CONTOURING BY UNIT VECTOR FIELD REGRESSION

Sarim Ather[†]

Amir Jamaludin*

Timor Kadir[‡]

Rhydian Windsor*

* Visual Geometry Group, Department of Engineering Science, University of Oxford
[†] Oxford University Hospitals NHS Foundation Trust
[‡] Plexalis Ltd

ICAAIIS LIU

ABSTRACT

This work introduces a simple deep-learning based method to delineate contours by 'walking' along learnt unit vector fields. We demonstrate the effectiveness of our pipeline on the unique case of open contours on the task of delineating the sacroiliac joints (SIJs) in spinal MRIs. We show that: (i) 95% of the time the average root mean square error of the predicted contour against the original ground truth is below 4.5 pixels (2.5mm for a standard T1-weighted SIJ MRI), and (ii) the proposed method is better than the baseline of regressing vertices or landmarks of contours.

Index Terms— CNN, MRI, Spine, SIJ, Sacroiliac Joint, Vector Field

1. INTRODUCTION

Contouring objects is a very important step in various medical image analysis tasks. Currently, one common approach is to predict a segmentation map of the object and then extract the map's edges. However, this approach has limitations. Firstly, the output segmentations are not necessarily a single interconnected volume and thus additional post-processing is required before finding edges, which can introduce errors (e.g. by removing additional volumes). Secondly, this method does not allow for detecting open contours. An alternative approach is to treat pixels along the open contour as segmentation targets. However, this approach often leads to small, challenging segmentation targets. Furthermore, these approaches do not guarantee a unique solution or easily allow for sub-pixel precision contours in both the open and closed settings.

Therefore, in this paper, we propose a new method to delineate contours, avoiding these limitations. This is done by 'walking' along a learnt vector field. Along the contour, the field should point parallel to the contour, whereas outside the contour the field should point to the nearest contour point. To demonstrate the effectiveness of this method, we apply it to a novel task; delineating the sacroiliac joint (SIJ) boundary in clinical MRI scans.

Sacroiliac Joint Delineation. The SIJ is the joint between the sacrum of the spine and the ilium bones of the pelvis. There are two SIJs per person, one on the left and one on the right.



Fig. 1. Overview of the contouring pipeline on an example SIJ MRI. The model outputs two vector fields, one for both the left (red) and right (green) SIJs. Each vector field is shown as a gradient map of the angle (in degrees) of the vector at that point. These vector fields are then used to extract contours for both the SIJs, shown in the bottom left panel.

MR imaging is typically done to look at the inflammation of the SIJ, or sacroiliitis, which is one of the causes of low back pain and part of the diagnosis for ankylosing spondylitis (AS). In AS, the severity of SIJ inflammation is used to assess disease progression. AS Grading systems often refer to specific regions surrounding the SIJ [1], which makes SIJ detection a must. Since the SIJ is defined as the space between two bones, we follow the approach suggested by [2] and delineate each SIJ as an individual open contour, which is beneficial for the further downstream task of grading the SIJ.

Related Work. There have been multiple works on detecting or segmenting parts of the spine in spinal medical imaging across several imaging modalities, e.g. intervertebral discs [3] and vertebral bodies in MRI [4] and CT scans [5] as well as the whole spine in DXA scans [6, 7]. However, there has been relatively little research on detecting the SIJ and related downstream tasks, for example, inflammation prediction or quantifying structural changes. The closest work to date on SIJ delineation is [8]. However, this method focuses on the classification of sacroiliitis and requires manual annotation to locate the SIJ region. Another closely related work is [9],



Fig. 2. The Unit Vector Field (UVF): (a) a slice of an SIJ MRI with annotated landmarks in red delineating the left SIJ (with respect to the patient), (b) the resulting target UVF, overlaid on top of a gradient map of the field's direction in degrees.

where the authors propose a method to detect changes in the SIJ. However, this is done without explicitly focusing on the SIJ region, instead taking the whole slice of an SIJ MRI as input. We propose that by delineating the SIJ, models can focus on the exact region of the disease without additional noise from surrounding anatomical structures.

