Can AI generate medical image data with clinical value?

Hazrat Ali, Christer Gronlund, Zubair Shah

Hamad Bin Khalifa University, Doha, Qatar. Umea University, Umea, Sweden.

Summary

Many researchers studied Generative Adversarial Networks (GANs) for producing synthetic CT scans and X-Ray images to improve the performance of AI-based models. It is not well explored how good GANs-based methods performed to generate reliable synthetic data. This work analyzed **43** published studies that reported GANs for synthetic data generation. Many of these studies suffered:
(1) data bias.
(2) lack of reproducibility.
(3) lack of feedback from the radiologists or other domain experts.
(4) lack of reproducibility.
(5) lack of clinical insights.

This work highlights research hotspots in countering the data scarcity problem, identifies various issues as well as potentials, and provides recommendations to guide future research. These recommendations might be useful to improve acceptability for the GANs based approaches for data augmentation as GANs for data augmentation are increasingly becoming popular in the AI and medical imaging research community.

Methods

Databases:
Pubmed, Scopus, IEEEXplore, and Google Scholar
Retrieved studies: 348 studies
Excluded: 208 studies + 81 duplicates.
Excluded: 16 studies in full text reading phase.
Total included studies: 43 studies
Criteria:
GANs for data augmentation of lung CT or X-rays.

Contact

Hazrat Ali
Email: haali2@hbku.edu.qa, hazrat.ali@live.com
www.alihazrat.weebly.com

Key question

Can the studies be translated to clinical applications?
We believe with the existing shortcomings in the studies, GANs based studies are not yet ready to be translated to clinical practice. Nevertheless, despite the shortcomings, the potential impact of using GAN-based methods to improve the training of AI models for COVID-19 diagnosis cannot be denied. Therefore, future transformation to clinical applications is not futile.

Key challenges identified in the studies reporting GANs based data augmentation

Key suggestions

➢ Early feedback from radiologists.
➢ Follow guidelines of CLAIM, assessment tools like PROBAST.
➢ Perform adversarial tests – For example, test on COVID-19 free data.
➢ Facilitate reproducibility.
➢ Use large data.

Conclusion

➢ We identified many areas where the research on using GANs-based methods for data augmentation in COVID-19 could be improved.
➢ We believe that the findings of this analysis of studies reporting GANs-based synthetic data complement findings of existing review by [48] who reported the many pitfalls in the use of machine learning methods for COVID-19, and an earlier study by [49] that identified the risk of bias in the reported AI models.
➢ The analysis in this work will help the readers to understand the limitations of published studies and design better studies in the future to overcome the shortcomings.

References

1 M. Roberts et al., "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans," Nat. Mach. Intell., vol. 3, no. 3, pp. 199–217, 2021.

2 B. G. Santa Cruz, M. N. Bossa, J. Sölter, and A. D. Husch, "Public Covid-19 X- ray datasets and their impact on model bias–A systematic review of a significant problem," Med. Image Anal., vol. 74, p. 102225, 2021.

3 Leveraging GANs for data scarcity of COVID-19: Beyond the hype, CVPR workshop on Generative Models for Computer Viision, June 2023, Canada.