IVIM-Morph: Motion-compensated quantitative Intra-voxel Incoherent Motion (IVIM) analysis for functional fetal lung maturity assessment from diffusion-weighted MRI data

Noga Kertes[a,1], Yael Zaffrani-Reznikov[a,1], Onur Afacan[b], Sila Kurugol[b], Simon K. Warfield[b], Moti Freiman[a,*]

[a]*Faculty of Biomedical Engineering, Technion, Haifa, Israel*
[b]*Boston Children's Hospital, Boston, MA, USA*

Abstract

Quantitative analysis of pseudo-diffusion in diffusion-weighted magnetic resonance imaging (DWI) data shows potential for assessing fetal lung maturation and generating valuable imaging biomarkers. Yet, the clinical utility of DWI data is hindered by unavoidable fetal motion during acquisition. We present IVIM-morph, a self-supervised deep neural network model for motion-corrected quantitative analysis of DWI data using the Intra-voxel Incoherent Motion (IVIM) model. IVIM-morph combines two sub-networks, a registration sub-network, and an IVIM model fitting sub-network, enabling simultaneous estimation of IVIM model parameters and motion. To promote physically plausible image registration, we introduce a biophysically informed loss function that effectively balances registration and model-fitting quality. We validated the efficacy of IVIM-morph by establishing a correlation between the predicted IVIM model parameters of the lung and gestational age (GA) using fetal DWI data of 39 subjects. Our approach was compared against six baseline methods: 1) no motion compensation, 2) affine registration of all DWI images to the initial image, 3) deformable registration of all DWI images to the initial image, 4) deformable registration of each DWI image to its preceding image in the sequence, 5) iterative deformable motion compensation combined with IVIM model parameter estimation, and 6) self-supervised deep-learning-based deformable registration. IVIM-morph exhibited a notably improved correlation with gestational age (GA) when performing in-vivo quantitative analysis of fetal lung DWI data during the canalicular phase. Specifically, over 2 test groups of cases, it achieved an R_f^2 of 0.44 and 0.52, outperforming the values of 0.27 and 0.25, 0.25 and 0.00, 0.00 and 0.00, 0.38 and 0.00, and 0.07 and 0.14 obtained by other methods. IVIM-morph shows potential in developing valuable biomarkers for non-invasive assessment of fetal lung maturity with DWI data. Moreover, its adaptability opens the door to potential applications in other clinical contexts where motion compensation is essential for quantitative DWI analysis. The IVIM-morph code is readily available at:`https://github.com/TechnionComputationalMRILab/qDWI-Morph`.

1. Introduction

Congenital pulmonary hypoplasia (PH) is a congenital abnormality marked by insufficient growth of the lung parenchyma Lakshminrusimha and Keszler (2015). This condition can result in significant and potentially fatal physiological impairments, including respiratory distress syndrome and transient tachypnea of the newborn Ahmed and Konje (2021). Approximately 10-15% of newborn deaths are caused by PH Kumar and Burton (2007).

Currently, the methods employed for the antenatal diagnosis of pulmonary hypoplasia include amniocentesis, prenatal ultrasound (US), and MRI. Amniocentesis involves extracting a small volume of amniotic fluid to as-

*Corresponding author: Moti Freiman, Tel.: +972-77-887-4147; Email: moti.freiman@technion.ac.il

[1]Equal contribution.

sess surfactant protein levels, which are considered an indicator of fetal lung maturity Rome et al. (1975). Prenatal ultrasound is a widely utilized technique for evaluating fetal lung maturity. However, it is primarily used to assess fluid parameters Beck et al. (2015) or to estimate fetal lung volume Moeglin et al. (2005). Similarly, fetal lung MRI can estimate fetal lung volume Ward et al. (2006). Nonetheless, these modalities do not provide sufficient insight into lung function and are, therefore, suboptimal for assessing functional fetal lung maturity and PH Avena-Zampieri et al. (2022).

Diffusion-weighted MRI (DWI) is a non-invasive imaging modality that is highly sensitive to the random motion of water molecules, which is primarily due to the water's thermal energy. In living tissues, the motion of water molecules is influenced and restricted by interactions with cell membranes and macromolecules. Moreover, the motion of water molecules is more confined in tissues with higher cellular density, while the motion of water molecules is less restricted in areas of low cellularity. In DWI, the random displacement of individual water molecules leads to signal attenuation when magnetic field encoding gradient pulses are applied at varying magnitudes and durations known as the "b-value" Koh and Collins (2007).

Quantitative analysis of DWI (qDWI) using multi-compartment signal decay models such as the Intravoxel Incoherent Motion (IVIM) Model Iima and Le Bihan (2016), can provide a separate assessment of diffusion and pseudo-diffusion in tissue. This approach allows for more precise imaging biomarkers that capture the key characteristics of functional lung maturity and PH such as the formation of a dense capillary network, an increase in pulmonary blood flow, a reduction in extracellular space, and an increase in tissue perfusion Ercolani et al. (2021); Korngut et al. (2022); Jakab et al. (2017).

However, the inevitable motion of the fetus during lengthy DWI acquisitions generally leads to inaccurate and unreliable quantitative analysis of diffusion and pseudo-diffusion, which effectively renders these imaging biomarkers of little utility in assessing functional lung maturity and PH in the clinical setting. Fig. 1 illustrates the deviations of the observed DWI signal acquired at different b-values from the signal decay model. For instance, Afacan et al. (2016) reported that nearly 40% (26 out of 65 cases) in their study cohort had severe motion artifacts, which essentially prevented the functional assessment of lung maturity with DWI. Thus, there is a critical need to develop methods for qDWI analysis that are robust to the presence of inter-volume motion in fetal DWI data.

Image registration algorithms have been previously used to address inter-volume motion before qDWI analysis. For instance, Guyader et al. Guyader et al. (2015) demonstrated improved accuracy and reliability in apparent diffusion coefficient (ADC) qDWI analysis of abdominal organs when employing initial motion correction, as opposed to qDWI analysis conducted without any motion correction. However, the registration of DWI images obtained using varying b-values may lead to suboptimal accuracy owing to differences in image contrast caused by varying sensitivities to diffusion and pseudo-diffusion effects. Registration of high b-value images, which have a low signal-to-noise ratio (SNR) by nature, also poses a significant challenge. Moreover, it is worth noting that optimization processes typically optimize loss functions tied to pairwise metrics, such as Dice similarity or intensity dissimilarity. These metrics, inherently designed for pairwise comparisons, possess limitations in their ability to comprehensively address motion across the entire set of DWI images concurrently. Furthermore, their primary focus tends to be on aligning image edges, rather than ensuring precise alignment of the observed signal decay within regions of interest with the signal decay model.

In the context of abdominal imaging, Kurugol et al. (2017) introduced an iterative motion correction model to address the differences in image contrast in the DWI images by registering images and estimating parameters with the IVIM model. Similarly, Sanz-Estébanez et al. (2018) simultaneously compensates for motion and performs qDWI analysis using a mono-exponential signal decay model. However, these techniques involve an iterative application of the image registration and model fitting steps. Unfortunately, the iterative process may lead to suboptimal results due to convergence to local minima. Additionally, the computational demand and processing time associated with such methods renders them impractical for clinical use with large datasets. Recently, Kornaropoulos et al. (2022) presented a novel approach for joint motion correction and quantitative analysis of prostate DWI data using the mono-exponential signal decay model. Specifically, they used a Markov-Random-Field (MRF) technique to simultaneously optimize a mo-

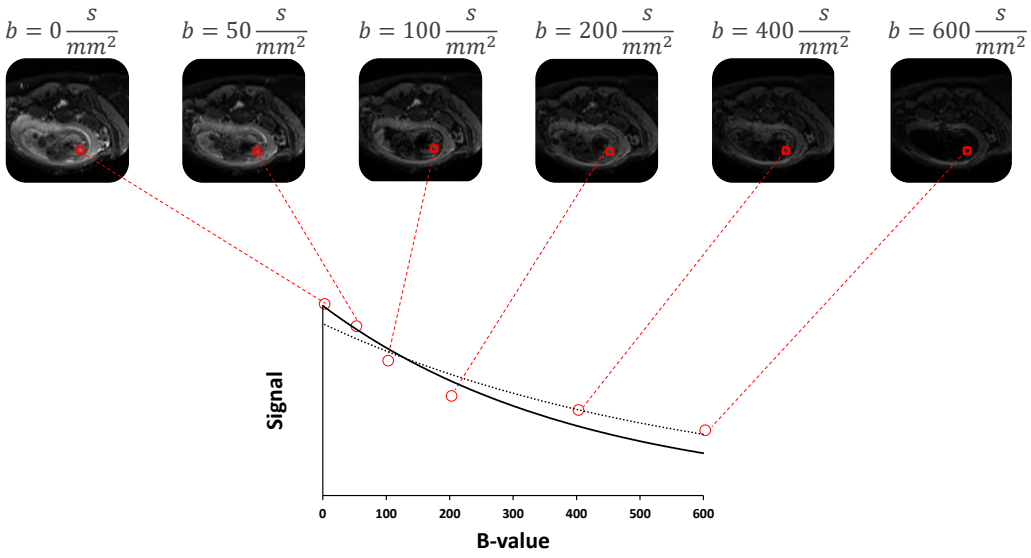

Figure 1: Fetal DWI data acquired with varying b-values. Fetal motion causes the observed signal (red circles) to deviate from the expected signal decay model (solid line). Fitting the model to the observed signal without accounting for motion may lead to an incorrect estimate of the model parameters (dashed line).

tion correction and qDWI model fitting problem. However, this method necessitates the discretization of both domains, and it is computationally intensive. Thus, clinically viable methods for qDWI analysis that are resilient to inter-volume motion artifacts are urgently needed.

