A comparative analysis to assess anatomical illustrations via three AI-Driven Text-to-Image generators

Ayesha Sikandar, Hafiza Moatter Bahar

Department of Biological Sciences, Faculty of Science and Technology, Virtual University of Pakistan

SUMMARY

In recent years the popularity of artificial intelligence (AI) has increased rapidly. While the machine itself can be built for deep learning in order to enhance its own efficiency through sophisticated networks, artificial intelligence (AI) uses machine-learning models to store, calculate, analyze, and even enhance extensive volumes of data that have to be retrieved whenever needed. Text-to-image AI models like Midjourney, Microsoft Bing Image Creator Powered by DALL- E, and Craiyon can generate artistic and impressive images. In this study, anatomical representations of the human ribs, brain, and lungs were produced by evaluating the above three AI-powered text-to-image producers. The generators were evaluated based on how well they represented the basic structure of ribs including correct number of ribs, false and floating ribs, sulcus and gyrus, as well as the structure of the cerebellum and thalamus. Not a single generator generated anatomically correct anatomical structures. It is required to add accurate images to the training databases to increase the accuracy. The study highlights the continued need of human medical illustrators, particularly in guaranteeing the availability of precise and understandable pictures. Improving their accuracy necessitates augmenting the training databases with anatomically correct images to help students.

Keywords: Anatomy – Anatomical structures – AI model – Text to image – Brain – Ribs and Lungs

INTRODUCTION

The aim of this study is to determine the usability of images in education that are produced using artificial intelligence. Educators of anatomy are frequently in the forefront of implementing cutting-edge and novel technology, such artificial intelligence (AI), into their curricula. Hard lessons gained from the application of AI tools in other fields, even though it may present new chances for anatomical teaching (Lazarus et al., 2022). Creation of images through text has made an abrupt change towards becoming a prevailing phenomenon. With the system of text-to-image creation, anyone can create artworks and digital images. This arises the question whether text-to-image synthesis is creative (Oppenlaender, 2022). Current techniques for generating text to images offer a fascinating way to convert between the im-

Corresponding author:
Ayesha Sikandar. Department of Biological Sciences, Faculty of Science and Technology, Virtual University of Pakistan. E-mail: ayesha.sikandar@vu.edu.pk

Submitted: January 15, 2024. **Accepted:** April 1, 2024

https://doi.org/10.52083/DLQJ8889

age and text domains. The generated image fidelity and text relevancy have gradually improved thanks to these techniques, but their applicability and quality are still constrained by a number of critical gaps that need to be filled. (Gafni et al., 2022) The process of creating visuals from written descriptions is known as text-to-image synthesis. Creating images on its own is a difficult task. Text and picture production together increase complexity because the data have to be merged from two distinct modalities. A similar strategy is used in the majority of current efforts on text-to-image synthesis when it comes to neural architectures (Souza et al., 2020). For medical image analysis to be used for both diagnostic and interventional purposes, robust and quick identification of anatomical structures is required. Currently available anatomy detection methods are mostly based on machine-learning approaches that use massively annotated image databases to understand how the collected anatomy should look. The use of less-than-ideal feature engineering approaches and, most significantly, computationally inadequate search algorithms for anatomy detection are the shortcomings of these technologies. (Ghesu et al., 2019).

Medical imaging has made substantial use of AI algorithms for pathology diagnosis, segmentation, and identification of anatomical structures. Research has demonstrated the efficacy of AI-based techniques in the interpretation of radiological images, such as CT scans, X-rays and MRIs for accurate anatomical localization and anomaly detection (Litjens et al., 2017).

Early uses of AI in the medical field mostly concentrates on automating medical tasks; more recently, prognostication and risk prediction have gained attention. Numerous studies look into how AI might help physicians with routine chores, helping with workflow optimization, quantification, and reporting.

Nevertheless, many clinical AI applications are still mostly utilized in research settings and are not yet widely adopted in clinical practice (Greenstein, 2018).

