DeDA: Deep Directed Accumulator

Hang Zhang[1][✉], Rongguang Wang[2], Renjiu Hu[1], Jinwei Zhang[1],
and Jiahao Li[1]

[1] Cornell University, Ithaca, USA
hz459@cornell.edu
[2] University of Pennsylvania, Philadelphia, USA

Abstract. Chronic active multiple sclerosis lesions, also referred to as rim+ lesions, are characterized by a hyperintense rim observed at the lesion's edge on quantitative susceptibility maps. Despite their geometrically simple structure, characterized by radially oriented gradients at the lesion edge with a greater gradient magnitude compared to non-rim+ (rim-) lesions, recent studies indicate that the identification performance for these lesions is subpar due to limited data and significant class imbalance. In this paper, we propose a simple yet effective image processing operation, deep directed accumulator (DeDA), which provides a new perspective for injecting domain-specific inductive biases (priors) into neural networks for rim+ lesion identification. Given a feature map and a set of sampling grids, DeDA creates and quantizes an accumulator space into finite intervals and accumulates corresponding feature values. This DeDA operation can be regarded as a symmetric operation to the grid sampling within the forward-backward neural network framework, the process of which is order-agnostic, and can be efficiently implemented with the native CUDA programming. Experimental results on a dataset with 177 rim+ and 3986 rim- lesions show that 10.1% of improvement in a partial (false positive rate < 0.1) area under the receiver operating characteristic curve (pROC AUC) and 10.2% of improvement in an area under the precision recall curve (PR AUC) can be achieved respectively comparing to other state-of-the-art methods. The source code is available online at https://github.com/tinymilky/DeDA.

Keywords: Directed accumulator · Neural networks · Multiple sclerosis · Quantitative susceptibility mapping

1 Introduction

Over the past decade, we have observed the substantial success of Convolutional Neural Networks (CNNs) in a multitude of grid-based medical imaging applications, such as magnetic resonance imaging (MRI) reconstruction [20,27] and lesion segmentation [13,25]. Despite the effectiveness of general inductive biases like translation equivariance [15] and locality [16], the diverse nature of
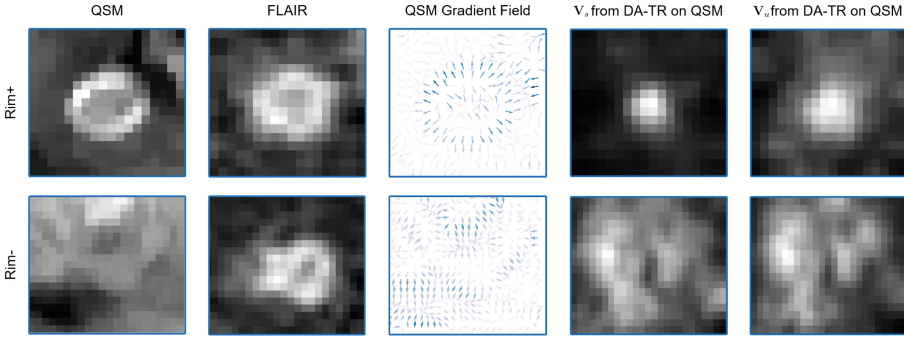
Fig. 1. Visual illustration of the difference between a rim+ and a rim- lesion. QSM image patches represent the lesion magnetic susceptibility, while Fluid Attenuated Inversion Recovery (FLAIR) image patches pinpoint the exact location of the lesions. The gradient field map of the QSM images presents normalized gradient vectors (the darker the blue, the larger the gradient vector's magnitude). The two rightmost columns display gradient magnitude maps $\mathbf{V}_s$ and QSM value maps $\mathbf{V}_u$ processed by DA-TR (see Sect. 2.2). Notably, the rim+ lesion exhibits structured patterns in the accumulator space by aggregating feature values along gradients, a characteristic absent in the rim- lesion. (Color figure online)

diseases represented in medical images necessitates highly domain-specific knowledge. Consequently, the question of how to incorporate domain-specific inductive biases, or priors, beyond general ones into neural networks for medical image processing remains an open challenge.

