brain tumor MRI scans. Overall, the three datasets contained 7,023 images, which comprised 1,621, 1,645, 2,000, and 1,757 images of the glioma, meningioma, normal tissue, and pituitary tissue classes, respectively.

This merged dataset, which brings together three separate sources, was carefully preprocessed prior to its use in our study to reduce discrepancies among the datasets and promote a more stable learning process. First, images with resolutions below  $224 \times 224$  px, which might not contain sufficient information for effectively training deep learning models, were removed. Second, all the images were converted into lossless and uncompressed BMP formats to minimize potential information loss during dataset refinement. Third, channel standardization was applied to all the images, and sizes of the images was standardized to a resolution of  $224 \times 224$  px. In this process, image resizing was performed using the area interpolation method to maximally preserve information. After implementing these steps, the distribution of the images was adjusted to achieve a balance among all the classes.

Using the class with the lowest number of images (1,200 normal tissue images) as a reference, 1,200 images were randomly selected from the dataset for each of the other classes. Consequently, each class in the constructed dataset contained 1,200 images. This dataset was randomly divided to create 3-fold sets with a 5:5 ratio for training and testing. Validation data within the training set were used in an 8:2 ratio. The final composition of the dataset that was used in this study is presented in Table 1.

### B. ENHANCEMNET TECHNIQUES

#### 1) NOISE REMOVAL FROM MRI SCANS OF BRAIN TUMORS

Since MRI utilizes a powerful magnetic field to generate images, nonuniform magnetic fields or external noise occurring around the human body during imaging can introduce unnecessary noise into images [17]. The presence of such noise in an MRI scan used for training deep learning models can potentially distort the features of objects or even impact

TABLE 1. Dataset composition used in the proposed study.

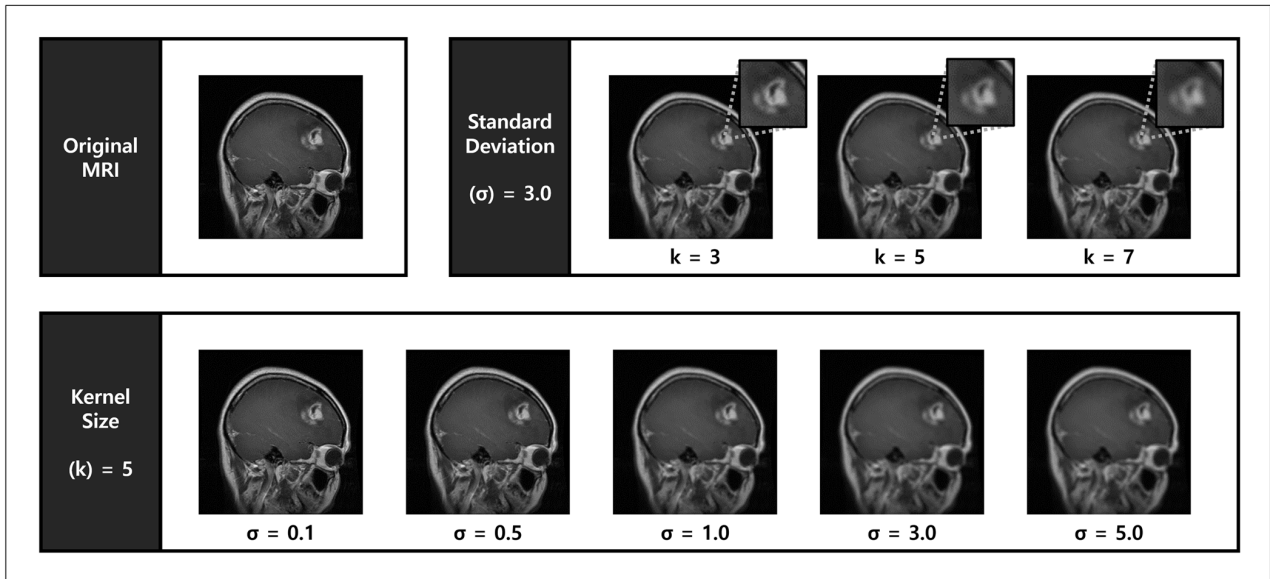
Dataset	Train	Validation	Test
Glioma	480	120	600
Meningioma	480	120	600
Normal	480	120	600
Pituitary	480	120	600
<b>Total</b>	1,920	480	2,400

the training results. This concern can be addressed by removing any noise generated during medical image acquisition, thereby assisting deep learning models in capturing object-specific features [18].

