The potential of AI text to video generation in medical education for neurologists

UCL Education Conference 2024 Subtheme: The modern and future Higher Education classroom

Noronha A[1], Davis D[1], Rahman T[1], Jain D[1], Ogundiya O[1], Vijayanathan J[1], Khusakul P[1], Fleming F[1], Aldridge C[1], Valnarov-Boulter I[1], Malik Z[1], Foakes A[3], Young TM[2]

[1] Medical student on iBSc Clinical Neurology & Brain Sciences degree at Queen Square Institute of Neurology, UCL

[2] Course director of Clinical Neurology & Brain Sciences iBSc at Queen Square Institute of Neurology, UCL (Corresponding author: t.young@ucl.ac.uk)

[3]Senior administrator, Clinical Neurology by distance learning MSc/PG Diploma/ PG Cert at Queen Square Institute of Neurology, UCL

Figure 1: The Queen Square Institute of Neurology (on the left of image)

Introduction to AI generated images part 1:

- Medical images are fundamental in medical education, allowing not only clear illustrations of specific medical conditions, but also aiding knowledge transfer through multimedia delivery.
- Traditional use of medical images in medical education has been based on patient photographs. However, such an approach has a number of potential limitations including the important aspects of obtaining adequate consent and confidentiality (Hill 2006).
- A potential alternative approach to traditional photography has emerged through generative tools using artificial intelligence (AI)
- The rapid introduction of artificial intelligence (AI) has impacted many of areas of professional life, including medical education.
- One key area within AI has been that of tools that allow rapid production of novel images from text prompts.

Introduction to AI generated images part 2

- One of our co-authors (Prof Young) has recently published an example of how such AI text to image tools can be used to produce photographic quality novel images which could be used in medical education (Figure 2):

Figure 2: A novel image to illustrate clinical features of a common medical condition (hypothyroidism) produced using an AI text to image generation tool not based on a patient (Kumar et al 2024)

- Whilst such still medical images may prove to be very useful in medical education, increasingly education itself has relied more on the use of videos in addition to – or instead of - still images.
- In our presentation we consider how AI might potentially play an important role in the generation of future medical education videos.

The use of non-AI generated videos in neurology education: evidence base

- **Video-based training improves the accuracy of seizure diagnosis**
- (Seneviratne et al. 2014) This study looked at improving 103 medical students' ability to identify the diagnostic differences between epileptic seizures and psychogenic seizures. Twenty video clips of unknown length were shown once to each student in the first stage of the study (Seneviratne et al. 2014). In the second stage, a single one-hour lecture was given on diagnosis and classification of ES and PNES. No videos were used in the third or fourth stages. In three months, medical students achieved results comparable to that of emergency medicine trainees. However, it remains uncertain to what extent students worked outside of these teaching sessions and there was a high (>50%) attrition rate in this study.

- **The effects of video instruction on a neuroscience intensive care unit nursing skills in case presentations and neurological examinations**
- 70 nurses working in a neuroscience intensive care unit ward were assessed on their ability to present clinical details of a patient as well as conduct a neurological examination and given a score out of 100 based on the examiner's consensus (Lau et al. 2021). They were split up into a control group (n = 37) and intervention group (n = 33). A 5-minute video was given to the intervention group and 3 weeks later the intervention group score increased from 54 to 60, whilst the control group remained the same at 67 (Lau et al. 2021). Due to a lack of objective ways of measuring performance, human or unconscious bias could affect the results of this investigation.

