Topic: Integrating Multi-Modal Deep Learning Techniques for Enhanced Segmentation and Analysis in Medical Imaging

professor's name: dr.Eslami

Fateme Gateh

ABSTRACT

In the 4 reviewed articles, the following topics have been discussed: **Few-Shot Medical Segmentation:** Few-shot medical segmentation involves learning to segment a new organ using minimal annotated support images. Existing methods often overlook global anatomy correlations between support and query sets. To address this, we propose a novel Cross Attention network that integrates global information from both parts, achieving significant improvement over metric learning-based methods on challenging datasets.**Automated Bone Marrow Segmentation:** Manual assessment of bone marrow signal is time-consuming. In our study, we explore using deep learning for automated bone marrow segmentation in children and adolescents. While model performance is acceptable, further improvement requires larger, balanced datasets and validation by radiologists.**Parallel Deep CNN for Brain Tumor Classification:** Convolutional neural networks (CNNs) can suffer from overfitting. Our novel Parallel Deep CNN (PDCNN) extracts global and local features from two parallel stages, mitigating overfitting. The PDCNN achieves high accuracies on three MRI datasets, outperforming state-of-the-art techniques.**Automated Fibrosis Segmentation in CMR Imaging:** Visualizing fibrosis in cardiac MRI (CMR) is crucial. We propose a deep learning solution that accurately segments myocardium and scar/fibrosis from contrast-enhanced CMR images. Our 3-stage neural network achieves expert-level performance, allowing direct calculation of clinical measures. This approach could extend to other imaging modalities and patient pathologies.

Methods

The methods examined in 4 articles are as follows:

STUDY 1

"Data sources included 155 two-dimensional late gadolinium-enhanced cardiac magnetic resonance (LGE-CMR) patient scans (with 1124 slices) and 246 synthetic 'LGE-like' scans (with 1360 slices). These scans were obtained from cine CMR using a novel style-transfer algorithm. First, it identified the left ventricle (LV) region of interest. Finally, it segmented the ROI into viable regions , we post-processed the segmentation results to ensure anatomical accuracy. These segmentations allowed us to directly compute clinical features, such as LV volume and scar burden."This study demonstrates an innovative approach to automated LGE-CMR image segmentation, providing accurate myocardium and scar/fibrosis segmentations for clinical analysis. The potential applications extend beyond LGE-CMR and could be adapted to other imaging modalities and patient pathologies.

STUDY 2

1. Brain Tumor Classification Using Convolutional Neural Networks (CNNs):

- Brain tumor diagnosis is crucial for effective medical intervention and patient survival.

- Traditional methods involve manually inspecting MRI brain images, which is time-consuming and prone to human errors.

- Researchers have turned to CNNs for automated brain tumor classification.

- A parallel architecture of two CNNs, known as the Parallel Deep Convolutional Neural Network (PDCNN), is proposed.

- Here's how it works:

    - Input layer receives preprocessed brain MRI images (reduced to 32x32 pixels and converted to grayscale).

    - Data augmentation generates new images from existing ones.

    - The dataset is split into training and validation sets.

    - PDCNN classifies input images using local, global, merging, and output routes.

    - The softmax function is used for brain tumor categorization.

2. Quantum Convolutional Neural Networks (QCNNs):

- Integrating QCNNs into medical diagnostics is transformative for brain tumor classification.
- These models are specifically designed to identify and classify brain cancer images with high precision.

3. Data Augmentation and Image Processing:

- In addition to CNNs, data augmentation and image processing techniques can enhance brain tumor classification.
- These methods help improve model performance by creating variations of existing images and reducing complexity.

STUDY 3

- In this study, knee images from 95 healthy patients and children with nonbacterial osteomyelitis between the ages of 6 and 18 years were selected in a long-term multicenter study.
- Bone marrow signal in MRI images is divided into three staining intensity levels using T2-weighted Dixon images:
  - 1: Low increase
  - 2: Average increase
  - 3: Strong increase until like a liquid signal
- A Convolutional Neural Network (CNN) is trained using 85 images to perform bone marrow segmentation.
- Four readers have manually segmented the test images and calculated the actual segmentation value using the STAPLE method.

