A Hybrid CNN-Transformer for Accelerated and Accurate Ktrans Map Reconstruction in DCE-MRI

Akito Yamauchi, ay2668@columbia.edu
Zongyu Li, zl3372@columbia.edu
Jia Guo jg3400@columbia.edu

Objective

To develop a deep learning–driven approach that:

- **Accelerate** Ktrans reconstruction
- Enables robust reconstruction under **noise**, **motion** (simulated with torchIO), and **low GBCA**

Introduction

Utility of Ktrans

The volume transfer constant (**Ktrans**) measures microvascular permeability and blood flow, holding significant diagnostic and prognostic value.
Example: glioma grading [1].

Limitation in Current Method

The conventional **Tofts model** requires **over an hour** to process a mouse brain scan and is **highly sensitive to noise and motion**. It also requires **gadolinium-based contrast agents (GBCAs)**, which increase the risk of **nephrogenic systemic fibrosis (NSF)** in individuals with renal impairment [2].

Preliminary Study

Lee et al. introduced a **spatiotemporal CNN-based approach** to accurately reconstruct Ktrans maps using only 3% of the standard GBCA dose [3].

Method (Data)

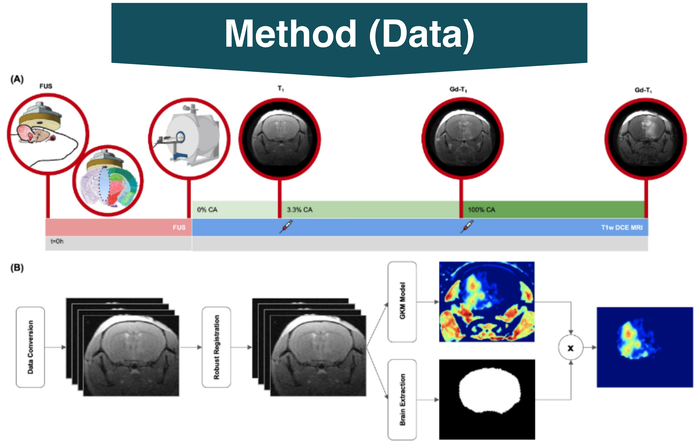


Fig. 1. (A) The blood-brain barrier (BBB) was first opened using focused ultrasound (FUS). We then injected 3.3% of the GBCA dose and acquired T1-weighted DCE-MRI, followed by the remaining 96.7% injection and a second DCE-MRI scan. A total of 10 mice were scanned.
(B) Raw DCE-MRI data were converted from DICOM to NIFTI, and within-subject robust registration was performed. Ground-truth Ktrans maps were generated using the Tofts model, and the whole brain was extracted using manually labeled brain masks. Noise and motion were artificially introduced with torchIO.

Method (Model)

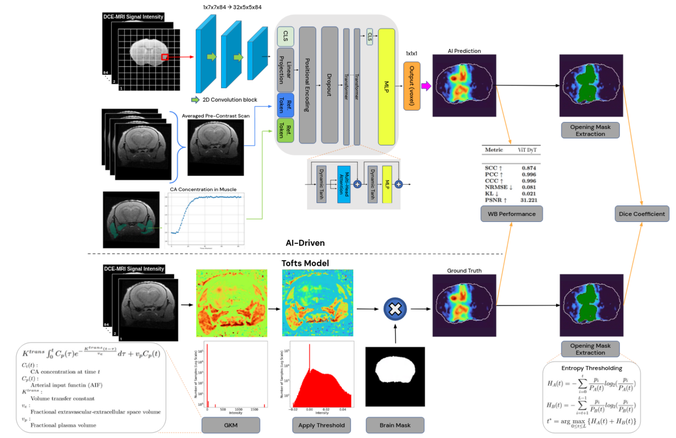


Fig. 2. We extracted 7×7×84 (x, y, time) patches with stride one. The hybrid model first extracted local spatial features using a 2D CNN, followed by a transformer encoder to capture global spatiotemporal relationships. The model was implemented in PyTorch using:

- Adam optimizer (learning rate = 1e-3)
- Mean Absolute Error (MAE) loss
- Batch size = 256

Training used five-fold cross-validation with patch-based learning to address limited data. All models were trained and evaluated on RTX2080Ti.

