Received: 20 January 2022 | Revised: 17 October 2022 | Accepted: 19 October 2022

DOI: 10.1002/mp.16081

RESEARCH ARTICLE

MEDICAL PHYSICS

VTDCE-Net: A time invariant deep neural network for direct estimation of pharmacokinetic parameters from undersampled DCE MRI data

Aditya Rastogi | **Arindam Dutta** | **Phaneendra Kumar Yalavarthy**

Computational and Data Sciences, Indian Institute of Science, Bengaluru 560012, India

Correspondence
Phaneendra Kumar Yalavarthy, Computational and Data Sciences, Indian Institute of Science, Bengaluru - 560012, India.
Email: yalavarthy@iisc.ac.in

Funding information
Prime Minister Research Fellowship (PMRF); Science and Engineering Research Board, Grant/Award Number: CRG/2018/000672

Abstract

Purpose: To propose a robust time and space invariant deep learning (DL) method to directly estimate the pharmacokinetic/tracer kinetic (PK/TK) parameters from undersampled dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) data.

Methods: DCE-MRI consists of 4D (3D-spatial + temporal) data and has been utilized to estimate 3D (spatial) tracer kinetic maps. Existing DL architecture for this task needs retraining for variation in temporal and/or spatial dimensions. This work proposes a DL algorithm that is invariant to training and testing in both temporal and spatial dimensions. The proposed network was based on a 2.5-dimensional Unet architecture, where the encoder consists of a 3D convolutional layer and the decoder consists of a 2D convolutional layer. The proposed VTDCE-Net was evaluated for solving the ill-posed inverse problem of directly estimating TK parameters from undersampled $k-t$ space data of breast cancer patients, and the results were systematically compared with a total variation (TV) regularization based direct parameter estimation scheme. In the breast dataset, the training was performed on patients with 32 time samples, and testing was carried out on patients with 26 and 32 time samples. Translation of the proposed VTDCE-Net for brain dataset to show the generalizability was also carried out. Undersampling rates (**R**) of 8×, 12×, and 20× were utilized with PSNR and SSIM as the figures of merit in this evaluation. TK parameter maps estimated from fully sampled data were utilized as ground truth.

Results: Experiments carried out in this work demonstrate that the proposed VTDCE-Net outperforms the TV scheme on both breast and brain datasets across all undersampling rates. For $\mathbf{K_{trans}}$ and $\mathbf{V_p}$ maps, the improvement over TV is as high as 2 and 5 dB, respectively, using the proposed VTDCE-Net.

Conclusion: Temporal points invariant DL network that was proposed in this work to estimate the TK-parameters using DCE-MRI data has provided state-of-the-art performance compared to standard image reconstruction methods and is shown to work across all undersampling rates.

KEYWORDS
DCE MRI, inverse problems, medical image reconstruction, permeability imaging and magnetic resonance imaging, pharmacokinetic modeling

1 | INTRODUCTION

Modern medical imaging methods provided an accurate and noninvasive (or minimally invasive) way of investigating anatomy and anatomical manifestations of various diseases and injuries. The tissue physiological or metabolical characterization is vital for accurate diagnosis of the disease. Recent advancements in the field of quantitative imaging methods like computed tomography (CT) perfusion,[1] dynamic contrast enhanced (DCE) magnetic resonance imaging (MRI),[2] positron emission tomography,[3] and diffusion tensor imaging[4] enabled this characterization. The DCE-MRI is a promising technique, where a gadolinium-based T1 shortening contrast agent[5] is injected into the body, followed by repeated acquisition of T1-weighted images of the region of interest after a fixed time interval. This 4D (3D + time) data processing results in physiological characterization of investigated tissue. In DCE-MRI, one takes advantage of blood capillaries surrounding the unhealthy tissue behaving differently than that of surrounding healthy tissue. This leads to change in the capillary permeability, that is, the rate at which the contrast agent permeates through the capillary and gets accumulated in extracellular extravascular space (EES). For example, in the case of malignant tumors, the blood capillaries formed are leaky and display hyperpermeability when compared to normal capillaries. By comparing such tracer kinetic (TK) parameters, clinicians can provide a better prognosis and treatment plan. Additionally, it is useful in assessing the efficacy of the treatment. To estimate these quantitative parameters, pharmacokinetic models are utilized, most popular being Tofts,[6] extended-Tofts or eTofts,[7] and Patlak model.[8] Patlak model parameterizes the tissue using two quantitative parameters, namely, $\mathbf{K_{trans}}$ and $\mathbf{V_p}$. $\mathbf{K_{trans}}$ denotes the rate at which contrast agent permeates from capillaries into EES and $\mathbf{V_p}$ measures the volume fraction of blood capillaries in each tissue (voxel). eTofts model utilizes an additional parameter called $\mathbf{k_{ep}}$, which denotes the rate at which contrast agent is cleared from the EES.

The raw data captured by the MR scanner are the Fourier transform of the desired anatomical image. These raw data are also called as *k* space data or Fourier space data. When these *k* space data are captured repeatedly as a function of time, they are termed $k-t$ space data. Repeated acquisitions as a function of time increases the overall scan time which makes DCE-MRI less appealing both ergonomically and economically. Additionally, to accurately estimate the quantitative parameters, the scans should have high spatial and temporal resolution, although one comes at the cost of other. Earlier works utilized principles of compressive sensing to undersample $k-t$ space below Nyquist threshold and still recover high-quality MR images.[9] By reducing acquisition time of each 3D volume, temporal resolution was improved and overall scan time was reduced. The methods to estimate TK parameters from undersampled $k-t$ data can be classified into indirect and direct methods. In indirect techniques, the anatomical image is reconstructed first using the principles of compressive sensing from which, as a second step, TK parameters are estimated. Experiments by Smith et al.,[10] Feng et al.,[11] and Rosenkrantz et al.[12] have shown that undersampling rates (**R**) of 4-28.7× can be achieved without compromising the diagnostically relevant information. They also utilized priors such as TV or wavelet transform, induced by $L_{1/2}$ or L_1 norms as sparsity constraints. In ref. [13], authors experimented with deep learning (DL)-based indirect parameter estimation using two popular model based DL techniques namely MoDL[14] and ISTA-Net[15] and compared the results with direct iterative schemes. The main disadvantage of indirect techniques like ISTA-Net and MoDL are that they provide reconstruction of anatomical images and not TK parameters. To estimate TK parameters from anatomical images additional information like AIF, T1 maps, and scan acquisition parameters are needed for which additional data acquisition and computation are required. On the contrary, in this work, we proposed an end-to-end network for reconstruction of TK parameters from undersampled k-t space data without any requirement of these parameters.

Contrary to the indirect methods, direct methods estimate the TK parameters from undersampled data by solving a nonlinear inverse problem without reconstructing the anatomical images as an intermediary step. Guo et al.[16,17] and Dikaois et al.[18] showed that direct reconstruction methods perform better than indirect methods. In their work,[18] the authors used Bayesian inference on prostate cancer dataset to achieve an undersampling rate of 4×. More recent works (refs. [16,17]) validated direct estimation methods for higher undersampling rate of 100× and gave the flexibility to incorporate any prior. Experiments with the breast data[13] showed that at low undersampling rates, DL-based indirect methods perform significantly better; however, at higher undersampling rate, the iterative direct techniques using total variation (TV)[19] regularization perform significantly better. However, direct iterative techniques are computationally expensive and require manual tuning of hyper parameters. Recent investigations addressed these limitations by using DL-based direct estimation techniques. Bliesener et al.[20] estimated $\mathbf{K_{trans}}$ from fully sampled data using DL at each pixel individually using one-dimensional (1D) convolution. However, instead of image as input, they utilized concentration maps and arterial input function (AIF) as input. Cagdas et al.[21,22] estimated parameters using dilated convolution and fully connected layers with temporal dimension are channels. Kettelkamp et.al[23] modified the network proposed by Cadgas et al. by incorporating AIF as an input information. However, as

these networks use fully convolutional layers and treat temporal dimension as number of channels, these networks are not invariant to change in temporal or spatial dimensions and require retraining once these dimensions vary. Moreover, Kettelkamp et.al[23] estimated TK parameters from fully sampled DCE-MRI data—not from undersampled DCE-MRI data—with explicit knowledge of AIF. From undersampled DCE-MRI data, AIF cannot be accurately estimated. The proposed architecture addressed both the issues of invariance to spatiotemporal dimension as well as the requirement of AIF for TK parameter estimation.

