The ACPSEM Medical Image Registration Special Interest Group (MIRSIG) Online Webinars

The **current seminar** 1200, Tue 10th November 2020, will be chaired by Laurel Schmidt

- Talk 1: **Evaluating DLCExpert: Mirada's deep learning contouring model"**

Presented by Eddie Gibbons (RT)

Webinar activities!!

-Use the "Q&A" to ask questions!

Live Poll!

Poll information will be used to confirm CPD, so it is important to participate!

Post webinar survey!

Please answer survey when email is sent

Seminar material available online!

Please see https://www.acpsem.org.au/About-the-College/Special-Interest-Groups/MIRSIG

Be more involved!

1. MIRSIG welcomes professions from all disciplines, including radiation therapists and radiation oncologists
2. Sign up to the MIRSIG mailing list (https://www.acpsem.org.au/Home , click myACPSEM, click speciality groups, tick MIRSIG)
3. Join MIRSIG as a member, email mirsig@acpsem.org.au

Evaluating DLC*Expert*:
Mirada's Deep Learning Contouring Model

Eddie Gibbons
Radiation Therapist
Deformable Registration Project Officer
Mid North Coast Cancer Institute
Port Macquarie

Health
Mid North Coast
Local Health District

MIRADA The Imaging Software People

DLCExpert™ Deep Learning Contouring

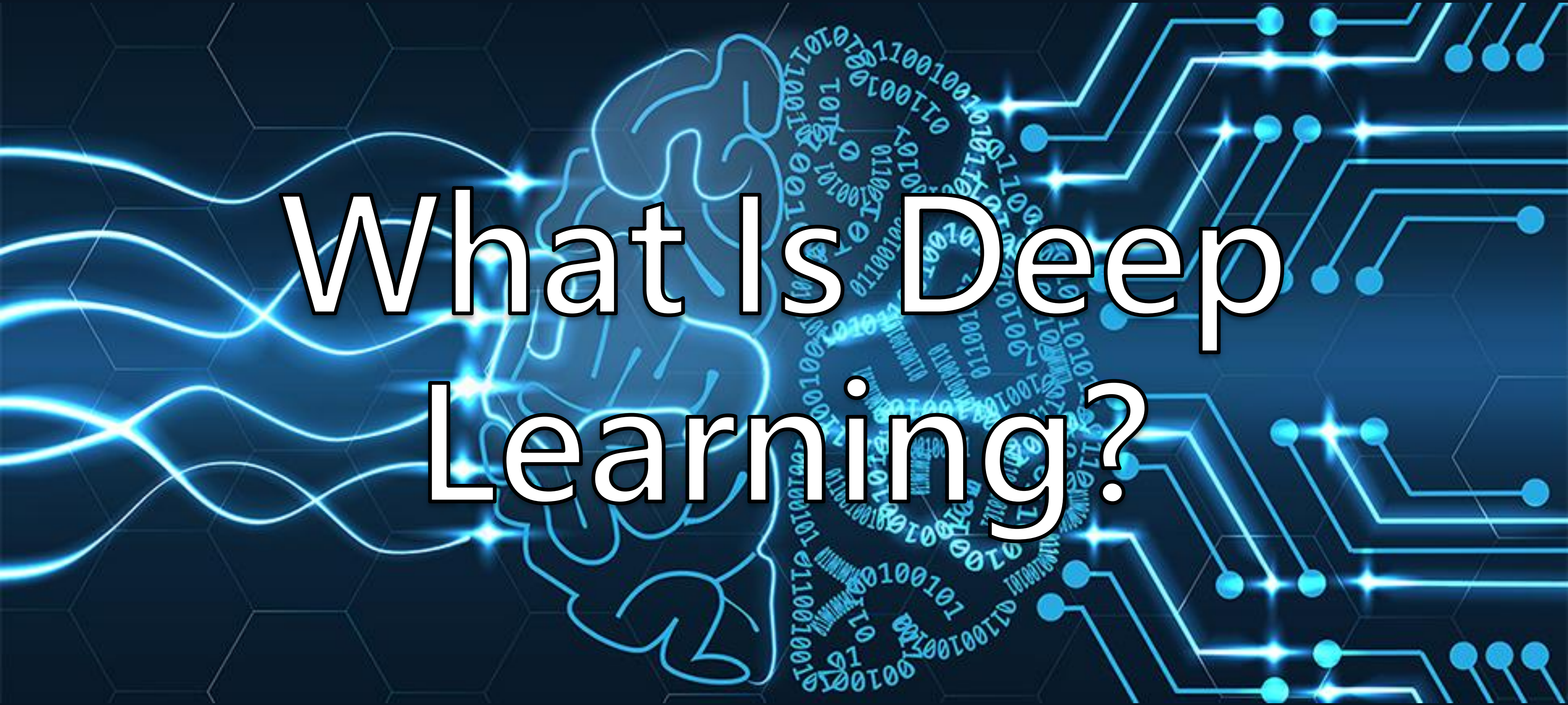
MIRSIG

Aboriginal artwork produced by Gumbaynggir artist Brentyn Lugnan.

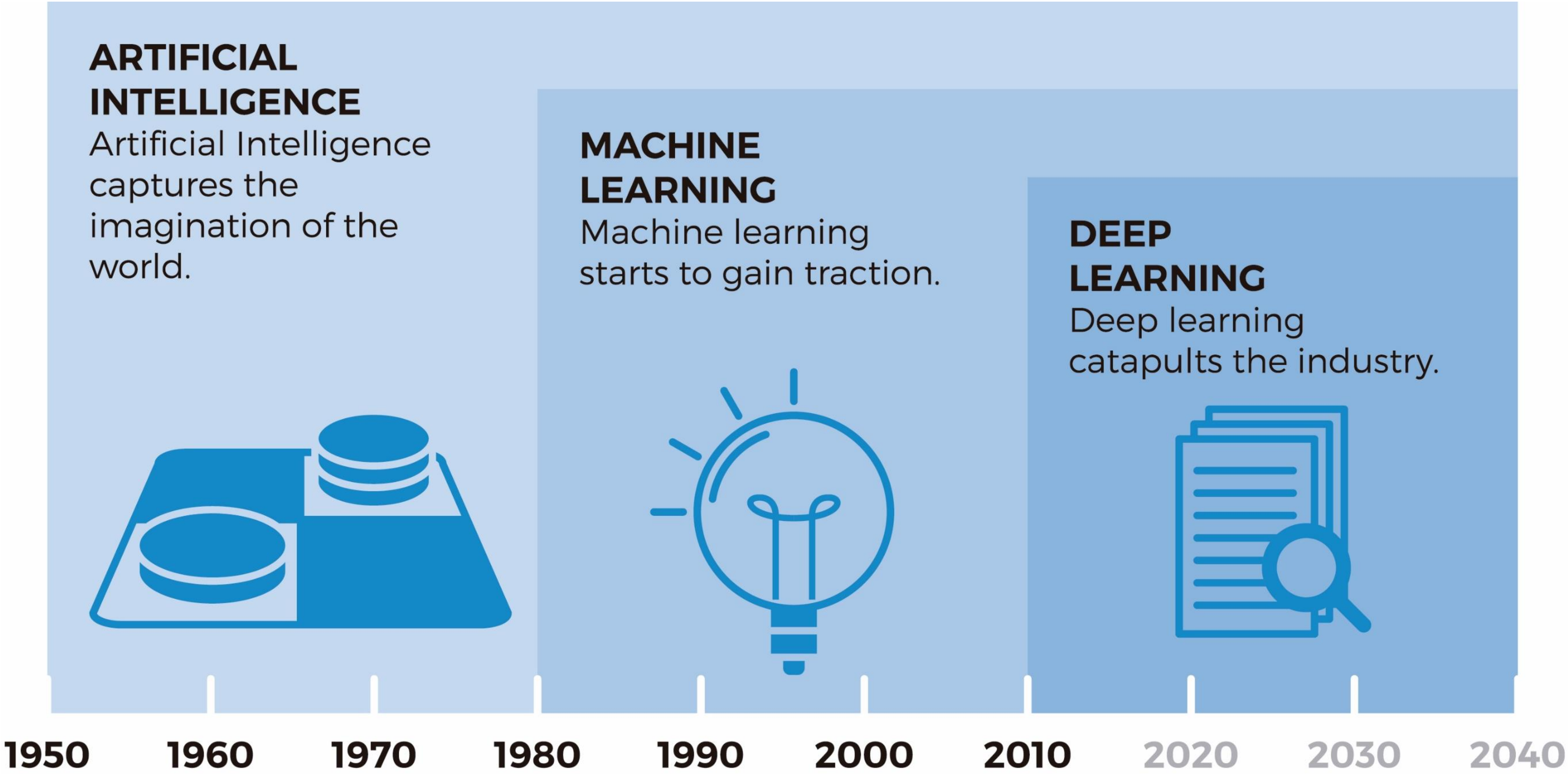
Learning Objectives

- Understand the operation of deep learning contour algorithms
- Identify the clinical benefits and limitations associated with auto-contouring
- Critique DLC*Expert* against current atlas-based contouring methods
- Describe strategies to incorporate deep learning models into a clinical workflow

What Is Deep Learning?

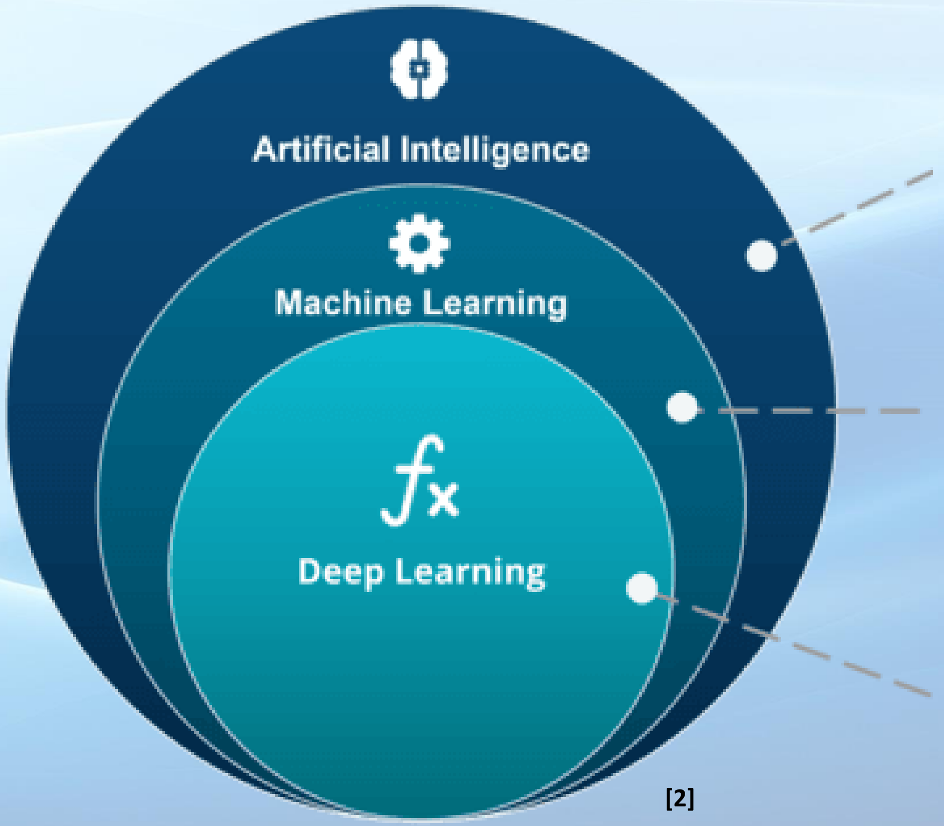


What Is Deep Learning?



[1]

How Does AI Work?

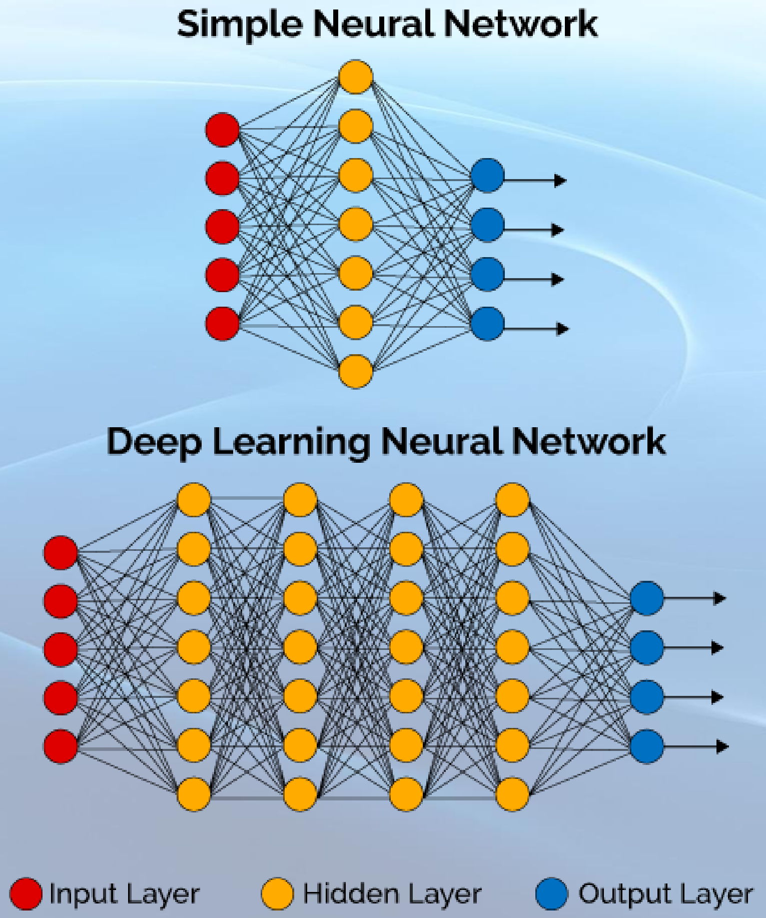


[2]

ARTIFICIAL INTELLIGENCE
A technique which enables computers to mimic human behaviour

MACHINE LEARNING
Subset of AI which uses statistical methods to enable machines to improve with experience

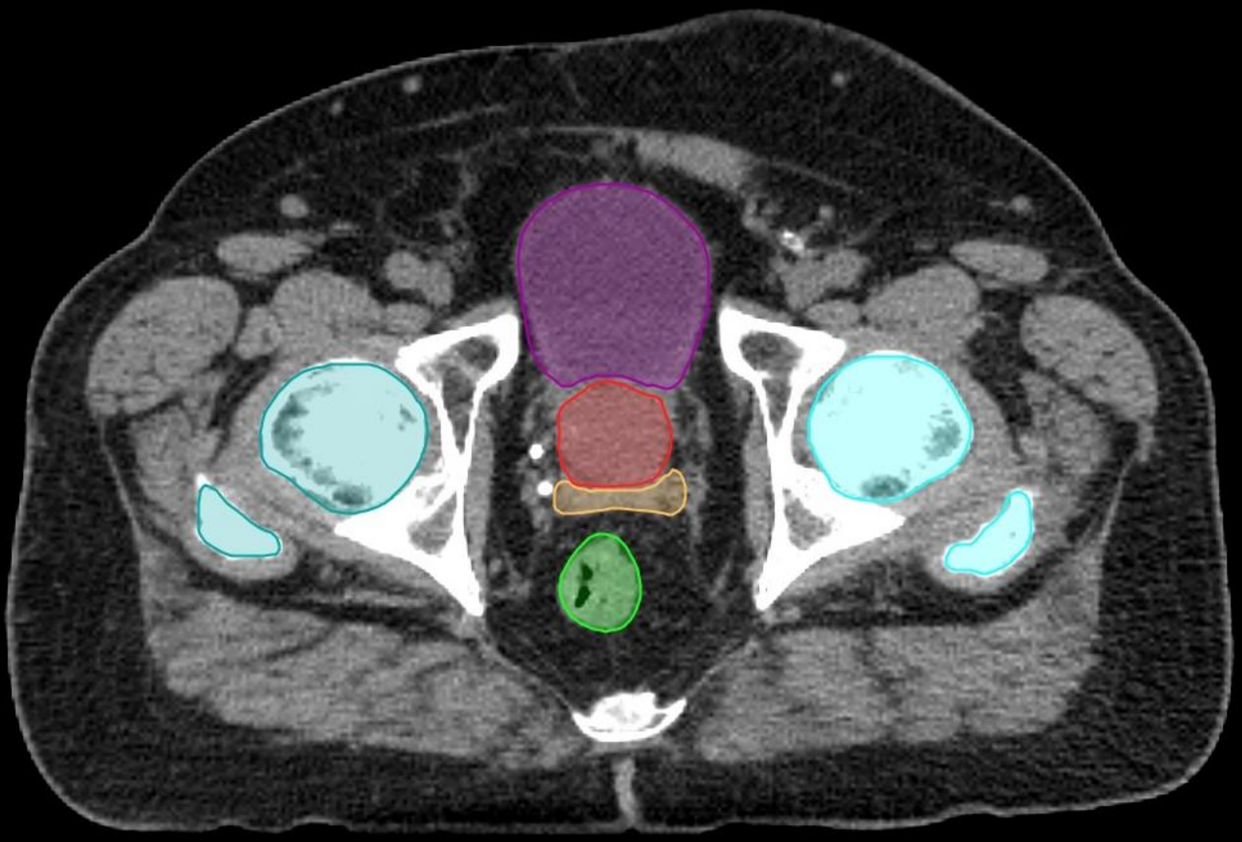
DEEP LEARNING
Subset of machine learning which makes the computation of multi-layer neural networks feasible. Independent predictions can be made without human input