Recently, within the field of anatomical fetal MRI reconstruction, several deep-learning models have emerged. Cordero-Grande et al. Cordero-Grande et al. (2022) introduced a deep generative prior model, while Xu et al. Xu et al. (2023) adopted an implicit neural representation approach to achieve motion-robust volumetric reconstruction of anatomical fetal MRI data. In a related context, Davidson et al. Davidson et al. (2022) utilized slice-to-volume deformable image registration to extract reliable 3D measurements of fetal lung volume from fetal MRI data. However, it's worth noting that all these studies primarily focus on slice-to-volume registration within the scope of a single-volume anatomical MRI, without considering potential motion between the different volumes required for quantitative DWI analysis in functional fetal lung maturation assessment.

In this study, we tackle this challenge with the introduction of a self-supervised Deep Neural Network (DNN) framework named "IVIM-Morph." This approach addresses simultaneous motion compensation and bi-exponential IVIM model parameter estimation. Our model comprises two key sub-networks: the first focuses on estimating deformation fields for motion correction, while the second predicts IVIM model parameters based on the motion-corrected data. To ensure the consistency of DWI signal decay with the IVIM model, we introduce an innovative, physics-based loss function. This loss function penalizes signal decays that deviate from the expected IVIM model behavior, thus maintaining physical plausibility. Importantly, our DNN model significantly reduces computation time compared to conventional methods.

We assessed the anatomical registration accuracy of our method by manually delineating one lung in each DW image from 5 cases with severe motion artifacts and 5 cases

with moderate motion artifacts. We then evaluated the alignment of the masks before and after registration using IVIM-Morph in comparison to various registration techniques. Additionally, we have showcased the clinical importance of using IVIM-Morph for reliable IVIM parameter estimation in the presence of motion by illustrating its capability to enhance the correlation between the predicted perfusion fraction parameter (f) in the fetal lung and gestational age (GA) through an analysis of 39 clinical fetal DWI datasets.

Our study delivers the following key contributions:

- We offer a self-supervised deep-learning-based mathematical framework for concurrently estimating motion correction and signal decay model parameters.
- We introduce an innovative registration loss function, guaranteeing physically sound deformation fields that align with the signal decay model.
- We present a comprehensive assessment of our approach, encompassing registration accuracy and its clinical applications in evaluating fetal lung functional maturity.
- We will make our code repository, facilitating motion-compensated IVIM analysis of DWI data, accessible to the public.

2. Background

The bi-exponential IVIM model describes the DWI signal attenuation at a particular voxel relative to the baseline signal as a function of the b-value used during the acquisition Iima and Le Bihan (2016):

$$f_{IVIM}(b_i, f, D^*, D) = S_0 \left(f \cdot e^{-b_i(D+D^*)} + (1-f) \cdot e^{-b_i \cdot D} \right) \quad (1)$$

where S_0 is the baseline signal obtained without applying any diffusion-synthesized gradients; D is the diffusion coefficient; D^* is the pseudo-diffusion coefficient; b_i is the b-value used during the acquisition; and f is for the perfusion fraction Federau (2017); Iima (2021).

The estimation of the IVIM model parameters from the DWI data acquired with multiple b-values $B = \{b_i\}_{i=0}^{N}$, is commonly done by solving the following least-squares problem:

$$\widehat{D}, \widehat{D^*}, \widehat{f} = \underset{D, D^*, f}{\operatorname{argmin}} \sum_{b_i \in B} \| S(b_i) - f_{IVIM}(b_i, f, D^*, D) \|^2 \quad (2)$$

Supplementary regularization terms are frequently incorporated to enhance estimation robustness in the presence of noise and improve clinical diagnostic accuracy Freiman et al. (2013); Orton et al. (2014); Spinner et al. (2021); Vidić et al. (2019).

In the past few years, state-of-the-art, DNN-based methods were introduced for IVIM parameter estimation. Bertleff et al. Bertleff et al. (2017) demonstrated the ability of supervised DNN to predict the IVIM model parameters from low SNR DWI data. Barbieri et al. Barbieri et al. (2020) proposed an unsupervised, physics-informed DNN (IVIM-NET) with results comparable to Bayesian methods with further optimizations by Kaandorp et al. Kaandorp et al. (2021) (IVIM-NET$_{\text{optim}}$). Zhang et al. Zhang et al. (2019) used a multi-layer perceptron with an amortized Gaussian posterior to estimate the IVIM model parameters from fetal lung DWI data. Recently, Vasylechko et al. Vasylechko et al. (2022) used unsupervised convolutional neural networks (CNN) to improve the reliability of IVIM parameter estimates by leveraging spatial correlations in the data.

Nevertheless, all these algorithms presuppose spatial alignment among the different b-value images, rendering them unsuitable for direct application in estimating IVIM model parameters for fetal DWI data, given the inevitable fetal motion during acquisition Afacan et al. (2016).

3. Method

We address the challenge of estimating the IVIM model parameters while compensating for motion artifacts by presenting a self-supervised DNN-based framework for simultaneous motion compensation and IVIM model parameters estimation. Specifically, we aim to find the optimal values for the IVIM model parameters D, D^*, and f, as well as the set of transformations $\Phi = \{\phi_i\}_{i=0}^{N}$ that align the observed DWI signals to the model predictions. The

joint optimization problem is formulated as follows:

$$\widehat{\Phi}, \widehat{D}, \widehat{D^*}, \widehat{f} = \underset{\Phi, D, D^*, f}{\operatorname{argmin}} \sum_{i=0}^{N} \|\phi_i \circ S(b_i) - f_{IVIM}(b_i, f, D^*, D)\|^2 \tag{3}$$

where ϕ_i represents the spatial transformation that aligns the i-th DWI signal to the reference signal (i.e., the model prediction at the i-th b-value).

However, direct optimization of this equation is challenging due to the huge number of unknowns associated with the combination of the bi-exponential IVIM signal decay model and the set of free-form deformations.

Instead, we frame this optimization problem as an estimation of the weights of a DNN that predicts IVIM model parameters and spatial transformations from the observed DWI signals and b-values. Specifically, we will minimize the following objective function:

$$\widehat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{i=0}^{N} \|f_\Theta[0]_i \circ S(b_i) - f_{IVIM}(b_i, f_\Theta[1])\|^2 \tag{4}$$

where Θ are the parameters of the neural network; f_Θ is the forward pass of the DNN function; and $f_\Theta[0]$ and $f_\Theta[1]$ represent the DNN outputs corresponding to the set of spatial transformations Φ and the IVIM parameters D^*, f, and D, respectively.

3.1. Network architecture

Fig. 2 presents the architecture of our IVIM-Morph network. It is comprised of two components: a quantitative IVIM (qIVIM-CNN) prediction network and an image registration network. The qIVIM-CNN network is responsible for predicting the IVIM parameters from the DWI data, while the image registration network is responsible for predicting the set of transformations that align each DWI image with the corresponding image reconstructed from the IVIM parameters with Eq. 1. We describe each component in the following sections.