To address this limitation, this work propose a 2.5D Unet[24]-based architecture for direct estimation of TK parameters from undersampled $k-t$ space. The network consists of three-dimensional (3D) convolutional layers for the encoder part and two-dimensional (2D) convolution layers for the decoder part. The bottleneck consists of an averaging layer along the temporal axis to convert 3D data into 2D data. As the network consists only of convolutional layers and has averaging layers along temporal dimensional, the network becomes invariant to changes in temporal and spatial dimensions while training and testing. This network has been named as VTDCE-Net for variable time and space DCE network. The proposed VTDCE-Net was trained and tested on breast cancer patients for undersampling rate of 8×, 12× and 20× using RGA undersampling pattern.[11] The results were compared with a direct iterative parameter estimation technique as performed in ref. [13] using PSNR, SSIM,[25] and high-frequency error norm (HFEN[26]) as metrics. The proposed network was also tested and trained on brain images to evaluate the generalization capabilities. Moreover, utility and importance of time invariant networks were demonstrated (in Section 3.5) by comparing the proposed network results with another noninvariant network (called as DCE-Net), which is also based on Unet architecture. The main difference in VTDCE-Net and DCE-Net is that the DCE-Net is not flexible to variance in temporal dimension. DCE-Net treats the temporal dimensions as channels, and a 2D Unet is where the output channel dimension is two in order to give the two TK parameters. The main objective of the experiment performed in Section 3.5 was to show that resampling the input to match the data dimensions on which the network was trained does not give the most optimal results.

2 | THEORY AND METHODS

Notation wise, small boldface alphabets such as $\mathbf{x}$ are for vectors and capital boldface alphabets like $\mathbf{A}$ are for matrices. The $\mathbf{r}$ describes the spatial location in 3D volume in image domain, that is, $\mathbf{r} \in \{x, y, z\}$ and $\mathbf{r}$ was used to denote location in 3D volume of frequency domain, that is, $\hat{\mathbf{r}} \in \{k_x, k_y, k_z\}$. The small boldface $\mathbf{k}(\hat{\mathbf{r}}, t)$ was utilized to denote fully sampled $k-t$ space data and small boldface $\hat{\mathbf{k}}(\hat{\mathbf{r}}, t)$ to denote under sampled $k-t$ space data.

Patlak model[8] was utilized for parameter estimation from DCE MRI images as it is a popular two compartmental model, which parameterizes the tissue with two TK parameters, namely, $\mathbf{K_{trans}}$ and $\mathbf{V_p}$. The forward model was defined with these two TK parameters as inputs and the undersampled $k-t$ space as output. The forward model is shown in Figure 1 and is mathematically defined as shown in Equation (1):

$$\hat{\mathbf{k}}(\hat{\mathbf{r}}, t) = \hat{f}(\mathbf{K_{trans}}(r), \mathbf{V}_p(r)) = \mathbf{U}(\hat{\mathbf{r}}, t) \odot f(\mathbf{K_{trans}}(\mathbf{r}), \mathbf{V}_p(r)). \tag{1}$$

In Equation (1), $f(\cdot)$ maps the inputs to the fully sampled $k-t$ data and $\mathbf{U}(\hat{\mathbf{r}})$ is the undersampling map. $\odot$ signifies element-wise multiplication. The forward models $\hat{f}(\cdot)$ and $f(\cdot)$ are nonlinear functions and the inverse problem is ill-posed. More detailed information on them can be found in ref. [13]. Equation (1) can be solved using a DL-based technique as shown below:

$$\hat{\mathbf{K}}_{\mathbf{trans}}(\mathbf{r}), \hat{\mathbf{V}}_{\mathrm{p}}(r) = \tilde{f}(\mathrm{iFFT2D}(\hat{f}(\mathbf{K_{trans}}(\mathbf{r}), \mathbf{V}_{\mathrm{p}}(r))). \tag{2}$$

Here, $\hat{\mathbf{K}}_{\mathbf{trans}}(\mathbf{r})$ and $\hat{\mathbf{V}}_{\mathrm{p}}(r)$ are the estimated maps using neural network $\tilde{f}(\cdot)$. This work proposed a neural network that is invariant to the temporal dimension of the input data using a 2.5D Unet architecture, where the encoder is 3D and decoder is 2D. The architecture is explained in the next subsection.

2.1 | VTDCE-net architecture

The architecture of the network is shown in Figure 2 and consisted of an encoder with 3D convolutions and a decoder with 2D convolutions. The input to the network was a time series data where each slice (e.g axial) has "T" temporal acquisitions. For a single axial slice, the input was a three-dimensional data (Height×Width×T), whereas the output to the network were two-dimensional TK parameter maps $\mathbf{K_{trans}}$ and $\mathbf{V_p}$ each having dimensions Height×Width. The objective was for the network to be invariant to the number of input time steps and give a two-dimensional reconstruction. To do this at each level of the 3D encoder, just before downsampling, a five-dimensional data (Batch×Channels×Height×Width×T)—which is the output of the 3D convolution layer of encoder—was reduced to four-dimensional data—input to the 2D convolution layer of decoder—by taking mean along the temporal ("T") dimension. It reduced the input from (Batch×Channels×Height×Width×T) to (Batch×Channels×Height×Width×1), where the last dimension was later squeezed out. This allowed the

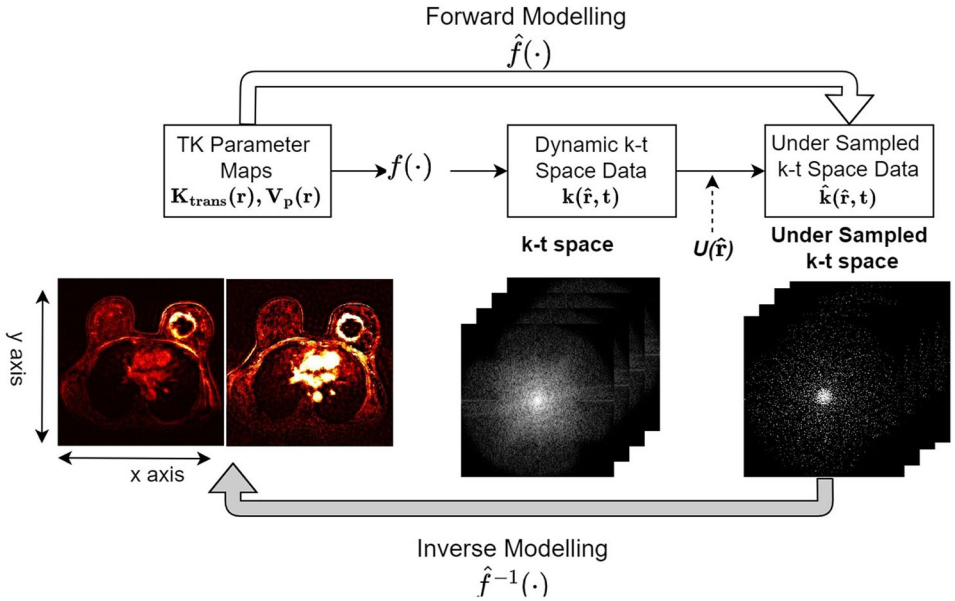

FIGURE 1 Block diagram of forward and inverse models for estimating the Tracer-kinetic parameters using Patlak model.[8] Randomized golden-angle radial scheme was utilized for obtaining undersampling mask $\mathbf{U(r, t)}$. The top-row presents the steps in forward modeling. The function $f(\cdot)$ is a non-linear function which maps the tracer-kinetic parameters onto the k-t space of anatomical images. $\hat{f}(\cdot)$ maps the tracer-kinetic parameters onto the undersampled k-t space of anatomical images, that is, the data acquired by Fast-MRI scanner. Black dashed arrows represent the parameters required for reconstruction. Grey shaded arrow depicts the inverse problem