In this study, we strive to answer this question by addressing the identification problem associated with a specific type of multiple sclerosis (MS) lesion, referred to as a chronic active lesion, or rim+ lesion. Histopathology studies characterize rim+ lesions by an iron-rich rim of activated macrophages and microglia [2,6,9,14]. These lesions are visible with in-vivo quantitative susceptibility mapping (QSM) [7,22] and phase imaging techniques [1,2]. Notably, they display a paramagnetic hyperintense rim at the edge (see Fig. 1). Despite several efforts to tackle the issue [4,18,24], a clinically reliable solution remains elusive.

Given the limited amount of data and high class imbalance, it's more advantageous to explicitly incorporate domain knowledge into the network as priors. As illustrated in Fig. 1, rim+ lesions distinguish themselves from rim- lesions in three primary aspects. Firstly, rim+ lesions exhibit a hyperintense ring-like structure at the lesion's edge on QSM. Secondly, a higher magnitude of gradients is observed near the edge of rim+ lesions, a feature not present in rim- lesions. Lastly, rim+ lesions are characterized by radially oriented gradients at the edge, whereas rim- lesions lack such structured orientations.

In this work, we introduce the **De**ep **D**irected **A**ccumulator (DeDA), a novel image processing operation. DeDA, symmetric to grid sampling within a neural network's forward-backward framework, explicitly encodes the aforementioned prior information. Given a feature map and sampling grids, DeDA creates an accumulator space, quantizes it into finite intervals, and accumulates feature

values. DeDA can also be viewed as a generalized discrete Radon transform, as it accumulates values between two discrete feature spaces. Our contributions are twofold: Firstly, we present DeDA, a simple yet powerful method that augments neural networks' representation capacity by explicitly incorporating domain-specific priors. Secondly, our experimental results on rim+ lesion identification demonstrate a notable improvement of 10.1% in partial area under the receiver operating characteristic curve (pROC AUC) and a 10.2% improvement in area under the precision recall curve (PR AUC), outperforming existing state-of-the-art methods.

2 Methodology

Numerous signal processing techniques, including the Fourier transform, Radon transform, and Hough transform, map discrete signals from image space to another functional space. We call this new space accumulator space, where each cell's value in the new space constitutes a weighted sum of values from all cells in the original image space. For our purposes, an appealing feature of the accumulator space is that local convolutions within it, like those in Hough and sinogram spaces, result in global aggregation of structural features, such as lines, in the feature map space. This proves beneficial for incorporating geometric priors into neural networks. Differing from attention-based methods, this convolution in accumulator space explicitly captures long-range information through direct geometric prior parameterization.

2.1 Differentiable Directed Accumulation

The process of transforming an image to an accumulator space involves a critical step, directed accumulation (DA), in which a cell from the accumulator space is pointed by multiple cells from the image space. Figure 2, Eq. (1) and Eq. (3) have shown that this DA operation is a symmetric operation to the grid sampling [12] within the forward-backward learning framework, where the backward pass of DA possesses the same structure as the forward pass of grid sampling if only one sampling grid is given, and vice versa for the forward pass. In addition, DA is further generalized to allow multiple sampling grids to accumulate values from the source feature map. Here we briefly review the grid sampling method and then derive the proposed DeDA.

