



Universiteit
Leiden
The Netherlands

Automated quality assurance of deep learning contours in head-and-neck radiotherapy

Mody, P.P.

Citation

Mody, P. P. (2026, January 22). *Automated quality assurance of deep learning contours in head-and-neck radiotherapy*. Retrieved from <https://hdl.handle.net/1887/4287843>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4287843>

Note: To cite this publication please use the final published version (if applicable).

Automated Quality Assurance of Deep Learning Contours in Head-and-Neck Radiotherapy

Prerak Mody

Colophon

About the cover:

The cover page was designed by Prerak Mody. On the front cover, you see an ensemble of contours in the middle represented by swirling strokes. The individual contours are depicted with various colors and represent the central theme of this thesis i.e., contouring of regions of interest within medical scans. These contours are made in collaboration by two parties: a human (clinician) and an AI. The AI (i.e., deep learning) has been anthropomorphized and thus depicted via a human-like hand with a mesh design. The human and AI hands are collaborating on the contour which is the motivating factor for this thesis.

This inside cover envisions the future of radiotherapy, showcasing a clinician (in white coat) editing AI-generated contours of the head-and-neck area in real-time. As the patient is positioned on the linear accelerator gantry, the clinician uses automated quality assurance tools to quickly adjust the treatment plan based on the patient's latest anatomical scans, ensuring highly precise and adaptive treatment delivery. A technician (in blue overalls) helps the patient position themselves on the gantry.

Google Gemini, a visual language model (VLM) was used to assist with these designs.

Automated Quality Assurance
of Deep Learning Contours
in Head-and-Neck Radiotherapy
Prerak Mody

ISBN: 978-94-6522-967-6

Thesis layout & cover designed by Prerak Mody
Printed by Ridderprint, the Netherlands

© 2026 Prerak Mody, Leiden, the Netherlands

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

Automated Quality Assurance of Deep Learning Contours in Head-and-Neck Radiotherapy

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van (waarnemend) rector magnificus,
volgens besluit van het college voor promoties
te verdedigen op donderdag 22 januari 2026
klokke 10:00 uur

door

Prerak Mody
geboren te Mumbai, Maharashtra, India
in 1993

Promotor: Prof. dr. ir. M. Staring
Prof. dr. ir. B.P.F. Lelieveldt

Leden promotiecommissie: Prof. dr. C.R.N. Rasch
Prof. dr. A. Mukhopadhyay
Technical University, Darmstadt
Prof. dr. J.P.W. Pluim
Eindhoven University of Technology, Eindhoven
Prof. dr. ir. C.A.T. van den Berg
University Medical Center, Utrecht

The research in this thesis was performed at the Division of Image Processing (LKEB),
Department of Radiology of Leiden University Medical Center, The Netherlands.

The research in this thesis was funded by Varian, a Siemens Healthineers Company, through
the HollandPTC-Varian Consortium (grant id 2019022)

The printing of this thesis was funded by:
Medis Medical Imaging
Library of Leiden University
Bontius Stichting/LUMC ResearchFoundation

Contents

List of abbreviations	v
1 Introduction	1
1.1 Head-And-Neck Anatomy	1
1.2 Radiotherapy Workflow and Automation Opportunities	2
1.3 Deep Learning for Auto-Contouring and its Limitations	4
1.4 Quality Assessment of Contour Automation	4
1.5 Thesis outline	6
2 Large-scale dose evaluation of deep learning organ contours in head-and-neck radiotherapy by leveraging existing plans	9
2.1 Introduction	11
2.2 Materials and methods	11
2.2.1 Data acquisition	11
2.2.2 Automated Contours	12
2.2.3 Treatment Planning Protocol	13
2.2.4 Automated Treatment Planning	13
2.2.5 Geometric Evaluation	14
2.2.6 Dose and NTCP Evaluation	14
2.3 Results	16
2.3.1 Geometric evaluation	16
2.3.2 Dose evaluation	16
2.4 Discussion	18
2.5 Acknowledgement	22
2.6 Appendix	22
2.6.1 Data Acquisition	22
2.6.2 Automated Contours	22
2.6.3 Automated Planning	24
2.6.4 Organ Dose Metrics	27