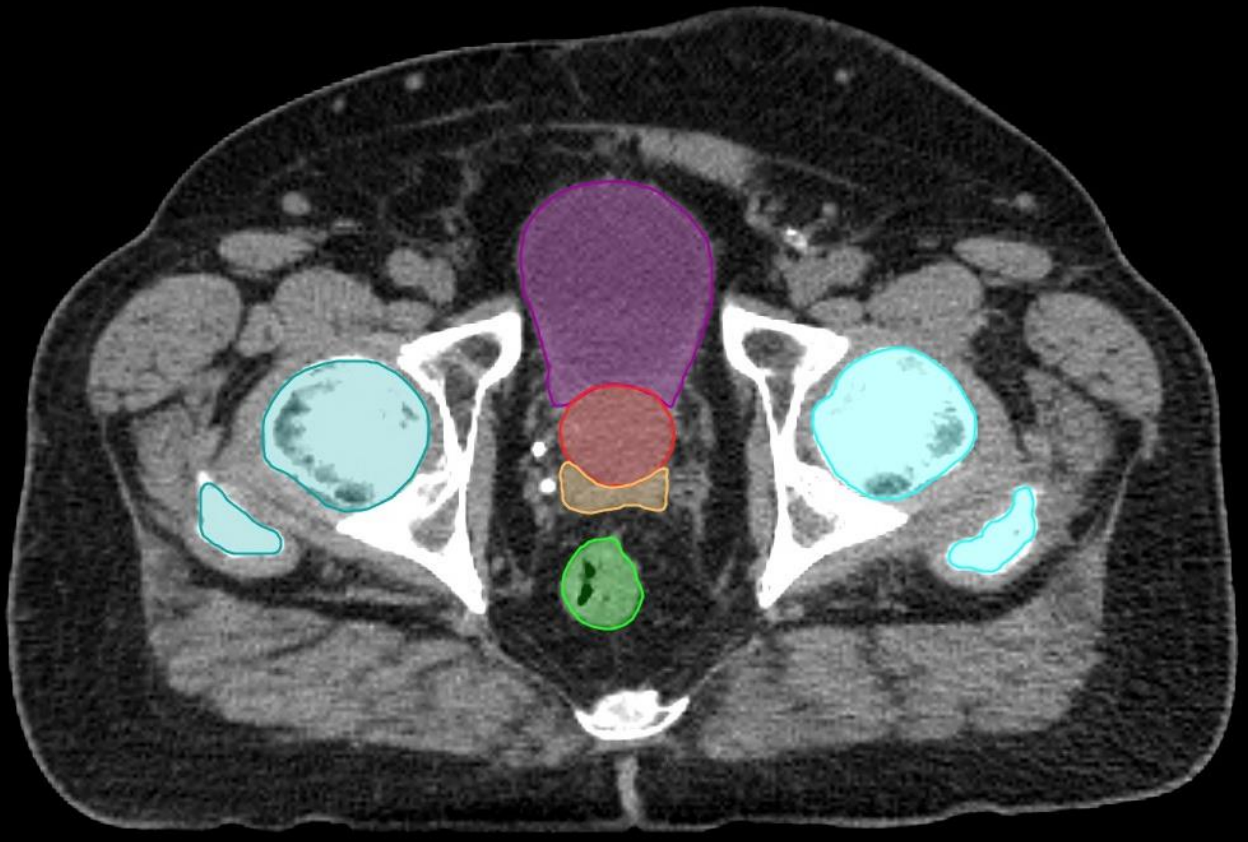
[3]

Health
Mid North Coast
Local Health District

Auto-Contouring with Deep Learning



Deep Learning Drawn



Human Drawn

Health
Mid North Coast
Local Health District

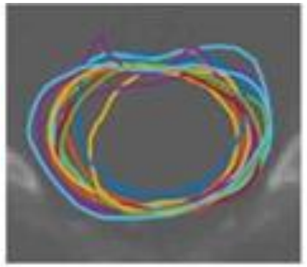
Contour List: Prostate Seminal Vesicles Bladder Rectum L Femoral Head R Femoral Head

Clinical Benefits of Auto-Contouring

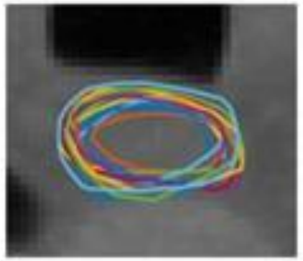
- Improved Contour Consistency
- Time Saving Efficiencies
- Reduced Inter/Intra Observer Variation

Inter-Observer Variation

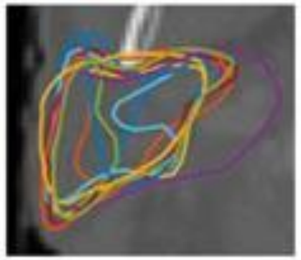
Brainstem



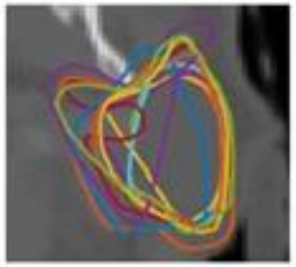
Oesophagus

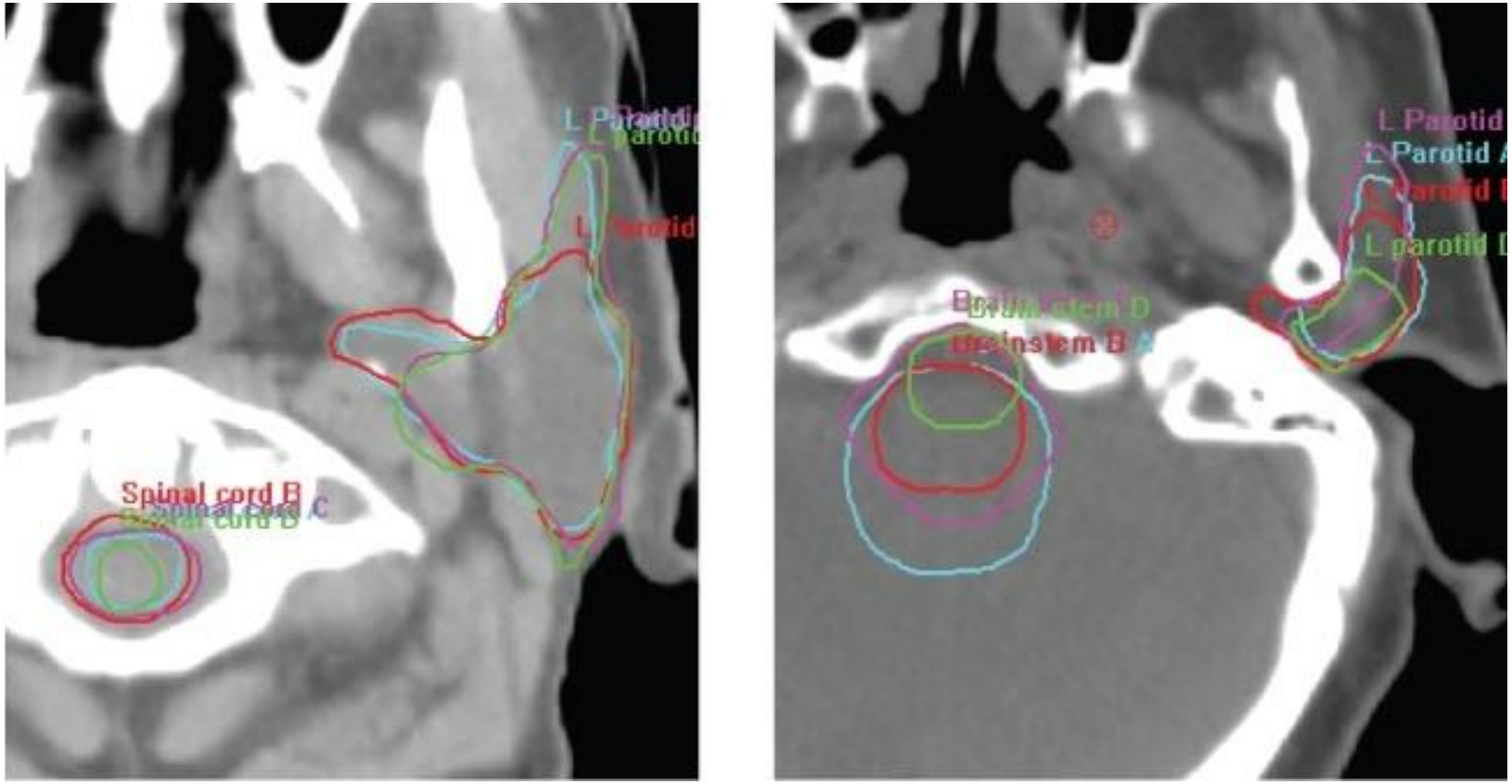


Right Parotid



Left Parotid





[4] Mukesh, M. et al. (2012). "Interobserver variation in clinical target volume and organs at risk segmentation: Can segmentation protocols help?" The British Journal of Radiology, 85, 530-536

Clinical Limitations of Auto-Contouring

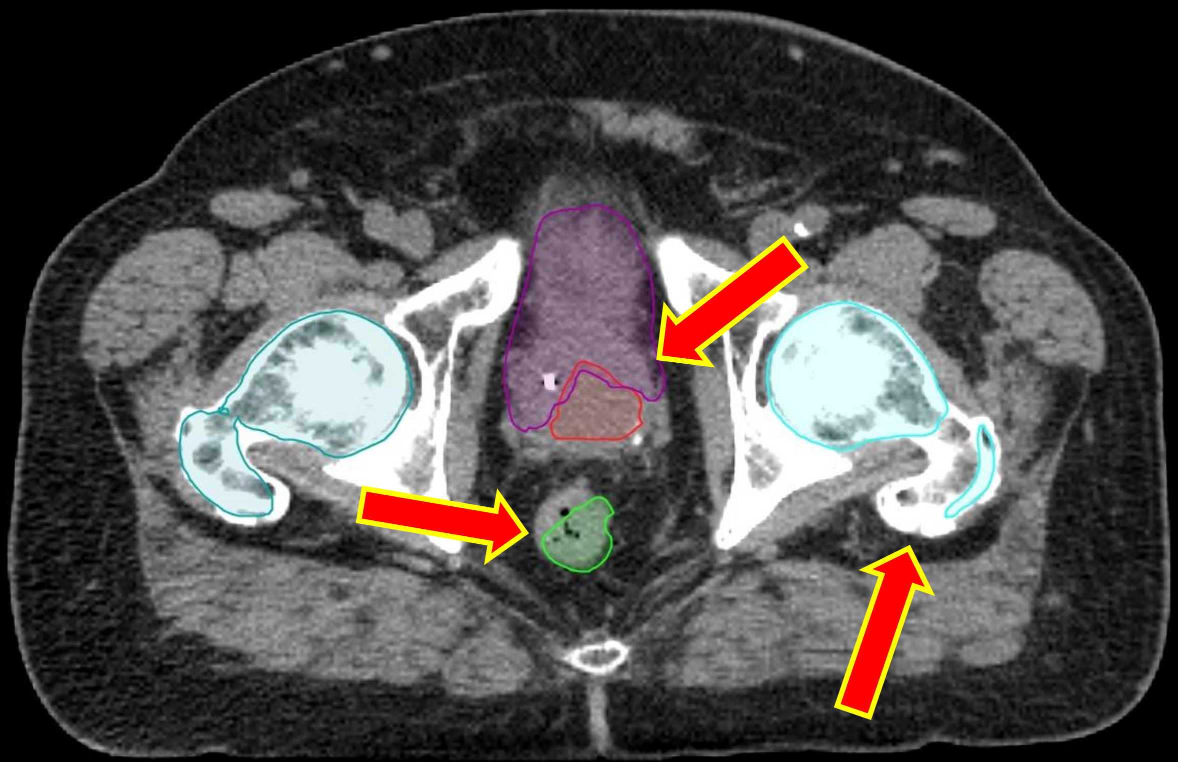
- Subject to Input Training Data
- Accuracy Cannot Consistently Match Human Performance
- Can Struggle to Adapt to Non-Standard Situations

Atlas vs Deep Learning

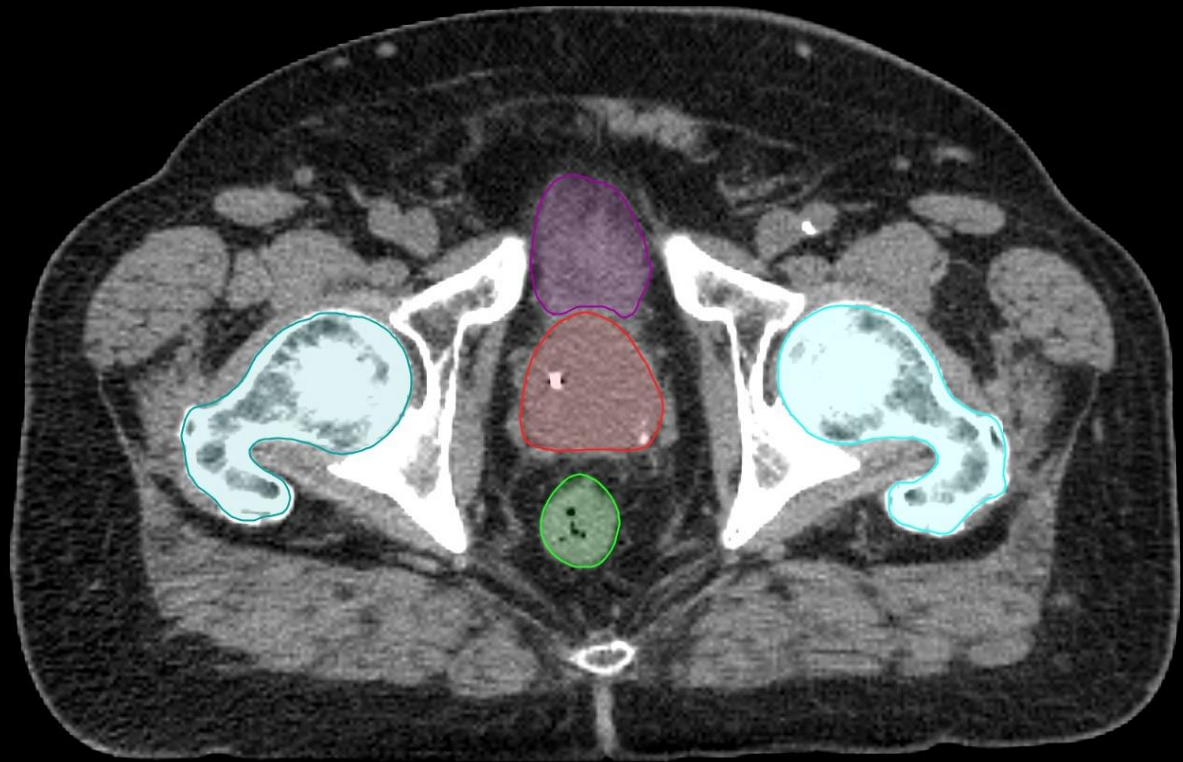
	Multi-Atlas	Deep Learning
Degrees of Freedom	$\sim 1\times10^5$	$\sim 1\times10^6$-1×10^8
Data Quantity	10-20 curated datasets	Hundreds or thousands of curated datasets
Reference Cases	Can only account for a limited number of scenarios	Better at accounting for a wider range of scenarios
Image Registration	Reliant on image registration	No image registration required
Data Library Required	Yes	No
Active Learning	No	No

[5] Aljabar P, Gooding M. [Mirada White Paper]. (2017). “The cutting edge: Delineating contours with Deep Learning”

Atlas vs Deep Learning



Multi-Atlas



Deep Learning

Contour List: Prostate Bladder Rectum L Femoral Head R Femoral Head

DLC*Expert*™

Deep Learning Contouring

DLC*Expert* is Mirada's deep learning contouring solution

Released 2018 – First commercially available AI based auto-contouring application

Mirada claim that DLC generates structures that are of similar clinical acceptance to human drawn contours

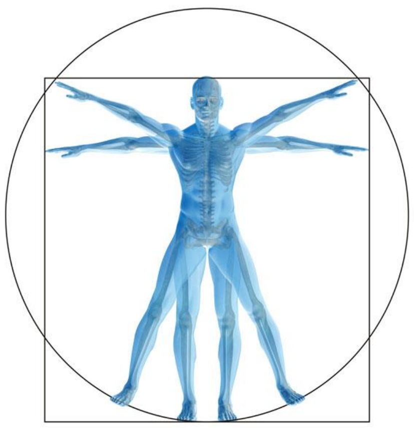
[6] Albano, A. [Press Release]. "Mirada releases DLC*Expert* – First commercially available Artificial Intelligence (AI) autocontouring software for radiation oncology", Mirada-Medical.com, 20/02/2018

DLC Expert™

Deep Learning Contouring

DLC Models:

- Head & Neck
- Thorax
- Prostate
- Supine Breast



	Head & Neck	Thorax	Prostate	Breast
# of Training Datasets	698	576	437	361
Data Type	Supine, non-contrast CT	Supine, non-contrast CT	Supine, non-contrast CT	Supine, non-contrast CT
Pixel Spacing	0.938 (480mm FOV)	0.98 (500mm FOV)	0.98 (500mm FOV)	0.98 (500mm FOV)