We applied a Gaussian filter to remove noise from MRI scans of brain tumors. When a Gaussian filter is applied to an image, any existing noise is reduced by minimizing the differences between each pixel and its neighboring pixels. Additionally, the overall image is smoothed, preventing the deep learning model from focusing excessively on specific areas during training. A Gaussian filter is based on a normal distribution and can be adjusted using parameters such as kernel size ( $k$ ) and standard deviation ( $\sigma$ ). The kernel size ( $k$ ) determines the filter dimensions, with a larger kernel size indicating that the filter accounts for more surrounding pixels. Regarding the standard deviation ( $\sigma$ ), it determines the weight distribution of the filter, with a larger standard deviation indicating a wider Gaussian distribution, which results in a stronger blurring effect [19]. In this study, we conducted a performance comparison under various conditions. We applied a combination of kernel ( $k$ ) values of 3, 5, and 7 along with standard deviations ( $\sigma$ ) ranging from 0.1 to 5.0 to obtain the optimal filter specifications for enhancing the classification performances of deep learning models for brain tumor MRI scans. Fig. 2 shows the blurring intensity of the Gaussian filter, which was applied to the MRI scan of a brain tumor, according to different kernel sizes and standard deviations.

#### 2) PATTERNED-GRIDMASK

GridMask was applied during the training process to enhance the performances of the deep learning models for classifying the MRI scans of brain tumors. This technique overcame the limitations of conventional information deletion methods such as random erasing, cutout, and hide-and-seek, which either covered objects completely or eliminated only the background [20], [21], [22], [23]. Additionally, it deleted regions corresponding to the uniformly distributed rectangular areas in an image. The deleted rectangular areas were filled with zeros and contained no relevant information. When trained on images obscured by the rectangular regions of GridMask, the deep learning models could perform



**FIGURE 2.** Blurring intensity of gaussian filter according to kernel size and standard deviation.

classifications even when there was any loss of image data, resulting in them exhibiting enhanced generalization and robustness against various patterns and transformations. Overall, GridMask yielded a higher performance while being more straightforward in its approach and computationally less expensive than commonly used augmentation techniques, exhibiting an improved generalization capability and preventing overfitting [24].

The size and density of the rectangular regions obscuring the image were determined by the three parameters  $d_1$ ,  $d_2$ , and ratio. When a rectangle with one side ranging from  $d_1$  to  $d_2$  was generated, the size and density of the obscured area within the rectangular region were determined by the ratio that represents the proportion of unobscured areas. Since the performance varied based on the parameter values, the appropriate values needed to be applied depending on the employed dataset. Therefore, we conducted additional experiments for the four types based on these three parameters. We adjusted the parameters based on the size of the rectangular area and the nature of the distribution. For small rectangular regions with a narrow distribution, the parameters were  $d_1 = 14$ ,  $d_2 = 28$ , with a ratio of 0.5. When these small regions had a wide distribution, we set  $d_1$  to 28,  $d_2$  to 56, and the ratio to 0.75. In larger rectangular regions with a narrow distribution, the parameters were  $d_1 = 28$ ,  $d_2 = 56$ , with a ratio of 0.5. When the distribution was wide in these larger regions, the settings were adjusted to  $d_1 = 56$ ,  $d_2 = 112$ , and a ratio of 0.75. Examples of GridMask being applied for different rectangular areas and distributions of images are shown in Fig. 3.