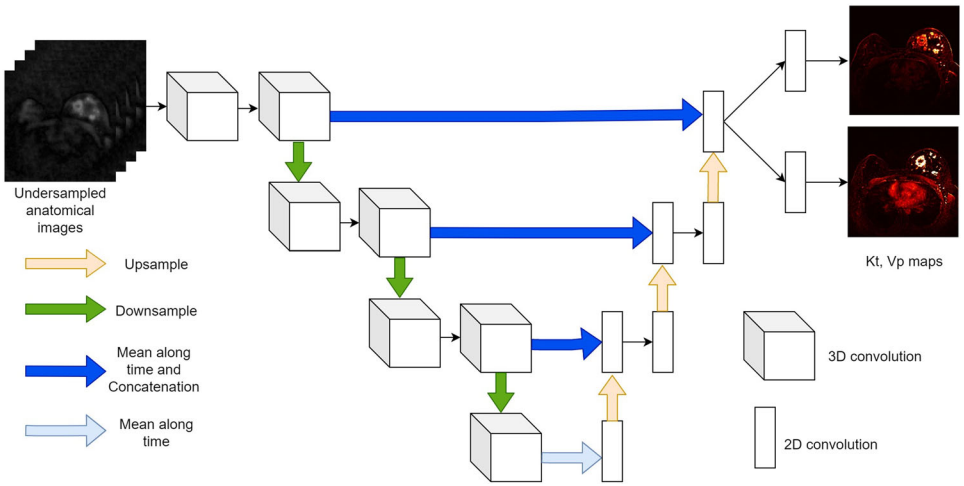

FIGURE 2 This figure depicts the architecture of VDCTE-Net. The encoder consists of 3D convolutions and downsampling and the decoder consists of 2D convolutions and upsampling. Sky blue denotes mean along the time axis, which converts 3D tensor into 2D tensor. The violet arrow denotes mean along time axis and concatenation. The last layer of Unet bifurcates into two branches one learns $\mathbf{K_{trans}}$ and the other learns $\mathbf{V_p}$

network to be trained and tested on different number of time samples as shown in our experiments. Our experiments showed that the results of the network were robust to change in time samples with this technique because the average was taken along the temporal dimension of the feature maps which are calculated using trained convolution layers. Moreover, the downsampling part along the Height, Width, and Temporal dimensions of the Unet architecture compressed the data and provides salience representation to have a robust quantitative imaging. Three levels of downsampling (upsampling) using maxpooling (transpose convolution) were utilized. The last layer of decoder was bifurcated into two channels to learn the two parameter maps. All convolutional filters (3D and 2D) had a filter-size of 3 with stride of 1. The convolutions were followed by batch-normalization and ReLU activation function (apart from the output layers).

2.2 | Datasets

For the experiments in this work, breast and brain DCE MRI datasets were utilized. The details of the same are in the next subsections.

2.2.1 | Breast dataset

Experiments on breast tissues were performed using QIN Breast DCE-MRI dataset,[27] which was available on The Cancer Imaging Archive (TCIA).[28] It contains data from a longitudinal study to assess the response to neoadjuvant chemotherapy.[29] Images were acquired at two points: before the first round of treatment (V1) and after the first round (V2) using Siemens 3T system with a body coil and a four-channel bilateral phased-array breast coil. The images were fat-saturated and were acquired using a 3D gradient echo-based TWIST sequence.[30] The data acquisition parameters included flip angle of 10°, TE of 2.9 ms, and TR of 6.2 ms, respectively. The field of view was 30–34 cm with 320×320 in-plane matrix size and a slice thickness of 1.4 mm. The dataset consisted of 4D DCE-MRI acquisition from 10 patients, with each patient being scanned before and after first round of neoadjunct chemotherapy. There are 26–32 image volumes, each containing 120–128 slices with a temporal resolution of 18–20 s and a total acquisition time of ∼10 min. A gadolinium-based contrast agent Gd-HP-D03A was administered with a dose of 0.1 mmol/kg of body mass followed by 20 ml saline flush at a speed of 2 ml/s using a programmable injector.

2.2.2 | Brain dataset

To show the robustness of VTDCE-Net, the network was also trained and tested on brain dataset. The brain DCE MRI dataset was made available by The Reference Image Database to Evaluate Therapy Response (Rider) NEURO MRI project[31] on TCIA[28] and consisted of scans of 19 patients with recurrent glioblastoma. Dynamic images were obtained during the intravenous injection of 0.1 mmol/kg of Magnevist at 3 ccs/s, started 24 s after the scan had begun on a 1.5 T scanner. The images were acquired using a 3D FLASH technique, with a flip angle of 25 degrees, TE of 1.8 ms, TR of 3.8 ms with voxel size of 1×1×5 mm, and in-plane matrix size of 256×256 . Each volume has 16 slices, obtained every 4.8 s with a total of 65 time samples.

2.3 | Implementation

The training of VTDCE-Net was performed for undersampling rate (**R**) of 8×, 12×, and 20× using radial golden angle (RGA)[11] undersampling pattern. RGA scheme follows a radial trajectory for data acquisition where the angle between two consecutive radial spokes is given by the golden ratio (∼ 1.618). The angle between two consecutive radial spokes is given as $\mathrm{mod}\left(\frac{360}{1.618} \times n, 360\right)$ for full spoke, where n is the number of spokes. The major advantage of RGA over uniform radial coverage is that in RGA, each new spoke always fills the largest gap of the previous coverage, and all spokes never repeat each other.[32] For each **R**, different networks were trained and tested. For brain dataset, training and testing were performed only for $\mathbf{R} = 20\times$. TV-based estimation scheme was also implemented for comparison as mentioned in ref. [13]. The implementation details of VTDCE-Net and TV scheme are provided below.

2.3.1 | VTDCE-net

For training on breast dataset, 492 3D images (2D + time) were utilized from six patients as training data and 82 3D images from two patients were used for validation. The input to the network was 3D data and outputs were $\mathbf{K_{trans}}$ and $\mathbf{V_p}$ maps. All samples of training and validation dataset had 32 time samples. Mean square error between reconstructed ground truth maps along with mean square error between their gradients was used as loss function, and ADAM[33] was used with parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$. The learning rate was kept at 1×10^{-4}, and training was performed for 200 epochs with a batchsize of four and patchsize of 128×128. For training, only those axial slices that had a visible tumor were utilized. For testing, unseen images of two patients were used. From the brain dataset of 19 patients, we used 10 for training (401 samples), 5 for validation, (71 samples) and 4 for testing.

2.3.2 | Total variation scheme

For implementing TV scheme, limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm (l-BFGS) algorithm[34] was utilized. The regularization parameters were tuned for each undersampling rate **R** and were kept constant for all patients.

2.4 | Computational implementation

All computations were carried out on a Linux workstation with an Intel i9 processor with 2.10 GHz clock speed, having 128 GB RAM and a Quadro RTX 8000 GPU with 48 GB memory. The pre-processing and post-processing steps were performed in MATLAB and implemented the neural networks in Pytorch. The