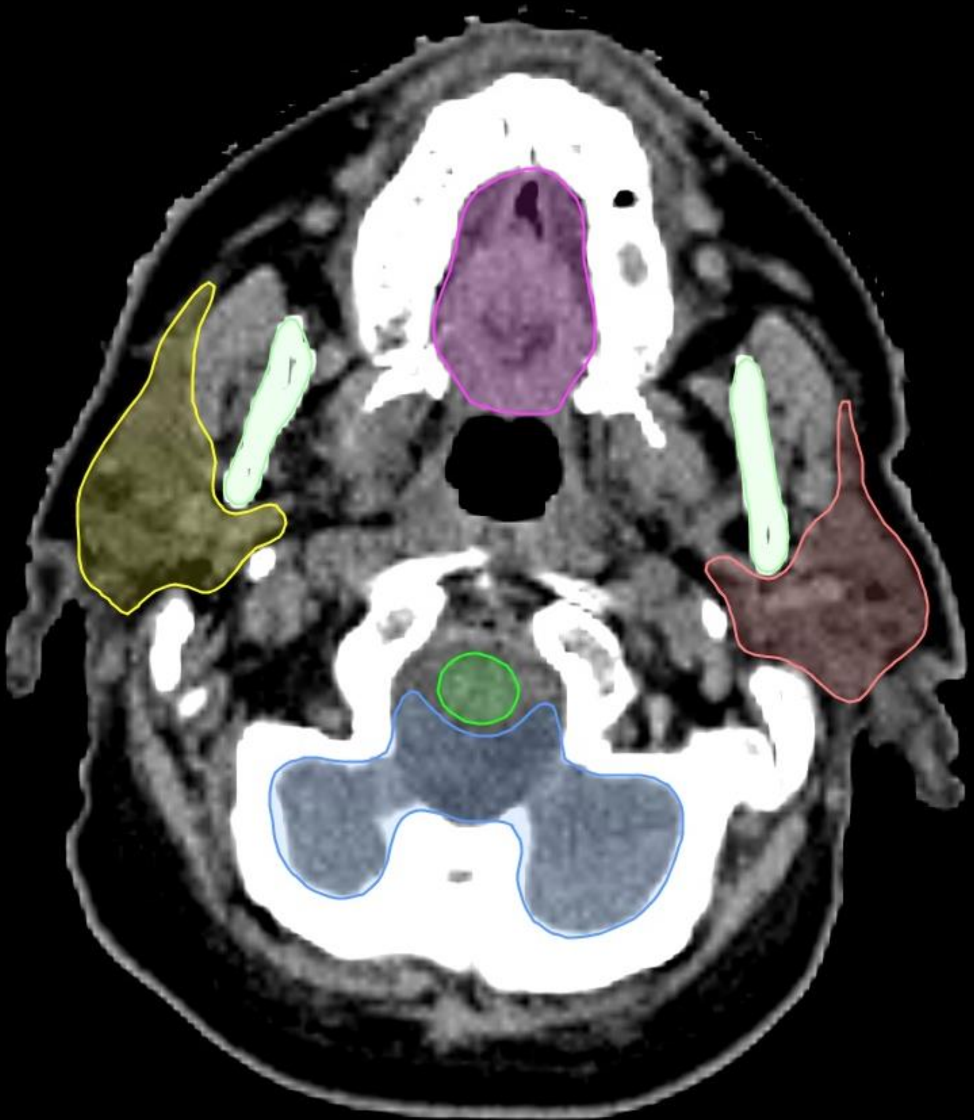
Head & Neck Model

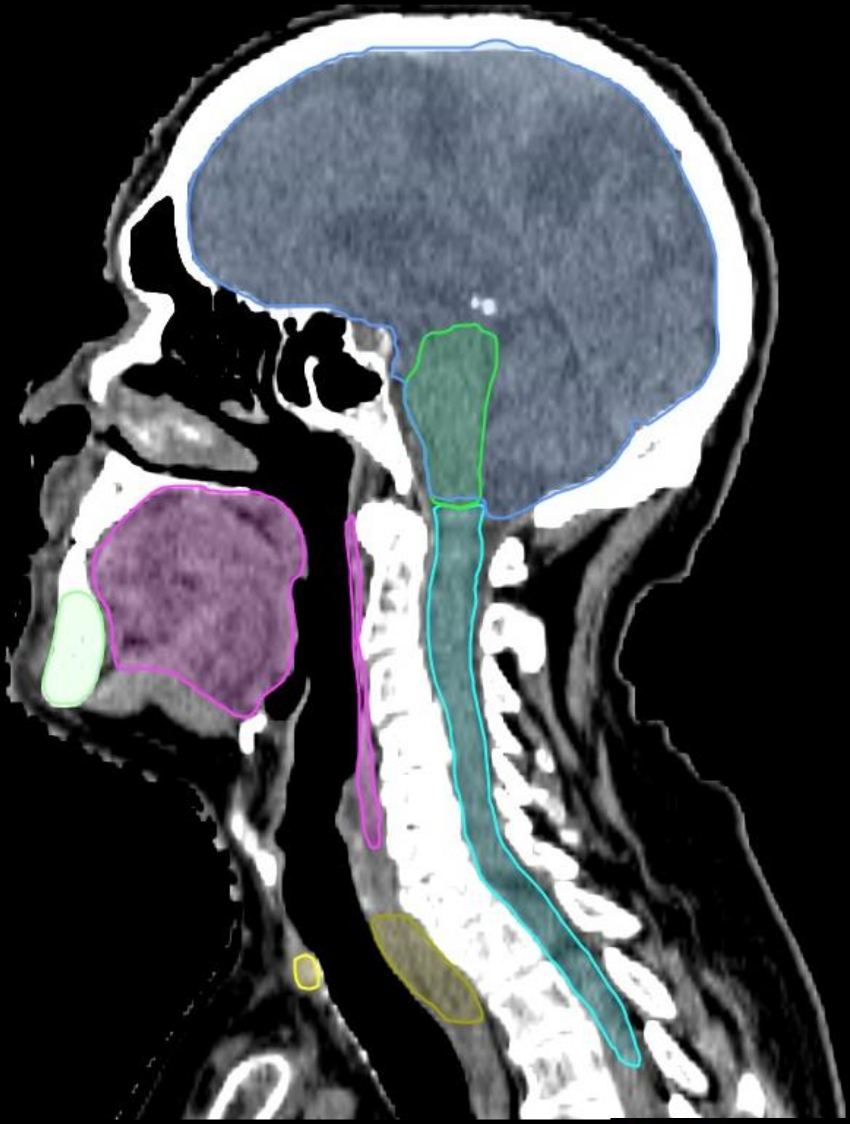
Contour List:

Brain
L Parotid
R Parotid
Oral Cavity
Spinal Cord
Brainstem
Mandible
Thyroid
Pharynx Constrictor
Oesophagus

Not Pictured:

L Submandibular
R Submandibular
L Carotid
R Carotid
L Arytenoid
R Arytenoid
Cerebellum
Cerebrum
L Buccal Mucosa
R Buccal Mucosa
Cricoid Cartilage
Glottis
Supra Glottis





Health
Mid North Coast
Local Health District

DLC Expert™
Deep Learning Contouring

Thorax Model

Contour List:

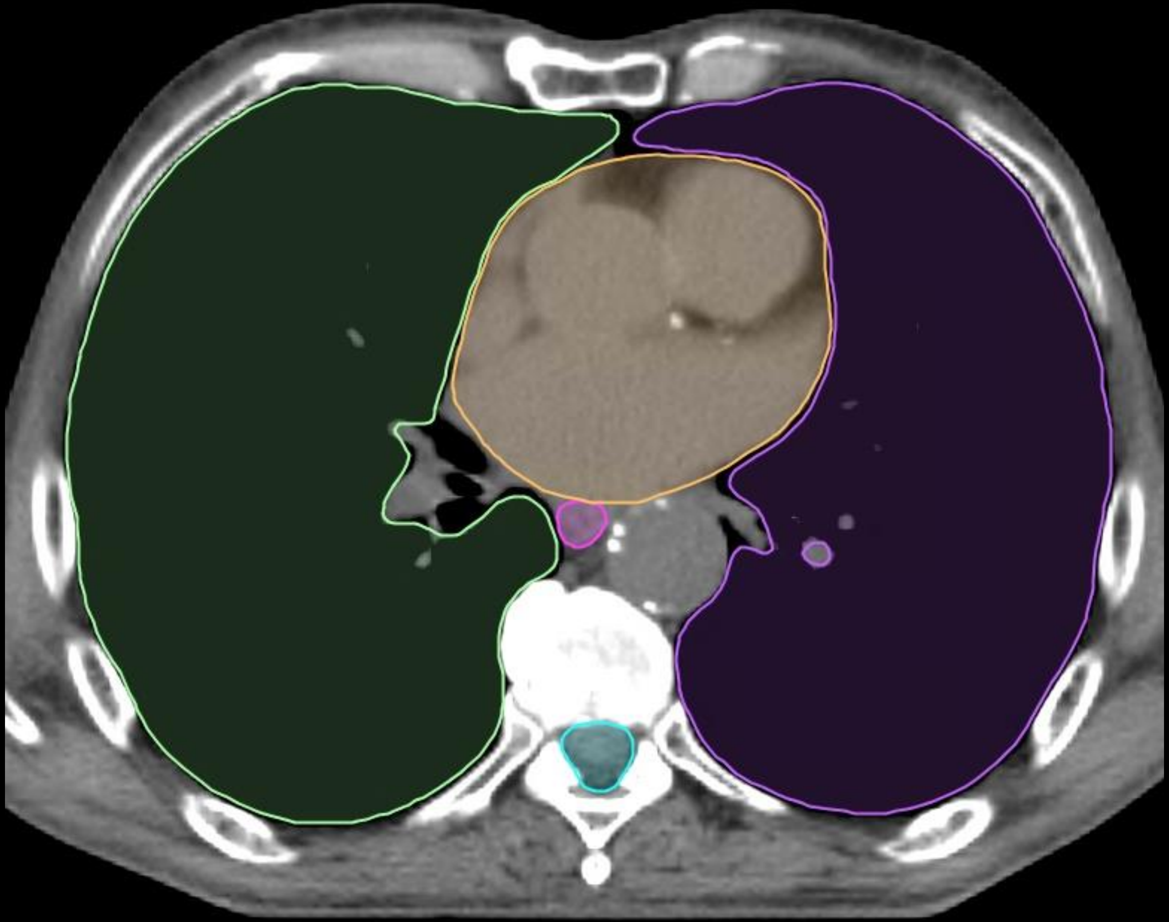
L Lung

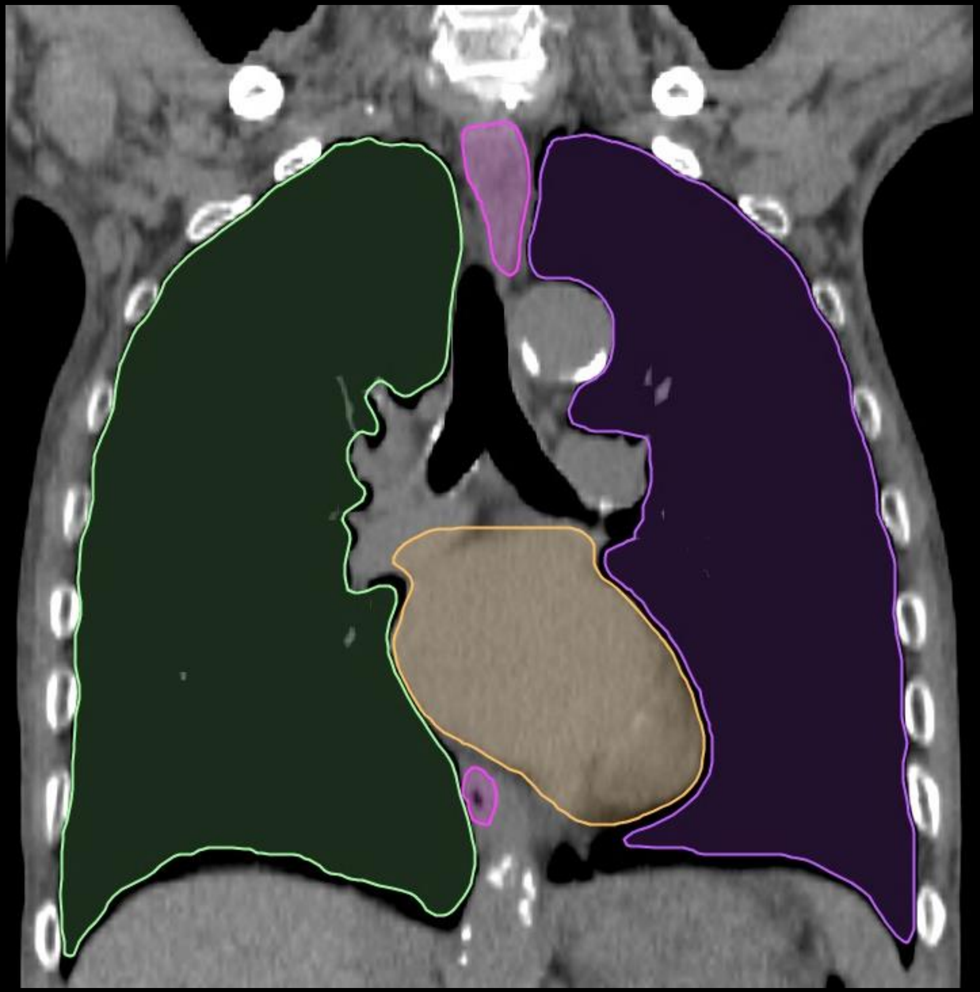
R Lung

Heart

Spinal Canal

Oesophagus





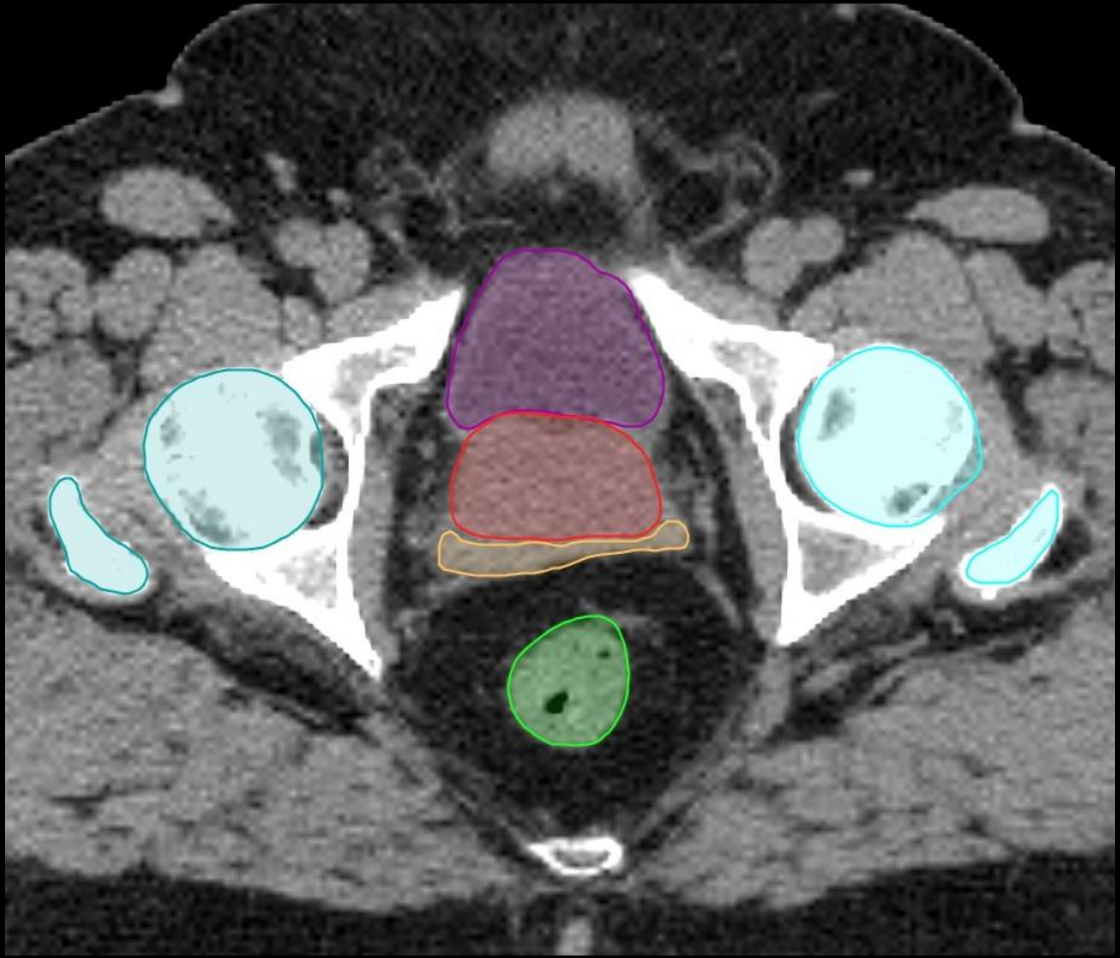
Health
Mid North Coast
Local Health District

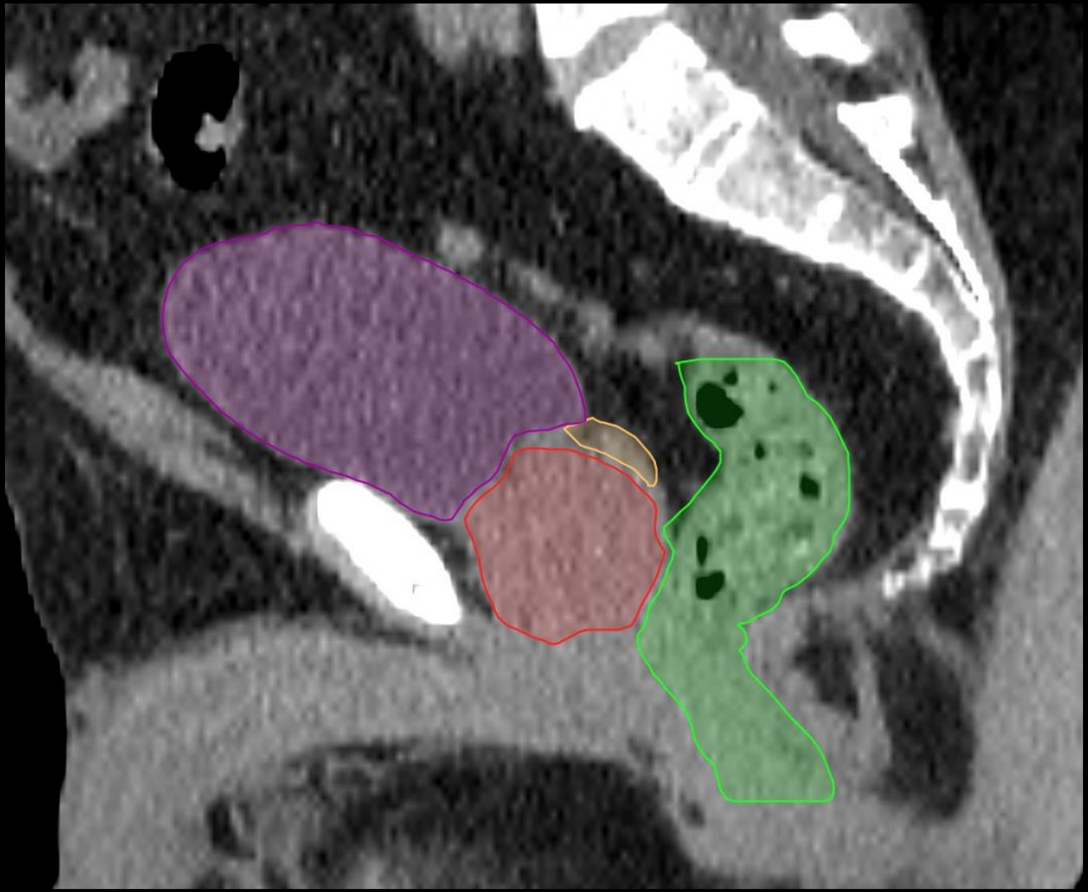
DLC Expert™
Deep Learning Contouring

Prostate Model

Contour List:
Prostate
Seminal Vesicles
Bladder
Rectum
L Femoral Head
R Femoral Head

