

# Advancing ASL Kidney Image Registration: A Tailored Pipeline with Groupwise VoxelMorph

**Anne Oyarzun-Domeño**<sup>1,2</sup>

ANNE.OYARZUN@UNAVARRA.ES

<sup>1</sup> *Electrical Electronics and Communications Engineering, Public University of Navarre, 31006, Pamplona, Spain.*

<sup>2</sup> *IdiSNA, Health Research Institute of Navarre, 31008, Spain.*

**Izaskun Cia**<sup>1</sup>

IZASKUN.CIA@UNAVARRA.ES

**Rebeca Echeverria-Chasco**<sup>2,3</sup>

RECHEVERRIAC@UNAV.ES

<sup>3</sup> *Department of Radiology, Clínica Universidad de Navarra, 31008, Pamplona, Spain.*

**María A. Fernández-Seara**<sup>2,3</sup>

MFSEARA@UNAV.ES

**Paloma L. Martín-Moreno**<sup>2,4</sup>

PLMARTIN@UNAV.ES

<sup>4</sup> *Department of Nephrology, Clínica Universidad de Navarra, 31008, Pamplona, Spain.*

**Nuria Garcia-Fernandez**<sup>2,4</sup>

NRGARCIA@UNAV.ES

**Gorka Bastarrika**<sup>2,3</sup>

BASTARRIKA@UNAV.ES

**Javier Navallas**<sup>1,2</sup>

JAVIER.NAVALLAS@UNAVARRA.ES

**Arantxa Villanueva**<sup>1,2,5</sup>

AVILLA@UNAVARRA.ES

<sup>5</sup> *Institute of Smart Cities (ISC), Health Research Institute of Navarre, 31006, Pamplona, Spain.*

## Abstract

Arterial spin labeling (ASL) provides a non-invasive assessment of renal blood flow, but it faces difficulties due to motion artifacts and the effects of blood inflow. This work introduces GVox, a deep learning-based motion correction (MoCo) framework tailored for ASL imaging. GVox extends VoxelMorph, incorporating cortical signal enhancement as metric to optimize and groupwise inference as main contribution. Proposed GVox demonstrates superior performance compared to the baseline Elastix, with significantly improved image similarity and computational efficiency.

**Keywords:** Arterial spin labeling, deep learning-based image registration, cortical signal.

## 1. Introduction

ASL is a magnetic resonance imaging (MRI) technique that enables the characterization of renal blood flow by magnetically label arterial blood water spins in a non-invasive way (Nery et al., 2018). *Control* images contain static tissue signal while *label* images contain both static tissue signal and tagged blood signal. Perfusion (or ASL) maps are computed by subtracting control and label images voxel by voxel. Added to the motion caused by the anatomical displacement of the kidney (due to respiratory and motion of the patient) is the presence of blood inflow that expands through the kidney, that being the case of misregistered image sequences and subsequent post-processing errors. Recent deep learning-based image registration (DLIR) approaches use convolutional neural networks (CNN) for the estimation of voxel or pixel-wise spatial correspondences (Zöllner et al., 2009), such as VoxelMorph (Balakrishnan et al., 2019). DLIR approaches model the function  $g_{\theta}(I_f, I_m) = \phi$ .  $I_f$  and  $I_m$  are the fixed and moving images over a  $n$ -D spatial domain  $\Omega \in \mathbb{R}^n$ , respectively,  $\phi$  is the deformation field, and  $\theta$  are the learnable parameters of function  $g$  that compute

the optimal deformation field  $\phi$  for each image pair  $(I_f, I_m)$ . The model warps  $I_m(p)$  into  $I_m(\phi(p))$  by spatial transformer network (STN) (Jaderberg et al., 2015) based module, considering  $p$  as a pixel in a 2D image. The spatial transformation enables the calculation of dissimilarity loss function named as  $F_{sim}$  and updates  $\theta$ .

**Aim.** We propose the following pipeline for ASL MoCo on healthy kidneys: 1) to extend VoxelMorph to provide a tool tailored to the unique needs of ASL, 2) to introduce a cortical enhancement based measurement as a customized loss metric used as a complement to the main similarity loss function, and 3) to implement groupwise inference stage by using principal component analysis (PCA).

## 2. Method

**Data.** The dataset consists of 48 healthy kidney ASL studies with acquisition matrix of  $96 \times 96$ . Each study contains a  $M_0$  proton density or reference image and a set (8-25 pairs) of control/labels ASL pairs. Manually drawn cortical masks are used for training. Images are intensity-normalized  $\{0 - 1\}$  and affinely registered with Elastix before training.

