



## OPEN ACCESS

## EDITED BY

Dong Zeng,  
Southern Medical University, China

## REVIEWED BY

Sophie Adler,  
University College London, United Kingdom  
Dezhi Cao,  
Shenzhen Children's Hospital, China  
Xiaoming Jiang,  
Chongqing University of Post and  
Telecommunications, China

## \*CORRESPONDENCE

Lixin Cai  
✉ pufhpecclx@163.com  
Guixia Kang  
✉ gxkang@bupt.edu.cn

<sup>†</sup>These authors have contributed equally to this work

RECEIVED 24 May 2025

ACCEPTED 29 August 2025

PUBLISHED 08 October 2025

## CITATION

Li Q, Yu H, Zhang M, Yuan X, Cai L and Kang G (2025) LEM-UNet: an edge-guided network for 3D multimodal images segmentation in focal cortical dysplasia.  
*Front. Neurosci.* 19:1634606.  
doi: 10.3389/fnins.2025.1634606

## COPYRIGHT

© 2025 Li, Yu, Zhang, Yuan, Cai and Kang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# LEM-UNet: an edge-guided network for 3D multimodal images segmentation in focal cortical dysplasia

Qiunan Li<sup>1†</sup>, Hao Yu<sup>2†</sup>, Manli Zhang<sup>1</sup>, Xiaotong Yuan<sup>1</sup>, Lixin Cai<sup>2\*</sup> and Guixia Kang<sup>1\*</sup>

<sup>1</sup>School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China, <sup>2</sup>Pediatric Epilepsy Center, Peking University First Hospital, Beijing, China

**Introduction:** Focal cortical dysplasia (FCD) is one of the common causes of refractory epilepsy. The subtle and indistinct edge of FCD lesions pose considerable challenges for accurate lesion localization. Therefore, we propose an edge guided segmentation network based on Laplacian pyramid to improve the localization performance of FCD lesions.

**Methods:** This is a retrospective study evaluated on two independent datasets. The proposed Laplacian Edge Mix UNet (LEM-UNet) builds upon the MedNeXt baseline and incorporates the Laplacian Edge Attention (LEA) block and the Multi-strategy Feature Fusion (MFF) block. LEA block captures lesion details and edge information during the encoding phase by integrating Laplacian pyramid feature maps with an attention mechanism, while MFF block fuses edge features with high level features during the decoding phase.

**Results:** The model's performance was assessed through 5-fold cross-validation across both Open and Private Datasets, demonstrating superior performance. The average Dice Coefficient achieved 0.452 and 0.597 on the Open and Private Datasets, respectively, representing improvements of 2.40% and 2.90% compared to the baseline model.

**Discussion:** The results demonstrate the importance of focusing on lesion edge in the FCD segmentation task. The integration of the Laplacian pyramid enhances the model's ability to capture lesions with blurred edge and subtle features. LEM-UNet exhibits significant advantages over current FCD segmentation algorithms. The source code and pre trained model weights are available at <https://github.com/simplify403/LEM-UNet>.

## KEYWORDS

focal cortical dysplasia, multimodal medical imaging, deep learning, edge information, medical image segmentation

## 1 Introduction

Epilepsy is a prevalent chronic neurological disorder characterized by the abrupt, abnormal discharge of brain neurons (Fisher et al., 2005). Most epilepsy patients can control seizures using medications, but one-third of patients develop drug-resistant epilepsy (DRE). Prolonged epileptic seizures can lead to severe consequences such as brain damage, decreased quality of life, and even premature death (Avakyan et al., 2017). Focal cortical dysplasia (FCD), a malformation of cortical development, is the most common cause of DRE (Kabat and Król, 2012). Clinically, an effective treatment for patients with

DRE is surgical resection of FCD lesions. About 70% of patients experience seizure relief after surgery (Wagstyl et al., 2022). A favorable outcome is associated with the complete resection of FCD lesions (Willard et al., 2022).

In clinical diagnosis, physicians can identify FCD lesions by analyzing neuroimaging results. Magnetic Resonance Imaging (MRI) is a medical imaging technique that uses the principles of nuclear magnetic resonance to produce detailed images of the internal structures of the human body. On MRI scans, characteristic manifestations of FCD include increased cortical thickness, blurring of the gray/white matter junction, the presence of a transmantle sign, abnormal cortical folding pattern, and increased signal intensity on FLAIR/T2-weighted MRI (Urbach et al., 2022) as depicted in Figure 1. Positron Emission Tomography (PET) is a nuclear medicine imaging technique used to observe biological processes and functional activities within the human body. Between seizures, PET images using fluorodeoxyglucose show hypometabolism in areas of gray matter tissue that are associated with the epileptogenic region (Baete et al., 2004). However, the identification of FCD lesions in clinical practice remains a formidable challenge. Firstly, FCD lesions are often subtle, exhibit diverse morphological features, and possess poorly defined edge, making them difficult to detect through routine visual inspection. Secondly, the analysis of high-resolution 3D imaging datasets is time-consuming, requires specialized expertise, and can be subjective in certain cases. Therefore, there is an urgent need for precise and efficient computational techniques to assist in the localization of FCD lesions.

In the early stages, the task of FCD lesion detection primarily focused on the extraction of cortical morphological features, with the main approaches including voxel-based morphometry (Mechelli et al., 2005) and surface-based morphometry (Thesen et al., 2011). The extracted voluminous cortical morphological features were then input into machine learning models to distinguish between normal and pathological tissues, such as David et al. (2021); Jin et al. (2018); and Hong et al. (2014). However, these methods exhibit several limitations: the computation of features is complex and extremely

time-consuming, and some basic features are insufficient to accurately differentiate FCD lesions.

Currently, deep learning-based models have achieved remarkable success in various medical imaging tasks, significantly advancing disease detection and diagnosis. Convolutional Neural Networks (CNNs) are the most widely used model for the detection of FCD lesions, with a lot of work is based on it, such as Gill et al. (2021) and Wang et al. (2020). Enhanced model designs have led to improved detection performance. Dev et al. (2019) trained a U-Net using 2D slices derived from 3D FLAIR-weighted sequence images. Niyas et al. (2021) proposed a 3D Res-UNet model for segmenting FCD lesions from MRI volumes, leveraging inter-slice information from 3D MRI data to achieve superior segmentation performance. Thomas et al. (2020) introduced a Multi-Res-Attention UNet to address the significant semantic gap in feature mapping caused by long-range connections between the encoder and decoder layers, resulting in a higher FCD detection rate. The nnU-Net framework has been proposed as a self-adaptive solution for diverse medical image segmentation tasks (Isensee et al., 2018). Spitzer et al. (2023) developed a graph-based nnU-Net for segmenting FCD lesions using surface-based cortical data, achieving a 22%–27% improvement in specificity compared to baseline methods. Zhang S. et al. (2024) employed a 3D full-resolution nnU-Net for automatic lesion segmentation, demonstrating strong performance in FCD lesion detection with a sensitivity of 0.73. Zhang X. et al. (2024) integrated multi-scale transformers into CNN-based encoding and decoding structures to overcome the limitations of local receptive fields in CNN models and successfully identified lesions in 82.4% of patients.

The aforementioned methods have significantly contributed to the detection of FCD lesions. However, traditional methods based on manual extraction of cortical morphological features combined with machine learning exhibit several limitations: (1) feature extraction is computationally complex and time-consuming, (2) basic features are insufficient to fully capture lesion characteristics. Transformer-based models, while powerful, demand substantial computational resources and are challenging to train on small datasets. CNN-based models have demonstrated significant potential in the field of lesion segmentation due to their

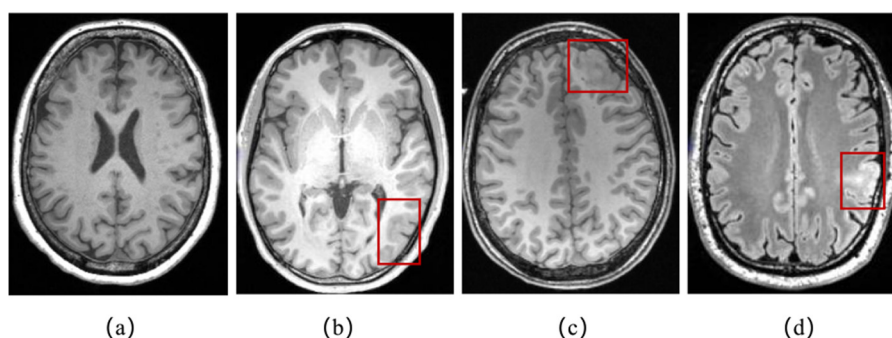


FIGURE 1

(a) A T1-weighted MRI slice of normal brain; (b) abnormal gray matter thickening in a T1-weighted MRI slice in red box; (c) blurred junction of gray and white matter in a T1-weighted MRI slice in red box; (d) increased signal intensity in FLAIR in red box.

excellent local perception capabilities. However, they still exhibit limitations when applied to FCD lesion segmentation tasks. FCD lesions are characterized by blurred edge, making it challenging for networks to accurately capture and identify edges and subtle lesions. Integrating edge detection methods into convolutional models opens new avenues to address these challenges. In other medical image segmentation tasks, models incorporating edge detection methods have shown exceptional performance, as seen in [Zhu et al. \(2023\)](#) and [Lu et al. \(2023\)](#). However, there is a paucity of relevant research in the task of FCD lesion segmentation.