Existing use of non-AI generated videos in neurology education

- **Searches were conducted on the open access web pages of some key national and international neurology websites for the following associations : Association of British Neurologists, The American Academy of Neurology , American Neurological Association (ANA), World Federation of Neurology and Association of British Neurologists**
- **The number , length and content of neurology teaching videos which were available to the public was determined . The following were the results found:**
- Association of British Neurologists : Yielded 31 videos ranging from 5 minutes to 1 hour 6 minute in length. The majority of the videos were between 40-50 minutes. Topics as diverse as stroke, headache and central nervous system infections were covered.
- The European Academy of Neurology : 74 videos were found, with the longest being a conference recording which lasted 3 hours and 41 minutes, and the shortest being a video about brain health which was 3 minutes long. 81 podcasts had also been published, although these required subscriptions to access.
- The American Academy of Neurology: There were ten videos that were available on the website, ranging from 51 minutes to 1 hour 1 minute, the majority of which were around 1 hour long. These covered topics such as epilepsy , pain medicine and vascular neurology.
- American Neurological Association (ANA): 5 educational videos were identified ranging from 3 minutes to 1 hour 7 minutes, covering topics such as stroke therapeutics , autoimmune CNS disorders and CSF venous fistulas. 27 podcasts were also identified which were around 20 minutes long each.
- World Federation of Neurology : 91 videos were identified , ranging from 3 minutes to 1 hour 43 minutes in length- with the majority of videos being under an hour. These covered topics such as the global burden of neurological diseases, management strategies for dementia and stroke prevention.
- **Overall, 211 videos were identified across all websites , their lengths ranged from 3 minutes to 3 hours and 41 minutes. These demonstrate the widespread current use of videos for neurology education.**

AI text to video generation

- AI text to video generation arises from the concept of generative AI which itself can be traced back to the creation of the first chatbot called Eliza in 1966, used to simulate conversations with a psychotherapist (Coheur 2020).

- Since then, the development of neural networks, natural language processing, computer vision, and machine learning have allowed other forms of generational AI to be created such as AI text to video (White 2023).

- Notably, the use of Generative Adversarial Networks(GANs) since 2014 facilitates the generation of realistic outputs by pitting a generator against a discriminator. This adversarial training forces the generator to improve(Goodfellow et al 2014).

- Whilst GANs were initially used to generate realistic text responses, Diedrik Kingma and Max Welling introduced variational autoencoders to have images and videos also as output (Kingma and Welling, 2019)

- The intersection of AI text generation with computer vision techniques has been pivotal in enhancing the visual quality of generated videos. Models like CLIP (Contrastive Language-Image Pre-training) , introduced by OpenAI in 2021, demonstrated the power of joint training large-scale datasets containing both text and image pairs, improving the model's understanding of textual descriptions in relation to visual content(openai.com, 2021).

- Determining the first text to video generator is difficult due to incremental advancements in the field. However, a notable early example that laid the foundation for subsequent developments is NVIDIA's Vid2vid, introduced in 2018 (Mitrev 2018)

The use of AI generated videos in medical education

- Whilst we are focussing on neurology education videos, when it comes to AI text to video production, there is sparse literature to date on its use in *any* aspect of medical education.

- Some potential barriers to the use of such techniques are apparent-the scientific journal Nature has stated that it would not publish work (image or video) produced by AI (Nature 2023).

- Despite this, the potential of AI generation and video combinations have been demonstrated to be effective in surgical simulations (Perumalla et al 2023).

- We therefore sought to explore the potential of AI text to video generation for neurology education.

Aims

- We set out to use widely available current AI text to video generative tools to produce short neurology education videos.
- Unlike with traditional lecture lengths in universities (close to an hour's duration), it has been shown that the most effective period of the human attention span lasts only 10 to 15 minutes (Stuart and Rutherford, 1978; Eze and Edward, 2017).
- Short duration videos have proof of effectiveness in science knowledge transfer (Nong et al 2022).
- We therefore aimed for short videos based around neurology education.
- Two common topics for neurology education videos were chosen: Important headache types to identify in clinical practice ('red flag' headaches) and education on the neurology examination

Methods

- Relevant literature searches were performed by our study group for use throughout this presentation.
- Regarding the text to video tools, Google searches were initially performed to identify widely available AI text to video generative tools.
- These tools were then investigated further by our study group with sample videos made where free options were available.
- Samples of these tools were then used in this presentation to produce proof of concept videos which we will show in this presentation. Where only free registration was used, videos were only included in this presentation where additional approval of use was obtained from the relevant company.