- The performance of the model and the readers have been evaluated using the Dice similarity coefficient and with specified agreement.

Study 4

The transformer-based Cross Attention module is leveraged to strengthen the fusion of support and query data. and Through the comparable results on two challenging datasets (abdominal segmentation dataset CHAOS and cardiac segmentation dataset MS-CMRSeg), we proved that the traditional meta-learning based methods still have great potential when strengthening the information exchange between two branches. We achieve remarkable improvement in the proposed method compared with currently dominant metric learning-based methods.

Results

The results of the above research are as follows:

Study 1:

The researchers propose a deep learning (DL) approach for automatically and accurately segmenting the myocardium and scar/fibrosis regions in late gadolinium-enhanced cardiac magnetic resonance (LGE-CMR) images. Additionally, they extract anatomical features, such as scar burden and ventricular volume. The DL model consists of three subnetworks, each with specific tasks:

1. The first subnetwork addresses class imbalance between the region of interest (ROI) and background.

2. The second subnetwork delineates the endocardium and epicardium.

3. The third subnetwork ensures anatomical correctness for both individual slices and entire volumes.

The proposed ACSNet outperforms manual expert segmentations and performs well on diverse input data with varying scar distribution patterns from different imaging centers and MRI machines. Furthermore, the technology enables the extraction of clinical anatomical covariate data, potentially enhancing the prognostic utility of LGE-CMR.The predicted LV and scar segmentations achieved 96% and 75%

balanced accuracy, respectively, compared to trained expert segmentations. The mean scar burden difference between manual and predicted segmentations was 2%.

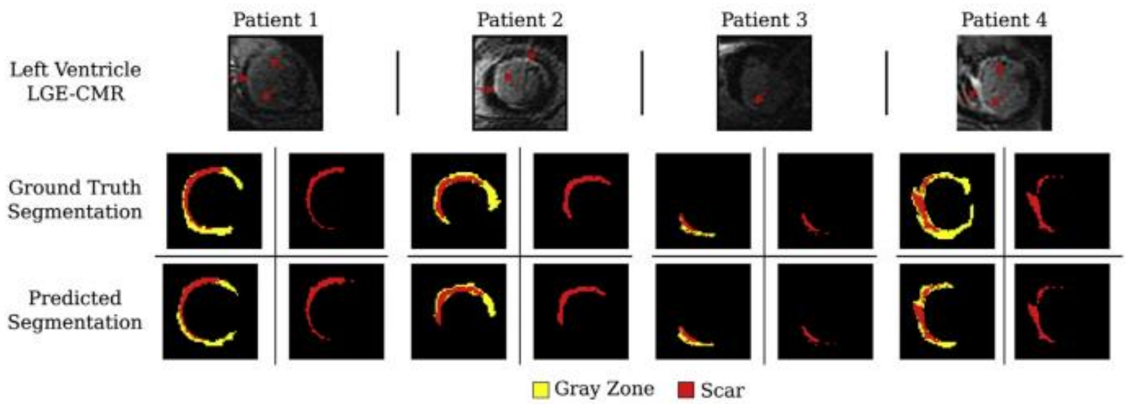


Figure 1 Scar segmentation results. Segmentations of enhancement regions from the myocardium segmentation network represent gray zone (yellow) and scar (red). The first row shows the original scan, the middle row shows the ground truth scar and gray zone segmentations, and the bottom row shows the predicted segmentations. LGE-CMR 5 cardiac magnetic resonance imaging with late gadolinium enhancement.

