

ChatGPT and medical education: a new frontier for emerging physicians

ChatGPT et l'éducation médicale : une nouvelle frontière pour les médecins émergents

Ethan Waisberg,^{1,2} Joshua Ong,³ Mouayad Masalkhi,⁴ Nasif Zaman,⁵ Sharif Amit Kamran,⁵ Prithul Sarker,⁵ Andrew G Lee,^{6,7} Alireza Tavakkoli⁵

¹Department of Ophthalmology, University of Cambridge, Cambridge, United Kingdom; ²Moorfields Eye Hospital, NHS Foundation Trust, London, United Kingdom; ³Department of Ophthalmology and Visual Sciences, University of Michigan Kellogg Eye Center, Michigan, USA; ⁴University College Dublin School of Medicine, Belfield, Dublin, Ireland; ⁵Human-Machine Perception Laboratory, Department of Computer Science and Engineering, University of Nevada, Nevada, USA; ⁶The Houston Methodist Research Institute, Houston Methodist Hospital, Texas, USA; ⁷Center for Space Medicine, Baylor College of Medicine, Texas, USA

Correspondence to: Ethan Waisberg, University of Cambridge, Cambridge, United Kingdom; email: ethan.waisberg@gehkl.nhs.uk

Published ahead of issue: Aug 23, 2023; CMEJ 2023 Available at <https://doi.org/10.36834/cmej.77644>

© 2023 Waisberg, Ong, Masalkhi, Zaman, Sarker, Lee, Tavakkoli; licensee Synergies Partners. This is an Open Journal Systems article distributed under the terms of the Creative Commons Attribution License. (<https://creativecommons.org/licenses/by-nc-nd/4.0>) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited.

Introduction

Since its release in late November 2022, ChatGPT (Open AI, USA) has rapidly gained attention for its detailed answers and human-like writing ability (Figure 1).¹ GPT refers to a generative pre-trained transformer architecture, in which a transformer neural network generates text. ChatGPT is a large language model (LLM) which generates human-like text from deep-learning techniques, in which a large dataset is analyzed to infer the relationship of words. ChatGPT has already been capable of tasks such as generating artificial intelligence (AI algorithms), or generating images based on patient descriptions of complex neuro-ophthalmic visual phenomena.²

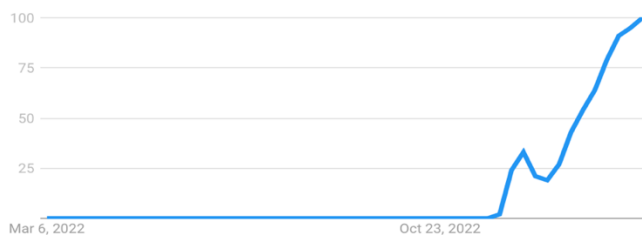


Figure 1. Worldwide Google searches for ChatGPT. Since its November 2022 release, searches for ChatGPT and its uses have grown exponentially.

AI is any technique leveraging machines to mimic human intelligence.³ AI technologies are rapidly evolving and are expected to revolutionize the field of medicine. AI can make medicine more efficient and safer when applied correctly: such as with precision medicine, improved diagnostic imaging or preventing medication errors. As these technologies reach clinical use, skills to interpret and use AI in a medical setting will become essential for doctors. The World Medical Association now advocates for the medical curriculum to foster an improved understanding of healthcare AI for medical students.⁴ This knowledge will allow for a more rapid and informed implementation of AI technologies in the future. As well, physicians must be informed for the potential biases and errors that AI algorithms can produce and how to mitigate these effects.

We asked ChatGPT to advise on how to build an AI to classify optical coherence tomography (OCT) images (Figure 2). As a combined health and computer science team that conducts NASA-funded machine learning research, we examined this recommendation.⁵⁻⁷ The general recommended steps to design an AI to classify OCT images are correct, which includes: collecting the OCT images, preprocessing data, labelling, splitting, training, testing, refining and deploying the AI.



Figure 2. Generated from ChatGPT from the text prompt “how can I build an AI to classify OCTs.”