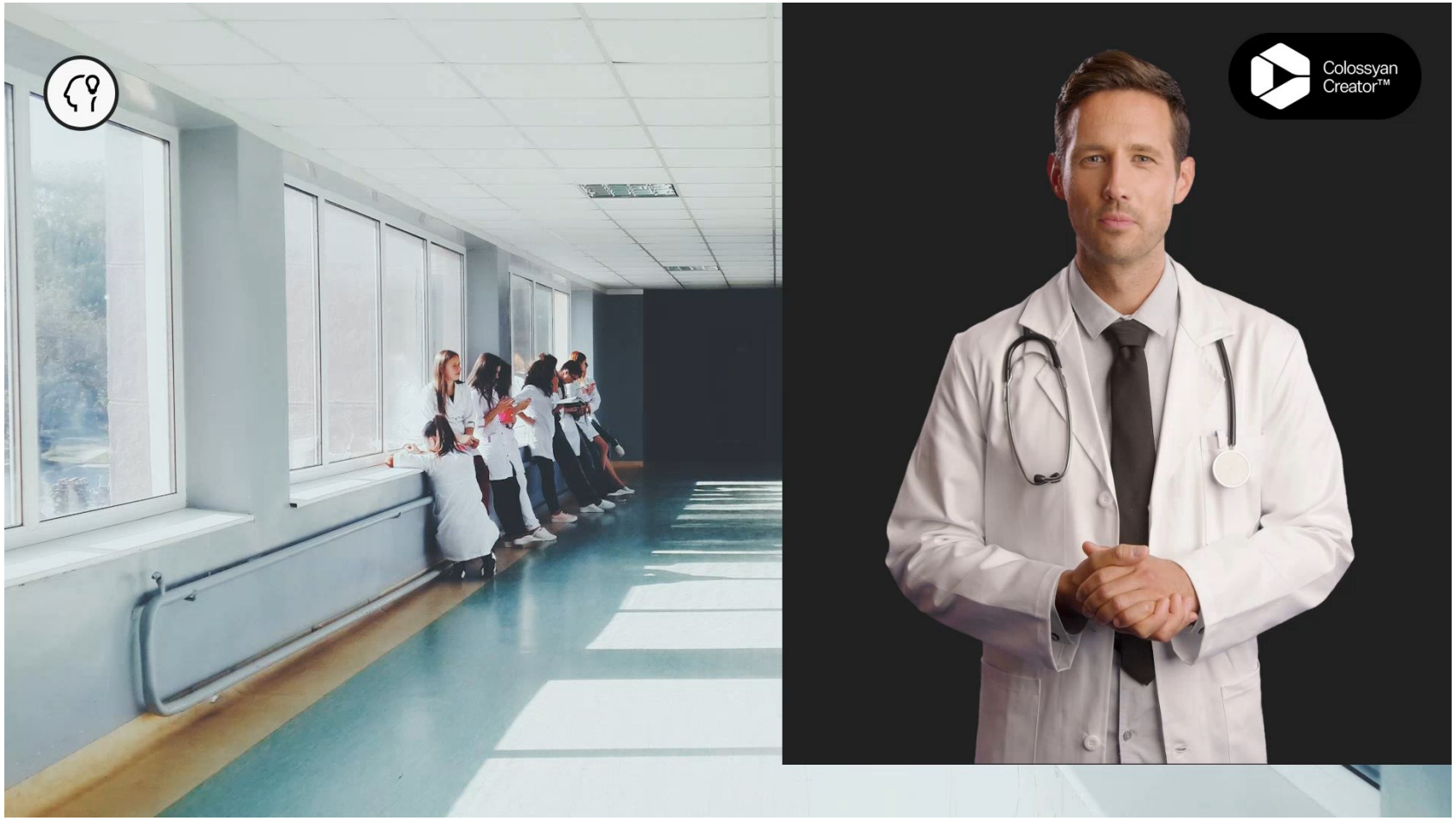
Results:

- We identified 17 AI text to video generation tools.
- With a sample of these we generated short neurology videos which we will show during our presentation
- Most tools included a free initial registration option with which a limited number of free videos could be made (typically with watermark present).
- Most tools gave the option of an avatar (with different genders and ethnicities being available. The voice could be chosen from a variety of accents.
- Some tools allowed for generation of a short video from a brief text prompt. In general, we found greater accuracy was obtained by inputting more detailed script (which could be generated from Chat GPT).
- All but one of the AI text to video generation tools that we trialled utilised short sections of stock conventionally produced video which was then combined with sound as still images as needed to create a continuous video. The linking of such stock videos was not always smooth.
- Limitations included lack of variety of both stock videos and avatar voices as well as some jumping of sound and video shots in places.