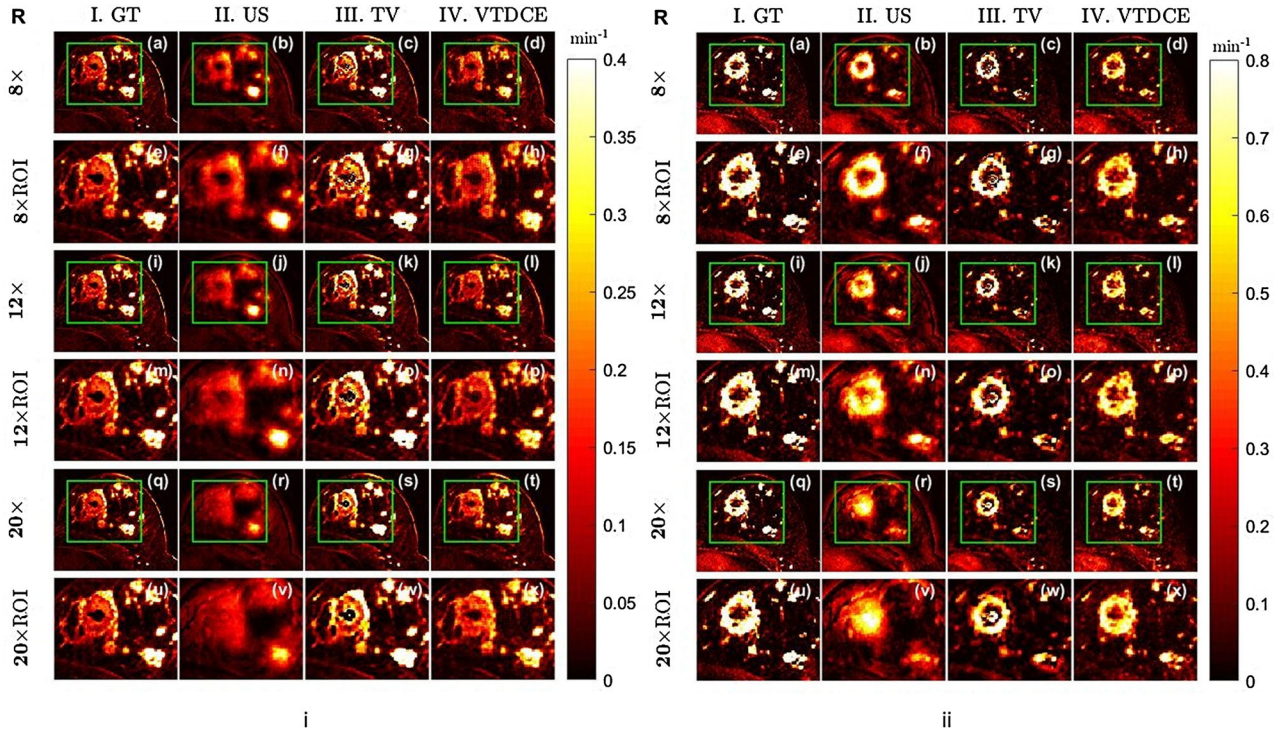

FIGURE 3 This figure illustrates the $\mathbf{K_{trans}}$ maps (left) and $\mathbf{V_p}$ maps (right) reconstructed using different reconstruction techniques and using different undersampling rates **R**. GT stands for ground truth, "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using total variation regularized scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. The even rows shows the zoomed in ROI bounded by the green box in figure above. In (i), it can be seen that as undersampling rate increases, the information in undersampled image (US) decreases. Both TV and VTDCE-Net are able to reconstruct the $\mathbf{K_{trans}}$ maps, but TV reconstruction shows hyper-intensive region toward the core of the tumor (which is generally not true as the core of tumor consists of dead tissue), whereas VTDCE reconstruction resembles more to the ground truth (GT). Similarly for $\mathbf{V_p}$ maps in (ii), one can see that VTDCE-Net provides improved performance than TV for all **R**; however, the maps are hypointense compared to the ground truth.

TV reconstruction algorithms were implemented on MATLAB 2018b using the parallel computing toolbox on 12 threads. Implementation of the proposed method, including the developed code, was made available as open-source at https://github.com/adityarastogi2k12/VTDCE.

3 | RESULTS

3.1 | Breast dataset

In this subsection, comparison of the performance of VTDCE-Net and TV parameter estimation method using PSNR, SSIM,[25] and HFEN[26] as metrics was performed. Note that the HFEN provides a measure of the fidelity at high spatial frequencies with lower-value indicating higher performance of reconstruction algorithm in terms of recovery of edge information.[26] The comparison was carried out on multiple slices of two patients "A" and "B." Patient "A" had 26 time samples, whereas patient "B" had 32 time samples. Network was trained and tested for $\mathbf{R} = 8\times$, $12\times$ and $20\times$. As previously mentioned, the training was completed on data with 32 time samples. Figure 3 shows $\mathbf{K_{trans}}$ and $\mathbf{V_p}$ maps estimated from a representative slice of patient "A" using different methods. "GT" stands for ground truth, "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, "TV" represents maps estimated using TV scheme, and "VTDCE" represents estimated maps using the proposed VTDCE-Net. In Figure 3-i, it can be seen that TV-estimated maps have hyper-intensive regions near the core of the tumor, which is not visible in the ground truth. This is generally not true because hypoxia and nutrient deficiency in the case of large solid tumors—due to insufficient blood supply—lead to necrosis. Figure 3-ii also shows that TV reconstruction results in hyperactivity in the necrotic region. $\mathbf{V_p}$ signifies the volume fraction of blood vessels/capillaries and therefore cannot show hyperactivity in the necrotic region. This artifact is not present in VTDCE-Net reconstruction. Table 1 shows the performance of the above-mentioned parameter estimation methods for multiple slices of two patients in reconstruction of $\mathbf{K_{trans}}$. The methods are compared using PSNR, SSIM, and HFEN for multiple undersampling rates. From

TABLE 1 This table shows PSNR, SSIM, and HFEN metrics for $\mathbf{K_{trans}}$ estimated using VTDCE and TV for $\mathbf{R} = 8\times, 12\times, 20\times$ for all slices of the Breast data of Patient "A" and Patient "B." "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using total variation regularized scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. From the table, it can be seen that proposed VTDCE performs better in terms of all metrics across both patients for all **R**.

		US			VTDCE			TV		
PAT	R	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN
PAT A	8×	34.74 ± 0.80	0.91 ± 0.01	62.91 ± 4.51	39.30 ± 0.84	0.95 ± 0.01	23.79 ± 2.09	38.28 ± 0.69	0.94 ± 0.01	24.04 ± 2.78
	12×	33.80 ± 0.79	0.89 ± 0.01	75.76 ± 5.06	39.66 ± 0.66	0.95 ± 0.00	23.65 ± 1.57	37.76 ± 0.60	0.95 ± 0.01	27.04 ± 3.13
	20×	32.55 ± 0.75	0.86 ± 0.01	96.25 ± 6.79	39.46 ± 0.74	0.94 ± 0.01	26.07 ± 1.90	37.05 ± 0.65	0.93 ± 0.01	31.35 ± 4.01
PAT B	8×	37.39 ± 0.55	0.91 ± 0.01	55.9 ± 4.11	39.53 ± 0.56	0.94 ± 0.01	30.68 ± 1.49	38.98 ± 0.67	0.95 ± 0.01	30.9 ± 2.75
	12×	36.54 ± 0.56	0.89 ± 0.01	64.32 ± 5.49	40.97 ± 0.56	0.95 ± 0.00	30.47 ± 2.03	38.98 ± 0.77	0.94 ± 0.01	31.99 ± 2.39
	20×	35.21 ± 0.62	0.86 ± 0.01	79.06 ± 5.89	38.95 ± 0.54	0.92 ± 0.00	33.75 ± 1.34	38.82 ± 0.81	0.93 ± 0.01	34.65 ± 2.21

TABLE 2 This table shows PSNR, SSIM, and HFEN metrics for $\mathbf{V_p}$ estimated using VTDCE and TV for $\mathbf{R} = 8\times, 12\times, 20\times$ for all slices of the Breast data of Patient "A" and Patient "B." "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using total variation regularized scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. From the table, it can be seen that VTDCE performs better in terms of both PSNR and SSIM across both patients and undersampling rates.

		US			VTDCE			TV		
PAT	R	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN
PAT A	8×	25.99 ± 0.48	0.64 ± 0.01	51.19 ± 5.29	29.21 ± 0.60	0.78 ± 0.01	30.18 ± 2.01	26.95 ± 0.85	0.74 ± 0.02	42.45 ± 3.84
	12×	25.12 ± 0.47	0.61 ± 0.01	60 ± 6.57	28.69 ± 0.65	0.76 ± 0.01	34.02 ± 1.94	26.53 ± 0.79	0.69 ± 0.02	46.01 ± 4.51
	20×	23.66 ± 0.49	0.57 ± 0.01	76.58 ± 9.01	27.14 ± 0.76	0.70 ± 0.01	36.89 ± 1.91	25.71 ± 0.82	0.62 ± 0.03	56.84 ± 8.49
PAT B	8×	25.05 ± 0.42	0.63 ± 0.01	33.1 ± 2.94	28.59 ± 0.36	0.79 ± 0.01	17.84 ± 0.88	23.84 ± 0.47	0.70 ± 0.02	27.79 ± 2.26
	12×	24.28 ± 0.44	0.60 ± 0.01	38.03 ± 3.58	28.52 ± 0.33	0.79 ± 0.01	19.71 ± 1.01	23.63 ± 0.46	0.66 ± 0.01	29.84 ± 2.77
	20×	22.93 ± 0.46	0.56 ± 0.01	47.47 ± 4.62	26.24 ± 0.41	0.68 ± 0.02	24.81 ± 0.94	23.28 ± 0.44	0.60 ± 0.01	33.03 ± 3.38

TABLE 3 This table shows PSNR and SSIM metrics for $\mathbf{K_{trans}}$ estimated using VTDCE and TV for $\mathbf{R} = 8\times, 12\times, 20\times$ in the R0I region for all slices of the Breast data of patients "A" and "B." "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using total variation regularized scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. From the table, it can be seen that VTDCE performs better in terms of PSNR for both patients and all **R** and in terms of SSIM, it also performs better than TV in most cases.

