Registration of Longitudinal Spine CTs for Monitoring Lesion Growth

Malika Sanhinova[a,b], **Nazim Haouchine**[a,c], **Steve D. Pieper**[d], **William M. Wells III**[a,c], **Tracy A. Balboni**[a,c,e], **Alexander Spektor**[a,e], **Mai Anh Huynh**[a,e], **Jeffrey P. Guenette**[a,c], **Bryan Czajkowski**[a,c], **Sarah Caplan**[a,e], **Patrick Doyle**[a,e], **Heejoo Kang**[a,e], **David B. Hackney**[a,b], **Ron N. Alkalay**[a,b]

[a]Harvard Medical School, Boston, MA, USA

[b]Beth Israel Deaconess Medical Center, Boston, MA, USA

[c]Brigham and Women's Hospital, Boston, MA, USA

[d]Isomics, Inc., Cambridge, MA, USA

[e]Dana-Farber Cancer Institute, Boston, MA, USA

Abstract

Accurate and reliable registration of longitudinal spine images is essential for assessment of disease progression and surgical outcome. Implementing a fully automatic and robust registration is crucial for clinical use, however, it is challenging due to substantial change in shape and appearance due to lesions. In this paper we present a novel method to automatically align longitudinal spine CTs and accurately assess lesion progression. Our method follows a two-step pipeline where vertebrae are first automatically localized, labeled and 3D surfaces are generated using a deep learning model, then longitudinally aligned using a Gaussian mixture model surface registration. We tested our approach on 37 vertebrae, from 5 patients, with baseline CTs and 3, 6, and 12 months follow-ups leading to 111 registrations. Our experiment showed accurate registration with an average Hausdorff distance of 0.65 mm and average Dice score of 0.92.

1. DESCRIPTION OF PURPOSE

Longitudinal lesion monitoring plays a pivotal role in observing changes of metastatic vertebrae and disease progression over time. By consistently evaluating the size, shape, and characteristics of lesions, healthcare professionals can gain valuable insights into the effectiveness of treatments, predict potential complications, and make informed decisions about patient care. The primary objective of this research is to conduct a longitudinal study focused on lesion growth in metastatic spines. Our approach allows for a comprehensive understanding of how metastatic lesions evolve and impact a patient's clinical course, ultimately leading to more personalized and effective medical interventions. The accurate tracking of lesions necessitates the alignment of 3D images captured at distinct time points. An initial CT scan establishes a baseline at 0 months, and follow-up scans at 3, 6, and 12-month are registered to this baseline. Registration is required due to altered parameters of the CT scan environment, such as origin, resolution, reconstruction methods, field of

view, etc. This registration process enables lesion comparison and observation of the lesion development over time.

Several notable contributions have been made in the field of advancing medical image registration techniques, each with its own limitations that provide valuable insights for further refinement and development. The methods described by Hille et al.,[1] Cai et al.[2] and Gueziri et al.[3] address medical image registration involving MRI, CT, and ultrasound during spine interventions. These techniques are primarily utilized to fuse intraoperative images of lower resolution with pre-operative high-resolution scans, resulting in improved image quality—typically outperforming the outcomes of purely intraoperative CT imaging. However, while such registration methods exhibit speed and precision for surgical interventions, they presume a consistent appearance of vertebrae pre- and post-intervention, which might not align with cases involving lesions. Balakrishnan et al.[4] and Hu et al.[5] developed state-of-the-art learning-based registration strategies that integrate weakly-supervised segmentations during training, by incorporating them into the loss function. Zhao et al.[6] introduced SpineRegNet which allows for affine-elastic deformation field estimation for spine scans. By registering MRI to CT scans, the method combines rigid vertebral and elastic disk registration, ensuring the preservation of spine biomechanics. The framework integrates multiple modules for flexible spinal movement, fusion of multiple deformation fields, and preservation of vertebrae rigidity.

Previous methods are focused on pre-operative to intraoperative registration. To the best of our knowledge, longitudinal registration was first addressed by Glocker et al. in 2014.[7] To overcome the initialization challenges of standard registration techniques for cases with small overlap, the authors proposed a registration method that includes a prior learning-based classification to estimate vertebrae centroids. This additional semantic information significantly improved registration compared to other initialization techniques, while applying the same intensity-based registration components. However, using intensity-based registration methods could be limiting for cases of metastatic spines. Metastatic vertebrae can undergo significant changes within a short time period, leading to dramatic alterations in their appearance and, consequently, intensity. This could hinder the accuracy and effectiveness of the intensity-based registration process.