3.1.1. Quantitative IVIM model fitting sub-network

The qIVIM-CNN is based on a Unet-like architecture Ronneberger et al. (2015) with three parallel decoders, one for each of the IVIM parameters (D, D^*, and f) Vasylechko et al. (2022). To ensure physically plausible IVIM model parameter estimates, we used a Sigmoid activation function at the output of each decoder Kaandorp et al. (2021):

$$P = P_{min} + Sigmoid(X_P) \times (P_{max} - P_{min}) \tag{5}$$

where P denotes any of the IVIM model parameters (D, D^*, f); P_{min} and P_{max} are the prior bounds on the parameter, and X_P is the output parameter map from the corresponding Unet decoder. Table 1 provides a summary of the boundaries used to constrain the estimates of IVIM model parameters. These boundaries were determined through an IVIM analysis of cases from our database without any significant motion observed using a segmented-least-squares approach Gurney-Champion et al. (2018) followed by non-linear trust-region-reflective optimization (SLS-TRF) Branch et al. (1999).

Table 1: Prior bounds on the IVIM parameters

Parameter	$D(\frac{mm^2}{sec})$	f (%)	$D^*(\frac{mm^2}{sec})$
minimum	0.0003	7	0.006
maximum	0.0032	50	0.15

3.1.2. Registration sub-network

We have utilized the Voxel-Morph DNN architecture, renowned for deformable medical image registration Balakrishnan et al. (2019); Dalca et al. (2019, 2018), as the foundation for our image registration sub-network. It predicts the deformation fields ($\Phi = \{\phi_i\}_{i=1}^N$) between the acquired DWI data ($\{S(b_i)\}_{i=1}^N$) to the corresponding model-reconstructed images ($\{R(b_i)\}_{i=1}^N$). The registration is performed between corresponding acquired b-values images and predicted model images such that the moving image $S(b_i)$ is registered to the fixed image $R(b_i)$. Through the registration of original images to those reconstructed by a predicted model, the network can successfully mitigate variations in the contrast between b-value images. Further, this allows for the utilization of physical prior knowledge via the IVIM model, which characterizes expected signal decay behavior.

3.2. Bio-physically-informed loss function

We introduce an innovative loss function comprising a weighted combination of the following three terms:

$$\mathcal{L} = \alpha_1 \mathcal{L}_{fit} + \alpha_2 \mathcal{L}_{smooth} + \alpha_3 \mathcal{L}_{sim} \tag{6}$$

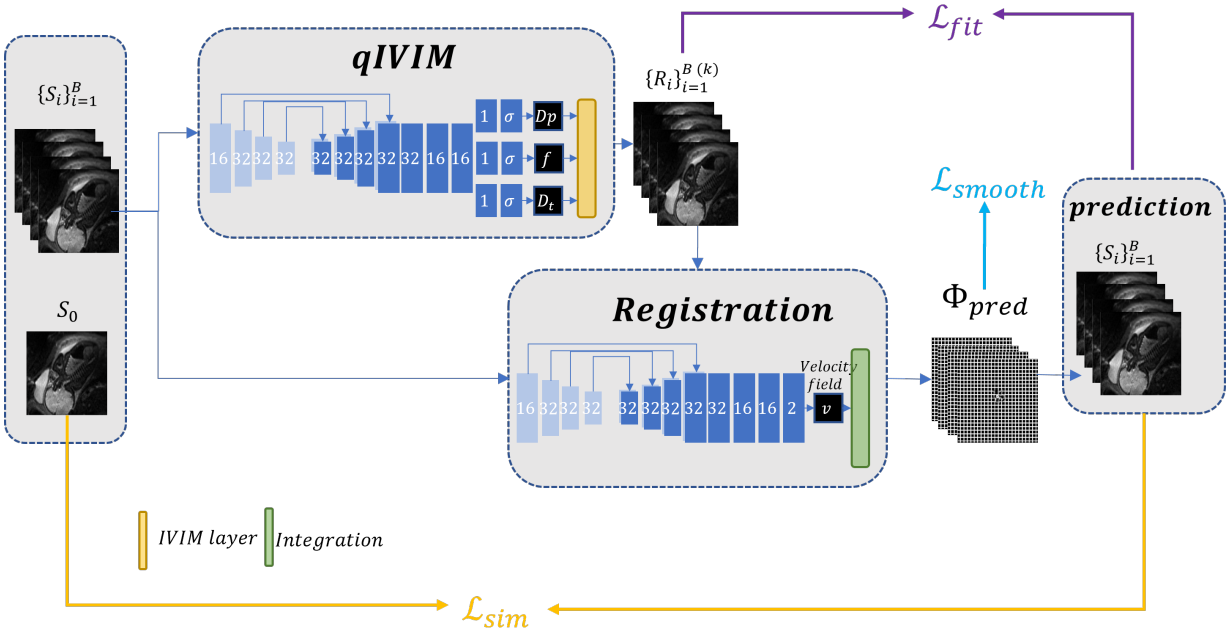

Figure 2: The architecture of the IVIM-Morph network, comprises two sub-networks: a quantitative IVIM (qIVIM), a convolutional neural network (CNN), and an image registration sub-network. The qIVIM-CNN sub-network extracts IVIM parameters from the DWI data, while the image registration sub-network aligns each b-value image with the corresponding image reconstructed by the IVIM parameters.

The model fitting loss ($\mathcal{L}_{fit}$) drives the IVIM-Morph to generate deformation fields that minimize the disparity between acquired images and model-generated images (Eq. 1). This guarantees a physically plausible representation of signal decay across the b-value axis, leading to enhanced precision in IVIM parameter maps. Specifically, $\mathcal{L}_{fit}$ is a weighted version of the standard error of the regression (WSER) between the models' prediction of the Diffusion-Weighted images ($\{R(b_i)\}_{i=1}^{N}$) and the corresponding deformed images ($\{\phi_i \circ S(b_i)\}_{i=1}^{N}$) that accounts for potential bias in high b-value images that have a low signal as follows:

$$\mathcal{L}_{fit} = \sqrt{\frac{\sum_{i=1}^{N} w_i \cdot (\phi_i \circ S(b_i) - R(b_i))^2}{\sum_{i=1}^{N} w_i} \cdot \frac{1}{N-p-1}} \quad (7)$$

The parameter N is the number of observations, which is equal to the number of b-values used during the scan. The parameter p denotes the number of unknowns in the model, which is three in our case (D, D^*, and f), and the weight (w_i) is defined as $log(b_i + 1) + 1$. We normalized the WSER by the average intensity value of the motion-compensated set.

The smoothness term ($\mathcal{L}_{smooth}$) encourages the creation of deformations that are both realistic and invertible. This loss penalizes for a large L_1 norm of the gradients of the velocity field u Balakrishnan et al. (2019):

$$\mathcal{L}_{smooth}(\Phi) = \frac{1}{N}\frac{1}{|\Omega|}\sum_{i=1}^{N} \|\nabla u_i\|^2 \quad (8)$$

where Ω is the image spatial domain.

Lastly, the registration loss ($\mathcal{L}_{sim}$) promotes the alignment of deformation fields, aligning the Diffusion-Weighted images ($S(b_i), i \in \{1, ..., N\}$) with the baseline image S_0, by calculating their local normalized cross-correlation (NCC) Balakrishnan et al. (2019)

$$\mathcal{L}_{sim}(S(b_0), \Phi \circ S) = \frac{1}{N}\frac{1}{|\Omega|}\sum_{i=1}^{N} NCC(S(b_0), \phi_i \circ S(b_i)) \quad (9)$$

3.3. Implementation details

We implemented our models on Visual Studio Code 1.79.2, Python 3.8.12 with PyTorch 1.13.0, and CUDA 11.8. We applied our suggested methods with a batch size of one, meaning that each batch consists of data from one patient with size: $n_b \times n_x \times n_y$, where n_b is the number of b-values used for scanning the patient, and $n_x \times n_y$ is the image shape. We used an Adam optimizer with an initial learning rate of 10^{-3} with a "reduce on plateau" learning rate decay scheduler. All the calculations in this study were carried out on a Linux machine equipped with a Tesla V100-PCIE-32GB GPU. The CPU in use was an Intel(R) Xeon(R) Gold 6230 CPU, operating at 2.10GHz.

4. Experiments

4.1. Data

We used a legacy fetal DWI dataset in this study Afacan et al. (2016). DWI data were acquired on a Siemens 3T Skyra scanner using an 18-channel body matrix coil. The imaging technique used was a multi-slice, single-shot echo-planar imaging (EPI) sequence for obtaining diffusion-weighted scans of the lungs. Each scan had an in-plane resolution of 2.5mm × 2.5mm and a slice thickness of 3mm. The echo time was set at 60ms and the repetition time ranged from 2s to 4.4s, depending on the number of slices required to cover the lungs. Each patient underwent scanning with 6 different b-values (0, 50, 100, 200, 400, 600 sec/mm^2) in both axial and coronal planes with 6 gradient directions. Trace-weighted images were exported from the scanner A ROI was manually drawn for each case in the right lung at b_0 Afacan et al. (2016).