Grid Sampling: Given a source feature map $\mathbf{U} \in \mathbb{R}^{C \times H \times W}$, a sampling grid $\mathbf{G} \in \mathbb{R}^{2 \times H' \times W'} = (\mathbf{G}^x, \mathbf{G}^y)$ specifying pixel locations to read from $\mathbf{U}$, and a kernel function $\mathcal{K}()$ defining the image interpolation, then the output value of a particular position (i, j) at the target feature map $\mathbf{V} \in \mathbb{R}^{C \times H' \times W'}$ can be written as follows:

$$\mathbf{V}_{ij}^c = \sum_n^H \sum_m^W \mathbf{U}_{nm}^c \mathcal{K}(\mathbf{G}_{ij}^x, n) \mathcal{K}(\mathbf{G}_{ij}^y, m), \qquad (1)$$
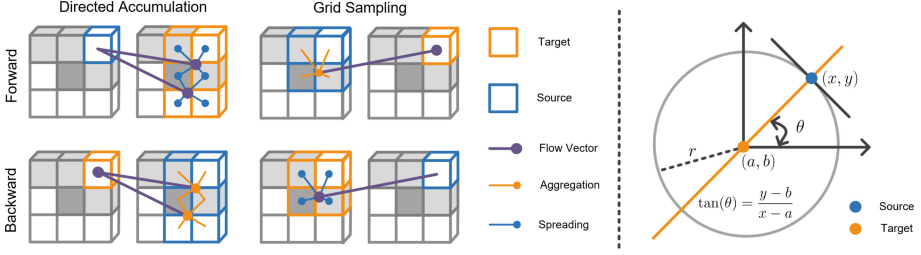
Fig. 2. Visual illustration of the proposed method. The left panel shows differences between the grid sampling and the proposed DeDA using bilinear sampling kernel. The right panel shows the schematic for rim parameterization, where the knowledge of a triple (x, y, θ) is mapped to a straight line (marked in orange) in the accumulator space. (Color figure online)

where the kernel function $\mathcal{K}()$ can be replaced with any other specified kernels, e.g. integer sampling kernel $\delta(\lfloor \mathbf{G}_{ij}^x + 0.5 \rfloor - n) \cdot \delta(\lfloor \mathbf{G}_{ij}^y + 0.5 \rfloor - m)$. Here $\lfloor x + 0.5 \rfloor$ rounds x to the nearest integer and $\delta()$ is the Kronecker delta function. The gradients with respect to $\mathbf{U}$ and $\mathbf{G}$ for back propagation can be defined accordingly [12].

DeDA: Given a source feature map $\mathbf{U} \in \mathbb{R}^{C \times H \times W}$, a target feature map $\mathbf{V} \in \mathbb{R}^{C \times H' \times W'}$, a set of sampling grids $\mathcal{G} = \{\mathbf{G}[k] \in \mathbb{R}^{2 \times H \times W} = (\mathbf{G}^x[k], \mathbf{G}^y[k]) \mid k \in \mathbb{Z}^+, 1 \leq k \leq N\}$ ($N \geq 1$ denotes the number of grids), and a kernel function $\mathcal{K}()$, the output value of a particular position (i, j) at the target feature map $\mathbf{V}$ can be written as follows:

$$\mathbf{V}_{ij}^c = \sum_k^N \sum_n^H \sum_m^W \mathbf{U}_{nm}^c \mathcal{K}(\mathbf{G}_{nm}^x[k], i) \mathcal{K}(\mathbf{G}_{nm}^y[k], j). \tag{2}$$

It is worth noting that the spatial dimension of the grid $\mathbf{G}[k]$ should be the same as that of $\mathbf{U}$, but the first dimension of $\mathbf{G}[k]$ can be an arbitrary number as long as it aligns with the number of spatial dimensions of $\mathbf{V}$, e.g. if given $\mathbf{U} \in \mathbb{R}^{H \times W}$ and $\mathbf{G}[k] \in \mathbb{R}^{3 \times H \times W}$, it is expected that $\mathbf{V} \in \mathbb{R}^{H' \times W' \times D'}$. Basically, the DeDA operation in Eq. (2) performs a tensor mapping by $\mathcal{D} : (\mathbf{U}, \mathcal{G}; \mathcal{K}) \to \mathbf{V}$, where $\mathcal{K}$ is the sampling kernel. For simplicity, function $\mathcal{D}()$ will be used to denote the DeDA forward for the rest of the paper.

