

Development and validation of a machine learning system for automated routine 2-dimensional morphometric measurements on female pelvic MRI

Sam White¹, Steven Knox^{2,3}, Jodie Avery¹, Louise Hull¹, Hu Wang⁴, Yuan Zhang⁴, Gustavo Carneiro^{4,5}, Minh-Son To^{2,4,6}



¹Robinson Research Institute, University of Adelaide, Australia

²South Australia Medical Imaging, SA Health, Australia

³Benson Radiology, South Australia, Australia

⁴Australian Institute for Machine Learning, University of Adelaide, Australia

⁵Centre for Vision, Speech and Signal Processing, University of Surrey, United Kingdom

⁶Flinders Health and Medical Research Institute, Flinders University, Australia



Background

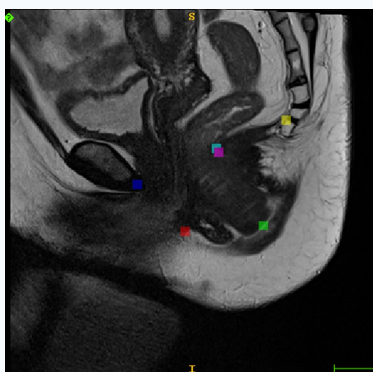
- Morphometric measurements in 2D and 3D assist in diagnosis and assessment of pathology on pelvic MRI. e.g.
 - ⇒ The pubococcygeal line represents the level of the pelvic floor for grading pelvic organ prolapse [1].
 - ⇒ A short anogenital distance is considered a strong diagnostic marker of endometriosis [2].
- Automating extraction of morphometric measurements can standardise and improve the efficiency of reporting.
- The purpose of this study was to develop a deep learning system to automatically perform routine morphometric measurements on female pelvic MRI.

Methods

- 300 sagittal T2 turbo spin echo (TSE) female pelvic MRI scans from the South Australia Medical Imaging (SAMI) database were extracted and split into a training/validation (200 and 50 scans) and test (50 scans) sets.
- Training/validation scans were annotated (SW) with ITK-SNAP [3] to define anogenital distance, pubococcygeal (PC) line, H-line, M-line.
- We utilised and compared image classification networks with various backbones pre-trained on the ImageNet database [4] and compared with landmark detection models based on the U-Net architecture [5].
- Data augmentation to increase training data diversity included rotations ($\pm 30^\circ$), brightness ($\pm 20\%$) and contrast ($\pm 20\%$) scaling.
- Testing was performed by comparing annotations predicted by the models to those made by a Clinical Radiologist (SK).
- Measurement accuracy was assessed by comparing mean absolute error (MAE) between the ImageNet and U-Net models.

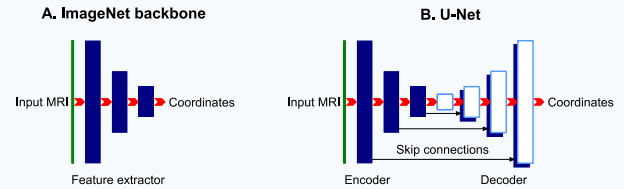
Results

1. Anatomical landmarks

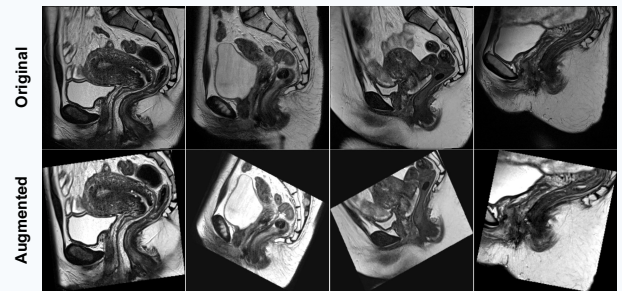


- Fourchette (Red)
- Centre anus (Green)
- Posterior aspect inferior border pubic symphysis (Dark blue)
- Anterior final coccygeal joint (Yellow)
- Posterior aspect anorectal junction (Light blue)
- Where M-line meets pubococcygeal line (Purple)

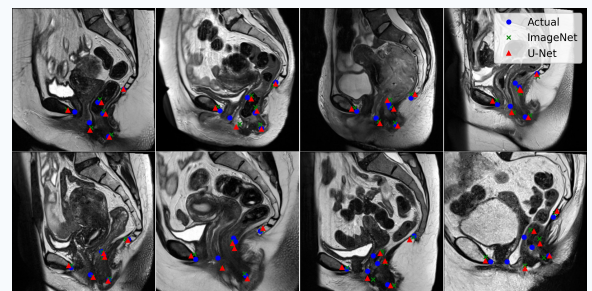
2. Neural network architectures



3. Data augmentation



4. Example predictions



5. Model performance

Landmark	Validation set			Test set		
	ImageNet MAE (mm)	U-Net MAE (mm)	P-value	ImageNet MAE (mm)	U-Net MAE (mm)	P-value
1 Fourchette	4.55	4.82	0.48	6.86	6.57	0.30
2 Cent anus	2.76	1.69	<0.001	3.25	2.25	<0.001
3 Pub symph	2.55	0.90	<0.001	3.32	4.85	<0.001
4 Coccyg jint	2.390	1.89	0.0052	3.02	2.71	0.37
5 Anorect junc	3.18	2.88	0.23	4.71	4.68	0.91
6 M-line, PC	2.37	2.69	0.11	3.25	3.63	0.092

Conclusions

- Deep learning convolutional network models show promising performance in the automated detection and localisation of important anatomical landmarks on female pelvic MRI.
- Validation on external, multi-centre datasets containing a broad range of pathologies will enable deployment in clinical radiology workflows.

References

- Tomsic MV et al, *Appl Radiol*, 2017;46(8):21-27
- Crestani A et al, *Hum Reprod Open*, 2021(1):hoab003
- Yushkevich PA et al, *Annu Int Conf IEEE Eng Med Biol Soc*, 2016:3342-3345
- Deng J et al, *Proc IEEE Conf CVPR*, 2009:248-255
- Ronneberger O et al, *MICCAI 2015*, Lecture Notes in Computer Science, vol 9351

Funding and acknowledgements

This work received funding from the Australian Government through the Medical Research Futures Fund: Primary Health Care Research Data Infrastructure Grant 2020 and from Endometriosis Australia.

Ethics approval for this study was granted by the Southern Adelaide Clinical HREC.

