



**HAL**  
open science

# Enhancing Fluoroscopy-Guided Interventions: a Neural Network to Predict Vessel Deformation without Contrast Agents

Francois Lecomte, Valentina Scarponi, Pablo A Alvarez, Juan M Verde,  
Jean-Louis Dillenseger, Eric Vibert, Stéphane Cotin

## ► To cite this version:

Francois Lecomte, Valentina Scarponi, Pablo A Alvarez, Juan M Verde, Jean-Louis Dillenseger, et al.. Enhancing Fluoroscopy-Guided Interventions: a Neural Network to Predict Vessel Deformation without Contrast Agents. The Hamlyn Symposium on Medical Robotics, Jun 2023, London, United Kingdom. pp.75-76, 10.31256/HSMR2023.38 . hal-04356336

**HAL Id: hal-04356336**

**<https://hal.science/hal-04356336>**

Submitted on 20 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Enhancing Fluoroscopy-Guided Interventions: a Neural Network to Predict Vessel Deformation without Contrast Agents

François Lecomte<sup>1</sup>, Valentina Scarponi<sup>1</sup>, Pablo A. Alvarez<sup>1</sup>, Juan M. Verde<sup>2</sup>,  
Jean-Louis Dillenseger<sup>3</sup>, Eric Vibert<sup>4</sup>, and Stéphane Cotin<sup>1</sup>

<sup>1</sup>Inria, CNRS, Univ Strasbourg, Strasbourg, France

<sup>2</sup>Department of Image-Guided Liver Interventions, IHU-Strasbourg, Strasbourg, France

<sup>3</sup>Univ Rennes, Inserm, LTSI - UMR 1099, Rennes, France

<sup>4</sup>AP-HP Hôpital Paul-Brousse, Centre Hépatobiliaire, Villejuif, France

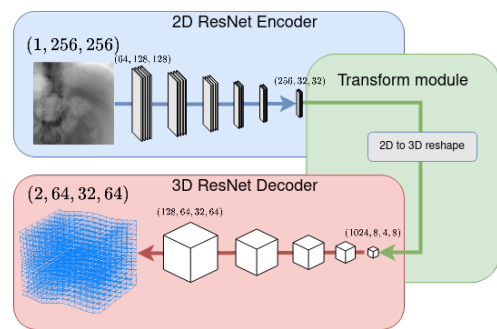
francois.lecomte@inria.fr

## INTRODUCTION

Image-guided procedures have experienced a rapid increase in popularity in recent years. The advancements in medical imaging technology have led to a shift in medical images from being used primarily for diagnosis to being a critical tool in theragnostic and therapeutic procedures. This shift has resulted in the emergence of new fields such as interventional radiology (IR), therapeutic endoscopy (TE), and minimally invasive image-guided surgery (IGS), with an increasing number of professionals adopting these techniques in their clinical practices due to improved outcomes [1]. One of the most widely used imaging methods in these procedures is X-ray-based imaging, including computed tomography (CT), 2D C-arm fluoroscopy, and cone-beam CT scans. These procedures typically require the use of contrast agents (CA) to visualize soft tissues with high definition and contrast. However, the use of CA presents several challenges, including the limited volumes that can be used and the toxicity of the agents when they are injected intravascularly [2]. The CA also follows the patient's hemodynamics, leading to transient visualization and asynchronous image guidance. In this paper, we aim to address the technical issues related to contrasted X-Ray images in image-guided therapy. We propose a deep learning approach that will allow for the visualization of vessels during image-guided procedures without the need for contrast agents, making these procedures safer, and more efficient, while providing real-time guidance.

## MATERIALS AND METHODS

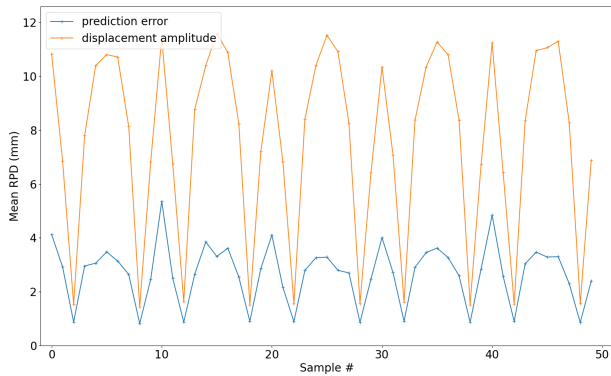
Before the intervention, a CT scan of the patient is acquired, vessels and other structures of interest are segmented and the planning of the intervention is performed. At this stage, the poses of both the C-arm and the patient are determined for the intervention. Then, at treatment time, the C-arm and the patient are positioned as per planning. However, due to breathing induced anatomical deformations, segmented vessels cannot be superposed to the fluoroscopic images. To compensate for these deformations, a non-rigid registration method is required.



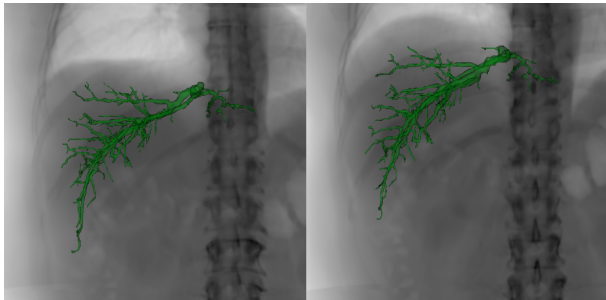
**Fig. 1** Each block in the Encoder downscales the feature maps and increases their number by a factor of 2. In the Decoder, this is reversed. The last decoder layer transforms the 128 feature maps into a 2-channel 3D image.