2.6.5	NTCP	29
2.6.6	Visual Results	30
3	Comparing Bayesian Models for Organ Contouring in Head and Neck Radiotherapy	31
3.1	Introduction	33
3.2	Method	34
3.2.1	Data	34
3.2.2	Neural Architecture	34
3.2.3	Training and Inference	35
3.2.4	Uncertainty	36
3.2.5	Evaluation	38
3.3	Results	39
3.3.1	Volumetric Performance	39
3.3.2	Expected Calibration Error	39
3.3.3	Region - Accuracy vs Uncertainty	39
3.4	Discussion and conclusion	41
3.5	Acknowledgements	44
4	Improving Uncertainty-Error Correspondence in Deep Bayesian Medical Image Segmentation	45
4.1	Introduction	47
4.2	Related Works	49
4.2.1	Epistemic and aleatoric uncertainty	49
4.2.2	Uncertainty use during training	50
4.2.3	Model calibration	50
4.3	Methods	52
4.3.1	Neural Architecture	52
4.3.2	Training Objectives	52
4.3.3	Evaluation	55
4.4	Experiments and Results	56
4.4.1	Datasets	56
4.4.2	Experimental Settings	56
4.4.3	Results	57
4.5	Discussion	60
4.5.1	Discriminative and Calibrative Performance	60
4.5.2	Uncertainty-Error Correspondence Performance	60
4.5.3	Future Work	64
4.6	Conclusion	64

4.7	Acknowledgement	65
4.8	Appendix	66
4.8.1	Segmentation "Failures" and "Errors"	66
4.8.2	Weightage of AvU loss	66
4.8.3	Hyperparameter selection	67
4.8.4	Visual Results	74
4.8.5	Head-And-Neck CT	74
4.8.6	Prostate MR	74
4.8.7	BayesH model	75
5	Manual Brush vs AI Pencil: Evaluating tools for auto-contour refinement of head-and-neck tumors on CT+PET scans	77
5.1	Introduction	79
5.2	Materials and methods	79
5.2.1	Dataset	81
5.2.2	Auto-contour and contour-refinement model training	81
5.2.3	Model (auto-contour and contour-refinement) validation	81
5.2.4	Web-based tool	82
5.2.5	User Cohort	82
5.3	Results	82
5.4	Discussion	85
5.4.1	Time for auto-contour refinement	86
5.4.2	Contour Consistency	86
5.4.3	Tooling	87
5.4.4	Future Work	87
5.5	Conclusion	87
5.6	Acknowledgement	87
5.7	Appendix	88
5.7.1	Dataset	88
5.7.2	Auto-contouring and contour-refinement models	88
5.7.3	Web interface	88
5.7.4	Additional results on user effort	92
6	Summary, discussion and future work	95
6.1	Thesis Summary	95
6.2	Chapter Recapitulations	95
6.2.1	Chapter 2	95
6.2.2	Chapter 3	96
6.2.3	Chapter 4	96

6.2.4	Chapter 5	97
6.3	Discussion and future work	97
6.4	General conclusions	99
7	Samenvatting, discussie en toekomstig werk	101
7.1	Samenvatting van de dissertatie	101
7.2	Hoofdstuk Samenvattingen	101
7.2.1	Hoofdstuk 2	101
7.2.2	Hoofdstuk 3	102
7.2.3	Hoofdstuk 4	102
7.2.4	Hoofdstuk 5	102
7.3	Discussie en toekomstig werk	103
7.4	Algemene conclusies	105
	References	107
	List of publications	123
	Acknowledgements	125
	Curriculum Vitae	127

List of abbreviations

QA	quality assessment
CT	computed tomography
PET	positron emission tomography
HN	head-and-neck
RT	radiotherapy
OAR	organ at risk
DL	deep learning
CTV	clinical target volume
RoI	region of interest
NTCP	normal tissue complication probability
AI	artificial intelligence
AvU	accuracy-vs-uncertainty
ECE	expected calibration error
R-AvU	region-based accuracy-vs-uncertainty
CE	cross entropy
PTV	planning target volume
DICE	Sørensen–Dice coefficient
D_{mean}	mean dose
D0.03cc	dose to 0.03 cubic centimeters of tissue
EUD	equivalent uniform dose
MaxEUD	maximum equivalent uniform dose
OOD	out of distribution