The data set consists of 39 cases with different levels of misalignment between the different b-value image volumes. For each subject, we chose only one slice where the ROI in the right lung was labeled. The images were then cropped to a shape of 96×96 and normalized by the 0.99 quantiles of the DWI image acquired without diffusion gradients (b-value=0 sec/mm^2).

To ensure the reproducibility of our findings, we established two distinct, non-overlapping groups of 16 cases each for hyperparameter tuning. The composition of these groups was planned to encapsulate a wide array of gestational ages, thereby encompassing nearly the full breadth of ages present in our dataset. We conducted hyperparameter tuning for each group independently, as detailed in Section 4.3. The remaining 23 cases, that were left out in each group are designated as the test cases for each group. Our primary findings and analysis will be conducted on these specific cases.

In addition, we chose a sample of 10 cases for analysis. This sample included 5 cases exhibiting severe motion artifacts and another 5 with only minor motion artifacts. For each of these cases, we conducted a manual segmentation of one lung in the different b-value images.

4.2. Baseline methods

We compared our method to six baseline methods as follows:

1. Quantifying IVIM parameters using the non-linear SLS-TRF approach Gurney-Champion et al. (2018); Branch et al. (1999), without utilizing any motion compensation.
2. Registering all b-value images to the b=0 sec/mm^2 image using SyN registration Avants et al. (2008), followed by quantifying IVIM parameters using the SLS-TRF algorithm.
3. Registering all b-value images to the b=0 sec/mm^2 image using affine registration Avants et al. (2008), followed by quantifying IVIM parameters using the SLS-TRF algorithm.
4. Registering each b-value image to the previous image using SyN registration. For example, we register the b=50 sec/mm^2 to b=0 sec/mm^2 image and then register b=100 sec/mm^2 to the result.
5. Iteratively quantifying IVIM parameters and registering each b-value image to the corresponding model image Kurugol et al. (2017).
6. Unsupervised VoxelMorph-based registration Balakrishnan et al. (2019) of all b-value images to the b=0 sec/mm^2 image, followed by quantifying IVIM parameters using the SLS-TRF algorithm.

4.3. Hyper-parameters tuning

For hyperparameter optimization, we implemented a grid search strategy, with a primary focus on determining the appropriate weights for the loss terms denoted as α_1, α_2, and α_3. We selected the values for these hyperparameters as follows: α_1 was varied within the range [0.5,

1, 5, 10], α_2 within [0.015, 0.03], and α_3 within [0.1, 0.8, 5].

The tuning process was conducted separately for two distinct groups, each comprising 16 cases, and was performed independently. The criterion used for selecting the optimal hyperparameters was based on the correlation between the IVIM parameter f and gestational age during the canalicular stage of fetal development (GA < 26 weeks).

4.4. Lung Segmentation Evaluation

We evaluated the anatomical registration accuracy of our IVIM-Morph in comparison to the different registration approaches outlined in Section 4.2 for cases with different levels of motion. These techniques were utilized to assess the alignment of images S_i (for $i > 0$) with the reference image S_0. Our experiment involved 10 selected cases, each of which included manual segmentation of one lung. The effectiveness of these alignment methods was quantitatively assessed using the Dice score metric. This evaluation was conducted both prior to and following the application of registration.

4.5. NCC loss contribution to the registration

We carried out an in-depth ablation study to thoroughly understand the impact and significance of the NCC loss on the registration process. To achieve this, we maintained constant values for certain parameters, setting $\alpha_1 = 1$ and $\alpha_2 = 0.015$. This controlled setup allowed us to isolate and examine the influence of the NCC loss more effectively. We conducted this experiment by repeatedly executing the experiment outlined in Section 4.4, but with a key variation each time: we altered the value of α_3 for each iteration. By systematically changing α_3 while keeping the other parameters fixed, we were able to observe how variations in the NCC loss component affected the overall registration performance. The series of experiments under varying α_3 conditions were instrumental in gauging the sensitivity and responsiveness of our registration process to changes in the NCC loss.

4.6. Clinical impact: Functional fetal lung maturity assessment

We assessed the performance of our proposed method by examining its correlation with the GA and the perfusion fraction parameter (f) in the IVIM model. This parameter indirectly represents the proportion of the capillary network within the tissue. As established in prior research, the f parameter exhibits a substantial increase with advancing gestational age in the fetal lung Ercolani et al. (2021); Korngut et al. (2022). We conducted the analysis on the two group test cases (23 cases each). For each case, we used IVIM-Morph to compute the IVIM parameter maps. Subsequently, we calculated the average parameter value in the lung for each case and evaluated the correlation between each parameter and GA separately for the canalicular and saccular phases, as suggested by Korngut et al. (2022).

5. Results

5.1. Hyper-parameters tuning

The optimal hyperparameters for group 1 are: $\alpha_1 = 10, \alpha_2 = 0.015, \alpha_3 = 0.1$, and for group 2 are: $\alpha_1 = 0.5, \alpha_2 = 0.015, \alpha_3 = 0.8$.

5.2. Lung Segmentation Evaluation

Lung segmentation evaluation results are presented in Fig. 4. The mean dice coefficient for each compared method is plotted as a boxplot, separately for the major and minor motion cases. We calculated the dice twice, one time using the optimal hyperparameters of group 1 and one time using the optimal hyperparameters of group 2. We also plotted the mean dice coefficient before applying registration. The mean dice before registration in the minor motion cases is 0.878 ± 0.036 and in the major motion cases is 0.771 ± 0.040, which is expected based on the cases' motion level. For cases involving major motion, IVIM-Morph succeeded in enhancing the dice coefficient achieving superior results for group 2 (dice = 0.854 ± 0.038) than group 1 (dice = 0.812 ± 0.046). Conversely, in scenarios with minor motion, IVIM-Morph, employing both sets of hyperparameters, consistently maintained a high dice coefficient.

5.3. NCC loss contribution to the registration

Figure 3 displays the outcomes of the experiment investigating the impact of NCC loss. In scenarios with minor motion, the choice of α_3 value seems to have a negligible effect on the dice coefficient achieved post-application

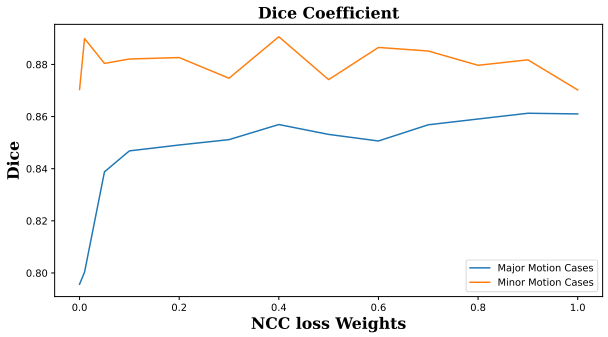

Figure 3: Average dice coefficient computed for lung segmentation between S_0 and the deformed lung segmentation in S_i (for $i > 0$), utilizing IVIM-Morph in 10 specifically chosen cases with varying α_3 values. The blue line indicates the cases characterized by major motion, whereas the orange line corresponds to the cases with minor motion.

Table 2: Methods running times

Method	Time (s)	Machine
SLS-TRF	57.93 ± 1.26	CPU
Affine - Reg to b0	53.32 ± 3.09	CPU
SyN - Reg to b0	54.04 ± 2.14	CPU
RSyN - Reg to next b	53.11 ± 2.29	CPU
Iterative SyN-TRF	261.16 ± 67.61	CPU
VoxelMorph + SLS - TRF	84.738 ± 7.23	CPU+GPU
IVIM-Morph	52.69 ± 1.90	GPU

of IVIM-Morph. Conversely, in cases of major motion, a lower weighting on NCC loss is observed to yield suboptimal dice scores. This finding highlights the increased importance of NCC loss in the optimization process, particularly in instances where major motion artifacts are present.