Using the planning CT scan of the patient, we generate a dataset of Digitally Reconstructed Radiographs (DRRs) and displacement field pairs. This dataset is then used to train a neural network to recover a displacement field from a single fluoroscopic image. At the time of the intervention, the C-arm is adjusted to the planned pose, an X-ray image is acquired, and the network predicts in real-time the 3D deformation field which is used to warp the segmented structures to visually augment the fluoroscopy. Our fully convolutional network architecture is detailed in 1. Features are extracted from the 2D input image and then reshaped into 3D features. Finally, the 3D features are decoded into a 3D displacement field. This displacement field represents the non-rigid transform of the CT to the anatomy visible in the fluoroscopic image.

We assume that the deformations at the time of the intervention can be modeled as smooth and invertible displacement fields, preventing unrealistic deformations such as self intersection and excessive stretching or compression. To obtain such deformations, we model the displacement field as a sum of Gaussian Kernels. Specifically, the displacement field is integrated over  $n = 100$  timesteps, while checking that smoothness and invertibility constraints are preserved after each timestep. The parameters of the Gaussian Kernels are randomized



**Fig. 2** The average RPD error of our method on the testing data against the average RPD displacement.



**Fig. 3** Augmented DRR at full inspiration (left), and full expiration (right), with the predicted hepatic veins position.

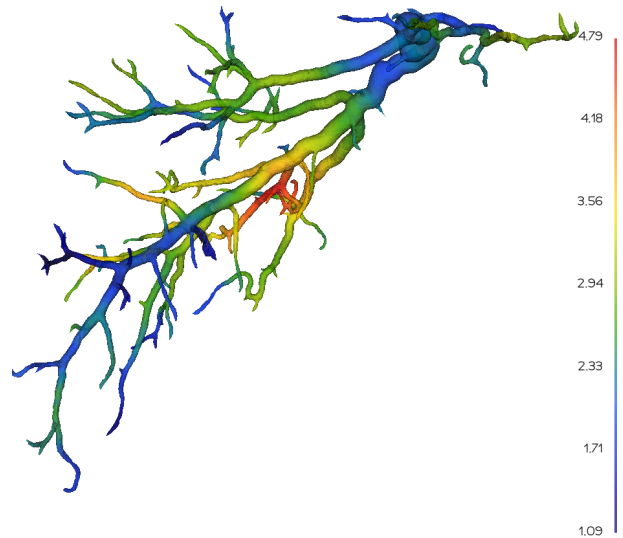
at each step, following the principle of Domain Randomization. Using the DeepDRR framework [3], one DRR is generated for each deformation of the pre-operative CT.

## RESULTS

A human liver CT obtained from a patient of the Paul Brousse hospital in Paris was used to generate a 10,000 sample dataset, split into 8,000 training samples and 2,000 validation samples. The maximum amplitude of deformation in the dataset was 22mm and 40mm in the LR direction and SI direction respectively.

The testing dataset was generated from the same 3D CT, this time using BSpline transforms tailored to mimic a breathing motion. Specifically, inhale and exhale phases and sliding motion of the organs against the bones were modeled. In this case, the maximum amplitude was 10mm and 25mm for the SI and LR directions. The dataset contains 5 inhale/exhale periods for a total of 50 samples. The accuracy of the network was measured via the reprojection distance (RPD) metric. Hepatic veins were deformed using the ground truth and the predicted displacement fields. The deformed mesh points were projected onto the image plane and the 2D RPD error was measured.

The mean RPD error on the testing dataset was  $2.7 \pm 1.9$  mm while the mean RPD displacement was  $7.7 \pm 3.9$  mm. Figure 4 shows the distribution of the error on the hepatic veins. Figure 3 shows an example of image augmentation by our method. A full video is available at <https://mimesis.inria.fr/project/augmented-fluoroscopy/>.



**Fig. 4** The average RPD error in mm on the hepatic veins position on the testing dataset.

## DISCUSSION

Even though the testing data were generated differently from the training data, the prediction of the network still reduced the error from 7.7 to 2.7mm. This validates our Domain Randomization approach, as the network learns to map a deformed fluoroscopy to a displacement field, and generalizes well on the testing data. This constitutes an advantage over other methods that might use a patient-specific motion prior obtained from a 4D-CT to train a neural network to predict deformations from a fluoroscopic image. A limitation arising from 2D fluoroscopy is that the displacement perpendicular to the image plane is not visible and thus cannot be predicted, but this is mitigated by the fact that the resulting out-of-plane error is also not visible on the augmented image.

## ACKNOWLEDGEMENT

The authors would like to thank Simon Rouze, MD for his valuable inputs during the development of this method. This work was funded by the French national research agency ANR (VATSOP project). The work of Juan Verde was supported by French state funds managed within the “Plan Investissements d’Avenir” and by the ANR (reference ANR-10-IAHU-02).

## REFERENCES

- [1] A. Epstein and al., “Impact of Minimally Invasive Surgery on Medical Spending and Employee Absenteeism,” *JAMA Surgery*, vol. 148, no. 7, pp. 641–647, 07 2013.
- [2] C. Mamoulakis and al., “Contrast-induced nephropathy: Basic concepts, pathophysiological implications and prevention strategies,” *Pharmacology Therapeutics*, vol. 180, pp. 99–112, 2017.
- [3] M. Unberath and al., “DeepDRR – A Catalyst for Machine Learning in Fluoroscopy-Guided Procedures,” *Lecture Notes in Computer Science*, vol. 11073, pp. 98–106, 2018.
- [4] N. Izumi and al., “Risk factors for distant recurrence of hepatocellular carcinoma in the liver after complete coagulation by microwave or radiofrequency ablation,” *Cancer*, vol. 91, pp. 949–956, 2001.
- [5] J. Tobin and al., “Domain randomization for transferring deep neural networks from simulation to the real world,” in *IROS, 2017*, pp. 23–30.