The rapid expansion of information and communication technology has led to substantial advancements in recent years, particularly in the software space. Artificial intelligence is without a doubt one of these advancements. Many artificial intelligence programs can be utilized for educational objectives, even though they can also be used for personal goals like amusement and usefulness. It is evident from a review of the area text that defining artificial intelligence is a challenging task. According to Wang (1995), the concept must be updated on a regular basis, because artificial intelligence is always evolving. According to Encyclopedia Britannica (2022), artificial intelligence is the capacity of a computer or computer-controlled robot to carry out tasks that are normally performed by intelligent entities. Furthermore, it has been suggested that the dimensions of artificial intelligence talents and capacities should take precedence over a precise definition (Schuett, 2019). Due to all of these factors, artificial intelligence is a phenomenon that spans multiple academic disciplines, including biology, linguistics, computer science, neuroscience, anthropology, and philosophy (Luckin et al., 2016). Barillot (1994) focuses on creating a knowledge base with a symbolic model of the anatomical structures of the brain and hypermedia tools for presenting and retrieving information related to the objects in order to enable efficient communication between the user and the system.

MATERIAL AND METHODS

The following prompts were entered during the study: “accurate anatomical illustration of the human ribs”, “accurate anatomical illustration of the human lungs”, and “detailed anatomical illustration of the human brain” into Microsoft Bing Image Creator Powered by DALL- E (https://www.bing.com/create). The Midjourney bot was also fed wiyh the same text prompts using https://discord.gg/midjourney discord service and Craiyon V3 (https://www.craiyon.com/). All the three platforms were accessed on October 30, 2023. These three platforms were chosen because of their popularity, algorithm quality, ease of use and support. The ethical considerations were also kept in mind. For the execution of results (ribs, brain, and lungs), these were chosen because the differences and deficiencies that arise in the anatomical

structures by the use of these image generators could easily be predicted and identified by the educators and students.

RESULTS

According to the results, the image generators can generate aesthetically pleasing anatomical images that could be included in blog posts, presentations, and other materials but did not meet the exact standards of anatomical accuracy. Exploring the anatomical structure of ribs and their connections to the spine and sternum the three generators were discovered to have consistent flaws. Important anatomical structures were missing from the images. The floating ribs (11th and 12th) are missing in all three images generated by the three different AI models (Fig. 1– (1 to 3)); on the other hand, the false ribs (8th,9th and 10th) that indirectly articulate with ribs are not illustrated properly in all three images either. The position of the prominent surface folds and grooves known as gyri and sulci, which are essential to comprehending the functional organization of the brain, were poorly depicted. Additionally, the structure of the cerebellum and the thalamus is poorly depicted, along with the medulla oblongata. Examining the intricate anatomical structure of the lungs, elucidating their bilateral placement within the thoracic cavity, the relationship between lobes, encompassing bronchi, bronchioles, and alveoli, are unclear. The number of cartilages in trachea is wrongly predicted and absent in all the images.

DISCUSSION AND CONCLUSION

Although the AI image generators look very attractive, overall, the analysis of the three AI- powered text-to-image generators showed that their