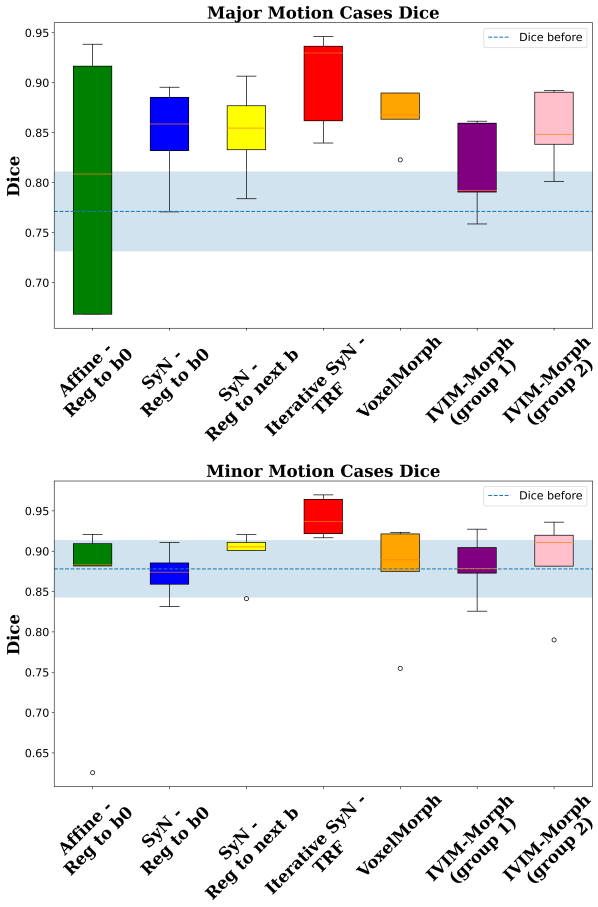

Figure 4: Results from the Lung Segmentation Evaluation Experiment. The bar plot at the top illustrates the dice coefficients for cases with significant motion, while the bottom bar plot shows the dice coefficients for cases with minimal motion. In both plots, the dashed line and the shaded area indicate the average and standard deviation of the dice scores prior to the application of registration.

5.4. *Clinical impact: Functional fetal lung maturity assessment*

Fig. 5 shows representative IVIM parameter maps generated by each method for a case with motion and for a case without motion. The IVIM-Morph method produced smoother parameter maps compared to the other methods. This can be attributed to the use of a CNN in the qDWI sub-network which leverages spatial correlations to estimate IVIM parameters more accurately and results in smoother parameter maps. In contrast, the other methods rely on traditional optimization and registration techniques, which may result in less accurate and noisier parameter maps.

Fig 6 presents the correlation analysis between the IVIM parameters as computed by each method and the GA. Our IVIM-Morph approach outperformed the other methods, achieving the highest correlation coefficient (0.44 for group 1 and 0.52 for group 2) for the f parameter in the canalicular phase. Figure 6 displays correlations derived from two test groups, where each group's

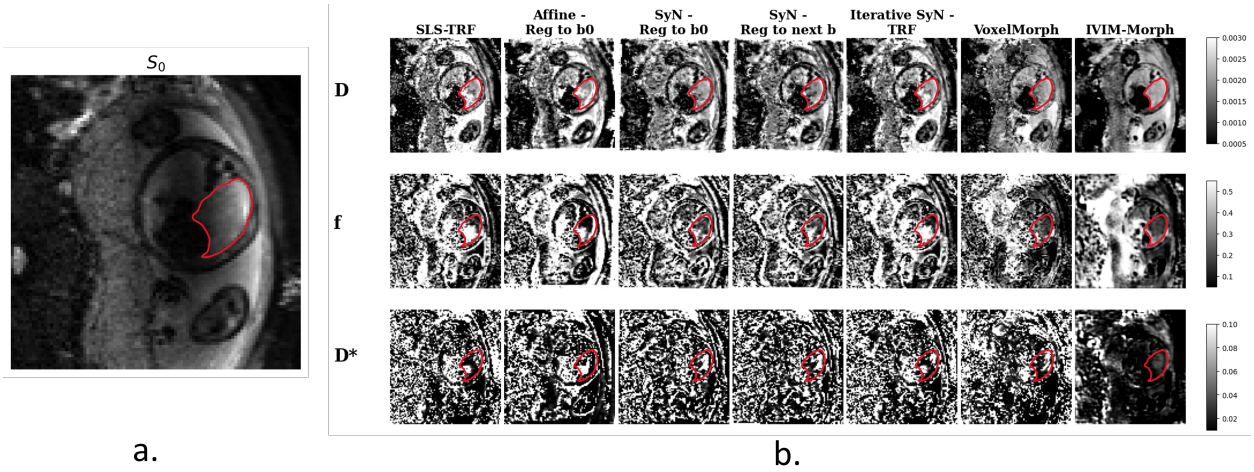

Figure 5: a.) Example of baseline S_0 image .b) IVIM parameter maps(D_p, D_t and F) of the same subject estimated using different methods, including SLS-TRF, registration to $S(b_0)$ using Affine registration, registration to $S(b_0)$ using SyN registration, registration to the previous b-value image using SyN registration, Iterative SyN-TRF, VoxelMorph and IVIM-Morph. The contours of one lung are displayed

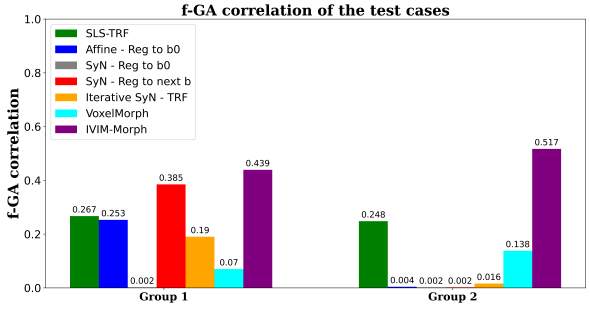

Figure 6: The correlations between f and the GA in the canalicular stage are calculated in two datasets: group 1 test cases and group 2 test cases.

test cases include those from the opposing original group (the 16 cases for the hyperparameter tune). Average IVIM parameters were computed in the ROI for each case across all evaluated methods, utilizing the best hyperparameters for each group as detailed in 5.1 for the IVIM-Morph calculations. Notably, IVIM-Morph demonstrates greater consistency in terms of correlation between the two groups, unlike other methods (except TRF-SLS and Syn - Reg to b_0), which showed varying correlations across the groups.

Supplementary material includes supplementary results and tables that summarize the correlations between different IVIM model parameters and GA.

6. Discussion and Conclusions

Accurately assessing IVIM parameters while addressing fetal movement is crucial for obtaining precise quantitative imaging biomarkers related to fetal lung development. In this study, we introduced IVIM-Morph, a self-supervised deep neural network approach designed for simultaneous motion compensation and quantitative DWI analysis using the IVIM model. Our method surpassed baseline approaches that consider motion and estimate IVIM model parameters, notably enhancing the correlation between the perfusion fraction parameter of the IVIM model (f) and gestational age (GA). While, our segmentation experiment results, as shown in Figure 4, indicate improved alignment through iterative model estimation and registration (SyN-TRF), as evidenced by the Dice coefficient. It is important to note that this alignment alone does not guarantee a physically plausible signal decay behavior across the entire b-value axis. Consequently, it may lead to inaccurate estimates of the IVIM parameters, as indicated by the poor correlation achieved between the pseudo-diffusion fraction parameter f and

gestational age (GA) using the SyN-TRF approach. In contrast, our IVIM-Morph approach strikes a better balance between precise boundary registration and maintaining a realistic signal decay behavior along the entire b-value axis. This equilibrium results in an improved correlation between the pseudo-diffusion fraction parameter f and GA and achieves comparable segmentation accuracy.

It's crucial to emphasize that, in contrast to previously proposed methods primarily addressing motion between slices within a single volume in anatomical fetal imaging, our IVIM-Morph focuses on addressing motion between volumes in quantitative DWI acquisitions that encompass multiple volumes. Additionally, in its current configuration, IVIM-Morph operates under the assumption of a single trace-weighted image per b-value, which is automatically generated by aggregating the various b-vector images into a single scalar map. Future enhancements may involve accommodating motion between the distinct b-vector images employed in the computation of the trace-weighted b-value image.

While our primary emphasis has been on the quantitative analysis of fetal lung DWI data, it's worth noting that the proposed approach holds potential applicability in various other quantitative DWI analysis domains that grapple with motion-related challenges. For instance, applications such as the detection and staging of liver fibrosis through IVIM analysis of abdominal DWI data Ye et al. (2020), the assessment of non-alcoholic fatty liver disease Guiu et al. (2012), and the identification of diffuse renal pathologies Caroli et al. (2018) can all derive benefits from our approach by effectively accounting for motion induced by processes like respiration between different b-value volumes.