Not Pictured:
Anus





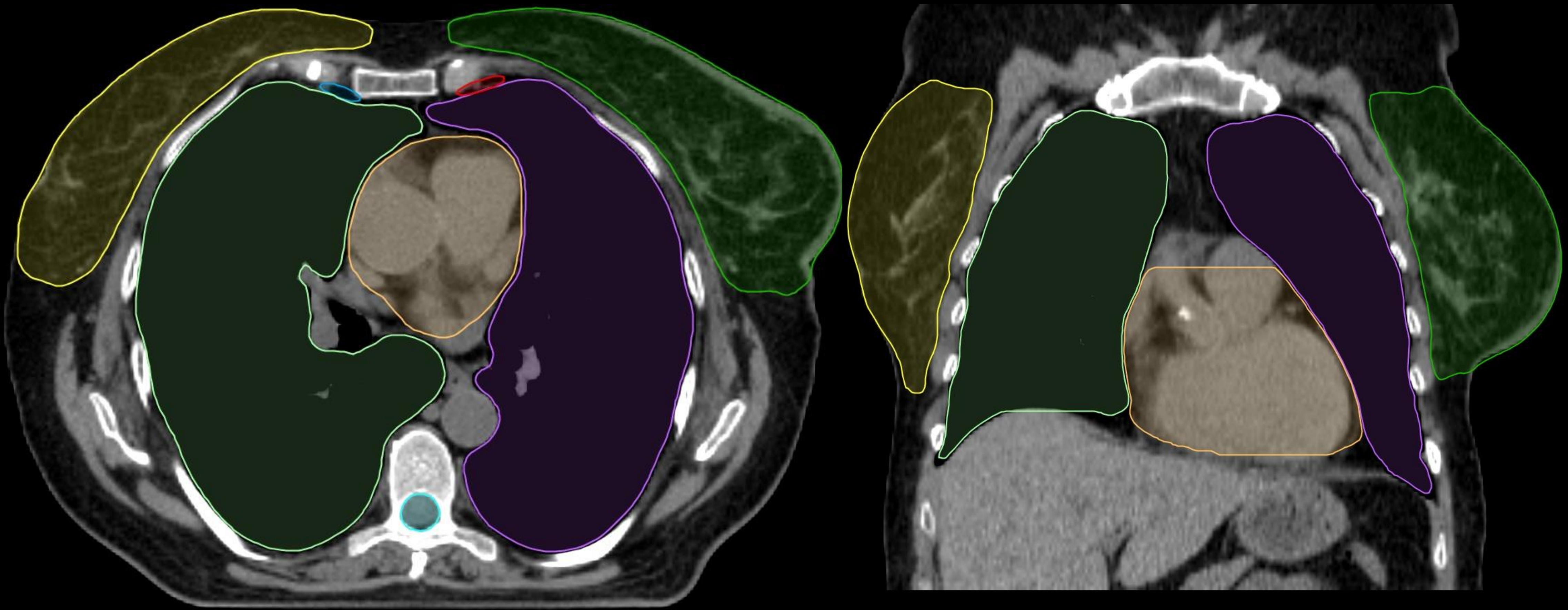
Supine Breast Model

Contour List:

L Lung
R Lung
Heart
Spinal Canal
R Breast
L Breast
L IMC
R IMC

Not Pictured:

Oesophagus



Health
Mid North Coast
Local Health District

DLC Expert™
Deep Learning Contouring

Custom Model Development

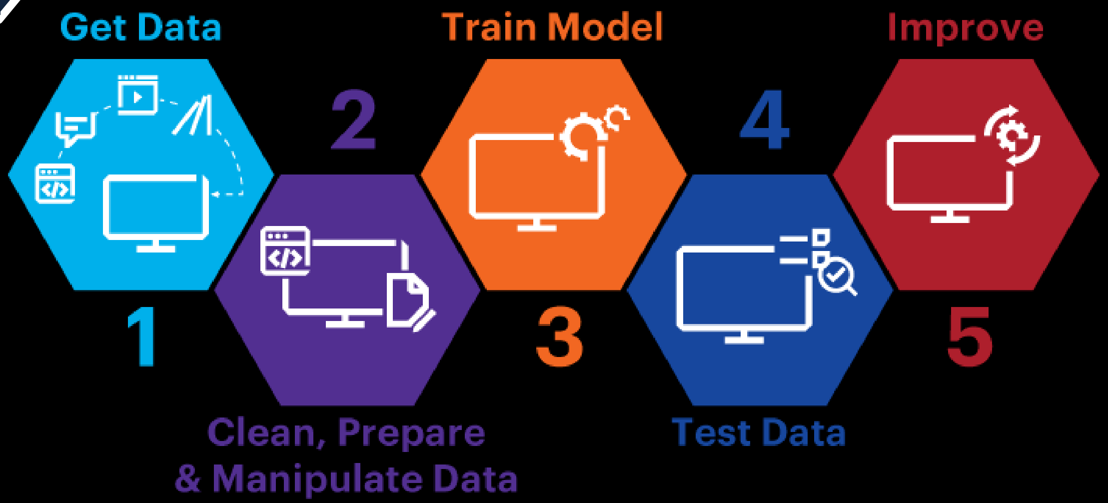
DLC models created by Mirada Science Team:

- On a remotely accessible local server or
- Datasets uploaded to Mirada HQ in Oxford

High-end system required (4x GPU's)

Min 150-200 consistently contoured datasets

4-6 week initial development process + testing + improved iterations



Custom vs Vendor Models

Custom Models	Vendor Models
Training data supplied by the **department**	Training data supplied by the **vendor**
Data curation lead by the **department**	Data curation lead by the **vendor**
Updates driven by the **department**	Updates driven by the **vendor**
Standardisation applied **locally**	Standardisation applied via **consensus guidelines**
Customisable naming & colour conventions	**Customisable** naming & colour conventions

Slide courtesy of Mirada Medical

DLC Workflow

1. CT Scan Completed
2. DICOM Data Sent To Mirada's Workflow Box™
3. CT Resampled If Pixel Size (FOV) +/- 5%
4. Dataset Contoured Using Appropriate Deep Learning Model
5. Post Processing Rules Applied
6. CT & Structure Set Pushed To TPS

Automated "One Click" Process

Average Time = ~11 minutes

System Requirements

DLC*Expert* Hardware Specifications

Item	Minimum Specification	Recommended Specification
Processor	Quad-core i7 or equivalent	2 x Quad-core Xeon or equivalent
Memory	16GB RAM	32GB RAM
Disk	250GB HDD (7200 RPM) 20GB available for system	250GB SSD 20GB available for system
Graphics Card	NVIDIA GeForce GTX 1060 or equivalent Graphics memory of 6GB CUDA Compute Capability: 3.5, 5.2, or 6.1	GeForce GTX 1080, equivalent or greater Graphics memory of 6GB or greater CUDA Compute Capability: 6.1
Operating System	Windows 10 (64-bit) Windows Server 2016	Windows Server 2016

Mirada offer an online DLC*Expert* trial via their website:

- De-identified datasets can be uploaded, processed, and retrieved for evaluation


PERFORMANCE
EVALUATION

DLC Evaluation Study

Research Question:

How do Deep Learning & Atlas-based contouring methods compare to clinical "gold-standard" RO contours?

Key Performance Indicators:

- Accuracy Calculation
- Qualitative Visual Assessment
- Time Benefit Analysis

DLC Evaluation Study

➢ Accuracy Calculation

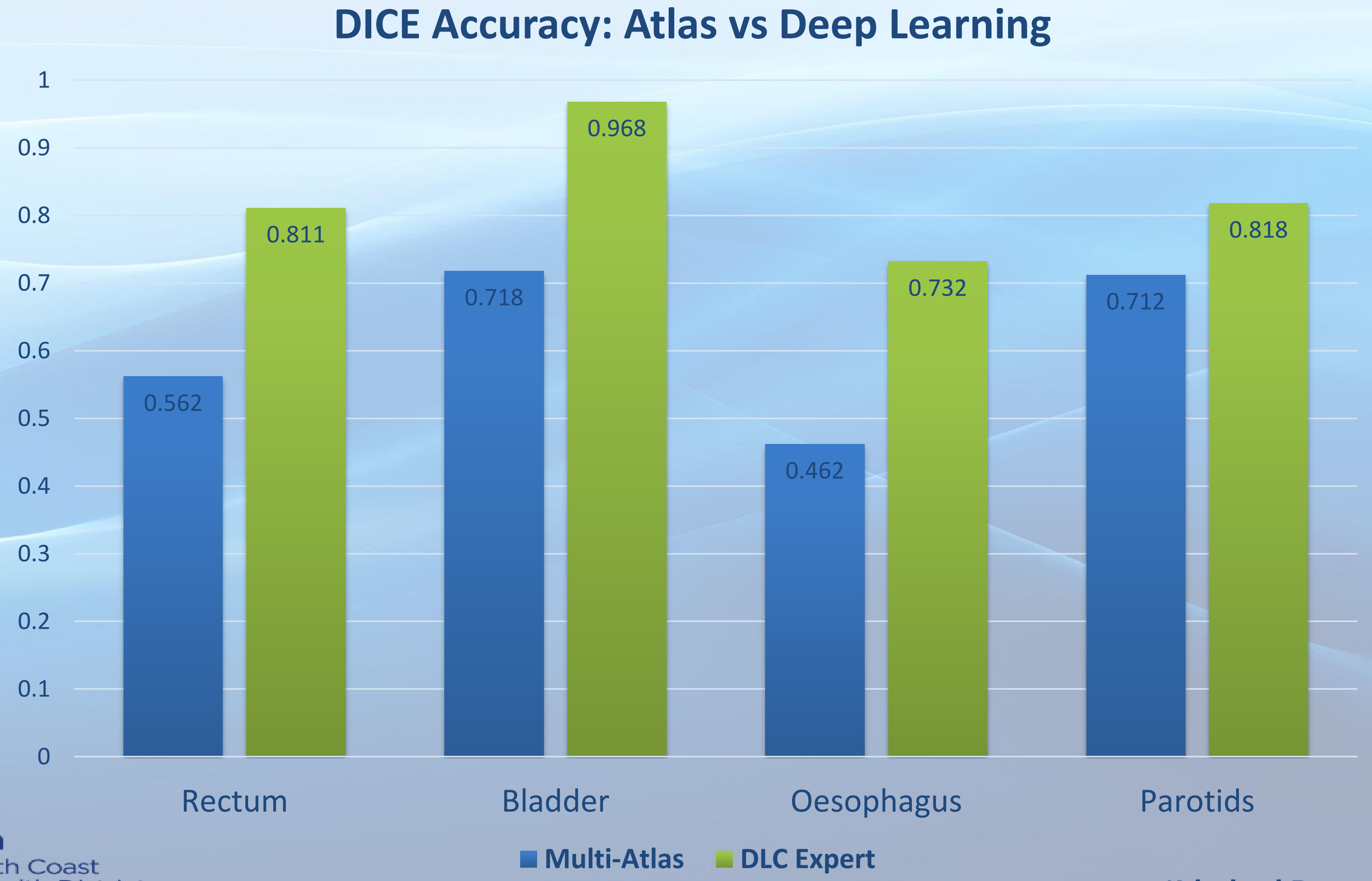
Unedited Deep Learning & Atlas contours measured using:

- Dice Similarity Coefficient (DICE)
- Hausdorff Distance

ProKnow software used to analyse data and calculate accuracy

Health
Mid North Coast
Local Health District

DLC Evaluation Study



Health
Mid North Coast
Local Health District

Limited Data - Preliminary Results

DLC Evaluation Study

➢ Visual Assessment

Classification	
Category	**Definition**
1	Accept contour as it is. Structure is very precise; ~**<1%** of cross-sectional slices require manual editing to meet clinical standards
2	Accept contour as it is. Minor edits to the structure may be required, however it is not clinically significant; ~**<10%** of cross-sectional slices require manual editing to meet clinical standards
3	Require contour to be corrected. Moderate edits to the structure are needed; ~**10-40%** of cross-sectional slices require manual editing to meet clinical standards
4	Require contour to be corrected. Major edits to the structure are needed; ~**>40%** of cross-sectional slices require manual editing to meet clinical standards

DLC Evaluation Study

➢ Visual Assessment

Average Qualitative Ranking (1-4)		
Anatomical Site	**Multi-Atlas**	**Deep Learning**
Prostate	3.0	2.04
Head & Neck	2.77	2.23

2	Accept contour as it is. Minor edits to the structure may be required, however it is not clinically significant; ~**<10%** of cross-sectional slices require manual editing to meet clinical standards
3	Require contour to be corrected. Moderate edits to the structure are needed; ~**10-40%** of cross-sectional slices require manual editing to meet clinical standards

DLC Evaluation Study

➢ **Time Benefit Analysis**

- Time measured to edit Deep Learning & Atlas contours to meet clinical standards
- Each contour will be delineated manually ***without*** the aid of auto-segmentation to give a baseline time result
- Structures will be timed individually to promote focus and limit the impact of potential distractions

DLC Evaluation Study

➢ Time Benefit Analysis

When compared to manual contouring ***without*** the aid of auto-segmentation, Deep Learning has shown the following time saving benefits:

