Effects of Deep Learning on radiology residents' performance in identifying esophageal cancer on CT

Yasaka et al., British Journal of Radiology, 2023
DOI: 10.1259/bjr.20220685

INTRODUCTION

Esophageal cancer is the mission of this research paper due to its infamous low survival rate (SR)[1], which is mainly because of delayed diagnosis. According to the American Cancer Society (ACS), the 5-year SR for localized esophageal cancer is only 47%[2]. However early detection could significantly improve outcomes, especially if the cancer is caught in its initial stages[2].

Ranked as the 8th most common cancer worldwide in 2020 by the World Health Organization (WHO)[3], esophageal cancer has a strong association with cigarette smoking[1]. While smoking is also a risk factor for other types of cancer, such as lung[4] and pancreatic cancers[5], this presents an advantage: incidental detection. Since CT scans are frequently used in clinics[6], esophageal cancer may be unintentionally identified in its early stage during scans for other conditions.

This is where Deep Learning (DL), specifically Convolutional Neural Networks (CNN), plays a big role in radiology[7]. Studies have shown that DL models can successfully detect, stage, and classify lesions from CT images[8-10]. By building and training these models, authors can improve the process of cancer detection. DL systems are best viewed as decision support services (DSS)[11] that improve accuracy and aid radiologists to make life-saving decisions.

METHODS AND MATERIALS

As shown in Fig1, Yasaka et al. built a deep learning model that spots signs of cancer from a single CT scan. Like usual in machine learning, the data was split into training, validation, and test sets. After training with labeled scans (both with and without cancer), radiologists could use the model to help detect possible cancer in new patient scans.

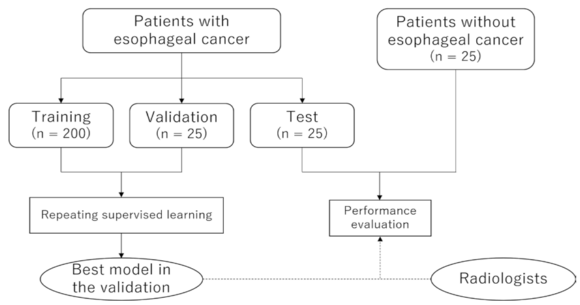

Fig1. Flowchart of the study

CT images were enhanced with contrast to make them easier for the model to understand. More tweaks were done using Python libraries like Pillow and pydicom. The lower parts of the images that didn't show the esophagus were cut out. To add some variety, the images were randomly flipped, rotated, slightly moved, or given noise.

Training and validation ran 15 times in each session, with the number of patients slowly increased by 10 each round. This created 15 models per session. They picked the model with the best results from Session 20 and used that one on the test set.

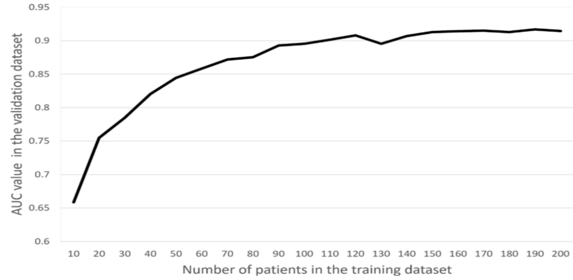

Fig3. Relationship between the training data and validation data

RESULTS AND DISCUSSION

The relation between the number of patients used in the training data and the performance of the CNN in the validation data set is shown in Fig3. As the number of patients in the training data set increased, CNN performance tended to improve. This trend became less obvious when the number of patients in the training data set exceeded 150. In the validation data set, the accuracy was 86%. While less-experienced readers benefited the most in diagnostic accuracy and confidence, a minor drop in specificity was observed for one reader, showing that AI helps but doesn't always improve every aspect equally.

In this paper, authors found that the DL model built may prove useful to the radiologists, specifically the less experienced readers in the field of esophageal cancer. Yasaka et al. managed to increase the R^2-score of the previously made model to detect esophageal cancer on CT images[12] from 84% to 92%. With all that said, it is imperative to remind ourselves that the fundamental role of deep learning models is to provide support to radiologists and radiology residents rather than replacing them in clinical settings.

To wrap it up, there were several limitations to the model that are worth mentioning:

1. Although the training split of the data was relatively small, its performance were capped once we reached about 150 patients, showing signs of diminishing returns and suggesting that adding more data might not improve the model's accuracy any more.

2. The cases with esophageal cancer were clearly visible on the images, meaning the model would probably not do that good noticing the cancer in its earlier stages.

3. Duo to how the images were preprocessed, the model isn't really capable of noticing patterns when patients were in lateral position when skanning.

CONCLUSION

In conclusion, the DL model was helpful mostly for radiologists with lesser experience. There also might be a risk that residents with little experience can lose confidence in diagnosing patients without the expert help of DL model, resulting in an abundance of reliance on artificial intelligence and deep learning technology.

REFERECNES

1. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. Lancet 2017; 390: 2383–96. https://doi.org/10.1016/ S0140-6736(17)31462-9

2. Esophagus Cancer. American Cancer Society Web site. Available from: https://www.cancer. org/cancer/esophagus-cancer.html (accessed 25 Jul 2021)

3. World Health Organization Web site. Available from: https://gco.iarc.fr/today/data/ factsheets/cancers/6-Oesophagus-fact-sheet. pdf (accessed 25 Jul 2021)

4. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. Eur Respir J 2016; 48: 889–902. https://doi.org/10.1183/13993003.00359- 2016

Hidalgo M. Pancreatic cancer. N Engl J Med 2010; 362: 1605–17. https://doi.org/10.1056/ NEJMra0901557

5. Mettler FA Jr, Thomadsen BR, Bhargavan M, Gilley DB, Gray JE, Lipoti JA, et al. Medical radiation exposure in the U.S. in 2006: preliminary results. Health Phys 2008; 95: 502–7. https://doi.org/10.1097/01.HP. 0000326333.42287.a2

6. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O. Deep learning with convolutional neural network in radiology. Jpn J Radiol 2018; 36: 257–72. https://doi.org/10.1148/ radiol.2017170706

7. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S. Liver fibrosis: deep convolutional neural network for staging by using gadoxetic acid-enhanced hepatobiliary phase MR images. Radiology 2018; 287: 146–55. https://doi.org/10.1148/radiol.2017171928

8. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S. Deep learning for staging liver fibrosis on CT: a pilot study. Eur Radiol 2018; 28: 4578–85. https://doi.org/10.1007/s00330- 018-5499-7

9. Yamashita R, Mittendorf A, Zhu Z, Fowler KJ, Santillan CS, Sirlin CB, et al. Deep convolutional neural network applied to the liver imaging reporting and data system (LIRADS) version 2014 category classification: a pilot study. Abdom Radiol (NY) 2020; 45: 24–35. https://doi.org/10.1007/s00261-019- 02306-7

10. European Society of Radiology (ESR). What the radiologist should know about artificial intelligence-an ESR white paper. Insights Imaging 2019; 10: : 44. https://doi.org/10. 1186/s13244-019-0738-2

11. Takeuchi M, Seto T, Hashimoto M, Ichihara N, Morimoto Y, Kawakubo H, et al. Performance of a deep learning-based identification system for esophageal cancer from CT images. Esophagus 2021; 18: 612–20. https://doi.org/10.1007/s10388-021- 00826-0