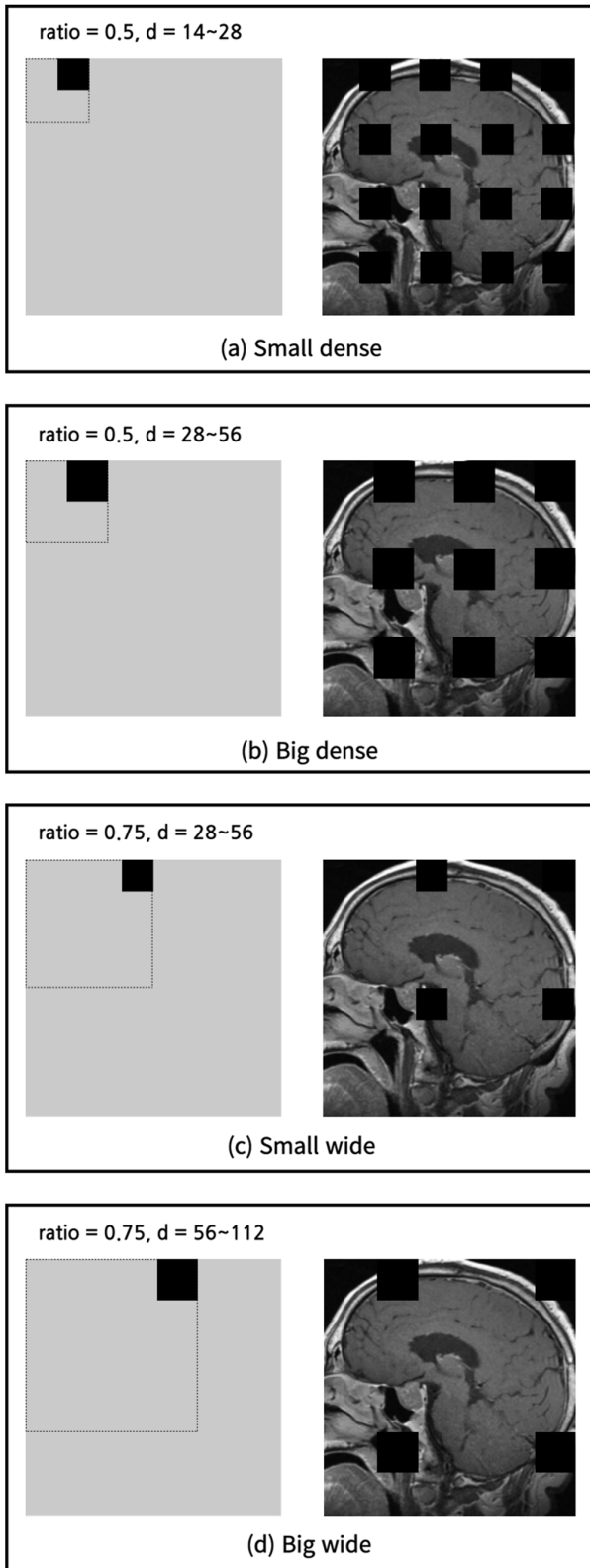
GridMask is a technique developed to achieve a reasonable balance between deletion and reserving of training data, drawing motivation from information deletion methods such as cutout and hide-and-seek (HaS) [25], [26]. Fundamentally, its aim is to enhance performance through a regulated deletion

that belongs neither to excessive deletion nor to excessive reservation. However, although the rectangular areas under standard GridMask follow a consistent distribution, the important details of an object can sometimes be hidden. In the context of brain tumor MRI, if the tumor is hidden, the image will appear normal, while the label still indicates a brain tumor. This discrepancy can cause significant loss during the training of the deep learning model, potentially leading to a decline in accuracy rather than improved generalization. Therefore, in the case of medical data, where features of specific areas are relatively important, even regulated deletion could become excessive deletion.

Consequently, in the Patterned-GridMask approach, instead of completely nullifying any information by filling the rectangular region with zeros, a mosaic pattern with 1-px intervals was applied. In this approach, values alternated between 0 and 1 at these intervals, with 0 representing the black regions with erased information and 1 representing the images that had retained their original features. While this approach is based on GridMask deletion, it is not about complete deletion. Instead, it can be seen as a distortion of certain areas, similar to the Patch Gaussian method, which aids in creating a more robust model [27]. This highlighted that the method retains essential information, even when the rectangular region obscures crucial object features. Consequently, it facilitates a more durable training process for deep learning models. The distinctions between the traditional GridMask and Patterned-GridMask approaches are depicted in Fig. 4.

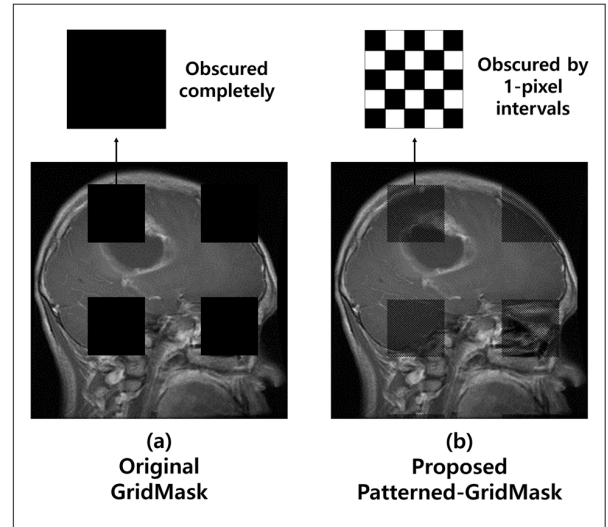
### C. DEEP LEARNING MODELS

In this study, we developed an enhanced brain tumor MRI classification CADx system; ViT-B/16 was used as the baseline deep learning model for performance evaluation.