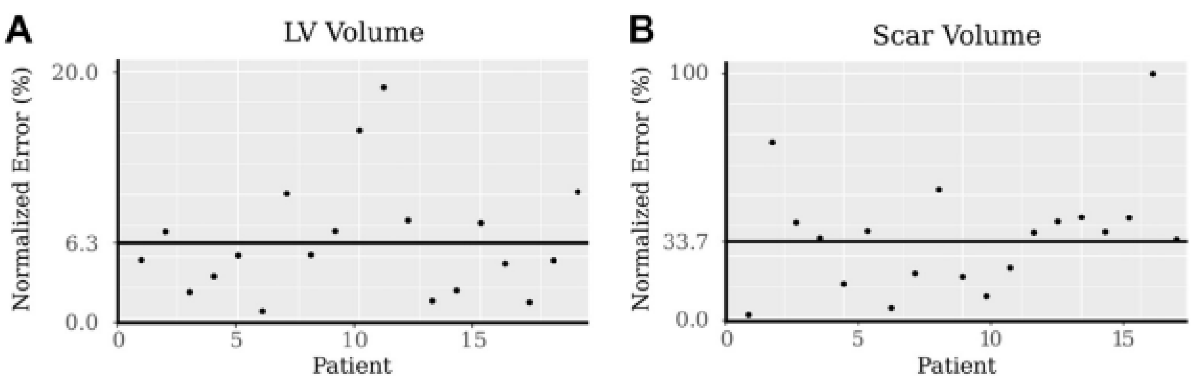


Figure 2 Scar and left ventricle (LV) volumes. LV (A) and scar (B) volume error is computed as the absolute error normalized by each respective volume. Each point represents the error in LV volume of a single segmented patient scan. The solid black line shows the mean.

Study 2:

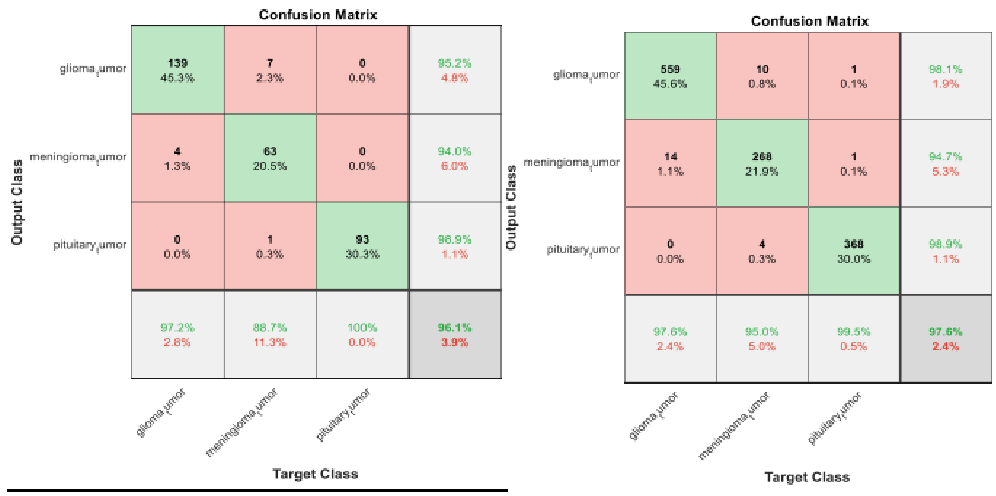
Three forms of MRI datasets are used to determine the effectiveness of the proposed method. The binary tumor identification dataset-I, Figshare dataset-II, and Multiclass Kaggle dataset-III provide accuracy of 97.33%, 97.60%, and 98.12%, respectively. The proposed structure is not only accurate but also efficient, as the proposed method extracts both low-level and high-level features, improving results compared to state-of-the-art techniques.