In conclusion, our study has showcased the clinical promise of evaluating functional fetal lung maturation non-invasively from fetal DWI data. our IVIM-Morph stands out as a means to markedly enhance the precision of non-invasive, quantitative fetal lung maturity assessment. Moreover, our proposed method can be readily extended to other clinical scenarios necessitating motion correction in the computation of quantitative MRI biomarkers. Collectively, our findings point to IVIM-Morph as a valuable asset in advancing fetal MRI, enhancing fetal health monitoring, and informing clinical decision-making.

Acknowledgments

This research was supported in part by the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel under award number 2019056, and by the National Institutes of Health (NIH) under award numbers R01 LM013608, R01 EB019483, R01 NS124212, and R01 NS121657. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health.

Declaration of Generative AI

During the preparation of this work, the author(s) used ChatGPT in order to improve readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Afacan, O., Gholipour, A., Mulkern, R.V., Barnewolt, C.E., Estroff, J.A., Connolly, S.A., Parad, R.B., Bairdain, S., Warfield, S.K., 2016. Fetal lung apparent diffusion coefficient measurement using diffusion-weighted MRI at 3 Tesla: Correlation with gestational age. Journal of Magnetic Resonance Imaging 44, 1650–1655. doi:10.1002/jmri.25294.

Ahmed, B., Konje, J.C., 2021. Fetal lung maturity assessment: A historic perspective and non–invasive assessment using an automatic quantitative ultrasound analysis (a potentially useful clinical tool). European Journal of Obstetrics & Gynecology and Reproductive Biology 258, 343–347.

Avants, B., Epstein, C., Grossman, M., Gee, J., 2008. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Medical Image Analysis 12, 26–41. URL: https://www.sciencedirect.com/science/article/pii/S1361841507000606, doi:https://doi.org/10.1016/j.media.2007.06.004. special Issue on The Third International Workshop on Biomedical Image Registration – WBIR 2006.

Avena-Zampieri, C.L., Hutter, J., Rutherford, M., Milan, A., Hall, M., Egloff, A., Lloyd, D.F., Nanda, S., Greenough, A., Story, L., 2022. Assessment of the fetal lungs in utero. American journal of obstetrics & gynecology MFM , 100693.

Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V., 2019. Voxelmorph: a learning framework for deformable medical image registration. IEEE Transactions on medical imaging 38, 1788–1800.

Barbieri, S., Gurney-Champion, O.J., Klaassen, R., Thoeny, H.C., 2020. Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted mri. Magnetic resonance in medicine 83, 312–321.

Beck, A.P.A., Araujo Junior, E., Leslie, A.T.F.S., Camano, L., Moron, A.F., 2015. Assessment of fetal lung maturity by ultrasound: objective study using gray-scale histogram. The Journal of Maternal-Fetal & Neonatal Medicine 28, 617–622.

Bertleff, M., Domsch, S., Weingärtner, S., Zapp, J., O'Brien, K., Barth, M., Schad, L.R., 2017. Diffusion parameter mapping with the combined intravoxel incoherent motion and kurtosis model using artificial neural networks at 3 t. NMR in Biomedicine 30, e3833.

Branch, M.A., Coleman, T.F., Li, Y., 1999. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. SIAM Journal on Scientific Computing 21, 1–23. doi:10.1137/S1064827595289108.

Caroli, A., Schneider, M., Friedli, I., Ljimani, A., De Seigneux, S., Boor, P., Gullapudi, L., Kazmi, I., Mendichovszky, I.A., Notohamiprodjo, M., et al., 2018. Diffusion-weighted magnetic resonance imaging to assess diffuse renal pathology: a systematic review and statement paper. Nephrology Dialysis Transplantation 33, ii29–ii40.

Cordero-Grande, L., Ortuño-Fisac, J.E., Del Hoyo, A.A., Uus, A., Deprez, M., Santos, A., Hajnal, J.V., Ledesma-Carbayo, M.J., 2022. Fetal mri by robust deep generative prior reconstruction and diffeomorphic registration. IEEE Transactions on Medical Imaging 42, 810–822.

Dalca, A.V., Balakrishnan, G., Guttag, J., Sabuncu, M., 2018. Unsupervised learning for fast probabilistic diffeomorphic registration. MICCAI: Medical Image Computing and Computer Assisted Intervention, LNCS 11070, 729–738.

Dalca, A.V., Balakrishnan, G., Guttag, J., Sabuncu, M., 2019. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. Medical Image Analysis 57, 226–236.

Davidson, J., Uus, A., Egloff, A., Van Poppel, M., Matthew, J., Steinweg, J., Deprez, M., Aertsen, M., Deprest, J., Rutherford, M., 2022. Motion corrected fetal body magnetic resonance imaging provides reliable 3d lung volumes in normal and abnormal fetuses. Prenatal Diagnosis 42, 628–635.

Ercolani, G., Capuani, S., Antonelli, A., Camilli, A., Ciulla, S., Petrillo, R., Satta, S., Grimm, R., Giancotti, A., Ricci, P., Catalano, C., Manganaro, L., 2021. IntraVoxel Incoherent Motion (IVIM) MRI of fetal lung and kidney: Can the perfusion fraction be a marker of normal pulmonary and renal maturation? European Journal of Radiology 139, 109726. doi:10.1016/J.EJRAD.2021.109726.

Federau, C., 2017. Intravoxel incoherent motion mri as a means to measure in vivo perfusion: A review of the evidence. NMR in Biomedicine 30, e3780.

Freiman, M., Perez-Rossello, J.M., Callahan, M.J., Voss, S.D., Ecklund, K., Mulkern, R.V., Warfield, S.K., 2013. Reliable estimation of incoherent motion parametric maps from diffusion-weighted mri using fusion bootstrap moves. Medical image analysis 17, 325–336.

Guiu, B., Petit, J.M., Capitan, V., Aho, S., Masson, D., Lefevre, P.H., Favelier, S., Loffroy, R., Vergès, B., Hillon, P., et al., 2012. Intravoxel incoherent motion diffusion-weighted imaging in nonalcoholic fatty liver disease: a 3.0-t mr study. Radiology 265, 96–103.

Gurney-Champion, O.J., Klaassen, R., Froeling, M., Barbieri, S., Stoker, J., Engelbrecht, M.R., Wilmink, J.W., Besselink, M.G., Bel, A., Van Laarhoven, H.W., et al., 2018. Comparison of six fit algorithms for the intravoxel incoherent motion model of diffusion-weighted

magnetic resonance imaging data of pancreatic cancer patients. PloS one 13, e0194590.

Guyader, J.M., Bernardin, L., Douglas, N.H., Poot, D.H., Niessen, W.J., Klein, S., 2015. Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion mr images of the abdomen. Journal of Magnetic Resonance Imaging 42, 315–330.

Iima, M., 2021. Perfusion-driven intravoxel incoherent motion (ivim) mri in oncology: applications, challenges, and future trends. Magnetic Resonance in Medical Sciences 20, 125.

Iima, M., Le Bihan, D., 2016. Clinical intravoxel incoherent motion and diffusion mr imaging: past, present, and future. Radiology 278, 13–32.

Jakab, A., Tuura, R., Kottke, R., Kellenberger, C.J., Scheer, I., 2017. Intra-voxel incoherent motion mri of the living human foetus: technique and test–retest repeatability. European radiology experimental 1, 1–11.

Kaandorp, M.P., Barbieri, S., Klaassen, R., van Laarhoven, H.W., Crezee, H., While, P.T., Nederveen, A.J., Gurney-Champion, O.J., 2021. Improved unsupervised physics-informed deep learning for intravoxel incoherent motion modeling and evaluation in pancreatic cancer patients. Magnetic resonance in medicine 86, 2250–2265.

Koh, D.M., Collins, D.J., 2007. Diffusion-weighted mri in the body: applications and challenges in oncology. American Journal of Roentgenology 188, 1622–1635.

Kornaropoulos, E.N., Zacharaki, E.I., Zerbib, P., Lin, C., Rahmouni, A., Paragios, N., 2022. Joint deformable image registration and adc map regularization: Application to dwi-based lymphoma classification. IEEE Journal of Biomedical and Health Informatics 26, 3151–3162.

Korngut, N., Rotman, E., Afacan, O., Kurugol, S., Zaffrani-Reznikov, Y., Nemirovsky-Rotman, S., Warfield, S., Freiman, M., 2022. Super-ivim-dc: Intra-voxel incoherent motion based fetal lung maturity assessment from limited dwi data using supervised learning coupled with data-consistency, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 743–752.