		US		VTDCE		TV	
PAT	R	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
PAT A	8 ×	29.06 ± 0.84	0.77 ± 0.02	33.36 ± 0.82	0.91 ± 0.01	32.68 ± 0.69	0.86 ± 0.01
	12 ×	28.06 ± 0.84	0.73 ± 0.02	34.46 ± 0.65	0.92 ± 0.00	32.22 ± 0.58	0.90 ± 0.00
	20 ×	26.76 ± 0.80	0.66 ± 0.02	34.02 ± 0.70	0.90 ± 0.00	31.58 ± 0.61	0.89 ± 0.01
PATB	8 ×	32.87 ± 0.76	0.817 ± 0.02	35.94 ± 1.43	0.93 ± 0.01	35.46 ± 2.65	0.92 ± 0.01
	12 ×	31.92 ± 0.56	0.78 ± 0.02	35.64 ± 1.46	0.92 ± 0.01	35.37 ± 2.73	0.91 ± 0.01
	20 ×	30.41 ± 0.52	0.72 ± 0.03	35.53 ± 1.24	0.91 ± 0.01	35.147 ± 2.80	0.90 ± 0.01

the table, it can be seen that VTDCE-Net performs better than TV for all undersampling rates for both patients in terms of PSNR and HFEN. In terms of SSIM, the VTDCE-Net performs at par or better than TV. Table 2 shows results for the estimation of $\mathbf{V_p}$ and it can be observed that VTDCE-Net performs better than TV scheme in terms of all metrics (PSNR, SSIM, and HFEN) by a margin of $\sim$ 3 dB in some cases. Tables 3 and 4 shows the performance of VTDCE-Net in the tumor region (ROI) of patients "A" and "B" for estimation of $\mathbf{K_{trans}}$ and $\mathbf{V_p}$ parameters. From the tables, it can be seen that the VTDCE-Net performs better than TV for both patients and for all undersampling rates by a margin of as much as $\sim$ 2.5 dB for $\mathbf{K_{trans}}$ and $\sim$ 3 dB for $\mathbf{V_p}$.

To examine if the proposed VTDCE-Net provided a significant improvement in performance of TK parameter estimation with TV-regularized algorithm, we performed a two tailed t-test for unequal variance for different undersampling rates. The null hypothesis was that the TV-based parameter estimation and the proposed VTDCE-Net have same mean performance in

TABLE 4 This table shows PSNR and SSIM metrics for $\mathbf{V_p}$ estimated using VTDCE and TV for $\mathbf{R} = 8\times, 12\times, 20\times$ in the R0I region for all slices of the Breast data of patients "A" and "B." "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using total variation regularized scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. From the table, it can be seen that VTDCE performs better in terms of PSNR and SSIM for both patients and all **R**.

		US		VTDCE		TV	
PAT	**R**	**PSNR**	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**
PAT A	8×	21.13 ± 0.43	0.48 ± 0.02	25.15 ± 0.73	0.71 ± 0.01	22.57 ± 1.11	0.69 ± 0.02
	12×	20.18 ± 0.49	0.43 ± 0.01	24.84 ± 0.82	0.71 ± 0.02	22.27 ± 1.06	0.65 ± 0.02
	20×	18.64 ± 0.61	0.37 ± 0.01	24.06 ± 0.77	0.67 ± 0.02	21.39 ± 1.10	0.55 ± 0.03
PAT B	8×	22.73 ± 0.54	0.54 ± 0.01	27.55 ± 0.49	0.76 ± 0.02	24.16 ± 0.67	0.65 ± 0.02
	12×	22.02 ± 0.52	0.51 ± 0.01	27.45 ± 0.40	0.76 ± 0.02	23.94 ± 0.72	0.62 ± 0.01
	20×	20.45 ± 0.52	0.44 ± 0.02	25.19 ± 0.46	0.64 ± 0.02	23.51 ± 0.62	0.55 ± 0.01

TABLE 5 This table shows the t-test for the $\mathbf{K_{trans}}$ maps of two patients for R 8×, 12×, 20×. The t-test is done for PSNR, and average values of PSNR for multiple slices of same patients are tabulated. The null hypothesis is that the performance of VTDCE-Net and TV are same in terms of mean PSNR. p-Value of < 0.05 shows significant difference in the performance of two methods. "Sig ?" denotes if the difference is significant to reject the null hypothesis (p-value < 0.05) or not.

		PSNR			Sig ?
PAT		**VTDCE**	**TV**	***p*-Value**	**(*p*-value < 0.05)**
PAT A	8×	39.30 ± 0.84	38.28 ± 0.69	$4.65\ e^{-5}$	Yes
	12×	39.66 ± 0.66	37.76 ± 0.60	$6.56\ e^{-12}$	Yes
	20×	39.46 ± 0.74	37.05 ± 0.65	$1.70\ e^{-12}$	Yes
PAT B	8×	39.53 ± 0.56	38.98 ± 0.67	0.0034	Yes
	12×	40.97 ± 0.56	38.98 ± 0.77	$5.22\ e^{-10}$	Yes
	20×	38.95 ± 0.54	38.82 ± 0.81	0.46	No

TABLE 6 This table shows the t-test for the $\mathbf{V_p}$ maps of two patients for R 8×, 12×, 20×. The t-test is done for PSNR, and average values of PSNR for multiple slices of same patients are tabulated. The null hypothesis is that the performance of VTDCE-Net and TV are same in terms of mean PSNR. p-Value of < 0.05 shows significant difference in the performance of two methods. "Sig ?" denotes if the difference is significant to reject the null hypothesis (p-value < 0.05) or not.

		PSNR			Sig ?
PAT		**VTDCE**	**TV**	***p*-Value**	**(*p*-value < 0.05)**
PAT A	8×	29.21 ± 0.60	26.95 ± 0.85	$3.72e^{-14}$	Yes
	12×	28.69 ± 0.65	26.53 ± 0.79	$3.53e^{-14}$	Yes
	20×	27.14 ± 0.76	25.71 ± 0.82	$5.27e^{-8}$	Yes
PAT B	8×	28.59 ± 0.36	23.84 ± 0.47	$1.70e^{-29}$	Yes
	12×	28.52 ± 0.33	23.63 ± 0.46	$2.48e^{-25}$	Yes
	20×	26.24 ± 0.41	23.28 ± 0.44	$3.27e^{-24}$	Yes

terms of PSNR. This test was conducted on averaged PSNR value for $R = 8\times, 12\times$, and $20\times$ for patients "A" and "B." Significance level $\alpha = 0.05$ was selected for testing. The results are shown for $\mathbf{K_{trans}}$ maps in Table 5 and for $\mathbf{V_p}$ maps in Table 6. Here, p denotes the probability that the difference in performance of these two methods happened by chance. A p-value of < 0.05 means that there is a significant difference in performance of these two methods. From the tables, it can be seen that for both parameter maps, the proposed VTDCE-Net performs significantly better than TV algorithm for both patients (except for the last reported result).

3.2 | Brain dataset

To show that the proposed VTDCE-Net can also work for other imaging scenarios, the performance of the proposed VTDCE-Net and TV scheme was compared on brain dataset using PSNR, SSIM, and HFEN as metrics. The comparison was carried out on multiple slices of four patients, all of which had 65 time samples. Network was trained and tested for $\mathbf{R} = 20\times$. The performance of the proposed VTDCE-Net and TV are illustrated in Figure 4, which shows the $\mathbf{K_{trans}}$ maps (Figure 4-i) and $\mathbf{V_p}$ maps (Figure 4-ii) reconstructed using different reconstruction techniques for undersampling rate $\mathbf{R} = 20\times$. "GT" stands for ground truth, "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using TV scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. The rows depict the patient name, and for each patient, two representative slices (z1 and z2) are shown. In Figure 4-i, it can be seen that undersampling leads to loss of information in parameter maps. $\mathbf{K_{trans}}$ map estimated by TV shows hyperintensive tumor regions along with higher permeability in non-tumorous regions inside the brain, whereas VTDCE estimation resembles more to the ground truth (GT). Similarly for $\mathbf{V_p}$ maps, it can be seen that proposed VTDCE-Net provide improved performance than TV for all patients.