To address these clinical needs and the shortcomings of existing algorithms, this study proposes an edge information-guided FCD lesion segmentation model. Based on the MedNext ([Roy et al., 2023](#)) network architecture, the main contributions of this paper are as follows:

1. We propose a Laplacian edge attention block to extract edge features from the T1 modality and integrate them with attention mechanisms, enabling spatial localization from global edges to FCD lesion edges.
2. We propose a multi-strategy feature fusion block to effectively combine high-frequency edge information with network decoding information.

## 2 Materials and methods

### 2.1 Materials

Our network was evaluated on two distinct datasets: the Open Dataset, provided by the Department of Epileptology at the University Hospital Bonn ([Schuch et al., 2023](#)) and our proprietary dataset, referred to as the Private Dataset. More detailed information on the two datasets is shown in [Table 1](#) and the brief introduction of the two datasets is as follows.

**Open dataset:** Open dataset consists of 85 people with FCD type II and 85 control persons. Data for each patient include preoperative T1, preoperative FLAIR and lesion ground truth. Lesion ground truth was annotated on FLAIR images by two neurologists. Because 7 patients underwent a FLAIR sequence with thick slices, we excluded these data and ended up with 78 3D cases.

**Private dataset:** Private dataset consists of 125 patients. Among them, 41.6% of the patients were diagnosed with FCD IIB pathologically, 24.8% with FCD IIA, and 18.4% with other types of malformation of cortical development, while the remaining small proportion was attributed to other etiological causes. Each patient contains preoperative 3-T T1, FLAIR, FDG-PET images, and lesion ground truth. Lesion ground truth was annotated on T1 images by the physician. This study has obtained approval from the Ethics Committee of Peking University First Hospital. All participants provided written informed consent for us to use their data for research purposes.

### 2.2 Methods

#### 2.2.1 Model architecture

Before detailing our method, we provide a concise overview of the MedNeXt network. MedNeXt enhances the nnU-Net framework by introducing several specialized modules: MedNeXt

TABLE 1 Subjects characteristics of FCD.

| Characteristics                       | Private dataset | Open dataset    |
|---------------------------------------|-----------------|-----------------|
| Number                                | 125             | 78              |
| Age of surgery(years, mean $\pm$ std) | 5.98 $\pm$ 4.15 | 6.30 $\pm$ 2.48 |
| Sex (male:female)                     | 78:47           | 44:34           |
| Hemisphere (left:right)               | 60:65           | 41:37           |
| Undergoing surgery                    | 125             | 46              |
| Pathology                             |                 |                 |
| FCD IA                                | 3               | 0               |
| FCD IB                                | 10              | 0               |
| FCD IIA                               | 31              | 15              |
| FCD IIB                               | 52              | 31              |
| FCD IIIB                              | 1               | 0               |
| MCD                                   | 23              | 0               |
| TSC                                   | 2               | 0               |
| Gliosis                               | 1               | 0               |
| MOGHE                                 | 2               | 0               |
| NA                                    | 0               | 32              |
| One year outcome                      |                 |                 |
| Seizure-free                          | 118             | 34              |
| Not seizure-free                      | 7               | 8               |
| NA                                    | 0               | 36              |

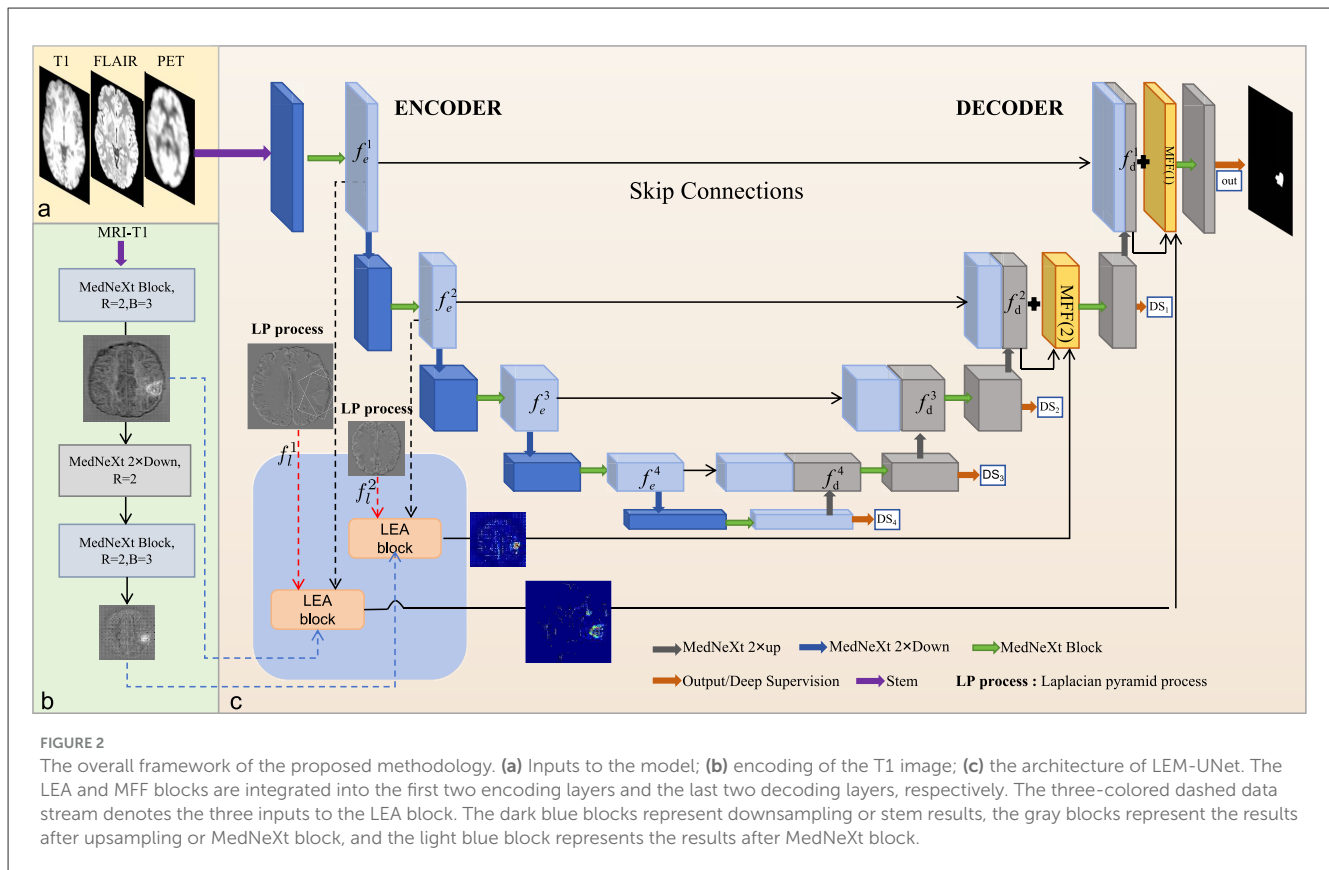
std, standard deviation; FCD, focal cortical dysplasia; MCD, malformation of cortical development; TSC, tuberous sclerosis complex; MOGHE, mild malformation of cortical development with oligodendroglial hyperplasia; NA, not applicable; Pathology classifications are based on the International League Against Epilepsy (ILAE) criteria recommendations for the neuropathologic workup of epilepsy surgery brain tissue.

Block, MedNeXt 2 $\times$ down Block, MedNeXt 2 $\times$ up Block, Stem Block and Output/Deep Supervision Block. Our approach adopts the MedNeXt network, adhering to the architectural components outlined in the original paper.

The proposed network architecture, illustrated in [Figure 2](#), comprises an encoder-decoder model, Laplacian edge attention blocks, and multi-strategy feature fusion blocks. The encoder-decoder model, derived from MedNeXt, supports composite scaling in depth, width, and receptive field to effectively extract contextual and high-level features from multimodal images. The Laplacian edge attention block integrates Laplacian pyramid features with encoded network features to enhance edge representation, directing the network's focus toward lesion edge. The multi-strategy feature fusion blocks combine edge features with decoded network features, employing diverse fusion strategies across decoding layers to retain edge information from shallow features. The subsequent sections elaborate on the structure and efficacy of the Laplacian edge attention block and the multi-strategy feature fusion block.

##### 2.2.1.1 Laplacian edge attention block

The segmentation of FCD lesions is particularly challenging due to their ambiguous edge. The LEA block addresses this by



enhancing edge information, enabling the network to better focus on lesion edge and thereby improving segmentation performance. Since edge features are inherently shallow, the LEA block is applied to the first and second layers of the encoding process to effectively capture texture details of lesion edges.

The LEA block proposed in this paper is shown in Figure 3. Each LEA block receives three inputs: encoded feature  $f_e^i$  from the main network branch, encoded feature  $f_{e-T1}^i$  from the T1 branch, and high-frequency features  $f_l^i$  from the Laplacian pyramid. These features generate an output feature map  $f_{edge}^i$  after being processed by the LEA block, which then enters the subsequent network for feature fusion processing. The following section will detail the inputs of the LEA block and its specific processing methods.

#### 2.2.1.1.1 Input of LEA block

The input of LEA block consists of three features.