**FIGURE 3.** Comparison of GridMask size and distribution determined by coverage area (ratio) and side length (d) in brain tumor MRI.

Different performance-enhancement techniques were validated across three deep learning models. The additional deep learning models that were used included the CNN-based



**FIGURE 4.** Characteristics of GridMask and Patterned-GridMask.

EfficientNetV2-M and TresNet-M and the transformer-based MaxViT-B. Overall, two CNN-based and two transformer-based deep learning models, including ViT-B/16, were used for performance comparison. We applied different combinations of techniques to the four deep learning models to evaluate the performance of each combination. All performance evaluations were further validated with 3-fold cross-validation for increased accuracy.

### 1) ViT-B/16

The vision transformer is a deep learning model that applies the transformer architecture, which was developed for natural language processing, to the field of image processing, in which it exhibits a remarkable classification performance. While conventional transformers accept 1-dimensional token embeddings as input sequences, the vision transformer processes 3-dimensional images by converting them into patch formats and then into 1-dimensional vectors. Additionally, position embeddings are applied to preserve the positional information in the images. The resulting final embedding vectors are passed through multi-headed self-attention layers to extract information, followed by a multilayer perceptron for image classification. The pivotal self-attention mechanism calculates attention weights by generating query, key, and value vectors for each input sequence and computes the similarity between the query and key. This technique captures the interdependencies among sequences, leading to a performance that is greater than that of conventional CNN architectures [28].

### 2) MAXViT-B

Conventional vision transformers use self-attention to learn the interdependencies among input data and detect distinctive patterns and relationships. The overall architecture of MaxViT resembled that of ViT, with the exception of the

block module, which comprised these three components: MBCConv, Block Attention, and Grid Attention.

MaxViT uses a novel approach called multi-axis attention, instead of the conventional self-attention mechanism, to handle the global context more efficiently for an entire image [29]. This approach addresses the issue of complexity regarding applying self-attention across the entire space in traditional ViT by dividing the computations into local attention and grid attention. Notably, this method allows the deep learning model to capture the overall image features regardless of the image size, while ensuring that local features are not overlooked.

### 3) EfficientNetV2-M

EfficientNetV2, which prioritizes a fast-training speed, is a successor to EfficientNetV1, which was designed to be smaller and faster. EfficientNetV2 introduces progressive learning, fused-MBCConv, and nonuniform scaling techniques to enhance the training speed. Progressive learning gradually increases the image size during training, thereby overcoming the need for small batch sizes when training with large images. Fused-MBCConv replaces the  $1 \times 1$  conv and  $3 \times 3$  depthwise conv of MBCConv with a single  $3 \times 3$  conv. Additionally, nonuniform scaling optimizes the training speed by progressively increasing the number of layers with stage advances, rather than uniformly scaling all the stages. By leveraging these techniques, the EfficientNetV2 deep learning model achieved an optimal performance with fewer parameters, exhibiting an ImageNet Top-1 accuracy of 83.9%, which indicated an efficient and effective performance [30], [31].

### 4) TResNet-M

ResNet, which was designed with iterative residual blocks, is a deep learning model that can achieve a high accuracy despite having a deep network. TResNet, which is based on ResNet50, aims to reduce graphical processing unit (GPU) processing while simultaneously enhancing accuracy [32]. To achieve this, the traditional Conv $7 \times 7$  stem is replaced with a stem composed of three Conv $3 \times 3$  layers, reducing the input resolution by four times and improving the accuracy [33]. Furthermore, to reduce the GPU memory usage and efficiently increase the batch size, the BatchNorm+ReLU layer is replaced with an Inplace-Activated BatchNorm layer. The leaky-rectified linear unit (ReLU) activation function is used instead of the traditional ReLU to achieve a higher accuracy at the same level of GPU memory consumption. In the block structure segment, BasicBlock layers with a larger receptive field are positioned in the initial two stages of the network, whereas bottleneck layers are placed in the last two stages to enhance both speed and accuracy.

## III. RESULTS

In this study, we used deep learning models to classify glioma, meningioma, normal tissue, and pituitary tissue in brain tumor MRI scans. We primarily chose ViT-B/16 for our

experiments owing to its decent performance and common use as a baseline deep learning model for classification. Additionally, we utilized the MaxViT-B, TresNet-M, and EfficientNetV2-M deep learning models to accurately evaluate the proposed method. All the deep learning models utilized in this study had been pretrained on ImageNet-21K.