Our contouring method has analogies to several works on shape representation using deep learning via implicit functions (e.g. [10, 11, 12]). In this case, rather than representing shapes as a binary mask over a regular grid of voxels, a model learns $f : \mathbb{R}^3 \to \mathbb{R}$, such that f(x, y, z) estimates the closest distance from point (x, y, z) to the object of interest's surface (signed distance functions), or whether (x, y, z) is occupied by the shape (occupancy functions). These methods allow for sub-pixel/voxel precision representations of surfaces. Though we validated our approach on SIJ MRIs, it is worth noting that open contours are widely used in other medical imaging tasks e.g. torso contour segmentation for better ECG interpretation [13], and reconstructing 3D meshed of the heart from 2D cardiac MRIs [14].

2. APPROACH OVERVIEW

Our method takes as input 2D images and outputs an array of vertices delineating the contour of interest. This is done by a two stage approach: (i) Firstly, a model to predicts *a unit vector field* (UVF) for the image. At location x, the UVF indicates the direction towards the nearest point on the contour of interest (ii) Secondly, we propose a method to extract open contours from this learned vector field. Our overall approach for the task of SIJ delineation can be seen in Figure 1.

2.1. Unit Vector Fields

The idea of contours and vector fields in combination is not a new one; for example, several early works in computer vision combined Snakes [15] with gradient vector flow [16], i.e. a vector field pointing towards object edges in a given im-



(a) Gaussian Heatmaps





(b) Gaussians + UVF



(c) Gaussian + Contour

(d) Original Image + Contour

Fig. 3. Following on from the example shown in Figure 2, alongside the UVF, we regress two 2D Gaussian heatmaps. (a) 2 Gaussians representing the start and end points of the contour, (b) the UVF overlaid on top of the Gaussians, (c) the contour which starts from the Gaussian now marked in Green and ends on the Gaussian marked in Blue, (d) the final contour for the left SIJ marked in Red.

age. However, instead of defining the vector field using object edges, we instead learn the unit vector field, $\hat{\mathbf{v}}_{i,j}$, where at each location in the vector field, (i, j), the field 'points' to the nearest vertex, i.e. annotated ground truth landmark, on the contour of the object. The unit vector field is made of two separate x and y components corresponding to the directions of the vectors in the field. To preserve the directionality of the contour, we impose a rule where vectors laying on top of the contour should 'point' to where the next vertex is expected. An example unit vector field can be seen in Figure 2.

2.2. Extracting Contours From Unit Vector Fields

The unit vector field alone does not obviously indicate where a contour starts and ends. We solve this by also predicting the start and end points with the same network that generates the unit vector field; this is done simultaneously as a separate output. We take inspiration from previous works [4, 17, 18] and regress two distinct Gaussian heatmaps for the start and end points respectively. Each Gaussian has a maximum value of 1 and a variance proportional to the area of the task-relevant object. In our case, we use the sacrum, i.e. the area which lies in between two SIJs. In the case where the contour is without a defined area of interest, we suggest scaling the Gaussian heatmap proportional to the length of the overall con-



(e) Mid-cor. T2-w (e) Mid-cor. STIR (f) Mid-cor. FS

Fig. 4. Example scans in the dataset with their marked-up annotated landmarks. (a), (b), and (c) are slices from the same T1-weighted scan at differing slice positions (anterior, mid-coronal, posterior) while (d), (e), and (f) are mid-coronal examples of different sequences in the dataset.

tour. The beginning of the contour is defined from the Gaussian heatmap designated as the start point. We then iteratively 'walk' following the direction in the UVF, $\hat{v}_{i,j}$, and the contour ends when approaching the second Gaussian heatmap i.e. end point. Each step is 1 unit in magnitude, although this could be adjusted to generate contours of varying fidelity. Figure 3 gives an example of how a contour is defined with the Gaussian heatmaps and the UVF. Since the UVF can be visualized, errors can be more easily interpreted. Though not shown in this work, a closed contour solution would not require heatmaps and could be found by simply searching for a loop in the UVF.

3. DATASET & TRAINING DETAILS

Dataset. The Oxford Sacroiliac Joint (**OSIJ**) dataset is a collection of SIJ MRIs from 339 patients that have undergone scanning for in the Oxford University Hospitals NHS trust. For experiments conducted in this work, the dataset is split into training (80%), validation (10%) and testing (10%) sets on a per subject basis (271:34:34). Each subject possesses an average of two sequences (typically T1, T2, STIR, and FS) resulting in a total of 793 scans. Each scan roughly consists of 20 2D slices, resulting in a total of 16,978 images.