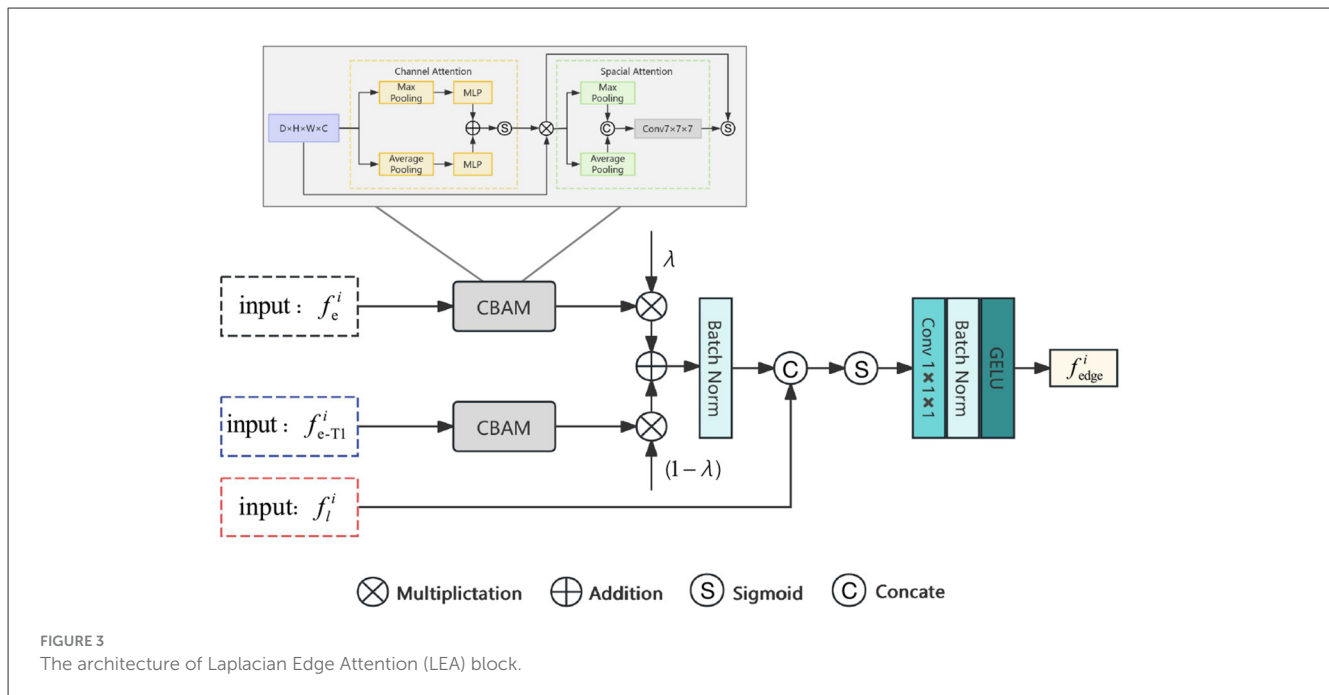
- Encoded feature  $f_e^i$ : Our backbone network continues to use the MedNext encoding-decoding model to process multimodal data, representing the feature maps extracted after MedNext block as  $f_e^i \in \mathbb{R}^{C_i \times D_i \times H_i \times W_i}$ . To obtain rich shallow-level features, the first and second layer feature maps after MedNext processing are used, namely  $f_e^1 \in \mathbb{R}^{C \times D \times H \times W}$  and  $f_e^2 \in \mathbb{R}^{2C \times \frac{D}{2} \times \frac{H}{2} \times \frac{W}{2}}$ , as shown by the black dotted data stream in Figure 2.
- Encoded feature  $f_{e-T1}^i$  from branch T1: The T1-weighted MRI contains rich tissue texture and structural information. We individually extracted the characteristics, still using the

MedNext network blocks, and obtained characteristics  $f_{e-T1}^1 \in \mathbb{R}^{C \times D \times H \times W}$  and  $f_{e-T1}^2 \in \mathbb{R}^{2C \times \frac{D}{2} \times \frac{H}{2} \times \frac{W}{2}}$ , as shown by the blue dotted data stream in Figure 2. Parameters  $R$  and  $B$  refer to paper (Roy et al., 2023).

- High frequency features  $f_l^i$ : The Laplacian pyramid is a multiresolution image representation method based on the Gaussian pyramid, capable of capturing edges and detail information of images at different scales (Burt and Adelson, 1987). In this paper, this classic edge detection technique is selected to highlight the high-frequency edge features and the detail features of the images. In the LEA block, T1-weighted MRI is used as input to construct the Laplacian pyramid, as T1-weighted MRI provides rich texture information and clear anatomical structures. The specific calculation can be seen in Equation 1.

$$\begin{aligned} N_k &= N, \text{ if } k = 0, \\ N_k &= d(gs(N_{k-1})), \text{ if } k \geq 1, \\ L_k &= N_k - up(N_{k+1}), \end{aligned} \quad (1)$$

where  $N$  represents the input image,  $gs$  represents the convolution operator with Gaussian filter,  $d$  is the  $2 \times$  downsampling operation, and  $up$  is the  $2 \times$  linear interpolation upsampling. We denote  $f_l^i = L_i(N)$ . As the image scale decreases, due to multiple processes of Gaussian filtering, upsampling, and downsampling, the high-frequency details in the feature maps significantly decrease. Therefore, the first and second layers of the Laplacian pyramid feature  $f_l^1 \in$



$\mathbb{R}^{1 \times D \times H \times W}$  and  $f_l^2 \in \mathbb{R}^{1 \times \frac{D}{2} \times \frac{H}{2} \times \frac{W}{2}}$  are selected as inputs for the LEA block, as shown by the red dotted data stream in Figure 2.

#### 2.2.1.1.2 Procedure of LEA block

At the  $i^{\text{th}}$  layer, the LEA block receives the three inputs mentioned above. At first,  $f_e^i$  and  $f_{e-T1}^i$  are each passed through a Convolutional Block Attention Module (CBAM) (Woo et al., 2018). The CBAM primarily consists of two components: channel attention and spatial attention, which serve to emphasize useful channel features and important spatial locations, respectively. Subsequently, the weights of the feature maps are adjusted to enhance the representation of salient features and  $f_{map}^i$  is obtained. The process is shown in Equation 2.

$$\begin{aligned} \beta &= 1 - \lambda, \\ f_{map}^i &= BN(CBAM(f_e^i) \times \lambda + CBAM(f_{e-T1}^i) \times \beta), \end{aligned} \quad (2)$$

where  $BN$  is the batch normalization,  $\lambda$  is a learnable parameter, with the purpose of adjusting the weights of  $f_e^i$  and  $f_{e-T1}^i$ , allowing the network to autonomously learn the complementary information between this two features.

The high frequency features  $f_l^i$  correspond to the overall edge information and texture features of the original image. In order to obtain edge information and detailed features related to the FCD lesions, we fuse  $f_l^i$  and  $f_{map}^i$ , further screening the features to ensure that the edges of the FCD lesions are highlighted, and finally obtain the edge attention feature  $f_{edge}^i$ , characterized as follows Equation 3.

$$\begin{aligned} map^i &= \sigma(\text{con}[f_l^i, f_{map}^i]), \\ f_{edge}^i &= g(BN(\text{Cov}(map^i))). \end{aligned} \quad (3)$$

In this context, the symbol  $\sigma$  signifies the sigmoid function,  $\text{Cov}$  signifies the  $1 \times 1 \times 1$  convolution,  $\text{con}$  is Concatenation operation and the symbol  $g$  represents GELU function.

Through the LEA block, we obtain the edge attention features for the first and second layers, denoted  $f_{edge}^1 \in \mathbb{R}^{C \times D \times H \times W}$  and  $f_{edge}^2 \in \mathbb{R}^{2C \times \frac{D}{2} \times \frac{H}{2} \times \frac{W}{2}}$ , respectively. These features are then inputted into the feature fusion block, where they are combined with the backbone network to achieve better segmentation performance.

#### 2.2.1.2 Multi-strategy feature fusion block

Multi-strategy feature fusion (MFF) block is shown in Figure 4. Our backbone network undergoes four downsampling and four upsampling operations. The upsampling features are now denoted as  $f_d^i$ , with the last upsampling result denoted as  $f_d^1$  and the penultimate upsampling result as  $f_d^2$ . Each MFF block accepts two inputs:  $f_d^i$  and  $f_{edge}^i$ , employing different fusion strategies at different decoder layers to maximize the use of edge information of the injury obtained from the LEA block. Define the output features of MFF block as  $f_{mff}^i$ .