A performance evaluation in terms of the accuracy and F1-score was conducted using a confusion matrix. Equations (1) and (2) mathematically represent each performance metric:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

$$F1 - score = \frac{Precision \times Recall}{2(Precision + Recall)} \quad (2)$$

where TP, TN, FP, and FN represent the instances where positive samples are correctly identified as positive, negative samples are correctly identified as negative, negative samples are incorrectly identified as positive, and positive samples are incorrectly identified as negative, respectively.

In this study, Gaussian filters were applied to remove noise from brain tumor MRI scans to enhance the performance. These filters differed in strength based on the kernel size and standard deviation. We applied various combinations of parameters to the brain tumor MRI scans to evaluate their impact on performance. Table 2 presents a comprehensive comparison of the performances at different kernel sizes and standard deviations. All the reported results represent averages of the 3-fold, with the best values denoted in bold. GridMask was applied to enhance the generalization performance of the deep learning model and improve its adaptability across various images. The performance of GridMask was influenced by the size and spacing of the rectangular regions in which the mask had been applied. We conducted four experiments in which the parameters mentioned in section II-B2, namely the sizes d1 and d2 and the density ratio of the rectangular distribution, were varied. GridMask was applied to the images processed by Gaussian filtering. The experimental results were computed as averages from the 3-fold experiments; they are presented in Table 3, with the values in bold indicating the highest performances.

Table 4 presents the original metrics of the ViT-B/16 model along with the performance metrics obtained after applying the Gaussian filter, GridMask, and the proposed Patterned-GridMask, while noting that the Patterned-GridMask was applied independently from the GridMask. The results revealed that applying the Gaussian filter and GridMask had yielded significant performance improvements. When the Gaussian filter was applied, the F1-score of 86.63 had increased to 90.18, indicating a substantial increase of 3.55. Moreover, the integration of GridMask with the images preprocessed using a Gaussian filter resulted in an F1-score of 91.93, which was 1.75 higher than the score obtained when just the Gaussian filter had been applied. Additionally, as shown in Table 4, combining the Gaussian filter with

**TABLE 2.** Performance comparison based on the application of gaussian filter (ViT-B/16).

Standard Deviation	Kernel					
	K = 3		K = 5		K = 7	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
$\sigma = 0.1$	86.00	86.63	85.43	86.08	84.43	85.25
$\sigma = 0.5$	88.93	89.09	<b>90.11</b>	<b>90.18</b>	90.00	90.03
$\sigma = 1.0$	89.25	89.59	85.11	85.24	87.96	88.16
$\sigma = 2.0$	86.92	87.14	88.85	88.92	85.63	85.87
$\sigma = 3.0$	84.92	84.92	88.32	88.36	85.46	85.67
$\sigma = 4.0$	84.92	84.89	87.60	87.78	87.51	87.63
$\sigma = 5.0$	87.43	87.54	88.53	88.70	83.35	83.51

**TABLE 3.** Performance comparison based on the size and distribution of GridMask (ViT-B/16).

Type of GridMask	Accuracy	F1-score
Small dense (d=14~28, ratio=0.5)	90.07	90.15
Big dense (d=28~56, ratio=0.5)	<b>91.85</b>	<b>91.93</b>
Small wide (d=28~56, ratio=0.5)	90.63	90.71
Big wide (d=56~112, ratio=0.5)	90.24	90.38

**TABLE 4.** Comparison of brain tumor classification performance based on applied methods (ViT-B/16).

Applied Method	Accuracy	F1-score
Original	86.00	86.63
Gaussian Filter	90.11	90.18
GridMask Only (Big dense)	89.51	89.62
Gaussian & GridMask	91.85	91.93
Ours (Gaussian & P-Grid)	<b>92.74</b>	<b>92.80</b>

GridMask resulted in a greater performance improvement than that obtained when applying just GridMask. Further experiments confirmed the same, indicating that applying both Gaussian filter and GridMask would be more effective for brain tumor classification. Moreover, Patterned-GridMask, which is our refined version of GridMask, demonstrated an exceptional performance, with an F1-score of 92.08.

To validate the effectiveness of the proposed method for brain tumor classification, we conducted additional experiments using CNN-based deep learning models, including EfficientNetV2-M and TResNet-M, and the transformer-based model MaxViT-B. These three deep learning models, which had been pre-trained on the same dataset as that used for training ViT-B/16, were selected as the deep learning models in the study. The overall performances of the deep learning models after applying different methods, including the proposed Patterned-GridMask, are presented in Table 5, where the highest performance is indicated in bold.