LUNGS
53%
time saved

HEART
28%
time saved

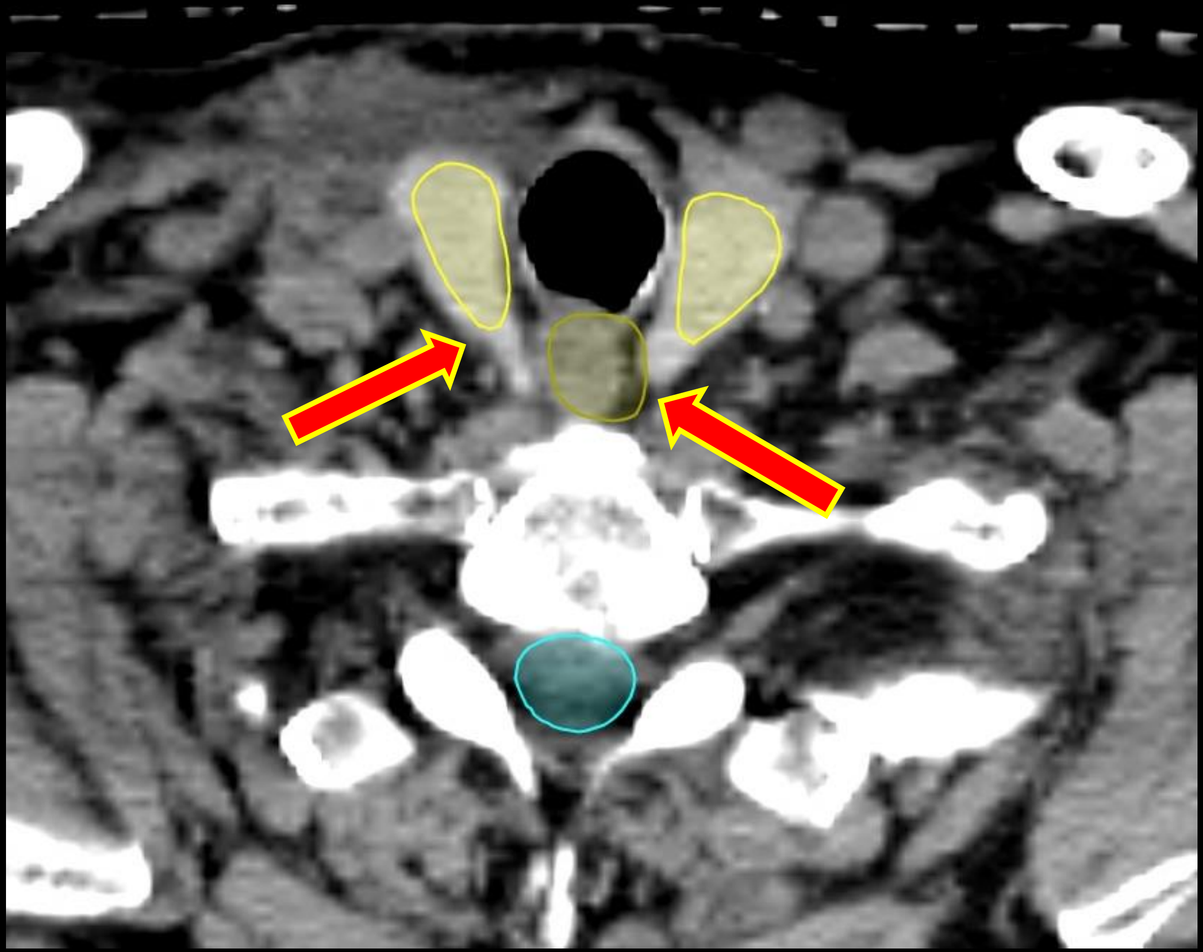
RECTUM
44%
time saved

BLADDER
64%
time saved

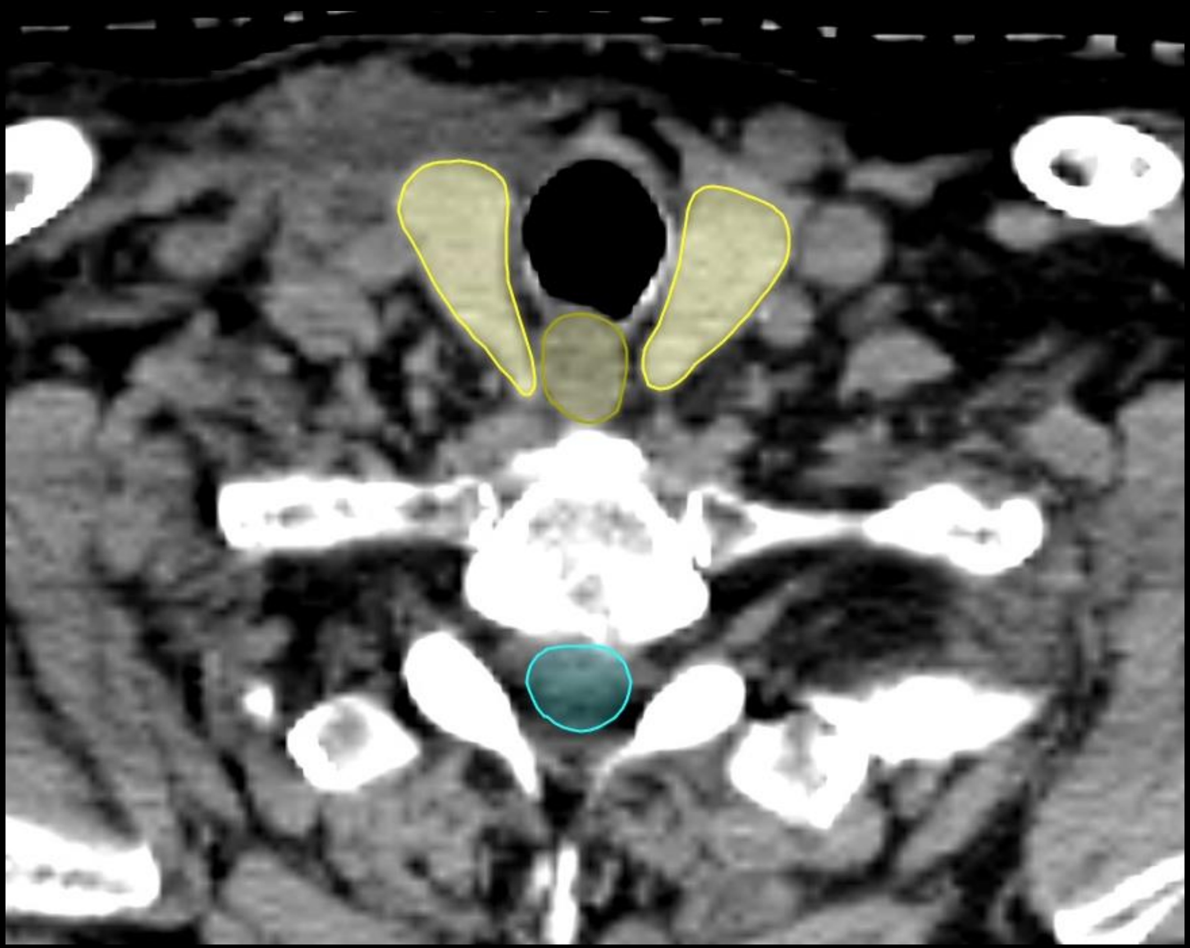
FEM HEADS
61%
time saved

Limited Data - Preliminary Results

Atlas vs DLC



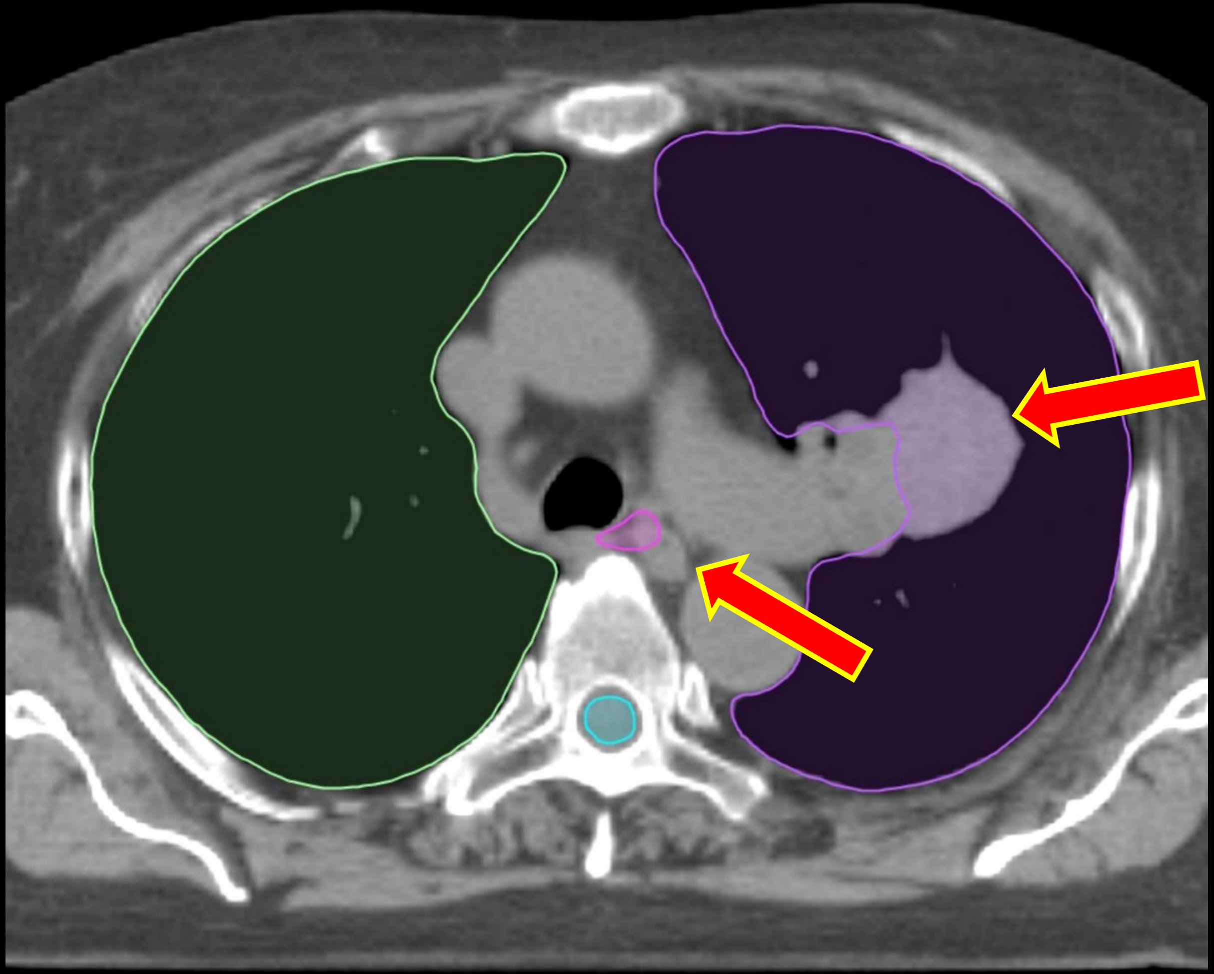
MNCCI H&N Atlas



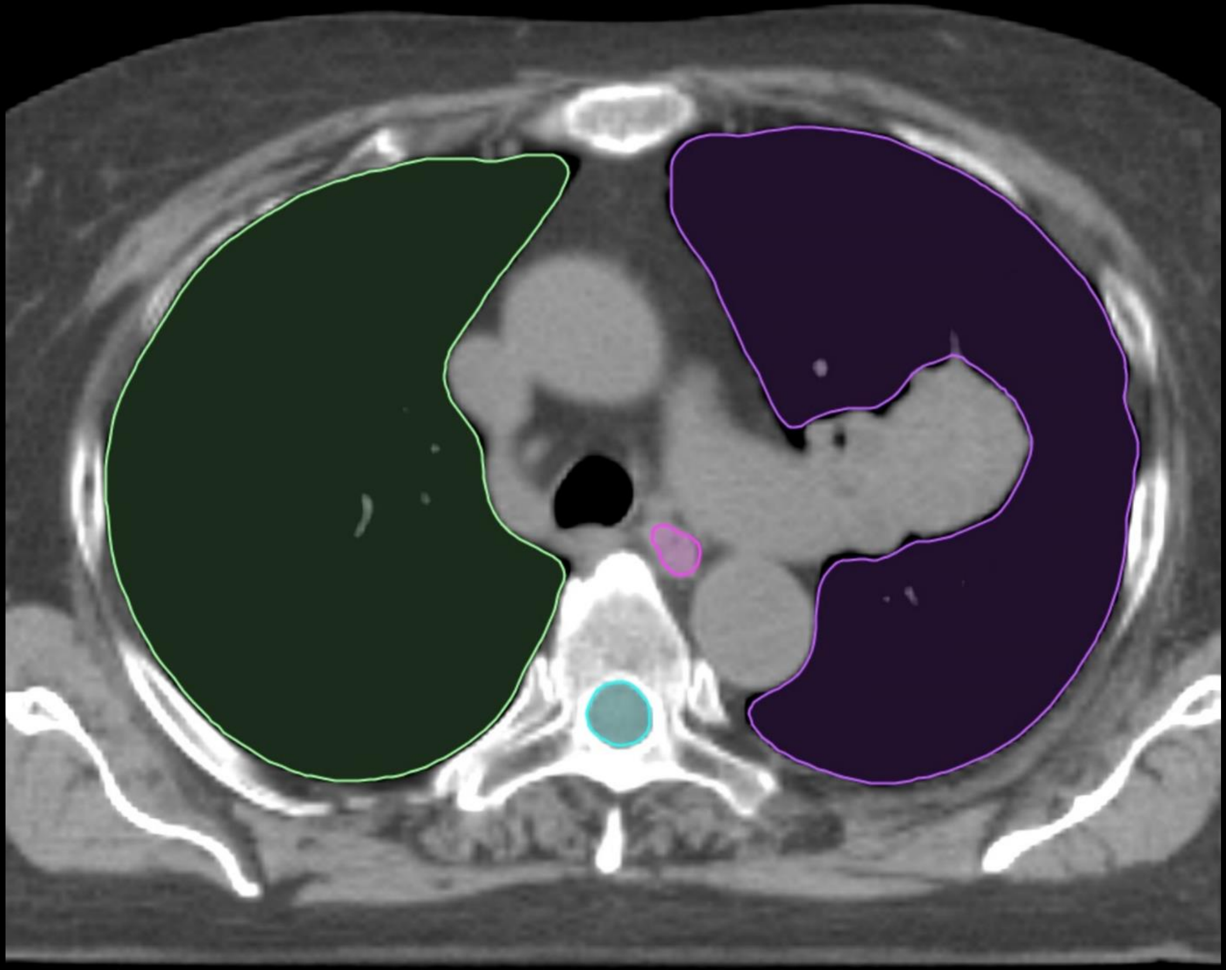
Deep Learning H&N Model

Contour List: Spinal Cord Thyroid Oesophagus

Atlas vs DLC



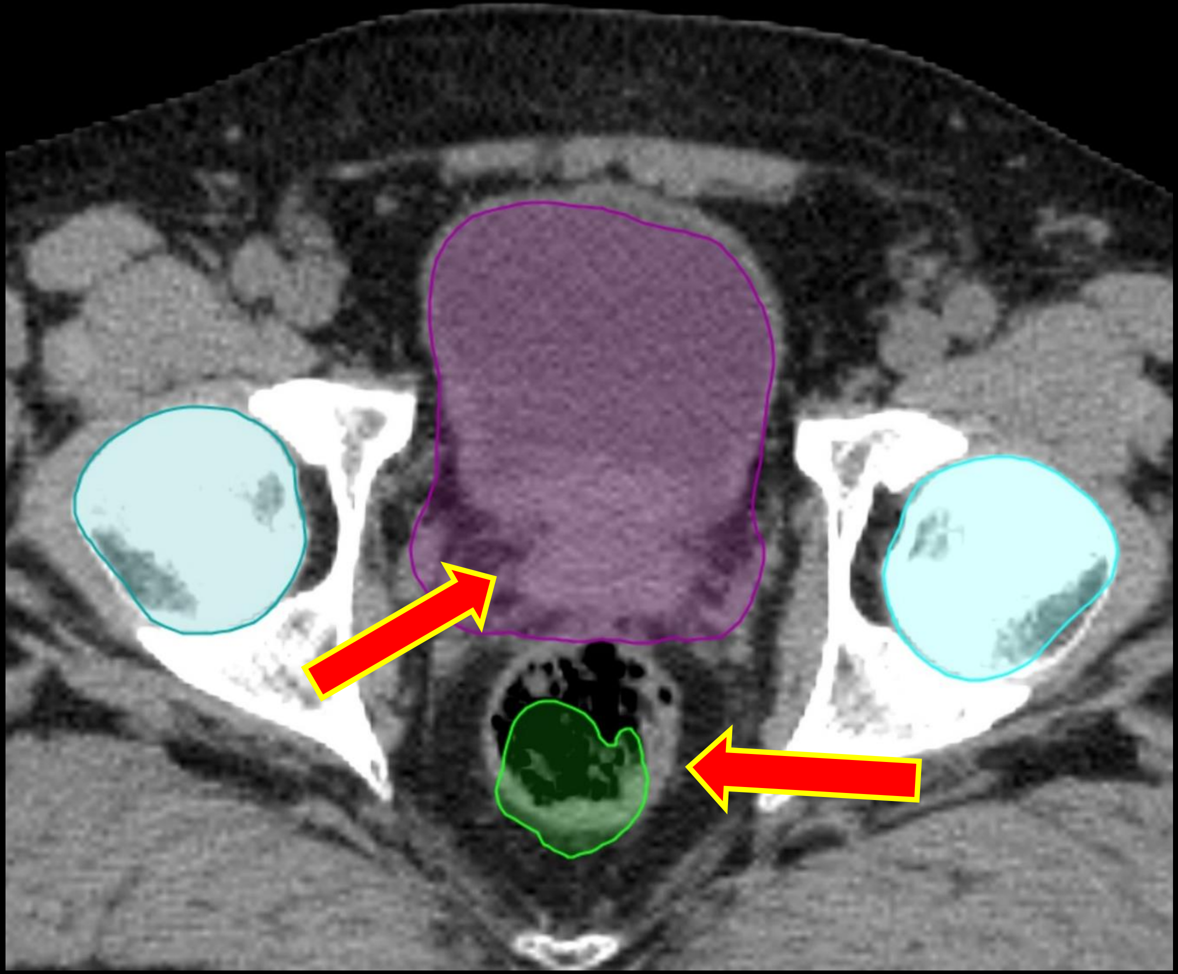
MNCCI Thorax Atlas



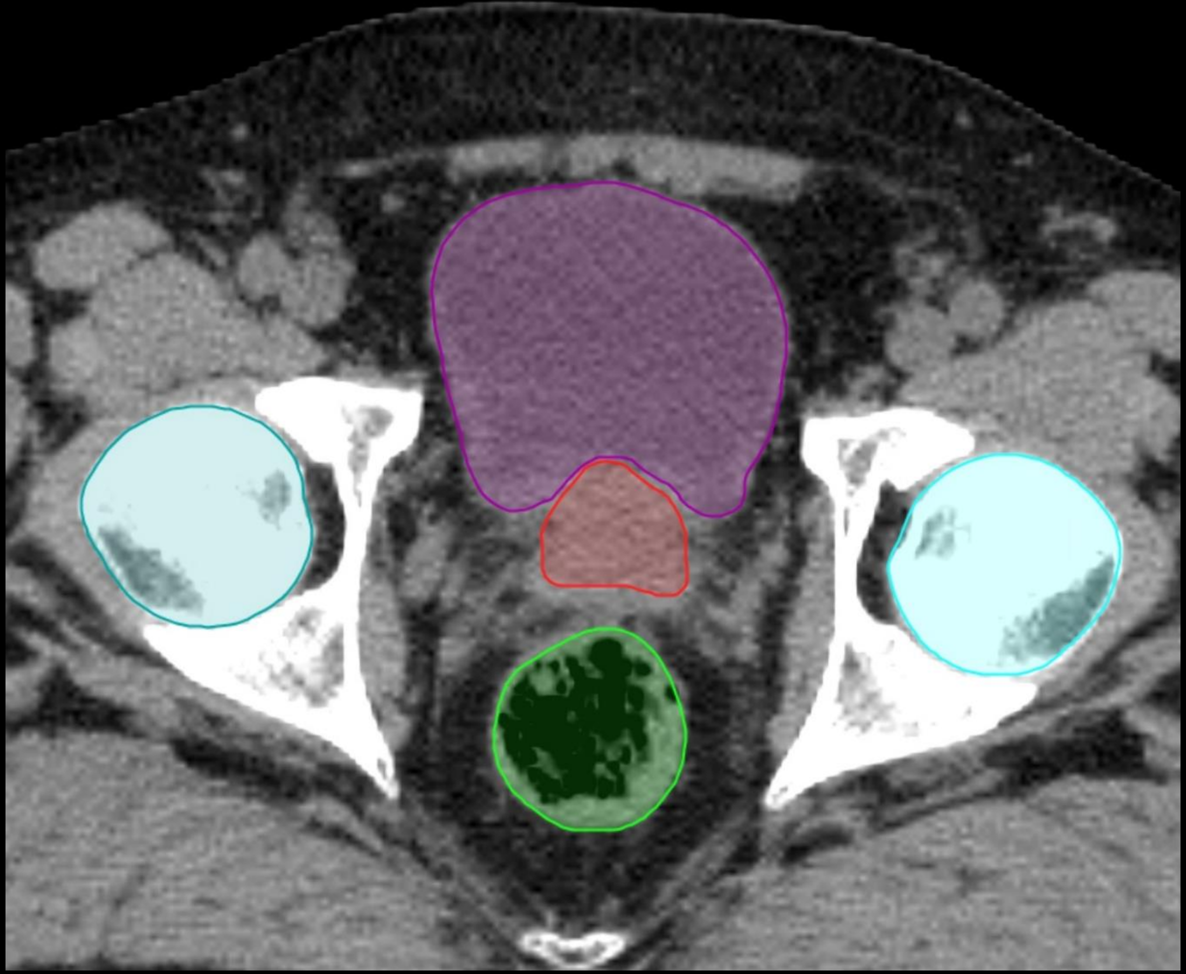
Deep Learning Thorax Model

Contour List: L Lung R Lung Spinal Cord/Canal Oesophagus

Atlas vs DLC



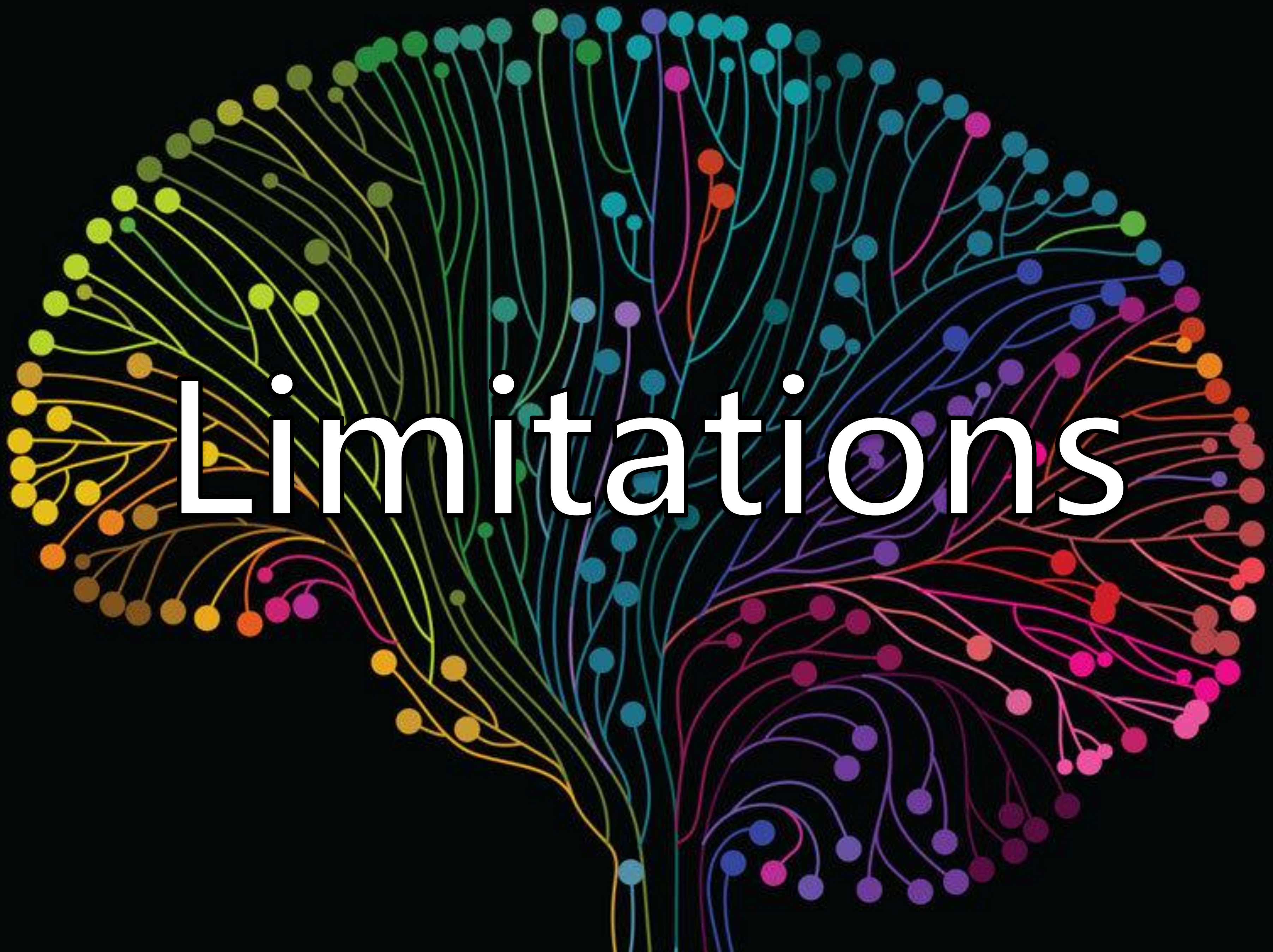