Our approach was developed to handle cases with large shape and intensity variations by utilizing segment-based registration. First, segmentation masks are predicted with a deep learning segmentation model that was trained on metastatic spines. Training set included spines with significant lesion progression and missing portions of vertebrae. Then, obtained segmentation masks are used to register follow-ups into a baseline on vertebra level. Rigid registration prevents deformation of bone structure, while segment-level registration ensures a unique transformation matrix per vertebra. Thus, the proposed two-step approach can be used for metastatic vertebrae undergoing significant shape and intensity changes.

2. METHODS

2.1 Multi-class Segmentation of Lesioned Vertebrae from CT

Let us define the training set $T^S = \{(\mathbf{I}_i, \mathbf{L}_j)\}i$ composed of CT images $\mathbf{I}_i$ of the vertebrae and their corresponding labels $\mathbf{L}_i$, with $i \in |T^S|$. We train a segmenter $S(\mathbf{I}; \theta_S)$ with θ_S being the learned parameters for network S. We use this to generate a per-voxel probability map $\mathbf{P}$ that will associate with each voxel of an image $\mathbf{I}$ the probability of that voxel being part of a vertebra level.

The network S is a deep CNN that follows a standard 3D U-Net architecture[8] with a Softmax final layer. Because we are interested in a joint segmentation and classification where we can outline the shape of the vertebra and identify its level, we optimize a *categorical cross-entropy* loss function using gradient descent over the parameters θ_S following:

$$L\left(\theta_S, T^S\right) = \sum_{i \in |T^S|} \|(I_i; \theta_S) - L_i\|_2$$

(1)

We used the nnU-Net[9] implementation, a self-adapting framework for 3D full-resolution image segmentation based on the U-Net architecture, to train the network (see Fig. 1).

Surface-based Longitudinal Registration—In order to perform the longitudinal CT registration, we used triangulated surfaces reconstructed from the segmentation masks resulting from S. These surfaces are reconstructed using a marching cubes algorithm, followed by an edge collapse-based incremental decimation.[10] We generate a surface for the baseline CT and the 3-months, 6-month and 12-month follow-ups CTs. We perform a pair-wise registration, and align the two surfaces by solving a probability density estimation problem.[10] The baseline surface represents Gaussian mixture model (GMM) centroids, and each follow-up surface represents data points. This can be solved by minimizing the negative log-likelihood function based on the GMM approach where we drive the surface of the model using implicit surface-to-surface forces, from baseline to follow-ups. Formally, assuming the N observations $\mathbf{x}$ and M GMM centroids $\mathbf{y}$, we aim at finding the transformation of the GMM centroid locations as $T(\mathbf{y}m; \mathbf{R}, \mathbf{t}, s) = s\mathbf{R}\mathbf{y}m + \mathbf{t}$, where $\mathbf{R}$ is a rotation matrix, $\mathbf{t}$ is a translation vector and s is a scaling parameter. The objective function to minimize is as follows:

$$E\left(\mathbf{R}, \mathbf{t}, s, \sigma^2\right) = \frac{1}{\sigma^2} \sum_{m,n=1}^{M,N} P(\mathbf{y}_m|x_n)\|x_n - s\mathbf{R}y_m - \mathbf{t}\|^2 + \frac{3N_p}{2}log\left(\sigma^2\right)$$

(2)

where $P(\cdot)$ denotes the GMM probability density function that smoothly weight the correspondences between the two surfaces. Finally, this function is optimized using an expectation maximization (EM) algorithm (see Fig. 1).

3. RESULTS

The segmentation framework was trained on 55 spines using a 5-fold cross-validation strategy. The CTs were acquired from 2 manufacturers in 2 hospitals with an image resolution of 0.35×0.35 mm - 1.0×1.0 mm in-plane and 0.5 mm - 1.5 mm thickness. We defined 18 classes across vertebral levels from C7 to L5, achieving a validation accuracy of 0.886 ($\pm$ 0.382). Testing the model on 5 unseen spines yielded an average accuracy of 0.943 ($\pm$ 0.239), indicating its robust generalization. Then, we used the trained segmentation model to predict the baseline and follow-up 3D surfaces for 5 patients without ground true segmentations.