**Registration pipeline** Our MoCo framework consists of separate training and testing stages (Figure 1a). *i. Training.* MoCo models are trained using VoxelMorph-2 architecture (Balakrishnan et al., 2019) in unique study (higher motion measured), Adam optimizer, learning rate of  $10^{-4}$ , 1000 epochs, 100 steps per epoch, a batch size of 8, and weighting regularization parameter ( $\lambda$ ) of 0.9 (experimentally set). The method is implemented on Python 3.7 using PyTorch 1.13 on GPU NVIDIA GeForce RTX 3060. Our models are trained using both unidirectional and bidirectional registration (Figure 1a). We use a customized loss function that measures  $F_{sim}$  between randomly selected  $I_f$  and  $I_m$  pairs, consisting of a weighted sum of normalized cross correlation ( $NCC$ ) and cortical temporal signal-to-noise ratio ( $tSNR$ ) (Equation 1).  $tSNR$  is calculated as the ratio of time series (random input images on training) average perfusion signal and its temporal standard deviation (SD), over the manually drawn ROI of the cortex.

$$F_{sim}(I_f, I_w, S_{I_f}, S_{I_w}) = -\omega_{ncc}NCC(I_f, I_w) - \omega_{tsnr}tSNR((S_{I_f}, S_{I_w})) \quad (1)$$

where  $I_w$  is  $I_m \circ \phi$ ,  $S_{I_f}$  and  $S_{I_w}$  are the cortical masks corresponding to  $I_f$  and  $I_w$ , being  $S_{I_w} = S_{I_m} \circ \phi$ .  $\omega_{ncc}$  and  $\omega_{tsnr}$  are the corresponding weights for  $NCC$  and  $tSNR$ , set as 0.9 and 0.1, respectively. *ii. Testing.* Intra-patient groupwise MoCo is achieved by evaluating trained models on unseen studies by iteratively warping  $N$  images into PCA-based template image  $T$ . The eigenvector  $V = (w_1, w_2, \dots, w_n)$  associated with the largest eigenvalue is used as weights for the construction of  $T$ . In each iteration, the process randomly designates a moving image and registers it with the template image  $T$ .  $T$  is initialized using  $N-1$  non-warped images from the unregistered sequence. As the inference advances, images from the unregistered sequence are systematically updated with warped, so it does  $T$ .

**Baseline.** We use Elastix (Klein et al., 2010) as the reference non-learning-based MoCo method due to its widespread use in ASL MoCo (Nery et al., 2019; Bones et al., 2019).

## 3. Results

We use mean structural similarity index (MSSIM) for MoCo evaluation, that assesses the structural similarity of images by comparing luminance, contrast, and structure (Sassi et al.,

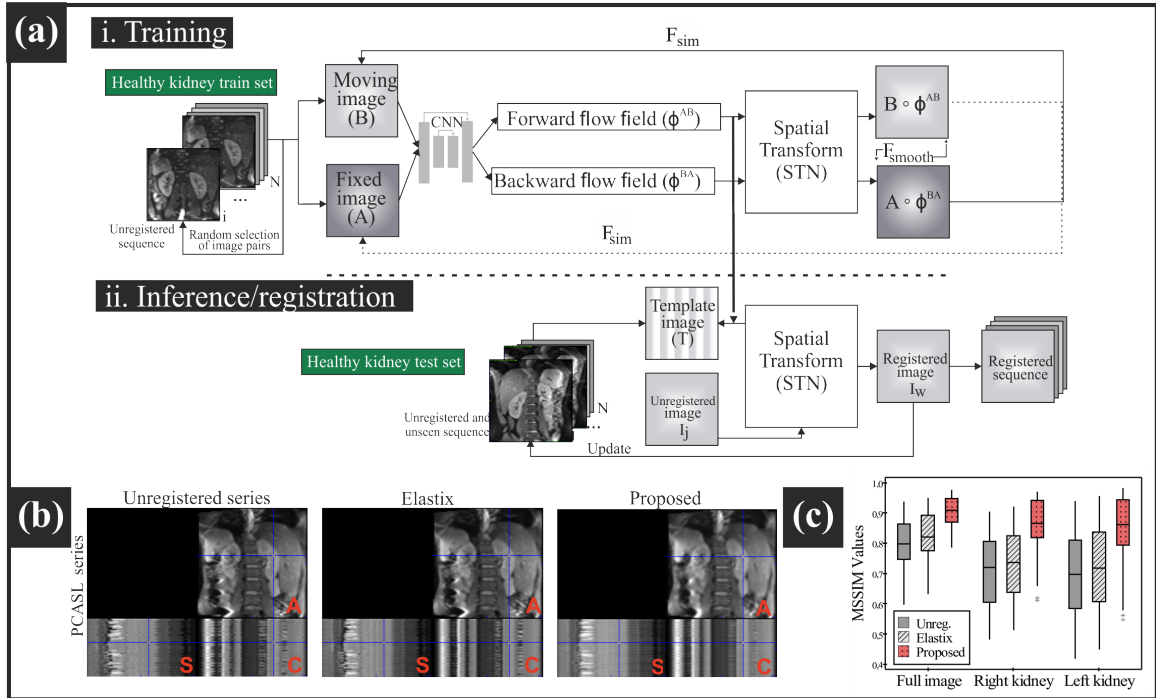


Figure 1: Overview of GVox. a) MoCo pipeline. b) MoCo effects on image alignment (axial (A), sagittal (S), and coronal (C) views). c) Boxplot showing MSSIM values.