In MFF (1), the inputs are processed as shown in Equation 4.

$$\begin{aligned} M^i &= f_d^i - \text{up}(d(g_s(f_d^i))), \\ f_{mff}^1 &= \sigma(M^1) \times f_{edge}^1, \end{aligned} \quad (4)$$

where  $M^i$  represents the high-frequency information obtained from the decoded feature  $f_d^i$  and it is designed to further suppress noise in the output of the LEA block. In MFF (2), the inputs are processed as shown in Equation 5.

$$\begin{aligned} I &= \sigma(\text{Cov}(f_{edge}^2)), \\ f_{mff}^2 &= \sigma(M^2) \times (f_{edge}^2 + \text{Cov}(d(f_{edge}^1))) \times I. \end{aligned} \quad (5)$$

We employ distinct fusion strategies in MFF block. The encoded features contain rich texture information in the shallow

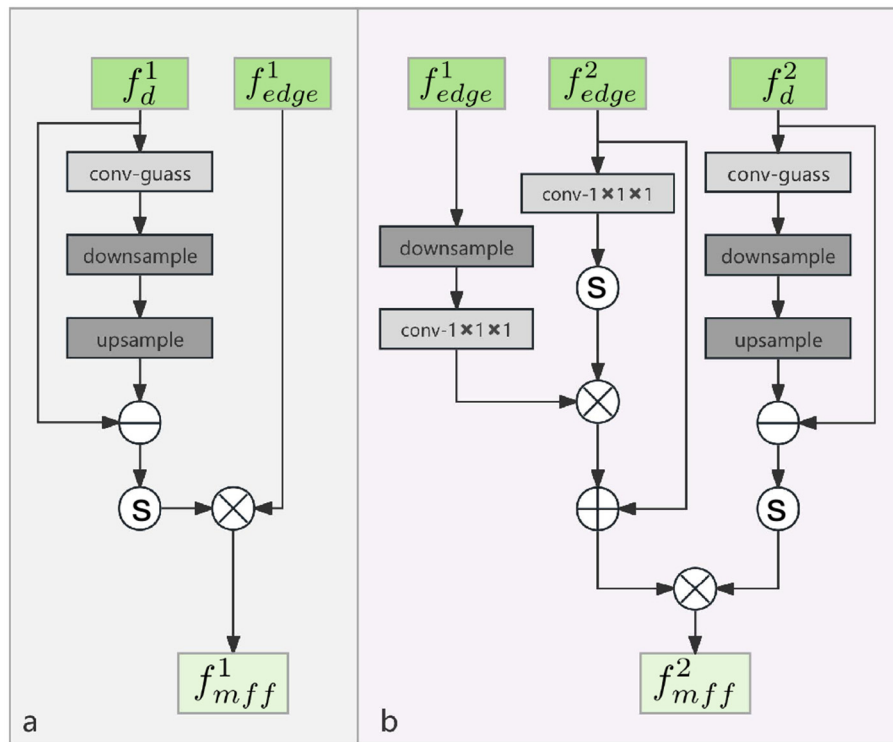


FIGURE 4  
The architecture of multi-strategy feature fusion (MFF) block. (a) MFF(1); (b) MFF(2).

layers, and as the network deepens, it increasingly focuses on complex high-level features and abstract representations. Similarly, in the Laplacian pyramid, with the increasing number of sampling iterations, the detail information in the feature maps is progressively lost.

Therefore, during the feature fusion phase, the edge information rich feature  $f_{edge}^1$  undergoes a  $2\times$  downsampling and is fed into the subsequent layer for further processing. The purpose of this operation is to make full use of the edge information from the shallow layers of the network and to minimize the loss of information. Finally, the fused feature  $f_{mff}^i$  from the encoding layer  $i^{th}$  is added to  $f_d^i$  for decoding.

### 2.2.2 Loss function

We employ a combination of Dice loss  $\mathcal{L}_{Dice}$  and Cross-Entropy loss  $\mathcal{L}_{CE}$  at the training stage of our network. It can be defined as follows in Equation 6.

$$\mathcal{L} = \sum_{i=1}^D w_{dice}^i * \mathcal{L}_{Dice}(f_{DS}^i, f^i) + w_{ce}^i * \mathcal{L}_{CE}(f_{DS}^i, f^i), \quad (6)$$

where  $D$  denotes the total number of layers, as we use deep supervision in our network.  $w_{dice}^i$  and  $w_{ce}^i$  are the weights of  $\mathcal{L}_{Dice}$  and  $\mathcal{L}_{CE}$  at the  $i^{th}$  decoding layer.  $f_{DS}^i$  is predicted feature map at the  $i^{th}$  decoding layer.  $f^i$  is Ground Truth (GT) at the  $i^{th}$  layer.

### 2.2.3 Evaluation metrics

To quantitatively compare our method with others, we employ five evaluation metrics.

Dice Coefficient (DC) is used to measure the overlap between the predicted segmentation and the ground truth. To assess individual performance, we set the DC thresholds at 0.0 and 0.22. A DC score greater than 0.0, the so-called “one voxel overlap,” has been used in other FCD detection studies (David et al., 2021; Spitzer et al., 2022; Gill et al., 2021), and a higher threshold of 0.22 has been shown to be effective in reproducing the precise positioning performance of expert raters (Walger et al., 2024). Precision (Pre) represents the proportion of samples that are actually positive among those predicted as positive by the model and reflects the accuracy of the positive predictions of the model. Recall (Rec) indicates the proportion of actual positive samples that are correctly predicted as positive by the model and reflects the model's coverage ability in the positive class. Intersection over Union (IoU) is used to measure the degree of overlap between the result of the segmentation and the ground truth. They are defined as Equation 7.

$$DC(G, P) = 2 * \frac{|G \cap P|}{|G| + |P|} = 2 * \frac{G \cdot P}{G^2 + P^2},$$

$$Pre = \frac{TP}{TP + FP},$$

$$Rec = \frac{TP}{TP + FN}, \quad (7)$$

$$IoU = \frac{G \cap P}{G \cup P},$$

$$Detection = \frac{N_d}{N_a} * 100\%,$$

where  $G$  and  $P$  denote the GT and the predicted results of the model, respectively.  $TP$  refers to the number of true positive voxels,  $FP$  refers to the number of false positive voxels and  $FN$  denotes the false negative voxels.  $N_d$  refers to the number of individuals with a dice value higher than the threshold of 0.0 or 0.22, while  $N_a$  denotes the total number of individuals. The higher the above metrics, the better the model performance.

Hausdorff Distance (HD) is a measure of similarity or difference between two sets of points, especially the edge difference between two shapes.  $HD_{95}$  uses a 95% quantile and is more robust, while HD is susceptible to outliers. A lower value  $HD_{95}$  indicates a higher degree of similarity and less difference between the two sets of points.  $HD_{95}$  is defined as Equation 8.

$$HD_{95}(G, P) = \max\{d_{95}(G, P), d_{95}(P, G)\} \quad (8)$$

Our experiment uses a five-fold cross-validation, so the above metrics are averaged.

## 2.2.4 Comparison models

We included a variety of comparison models, namely: UNETR++ (Shaker et al., 2024), CoTr (Xie et al., 2021), nnFormer (Zhou et al., 2023), Res-Unet (Diakogiannis et al., 2020), and nnUNet. UNETR++, nnFormer, and CoTr represent successful integrations of CNNs and transformers. These hybrid frameworks leverage the transformer's strong long-sequence modeling capabilities to compensate for CNN's limited receptive field and weak long-range dependencies. They have excellent performance in medical data sets such as brain tumors and organ segmentation. Both Res-Unet and nnUNet are pure CNN-based architectures. Res-Unet introduces residual connections in the encoder-decoder, effectively alleviating the problem of vanishing gradients and making it suitable for small-sample medical image segmentation. nnUNet is an automated medical image segmentation framework. It has achieved excellent performance in multiple medical segmentation challenges and exhibits strong generalization performance. By including pure CNNs models and hybrid models, we can better demonstrate the advantages of our proposed method in the results.

# 3 Results

## 3.1 Implementation details

Our method was implemented on Python 3.9.12 and Pytorch 2.1.1. Experiments were deployed based on the MedNeXt framework, using five-fold cross-validation and the data preprocessing operations built into the MedNeXt framework to perform format conversion, cropping, resampling, and standardization on the data, resulting in a data size of  $128 \times 128 \times 128$  with a voxel size of  $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ . The model training batch size was 2, the initial learning rate was  $10^{-3}$  and the

maximum epoch was 100. Frame size configuration was M, kernel size was 3,  $\lambda$  was 0.5 and the optimizer was AdamW.

For all comparison models, the training batch size, patch size and maximum epoch were the same as those of MedNeXt. The initial learning rate was  $10^{-2}$  and the optimizer was SGD.

## 3.2 Results on open and private datasets

First, we use five indicators to evaluate the predictions of various models, showing the overall and subtype results, respectively. Secondly, we compared the prediction results of different models in terms of individual detection rate and slice (axial, sagittal, and coronal) detection rate, enabling a comprehensive analysis of the results. Finally, we present and discuss the 16 negative cases in the Private Dataset separately.

**Open dataset result:** As listed in Table 2, compared to MedNeXet\_M3, our method improved the DC metric by 2.40%. LEM-UNet also achieved 2.00% of the DC score over the MedNeXet\_L3 which used more convolutional layers. Compared with other CNN-based models, LEM-UNet showed significant improvements in these five metrics. Among transformer-based models, our model exhibited a substantial advantage, with a 5.80% increase in the DC metric over the comparatively high-performing CoTr model. The results indicated that nnFormer and UNETR++ performed poorly in this dataset, struggling to capture the details and characteristic edge of the FCD lesions. In terms of the ability to identify positive class samples, our model excelled, with a notably improved ability to capture FCD lesions details after integrating LEA blocks.

**Private dataset result:** As shown in Table 2, all networks achieved a higher DC score compared to the Open dataset. One reason is that each case in this dataset comprised three modalities of raw data (T1, FLAIR, and FDG-PET), providing the base network with a sufficient amount of rich information. However, our network still demonstrated commendable performance, with a 2.90% improvement in DC metrics over MedNeXet\_M3. In terms of the ability to identify positive class samples, our network remained more prominent compared to other networks, with a 2.90% increase in Rec over MedNeXet\_M3. Secondly, compared to transformer-based networks, LEM-UNet still demonstrated significant advantages, particularly in terms of predictive precision, indicating that our network is more capable of focusing on the subtle details of the lesions.