Clearly, the transformer-based deep learning models, specifically ViT-B/16 and MaxViT, exhibited greater improvements in performance than the CNN-based deep learning models, namely EfficientNetV2-M and TResNet-M, did after applying the Gaussian filter. This was attributed to the inherent attention mechanism of transformer-based deep learning models, which leverages attention to capture the interactions among input sequence elements, facilitating feature extraction and global information comprehension. Consequently, applying a Gaussian filter was more effective in reducing the global noise of transformer deep learning models. Regarding CNN-based deep learning models, which are more specialized in capturing local features, they displayed relatively modest performance enhancements, which was attributed to the localized nature of their feature extraction.

Table 5 and Fig. 5 show that the combination of Patterned-GridMask and the Gaussian filter produced greater improvements in the performances of the four deep learning models than that produced by the combination of GridMask and the Gaussian filter. These improvements achieved by masking portions of the image can be likened to the effect of masked autoencoders, which train a deep learning model to predict obscured parts by masking certain sections of the input image [34]. Similarly, by masking certain parts of the objects, both GridMask and Patterned-GridMask can enable deep learning models to capture object features more effectively, providing a form of regularization that can

TABLE 5. Performance evaluation based on four classification models and applied methods.

Classification Model	Applied Method							
	Original		Gaussian Filter		Gaussian & GridMask		Ours (Gaussian & P-Grid)	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
ViT-B/16	86.00	86.63	90.11	90.18	91.85	91.93	<b>92.74</b>	<b>92.80</b>
MaxViT-B	95.83	95.86	96.31	96.31	96.85	96.87	<b>97.74</b>	<b>97.75</b>
TresNet-M	94.18	94.23	94.78	94.78	95.63	95.64	<b>96.86</b>	<b>96.89</b>
EfficientNetV2-M	93.64	93.65	93.25	93.26	94.49	94.51	<b>95.79</b>	<b>95.81</b>

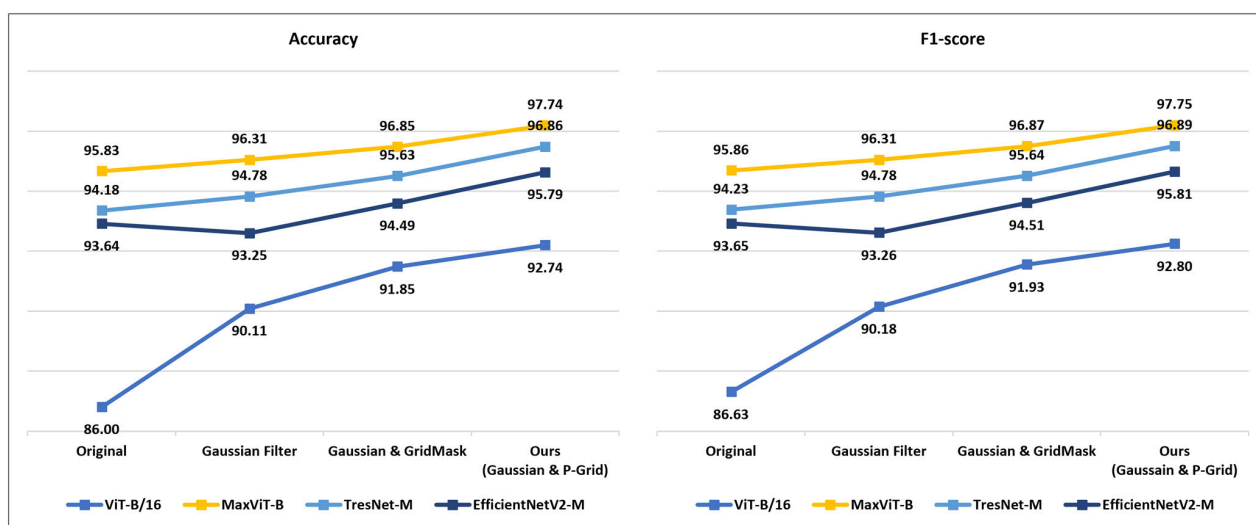


FIGURE 5. Performance comparison according to classification models and applied methods.

enhance the generalization capabilities of deep learning models. Compared to the approach of the original GridMask, which simply removes the information in rectangular patterns, Patterned-GridMask offers a more detailed approach that minimizes the loss of essential information when masking vital features, achieving a balance between maintaining the crucial details in an image and delivering the benefits of data augmentation. Overall, the combined application of the Gaussian filter and GridMask was effective in enhancing the brain tumor classification performance, and integrating Patterned-GridMask instead of GridMask further optimized the performance.