Table 7 shows the performance of above mentioned parameter estimation methods for multiple slices of four patients in reconstruction of $\mathbf{K_{trans}}$. From the table, it can be observed that the proposed VTDCE-Net performs better than TV for all patients in terms of PSNR and HFEN and is at par or better than TV scheme in terms

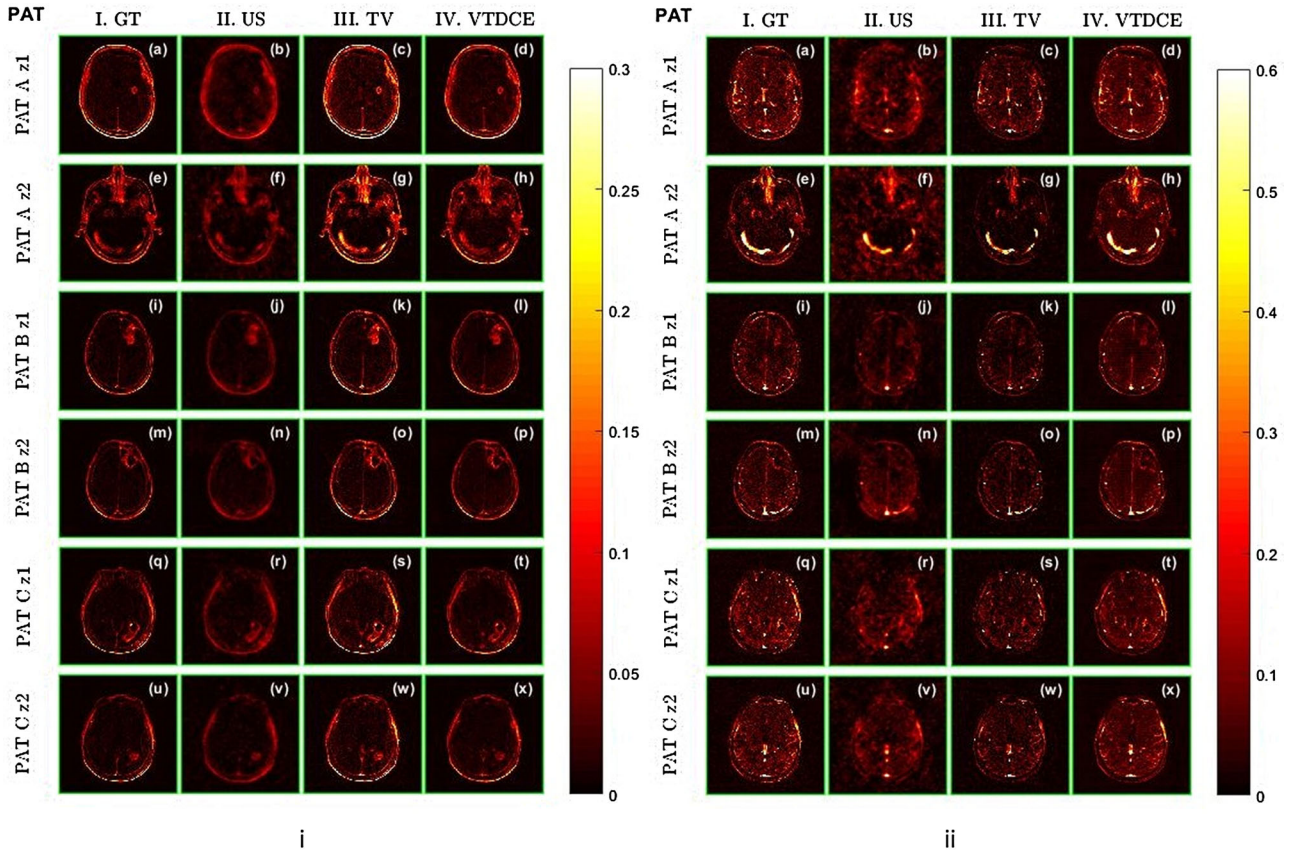

FIGURE 4 This figure illustrates the $\mathbf{K_{trans}}$ maps (left) and $\mathbf{V_p}$ maps (right) reconstructed using different reconstruction techniques for undersampling rate $\mathbf{R} = 20\times$. "GT" stands for ground truth, "US" stands for maps estimated from undersampled data using zero-filled $k - t$ space, TV represents maps estimated using total variation scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. The rows depict the patient name, and for each patient, two representative slices (z1 and z2) are shown. In (i) and (ii), it can be seen that undersampling leads to information loss. Both TV and VTDCE-Net are able to reconstruct the $\mathbf{K_{trans}}$ maps, but TV reconstruction shows hyper-intensive $\mathbf{K_{trans}}$ maps which show high permeability in non-tumorous region inside the brain, whereas VTDCE reconstruction resembles more to the ground truth (GT). Similarly, for $\mathbf{V_p}$ maps in (ii), one can see that VTDCE-Net provides improved performance than TV for all patients. Viewers are advised to see the figures in full brightness of their screen.

TABLE 7 This table shows PSNR, SSIM, and HFEN metrics for $\mathbf{K_{trans}}$ estimation using VTDCE and TV for $\mathbf{R} = 20\times$ for all slices of four patients. "US" stands for maps estimated from undersampled data using zero-filled $k - t$ space, TV represents maps estimated using total variation scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. From the table, it can be observed that VTDCE performs better in terms of both PSNR and SSIM in case of all patients.

	US			VTDCE			TV		
PAT	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN
PAT A	32.87 ± 0.93	0.79 ± 0.04	76.75 ± 3.43	38.50 ± 1.01	0.93 ± 0.02	21.78 ± 4.01	36.62 ± 2.52	0.92 ± 0.04	28.11 ± 3.75
PAT B	37.08 ± 1.87	0.87 ± 0.05	76.71 ± 10.39	41.38 ± 1.47	0.95 ± 0.01	24.98 ± 3.21	40.99 ± 0.86	0.95 ± 0.00	28.73 ± 1.07
PAT C	35.21 ± 1.05	0.84 ± 0.04	80.13 ± 9.03	39.31 ± 1.63	0.93 ± 0.02	28.69 ± 3.6	37.04 ± 3.84	0.88 ± 0.16	32.83 ± 4.68
PAT D	32.30 ± 0.78	0.79 ± 0.04	87.65 ± 4.47	38.34 ± 1.16	0.93 ± 0.01	23.25 ± 1.48	36.65 ± 2.13	0.91 ± 0.04	26.43 ± 2.07

of SSIM. Table 8 shows results for the estimation of $\mathbf{V_p}$, and similar to Table 7, it can be observed that the proposed VTDCE-Net performs better than TV scheme in terms of PSNR, SSIM, and HFEN by a margin of ∼ 2–4 dB in terms of PSNR and as much as 0.12 points in terms of SSIM. The proposed VTDCE-Net reduced the HFEN by as much as 44% over TV. Table 9 shows the performance of VTDCE-Net in the tumor region (ROI) of patients "A," "B," and "C" for estimation of $\mathbf{K_{trans}}$ parameter for undersampling rate of 20×. From the table, it can be seen that the VTDCE-Net performs better than TV for all three patients and by a margin of as much as ∼ 4 dB. Only results for $\mathbf{K_{trans}}$ parameter is shown in ROI because it is the more important parameter in tumor

TABLE 8 This table shows PSNR, SSIM, and HFEN metrics for $\mathbf{V_p}$ estimation using VTDCE and TV for $\mathbf{R} = 20\times$ for all slices of four patients. "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using total variation scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. From the table, it can be observed that VTDCE performs better in terms of PSNR for all patients and in terms of SSIM for three out of four patients.