**Classification results:** In LEM-UNet, the segmentation performance according to histopathology and seizure-free state is shown in Table 3. In the Private Dataset, the DC score for 13 subjects with FCD type I reached 0.655, whereas for 23 subjects with the other MCDs, the DC score was 0.620.

**Detection result:** We analyzed the detection rate results of different models. Table 4 showed the results of the detection rate at the individual and slice level, respectively. LEM-UNet still demonstrated superior performance. One of the major challenges in FCD lesion detection is the occurrence of false-positive clusters that do not have any overlap with the ground truth. Table 5 showed the statistical results on the individual level, which counts the false-positive, true positive, false negative, and ground truth clusters, respectively. All predicted results were subjected

TABLE 2 Performance metrics for different methods on Open and Private datasets.

| Dataset              | Open dataset         |                      |                      |                      |                        |
|----------------------|----------------------|----------------------|----------------------|----------------------|------------------------|
| Methods              | DC↑                  | Pre↑                 | Rec↑                 | IoU↑                 | HD <sub>95</sub> ↓     |
| UNETR++              | 0.367 ± 0.316        | 0.472 ± 0.406        | 0.362 ± 0.331        | 0.271 ± 0.247        | 28.189 ± 33.847        |
| CoTr                 | 0.394 ± 0.318        | 0.466 ± 0.382        | 0.397 ± 0.319        | 0.291 ± 0.244        | 33.997 ± 39.651        |
| nnFormer             | 0.343 ± 0.275        | 0.398 ± 0.345        | 0.376 ± 0.312        | 0.241 ± 0.207        | 59.206 ± 40.810        |
| Res-Unet             | 0.416 ± 0.310        | 0.528 ± 0.387        | 0.403 ± 0.324        | 0.310 ± 0.249        | 20.451 ± 26.830        |
| nnUNet               | 0.403 ± 0.316        | 0.524 ± 0.390        | 0.384 ± 0.328        | 0.300 ± 0.253        | 22.318 ± 29.766        |
| MedNeXt_L3           | 0.432 ± 0.312        | 0.522 ± 0.384        | 0.444 ± 0.338        | 0.325 ± 0.254        | 17.525 ± 27.450        |
| MedNeXt_M3           | 0.428 ± 0.316        | 0.543 ± 0.391        | 0.418 ± 0.333        | 0.323 ± 0.255        | 20.490 ± 26.986        |
| <b>LEM-UNet(our)</b> | <b>0.452 ± 0.310</b> | <b>0.546 ± 0.384</b> | <b>0.447 ± 0.338</b> | <b>0.346 ± 0.263</b> | <b>16.749 ± 26.830</b> |
| Dataset              | Private Dataset      |                      |                      |                      |                        |
| UNETR++              | 0.549 ± 0.269        | 0.569 ± 0.309        | 0.621 ± 0.299        | 0.432 ± 0.238        | 27.138 ± 21.695        |
| CoTr                 | 0.534 ± 0.246        | 0.569 ± 0.303        | 0.590 ± 0.281        | 0.399 ± 0.209        | 27.935 ± 22.156        |
| nnFormer             | 0.512 ± 0.237        | 0.497 ± 0.286        | 0.651 ± 0.261        | 0.376 ± 0.204        | 35.113 ± 24.957        |
| Res-Unet             | 0.556 ± 0.238        | 0.560 ± 0.286        | 0.604 ± 0.268        | 0.419 ± 0.212        | 26.513 ± 20.968        |
| nnUNet               | 0.560 ± 0.252        | 0.584 ± 0.291        | 0.616 ± 0.275        | 0.428 ± 0.227        | 25.846 ± 21.736        |
| MedNeXt_L3           | 0.579 ± 0.239        | 0.624 ± 0.288        | 0.631 ± 0.274        | 0.444 ± 0.221        | 24.851 ± 20.782        |
| MedNeXt_M3           | 0.568 ± 0.250        | 0.616 ± 0.294        | 0.630 ± 0.277        | 0.436 ± 0.230        | 24.535 ± 19.920        |
| <b>LEM-UNet(our)</b> | <b>0.597 ± 0.230</b> | <b>0.626 ± 0.278</b> | <b>0.659 ± 0.253</b> | <b>0.460 ± 0.217</b> | <b>22.763 ± 18.263</b> |

An upward arrow (↑) indicates that a higher value for the metric corresponds to better model performance, while a downward arrow (↓) signifies that a lower value indicates better performance. Bold values indicate the best performance value for each metric.

TABLE 3 Segmentation performance based on pathology and seizure freedom in Open and Private dataset.

|                  | DC↑           | Rec↑          | DC>0.0 (n) | Patients (n) |
|------------------|---------------|---------------|------------|--------------|
| Private Dataset  |               |               |            |              |
| FCD I            | 0.655 ± 0.203 | 0.704 ± 0.236 | 13         | 13           |
| FCD IIA          | 0.605 ± 0.247 | 0.680 ± 0.233 | 31         | 31           |
| FCD IIB          | 0.555 ± 0.239 | 0.650 ± 0.281 | 51         | 52           |
| FCD IIIB         | 0.806         | 0.742         | 1          | 1            |
| MCD              | 0.620 ± 0.217 | 0.591 ± 0.247 | 23         | 23           |
| Other            | 0.677 ± 0.117 | 0.789 ± 0.057 | 5          | 5            |
| Seizure freedom  |               |               |            |              |
| Seizure-free     | 0.600 ± 0.233 | 0.669 ± 0.250 | 117        | 118          |
| Not seizure-free | 0.554 ± 0.158 | 0.444 ± 0.207 | 7          | 7            |
| Open Dataset     |               |               |            |              |
| FCD IIA          | 0.357 ± 0.349 | 0.306 ± 0.319 | 9          | 15           |
| FCD IIB          | 0.578 ± 0.254 | 0.554 ± 0.282 | 27         | 31           |
| NA               | 0.375 ± 0.339 | 0.410 ± 0.377 | 21         | 32           |
| Seizure freedom  |               |               |            |              |
| Seizure-free     | 0.538 ± 0.293 | 0.491 ± 0.296 | 29         | 34           |
| Not seizure-free | 0.380 ± 0.323 | 0.362 ± 0.329 | 5          | 8            |
| NA               | 0.386 ± 0.338 | 0.424 ± 0.383 | 23         | 36           |

“Other” includes TSC, Gliosis, and MOGHE pathologies.

to the same connected component post-processing procedure. If retaining only the largest connected component did not yield an improvement in the DC score, the original predictions were preserved, and no non-largest connected components were removed. Independent connected components were identified using the standard 3D 26-connectivity criterion. The results demonstrate that our method exhibits strong performance in identifying true positive clusters, achieving the fewest false-positive clusters on the Open Dataset. In contrast, nnFormer predicted a relatively larger number of false-positive clusters on both datasets.

**MRI-negative result:** We included a discussion on MRI-negative patients. Table 6 displayed the results of different models on 16 MRI-negative patients from Private Dataset. These 16 patients appeared positive on PET. Most models experienced varying degrees of performance degradation on these difficult-to-identify MRI-negative patients. For example, the nnUNet’s DC score dropped by 6.00%, and the Res-Unet’s DC score dropped by 2.00%. However, the hybrid transformer architecture model appeared relatively stable. Our method showed only a slight performance degradation and fully identified all 16 MRI-negative patients.

### 3.3 Visualization comparison and discussion

The visualization results of two datasets are illustrated in Figure 5. The first two rows display the visualization results on the Private Dataset, while the last two rows show the visualization

TABLE 4 Comparison of individual detection rates and average detection rates across the three directional slices between LEM-UNet and other models across two Datasets.

| Dataset       | Individual   |              |                 |              | Slice        |              |                 |              |
|---------------|--------------|--------------|-----------------|--------------|--------------|--------------|-----------------|--------------|
|               | Open Dataset |              | Private Dataset |              | Open Dataset |              | Private Dataset |              |
|               | DC>0.0       | DC>0.22      | DC>0.0          | DC>0.22      | DC>0.0       | DC>0.22      | DC>0.0          | DC>0.22      |
| UNETR++       | 64.10        | 58.97        | 92.80           | 87.20        | 51.97        | 47.44        | 76.67           | 69.62        |
| CoTr          | <b>73.08</b> | 64.10        | 92.80           | 84.80        | 58.49        | 51.82        | 76.25           | 69.04        |
| nnFormer      | 70.51        | 62.82        | 97.60           | 85.60        | 55.41        | 50.25        | 81.15           | 70.91        |
| Res-UNet      | 71.79        | 66.67        | 96.00           | 88.00        | 58.27        | 52.15        | 78.58           | 70.76        |
| nnUNet        | 70.51        | 62.82        | 96.80           | 85.60        | 55.41        | 50.08        | 78.90           | 70.93        |
| MedNeXt_L3    | <b>73.08</b> | <b>67.95</b> | 97.60           | <b>91.20</b> | 61.06        | 56.18        | 78.63           | 71.88        |
| MedNeXt_M3    | 71.79        | 66.67        | 96.80           | 88.80        | 58.97        | 53.22        | 78.32           | 71.03        |
| LEM-UNet(our) | <b>73.08</b> | <b>67.95</b> | <b>99.20</b>    | 89.60        | <b>61.45</b> | <b>56.78</b> | <b>81.22</b>    | <b>74.16</b> |

Bold values indicate the best performance value for each metric.