#### IV. CONCLUSION

In this study, we proposed a CADx system algorithm optimized for brain tumor classification. The proposed Patterned-GridMask effectively overcame the limitations of the original GridMask, which obscured brain tumors. Applying Patterned-GridMask increased the performance of the baseline deep learning model (transformer-based

ViT-B/16) by over 6%. Additionally, this approach demonstrated the highest increase in performance on all the additional validation deep learning models. Additionally, the effectiveness of the proposed technique was validated using the transformer-based MaxViT-B and CNN-based TresNet-M and EfficientNetV2-M. Despite the high accuracy of pre-existing models, the proposed Patterned-GridMask elicited consistent performance enhancements across all models, thus substantiating the effectiveness of this method. Overall, the proposed technique can yield excellent performance improvements in brain tumor diagnosis, minimizing the need for invasive procedures and improving the survival rates of people that are afflicted with it. Furthermore, it can be used as a reference in developing other techniques that can improve the classification performances of deep learning models in fields related to medical diagnosis.

Although the proposed method exhibited significant performance improvements, there are still areas for further research. To address the limitations of scarce medical data, we plan to implement various augmentation techniques. Our



focus will not be on basic augmentation, but on methods that maintain the unique characteristics of brain tumor images. Second, we intend to apply the proposed method to other classification models to verify its performance. Since it showed promising results when applied to the four deep learning models in the study, we wish to also confirm its effectiveness when applied to other deep learning models. Finally, we plan to verify whether our method can produce similar performance improvements on medical datasets other than those of brain tumors, which includes investigating the optimal parameter values for these datasets.