	US			VTDCE			TV		
PAT	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN	PSNR	SSIM	HFEN
PAT A	25.49 ± 0.95	0.52 ± 0.04	65.76 ± 6.33	29.37 ± 0.68	0.74 ± 0.01	27.11 ± 3.3	25.11 ± 2.04	0.70 ± 0.04	48.49 ± 7.74
PAT B	28.24 ± 1.43	0.60 ± 0.07	71.27 ± 7.52	31.83 ± 1.08	0.77 ± 0.03	30.62 ± 5.06	29.95 ± 0.87	0.65 ± 0.01	45.85 ± 6.78
PAT C	26.32 ± 1.34	0.55 ± 0.05	69.45 ± 7.13	29.85 ± 0.86	0.74 ± 0.01	30.22 ± 2.84	26.18 ± 1.91	0.71 ± 0.03	45.92 ± 3.49
PAT D	25.17 ± 1.03	0.50 ± 0.05	72.55 ± 4.33	29.72 ± 0.70	0.74 ± 0.02	27.52 ± 1.24	26.46 ± 1.40	0.62 ± 0.02	49.19 ± 5.45

TABLE 9 This table shows PSNR and SSIM metrics for $\mathbf{K_{trans}}$ estimation using VTDCE and TV for $\mathbf{R} = 20\times$ in **ROI** for all slices of three patients. "US" stands for maps estimated from undersampled data using zero-filled $k-t$ space, TV represents maps estimated using total variation scheme, and VTDCE represents estimated maps using the proposed VTDCE-Net. From the table, it can be observed that VTDCE performs better in terms of both PSNR and SSIM in case of all patients.

	US		VTDCE		TV	
PAT	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
PAT A	35.71 ± 1.95	0.83 ± 0.03	38.06 ± 2.08	0.88 ± 0.04	37.23 ± 2.85	0.86 ± 0.05
PAT B	36.15 ± 2.08	0.84 ± 0.03	39.20 ± 1.28	0.91 ± 0.01	37.09 ± 2.32	0.89 ± 0.01
PAT C	30.01 ± 1.15	0.75 ± 0.05	34.09 ± 1.20	0.84 ± 0.03	29.92 ± 3.37	0.79 ± 0.14

TABLE 10 Repeated held-out four-fold cross validation for results for $\mathbf{K_{trans}}$ parameter estimation of patients "A" and "B"

	Fold 1		Fold 2		Fold 3		Fold 4	
PAT	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
PAT A	39.66 ± 0.66	0.95 ± 0.00	39.51 ± 0.78	0.95 ± 0.01	39.83 ± 0.49	0.96 ± 0.00	39.27 ± 0.53	0.94 ± 0.01
PAT B	40.97 ± 0.56	0.95 ± 0.00	41.39 ± 0.79	0.95 ± 0.00	40.66 ± 0.71	0.94 ± 0.01	40.42 ± 0.61	0.94 ± 0.01

region than $\mathbf{V_p}$. Moreover, results for only three patients are shown in Table 9, as the tumor region in patient "D" could not be properly identified.

3.3 | Cross validation

To demonstrate the robustness of the proposed architecture to training data, a repeated held-out fourfold cross validation study was performed on VTDCE-Net for 12× undersampling rate on breast dataset. In each of them, the validation data (data of two patients) was swapped with two different patients from training data. The results in terms of averaged PSNR (in dB) and SSIM of the test data are shown in Table 10 for $\mathbf{K_{trans}}$ and Table 11 for $\mathbf{V_p}$. It was found that the performance of proposed VTDCE-Net is similar in all cases, which suggests that the network has not overfitted and provides a good generalization.

3.4 | Runtime

The runtimes of VTDCE-Net and TV scheme for $\mathbf{K_{trans}}$ and $\mathbf{V_p}$ map estimation of a single axial slice are tabulated in Table 12 (rounded off to nearest second). From the table, one can deduce that the VTDCE-Net, which is a neural network based algorithm, is the faster algorithm to reconstruct the tracer kinetic maps.

3.5 | Comparison with time-variant method

The novelty of proposed VTDCE-Net compared to existing networks is that it is able to handle multiple time samples. As there is no fixed standard defining the number of volumetric acquisitions, the clinicians determine this either based on their domestic protocols or based on the patient history. The latter is important because for accurate estimation of tracer-kinetic parameters, the contrast agent concentration profile should show early rise and a long tail. This might not be true for patients with slow perfusion near the region of interest due to higher mean transit time. In such cases, more number of time samples need to be acquired. Similarly, there are cases in which the clinicians have to reduce the number of time samples in order to reduce scan time such as in the case of late-stage Alzheimer's and dementia patients, who cannot stay steady for long. One

TABLE 11 Repeated held-out four-fold cross validation for results for $\mathbf{V_p}$ parameter estimation of patients "A" and "B"

PAT	Fold 1		Fold 2		Fold 3		Fold 4	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
PAT A	28.69 ± 0.65	0.76 ± 0.01	28.26 ± 0.71	0.75 ± 0.01	28.85 ± 0.69	0.76 ± 0.00	28.47 ± 0.63	0.76 ± 0.01
PAT B	28.52 ± 0.33	0.79 ± 0.01	28.90 ± 0.72	0.80 ± 0.01	28.76 ± 0.52	0.79 ± 0.00	29.11 ± 0.63	0.80 ± 0.00

TABLE 12 This table shows the approximate reconstruction time for $\mathbf{K_{trans}}$ and $\mathbf{V_p}$ for VTDCE-Net and TV scheme for a single slice.

Method	Reconstruction time (s)
TV	65
VTDCE-Net	0.02

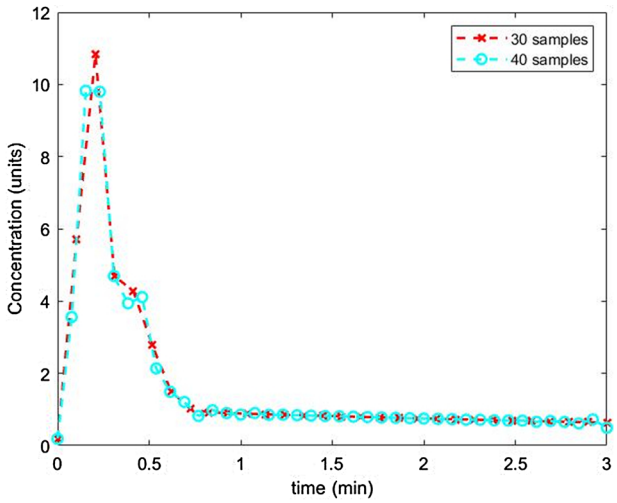

FIGURE 5 Concentration profile at location where $\mathbf{V_p} = 1$. Red curve denotes the original sampling of 30 points in a span of 3 min. The cyan curve shows the resampled points.

solution is to resample the data to match the temporal dimension on which the network is trained. However, the experiments in this work have shown that it is not the best strategy. Figure 5 shows the simulated concentration profile as a function of time from a region where $\mathbf{V_p} = 1$. The red curve denotes the original profile sampled 30 times in a duration of 3 min. The cyan curve shows the same profile resampled to have 40 time points in 3 min using a polyphase anti-aliasing filter. From the graph, it can be observed that the peak concentration and the wash-in[35] slope of the resampled profile are different from the original profile. This can lead to difference in the estimated parameter maps. To demonstrate this effect, another Unet was trained to estimate tracer-kinetic parameters from undersampled DCE MRI data of breast tissue for $\mathbf{R} = 8\times$. However, it treated the temporal dimensions as number of channels and therefore was not flexible to temporal samples. Similar to VTDCE-Net, three levels of downsampling/upsampling were performed and used the same loss function. This

TABLE 13 This table shows the performance of DCE-Net and VTDCE-Net for estimation of $\mathbf{K_{trans}}$ of patient "A" for undersampling rate of 8×. For DCE-Net, the time samples were upsampled to 32 from original acquisition of 26 to match the training data temporal dimension. The reconstructed distribution was not shown here.

Method	PSNR (dB)
DCE-Net	38.1
VTDCE-Net	39.30

TABLE 14 Performance of network for $\mathbf{K_{trans}}$ estimate when trained on higher undersampling **R** and tested for lower **R** for PAT "A" of breast dataset. Compared to Table 1, the performance deteriorates significantly. This shows that the network does not work well when trained and tested for different undersampling rates.

Trained/ Tested	8×		12×	
	PSNR	SSIM	PSNR	SSIM
20×	32.27	0.88	34.44	0.92
12×	35.80	0.93		

network was called as DCE-Net and it was tested on patient "A" by resampling the time points from 26 to 32, and the results are tabulated in Table 13. It can be seen that VTDCE-Net, which is flexible to the number of time samples, gives better performance than a network which requires resampling. This shows that resampling is not an ideal strategy and it is useful to have a network which is invariant to number of time samples.