TABLE 5 Comparisons of the number of ground GT lesions, TP lesions, FP lesions, and FN lesions among models different on two Datasets.

|                 | TP lesions | FP lesions | FN lesions | GT lesions |
|-----------------|------------|------------|------------|------------|
| Private Dataset |            |            |            |            |
| UNETR++         | 116        | 42         | 9          | 125        |
| CoTr            | 116        | <b>9</b>   | 9          | 125        |
| nnFormer        | 122        | 101        | 3          | 125        |
| Res-UNet        | 123        | 31         | 2          | 125        |
| nnUNet          | 121        | 34         | 4          | 125        |
| MedNeXt_L3      | 122        | 23         | 3          | 125        |
| MedNeXt_M3      | 121        | 57         | 4          | 125        |
| LEM-UNet(our)   | <b>124</b> | 17         | <b>1</b>   | 125        |
| Open Dataset    |            |            |            |            |
| UNETR++         | 50         | 45         | 28         | 78         |
| CoTr            | <b>57</b>  | 78         | <b>21</b>  | 78         |
| nnFormer        | 55         | 109        | 23         | 78         |
| Res-UNet        | 56         | 30         | 22         | 78         |
| nnUNet          | 55         | 37         | 23         | 78         |
| MedNeXt_L3      | <b>57</b>  | 23         | <b>21</b>  | 78         |
| MedNeXt_M3      | 56         | 20         | 22         | 78         |
| LEM-UNet(our)   | <b>57</b>  | <b>13</b>  | <b>21</b>  | 78         |

TP, true positive; FP, false-positive; FN, false negative; GT, ground truth. Bold values indicate the best performance value for each metric.

results on the Open Dataset. The ground truth is set to white, while the model's predicted results are set to red and superimposed upon the ground truth for comparison. We observe that transformer-based models exhibit a significantly lower ability to capture positive samples in this task compared to CNN-based models. They struggle to correctly identify FCD lesions with blurred edges and subtle

features, even resulting in missed detections, as seen in the CoTr prediction results. By comparing all visualized outcomes, LEM-UNet demonstrates superior performance in capturing the lesion details. Furthermore, our network, by enhancing the focus on edge information, becomes more sensitive to lesion edges, increasing the number of true positive identifications while reducing the likelihood of false positives.

The heat map shown in Figure 6 further illustrates the network focus when processing features. In the LEA block, it has a stronger focus on the edges of the lesion, with a color gradient from blue to red indicating low to high intensity.

### 3.4 Ablation study

In this section, we underscore the significance of each proposed block by conducting a series of ablation studies. Before proceeding with the ablation experiments, a brief discussion will be held on the initialization value of the learnable parameter  $\lambda$ . The experiments will be conducted on the Private Dataset, and the results are shown in Table 7.

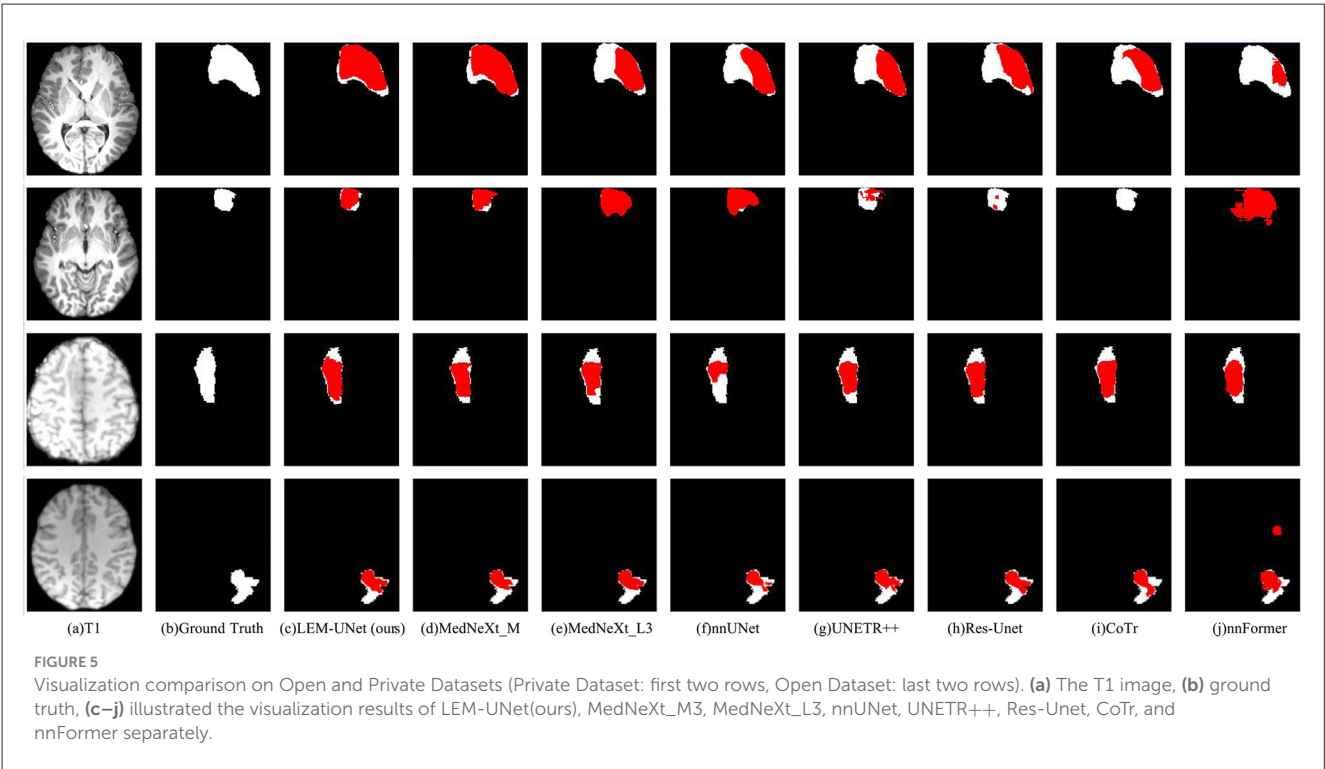
Performance is evaluated based on three metrics: DC, Pre, and Rec. It is observed that when  $\lambda = 0.5$ , the model exhibited superior comprehensive performance, with a higher Pre and DC score compared to when  $\lambda = 0.3$  or  $\lambda = 0.7$ . Next, we proceed with the deployment of ablation studies. Within the context of the LEM-UNet framework, the LEA and MFF blocks are of paramount importance. Given that one of the inputs to the MFF block is the output of the LEA block, this paper evaluates the segmentation outcomes when the MFF and LEA blocks are sequentially removed, in order to validate the performance of each block. We conduct these assessments on the MedNeXt\_M3 baseline with a parameter setting of  $\lambda = 0.5$ , using the Private Dataset, and present the results, as detailed in Table 8.

In LEM-UNet, the incorporation of the edge attention module improves model capabilities but also introduces additional resource overhead. We conducted a comprehensive comparison of the

TABLE 6 The evaluation results of MRI-negative patients on different models.

| Methods       | DC↑                  | Pre↑                 | Rec↑                 | IoU↑                 | HD <sub>95</sub> ↓     | DC>0.0    | FP lesions |
|---------------|----------------------|----------------------|----------------------|----------------------|------------------------|-----------|------------|
| UNETR++       | 0.551 ± 0.244        | 0.586 ± 0.301        | 0.602 ± 0.272        | 0.415 ± 0.221        | 32.477 ± 30.586        | 15        | 18         |
| CoTr          | 0.533 ± 0.224        | 0.559 ± 0.280        | 0.605 ± 0.298        | 0.388 ± 0.179        | 33.176 ± 26.552        | 15        | 2          |
| nnFormer      | 0.507 ± 0.210        | 0.529 ± 0.301        | <b>0.633 ± 0.211</b> | 0.363 ± 0.176        | 42.413 ± 30.991        | <b>16</b> | 34         |
| Res-Unet      | 0.536 ± 0.235        | 0.572 ± 0.276        | 0.563 ± 0.287        | 0.394 ± 0.186        | 35.872 ± 31.909        | 14        | 5          |
| nnUNet        | 0.500 ± 0.224        | 0.540 ± 0.289        | 0.549 ± 0.290        | 0.360 ± 0.192        | 35.380 ± 24.813        | <b>16</b> | 11         |
| MedNeXt_L3    | 0.555 ± 0.248        | 0.594 ± 0.288        | 0.594 ± 0.284        | 0.419 ± 0.224        | 32.904 ± 31.399        | <b>16</b> | 10         |
| MedNeXt_M3    | 0.560 ± 0.196        | 0.613 ± 0.254        | 0.607 ± 0.233        | 0.413 ± 0.192        | 33.912 ± 30.913        | <b>16</b> | 12         |
| LEM-UNet(our) | <b>0.591 ± 0.199</b> | <b>0.673 ± 0.267</b> | 0.613 ± 0.236        | <b>0.445 ± 0.193</b> | <b>27.542 ± 26.472</b> | <b>16</b> | 3          |

Bold values indicate the best performance value for each metric.



parameter counts and computational complexity between LEM-UNet and other models, as shown in Table 9. Compared to the baseline model MedNeXt\_M3, LEM-UNet exhibits higher computational complexity but achieves superior segmentation performance. Our approach strikes a better balance between computational resource utilization and performance enhancement.