To allow back propagation for training networks with DeDA, the gradients with respect to $\mathbf{U}$ are derived using the chain rule as follows:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{V}_{nm}^c} \frac{\partial \mathbf{V}_{nm}^c}{\partial \mathbf{U}_{ij}^c} = \sum_k^N \sum_n^{H'} \sum_m^{W'} \mathbf{A}_{nm}^c \mathcal{K}(\mathbf{G}_{ij}^x[k], n) \mathcal{K}(\mathbf{G}_{ij}^y[k], m). \tag{3}$$

The gradient tensor with respect to $\mathbf{V}$ is $\mathbf{A}$. The structure of Eq. (3) reduces to Eq. (1) when $N = 1$, indicating DeDA's symmetry with grid sampling. Given identical transformations for each channel c in DeDA's forward and backward passes, we denote the feature map with spatial dimensions alone henceforth.
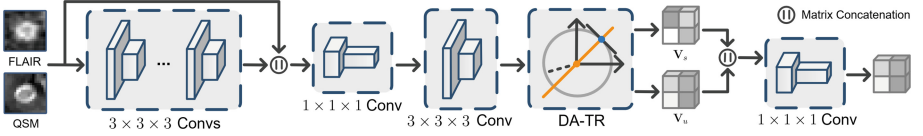
Fig. 3. Schematic of the network layer for DA-TR. Conv denotes a convolutional layer, and each of these layers consists a $3\times3\times3$ or $1\times1\times1$ convolution, a batch normalization, and a ReLU activation.

2.2   DeDA-Based Transformation Layer for Rim Parameterization

In this section, we derive DeDA-based transformation and its convolution layers for rim parameterization. As shown in Fig. 1, a rim+ lesion can be characterized by a hyperintense rim at the lesion edge on QSM and differs from a rim- lesion in both image intensities and gradients at the edge. To account for both image intensities and gradients, the rim is parameterized as $\tan(\theta) = \dfrac{y - b}{x - a}$, where (a, b) are parameters of coordinates for the rim center in the accumulator space and θ represents the gradient direction at (x, y) in the image space. As can be seen from the right panel of Fig. 2, mapping a single (x, y, θ) to the accumulator space produces a straight line, and thus coordinates of the rim center can be identified by the intersection of many of these lines.

DeDA Transformation of the Rim: Given a source feature map $\mathbf{U} \in \mathbb{R}^{H \times W}$, the magnitude of image gradients can be obtained as follows $\mathbf{S} = \sqrt{\mathbf{U}_x \odot \mathbf{U}_x + \mathbf{U}_y \odot \mathbf{U}_y}$, where $\odot$ denotes the Hadamard product, $\mathbf{U}_x = \dfrac{\partial \mathbf{U}}{\partial x}$, and $\mathbf{U}_y = \dfrac{\partial \mathbf{U}}{\partial y}$. The image gradient tensor $\mathbf{U}_x$ and $\mathbf{U}_y$ can be efficiently computed using convolution kernels such as the Sobel operator. Normalized gradients can be obtained by $\hat{\mathbf{U}}_x = \dfrac{\mathbf{U}_x}{\mathbf{S} + \epsilon}$ and $\hat{\mathbf{U}}_y = \dfrac{\mathbf{U}_y}{\mathbf{S} + \epsilon}$, where ϵ is a small real value to avoid zero denominator. The mesh grids of $\mathbf{U}$ are denoted as $\mathbf{M}_x$ (value range: $(0, H - 1)$) and $\mathbf{M}_y$ (value range: $(0, W - 1)$). We can then generate a set of sampling grids as follows:

$$\mathcal{G} = \{\mathbf{G}[k] = (\mathbf{G}^x[k], \mathbf{G}^y[k]) \mid k \in \mathbb{Z}^+, 1 \le k \le N\}, \tag{4}$$

