Robust Triangle-Tree based Registration on Whole Slide Images

Christian Marzahl CMARZAHL@GESTALTDIAGNOSTICS.COM
Gestalt Diagnostics, Spokane, United States

Jason Lunder
Gestalt Diagnostics, Spokane, United States

Brian Napora
Gestalt Diagnostics, Spokane, United States

Editor:

Abstract

The registration of whole slide images (WSIs) provides the basis for aligning corresponding regions across multiple immunohistochemistry (IHC) stains with hematoxylin & eosin (H&E)-stained WSI in digital pathology. To this date, WSI registration is a very challenging task due to different stains, digitization processes, image size, image quality or elastic tissue deformations. We introduce a novel, multi-scale registration approach based on a triangulation tree structure with attention-based key-points (KP) matching methods in the following paper. We validated our method on 100 WSIs from the ACROBAT challenge validation set with four different stainings. Our approach reached a validation score of 149.08 in the ACROBAT challenge.

Keywords: Registration, Microscopy, Pathology

1. Introduction

Highly accurate registration of multiple slides with different types of stain, from the same histology block is a valuable tool for research, and clinical applications. Achieving this goal is very challenging due to the characteristics of digital pathology WSI. These images frequently are very large in file size, with multiple types of stain, digitized with different scanners, and frequently have tissue deformations in the manual preparation along with other artefacts like bubbles, pen markers or tissue tearing. For further details regarding state-of-the-art methods and limitations, please consult the challenge website (Weitz et al.).

Our proposed method extends the quad-tree (QT)-based registration methods proposed by (Marzahl et al., 2021) in two characteristics. First, we replaced the available classical KP extractor Scale Invariant Feature Transform (SIFT) with the deep learning-based method SuperPoint (DeTone et al., 2018) and attention-based KP matching. Second, we replaced the QT with a triangulation-based approach where we performed a Delaunay triangulation (Boris, 1934) based on the found KP and only redefined our registration recursively in areas represented as triangles which contain landmarks to register.

©

1.1 Datasets

The training dataset contains a total of 750 cases, the validation set 100 cases and the test set 302 cases. For the validation and test set, the landmarks to register are manually placed on the IHC slides by the study organizer and have to be automatically transferred to the H&E slides. A detailed description can be accessed at the challenge homepage (Weitz et al.).

2. Methods

Any image registration method aims to find a transformation T between a moving source image I and a fixed target image I'. Our registration approach is based on a tree-structure backbone which recursively divides the WSI into region of interest (RoI) with successively higher resolution levels. For each source/target RoI pair we extract matching KP via deep learning-based methods to estimate a transformation matrix. This results in a piece-wise affine approximation of any non-linear deformation. As the KP detector, we support a variety of extractors like SIFT (Lowe, 2004), Oriented FAST and Rotated BRIEF (ORB) (Rublee et al., 2011) or SuperPoint (DeTone et al., 2018) with matching methods like OpenGlue (Viniavskyi et al., 2022), SuperGlue (Sarlin et al., 2020) or LoFTR (Sun et al., 2021). In the following section, we will describe our method in detail.

2.1 Key-point extraction and matching

The basis for the creation of a tree-based representation of the registration is build by matching KP on the source and target images. The previous publications (Marzahl et al., 2021), showed that classical KP methods such as SIFT or ORB find an insufficient number of matching points on consecutive slices from the same block. Therefore, deep-learning-based KP descriptors like SuperPoint were combined with attention-based matching techniques like SuperGlue, OpenGlue or LoFTR to overcome this limitation. Fine-tuning on the provided pathology training data was not performed, and the published weights of the authors were used. The matching approach with the highest number of matched KPs per source target patch was used at inference time.

2.2 Triangle tree-based registration approach

Tree-based data representations are characterized by their recursive definition, where each segment on the previous level is divided into several sub-segments with a higher spatial resolution. This hierarchical structure makes them ideal candidates to approximate WSIs with their pyramidal approach when creating images at multiple magnification levels. Using this data structure as the backbone, our registration algorithm is recursively applied to higher magnification levels of the source and target WSIs and can be described as follows:

1. In the first step, the target image is rotated in 22.5-degree steps from 0 to 337.5, and the angle with the most matching KPs between the source/target image is taken as the initial rotation angle. We motivate this by to the property of deep-learning-based KP methods to be rotation invariant only up to a certain angle (Pautrat et al., 2020).