For the annotations of the contour of the SIJs, an expert was tasked with marking the landmarks (vertices) that best define both left and right SIJs through every slice in a given scan. The number of landmarks varies depending on the view of the SIJ; typically, mid-coronal SIJs cover a bigger image area demanding a larger number of landmarks and vice versa. The number of landmarks per slice ranges from 2 to 21.

Training Details. The experiments in this work were con-



Fig. 5. Cumulative test set error distribution (measured in pixels). Baseline is in blue and contouring via UVF is in red.

ducted using a simple U-Net architecture [19]. For each contour, the network predicts 2 Gaussian heatmaps and 2 components (x and y direction) of the unit vector field; separate contours were predicted for each of the two SIJ (left and right). The SIJs are not guaranteed to be inside the field-of-view of the scans and as such these cases were kept in the training set to suppress false positives. The scans were typically squares in shape; thus, they are bi-cubically re-sampled to 224×224 pixels. Slices that were not square were padded with zeros prior to re-sampling so as to not change the aspect ratio.

The network is trained using an Adam optimiser [20] with $\beta_1 = 0.9$, $\beta_2 = 0.999$ and a learning rate of 10^{-3} until convergence. Several augmentations were applied during training, namely: (a) translation $\pm 20\%$, (b) scale $\pm 20\%$, (c) rotation $\pm 15^{\circ}$, (d) left/right flips, (e) additive Gaussian noise, and (f) Gaussian blur. A combination of L2-loss, for the UVF, and weighted L2-loss (see [4]), for the Gaussian heatmaps, is used to train the network.

4. PERFORMANCE EVALUATION & RESULTS

For comparison we compare against a baseline network trained to predict 21 Gaussian heatmaps for each SIJ, 21 being the maximum number of landmarks in the dataset. We find this to be the simplest naïve solution to predict landmarks using a similar U-Net architecture as our proposed UVF approach. Samples with a lower number of annotated points were up-sampled via linear interpolation. At test time, each prediction is compared against the ground truth landmarks of the contour and the root mean square (RMS) error is cal-



(a) Example with GT annotations



Fig. 6. Quantitative result of the baseline against the proposed method on a test set example. Green contours highlight the right SIJ and red contours highlight the left; GT in yellow. (b) and (d) are from the baseline model while (c) and (e) are contours using UVF. Baseline predictions are sparse, with 21 landmarks for each contour, resulting in more aliasing.

culated from the closest points between the prediction and ground truth.

Results for both networks are shown in Figure 5 and Table. 1. Contouring by UVF overall works slightly better than the baseline ranging from 0.14 to 0.35 difference in RMS pixel error up to 95% of the data in the test set. This might not seem like a huge amount, but Figure 6 highlights that there is lower aliasing when looking at the contours using via UVF compared to just predicting landmarks via heatmaps. In general, 95% of the test set have a lower than 4.5 pixel error which for our purposes is adequate for further downstream tasks e.g. defining an ROI for SIJ oedema classification. Figure 7 shows results on several examples both from **OSIJ**'s test





(c) OSIJ Test Sample 2







(b) rID: 73884



(d) rID: 75292



(f) rID: 154033



set and to images extracted from Radiopaedia.

Data Proportion	0.1	0.3	0.5	0.7	0.9	0.95
Baseline Error	0.52	1.00	1.41	2.00	3.40	4.45
UVF Error	0.38	0.72	1.15	1.76	3.10	4.10

Table 1. Table of RMS per proportion of data in the test set.

5. CONCLUSION

In this paper, we presented a pipeline to contour, focusing more on open contours but applicable to closed contours as well, objects in images and demonstrated its use to delineate SIJs in coronal spinal MRIs. Overall, the performance is better than the naive baseline of predicting landmarks of contours and is applicable to other contouring problems in medical image analysis.

6. COMPLIANCE WITH ETHICAL STANDARDS

The scans in the dataset were sourced from retrospective scans in the Oxford University Hospitals approved by the Health Research Authority (IRAS Project ID 207858).

7. ACKNOWLEDGMENTS

The authors would like to thank Aimee Readie and Gregory Ligozio for their useful discussion and feedback during the research of this paper. Rhydian Windsor is supported by CRUK as part of the EPSRC AIMS CDT (EP/L015897/1).