where $\mathbf{G}[k] \in \mathbb{R}^{2 \times H \times W}$, $\mathbf{G}^x[k] = k\hat{\mathbf{U}}_x + \mathbf{M}_x$, $\mathbf{G}^y[k] = k\hat{\mathbf{U}}_y + \mathbf{M}_y$, and $N = \max(H, W)$. Now the DeDA-based transformation of Rim (DA-TR) can be formulated as $\mathbf{V}_s = \mathcal{D}(\mathbf{S}, \mathcal{G}; \mathcal{K})$ and $\mathbf{V}_u = \mathcal{D}(\mathbf{U}, \mathcal{G}; \mathcal{K})$, where the integer sampling kernel is used. It is worth noting that feature and gradient magnitude values are accumulated separately due to differences of image intensity and gradients between rim+ and rim- lesions (see Fig. 1).

Network Layer for DA-TR: To gain more representation ability and capture long-range contextual information, DA-TR is applied to both intermediate feature maps and original images. As can be seen from Fig. 3, image patches of

lesions are processed through a set of convolutional layers with each consisting of a $3 \times 3 \times 3$ or $1 \times 1 \times 1$ convolution, a batch normalization [11] and a ReLU activation function, followed by a DA-TR layer and a $1 \times 1 \times 1$ convolutional layer. The first $1 \times 1 \times 1$ conv layer is used to fuse feature maps and original image patches for better feature embedding, and the second one is used to fuse DeDA transformed gradient magnitude maps $\mathbf{V}_s$ and feature maps $\mathbf{V}_u$. It is worth noting that only in-plane rims are observed, and thus the DA-TR is performed on the 2D feature map slices along the axial direction.

3   Experiments and Results

For fair and consistent comparison, the dataset applied in the previous work [24] was asked for and used to demonstrate the performance of the proposed DeDA-based rim parameterization DA-TR. A total of 172 subjects were included in the dataset, and 177 lesions were identified as rim+ lesions and 3986 lesions were identified as rim- lesions, please refer to [24] for more details about the image acquisition and pre-processing.

3.1   Comparator Methods and Implementation Details

Comparator Methods: Three methods have been developed so far for rim+ lesion identification, of which APRL [18] and RimNet [4] are on phase imaging and QSMRim-Net [24] is on QSM. In comparison with these methods, we use QSM along with T2-FLAIR images as the network inputs for RimNet and QSMRimNet, and use the QSM image to extract first-order radiomic features for APRL. Furthermore, we applied residual networks (ResNet) [10], vision transformer (ViT) [8], Swin transformer [17], and Nested transformer [28] as backbone architecture for our application, and determined that ResNet with 18 convolution layers works the best. Transformer-based networks with fewer inductive biases rely heavily on the use of a large training dataset or depends strongly on the feature reuse [19], as a result, these networks as well as CNNs with deeper structures are prone to overfit small datasets.

Implementation Details: A stratified five-fold cross-validation procedure was applied to train and validate the performance, and all experiments including ablation study were carried out within this setting. Each lesion was cropped into patches with a fixed size of $32 \times 32 \times 8$ voxels. Random flipping, random affine transformation and random Gaussian blurring were used to augment our data. More details of the training procedure can be found out in the supplementary materials.

3.2   Results and Ablation Study

Lesion-wise Results: To evaluate the performance of each method and produce clinically relevant results, pROC curves with false positive rates (FPRs) in the range of $(0, 0.1)$ and PR curves of the different validation folds were interpolated using piece-wise constant interpolation and averaged to show the overall