MNCCI Prostate Atlas



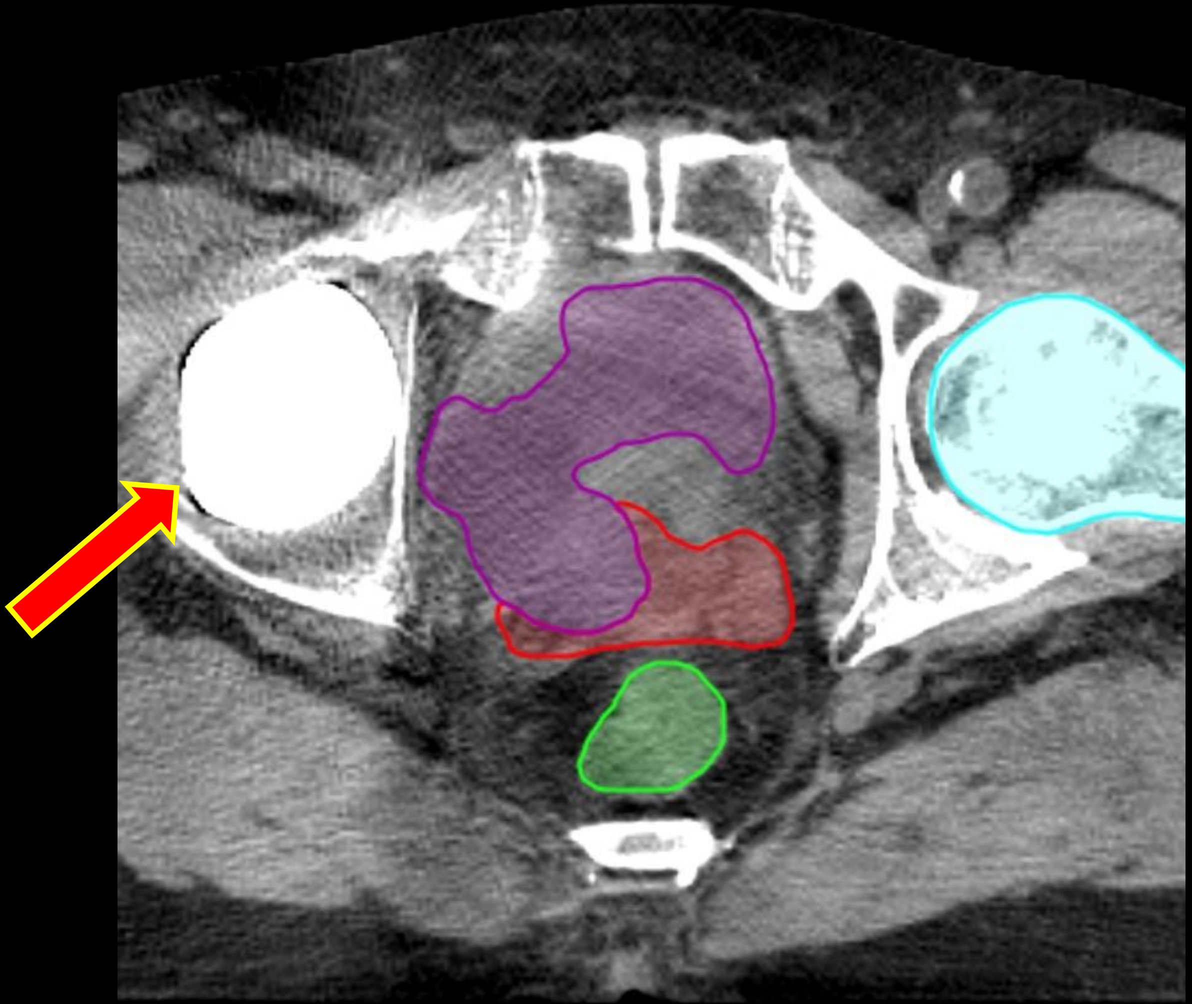
Deep Learning Prostate Model

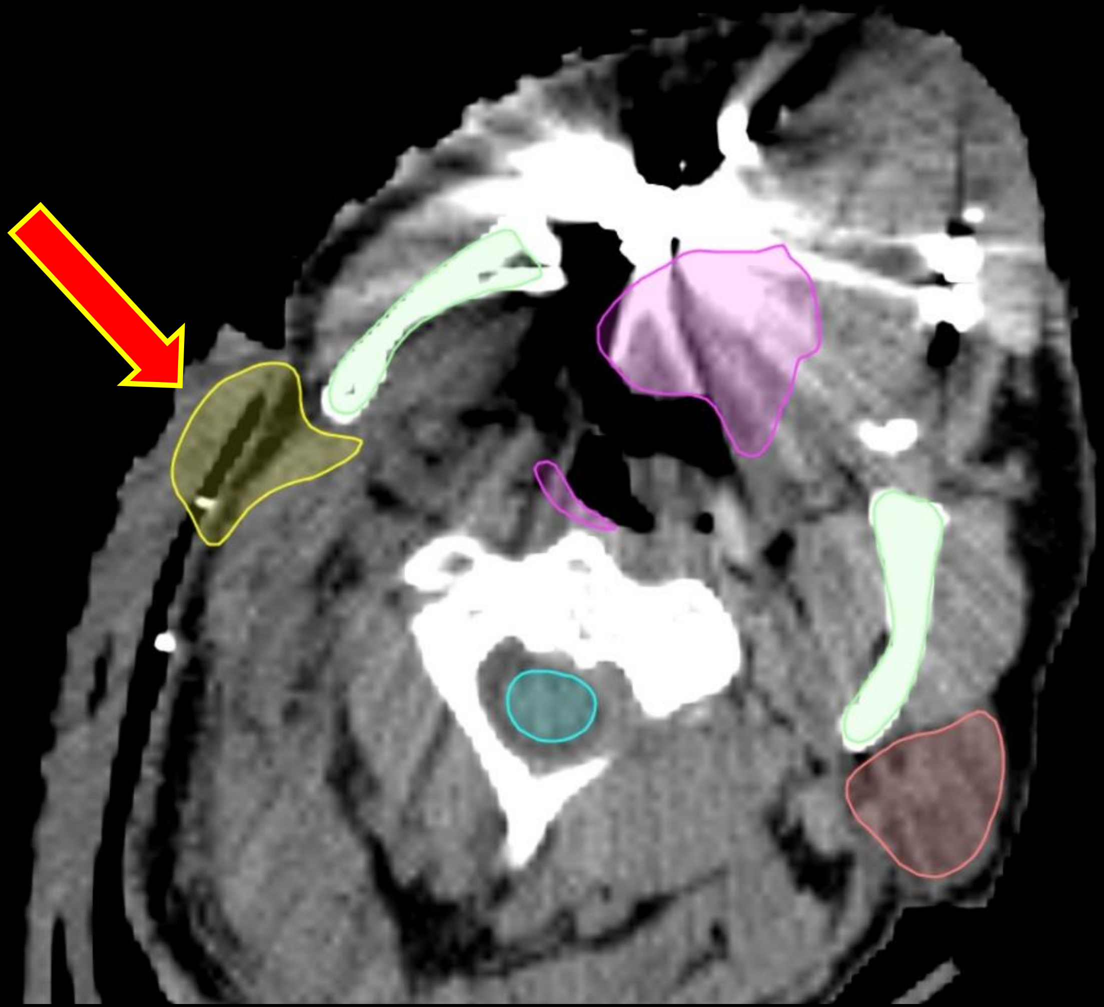
Contour List: Bladder Rectum Prostate L Femoral Head R Femoral Head

Limitations



Limitations

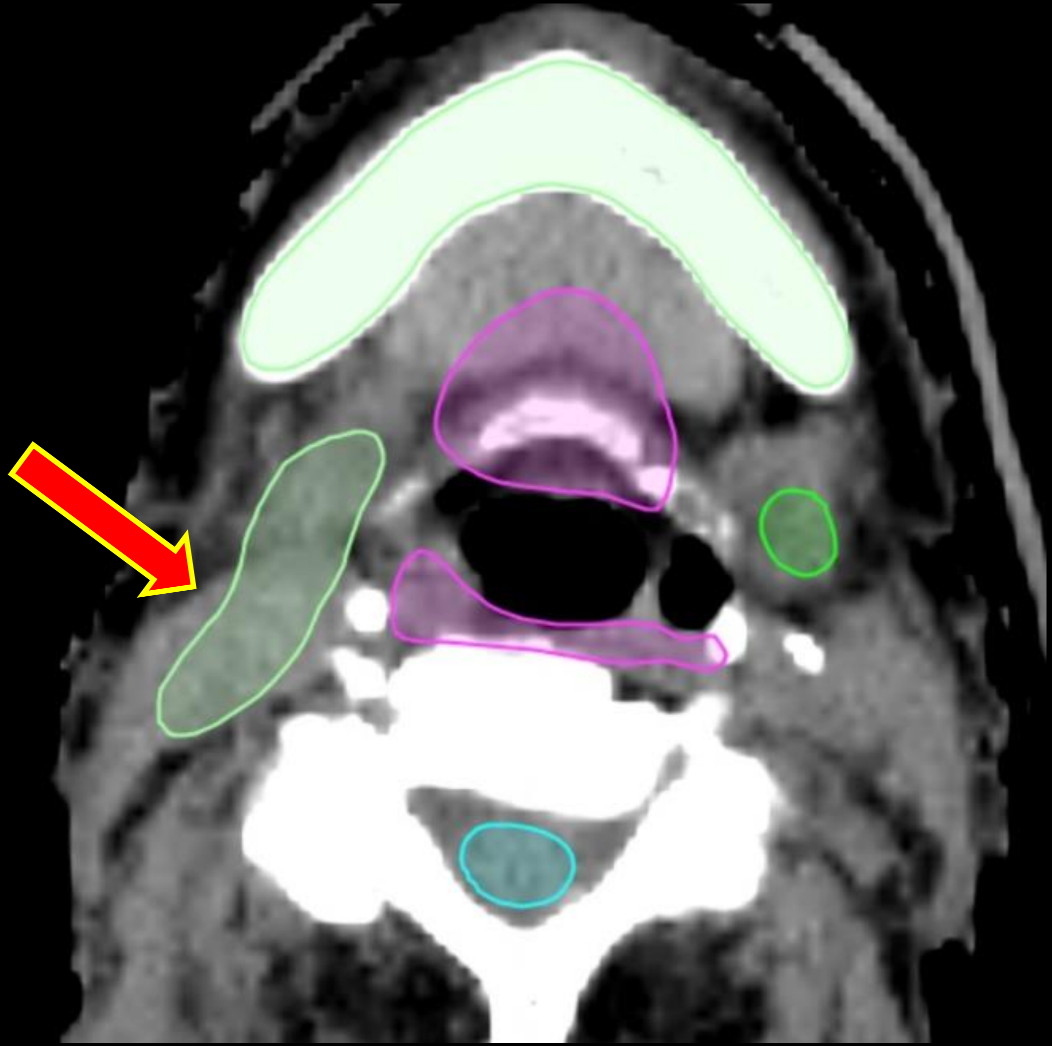




Health
Mid North Coast
Local Health District

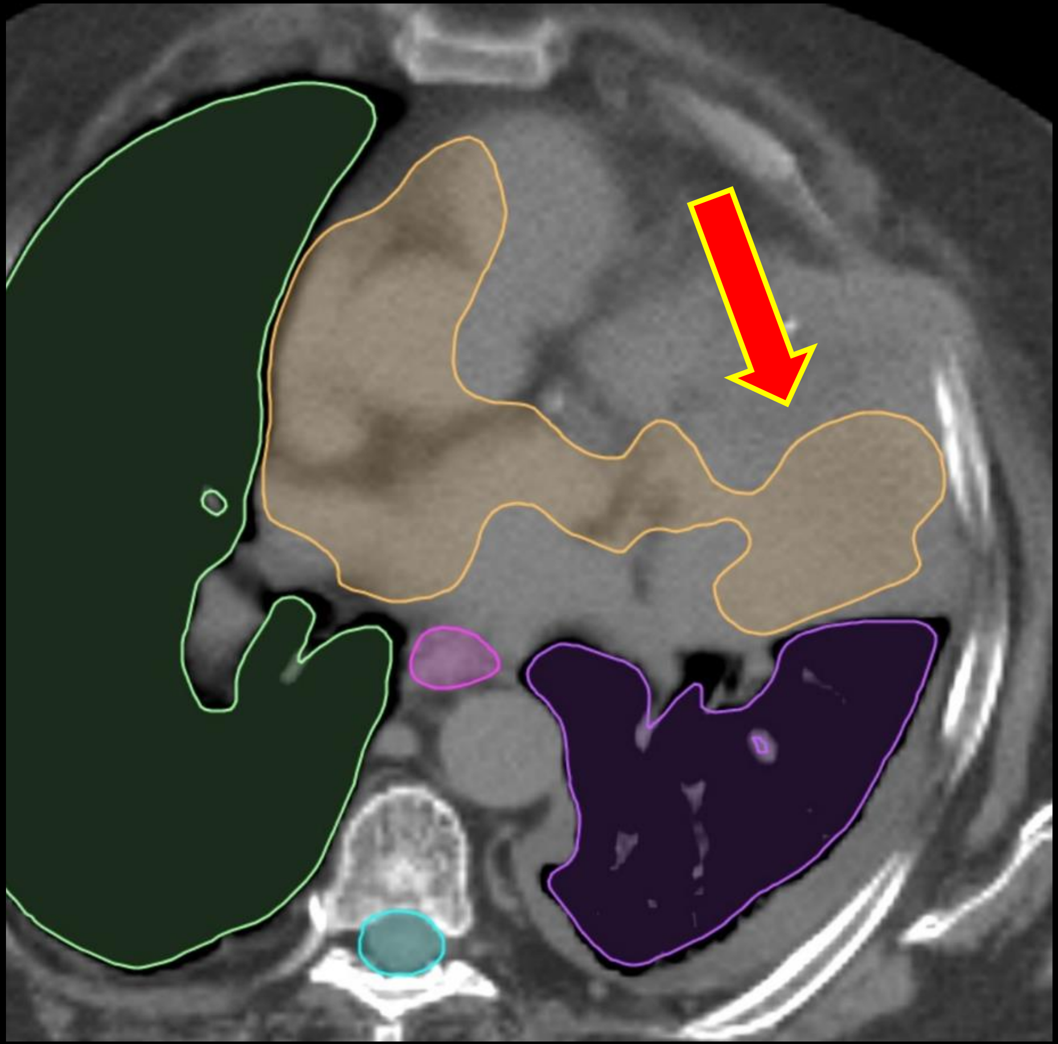
Contour List: Prostate Rectum Bladder L Femoral Head L Parotid R Parotid Oral Cavity Spinal Cord Mandible Pharynx Constrictor

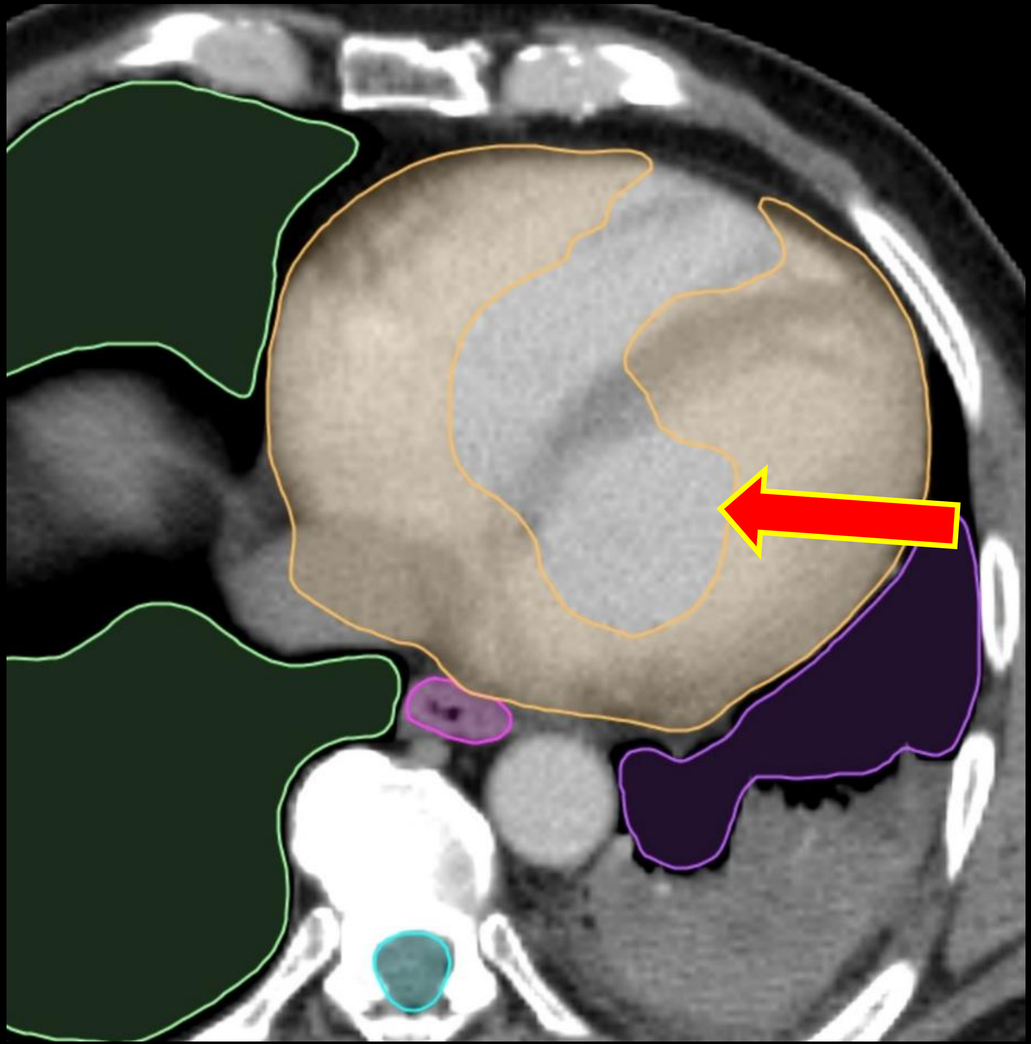
Limitations



Contour List: Prostate Rectum Bladder Seminal Vesicles Oral Cavity Spinal Cord Mandible L Submandibular R Submandibular Pharynx Constrictor