Results: AI Text to video generation using the tool: Video gen with permission to display for this conference

Introduction to Neurology text to video amended.mp4

Results: AI Text to video generation using the tool: Colossyan with permission to display for this conference



https://app.colossyan.com/share/4lwtp5xq

Results: AI Text to video generation using the tool: Colossyan and also Synthesia with permission to display for this conference

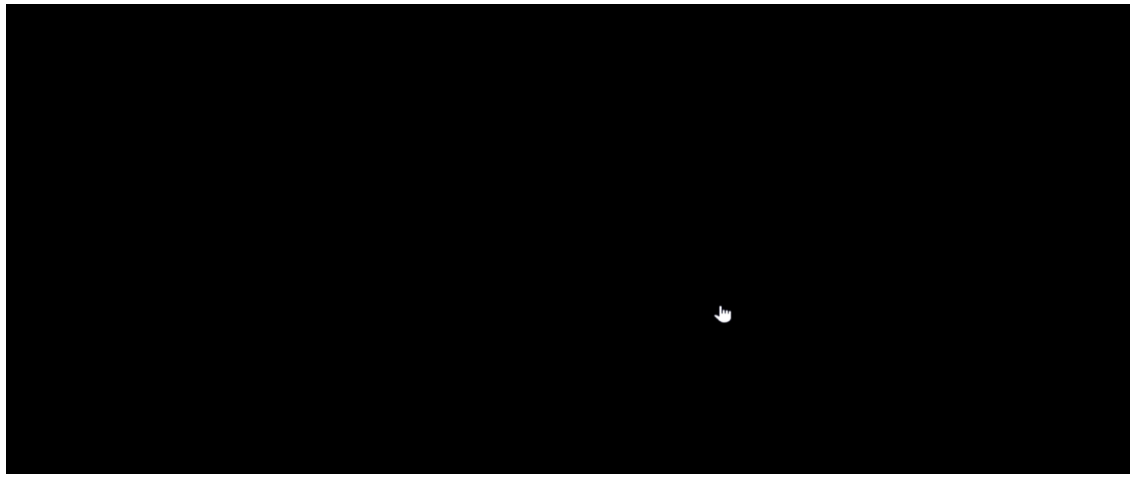
Our video example links:

https://app.colossyan.com/share/e9tgboe2

https://share.synthesia.io/2fc1de13-5655-433e-b8e0-b81f9ba89690

Results: AI Text to video generation using the tool: Midjourney; usage here allowed within terms of paid subscription

- This is the only example we are able to make of a text to novel video (not using stock videos). Whilst clearly being much more limited in terms of length and detail this approach does have the potential advantage of producing a wider range of novel material if there is further development.



Text Prompt to Midjourney 19/1/2024: ‘Patient walking with a walking stick photographic hyperrealistic quality ‘

Midjourney alpha iteration of version 5; Publisher Midjourney via Discord URL of the AI system. https://www.midjourney.com/home/

https://cdn.midjourney.com/6a60ca78-caf7-4620-a674-1e78d8c88802/video.mp4

Conclusion

- AI generative tools to produce novel photographic images from text inputs have been a widely available for a number of years. Despite this, use of this striking new technology is only just beginning to be explored for medical education.
- AI text to video tools appear to be at an earlier stage than for AI text to image tools. In keeping with this there is currently little in the literature to show evidence of the use of such AI generated videos in medical education.

- Neurology is often viewed a s a difficult-even intimidating –subject by learners and yet remains important to teach not only because neurology conditions are common but also because neurology is a shortage specialty

- We explored in this presentation several widely available AI text to video generative tools and used these to produce proof of concept short videos on neurology education.