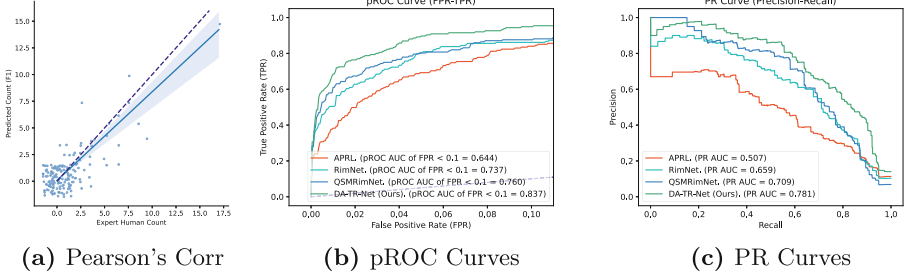**(a)** Pearson's Corr   **(b)** pROC Curves   **(c)** PR Curves

Fig. 4. The predicted count of rim + lesions from DA-TR-Net versus the expert human count is shown in (a), where points in the plot have been jittered for better visualization. The pROC and PR curves for the proposed and other comparator methods are shown in (b) and (c), where AUC denotes the area under the curve.

performance at the lesion level. For each curve, AUC was computed directly from the interpolated and averaged curves. The binary indicators of rim+/rim-lesions were generated by thresholding the model probabilities to maximize the F_1 score, where $F_1 = 2 \cdot \dfrac{precision \cdot sensitivity}{precision + sensitivity}$. In addition, accuracy, F1 score, sensitivity, specificity, and precision were used to characterize the performance of each method. Table 1 and Fig. 4 show the lesion-wise performance metrics of the proposed methods in comparison with the other methods. DA-TR-Net outperformed the other competitors in all evaluation metrics. With a slightly higher overall accuracy and specificity with other methods, DA-TR-Net resulted in a 5.5%, 15.4% and 39.4% improvement in F_1 score, 10.1%, 13.6% and 30.0% improvement in pROC (FPR < 0.1) AUC, and 10.2%, 18.5% and 54.0% improvement in PR AUC compared to QSMRimNet, RimNet and APRL, respectively.

Subject-wise Results: We also evaluated the performance at the subject-level. Pearson's correlation coefficient was used to measure the correlation model predicted count and human expert count. Mean Squared Error (MSE) was also used to measure the averaged accuracy for the model predicted count. Figure 4a shows the scatter-plot for the predicted count v.s. the human expert count, along with the identity line, and the Pearson's correlation coefficient (ρ) for DA-TR-Net was $\rho = 0.93(95\%CI : 0.90, 0.95)$ As can be seen from Table 1, the Pearson's correlations and MSE for the proposed DA-TR-Net was found higher

Table 1. Results of the proposed and other methods using a stratified five-fold cross-validation scheme. The best performing metric is bolded.

Method	Accuracy	F_1	Sensitivity	Specificity	Precision	ROC AUC	pROC AUC	PR AUC	ρ (95%CI)	MSE
APRL [18]	0.954	0.538	0.627	0.969	0.470	0.940	0.644	0.507	0.68 (0.59,0.75)	3.16
RimNet [4]	0.970	0.650	0.655	0.984	0.644	0.950	0.737	0.659	0.75 (0.67,0.81)	2.41
QSMRimNet [24]	0.977	0.711	0.667	0.991	0.761	0.939	0.760	0.709	0.89 (0.86,0.92)	1.00
DA-TR-Net (Ours)	**0.980**	**0.750**	**0.712**	**0.992**	**0.792**	**0.975**	**0.837**	**0.781**	**0.93(0.90,0.95)**	**0.69**

Table 2. Ablation study on the effects for each component in DA-TR. Multiple check marks for sets of N denote the union of the checked sets. Pre-Convs denotes a convolution block with six $3 \times 3 \times 3$ convolution layers.