2008). MSSIM is calculated in the whole image, and in the right and left kidney encompassing bounding box, separately. For statistical analysis, we use ANOVA with Dunnett pairwise comparison, where unregistered series is set as the control group. MSSIM was significantly higher ( $p < 0.05$ ) in GVox compared to both unregistered series and Elastix. Regarding the evaluation of registration direction, none statistical ( $p > 0.05$ ) difference was found in MSSIM metric between unidirectional and bidirectional training frameworks. Mean registration times for Elastix and GVox are 76 and 9 secs per epoch, respectively. Thus, GVox demonstrates higher computational efficiency. Qualitative results (Figure 1b) demonstrate excellent performance of GVox in sequentially aligning images.

#### 4. Conclusions

Our GVox method shows significant promise in renal ASL MoCo, preserving anatomical structure while delivering efficient runtimes and superior performance in image similarity. These findings suggest the potential of GVox to enhance the accuracy and reliability of medical MoCo for renal imaging. Moreover, groupwise registration allows for the inclusion of full ASL data series. As future work, our method shows potential for enhancing renal imaging accuracy and reliability, particularly in challenging scenarios such as transplant patients and chronic kidney disease patients with inconsistent breathing patterns.

#### Acknowledgments

Project PC181-182 RM-RENAL and predoctoral grant 0011-0537-2021-000050, by the Department of University, Innovation and Digital Transformation, Government of Navarra.

## References

- Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, John Guttag, and Adrian V. Dalca. Vox-elMorph: A Learning Framework for Deformable Medical Image Registration, September 2019.
- Isabell K. Bones, Anita A. Hartevelde, Suzanne L. Franklin, Matthias J. P. van Osch, Jeroen Hendrikse, Chrit T. W. Moonen, Clemens Bos, and Marijn van Stralen. Enabling free-breathing background suppressed renal pCASL using fat imaging and retrospective motion correction. *Magnetic Resonance in Medicine*, 82(1):276–288, July 2019. ISSN 1522-2594. doi: 10.1002/mrm.27723.
- Max Jaderberg, Karen Simonyan, Andrew Zisserman, and koray kavukcuoglu. Spatial Transformer Networks. In *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.
- Stefan Klein, Marius Staring, Keelin Murphy, Max A. Viergever, and Josien P. W. Pluim. Elastix: A toolbox for intensity-based medical image registration. *IEEE transactions on medical imaging*, 29(1):196–205, January 2010. ISSN 1558-254X. doi: 10.1109/TMI.2009.2035616.
- Fabio Nery, Isky Gordon, and David L. Thomas. Non-Invasive Renal Perfusion Imaging Using Arterial Spin Labeling MRI: Challenges and Opportunities. *Diagnostics (Basel, Switzerland)*, 8(1):E2, January 2018. ISSN 2075-4418. doi: 10.3390/diagnostics8010002.
- Fabio Nery, Enrico De Vita, Chris A. Clark, Isky Gordon, and David L. Thomas. Robust kidney perfusion mapping in pediatric chronic kidney disease using single-shot 3D-GRASE ASL with optimized retrospective motion correction. *Magnetic Resonance in Medicine*, 81(5):2972–2984, May 2019. ISSN 1522-2594. doi: 10.1002/mrm.27614.
- Olfa Ben Sassi, Tijani Delleji, Abdelmalik Taleb-Ahmed, Imed Feki, and Ahmed Ben Hamida. MR Image Monomodal Registration Using Structure Similarity Index. In *2008 First Workshops on Image Processing Theory, Tools and Applications*, pages 1–5, November 2008. doi: 10.1109/IPTA.2008.4743741.
- Frank G. Zöllner, Rosario Sance, Peter Rogelj, María J. Ledesma-Carbayo, Jarle Rørvik, Andrés Santos, and Arvid Lundervold. Assessment of 3D DCE-MRI of the kidneys using non-rigid image registration and segmentation of voxel time courses. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, 33(3):171–181, April 2009. ISSN 1879-0771. doi: 10.1016/j.compmedimag.2008.11.004.