Limitations



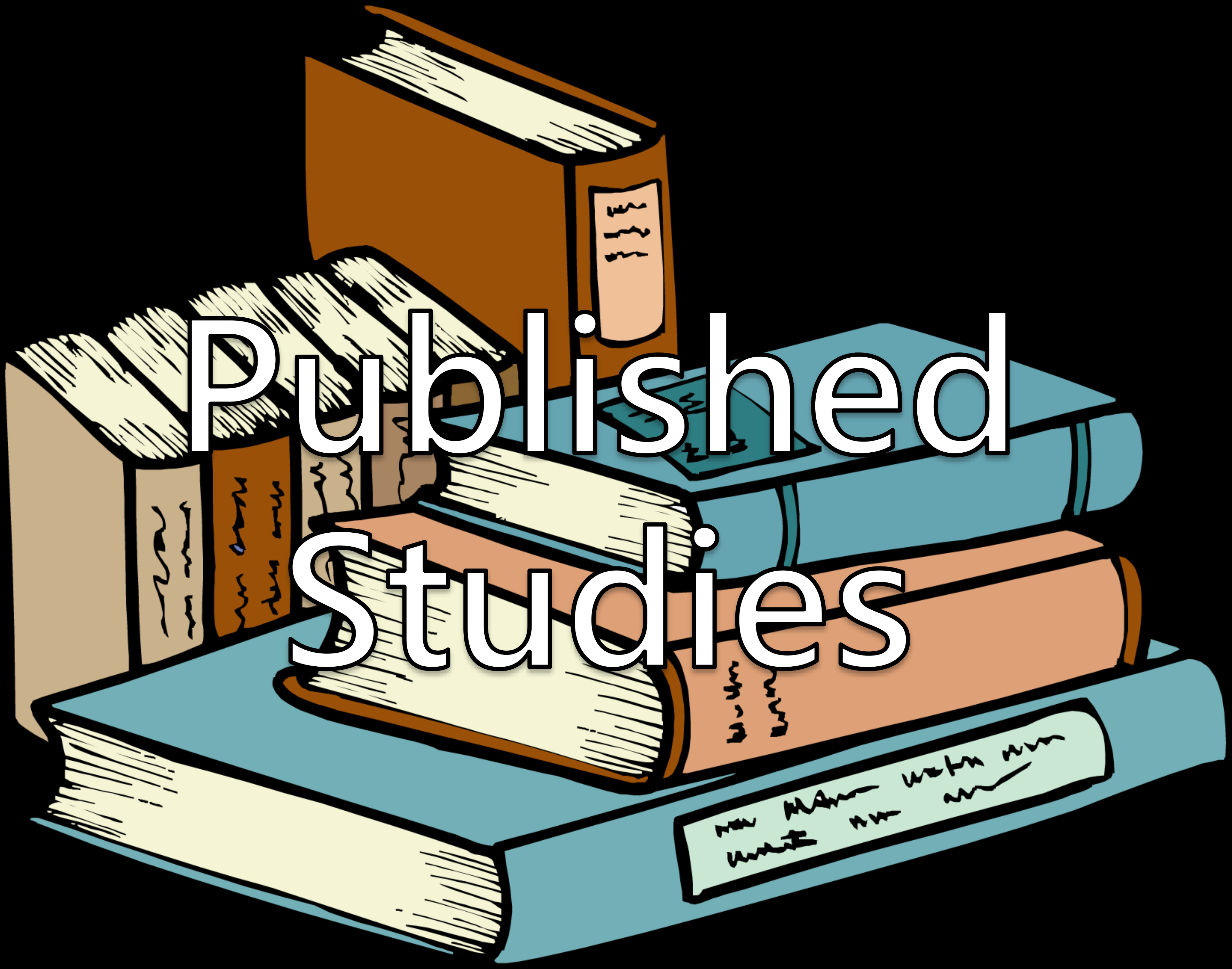


Contour List: Spinal Canal L Lung R Lung Oesophagus Heart

Clinical Limitations of Auto-Contouring

- Subject to Input Training Data
- Accuracy Cannot Consistently Match Human Performance
- Can Struggle to Adapt to Non-Standard Situations

Published Studies



Published Studies

Performance Comparison For H&N Auto-Contours

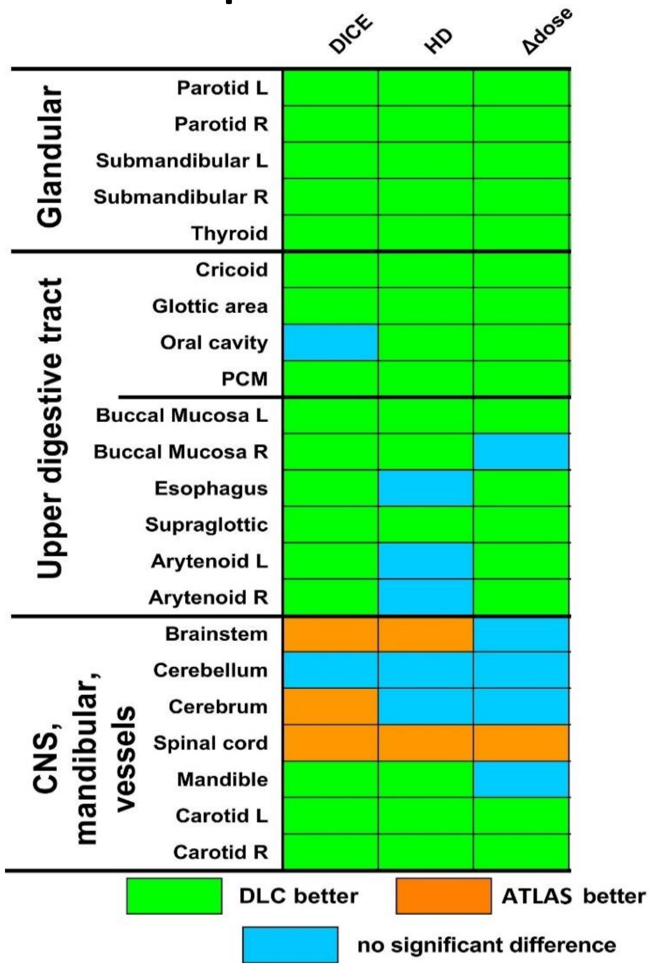
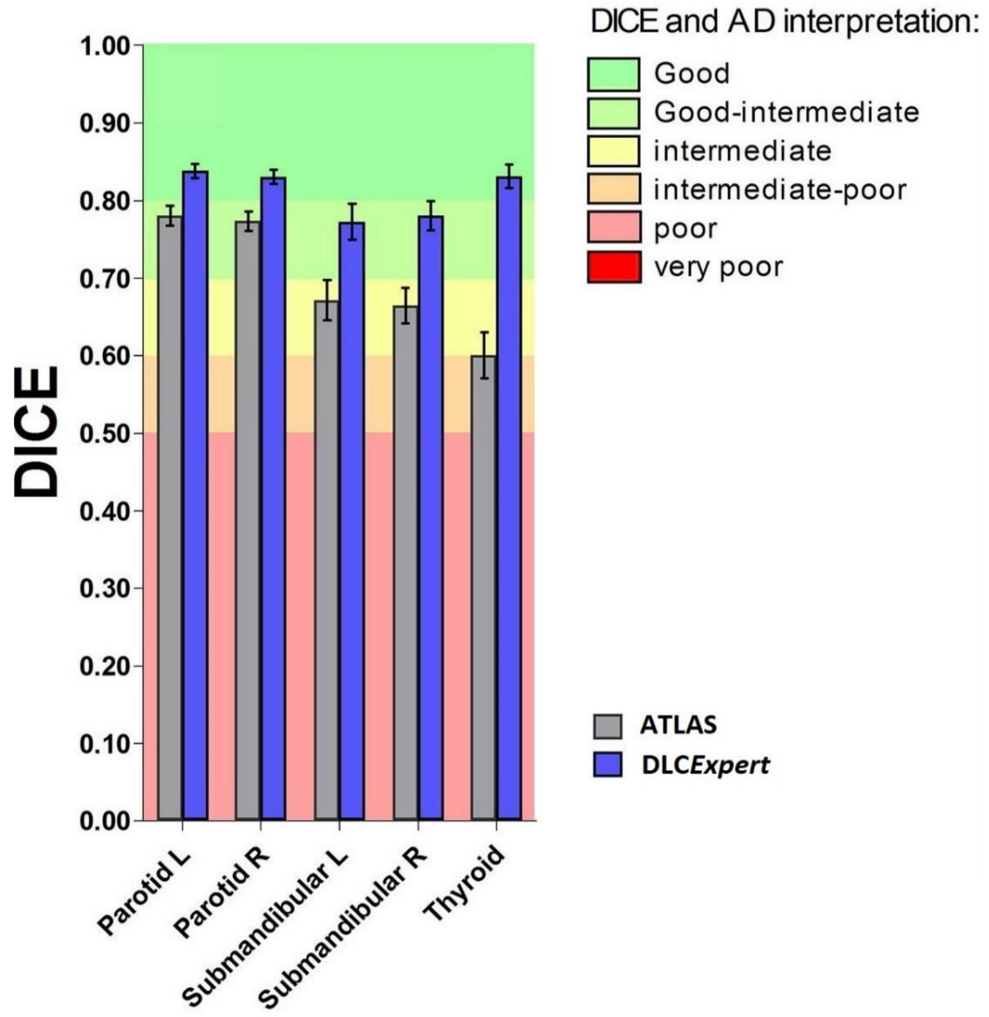


Fig. 7. Overview of the results of all HN OARs. Green indicates that DLC is significantly better than ABAS, orange that ABAS is significantly better than DLC and blue indicates that there is no significant difference.

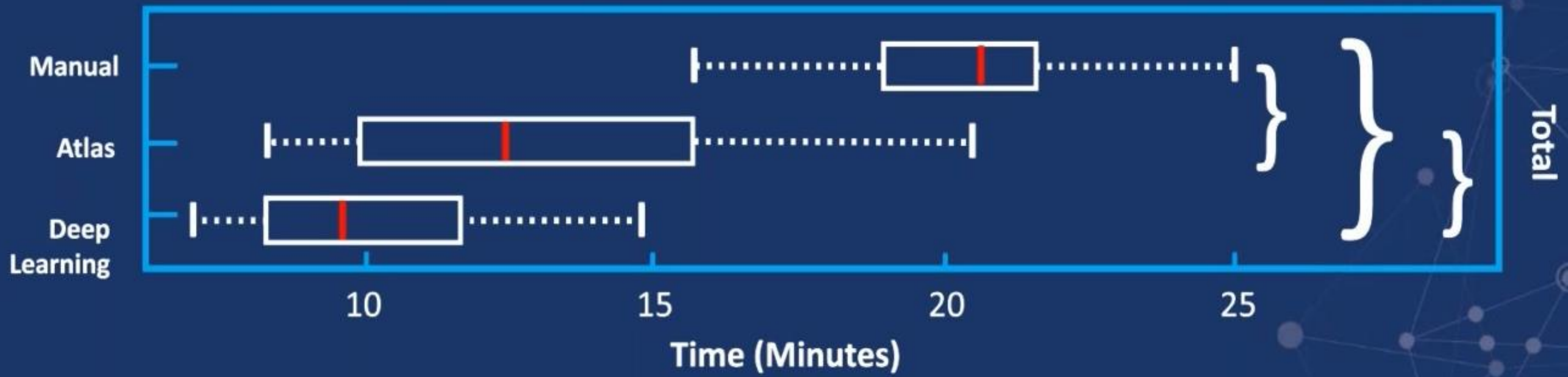
DICE Scores For Glandular H&N OAR's



[7] Van Dijk, L. et al. (2019). "Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring". Journal of Radiotherapy and Oncology, 142, 115-123

Time

Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer



66%

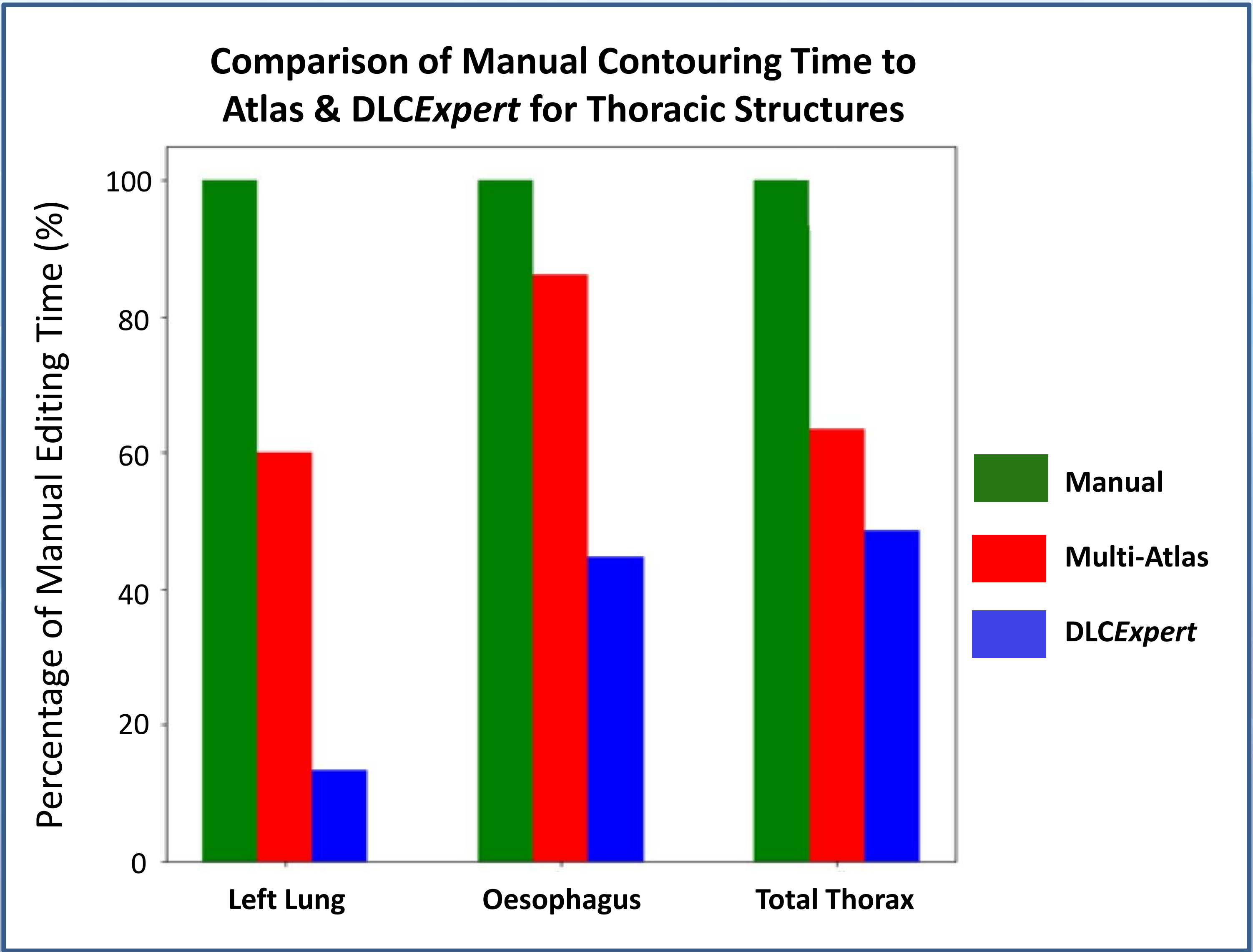
Time savings for thorax patients

Slide courtesy of Mirada Medical

DLCExpert
Deep Learning Contouring

[8] Lustberg, T., van Soest, J., Gooding, M., Peressutti, D., Aljabar, P., van der Stoep, J., ... & Dekker, A. (2018). Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. *Radiotherapy and Oncology*, *126*(2), 312-317.

Published Studies



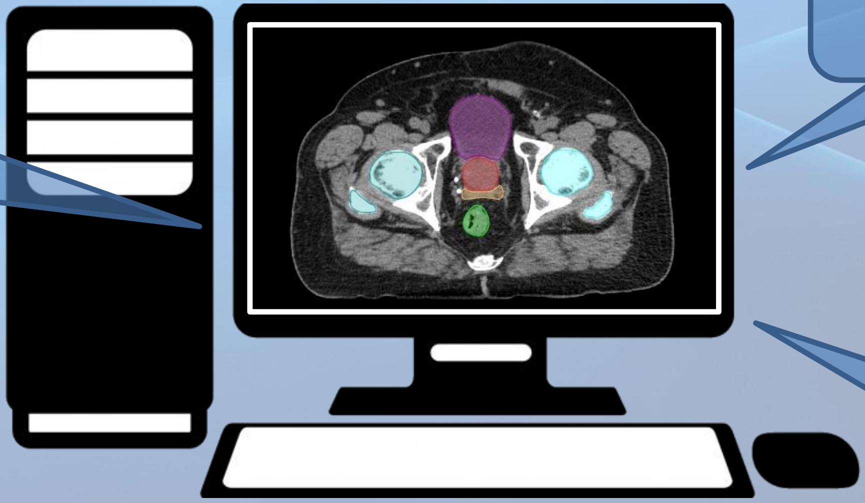
[9] Peressutti, D. et al. (2018). "Evaluation of DLCExpert for Contouring of Thoracic Organs-At-Risk". Mirada-Medical.com

Published Studies

Qualitative assessment of Deep Learning contours: Turing Test

- Classical Test of AI
- Try for yourself *http://www.autocontouring.com*

How was this contour drawn? Human or computer?

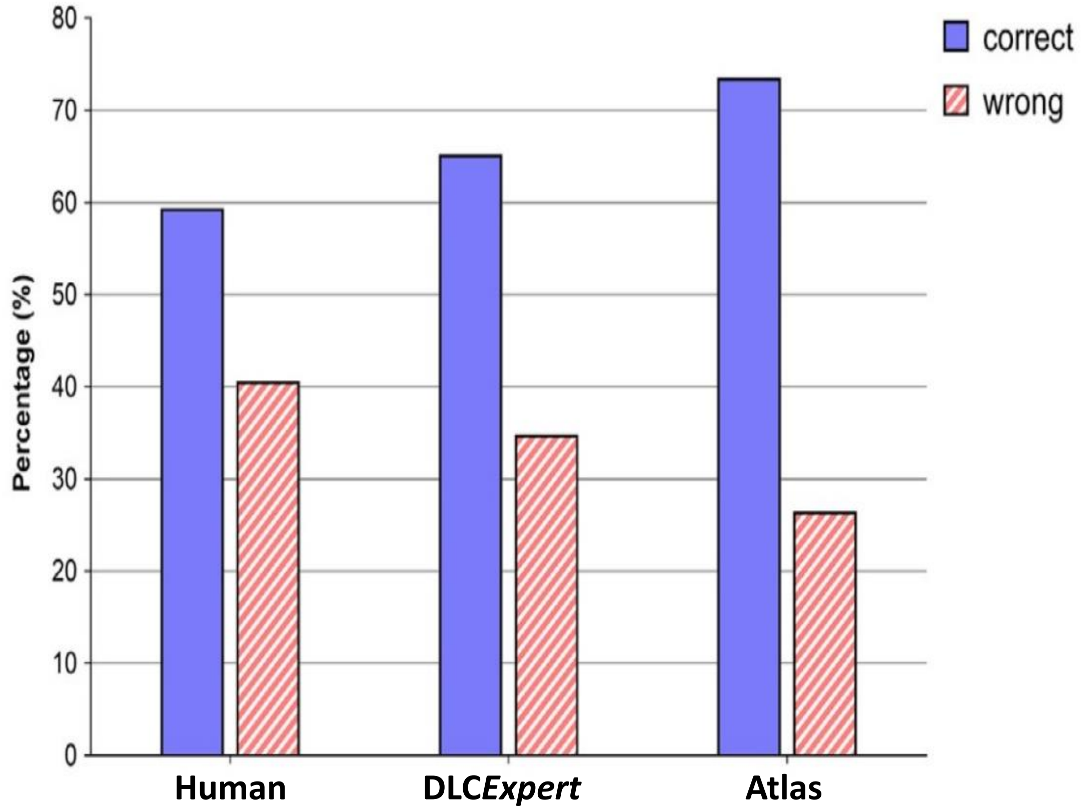


Is this contour clinically acceptable?