4 Discussion

The segmentation of FCD lesions is highly complex, as their characteristics vary across different imaging modalities and are often characterized by blurred edge, posing significant challenges to the segmentation task. The success of FCD lesion resection surgery heavily depends on the precision of the resection area. Therefore, we propose integrating edge enhancement methods

into convolutional networks, aiming to improve segmentation performance by enhancing high-frequency edge information of lesions, thereby providing a reliable reference for the clinical diagnosis of FCD lesions.

From the experimental results, LEM-UNet demonstrates superior performance, outperforming comparison models across multiple key metrics, as shown in Table 2. Furthermore, our method exhibits significant advantages in individual and slice detection rates, achieving 99.20% and 81.22% (DC > 0.0) on the Private Dataset, respectively. Based on visualization of feature maps, the LEA block demonstrates the ability to capture high-frequency edge information of lesions, as illustrated in Figure 6, although these edge cues are not entirely closed. According to the ablation study, the integration of the Laplacian pyramid with CBAM effectively enhances the network's focus on high-frequency edge information of FCD lesions, which is one of the key

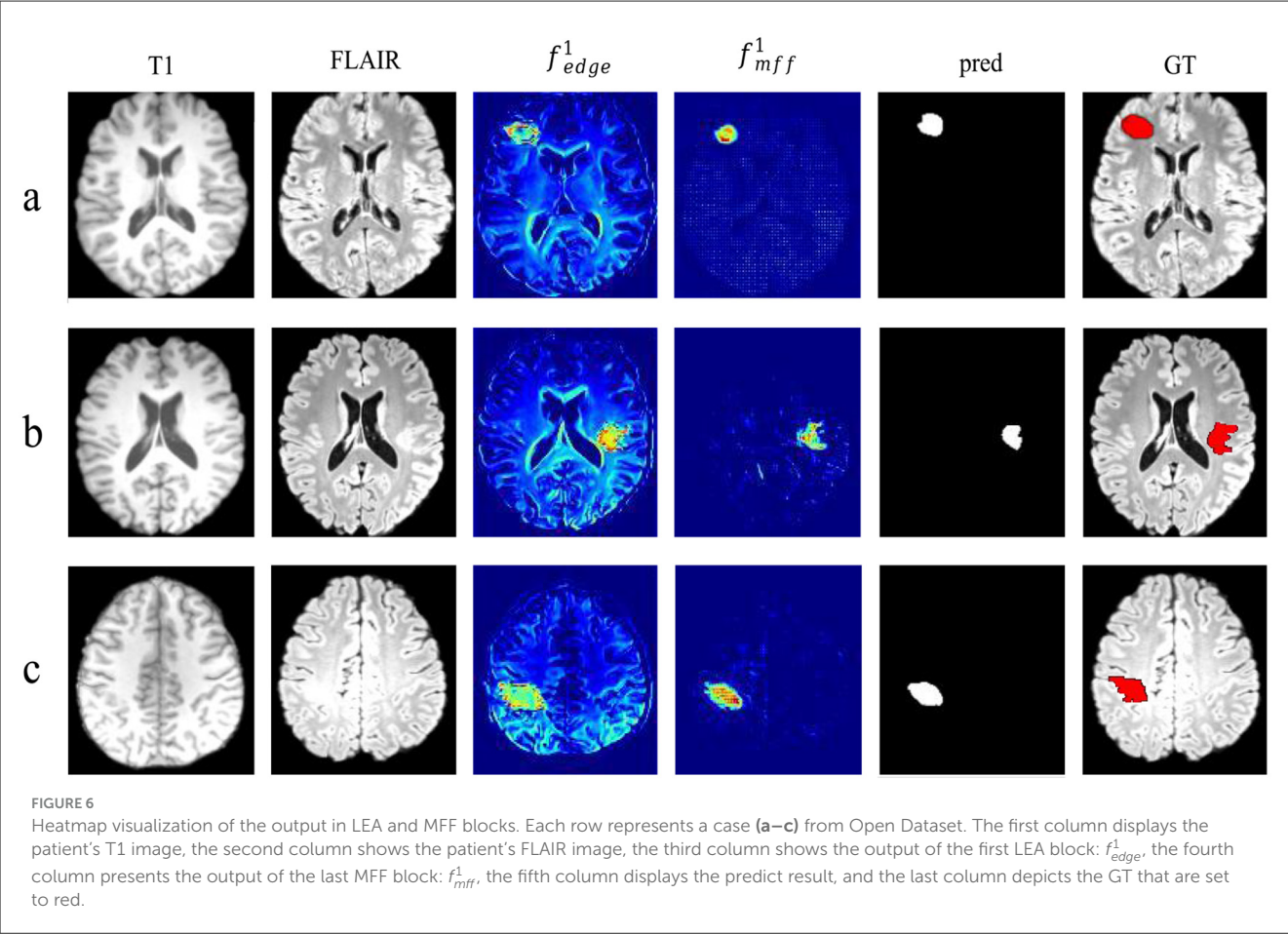


TABLE 7  $\lambda$  initialization in LEM-UNet on Private Dataset.

| $\lambda$ | DC $\uparrow$ | Pre $\uparrow$ | Rec $\uparrow$ |
|-----------|---------------|----------------|----------------|
| 0.7       | 0.589         | 0.631          | 0.635          |
| 0.5       | <b>0.597</b>  | 0.626          | <b>0.659</b>   |
| 0.3       | 0.591         | <b>0.634</b>   | 0.637          |

Bold values indicate the best performance value for each metric.

TABLE 8 Ablation study of proposed blocks for LEM-UNet on Private Dataset.

| Baseline | LEA block | MFF block | DC $\uparrow$ |
|----------|-----------|-----------|---------------|
| ✓        |           |           | 0.568         |
| ✓        | ✓         |           | 0.588         |
| ✓        | ✓         | ✓         | 0.597         |

factors contributing to the improved segmentation performance of LEM-UNet.

Deep learning-based methods have demonstrated promising performance in FCD lesion segmentation tasks (e.g., Thomas et al., 2020; Zhang X. et al., 2024; Niyas et al., 2021). However, compared to the excellent segmentation results in other tasks, such as brain tumor segmentation and polyp segmentation, there

TABLE 9 Comparison of computational complexity for different models on Private Dataset.

| Method         | Parameters (M) | GFLOPs | DC $\uparrow$ |
|----------------|----------------|--------|---------------|
| UNETR++        | 42.64          | 72.52  | 0.549         |
| CoTr           | 41.91          | 689.97 | 0.534         |
| nnFormer       | 37.67          | 63.68  | 0.512         |
| Res-Unet       | 12.08          | 678.97 | 0.556         |
| nnUNet         | 31.20          | 482.69 | 0.560         |
| MedNeXt_L3     | 61.78          | 500.00 | 0.579         |
| MedNeXt_M3     | 17.54          | 254.44 | 0.568         |
| LEM-UNet (our) | 17.66          | 336.83 | 0.597         |

is still under-segmentation, which may be closely related to the complexity of FCD lesions. Previous approaches primarily focused on feature fusion and enhancement, contributing to addressing segmentation challenges such as semantic gaps, spatial information loss, and limited receptive fields. However, they often overlooked edge features that are particularly meaningful for segmentation tasks. To address this, we introduced the Laplacian pyramid to process lesion edges, leveraging the characteristics of FCD lesions to reduce the false positive rate

in segmentation results. By incorporating edge analysis tailored to FCD lesion features, LEM-UNet maximizes the capture of morphological characteristics of lesion edges while focusing on deep abstract features, which is a key advantage of our method over existing algorithms.

Nevertheless, the proposed method still suffers from the inherent limitations of deep learning frameworks, such as its dependence on the accuracy of training data annotations. Additionally, the small dataset size restricts the model's generalizability and robustness. We performed five-fold cross-validation on both Open and Private Datasets and conducted a series of optimization experiments. However, due to the limited availability of FCD data, our method did not use scan data from another center as an independent test set to test the model performance, which is another limitation. Future work will integrate multi-center data to comprehensively enhance the model's generalization capability.

## 5 Conclusion

In this study, we introduce LEM-UNet, an advanced framework for 3D FCD lesions segmentation in multimodal medical images, fortified by the innovative LEA and MFF blocks. We designed the LEA block, which uses a 3D Laplacian pyramid to capture shallow edge information in images, combined with the CBAM to achieve precise localization from global information to local lesion information, further guiding the network to focus on lesion edges and detailed information. Additionally, we designed the MFF block, which combines high-frequency edge features with advanced prediction features, overcoming semantic differences between feature maps while further eliminating redundant information. We deploy these designed modules in the MedNeXt\_M3 framework, and the results evaluated through five metrics demonstrate that our method provides competitive performance, with overall metrics superior to other SOTA methods. This emphasizes the effectiveness of our proposed framework in the field of medical image segmentation.

## Data availability statement

The datasets presented in this article are not readily available because our Private Dataset is restricted by ethical requirements. Requests to access the datasets should be directed to Open Dataset: <https://openneuro.org/datasets/ds004199/>.

