Evaluation of DLCExpert™ for Contouring of Thoracic Organs-At-Risk

Peressutti D.[1], Aljabar P.[1], Lustberg T.[2], van Soest J.[2], Dekker A.[2], van Elmpt W.[2], Gooding M.[1]
[1] Mirada Medical Ltd., Oxford, UK, [2] Dept. of Radiation Oncology MAASTRO, Maastricht, ND

Introduction

Auto-contouring has been shown to save time and improve consistency of contouring between clinical staff [1]. However, despite advances in automatic contouring methods, automatically generated contours may still require significant editing before they are considered clinically acceptable.
In recent years, machine learning, and in particular deep learning [2], techniques have become the state-of-the-art for many classification tasks in computer vision and image processing. In this investigation, we assessed the performance of a deep learning-powered system DLCExpert™ for the automatic contouring of thoracic OARs for the treatment of lung cancer.

Methods

For this study, a set of 572 clinical cases, comprising a CT volume image and corresponding clinical contours, were acquired from the Department of Radiation Oncology MAASTRO, Maastricht University Medical Centre, The Netherlands. Evaluation was performed on the lungs, spinal canal, heart, mediastinum envelope and esophagus. The set of clinical cases was randomly divided into a training set (450), cross-validation set (56) and test set (66). The performance of DLCExpert was compared to an atlas-based auto-segmentation (ABAS) method (Mirada Workflow Box 1.4, Mirada Medical Ltd, Oxford, UK), which may be considered the current state-of-the-art for automatic contouring in radiotherapy. The multi-atlas ABAS employed a separate set of 20 atlases, contoured by the same institution, to contour the test images.

A qualitative evaluation of contours was made using a website (www.autocontouring.com) that blinded users (professionals in the field of thoracic radiotherapy) to the source of contours. Users were asked to grade the quality of the contours as if reviewing the contouring of a colleague. Furthermore, a quantitative evaluation against ground-truth clinical contours was performed by computing the Dice Similarity Coefficient (DSC) and root-mean-square distance (RMSD) between both sets of automatically generated contours and the manual clinical contours [3]. To evaluate the real clinical benefit, a time-saving evaluation was also performed on a subset of 20 test cases.

Results

Quantitative, qualitative and time-saving results are presented in Figures 1-3. Results from the performed evaluations show DLCExpert to outperform ABAS for the majority of OARs, except for the heart where performance of the two methods was similar.

Conclusions

This investigation has shown that DLCExpert has significantly outperformed ABAS methods for the automatic contouring of organs-at-risk in lung cancer. Results show improvements in accuracy as well as reductions in the contouring editing time compared to an ABAS method. A qualitative assessment has shown DLCExpert to achieve a contouring acceptance rate comparable to a clinical standard.

References

[1] Sharp, G. *et al.* "Vision 20/20: Perspectives on automated image segmentation for radiotherapy", Medical Physics, 41(5), 2014
[2] Deep Learning in Medical Image Analysis, Mirada white paper
[3] Peressutti, D. *et al.* "Deep Learning Contouring of Thoracic Organs-At-Risk", AAPM 2017

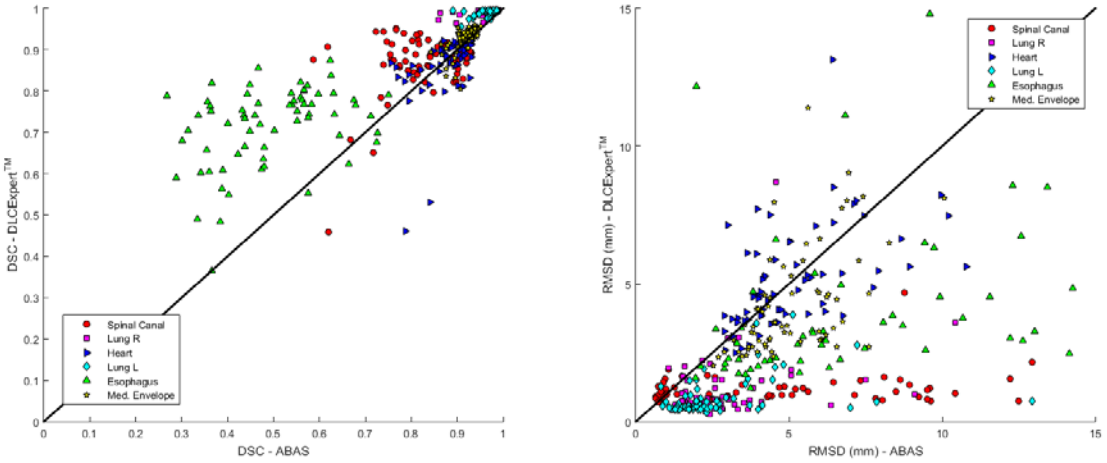

Figure 1. (Left) Comparison of ABAS (x-axis) and DLCExpert (y-axis) DSC values for all OARs. DLCExpert outperforms ABAS if the point lies above the bisector line. (Right) Comparison of ABAS (x-axis) and DLCExpert (y-axis) RMSD in mm for all OARs. DLCExpert outperforms ABAS if the point lies below the bisector line.