By registering the 37 predicted vertebrae individually, we performed 111 registrations, aligning 3, 6 and 12 months follow-ups with the baseline images. The mean Dice Similarity Coefficient achieved by our registration method is 0.92 ± 0.11. The Dice score provides a measure of the overlap between the baseline and a follow-up, indicating the accuracy of the registration process. The high mean Dice of 0.92 ± 0.11 signifies a precise alignment between the registered vertebrae. The average Hausdorff distance obtained from our method was measured at 0.65 mm (95% within 1.74 mm). The Hausdorff distance reflects the discrepancy between the points of two sets, indicating the extent of spatial dissimilarity between the registered vertebrae. Our method's low average Hausdorff distance suggests excellent alignment of the follow-ups into the baseline. The registration is computationally fast with an execution time below 5 seconds. Figure 2 represents our method's results, demonstrating an example of lesion growth over time. Here, an accurate segmentation mask and precise follow-up registration enabled longitudinal lesion tracking in a metastatic vertebrae, including size, shape and structure monitoring.

Importantly, there is a correlation between segmentation accuracy and subsequent registration performance. As evidenced by our experiments, improved segmentation quality directly influenced the registration Dice score and Hausdorff distance. When segmentation accurately identifies the anatomical boundaries, it provides a more precise representation of the target structures, resulting in improved alignment during the registration process.

4. CONCLUSION

In this paper we presented an automated deep-learning-based tool for registration of longitudinal spine CT. Analysis on the aligned CTs can be used to visually assess and quantify lesion growth and response to treatment. Further studies with larger datasets and different types of spinal diseases are warranted to validate the performance of these algorithms and their potential in predicting disease progression for improved treatment and management. Future work will first consist of providing shape analysis of the longitudinal lesion growth. We will also integrate the type of lesion in the classification model to further monitor and evaluate the lesion growth. To extend the evaluation of this approach we are preparing a larger dataset of longitudinal spine CTs.

ACKNOWLEDGEMENT

## REFERENCES

[1]. Hille G, Saalfeld S, Serowy S, and Tönnies K, "Multi-segmental spine image registration supporting image-guided interventions of spinal metastases," Computers in biology and medicine 102, 16–20 (nov 2018). [PubMed: 30236968]

[2]. Cai Y, Wu S, Fan X, Olson J, Evans L, Lollis S, Mirza SK, Paulsen KD, and Ji S, "A level-wise spine registration framework to account for large pose changes," International journal of computer assisted radiology and surgery 16, 943–953 (jun 2021). [PubMed: 33973113]

[3]. Gueziri HE and Collins DL, "Fast registration of CT with intra-operative ultrasound images for spine surgery," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11397 LNCS, 29–40 (2019).

[4]. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, and Dalca AV, "VoxelMorph: A Learning Framework for Deformable Medical Image Registration," IEEE Transactions on Medical Imaging 38, 1788–1800 (sep 2018).

[5]. Hu Y, Modat M, Gibson E, Li W, Ghavami N, Bonmati E, Wang G, Bandula S, Moore CM, Emberton M, Ourselin S, Noble JA, Barratt DC, and Vercauteren T, "Weakly-supervised convolutional neural networks for multimodal image registration," Medical Image Analysis 49, 1–13 (oct 2018). [PubMed: 30007253]

[6]. Zhao L, Pang S, Chen Y, Zhu X, Jiang Z, Su Z, Lu H, Zhou Y, and Feng Q, "SpineRegNet: Spine Registration Network for volumetric MR and CT image by the joint estimation of an affine-elastic deformation field," Medical Image Analysis 86, 102786 (may 2023). [PubMed: 36878160]

[7]. Glocker B, Zikic D, and Haynor DR, "Robust registration of longitudinal spine CT," Medical image computing and computer-assisted intervention : MICCAI … International Conference on Medical Image Computing and Computer-Assisted Intervention 17(Pt 1), 251–258 (2014). [PubMed: 25333125]

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

[9]. Isensee F, Jaeger PF, Kohl SA, Petersen J, and Maier-Hein KH, "nnu-net: a self-configuring method for deep learning-based biomedical image segmentation," Nature methods 18(2), 203–211 (2021). [PubMed: 33288961]