Table 1 Suggested tumor type categorization outcomes using multiclass Figshare dataset

Method	Original Dataset II	Augmented Dataset II	Measures			
			Accuracy (%)	Error (%)	Time (s)	Kappa
CNN	✓		95.80	4.20	833	0.933
		✓	97.90	2.10	1627	0.967
PDCNN	✓		96.10	3.90	4169	0.938
		✓	97.60	2.40	5803	0.962

Table 2 Suggested tumor type categorization outcomes using multiclass Kaggle dataset III.

Method	Original Dataset III	Augmented Dataset III	Measures			
			Accuracy (%)	Error (%)	Time (s)	Kappa
CNN	✓		94.10	5.90	629	0.919
		✓	97.70	2.30	783	0.968
PDCNN	✓		95.60	4.40	3465	0.940
		✓	98.12	1.88	6430	0.974



(a) (b)

Fig. 3. Confusion matrix of PDCNN using Figshare dataset-II (a) confusion matrix of the original dataset-based classification (b)confusion matrix of the original dataset with augmentation-based classification

Table 3 Comparison with existing techniques those used Kaggle and Figshare datasets

No	Authors	Methodology	Year	Dataset	Accuracy
1.	P. Afshar et al. [18]	Capsule Networks	2019	Figshare Dataset-II	90.89%
2.	Chirodip Lodh Choudhury et al. [26]	CNN	2020	Binary Dataset-I	96.08%
3.	H. H. Sultan et al. [25]	Resize + Augmentation + CNN + Hyperparameter Tuning	2019	Figshare Dataset-II	96.13%
4.	Suhib et al. [15]	Gray + Resize + Flatten + CNN	2020	Binary Dataset-I	96.7%
5.	Agus Eko Minarno et al. [20]	Resize + Augmentation + CNN + Hyperparameter Tuning	2021	Kaggle Dataset-III	96.00%
6.	Priyansh et al. [29]	CNN Based Transfer Learning Approach	2021	Binary Dataset-I	Resnet-50- 95%, VGG-16- 90%, Inception-V3-55%
7.	T.Rahmanet al. [30]	Resize + Gray + Augmen -tation + Binary + CNN	2022	Binary Dataset-I	96.9%
8.	**Proposed Structure**	**Resize + Gray scale conversion + Augmentation + PDCNN**	-	**Binary Dataset-I** **Figshare Dataset-II** **Kaggle Dataset-III**	**97.33%** **97.60%** **98.12%**

Study 3:

Consensus score of model performance showed acceptable results for all but one examination. Model performance and reader agreement had highest scores for level-1 signal (median Dice 0.68) and lowest scores for level-3 signal (median Dice 0.40), particularly in examinations where this signal was sparse.It is feasible to develop a deep-learning-based model for automated segmentation of bone marrow signal in children and adolescents. Our model performed poorest for the highest signal intensity in examinations where this signal was sparse. Further improvement requires training on larger and more balanced datasets and validation against ground truth, which should be established by radiologists from several institutions in consensus.

Table 4 Median and mean Dice similarity coefcient between ground truth and the segmentations performed by the artifcial intelligence (AI) model and the four readers

		Reader 1	Reader 2	Reader 3	Reader 4	AI model
Level-1 signal (turquoise)	Median (range)	0.80 (0.69–0.90)	0.83 (0.70–0.88)	0.73 (0.42–0.81)	0.75 (0.57–0.84)	0.68 (0.60–0.74)
	Mean	0.81	0.8	0.67	0.72	0.68
Level-2 signal (blue)	Median (range)	0.72 (0.35–0.87)	0.67 (0.23–0.84)	0.55 (0.29–0.78)	0.17 (0.01–0.53)	0.47 (0.25–0.62)
	Mean	0.7	0.64	0.55	0.17	0.45
Level-3 signal (yellow)	Median (range)	0.64 (0.20–0.87)	0.67 (0.52–0.93)	0.59 (0.44–0.75)	0.00 (0.00–0.89)	0.40 (0.00–0.71)
	Mean	0.6	0.69	0.59	0.11	0.38
Combined levels 2 + 3 signal	Median (range)	0.79 (0.33–0.86)	0.71 (0.19–0.89)	0.61 (0.38–0.78)	0.18 (0.01–0.73)	0.55 (0.24–0.74)
	Mean	0.7	0.69	0.6	0.21	0.5

