

RESEARCH

Open Access



Artificial intelligence development for detecting prostate cancer in MRI

Chalida Aphinives^{1*} and Potchavit Aphinives²

Abstract

Background: Artificial intelligence (AI) is the recently advanced technology in machine learning which is increasingly used to help radiologists, especially when working in arduous conditions. Microsoft Corporation offered a free-trial service calling Custom Vision to develop AI for images.

Results: This study included 161 prostate cancer images with 189 lesions from 52 patients. The 160-tag iteration presented the best performance: precision 20.0%, recall 6.3%, mean average precision (M.A.P.) 13.1%, and prediction rate 31.58%. The performance of a 1-h training was better than quick training, but was not different from a 2-h training.

Conclusion: Health personnel can easily develop AI for the detection of prostate cancer lesions in MRI. However, the AI development is further required, and the result should be interpreted along with radiologist.

Keywords: AI, Machine learning, Prostate cancer, MRI

Background

Prostate cancer is the 4th most common cancer in Thai men with 6467 new cases in 2018 or 7.6% of all new cancer cases in Thai men [1]. Prostate cancer is the 2nd most common cancer in both incidence and mortality of men worldwide [2, 3].

Several companies have developed computer-aided detection and diagnosis (CAD) for radiology since the late 1960s. But the real development and systematic research were begun in the early 1980s [4]. Artificial intelligence (AI), which is the recently advanced technology in machine learning, may improve CAD for radiology in clinical practice. In general, AI tasks included automated detection, localization of suspicious lesions, automated diagnostic classification, and prediction of the aggressiveness of cancer from prostate multi-parametric MRI [5–7]. Although AI recently causes concern that machine may replace human in the near future, these fears have occurred periodically among radiologists since the

first development of CAD. Nowadays, CAD and AI have proven their support role for radiologists, especially under arduous condition [8–10].

Microsoft Corporation introduced its cloud platform called Azure supplying over 100 services, some are free-trial and some are always free. The machine learning is a feature-based algorithm of the AI before the advent of deep learning (DL), which is the main algorithm for developing AI for medical imaging. Under the budget-constrained situation in the authors' hospital, an attempt was made to develop AI for detecting cancer lesions within MRI under "custom vision" which is one of the free-trial services from Azure. It was aimed to test the possibility of using this service; therefore, this is a pilot study conducted solely by clinicians with some guidance from one computer scientist.

Methods

This study was approved by the institution Ethics Committee for Human Research based on the Declaration of Helsinki and the ICH Good Clinical Practice Guidelines. No informed consent was needed because this was a

* Correspondence: chalida.aphinives@gmail.com

¹Department of Radiology, Faculty of Medicine, Khon Kaen University, 123 Group 16, Mittraparp Rd, Muang, Khon Kaen province 40002, Thailand
Full list of author information is available at the end of the article

Table 1 The “Performance Per Tag” of 5 iterations with a 1-h training

Dataset (tags)	Precision (%)	Recall (%)	M.A.P. (%)
1. 30	0.0	0.0	34.8
2. 60	0.0	0.0	11.1
3. 100	0.0	0.0	4.8
4. 130	0.0	0.0	4.0
5. 160	20.0	6.3	13.1

retrospective study of stored images in the hospital PACS database.

Radiological images of prostate cancer-proven patients who underwent multi-parameter MRI (mpMRI) during 2 years (2018–2019) were retrieved from the hospital PACS database. The scans were obtained by using the 3T MRI scanner (Achieva®, Philips Health Care) or 1.5T MRI scanner (Aera®, Siemens AG 2012), without endorectal coil. The mpMRI techniques of prostate gland included the following:

- Axial T1W, T2W in whole pelvis, small FOV 48.1 × 36 cm, 8-mm-slice thickness
- Coronal T2W or BTFE, 3-mm-slice thickness
- Thin slice axial, sagittal, and coronal image T2W TSE, large FOV 21.4 × 16 cm, 4-mm-slice thickness
- Diffusion image (B0-800-1000-1500), ADC map in small FOV 24 × 18 cm, 3-mm-slice thickness
- DCE serial dynamic contrast enhancement in small axial FOV, 3-mm-slice thickness

Only 161 prostate cancer images of T2W axial views with 189 hypointense signal lesions (PCa lesion) from 52 patients (PIRADS 4 or 5) were included in this study. All

lesions were located within transition or peripheral zones. All patients were biopsied by core needle instruments via TRUS.

The training processes were divided into 5 iterations of 30, 60, 100, 130, and 160 lesion datasets. The images were uploaded and every lesion manually tagged to help train the object detector. If an image has 3 PCa lesions, it added up to 3 tags in this dataset. After each training for 1 h, this AI was evaluated with testing a dataset from 10 different images that were not included in the training dataset. The testing dataset was composed of 19 PCa lesions.

