

Deep Neural Networks for Quality Assurance of Image Registration

S Bannister¹ⁱ, D Page¹ⁱⁱ, T Standen¹, A Dunne¹, J Rawling¹, C J Birch-Sykes¹, M Z Wilson², S Holloway², J McClelland³, Y Peters¹

¹ *University of Manchester*

² *Department of Medical Physics and Biomedical Engineering, UCL, London, UK*

³ *Centre for Medical Image Computing, UCL, UK*

Abstract

Patient-specific quality assurance of image registrations is needed to enable their use in adaptive radiotherapy. An automated method of assessing the quality of a registration between head and neck CT scans was investigated. Ground truth organ contours were propagated to the deformed image from the floating image and then compared with contours on the reference image by calculating Dice similarity. A 2D convolutional neural network was designed to predict Dice coefficients based on the reference image and the displacement vector field of the registration. The network was trained using axial slices of head and neck CT images to predict the metric for the spinal canal, parotid glands and brainstem. The network was able to predict the Dice coefficient for unseen images. For the spinal canal, 95% of predictions were within 0.208 of the true value, with an average absolute difference of 0.0811. For the left parotid gland, 95% were within 0.270 of the true value, with an average absolute difference of 0.0987. This demonstrates that convolutional neural networks can be trained to effectively predict similarity metrics which can be used to assess the quality of an automatically produced registration.

1. Introduction

Image registration has potential clinical applications in adaptive radiotherapy (Veiga et al., 2015), but it must be accurate to ensure anatomy and dose field are mapped to the correct locations for the purposes of plan adaptation or dose accumulation. Currently, methods are available to assess the general performance of registration algorithms (Yeap et al., 2017), but if such algorithms are to be used clinically, manual checking is required for each patient. Previously, some work has been done to automatically provide quality assurance for the registrations of individual patient images. One method propagates a ground truth contour from one image, then back-propagates it onto a second image for validation, requiring contours on two images (Beasley et al., 2016). Another method uses a reverse classifier to predict contours for images for which ground truth contours are available, and the accuracy of these predictions is used to estimate the true accuracy of the registration (Valindria et al., 2017).

In this project, an alternative method for patient-specific quality assurance of image registration was investigated. This method involves designing a convolutional neural network (CNN) that takes the reference image of a registration and the corresponding displacement vector field, which describes the transformation between the reference and floating images,

as inputs. The network predicts similarity metrics between the transformed and original organ contours, producing quantitative measures of registration performance in the absence of labels on one of the images. This would allow for registrations with poor predicted similarity metrics to be flagged for manual review. This project focuses on inter-patient registrations, as more inter-patient data is available to train a network on. This could later be extended to intra-patient registrations, as this would be the scenario in a clinical setting.

2. Methodology

The data for the project was obtained from the Head-Neck Cetuximab dataset of The Cancer Imaging Archive. This data contained three dimensional scans of the head and neck region and, for most patients, contours of the spinal canal, brainstem and left and right parotid glands (Bosch et al., 2015). Approximately 2000 sets of inter-patient deformable registrations were created from an initial 94 patient scans. The corresponding displacement vector fields were also generated and contours were propagated from the reference image to the deformed image. As the project investigated the effectiveness of this method on two dimensional images, all images were divided into two dimensional slices that could be used as inputs for the network.

Table 1: The order and output dimensions of the layers in the CNN architecture used for this project. The convolutional layers had filters of dimension 3×3 and a 1×1 stride, with 32, 64 and 128 filters respectively and the max pooling layers each had 2×2 pooling windows and 2×2 stride. Training was done using stochastic gradient descent.

Layer	Output dimensions
Input	$512 \times 512 \times 4$
2D Convolution	$510 \times 510 \times 32$
Batch Normalisation	$510 \times 510 \times 32$
2D Max Pooling	$255 \times 255 \times 32$
2D Convolution	$253 \times 253 \times 64$
Batch Normalisation	$253 \times 253 \times 64$
2D Max Pooling	$126 \times 126 \times 64$
2D Convolution	$124 \times 124 \times 128$
Batch Normalisation	$124 \times 124 \times 128$
Flatten	1968128
Dense	32
Batch Normalisation	32
Dense	1

The volume-based Dice coefficient was calculated for the deformable and reference contours for the supervision used during training. The technique was investigated for spinal canals, brainstems and parotid glands, chosen due to their close proximity in the neck region, where accurate dosage delivery could greatly reduce radiotherapy-induced complications. The same network architecture, displayed in Table 1, was used for each with a mean squared error loss function and a learning rate of 0.01. This network was purpose

built to predict Dice coefficients and optimized using spinal canal images. The behaviour of the network was investigated using 4×4 black box occlusion maps, illustrating which regions of the input images were most influential to the network output.