## Ethics statement

The studies involving humans were approved by the Ethics Committee of Peking University First Hospital (No. 102627). Written informed consent was obtained from the parents of all participants, allowing the use of their children's data for research purposes. The studies were conducted

in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

## Author contributions

QL: Investigation, Conceptualization, Writing – review & editing, Methodology, Validation, Formal analysis, Visualization, Data curation, Software, Writing – original draft. HY: Writing – original draft, Data curation, Validation, Methodology, Writing – review & editing, Formal analysis. MZ: Writing – review & editing, Methodology, Formal analysis, Data curation. XY: Writing – review & editing, Data curation. LC: Resources, Validation, Writing – review & editing, Conceptualization, Supervision, Methodology. GK: Methodology, Supervision, Validation, Conceptualization, Writing – review & editing, Funding acquisition.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by the State Key Program of the National Natural Science Foundation of China (82030037).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Avakyan, G., Blinov, D., Lebedeva, A., Burd, S., and Avakyan, G. (2017). Ilae classification of the epilepsies: the 2017 revision and update. *Epilepsy Paroxysmal Condit.* 9, 6–25. doi: 10.17749/2077-8333.2017.9.1.006-025
- Baete, K., Nuyts, J., Van Paesschen, W., Suetens, P., and Dupont, P. (2004). Anatomical-based fdg-pet reconstruction for the detection of hypo-metabolic regions in epilepsy. *IEEE Trans. Med. Imaging* 23, 510–519. doi: 10.1109/TMI.2004.825623
- Burt, P. J., and Adelson, E. H. (1987). “The laplacian pyramid as a compact image code,” in *Readings in Computer Vision* (Elsevier), 671–679. doi: 10.1016/B978-0-08-051581-6.50065-9
- David, B., Kröll-Seger, J., Schuch, F., Wagner, J., Wellmer, J., Woermann, F., et al. (2021). External validation of automated focal cortical dysplasia detection using morphometric analysis. *Epilepsia* 62, 1005–1021. doi: 10.1111/epi.16853
- Dev, K. B., Jogi, P. S., Niyas, S., Vinayagamani, S., Kesavadas, C., and Rajan, J. (2019). Automatic detection and localization of focal cortical dysplasia lesions in MRI using fully convolutional neural network. *Biomed. Signal Process. Control* 52, 218–225. doi: 10.1016/j.bspc.2019.04.024
- Diakogiannis, F. I., Waldner, F., Caccetta, P., and Wu, C. (2020). Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogram. Rem. Sens.* 162, 94–114. doi: 10.1016/j.isprsjprs.2020.01.013
- Fisher, R. S., Boas, W. V. E., Blume, W., Elger, C., Genton, P., Lee, P., et al. (2005). Epileptic seizures and epilepsy: definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (IBE). *Epilepsia* 46, 470–472. doi: 10.1111/j.0013-9580.2005.66104.x
- Gill, R. S., Lee, H.-M., Caldirou, B., Hong, S.-J., Barba, C., Deleo, F., et al. (2021). Multicenter validation of a deep learning detection algorithm for focal cortical dysplasia. *Neurology* 97, e1571–e1582. doi: 10.1212/WNL.000000000000102698
- Hong, S.-J., Kim, H., Schrader, D., Bernasconi, N., Bernhardt, B. C., and Bernasconi, A. (2014). Automated detection of cortical dysplasia type II in MRI-negative epilepsy. *Neurology* 83, 48–55. doi: 10.1212/WNL.0000000000000543
- Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., et al. (2018). NNU-net: self-adapting framework for u-net-based medical image segmentation. *arXiv preprint arXiv:1809.10486*.
- Jin, B., Krishnan, B., Adler, S., Wagstyl, K., Hu, W., Jones, S., et al. (2018). Automated detection of focal cortical dysplasia type ii with surface-based magnetic resonance imaging postprocessing and machine learning. *Epilepsia* 59, 982–992. doi: 10.1111/epi.14064
- Kabat, J., and Król, P. (2012). Focal cortical dysplasia-review. *Polish J. Radiol.* 77:35. doi: 10.12659/PJR.882968
- Lu, L., Chen, S., Tang, H., Zhang, X., and Hu, X. (2023). A multi-scale perceptual polyp segmentation network based on boundary guidance. *Image Vis. Comput.* 138:104811. doi: 10.1016/j.imavis.2023.104811
- Mechelli, A., Price, C. J., Friston, K. J., and Ashburner, J. (2005). Voxel-based morphometry of the human brain: methods and applications. *Curr. Med. Imag.* 1, 105–113. doi: 10.2174/15734050504038726
- Niyas, S., Vaisali, S. C., Show, I., Chandrika, T., Vinayagamani, S., Kesavadas, C., et al. (2021). Segmentation of focal cortical dysplasia lesions from magnetic resonance images using 3D convolutional neural networks. *Biomed. Signal Process. Control* 70:102951. doi: 10.1016/j.bspc.2021.102951
- Roy, S., Koehler, G., Ulrich, C., Baumgartner, M., Petersen, J., Isensee, F., et al. (2023). “Mednext: transformer-driven scaling of convnets for medical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer), 405–415. doi: 10.1007/978-3-031-43901-8\_39
- Schuch, F., Walger, L., Schmitz, M., David, B., Bauer, T., Harms, A., et al. (2023). An open presurgery MRI dataset of people with epilepsy and focal cortical dysplasia type II. *Sci. Data* 10:475. doi: 10.1038/s41597-023-02386-7
- Shaker, A. M., Maaz, M., Rasheed, H., Khan, S., Yang, M.-H., and Khan, F. S. (2024). Unetr++: delving into efficient and accurate 3D medical image segmentation. *IEEE Trans. Med. Imag.* 43, 3377–3390. doi: 10.1109/TMI.2024.3398728
- Spitzer, H., Ripart, M., Fawaz, A., Williams, L. Z., project, M., Robinson, E. C., et al. (2023). “Robust and generalisable segmentation of subtle epilepsy-causing lesions: a graph convolutional approach,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer), 420–428. doi: 10.1007/978-3-031-43993-3\_41
- Spitzer, H., Ripart, M., Whitaker, K., D’Arco, F., Mankad, K., Chen, A. A., et al. (2022). Interpretable surface-based detection of focal cortical dysplasias: a multi-centre epilepsy lesion detection study. *Brain* 145, 3859–3871. doi: 10.1093/brain/awac224
- Thesen, T., Quinn, B. T., Carlson, C., Devinsky, O., DuBois, J., McDonald, C. R., et al. (2011). Detection of epileptogenic cortical malformations with surface-based MRI morphometry. *PLoS ONE* 6:e16430. doi: 10.1371/journal.pone.0016430
- Thomas, E., Pawan, S., Kumar, S., Horo, A., Niyas, S., Vinayagamani, S., et al. (2020). Multi-res-attention uNet: a CNN model for the segmentation of focal cortical dysplasia lesions from magnetic resonance images. *IEEE J. Biomed. Health Inf.* 25, 1724–1734. doi: 10.1109/JBHI.2020.3024188
- Urbach, H., Kellner, E., Kremers, N., Blümcke, I., and Demerath, T. (2022). MRI of focal cortical dysplasia. *Neuroradiology* 64, 443–452. doi: 10.1007/s00234-021-02865-x
- Wagstyl, K., Whitaker, K., Raznahan, A., Seidlitz, J., Vértes, P. E., Foldes, S., et al. (2022). Atlas of lesion locations and postsurgical seizure freedom in focal cortical dysplasia: a meld study. *Epilepsia* 63, 61–74. doi: 10.1111/epi.17130
- Walger, L., Bauer, T., Kügler, D., Schmitz, M. H., Schuch, F., Arendt, C., et al. (2024). “Bridging the gap between human and artificial intelligence—an evaluation framework for computer-aided detection of brain lesions”, in *SSRN Scholarly Paper* (Rochester, NY). doi: 10.2139/ssrn.4692599
- Wang, H., Ahmed, S. N., and Mandal, M. (2020). Automated detection of focal cortical dysplasia using a deep convolutional neural network. *Comput. Med. Imag. Graph.* 79:101662. doi: 10.1016/j.compmedimag.2019.101662
- Willard, A., Antonic-Baker, A., Chen, Z., O’Brien, T. J., Kwan, P., and Perucca, P. (2022). Seizure outcome after surgery for MRI-diagnosed focal cortical dysplasia: a systematic review and meta-analysis. *Neurology* 98, e236–e248. doi: 10.1212/WNL.0000000000013066
- Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). “CBAM: convolutional block attention module,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, 3–19. doi: 10.1007/978-3-030-01234-2\_1
- Xie, Y., Zhang, J., Shen, C., and Xia, Y. (2021). “COTR: efficiently bridging CNN and transformer for 3D medical image segmentation,” in *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021: 24th International Conference, Strasbourg, France, September 27-October 1, 2021, Proceedings, Part III 24* (Springer), 171–180. doi: 10.1007/978-3-030-87199-4\_16
- Zhang, S., Zhuang, Y., Luo, Y., Zhu, F., Zhao, W., and Zeng, H. (2024). Deep learning-based automated lesion segmentation on pediatric focal cortical dysplasia ii preoperative MRI: a reliable approach. *Insights Imaging* 15:71. doi: 10.1186/s13244-024-01635-6
- Zhang, X., Zhang, Y., Wang, C., Li, L., Zhu, F., Sun, Y., et al. (2024). Focal cortical dysplasia lesion segmentation using multiscale transformer. *Insights Imaging* 15:222. doi: 10.1186/s13244-024-01803-8
- Zhou, H.-Y., Guo, J., Zhang, Y., Han, X., Yu, L., Wang, L., et al. (2023). NNFORMER: volumetric medical image segmentation via a 3d transformer. *IEEE Trans. Image Proc.* 32, 4036–4045. doi: 10.1109/TIP.2023.3293771
- Zhu, Z., He, X., Qi, G., Li, Y., Cong, B., and Liu, Y. (2023). Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI. *Inf. Fusion* 91, 376–387. doi: 10.1016/j.inffus.2022.10.022