The system presented the “Performance Per Tag” after the training process into 3 values:

1. *Precision* indicates the fraction of identified images that were correct. For example, if the model recognized lesions in 100 images, and 99 of them were actually had lesions, then the precision would be 99%.
2. *Recall* indicates the fraction of actual images that were correctly recognized. For example, if there actually were 100 images containing lesions, and the model recognized 80 of them, the recall would be 80%.
3. *Mean average precision (M.A.P.)* tells the overall precision of the object detector at finding lesion.

The clinical performance of this AI is presented with the amount and percentage of correct detections among 5 iterations of training.

Another factor that affects the AI performance should be the duration of training. One-hour training was used as a standard training process as previously mentioned. Then, “quick training” and “2-h training” iterations were

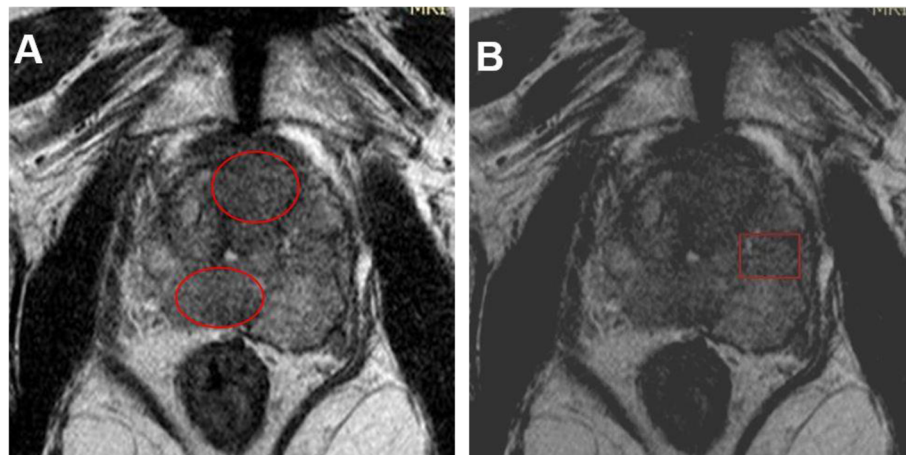


Fig. 1 The false-positive prediction from the 60-tag iteration: **a** radiologist identified 2 lesions at anterior transition zone and right peripheral zone (red circle) in a testing image. **b** AI predicted lesion (red rectangle) at the peripheral zone which was a false position

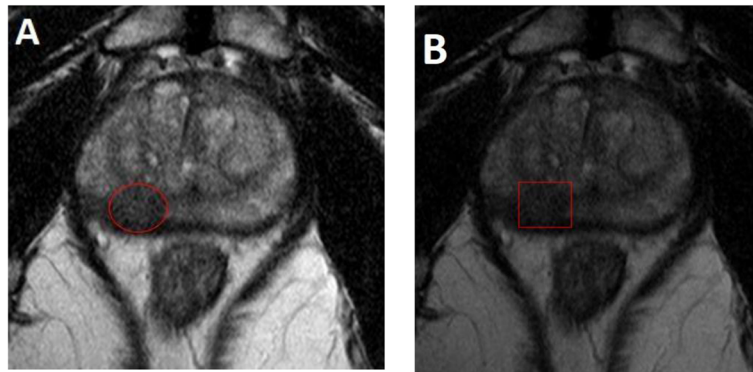


Fig. 2 The true prediction from 100-tag iteration: **a** radiologist identified 1 lesion (red circle) in a testing image. **b** AI predicted 1 lesion (red rectangle) at the true position

performed with the 160 lesion dataset and these performances in both “Performance Per Tag” and clinical performance were compared.

Results

This study included 161 prostate cancer images with 189 PCa lesions from 52 patients. The “Performance Per Tag” of 5 iterations of 30, 60, 100, 130, and 160 tags are presented in Table 1. Ten images with 19 PCa lesions were tested in each iteration. The false-positive prediction from the 60-tag iteration is shown in Fig. 1. The 100-tag iteration showed true-positive predictions in Figs. 2 and 3; however, only one out of three PCa lesions was predicted in Fig. 3. The clinical performance of each training with the same testing dataset (10 images with 19 PCa lesions) is presented in Table 2.

The “Performance Per Tag” was improved from the quick training iteration to the 1-h training iteration, but the 2-h training iteration showed the same values as the 1-h training iteration (Table 3). The clinical performance showed the same results as the “Performance Per Tag” (Table 4).

Discussion

Artificial intelligence (AI) is developed from computer algorithms to simulate intelligent behavior that is capable of learning, reasoning, problem-solving, and self-developing. One of the more sophisticated sets of algorithms is often referred to as deep learning (DL) which is developed from the machine learning (ML). The ML is the ability of an AI to extract information from raw data and to learn from experience [11–14]. Microsoft Corporation provides the free-trial service called “custom vision” which health care personnel can use to develop the AI in their daily practice, especially in radiology. This free-trial service, however, can be regarded as an ML level, while DL needs some additional programming. So DL was not included in this study.