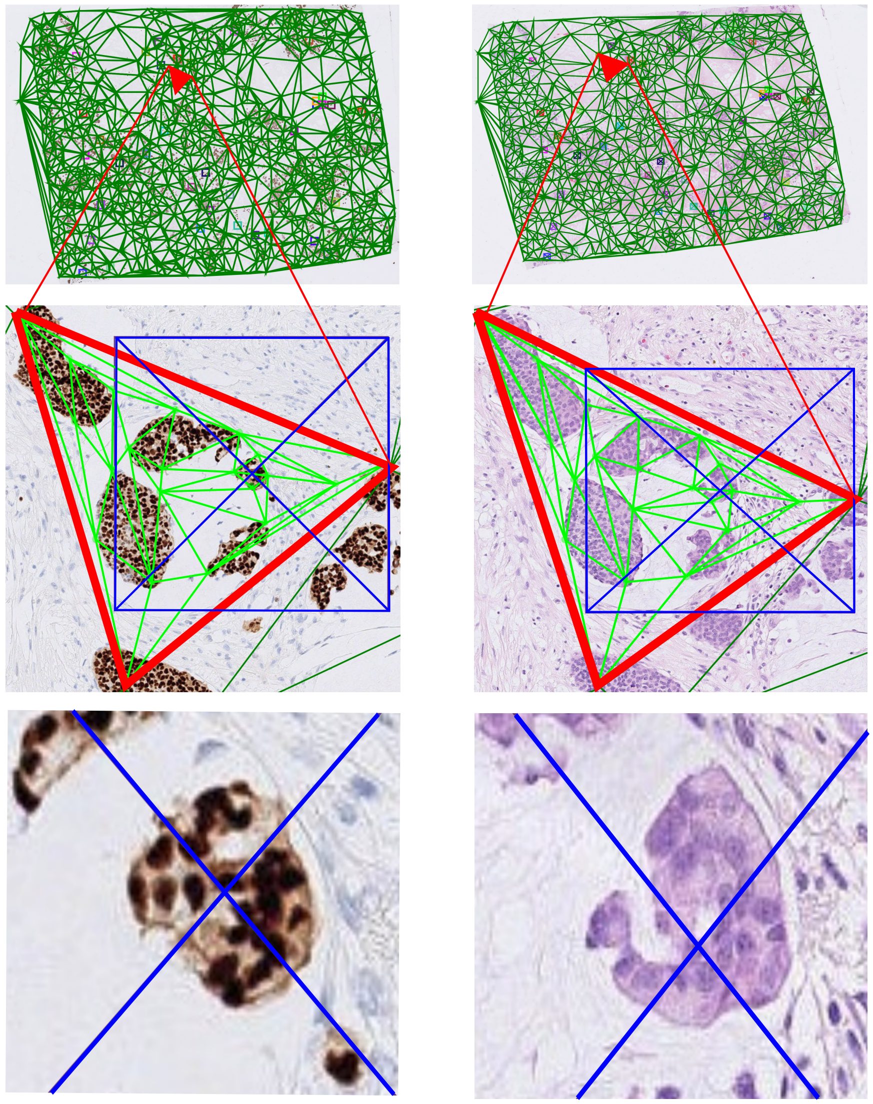

Figure 1: The top row represents the source (left) IHC WSI and target H&E WSI with the Delaunay triangulation (dark green triangles) based on the extracted KP. The second row visualises an example source and target triangle from tree level zero (red) with the next triangulation level (light green). The bottom row depicts the final level of the tree with the landmark (blue box) transferred from the source (IHC) to the target H&E region.

2. We extract a source and a target RoI from the source/target WSI. On tree level zero, this RoI represents the entire WSI, because it consists of a sub-region of the WSI in subsequent levels of the tree.

3. Keypoints are extracted and matched from source and target RoIs using deep-learning-based methods as described in section 2.1. To increase the number of matched KP, multiple patch-sizes are combined with contrast normalizing pre-processing as a type of test time augmentation followed by non-maximum suppression.

4. Delaunay triangulation (OPENCV implementation) is performed between the source KP and mapped to the target KP. Each triangle represents a potential new child note for the current tree level (see Fig. 1). In the next step, we compute the transformation matrix $T \in \mathbb{R}^{2\times 3}$ by computing the affine matrix (OPENCV implementation) with RANSAC (Fischler and Bolles, 1981) filtering.