Which of these two contours do you prefer?

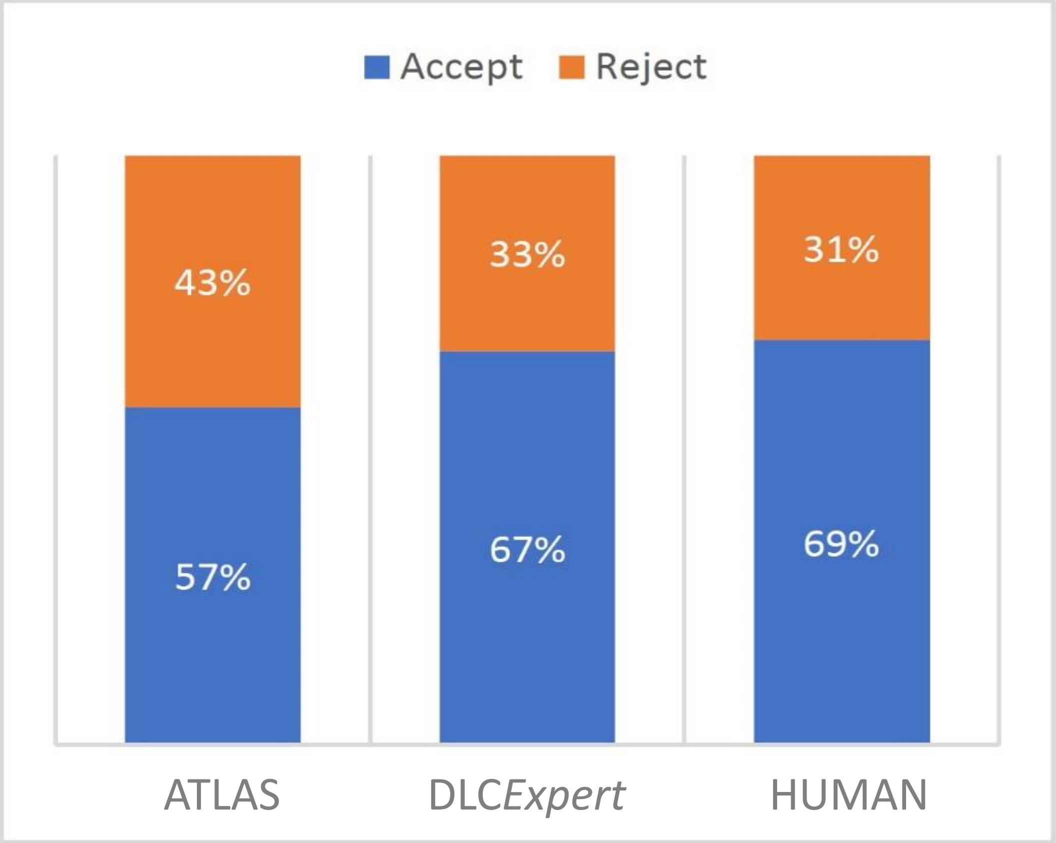
Published Studies

Turing Test For H&N Structures: "Human or Computer?"



[7] Van Dijk, L. et al. (2019). "Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring". Journal of Radiotherapy and Oncology, 142, 115-123

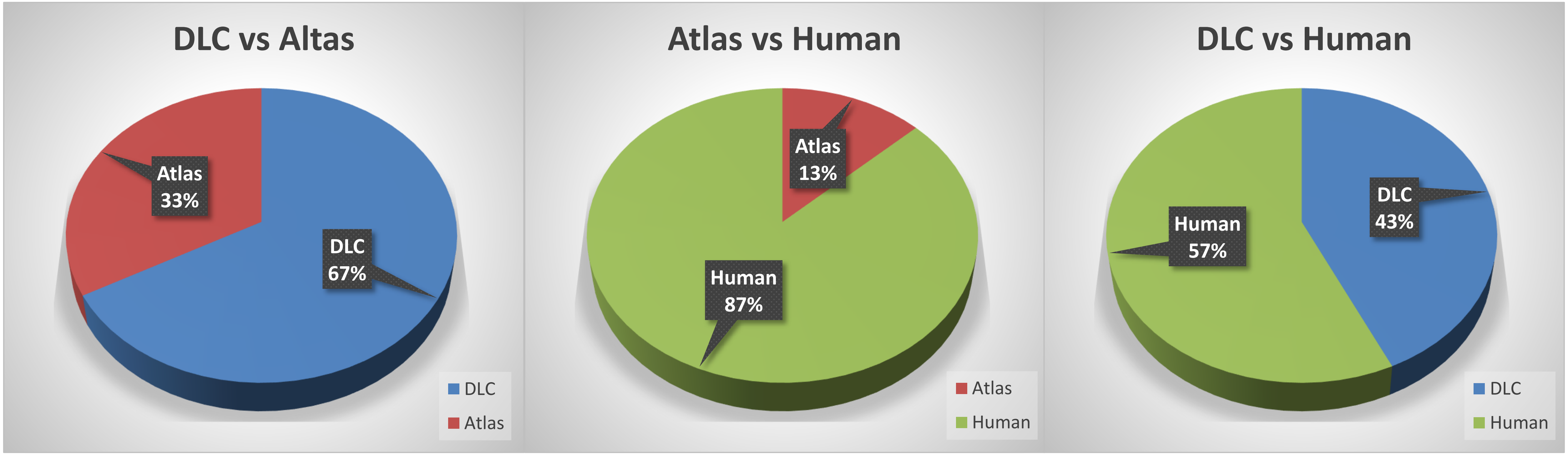
Qualitative Clinician Acceptance for Thoracic Structures



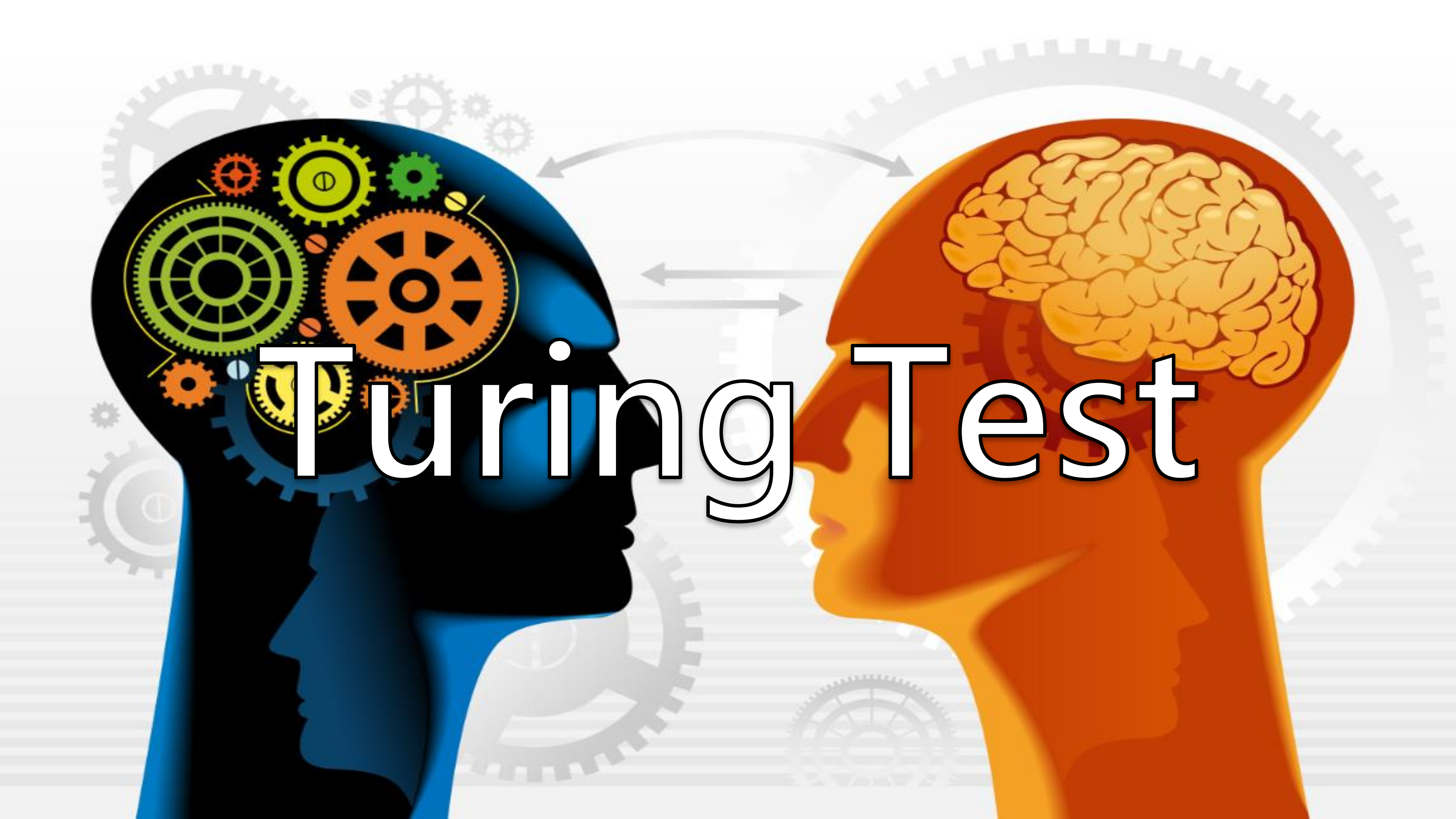
[9] Peressutti, D. et al. (2018). "Evaluation of DLCExpert for Contouring of Thoracic Organs-At-Risk". Mirada-Medical.com

Published Studies

Contour Preference In Blind Side-By-Side Comparison For Prostate Segmentation



[10] Gooding, M. et al. (2018). “Multi-centre evaluation of atlas-based and deep learning contouring using a modified Turing Test”. ESTRO 2018 Poster


Turing Test

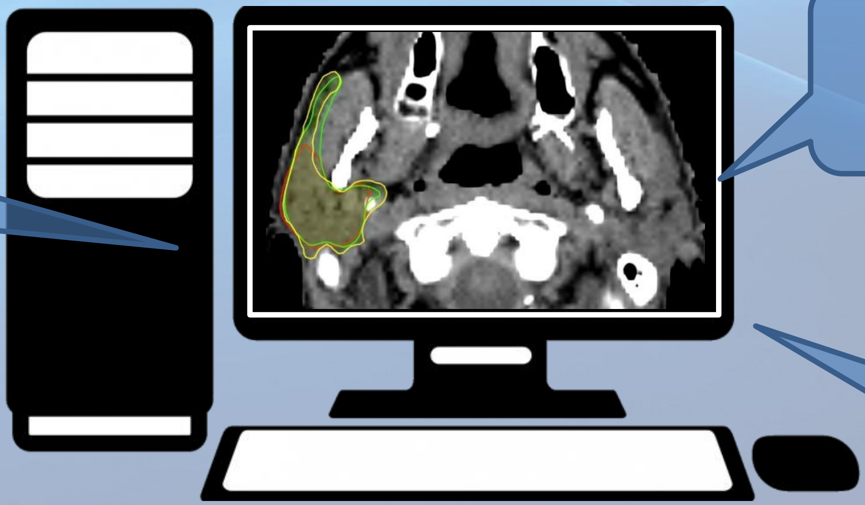
Modified Turing Test

➢ How This Works

You will choose between 3 blinded contours:

- Human Drawn
- Atlas Drawn
- Deep Learning Drawn

Which contour was drawn by a human?



Is this contour clinically acceptable?

Which of these contours do you prefer?

Conclusion

Advancements in AI technology has opened the door to a new "gold-standard" in auto-contouring

Deep learning models can generate superior contours compared to atlas-based methods, leading to tangible time-saving benefits

Auto-contouring can not consistently match human performance. Each structure requires review by a trained clinician

Future Developments

- Updating current Deep Learning models with new structures
- Structure set merging of multiple models & atlases
- Developing Deep Learning models based on MR datasets
- Deep Learning Deformable Image Registration

References

1. Gorini, M. (2020). “What is the Difference Between Machine Learning and Deep Learning?”. BiSmart. Accessed 14/10/20. https://blog.bismart.com/en/difference-between-machine-learning-deep-learning
2. Saravana, N. (2019). “How Deep Learning is Different from Machine Learning”. DataWider. Accessed 15/10/20. https://datawider.com/how-deep-learning-is-different-from-machine-learning/
3. Vazquez, F. (2017). “Deep Learning Made Easy with Deep Cognition”. Kdnuggets. Accessed 16/10/20. https://www.kdnuggets.com/2017/12/deep-learning-made-easy-deep-cognition.html
4. Mukesh, M. et al. (2012). “Interobserver variation in clinical target volume and organs at risk segmentation: Can segmentation protocols help?” The British Journal of Radiology, 85, 530-536
5. Aljabar P, Gooding M. (2017). [Mirada White Paper]. “The cutting edge: Delineating contours with Deep Learning”
6. Albano, A. (2018). [Press Release]. “Mirada releases DLC*Expert* – First commercially available Artificial Intelligence (AI) autocontouring software for radiation oncology”, Mirada-Medical.com
7. Van Dijk, L. et al. (2019). “Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring”. Journal of Radiotherapy and Oncology, 142, 115-123
8. Lustberg,T. et al. (2018). “Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer”. Journal of Radiotherapy and Oncology, 126(2), 312-317
9. Peressutti, D. et al. (2018). “Evaluation of DLCExpert for Contouring of Thoracic Organs-At-Risk”. Mirada-Medical.com
10. Gooding, M. et al. (2018). “Multi-centre evaluation of atlas-based and deep learning contouring using a modified Turing Test”. ESTRO 2018 Poster

The ACPSEM Medical Image Registration Special Interest Group (MIRSIG) Online Webinars

Questions and Answers from the November 2020 Webinar Chaired by Laurel Schmidt (Talk 1 by Eddie Gibbons)

Question 1: Does the system accept MR images for training?

Answer:

Yes, it is possible to train a deep learning model on MR datasets. Mirada participated in a contouring challenge at the AAPM meeting in 2019 which looked at deep learning auto-contouring on MR datasets. No commercial models have been released as of yet, but it is something that is being looked into. Body sites where MR auto-contouring may be of benefit could be the brain, prostate, rectum, and some H&N sites. I will also note that it is currently possible to create atlas-based auto-contouring models using MR images.

Question 2: How would deep learning work if users wanted some contours from CT and others from MRI dataset?

Answer:

This might be tricky, as I believe deep learning models can only be trained using an individual scan modality (i,e. only CT, or only MR). To get the desired outcome, the user would likely need two separate deep learning models. One for CT and one for MR. The deep learning structures from both models could then be merged/deformed onto the planning CT.

Question 3: How does DL work for nodal delineation (e.g. HN, breast etc.)?

Answer:

The models we currently have installed do not contour the nodal chains for H&N or breast patients. The reason being is that they were not trained to do so. However, I see no reason why this wouldn't work. If a deep learning model was developed which incorporated consistently contoured nodal structures in the training data, I believe the model would deliver the desired result. I am unsure if this has already been attempted, however I would assume that a deep learning based model would handle nodal structures better than atlas-based methods.

Question 4: Which clinical sites you feel DLC works better than ATLAS?

Answer:

For the four models we have installed (H&N, prostate, thorax, supine breast), I would choose deep learning over atlas for all of these sites. There are a few rare structures where atlas contouring can outperform deep learning (i.e. spinal cord, brainstem), however on the whole, deep learning is my preferred method of auto-contouring for each of the listed body sites.

ACPSEM Australasian College of Physical Scientists & Engineers in Medicine

MIRSIG

ACPSEM CPD Endorsed Activity

The ACPSEM Medical Image Registration Special Interest Group (MIRSIG) Online Webinars

Questions and Answers from the November 2020 Webinar Chaired by Laurel Schmidt (Talk 1 by Eddie Gibbons)

Question 5: How well does AI work for head & neck and what level of time-saving has been reported?

Answer:

Deep learning performs very well for head & neck cases. The study we have undertaken has not yet assessed time saving for H&N OAR's, although anecdotally it is providing better time saving benefits than atlas contouring. This is due to the increased accuracy it provides over atlas methods. For papers that have reported on DLC*Expert* for the H&N, see:

Van Dijk, L. et al. (2019). "Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring". Journal of Radiotherapy and Oncology, 142, 115-123

Brunenberg, E. et al. (2020). External validation of deep learning-based contouring of head and neck organs at risk". Physics and Imaging in Radiation Oncology. 15, 8-15

Question 6: Did you look into any other solutions from other vendors, if so, why did you choose Mirada?

Answer:

As we are current Mirada users, we pursued DLC*Expert* because we already have experience operating the software and have a good working relationship with the vendor. When we first began looking into deep learning segmentation, Mirada was the only vendor offering a commercial solution, which made the decision easy. We didn't look at any other solutions from other vendors because they weren't being offered at the time.

Question 7: Does Mirada provide the tools for customers to train their own models? Or does the data have to be sent to them?

Answer:

To the best of my knowledge, the actual training of the model is completed by Mirada staff. This can be done online by uploading CT data to a Mirada server via FTP, or by configuring a local database that can be remotely accessed by Mirada staff. I do not believe there is currently an option for the customer to perform the training themselves.

Question 8: How does Mirada DL compare with other commercial DL systems e.g. Raystation?

Answer:

I have not had any experience with other commercial deep learning segmentation systems unfortunately. I know a few other vendors now offer solutions, however I don't believe there are any studies that have looked at comparing the results from different applications. My view is that the most important factor when looking to achieve accurate results is the quality of the training data. I would assume that all vendors handle the model development process in a similar way, which shouldn't significantly impact the results.