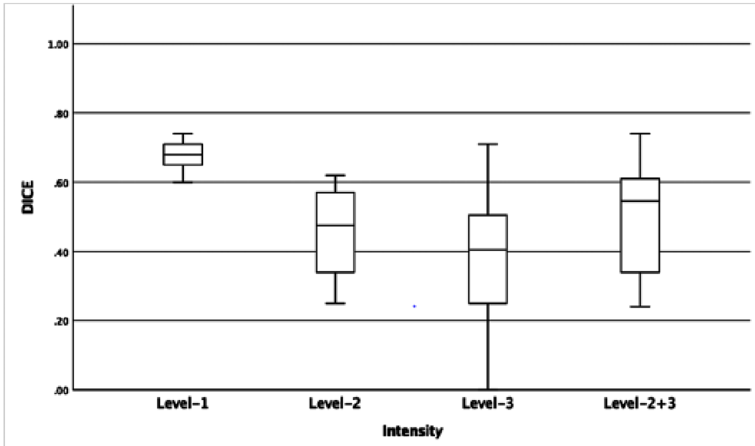


Fig. 4 Boxplot illustrates the performance of the artifcial intelligence (AI) model compared to ground truth for the diferent signal intensities expressed by the Dice similarity coefcient

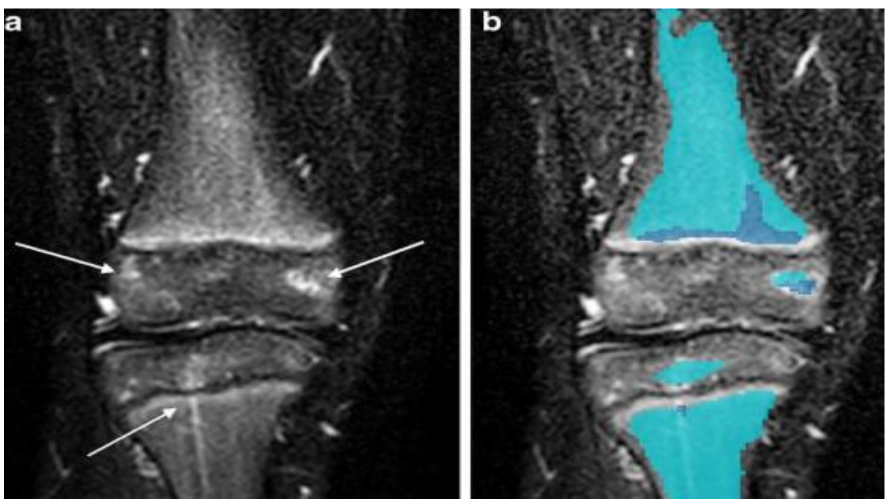


Fig. 5 a, b MRI, coronal T2-W Dixon water-only of the knee in an 8-year-old healthy and asymptomatic girl. Image (b) includes the segmentation mask performed by the model. The small foci of level-3 signal shown with arrows in image (a) are either missed (no color coding in image b) or incorrectly labelled with a lower intensity level by the model (coded with either blue or turquoise in image b)

Study 4:

1. U-Transformer: Self and Cross Attention for Medical Image Segmentation:

- Medical image segmentation is particularly challenging for complex and low-contrast anatomical structures. To address this, researchers have introduced the U-Transformer network.

- The U-Transformer combines a U-shaped architecture (similar to U-Net) with self-attention and cross-attention mechanisms from Transformers.

- Self-attention allows global interactions between encoder features, capturing long-range contextual information.

- Cross-attention in the skip connections refines spatial recovery in the U-Net decoder by filtering out non-semantic features.