5. The final step is evaluating if a landmark is located within a triangle. A new recursion level is created if a landmark is present, and the process restarts at step two. This refinement process is repeated until fewer than three matching KP are found, and no affine transformation can be calculated.

To calculate the position of a given landmark in the source image after registration, the final transformation matrix associated with the last tree level is used to transform the point to the target coordinate system.

3. Results

The validation and test set results were created on a NVIDIA GeForce 3080 mobile with 8 GB memory and an Intel i7-11800H with eight cores. On the validation set of the ACROBAT challenge, our approach found a mean of 335 KP (min=0, max=1049, std=191) for level zero, and only failed on the flipped WSI number 61 with zero found KP, reaching a score of 149.08. On the 302 test set WSI pairs, we found a mean of 342 KP (min=26, max=1111, std=168) for tree level zero and were able to register all WSI pairs. The registration time strongly depends on the number of KP found and reached maximum tree level in a mean of 216 seconds (min=33, max=341) for the fifty provided landmarks. Without focusing the registration on the 50 landmarks, the registration time is about 5 hours per WSI to reach the same level of accuracy.

4. Discussion and Conclusion

In this work, we have presented a novel registration method which adapts an existing QT-based registration to generalize on consecutive slides via deep learning-based KP matching methods in combination with test time augmentation. Furthermore, we applied Delaunay triangulation to focus only on RoI, which contain landmarks that are relevant for the challenge and therefore limit computational expense. The proposed method achieved a score of 149.08 on the 100 validation slides and showed its robustness on the test set, finding more than 3 KP matches on all 302 WSI and therefore enabling the automatic registration of all WSI. Besides the positive sides of the algorithm, there are some limitations to consider. The algorithm is designed to transfer a limited number of landmarks between the source and target image. It is computationally costly and time intensive to register complete whole slide images with the same level of accuracy as shown in the results section. Furthermore, the landmarks were placed by human annotators, which marked visually distinguishable regions, which is the same idea behind KP detection algorithms. This shared behaviour was made apparent by the observation that the key point algorithms find many KP very close to the set landmarks.

References

N Boris. Delaunay. sur la sphere vide. *Izvestia Akademia Nauk SSSR, VII Seria, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7:793–800, 1934.

Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Superpoint: Self-supervised interest point detection and description. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 224–236, 2018.

Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

Christian Marzahl, Frauke Wilm, Dressler Franz F., Lars Tharun, Sven Perner, Christof A. Bertram, Christine Kröger, Jörn Voigt, Robert Klopfleisch, Andreas Maier, Marc Aubreville, and Katharina Breininger. Robust quad-tree based registration on whole slide images. In Manfredo Atzori, Nikolay Burlutskiy, Francesco Ciompi, Zhang Li, Fayyaz Minhas, Henning Müller, Tingying Peng, Nasir Rajpoot, Ben Torben-Nielsen, Jeroen van der Laak, Mitko Veta, Yinyin Yuan, and Inti Zlobec, editors, *Proceedings of the MICCAI Workshop on Computational Pathology*, volume 156 of *Proceedings of Machine Learning Research*, pages 181–190. PMLR, 27 Sep 2021. URL `https://proceedings.mlr.press/v156/marzahl21a.html`.

Rémi Pautrat, Viktor Larsson, Martin R. Oswald, and Marc Pollefeys. Online invariance selection for local feature descriptors. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision – ECCV 2020*, pages 707–724, Cham, 2020. Springer International Publishing. ISBN 978-3-030-58536-5.

Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *2011 International conference on computer vision*, pages 2564–2571. Ieee, 2011.

Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. SuperGlue: Learning feature matching with graph neural networks. In *CVPR*, 2020. URL `https://arxiv.org/abs/1911.11763`.

Jiaming Sun, Zehong Shen, Yuang Wang, Hujun Bao, and Xiaowei Zhou. LoFTR: Detector-free local feature matching with transformers. *CVPR*, 2021.

Ostap Viniavskyi, Mariia Dobko, Dmytro Mishkin, and Oles Dobosevych. Openglue: Open source graph neural net based pipeline for image matching, 2022. URL `https://arxiv.org/abs/2204.08870`.

Philippe Weitz, Masi Valkonen, Leslie Solorzano, Johan Hartman, Pekka Ruusuvuori, and Mattias Rantalainen. Acrobat 2022 - grand challenge. URL https://acrobat.grand-challenge.org/.