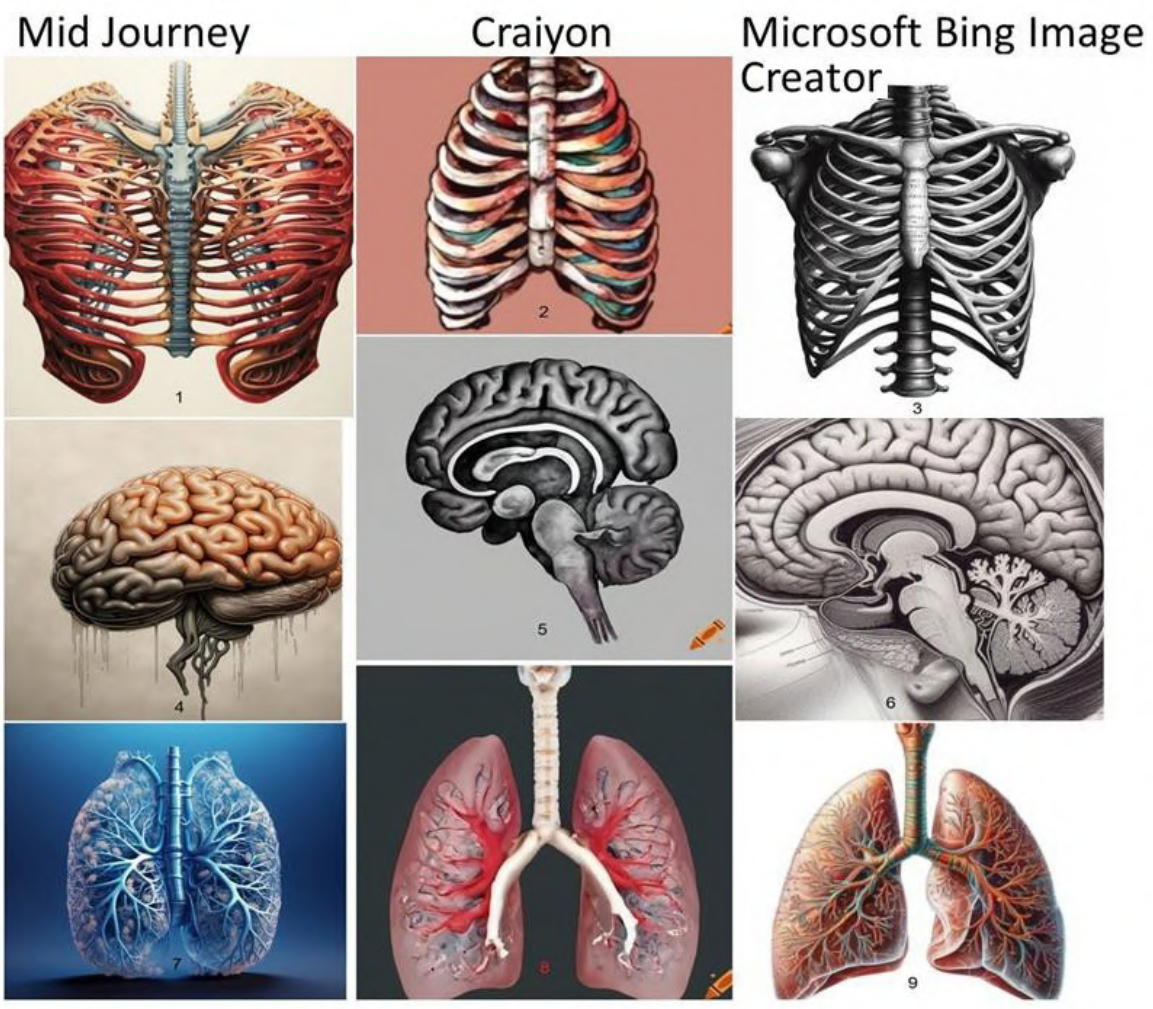

Fig. 1.- Anatomical illustrations of ribs, brain and lungs generated by AI text to image generators: Midjourney (1,4,7), Microsoft Bing Image Creator (3,6,9) and Craiyon (2,5,8).