In theory, more learning makes better AI performance, so the “Performance Per Tag” should improve gradually from 30, 60, 100, 130, and 160-tag iterations. Although the 160-tag iteration showed the best performance values, other iterations showed inconsistent values. The clinical performances improved gradually from 30, 100, 130, and 160-tag iterations, except for the 130-tag

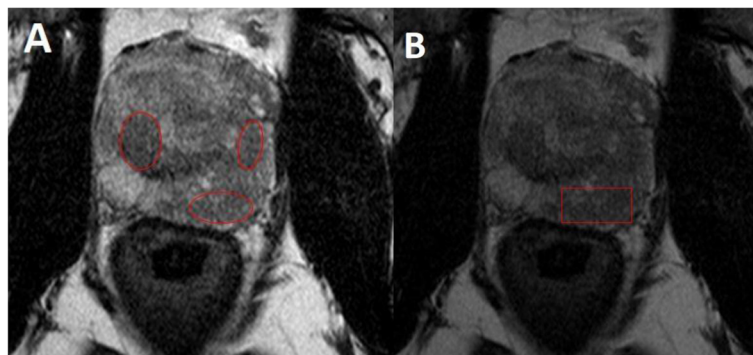


Fig. 3 The partial true prediction from 100-tag iteration: **a** radiologist identified 3 lesions (red circle) in a testing image. **b** AI predicted 1 lesion (red rectangle) at the true position

Table 2 The clinical performance of 5 training datasets

Test/train	30 tags	60 tags	100 tags	130 tags	160 tags
1 (3 lesions)	0	1	1	0	1
2 (1 lesion)	0	0	1	0	1
3 (3 lesion)	0	0	0	0	0
4 (2 lesion)	0	0	0	0	1
5 (1 lesions)	0	0	0	0	1
6 (2 lesions)	0	0	0	0	0
7 (1 lesions)	0	0	0	0	0
8 (1 lesions)	0	1	1	0	1
9 (2 lesions)	0	0	0	0	1
10 (3 lesions)	0	0	0	0	0
Total (19 lesions)	0	2	3	0	6
Percent	0.00	10.53	15.79	0.00	31.58

iteration which showed the results worse than the 100-tag iteration. Many discrete tag varieties, in which each variety had few tag patterns, may confuse the AI on the 130-tag iteration. With more tag patterns, the AI made better clinical performances with the best prediction rate at 31.58%.

The duration of training should affect the performance, and more sophisticated learning needed more time. The 1-h training model made better performances than a quick training model. The 2-h training model, however, was no different in performance from the 1-h training model. With only 160 tags, the AI needed 1 h to experience every pattern thoroughly. One more hour helped the AI learn nothing more. If more images were uploaded, a 2-h training may improve AI performance.

The accuracy and speed of the CAD/AI systems are dependent upon how their algorithms register data and how the system has been trained to learn effect calculation times [10]. The accuracy of detection for prostate cancer using CAD/AI system (43%) was comparable to standard ultrasound-guided biopsy (40%) [15]. Our study used the discrete images to train AI system which was less sophisticated than CAD/AI system, so it was not surprised to achieve low precision (20%) and recall (6.3%) which meant only 6.3 cases would be detected correctly from 100 positive cases. There is a need to perform additional studies with large data sets to improve the performance and impact of this system. Besides

Table 3 The “Performance Per Tag” of 3 different durations of training with a 160 lesion dataset

Duration	Precision (%)	Recall (%)	M.A.P. (%)
1. Quick	0.0	0.0	2.3
2. 1 hour	20.0	6.3	13.1
3. 2 hours	20.0	6.3	13.1

Table 4 The clinical performance of 3 different durations of training with a 160 lesion dataset

Test/train	Quick	1 hour	2 hours
1 (3 lesions)	0	1	1
2 (1 lesion)	1	1	1
3 (3 lesion)	0	0	0
4 (2 lesion)	0	1	1
5 (1 lesions)	0	1	1
6 (2 lesions)	0	0	0
7 (1 lesions)	1	0	0
8 (1 lesions)	0	1	1
9 (2 lesions)	0	1	1
10 (3 lesions)	0	0	0
Total (19 lesions)	2	6	6
Percent	10.53	31.58	31.58

radiologist, other clinicians should use the AI system with utmost consideration.

Conclusion

Health personnel can easily develop AI for the detection of PCa lesion in T2W MRI. AI can predict one third of PCa lesions correctly after training with only 160 images and the free-trial service. However, the AI development is further required, and the result should be interpreted along with radiologist.