- Experiments on abdominal CT-image datasets demonstrate significant performance gains compared to U-Net and local Attention U-Nets.

- U-Transformer provides interpretability features, making it a powerful tool for medical image segmentation.

2. Few-Shot Medical Image Segmentation with Cross Attention Transformer (CAT-Net):

- CAT-Net is another exciting approach based on cross masked attention Transformers.

- It addresses few-shot medical image segmentation by mining correlations between support and query images.

- CAT-Net focuses on useful foreground information, enhancing representation capacity.

- The network achieves impressive results by merging features from both support and query data.

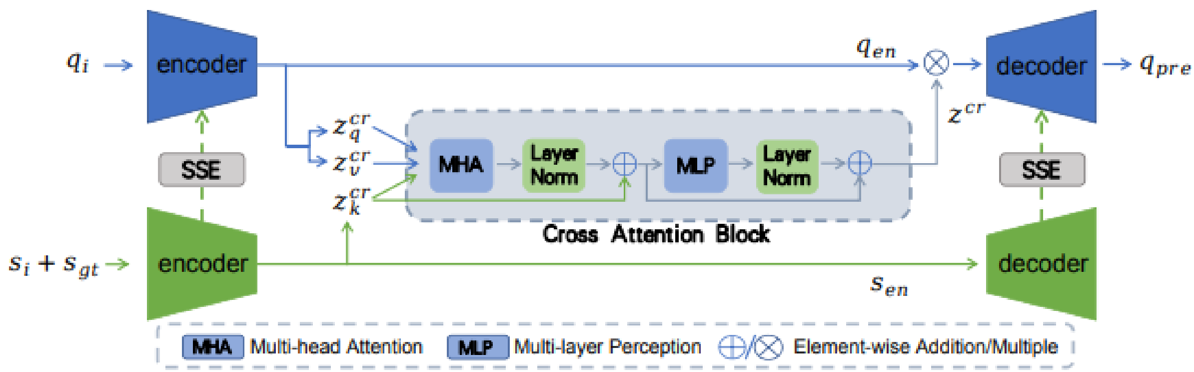


Fig. 6. The network architecture of our proposed method. It is mainly composed of three parts: The conditioner(green part) and the segmenter(blue part) extract and analyze the features of support images and query images, respectively, the Cross Attention Block in the middle fuses the global feature of two parts.

TABLE 5 COMPARISONS AGAINST OTHER METHODS

Method	MS-CMRSeg				CHAOS				
	LV-BP	LV-MYO	RV	Mean	L.kildney	R.kidney	Spleen	Liver	Mean
SE-Net [17]	69.92	44.71	65.43	59.69	62.11	61.32	51.80	27.43	50.66
CANet [31]	78.99	43.61	61.10	61.07	69.53	77.15	67.05	72.88	71.65
PPNet [32]	67.78	42.61	60.80	57.06	62.13	71.78	66.57	**73.12**	68.40
PANet [20]	80.20	45.67	66.95	64.27	53.45	38.64	50.90	42.26	46.33
Ours	72.49	**45.97**	**70.18**	62.88	**74.92**	71.89	**75.43**	64.57	**71.70**