Kumar, P., Burton, B.K., 2007. Congenital malformations: evidence-based evaluation and management. McGraw-Hill.

Kurugol, S., Freiman, M., Afacan, O., Domachevsky, L., Perez-Rossello, J.M., Callahan, M.J., Warfield, S.K., 2017. Motion-robust parameter estimation in abdominal diffusion-weighted MRI by simultaneous image registration and model estimation. Medical Image Analysis 39, 124–132. doi:10.1016/J.MEDIA.2017.04.006.

Lakshminrusimha, S., Keszler, M., 2015. Persistent pulmonary hypertension of the newborn. Neoreviews 16, e680–e692.

Moeglin, D., Talmant, C., Duyme, M., Lopez, A.C., 2005. Fetal lung volumetry using two- and three-dimensional ultrasound. Ultrasound in Obstetrics & Gynecology 25, 119–127. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/uog.1799, doi:10.1002/UOG.1799.

Orton, M.R., Collins, D.J., Koh, D.M., Leach, M.O., 2014. Improved intravoxel incoherent motion analysis of diffusion weighted imaging by data driven bayesian modeling. Magnetic resonance in medicine 71, 411–420.

Rome, R., Glover, J.I., Simmons, S., 1975. The benefits and risks of amniocentesis for the assessment of fetal lung maturity. BJOG: An International Journal of Obstetrics & Gynaecology 82, 662–668.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer. pp. 234–241.

Sanz-Estébanez, S., Pieciak, T., Alberola-López, C., Aja-Fernández, S., 2018. Robust estimation of the apparent diffusion coefficient invariant to acquisition noise and

physiological motion. Magnetic Resonance Imaging 53, 123–133.

Spinner, G.R., Federau, C., Kozerke, S., 2021. Bayesian inference using hierarchical and spatial priors for intravoxel incoherent motion mr imaging in the brain: Analysis of cancer and acute stroke. Medical image analysis 73, 102144.

Vasylechko, S.D., Warfield, S.K., Afacan, O., Kurugol, S., 2022. Self-supervised ivim dwi parameter estimation with a physics based forward model. Magnetic Resonance in Medicine 87, 904–914.

Vidić, I., Jerome, N.P., Bathen, T.F., Goa, P.E., While, P.T., 2019. Accuracy of breast cancer lesion classification using intravoxel incoherent motion diffusion-weighted imaging is improved by the inclusion of global or local prior knowledge with bayesian methods. Journal of Magnetic Resonance Imaging 50, 1478–1488.

Ward, V.L., Nishino, M., Hatabu, H., Estroff, J.A., Barnewolt, C.E., Feldman, H.A., Levine, D., 2006. Fetal lung volume measurements: determination with mr imaging—effect of various factors. Radiology 240, 187–193.

Xu, J., Moyer, D., Gagoski, B., Iglesias, J.E., Grant, P.E., Golland, P., Adalsteinsson, E., 2023. Nesvor: Implicit neural representation for slice-to-volume reconstruction in mri. IEEE Transactions on Medical Imaging .

Ye, Z., Wei, Y., Chen, J., Yao, S., Song, B., 2020. Value of intravoxel incoherent motion in detecting and staging liver fibrosis: A meta-analysis. World Journal of Gastroenterology 26, 3304.

Zhang, L., Vishnevskiy, V., Jakab, A., Goksel, O., 2019. Implicit modeling with uncertainty estimation for intravoxel incoherent motion imaging, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 1003–1007.

Supplementary Materials of the paper:

IVIM-Morph: Motion-compensated quantitative Intra-voxel Incoherent Motion (IVIM) analysis for functional fetal lung maturity assessment from diffusion-weighted MRI data

Noga Kertes, Yael Zaffrani-Reznikov, Onur Afacan, Sila Kurugol,
Simon K. Warfield, Moti Freiman

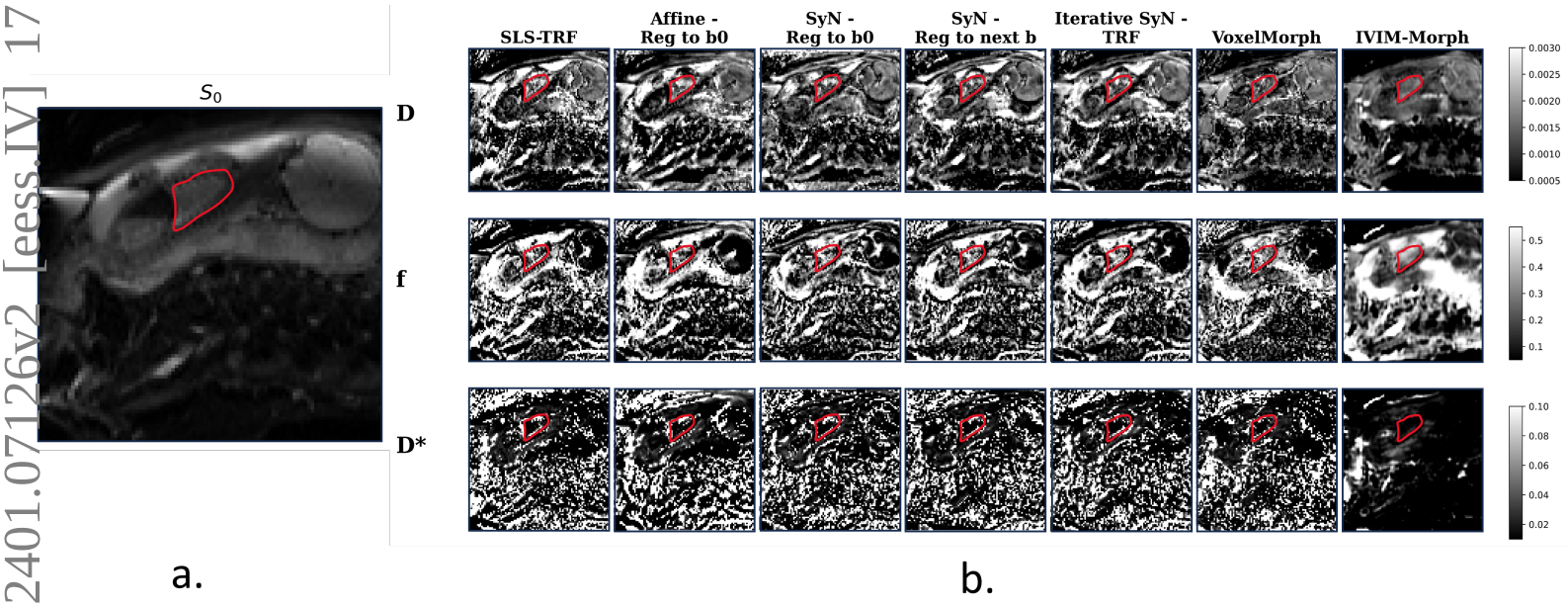

Figure 1: a.) Example of baseline S_0 image .b) IVIM parameter maps(D_p, D_t and F) of the same subject estimated using different methods, including SLS-TRF, registration to $S(b_0)$ using Affine registration, registration to $S(b_0)$ using SyN registration, registration to the previous b-value image using SyN registration, Iterative SyN-TRF, VoxelMorph and IVIM-Morph. The contours of one lung are displayed

Table 1: Group 1 tests cases IVIM parameters correlation with GA in the canicular and saccular stages

	Dt_can	Dt_sac	Dp_can	Dp_sac	Fp_can	Fp_sac
SLS-TRF	0.111	0.118	0.279	0.008	0.267	0.107
Affine-TRF reg to b0	0.022	0.089	0.052	0.004	0.253	0.202
SyN-TRF reg to b0	0.145	0.000	0.021	0.002	0.002	0.000
SyN-TRF reg to next b	0.297	0.018	0.023	0.001	0.385	0.009
Iterative SyN-TRF	0.203	0.124	0.004	0.001	0.190	0.048
IVIM-Morph	0.075	0.417	0.286	0.040	**0.439**	0.125
VoxelMorph	0.074	0.101	0.254	0.026	0.070	0.121

Table 2: Group 2 tests cases IVIM parameters correlation with GA in the canicular and saccular stages