Abbreviations

AI: Artificial intelligence; CAD: Computer-aided detection and diagnosis; DL: Deep learning; ML: Machine learning; MRI: Magnetic resonance imaging; mpMRI: Multi-parameter MRI; PACS: Picture archiving and communication system; PCa: Prostate cancer

Acknowledgements

We would like to thank Dr. Thanapong Intharath from Department of Statistics, Faculty of Science, Khon Kaen University, for AI system consultation. We would like to acknowledge Emeritus Professor James A. Will, University of Wisconsin-Madison, for editing the manuscript via Publication Clinic KKU, Thailand.

Authors' contributions

CA: conceptualization, methodology, validation, resources, writing- review and editing, and project administration; PA: methodology, investigation, data curation, and writing—original draft. The authors have read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

All data and material in this study are available for your request.

Declarations

Ethics approval and consent to participate

This study was approved by the Khon Kaen University Ethics Committee for Human Research based on the Declaration of Helsinki and the ICH Good Clinical Practice Guidelines with reference number HE621497. No informed

consent was needed because this was a retrospective study of stored images in the hospital PACS database.

Consent for publication

Not applicable

Competing interests

None.

Author details

¹Department of Radiology, Faculty of Medicine, Khon Kaen University, 123 Group 16, Mittraparp Rd, Muang, Khon Kaen province 40002, Thailand.

²Department of Surgery, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand.

Received: 15 January 2021 Accepted: 18 March 2021

Published online: 26 March 2021

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021. <https://doi.org/10.3322/caac.21660>.
2. Rawla P (2019) Epidemiology of prostate cancer. *World J Oncol*. 10(2):63–89. <https://doi.org/10.14740/wjon1191>
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 68(6):394–424. <https://doi.org/10.3322/caac.21492>
4. Doi K (2007) Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput Med Imaging Graph Off J Comput Med Imag Soc*. 31(4–5):198–211. <https://doi.org/10.1016/j.compmedimag.2007.02.002>
5. Harmon SA, Tuncer S, Sanford T, Choyke PL, Türkbey B (2019) Artificial intelligence at the intersection of pathology and radiology in prostate cancer. *Diagn Interv Radiol Ank Turk*. 25(3):183–188. <https://doi.org/10.5152/dir.2019.19125>
6. Mortensen MA, Borrelli P, Poulsen MH, Gerke O, Enqvist O, Ulén J et al (2019) Artificial intelligence-based versus manual assessment of prostate cancer in the prostate gland: a method comparison study. *Clin Physiol Funct Imaging*. 39(6):399–406. <https://doi.org/10.1111/cpf.12592>
7. Gamito EJ, Crawford ED (2004) Artificial neural networks for predictive modeling in prostate cancer. *Curr Oncol Rep*. 6(3):216–221. <https://doi.org/10.1007/s11912-004-0052-z>
8. Wildeboer RR, van Sloun RJG, Wijkstra H, Mischi M (2020) Artificial intelligence in multiparametric prostate cancer imaging with focus on deep-learning methods. *Comput Methods Programs Biomed*. 189:105316. <https://doi.org/10.1016/j.cmpb.2020.105316>
9. Hu X, Cammann H, Meyer H-A, Miller K, Jung K, Stephan C (2013) Artificial neural networks and prostate cancer—tools for diagnosis and management. *Nat Rev Urol*. 10(3):174–182. <https://doi.org/10.1038/nrurol.2013.9>
10. Nelson CR, Ekberg J, Fridell K (2020) Prostate cancer detection in screening using magnetic resonance imaging and artificial intelligence. *Open Artif Intell J* 6:1
11. Cuocolo R, Cipullo MB, Stanzione A, Ugga L, Romeo V, Radice L et al (2019) Machine learning applications in prostate cancer magnetic resonance imaging. *Eur Radiol Exp*. 3(1):35. <https://doi.org/10.1186/s41747-019-0109-2>
12. McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP et al (2018) Deep learning in radiology. *Acad Radiol*. 25(11):1472–1480. <https://doi.org/10.1016/j.acra.2018.02.018>
13. Suzuki K (2017) Overview of deep learning in medical imaging. *Radiol Phys Technol*. 10(3):257–273. <https://doi.org/10.1007/s12194-017-0406-5>
14. Saba L, Biswas M, Kuppli V, Cuadrado Godia E, Suri HS, Edla DR et al (2019) The present and future of deep learning in radiology. *Eur J Radiol*. 114:14–24. <https://doi.org/10.1016/j.ejrad.2019.02.038>
15. Yang X, Liu C, Wang Z, Yang J, Min HL, Wang L, Cheng KT (2017) Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric MRI. *Med Image Anal*. 42:212–227. <https://doi.org/10.1016/j.media.2017.08.006>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)