Result

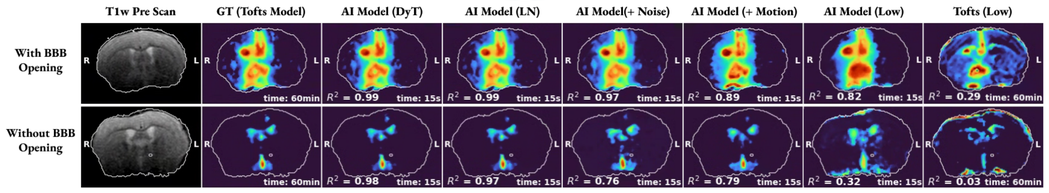


Fig. 3. Qualitative comparison of Ktrans maps reconstructed using the Tofts model. The top row shows slices from a subject with significant BBB opening (positive control), while the bottom row shows slices from a subject without BBB opening (negative control).

- Under ideal conditions (full GBCA dose, no noise/motion), the AI model predicted **nearly identical to the ground truth**, while **reducing reconstruction time from 60 minutes to 15 seconds.**
- Under simulated noise and motion, the AI-model maintained its performance
- With 3% GBCAs, the AI-driven approach outperformed the Tofts model, though performance in negative control cases was weaker.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Metric	AI DyT	AI LN	AI Noise	AI Motion	AI Low	Tofts Low
Positive Control						
SCC ↑	**0.874**	0.872	0.862	0.800	**0.618**	0.209
PCC ↑	**0.996**	0.995	0.987	0.946	**0.903**	0.182
CCC ↑	**0.996**	0.995	0.985	0.940	**0.885**	0.118
NRMSE ↓	**0.081**	0.086	0.153	0.316	**0.463**	2.509
KL ↓	**0.021**	0.029	0.072	0.098	**0.085**	0.154
PSNR ↑	**31.221**	30.543	25.704	19.624	**16.624**	9.637
Negative Control						
SCC ↑	**0.936**	0.534	0.408	0.511	**0.382**	0.209
PCC ↑	**0.988**	0.986	0.899	0.879	**0.570**	0.182
CCC ↑	**0.988**	0.986	0.868	0.851	**0.541**	0.118
NRMSE ↓	**0.147**	0.160	0.492	0.556	**1.032**	2.509
KL ↓	**0.006**	0.025	0.087	4.978	**0.611**	0.154
PSNR ↑	**36.578**	36.175	25.830	23.874	**17.903**	9.637

Table 1. Quantitative performance comparison:

- 100% GBCAs with Dynamic Tanh
- 100% GBCAs with Layer Normalization
- 100% GBCAs with noise + Dynamic Tanh
- 100% GBCAs with motion + Dynamic Tanh
- 3.3% GBCAs with Dynamic Tanh
- 3.3% GBCAs with Tofts model

The transformer with Dynamic Tanh outperformed layer normalization, especially in negative control cases (SCC). Results also demonstrate robustness under noise and motion. For 3.3% GBCAs, the AI model significantly outperformed the Tofts model.

Conclusion

The hybrid CNN-Transformer approach reduces Ktrans reconstruction time by **over 200-fold** while maintaining **high fidelity under noise, motion, and reduced GBCA doses**. This method significantly accelerates the workflow and suggests the potential for reducing GBCA use, thereby **lowering the risk of NSF** in patients with renal impairment. Future work includes scaling to larger datasets and extending the method to human DCE-MRI studies.

Reference

1. Zhao, Ming, et al. "Quantitative analysis of permeability for glioma grading using dynamic contrast-enhanced magnetic resonance imaging." Oncology letters 14.5 (2017): 5418-5426.
2. Shamam, Younus M., and O. De Jesus. "Nephrogenic systemic fibrosis." (2021).
3. Lee, Pin-Yu, et al. "Deep learning enables reduced gadolinium dose for contrast-enhanced blood-brain barrier opening." ArXiv (2023): arXiv-2301.