capacity to produce precise and in-depth anatomical representations was severely limited. A similar study was conducted by Noel (2023); according to him, none of the three generators could provide illustrations that were precise and detailed enough. The author investigated that how well different AI-powered text-to-image generators could produce precise and in-depth anatomical representations of the human heart, brain, and skull. The authors attempt to evaluate these generators' potential for automating the production of medical illustrations by looking at their output, and found out greater limitations in their execution. These flaws demonstrate how AI technology is now unable to replicate the degree of accuracy and precision needed in medical anatomical representations. While AI-powered text-to-image generators have potential for a variety of uses, including medical illustration, it is important to remember that more developments and improvements are needed to guarantee that the illustrations produced meet the standards of anatomical accuracy and detail needed for clinical practice, research, and education in medicine. Medical illustrators are trained to collaborate with others to create purposeful visualizations and to communicate information visually to a specific audience. They are an essential part of the community for anatomical education; their knowledge and contributions are still crucial to the effective delivery of anatomy instruction, and they should not be disregarded or replaced (Cornwall et al., 2023). AI's potential will be regulated with the support of an ethical framework that maintains human control. AI is more than just a cutting-edge technology that needs to be regulated. It is a strong force that is changing settings, interpersonal relationships, and daily routines. It is imperative that this power be used for good purposes in order to ensure the welfare of humanity. In this process, ethics is crucial because it guarantees that AI regulations maximize its promise while minimizing its risks (Taddeo and Floridi, 2018). There are still many unexamined ethical and societal implications of using AI (Goisauf and Abadía, 2022). It is important to recognize that even the most advanced AI image generators have their constraints. Artificial intelligence (AI) image generators frequently produce esoteric graphics rather than accurate representations, which could not always meet the demands of medical illustration (Noel, 2023). The growing prevalence and use of AI in education places a responsibility on educators to learn about the capabilities and limitations of the technology, even if it needs to require new technical skills (Cornwall et al., 2023). However, it is a complex concept that requires a careful balancing of the following factors: the purpose and character of the use and the nature of the copyrighted work.

REFERENCES

BARILLOT C, GIBAUD B, MONTABORD E, GARLATTI S, GAUTHIER N, KANELLOS I (1994) Information system to manage anatomical knowledge and image data about the brain. *In Visualization in Biomedical Computing*, 2359: 424-434.

CORNWALL J, KREBS C, HILDEBRANDT S, GREGORY J, PENNEFATHER P (2023) Considerations on the use of artificial intelligence in generating anatomical images: Comment on "Evaluating AI-powered text-to-image generators for anatomical illustration: A comparative study." *Anat Sci Educ*, 2023, doi: 10.1002/ase.2347.

GAFNI O, POLYAK A, ASHUAL O, SHEYNIN S, PARIKH D, TAIGMAN Y (2022) Make-A-Scene: Scene-Based Text-to-Image Generation with Human Priors. *In European Conference on Computer Vision*, 89-106.

GHESU F, GEORGESCU B, ZHENG Y, GRBIC S, MAIER A, HORNEGGER J, COMANICIU D (2019) Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. *IEEE Trans Pattern Anal Mach Intell*, 41(1): 176-189.

GOISAUF M, ABADÍA M (2022) Ethics of AI in radiology: a review of ethical and societal implications. *Frontiers in Big Data*, 5: 850383.

GREENSTEIN S, GULICK S (2018) Zebra medical vision. *Harvard Business School*, Case 619-014.

LAZARUS MD, TRUONG M, DOUGLAS P, SELWYN N (2022) Artificial intelligence and clinical anatomical education: Promises and perils. *Anat Sci Educ*, 17: 249-262.

LITJENS G, KOOI T, BEJNORDI BE, SETIO AAA, CIOMPI F, GHAFOORIAN M, VAN DER LAAK JA, VAN GINNEKEN B, SÁNCHEZ CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal*, 42: 60-88.

LUCKIN R, HOLMES W, GRIFFITHS M, FORCIER LB (2016) Intelligence unleashed: An argument for AI in education. *UCL Knowledge Lab: London, UK*.

NOEL GP (2023) Evaluating AI-powered text-to-image generators for anatomical illustration: A comparative study. *Anat Sci Educ*, doi: 10.1002/ase.2336.

OPPENLAENDER J (2022) The creativity of Text-to-Image generation. *Proc 25th Int Acad Mindtrek Conference*, pp 192-202.

SCHUETT J (2019) A legal definition of AI. *arXiv preprint arXiv:1909.01095*.

SOUZA D, WEHRMANN J, RUIZ D (2020) Efficient neural architecture for Text-to-Image synthesis. *2020 International Joint Conference on Neural Networks (IJCNN)*, pp 1-8.

TADDEO M, FLORIDI L (2018) How AI can be a force for good. *Science*, 361: 751-752.

WANG P (1995) On the working definition of intelligence. *Technical Report 94, Center for Research on Concepts and Cognition, Indiana University*, pp 1-32.