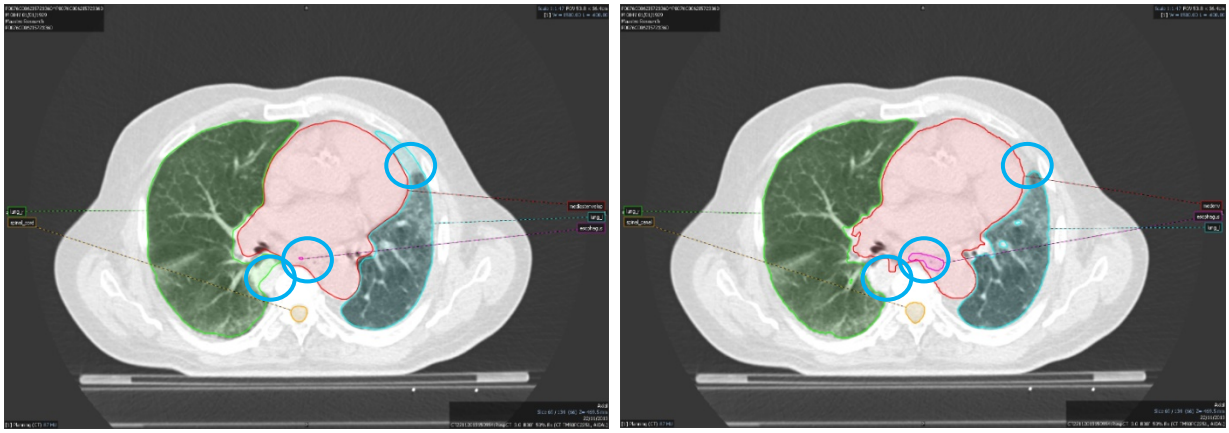

Figure 2. Visual comparison of ABAS (left) and DLCExpert™ (right) contours. Highlighted the regions where DLCExpert™ corrects for ABAS inaccuracies.

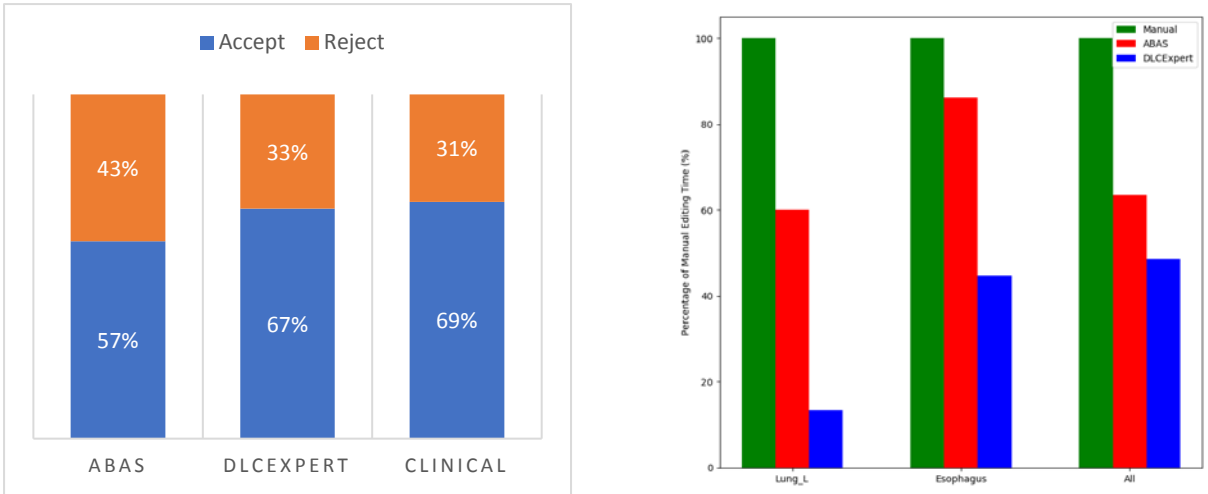

Figure 3. (Left) Results of the qualitative assessment through blinded test. DLCExpert shows a percentage of accepted contours (67%) closer to the clinical acceptance rate (69%) compared to ABAS (57%). Results are reported for 550 evaluated contours. (Right) Percentage of median manual editing time compared to manual contouring (green) is reported for ABAS (red) and DLCExpert (blue) for the left lung, esophagus and the aggregated results for all OARs. DLCExpert requires significantly less editing time than ABAS.