8. REFERENCES

- [1] M. Krohn, L. S. Braum, J. Sieper, I. Song, A. Weiß, J. Callhoff, C. E. Althoff, B. Hamm, and K. A. Hermann, "Erosions and fatty lesions of sacroiliac joints in patients with axial spondyloarthritis: Evaluation of different mri techniques and two scoring methods," *The Journal of Rheumatology*, vol. 41, no. 3, 2014.
- [2] W. P. Maksymowych, S. S. Dhillon, P. Chiowchanwisawakit, S. J. Pedersen, B. Martinez, M. Østergaard, and R. G. Lambert, "Development and validation of web-based training modules for systematic evaluation of active inflammatory lesions in the spine and sacroiliac joints in spondyloarthritis," *The Journal of Rheumatology Supplement*, vol. 84, pp. 48–57, 2009.
- [3] A. Jamaludin, M. Lootus, T. Kadir, and A. Zisserman, "Automatic intervertebral discs localization and segmentation: A vertebral approach," in *Computational Methods and Clinical Applications for Spine Imaging*, 2015.
- [4] R. Windsor, A. Jamaludin, T. Kadir, and A. Zisserman, "A convolutional approach to vertebrae detection and labelling in whole spine mri," in *Medical Image Computing and Computer Aided Intervention*, 2020.
- [5] R. Tao, W. Liu, and G. Zheng, "Spine-transformers: Vertebra labeling and segmentation in arbitrary field-ofview spine cts via 3d transformers," *MedIA*, vol. 75, 2022.
- [6] A. Jamaludin, T. Kadir, E. Clark, and A. Zisserman, "Predicting scoliosis in dxa scans using intermediate representations," in *Computational Methods and Clini*cal Applications for Spine Imaging, 2018.
- [7] E. Bourigault, A. Jamaludin, T. Kadir, and A. Zisserman, "Scoliosis measurement on dxa scans using a combined deep learning and spinal geometry approach," in *MIDL*, 2022.

- [8] M. C. Faleiros, M H Nogueira-Barbosa, V. F. Dalto, J. R. Ferreira Júnior, A. P. M. Tenório, R. Luppino-Assad, P. Louzada-Junior, R. M. Rangayyan, and P. M. de Azevedo-Marques, "Machine learning techniques for computer-aided classification of active inflammatory sacroiliitis in magnetic resonance imaging," *Advances in Rheumatology*, vol. 60, no. 1, May 2020.
- [9] K. K. Bressem, L. C. Adams, F. Proft, K. A. Hermann, T. Diekhoff, L. Spiller, S. M. Niehues, M. R. Makowski, B. Hamm, M. Protopopov, V. R. Rodriguez, H. Haibel, J. Rademacher, M. Torgutalp, R. G. Lambert, X. Baraliakos, W. P. Maksymowych, J. L. Vahldiek, and D. Poddubnyy, "Deep learning detects changes indicative of axial spondyloarthritis at MRI of sacroiliac joints," *Radiology*, Aug. 2022.
- [10] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger, "Occupancy networks: Learning 3d reconstruction in function space," in *CVPR*, 2019.
- [11] J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, "DeepSDF: Learning continuous signed distance functions for shape representation," in *CVPR*, 2019.
- [12] J. Chibane, T. Alldieck, and G. Pons-Moll, "Implicit functions in feature space for 3d shape reconstruction and completion," in *CVPR*, 2020.
- [13] H. Smith, A. Banerjee, R. Choudhury, and V. Grau, "Automated torso contour extraction from clinical cardiac mr slices for 3d torso reconstruction," in *EMBC*, 2022.
- [14] A. Banerjee, J. Camps, E. Zacur, C. M. Andrews, Y. Rudy, R. P. Choudhury, B. Rodriguez, and V. Grau, "A completely automated pipeline for 3D reconstruction of human heart from 2D cine magnetic resonance slices," *Philos. Trans. R. Soc. A*, vol. 379, 2021.
- [15] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models.," *IJCV*, vol. 1, no. 4, 1988.
- [16] C. Xu and J. Prince, "Snakes, shapes, and gradient vector flow," *IEEE TPAMI*, vol. 7, no. 3, pp. 359–369, 1998.
- [17] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Regressing heatmaps for multiple landmark localization using CNNs," in *MICCAI*, 2016.
- [18] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," in arXiv preprint arXiv:1904.07850, 2019.
- [19] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.
- [20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *ICLR*, 2015.