3. Results and Discussion

The technique was tested on spinal canals, brainstems and parotid glands. The predictions for the spinal canal and left parotid gland are depicted in Figure 1. For the spinal canal, training was conducted with 3619 slices and validation with 1552. The left parotid involved 793 slices for training and 341 for validation.

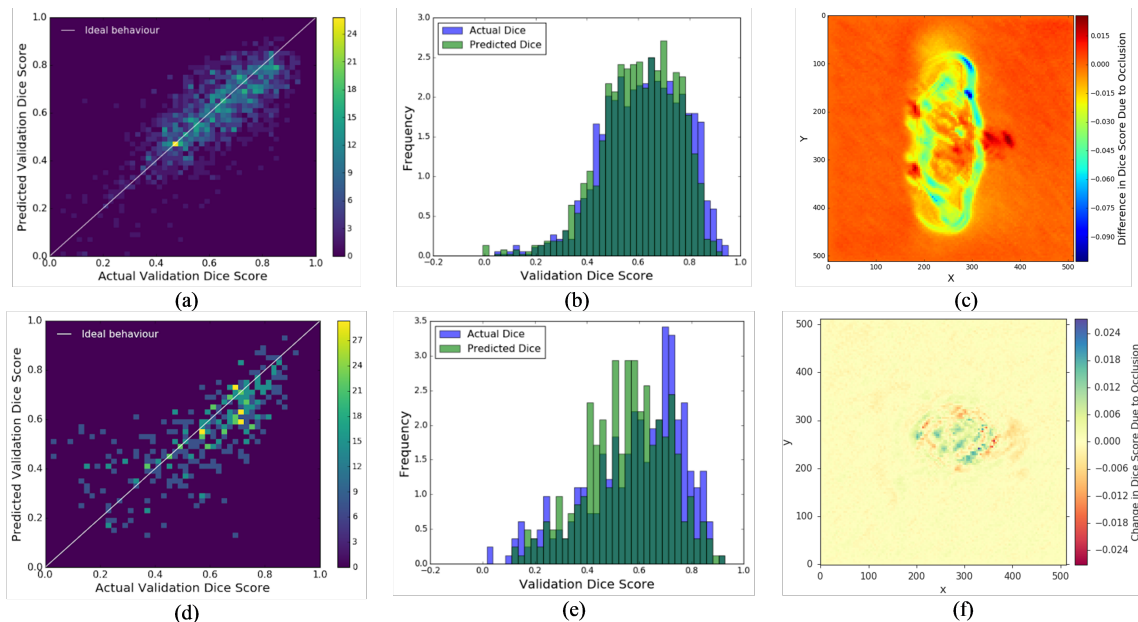


Figure 1: (a) A 2D histogram depicting the relationship between the predicted and true values of the Dice coefficient for the spinal canal validation data with a line of ideal behaviour. (b) A 1D histogram for the same spinal canal data. (c) An example of an occlusion map for the spinal canal. (d),(e) and (f) display the analogous results for the left parotid gland.

The absolute difference between the predicted and true Dice scores was calculated. For the validation data, the 95th percentile of this difference was 0.208 for the spinal canal and 0.270 for the left parotid. The mean difference was 0.0811 for the spinal canal and 0.0987 for the left parotid.

Conclusion: We have produced evidence that accurate predictions of Dice coefficients can be made by CNNs, demonstrating a potential solution to quality assurance concerns for automated image registration of patient scans.

Acknowledgments

We would like to acknowledge Global Challenge Network+ in Advanced Radiotherapy for establishing this project.

References

- W. J. Beasley et al. An automated workflow for patient-specific quality control of contour propagation. *Phys. Med. Biol.*, 61(24):8577–8586, December 2016. ISSN 0031-9155. doi: 10.1088/1361-6560/61/24/8577.
- Bosch et al. Data From Head-Neck_Cetuximab. The Cancer Imaging Archive, 2015. URL <http://doi.org/10.7937/K9/TCIA.2015.7AKGJUPZ>. Date Accessed: 13/12/2018.
- V. V. Valindria et al. Reverse Classification Accuracy: Predicting Segmentation Performance in the Absence of Ground Truth. *IEEE Trans. Med. Imaging*, 36(8):1597–1606, August 2017. ISSN 0278-0062. doi: 10.1109/TMI.2017.2665165.
- C. Veiga et al. Cone-beam computed tomography and deformable registration-based dose of the day calculations for adaptive proton therapy. *International Journal of Particle Therapy*, 2(2):404–414, October 2015. ISSN 2331-5180. doi: 10.14338/IJPT-14-00024.1.
- P L Yeap et al. Automatic contour propagation using deformable image registration to determine delivered dose to spinal cord in head-and-neck cancer radiotherapy. *Phys. Med. Biol.*, 62(15):6062–6073, July 2017.

Notes

ⁱsarah.bannister-2@student.manchester.ac.uk

ⁱⁱdenis.page@student.manchester.ac.uk