3.6 | Generalization for different undersampling rates

To test if a network trained for higher undersampling rate can estimate TK parameters accurately for lower undersampling rate, we performed an experiment on patient "A" of breast dataset and tabulated the results in Table 14. The table shows the performance of VTDCE-Net when trained for a higher undersampling rate and tested for a lower undersampling rate. When compared to Table 1, the performance of VTDCE-Net deteriorated when trained and tested for different undersampling rates. This was primarily because for different undersampling rates, different degrees of degradation/corruption were caused in the TK parameters. When the network was only trained for one undersampling rate, it only saw the degradation of that degree and the

weights were trained accordingly. Hence, the network did not generalize well for even lower undersampling rate while testing. However, from the table, one can see that the closer training **R** was to testing **R**, the better was the recovery.

4 | CONCLUSION

This work proposed a robust neural network for direct parameter estimation from undersampled DCE-MRI data that is invariant to spatiotemporal dimensions. Systematic comparison with another popular direct and robust parameter estimation technique, that is, TV-based reconstruction, was carried out. This work also demonstrates that the proposed VTDCE-Net outperforms TV-based reconstruction on breast dataset for undersampling rates of 8×, 12×, and 20×. To demonstrate the generalizability of the proposed network, training and testing on brain dataset for **R** = 20× were also performed. This work also showed that techniques like temporal resampling are not the best option when it comes to using neural networks for parameter estimation, especially in the physiological imaging methods like DCE-MRI. However, a major limitation of this architecture is that network cannot generalize across different undersampling rates as shown in Section 3.6, that is, a new network has to be trained for each undersampling rate. Another major limitation of this study is that it has utilized relatively a smaller dataset. Detailed experimentation on a much larger dataset is required for knowing the clinical utility. Our future work will focus on addressing this limitation.

ACKNOWLEDGMENTS

This work was supported by the Science & Engineering Research Board (SERB) core research grant (No. CRG/2018/000672). Aditya Rastogi acknowledges the Prime Minister Research Fellowshop (PMRF).

CONFLICT OF INTEREST

There are no conflicts of interest declared by the authors.

DATA AVAILABILITY STATEMENT

Both the breast and brain datasets were obtained from The Cancer Imaging Archive (https://wiki.cancerimagingarchive.net/display/Public/Wiki). The trained network has been made available for enthusiastic users as open source at https://github.com/adityarastogi2k12/VTDCE.

REFERENCES

1. Lui YW, Tang ER, Allmendinger AM, Spektor V. Evaluation of CT perfusion in the setting of cerebral ischemia: patterns and pitfalls *Am J Neuroradiol*. 2010;31:1552-1563.
2. Gribbestad IS, Gjesdal KI, Nilsen G, Lundgren S, Hjelstuen MH, Jackson A. An introduction to dynamic contrast-enhanced MRI in oncology. In: *Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Oncology*. Springer; 2005:1-22.
3. Rahmim A, Lodge MA, Karakatsanis NA, et al. Dynamic whole-body pet imaging: principles, potentials and applications. *Eur J Nucl Med Mol Imaging*. 2019;46:501-518.
4. Lagopoulos J. Diffusion tensor imaging: an overview. *Acta Neuropsychiatr*. 2007;19:127-128.
5. Yan Y, Sun X, Shen B. Contrast agents in dynamic contrast-enhanced magnetic resonance imaging. *Oncotarget*. 2017; 8:43491-43505.
6. Tofts PS, Kermode AG. Measurement of the blood-brain barrier permeability and leakage space using dynamic mr imaging. 1. fundamental concepts. *Magn Reson Med*. 1991;17:357-367.
7. Tofts PS, Brix G, Buckley DL, et al. Estimating kinetic parameters from dynamic contrast-enhanced t1-weighted MRI of a diffusable tracer: standardized quantities and symbols. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*. 1999;10: 223-232.
8. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab*. 1983;3:1-7. PMID: 6822610.
9. Jaspan ON, Fleysher R, Lipton ML. Compressed sensing MRI: a review of the clinical literature. *Br J Radiol*. 2015;88:20150487.
10. Smith DS, Welch EB, Li X, et al. Quantitative effects of using compressed sensing in dynamic contrast enhanced MRI. *Phys Med Biol*. 2011;56:4933.
11. Feng L, Grimm R, Block KT, et al. Golden-angle radial sparse parallel mri: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magn Reson Med*. 2014;72:707-717.
12. Rosenkrantz AB, Geppert C, Grimm R, et al. Dynamic contrast-enhanced mri of the prostate with high spatiotemporal resolution using compressed sensing, parallel imaging, and continuous golden-angle radial sampling: preliminary experience. *J Magn Reson Imaging*. 2015;41:1365-1373.
13. Rastogi A, Yalavarthy PK. Comparison of iterative parametric and indirect deep learning-based reconstruction methods in highly undersampled DCE-MR imaging of the breast. *Med Phys*. 2020;47:4838-4861.
14. Aggarwal HK, Mani MP, Jacob M. Modl: Model-based deep learning architecture for inverse problems. *IEEE Trans Med Imaging*. 2018;38:394-405.
15. Zhang J, Ghanem B. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2018:1828-1837.
16. Guo Y, Lingala SG, Zhu Y, Lebel RM, Nayak KS. Direct estimation of tracer-kinetic parameter maps from highly undersampled brain dynamic contrast enhanced MRI. *Magn Reson Med*. 2017;78:1566-1578.
17. Guo Y, Lingala SG, Bliesener Y, Lebel RM, Zhu Y, Nayak KS. Joint arterial input function and tracer kinetic parameter estimation from undersampled dynamic contrast-enhanced MRI using a model consistency constraint. *Magn Reson Med*. 2018;79:2804-2815.
18. Dikaios N, Arridge S, Hamy V, Punwani S, Atkinson D. Direct parametric reconstruction from undersampled (k, t)-space data in dynamic contrast enhanced MRI. *Med Image Anal*. 2014;18:989-1001.
19. Kojima S, Shinohara H, Hashimoto T, Hirata M, Ueno E. Iterative image reconstruction that includes a total variation regularization for radial MRI. *Radiol Phys Technol*. 2015;8:295-304.
20. Bliesener Y, Acharya J, Nayak KS. Efficient dce-mri parameter and uncertainty estimation using a neural network. *IEEE Trans Med Imaging*. 2019;39:1712-1723.

21. Ulas C, Das D, Thrippleton MJ, et al. Convolutional neural networks for direct inference of pharmacokinetic parameters: application to stroke dynamic contrast-enhanced MRI. *Front Neurol*. 2019;9:1147.
22. Ulas C, Tetteh G, Thrippleton MJ, et al. Direct estimation of pharmacokinetic parameters from DCE-MRI using deep CNN with forward physical model loss. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2018:39-47.
23. Kettelkamp J, Lingala SG. Arterial input function and tracer kinetic model-driven network for rapid inference of kinetic maps in dynamic contrast-enhanced MRI (AIF-TK-net). In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2020:1450-1453.
24. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*. Springer; 2015:234-241.
25. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP, et al. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process*. 2004;13:600-612.
26. Ravishankar S, Bresler Y. Mr image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans Med Imaging*. 2010;30:1028-1041.
27. Huang W, Li X, Chen Y, et al. Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2014.A2N1IXOX "2014".
28. Clark K, Vendt B, Smith K, et al. The cancer imaging archive (tcia): Maintaining and operating a public information repository. *J Digit Imaging*. 2013;26:1045-1057.
29. Huang W, Li X, Chen Y, et al. Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge. *Transl Oncol*. 2014;7:153-166.
30. Le Y, Kroeker R, Kipfer HD, Lin C. Development and evaluation of TWIST dixon for dynamic contrast-enhanced (DCE) MRI with improved acquisition efficiency and fat suppression. *J Magn Reson Imaging*. 2012;36:483-491.
31. Barboriak D. Data from rider_neuro_mri. *The Cancer Imaging Archive*. 2015;577.
32. Feng L. Golden-angle radial MRI: basics, advances, and applications. *J Magn Reson Imaging*. 2022;56:45-62.
33. Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. 2014.
34. Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. *Math Program*. 1989;45:503-528.
35. Nagel E, Klein C, Paetsch I, et al. Magnetic resonance perfusion measurements for the noninvasive detection of coronary artery disease. *Circulation*. 2003;108:432-437.

How to cite this article: Rastogi A, Dutta A, Yalavarthy PK. VTDCE-Net: A time invariant deep neural network for direct estimation of pharmacokinetic parameters from undersampled DCE MRI data. *Med Phys*. 2023;50:1560–1572. https://doi.org/10.1002/mp.16081