We then asked ChatGPT to code an AI to classify mammography scans (Figure 3).

```
python Copy code

import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense

# Define the CNN model architecture
model = Sequential()
model.add(Conv2D(32, (3, 3), activation='relu', input_shape=(224, 224, 3)))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(256, (3, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dense(1, activation='sigmoid'))

# Compile the model
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

# Load the dataset of mammography scans and split it into training and validation sets
# You will need to replace this with your own dataset
train_data = tf.keras.preprocessing.image_dataset_from_directory('path/to/train/directory',
                                                               image_size=(224, 224),
                                                               batch_size=32)
val_data = tf.keras.preprocessing.image_dataset_from_directory('path/to/validation/directory',
                                                             image_size=(224, 224),
                                                             batch_size=32)

# Train the model on the dataset
model.fit(train_data, epochs=10, validation_data=val_data)

# Use the model to predict breast cancer in a new mammography scan
# You will need to replace this with your own mammography scan
new_scan = tf.keras.preprocessing.image.load_img('path/to/new/scan', target_size=(224, 224))
new_scan = tf.keras.preprocessing.image.img_to_array(new_scan)
new_scan = new_scan.reshape(1, 1) + new_scan.shape
prediction = model.predict(new_scan)
print(prediction)
```

Figure 3. Generated from ChatGPT from the text prompt “write code for an AI to analyze mammography scans.”

ChatGPT coded a convolutional neural network with two fully connected layers and four convolutional layers to analyze mammography images and predict whether breast cancer is present in a scan. While this is a sample code and would need further adaptations prior to being implemented clinically, this represents an amazing start for a medical student’s AI journey.

ChatGPT has certain restrictions and drawbacks as an AI language model. Since a large corpus of material, including web pages, books, and other sources, was used to train ChatGPT, it may provide replies that are completely or partially identical to already published writings.

Furthermore, ChatGPT may not have been trained with the most recent data and may provide replies that are inaccurate or outdated. It is also possible that the model will not be able to comprehend the context of a query or discussion, which might result in misunderstandings and mistakes.

Other ChatGPT pitfalls have also been outlined in a previous study; those include self-plagiarism when asked a question many times, and a high degree of direct or “word-for-word” plagiarism from internet sources such as Wikipedia and LinkedIn.¹ A number of actions that may be taken in order to prevent or lessen these pitfalls. This includes validating the results obtained from ChatGPT by comparing it with the most up-to-date clinical guidelines/resources, giving more context when posing questions to or conversing with ChatGPT so that the model can better comprehend the context of the query.

Finally, despite the efforts to increase AI teaching worldwide, artificial intelligence is yet to be incorporated in the medical education process. In a study examining the exposure to artificial intelligence in Canadian medical education, 85% of respondents indicated that there was no formal educational opportunities regarding AI.⁸ All things considered, ChatGPT can potentially be a useful tool in medical education to improve students’ understanding of AI, but employing it beyond educational purposes should be approached with caution.

Conflicts of Interest: The authors declare no conflicts of interest.

Funding: NASA Grant [80NSSC20K183]: A Non-intrusive Ocular Monitoring Framework to Model Ocular Structure and Functional Changes due to Long-term Spaceflight.

References

1. Alser M, Waisberg E. Concerns with the usage of ChatGPT in academia and medicine: a viewpoint. *Am J Med Open*. Feb 2023;100036. <https://doi.org/10.1016/j.ajmo.2023.100036>
2. Waisberg E, Ong J, Masalkhi M, et al. Text-to-image artificial intelligence to aid clinicians in perceiving unique neuro-ophthalmic visual phenomena. *Ir J Med Sci*. Feb 14, 2023. <https://doi.org/10.1007/s11845-023-03315-8>
3. Waisberg E, Ong J, Paladugu P, et al. Challenges of artificial intelligence in space medicine. *Space Sci Technol*. 2022;2022:1-7. <https://doi.org/10.34133/2022/9852872>
4. World Medical Association. *WMA statement on augmented intelligence in medical care*. Oct 2019. Available from <https://www.wma.net/policies-post/wma-statement-on-augmented-intelligence-in-medical-care/> [Accessed Feb 28, 2023].
5. Waisberg E, Ong J, Kamran SA, et al. Transfer learning as an AI-based solution to address limited datasets in space medicine. *Life Sci Space Res*. 2023;36:36-38. <https://doi.org/10.1016/j.lssr.2022.12.002>
6. Waisberg E, Ong J, Zaman N, Kamran SA, Lee AG, Tavakkoli A. A non-invasive approach to monitor anemia during long-duration spaceflight with retinal fundus images and deep learning. *Life Sci Space Res*. 2022;33:69-71. <https://doi.org/10.1016/j.lssr.2022.04.004>
7. Waisberg E, Ong J, Paladugu P, et al. Advances in machine learning to detect preventable causes of blindness. *Eye*. Dec 10, 2022. <https://doi.org/10.1038/s41433-022-02354-2>
8. Pucchio A, Rathagirishnan R, Caton N, et al. Exploration of exposure to artificial intelligence in undergraduate medical education: a Canadian cross-sectional mixed-methods study. *BMC Med Educ*. 2022;22(1):815. <https://doi.org/10.1186/s12909-022-03896-5>

Published ahead of ISSN