[10]. Fedorov A, Khallaghi S, Sánchez CA, Lasso A, Fels S, Tuncali K, Sugar EN, Kapur T, Zhang C, Wells W, et al. , "Open-source image registration for mri–trus fusion-guided prostate interventions," International journal of computer assisted radiology and surgery 10, 925–934 (2015). [PubMed: 25847666]
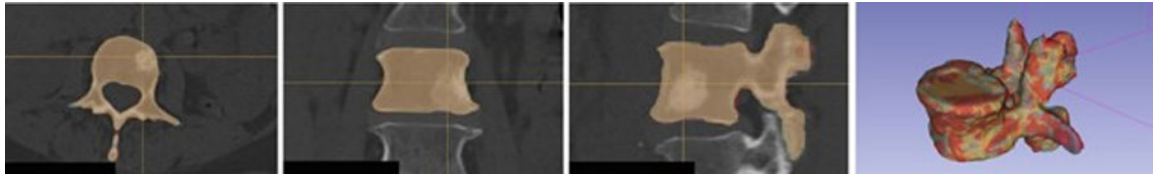
Figure 1:
Our two-step registration pipeline first classifies and segment the vertebra to generate 3D surfaces using, then use a surface-based registration approach to align the longitudinal CTs.
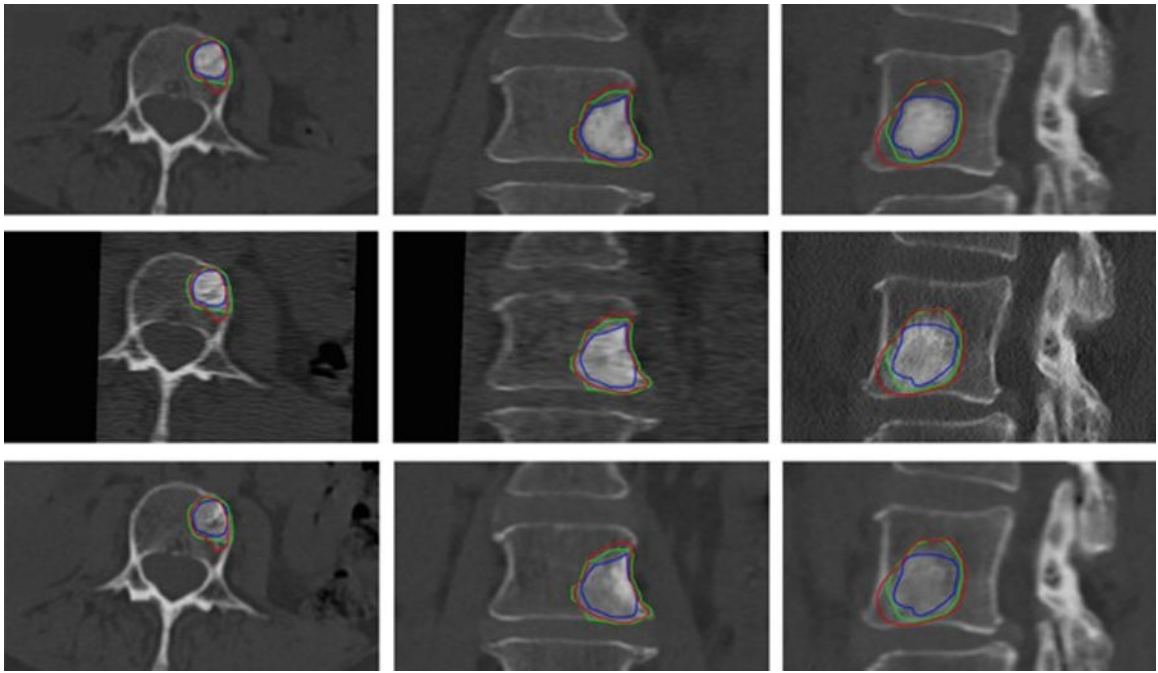
Figure 2:
Axial, coronal and sagittal planes showing the lesion progression in one case. The lesion is outlined over time (baseline in blue, 3 months in green, 6 months in red) and fused with the longitudinal CT set. Top to bottom: baseline, 3 and 6 months.