	Dt_can	Dt_sac	Dp_can	Dp_sac	Fp_can	Fp_sac
SLS-TRF	0.007	0.000	0.299	0.058	0.249	0.027
Affine-TRF reg to b0	0.007	0.042	0.360	0.216	0.004	0.007
SyN-TRF reg to b0	0.095	0.007	0.478	0.129	0.003	0.017
SyN-TRF reg to next b	0.023	0.000	0.445	0.056	0.002	0.021
Iterative SyN-TRF	0.000	0.002	0.420	0.063	0.016	0.001
IVIM-Morph	0.400	0.008	0.425	0.013	**0.517**	0.042
VoxelMorph	0.027	0.132	0.069	0.068	0.139	0.121

Group 1 Test Cases

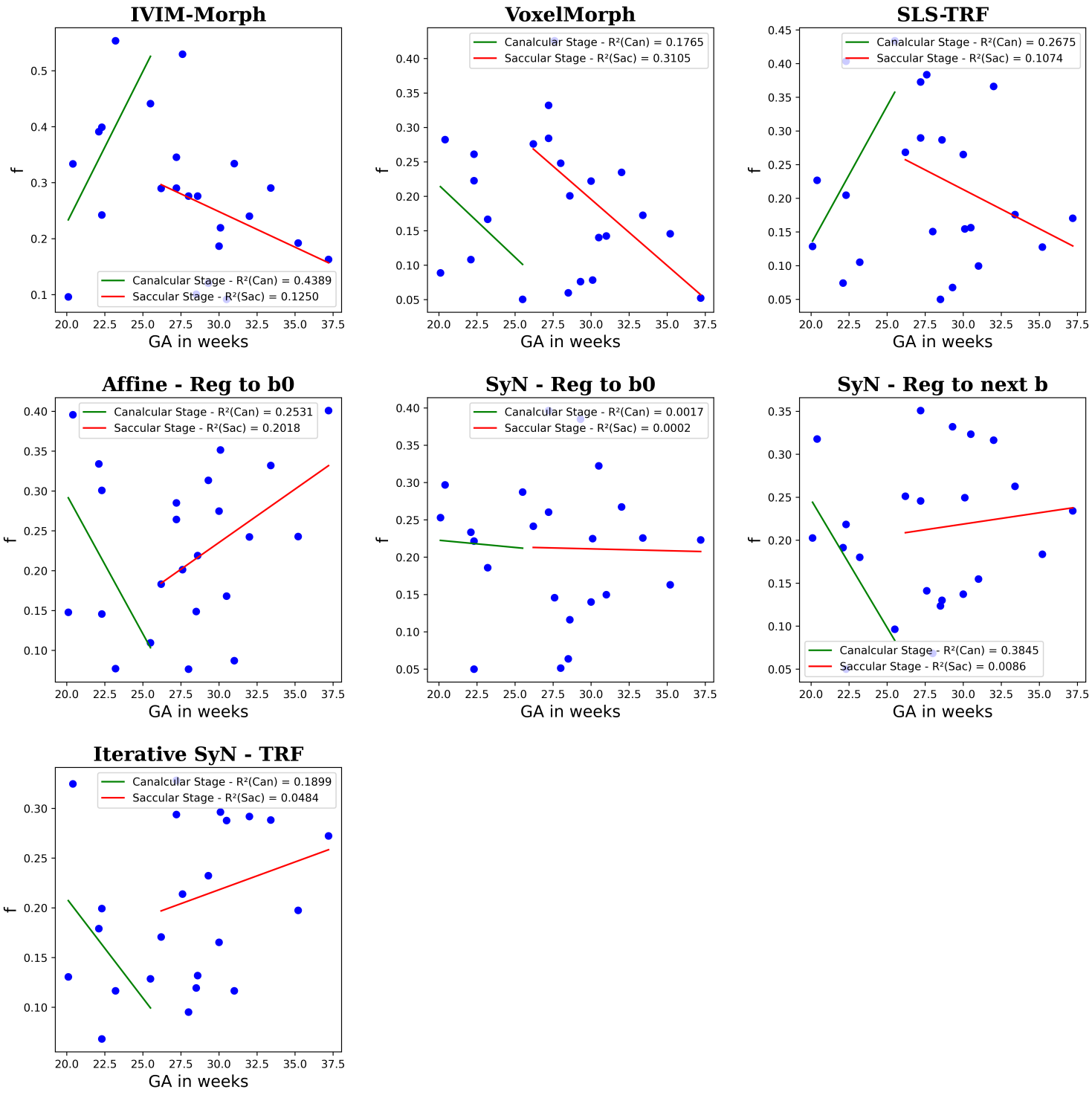

Figure 2: Group 1: Scatter plots depict the average f parameter values in the fetal lung across all methods evaluated. Separate linear fitting with GA was performed for both the canalicular and saccular stages.

Group 2 Test Cases

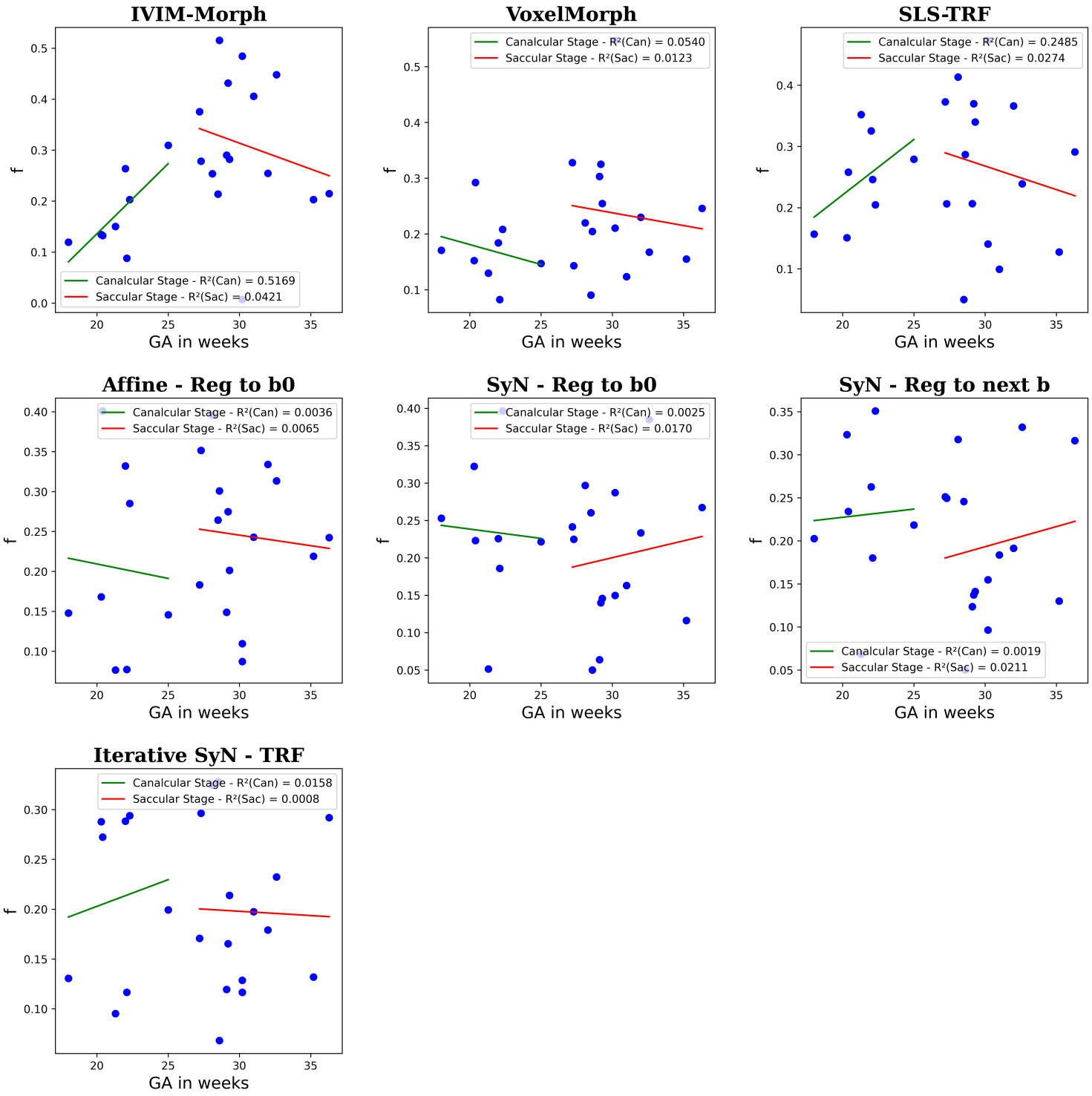

Figure 3: Group 2: scatter plots depict the average f parameter values in the fetal lung across all methods evaluated. Separate linear fitting with GA was performed for both the canalicular and saccular stages.