#	Pre-Convs	$\mathbf{V}_u$	$\mathbf{V}_s$	$N \in \{5, 7, 9\}$	$N \in \{11, 13\}$	$N \in \{15\}$	F_1	ROC AUC	pROC AUC (FPR< 0.1)	PR AUC
1	×	×	×	×	×	×	0.685	0.945	0.753	0.689
2	×	✓	✓	×	×	✓	0.701	0.971	0.790	0.720
3	✓	✓	×	×	×	✓	0.703	0.967	0.795	0.714
4	✓	✓	✓	×	×	✓	0.702	0.976	0.817	0.736
5	✓	✓	✓	×	✓	✓	0.727	0.975	0.825	0.743
6	✓	✓	✓	✓	✓	✓	0.750	0.975	0.837	0.781

than other competitors. This demonstrates that the performance of DA-TR-Net at the subject-level is statistically significantly higher than that of APRL, Rim-Net, and QSMRim-Net (Table 2).

Ablation Study: We conducted an ablation study to investigate the effects of each component accompanied with DA-TR. First, we examined the effects of applying the proposed DA-TR to the latent feature maps and raw images. Second, we examined the effects of using $\mathbf{V}_u$ and $\mathbf{V}_s$, because rim+ lesions differ from rim- lesions in both gradient magnitudes and values at the edge of the lesion. We then investigated how multi-radius rim parameterization can affect the results, as the size of rim+ lesions vary greatly with a radius from 5 to 15 among different subjects. Results from models #1, #2 and #4 show that the rim parametrization DA-TR is useful for rim+ identification, and DA-TR used in the latent feature map space performs even better. Comparing model #3 and #4, one can see that accumulating both gradient magnitudes and feature values is beneficial. The consistent performance improvement from model #4 to #5 and from model #5 to #6 has demonstrated the effectiveness of applying multi-radius rim parameterization. More results on backbone networks can be found in the appendix.

3.3   Discussions

Medical images often require processing of a primary target or region of interest (ROI), such as rims, left ventricles, or tumors. These ROIs frequently exhibit distinct geometric structures [26] or possess specific spatial relationships [25] with their surroundings. Capturing these characteristics poses a challenge for modern neural networks, especially given limited and imbalanced training data. While differentiable grid sampling [12] can tackle some of these issues within a certain scope, another major class involving transformations (e.g. Hough transform [3]) that necessitate directed accumulation is overlooked. Our proposed DeDA bridges this gap, enabling the use of image transformations with directed accumulation within a neural network. This allows for the parametrization of geometric shapes and the modeling of spatial correlations in a differentiable manner.

While the study focuses on rim+ lesion identification, the proposed DeDA can be extended to other applications. These include the utilization of polar transformation for skin lesion recognition/segmentation, symmetric circular transformation for cardiac image registration [23], parabola transformation for curvilinear structure segmentation [21], and high-dimensional bilateral filtering [5].

4   Conclusions

We present DeDA, an image processing operation that helps parameterize rim and effectively incorporates prior information into networks through a value accumulation process. The experimental results demonstrate that DeDA surpasses existing state-of-the-art methods in all evaluation metrics by a significant margin. Furthermore, DeDA's versatility extends beyond lesion identification and can be applied in other image processing applications such as Hough transform, bilateral grid, and Polar transform. We are excited about the potential of DeDA to advance numerous medical applications and other image processing tasks.

References

1. Absinta, M., et al.: Seven-tesla phase imaging of acute multiple sclerosis lesions: a new window into the inflammatory process. Ann. Neurol. **74**(5), 669–678 (2013)
2. Absinta, M., et al.: Persistent 7-tesla phase rim predicts poor outcome in new multiple sclerosis patient lesions. J. Clin. Investig. **126**(7), 2597–2609 (2016)
3. Ballard, D.H.: Generalizing the Hough transform to detect arbitrary shapes. Pattern Recogn. **13**(2), 111–122 (1981)
4. Barquero, G., et al.: RimNet: a deep 3D multimodal MRI architecture for paramagnetic rim lesion assessment in multiple sclerosis. NeuroImage Clinical **28**, 102412 (2020)
5. Chen, Jiawen, Paris, Sylvain, Durand, Frédo.: Real-time edge-aware image processing with the bilateral grid. ACM Trans. Graph. **26**(3), 103 (2007). https://doi.org/10.1145/1276377.1276506
6. Dal-Bianco, A., et al.: Slow expansion of multiple sclerosis iron rim lesions: pathology and 7 t magnetic resonance imaging. Acta Neuropathol. **133**(1), 25–42 (2017)
7. De Rochefort, L., et al.: Quantitative susceptibility map reconstruction from MR phase data using Bayesian regularization: validation and application to brain imaging. Magn. Reson. Med. Official J. Int. Soc. Magn. Reson. Med. **63**(1), 194–206 (2010)