- There are a number of potential challenges for AI generated images and videos including copyright and the need for greater clarity on use of the original images used to train the tools. There is also the potential risk that preexisting stereotypes might be perpetuated.
- However, given the ease with which these videos were produced we believe that this demonstrates a great potential future resource, not only for neurology education, but for medical education-and indeed education-more widely.

Thank you…

We are very grateful to Colossyan, Synthesia and Video gen for not only allowing us to present videos using a free account, but also in doing so in a very encouraging manner.

References

- Coheur, L. (2020). From Eliza to Siri and Beyond. *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp.29–41. doi:https://doi.org/10.1007/978-3-030-50146-4_3.
- Eze, C. and Misava, E. (2017). Lecture duration: A risk factor for quality teaching and learning in Higher Education. Integrity Journal of Education and Training, [online] 1, pp.1–5. Available at: https://ecommons.aku.edu/eastafrica_fhs_fhs/12/.
- Goodfellow, I., et al. (2014). "Generative Adversarial Nets." In Advances in Neural Information Processing Systems (NeurIPS). Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2 December 2014 Pages 2672–2680
- Kingma, D.P. and Welling, M. (2019). An Introduction to Variational Autoencoders. Foundations and Trends® in Machine Learning, 12(4), pp.307–392. doi:https://doi.org/10.1561/2200000056.
- Kumar A, Burr P, Young TM. Using AI Text-to-Image Generation to Create Novel Illustrations for Medical Education: Current Limitations as Illustrated by Hypothyroidism and Horner Syndrome. JMIR Med Educ. 2024 Feb 22;10:e52155. doi: 10.2196/52155. PMID: 38386400; PMCID: PMC10921331.
- Hill K. Consent, confidentiality and record keeping for the recording and usage of medical images. J Vis Commun Med. 2006 Jun;29(2):76-9. doi: 10.1080/01405110600863365. PMID: 16928590.
- Lau, K.H.V., Hamlyn, E., Williams, T.J., Qureshi, M.M., Mak, K., Mian, A., Cervantes-Arslanian, A., Zhu, S. and Takahashi, C. (2021). The Effects of Video Instruction on Neuroscience Intensive Care Unit Nursing Skills in Case Presentations and Neurological Examinations. *Journal of Neuroscience Nursing*, [online] 53(3), pp.129–133. doi:https://doi.org/10.1097/JNN.0000000000000591.

References continued

- Mitrev, D. (2018). Vid2Vid Explanation - Conditional GANs for Video-to-Video Synthesis. [online] Available at: https://neurohive.io/en/state-of-the-art/vid2vid-conditional-gans-for-video-to-video-synthesis/.
- Nature. Why Nature will not allow the use of generative AI in images and video. 2023 Jun;618(7964):214. doi: 10.1038/d41586-023-01546-4. PMID: 37286651.
- Nong, L., Liao, C., Ye, J.H., Wei, C., Zhao, C., Nong, W., 2022. The STEAM learning performance and sustainable inquiry behavior of college students in China. Front Psychol 13. https://doi.org/10.3389/FPSYG.2022.975515
- Perumalla C, Kearse L, Peven M, Laufer S, Goll C, Wise B, Yang S, Pugh C. AI-Based Video Segmentation: Procedural Steps or Basic Maneuvers? J Surg Res. 2023 Mar;283:500-506. doi: 10.1016/j.jss.2022.10.069. Epub 2022 Nov 24. PMID: 36436286; PMCID: PMC10368211.
- openai.com. (2021). CLIP: Connecting text and images. [online] Available at: https://openai.com/research/clip.
- Seneviratne, U., Ding, C., Bower, S., Craig, S., Leech, M. and Phan, T.G. (2014). Video-based training improves the accuracy of seizure diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, [online] 85(4), pp.466–470. doi:https://doi.org/10.1136/jnnp-2013-306618.
- Stuart, J. and Rutherford, R.J.D. (1978). MEDICAL STUDENT CONCENTRATION DURING LECTURES. The Lancet, [online] 312(8088), pp.514–516. doi:https://doi.org/10.1016/S0140-6736(78)92233-X.
- White, M. (2023). A Brief History of Generative AI. [online] Medium. Available at: https://matthewdwhite.medium.com/a-brief-history-of-generative-ai-cb1837e67106