## REFERENCES

- [1] Asan Med. Center. (4, 2023). *Medical Information*. [Online]. Available: <https://www.amc.seoul.kr/asan/healthinfo/body/bodyDetail.do?bodyId=15>
- [2] C. H. Rhee, "What is a brain tumor," *Mag. J. Korean Soc. Steel Construct.*, vol. 20, no. 3, pp. 121–123, 2008.
- [3] Korea Nat. Stat. Office. (2020). *Cancer Incidence and Mortality Statistics 2020*. [Online]. Available: <http://kostat.go.kr/>
- [4] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, "Cancer statistics 2023," *CA, A Cancer J. Clinicians*, vol. 73, no. 1, pp. 17–48, Jan. 2023.
- [5] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Phys. Technol.*, vol. 10, no. 3, pp. 257–273, Sep. 2017.
- [6] Z. Rasheed, Y.-K. Ma, I. Ullah, Y. Y. Ghadi, M. Z. Khan, M. A. Khan, A. Abdusalomov, F. Alqahtani, and A. M. Shehata, "Brain tumor classification from MRI using image enhancement and convolutional neural network techniques," *Brain Sci.*, vol. 13, no. 9, p. 1320, Sep. 2023.
- [7] H. Alsaifi, R. Guesmi, B. M. Alshammari, T. Hamrouni, T. Guesmi, A. Alzamil, and L. Belguesmi, "A novel data augmentation-based brain tumor detection using convolutional neural network," *Appl. Sci.*, vol. 12, no. 8, p. 3773, Apr. 2022.
- [8] A. A. Asiri, A. Shaf, T. Ali, M. Aamir, M. Irfan, S. Alqahtani, K. M. Mehdar, H. T. Halawani, A. H. Alghamdi, A. F. A. Alshamrani, and S. M. Alqhtani, "Brain tumor detection and classification using fine-tuned CNN with ResNet50 and U-Net model: A study on TCGA-LGG and TCIA dataset for MRI applications," *Life*, vol. 13, no. 7, p. 1449, Jun. 2023.
- [9] K. S. Kumar and A. Rajendran, "Deep convolutional neural network for brain tumor segmentation," *J. Electr. Eng. Technol.*, vol. 18, no. 5, pp. 3925–3932, Sep. 2023.
- [10] G. Murtaza, L. Shuib, G. Mujtaba, and G. Raza, "Breast cancer multi-classification through deep neural network and hierarchical classification approach," *Multimedia Tools Appl.*, vol. 79, nos. 21–22, pp. 15481–15511, Jun. 2020.
- [11] V. Göreke, "A novel deep-learning-based CADx architecture for classification of thyroid nodules using ultrasound images," *Interdiscipl. Sci. Comput. Life Sci.*, vol. 15, no. 3, pp. 360–373, Sep. 2023.
- [12] A. Gopinath, P. Gowthaman, L. Gopal, M. Abul Ala Walid, M. Manju Priya, and K. Keshav Kumar, "Enhanced lung cancer classification and prediction based on hybrid neural network approach," in *Proc. 8th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jun. 2023, pp. 933–938.
- [13] M. Nickparvar. (2021). *Brain Tumor MRI Dataset*. [Online]. Available: <https://doi.org/10.34740/KAGGLE/DSV/2645886>
- [14] J. Cheng. (2017). *Brain Tumor Dataset*. [Online]. Available: <https://doi.org/10.6084/m9.figshare.1512427.v5>
- [15] (2020). *Kaggle Competition: Brain Tumor Classification (MRI)*. [Online]. Available: <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>
- [16] (2020). *Kaggle Competition: Brain Tumor Detection*. [Online]. Available: <https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection>
- [17] E. Terreno, D. D. Castelli, A. Viale, and S. Aime, "Challenges for molecular magnetic resonance imaging," *Chem. Rev.*, vol. 110, no. 5, pp. 3019–3042, 2010.
- [18] M. Bhuiyan, M. A. A. Nasim, S. Saif, D. K. D. Gupta, M. J. Alam, and S. Talukder, "Online learning for X-ray, CT or MRI," 2023, *arXiv:2306.06491*.
- [19] I. T. Young and L. J. van Vliet, "Recursive implementation of the Gaussian filter," *Signal Process.*, vol. 44, no. 2, pp. 139–151, Jun. 1995.
- [20] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random erasing data augmentation," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 7, pp. 13001–13008.
- [21] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," 2017, *arXiv:1708.04552*.
- [22] K. K. Singh, H. Yu, A. Sarmasi, G. Pradeep, and Y. Jae Lee, "Hide-and-seek: A data augmentation technique for weakly-supervised localization and beyond," 2018, *arXiv:1811.02545*.
- [23] P. Chen, S. Liu, H. Zhao, X. Wang, and J. Jia, "GridMask data augmentation," 2020, *arXiv:2001.04086*.
- [24] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "AutoAugment: Learning augmentation policies from data," 2018, *arXiv:1805.09501*.
- [25] P. Kaur, B. S. Khehra, and E. B. S. Mavi, "Data augmentation for object detection: A review," in *Proc. IEEE Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2021, pp. 537–543.
- [26] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, Dec. 2022, Art. no. 100258.
- [27] R. Gontijo Lopes, D. Yin, B. Poole, J. Gilmer, and E. D. Cubuk, "Improving robustness without sacrificing accuracy with patch Gaussian augmentation," 2019, *arXiv:1906.02611*.
- [28] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.
- [29] Z. Tu, H. Talebi, H. Zhang, F. Yang, P. Milanfar, A. Bovik, and Y. Li, "MaxViT: Multi-axis vision transformer," in *Proc. 17th Eur. Conf. Comput. Vis. Cham, Switzerland: Springer*, 2022, pp. 459–479.
- [30] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 1–10.
- [31] M. Tan and Q. V. Le, "EfficientNetV2: Smaller models and faster training," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 10096–10106.
- [32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [33] T. Ridnik, H. Lawen, A. Noy, E. Ben, B. G. Sharir, and I. Friedman, "TRResNet: High performance GPU-dedicated architecture," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 1399–1408.
- [34] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, "Masked autoencoders are scalable vision learners," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 15979–15988.



**JI-HYEON LEE** is currently pursuing the B.S. and M.S. degrees with the Department of Electronics Engineering and the Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, South Korea.



**JUNG-WOO CHAE** received the B.S. degree from the Department of Electronics Engineering and the M.S. degree from the Department of BIT Medical Convergence from Kangwon National University, South Korea, in 2019 and 2021, respectively, where is currently pursuing the Ph.D. degree.



**HYUN-CHONG CHO** (Member, IEEE) received the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida, Gainesville, FL, USA, in 2009. From 2010 to 2011, he was a Research Fellow with the University of Michigan, Ann Arbor, MI, USA. From 2012 to 2013, he was a Chief Research Engineer with LG Electronics, South Korea. He is currently a Professor with the Department of Electronics Engineering and the Interdisciplinary BIT Medical Convergence, Kangwon National University, South Korea.

...