8. Dosovitskiy, A., et al.: An image is worth 16×16 words: transformers for image recognition at scale. In: International Conference on Learning Representations (2020)
9. Gillen, K.M., et al.: QSM is an imaging biomarker for chronic glial activation in multiple sclerosis lesions. Ann. Clin. Transl. Neurol. **8**(4), 877–886 (2021)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
11. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International Conference on Machine Learning, pp. 448–456. PMLR (2015)
12. Jaderberg, M., Simonyan, K., Zisserman, A., et al.: Spatial transformer networks. In: Advances in Neural Information Processing Systems, vol. 28 (2015)
13. Kamnitsas, K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **36**, 61–78 (2017)
14. Kaunzner, U.W., et al.: Quantitative susceptibility mapping identifies inflammation in a subset of chronic multiple sclerosis lesions. Brain **142**(1), 133–145 (2019)
15. Kayhan, O.S., Gemert, J.C.V.: On translation invariance in CNNs: convolutional layers can exploit absolute spatial location. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14274–14285 (2020)
16. Lenc, K., Vedaldi, A.: Understanding image representations by measuring their equivariance and equivalence. In: Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition, pp. 991–999 (2015)
17. Liu, Z., et al.: Swin transformer: hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 10012–10022 (2021)
18. Lou, C., et al.: Fully automated detection of paramagnetic rims in multiple sclerosis lesions on 3t susceptibility-based MR imaging. NeuroImage Clin. **32**, 102796 (2021)
19. Matsoukas, C., Haslum, J.F., Sorkhei, M., Söderberg, M., Smith, K.: What makes transfer learning work for medical images: feature reuse and other factors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9225–9234 (2022)
20. Muckley, M., et al.: Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. IEEE Trans. Med. Imaging **40**(9), 2306–2317 (2021)
21. Shi, T., Boutry, N., Xu, Y., Géraud, T.: Local intensity order transformation for robust curvilinear object segmentation. IEEE Trans. Image Process. **31**, 2557–2569 (2022)
22. Wang, Y., Liu, T.: Quantitative susceptibility mapping (QSM): decoding MRI data for a tissue magnetic biomarker. Magn. Reson. Med. **73**(1), 82–101 (2015)
23. Zhang, H., Hu, R., Chen, X., Wang, R., Zhang, J., Li, J.: DAGrid: directed accumulator grid. arXiv preprint arXiv:2306.02589 (2023)
24. Zhang, H., et al.: QSMRim-Net: imbalance-aware learning for identification of chronic active multiple sclerosis lesions on quantitative susceptibility maps. NeuroImage Clin. **34**, 102979 (2022)
25. Zhang, H., et al.: ALL-Net: anatomical information lesion-wise loss function integrated into neural network for multiple sclerosis lesion segmentation. NeuroImage Clin. **32**, 102854 (2021)
26. Zhang, H., et al.: Geometric loss for deep multiple sclerosis lesion segmentation. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), pp. 24–28. IEEE (2021)

27. Zhang, H., et al.: Efficient folded attention for medical image reconstruction and segmentation. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 10868–10876 (2021)
28. Zhang, Z., Zhang, H., Zhao, L., Chen, T., Arik, S.Ö., Pfister, T.: Nested hierarchical transformer: towards accurate, data-efficient and interpretable visual understanding. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, pp. 3417–3425 (2022)