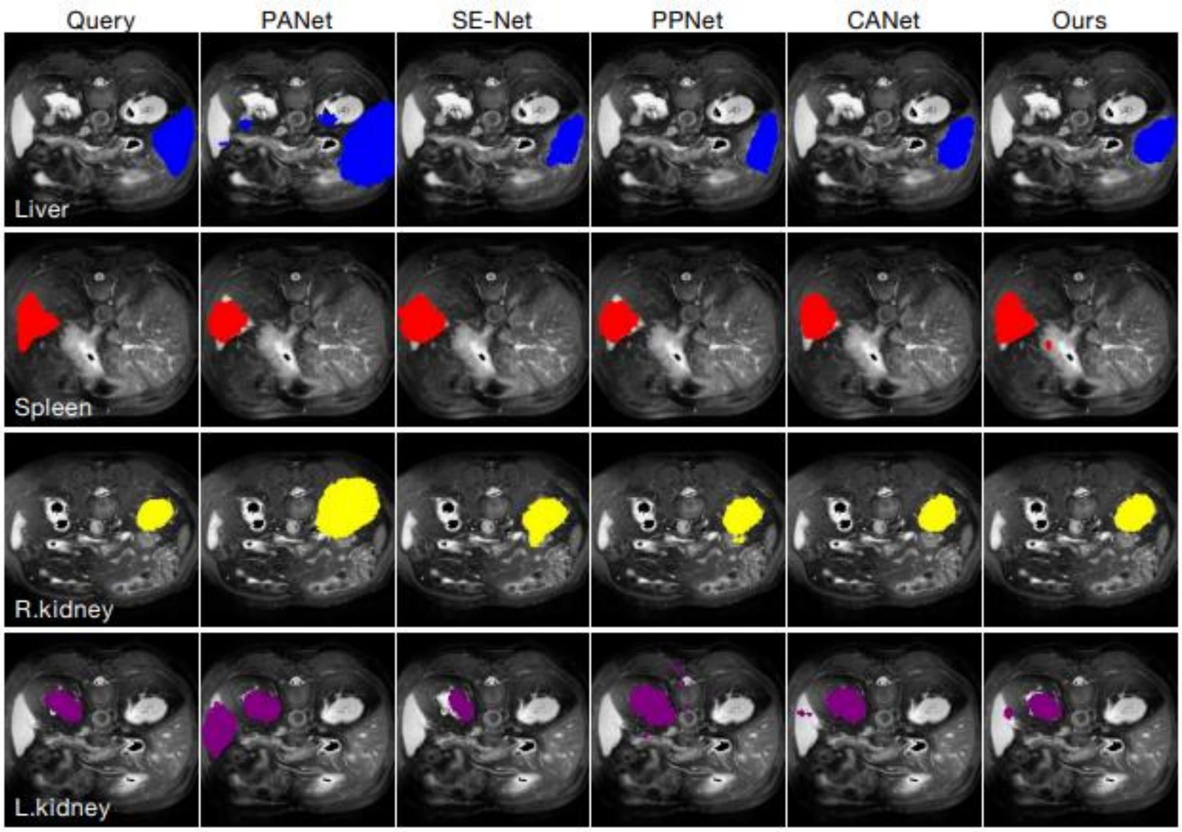


Fig. 7. Qualitative results of our method compared with other methods on abdominal MRI dataset CHAOS.

REFRENCE

1- Dan M. Popescu, PhD,* Haley G. Abramson, BS,† Rebecca Yu, BS,† Changxin Lai, BS,Julie K. Shade, PhD,* Katherine C. Wu, MD,*‡ Mauro Maggioni, PhD,*Natalia A. Trayanova, PhD, FHRS,(2012) ally informed deep learning on contrastenhanced cardiac magnetic resonance imaging for scar segmentation and clinical feature extraction

2- Takowa Rahman, Md Saiful Islam,(2023) MRI brain tumor detection and classification using parallel deep convolutional neural networks, Contents lists available at ScienceDirect Measurement: Sensors

3- Elisabeth von Brandis,Håvard B. Jenssen,Derk F. M. Avenarius, Atle Bjørnerud1, Berit Flatø,Anders H.Tomterstad,Vibke Lilleby, Karen Rosendahl,Tomas Sakinis,PiaK.K. Zadig,Lil-Sofe Ording Müller, Automated segmentation of magnetic resonance bone marrow signal: a feasibility study,(2021), Pediatric Radiology (2022) 52:1104–1114

4- Yao Niu, dZhuoran Li, Shaozi Li, (2022),Cross Attention with Transformer for Few-shot Medical Image Segmentation, 2022 12th International Conference on Information Technology in Medicine and Education (ITME)