



NewsLetter

自治医科大学 地域医療オープン・ラボ

2024
JUL
特別号

ChatGPT の耳鼻科領域への活用 「生成 AI の耳鼻咽喉科領域に対する精度、有効性」について明らかに

自治医科大学とちぎ子ども医療センター小児耳鼻咽喉科部門の野田昌生助教、甲州亮太大学院生、伊藤真人教授らは生成 AI の耳鼻咽喉科領域に対する有用性について検証を行いました。

今回その研究成果が「JMIR Medical Education」誌、「JMIR AI」誌に掲載されましたので、野田助教に研究内容について伺いました。

論文名: Performance of GPT-4V in Answering the Japanese Otolaryngology Board Certification Examination Questions: Evaluation Study

著者名: Masao Noda, Takayoshi Ueno, Ryota Kosu, Yuji Takaso, Mari Dias Shimada, Chizu Saito, Hisashi Sugimoto, Hiroaki Fushiki, Makoto Ito, Akihiro Nomura, Tomokazu Yoshizaki

掲載誌: 「JMIR Medical Education」 2024 Mar 28;10:e57054. doi: 10.2196/57054..

論文名: Feasibility of Multimodal Artificial Intelligence Using GPT-4 Vision for the Classification of Middle Ear Disease: Qualitative Study and Validation

著者名: Masao Noda, Hidekane Yoshimura, Takuya Okubo, Ryota Kosu, Yuki Uchiyama, Akihiro Nomura, Makoto Ito, Yutaka Takumi

掲載誌: 「JMIR AI」 JMIR AI 2024;3:e58342 doi: 10.2196/58342

Q1. 本研究内容について教えてください。

本研究では、生成 AI を用いて二つの研究系として耳鼻咽喉科領域での精度検証を行いました。

まず、日本耳鼻咽喉科専門医試験に対する GPT-4Vision(GPT-4V) のパフォーマンスを検討しました。2023 年度の耳鼻咽喉科専門医試験に含まれる全 100 問（画像問題を含む）に対する GPT-4V の回答能力と精度を評価しました。

テキストのみの入力では平均正解率が 24.7%でしたが、英訳と開発したプロンプトの追加により 47.3%まで有意に向上しました ($P<.001$)。テキストのみの入力における非応答率は平均 46.3%でしたが、英訳と開発したプロンプトの追加で 2.7%まで低減しました ($P<.001$)。画像問題を含む問題では、テキストのみの入力に比べて、テキストと画像の入力が正解率を向上させました。ただし、画像ベースの問題の正解率はテキストベースの問題よりも低く、GPT-4V の現段階での改善の余地が示唆されました。特に、画像解析能力に関しては、さらなる研究と開発が必要とされることがわかりました。

さらに、GPT-4V による中耳炎診断における有用性を検討するために、鼓膜画像と患者情報に対する GPT-4V のパフォーマンスを検証しました。主な中耳疾患（急性中耳炎、滲出性中耳炎、真珠腫性中耳炎、慢性穿孔性中耳炎）の 190 データに対する GPT-4V の回答能力と精度を評価しました。

鼓膜画像と患者情報を入力するマルチモーダル AI によって、診断分類精度は 82.1%を達成しました。疾患によって精度の高いものや低いものがあり、疾患特異性があることがわかりまし

た。また、医師との比較では、耳鼻咽喉科医よりは劣るものの、耳鼻咽喉科ではない医師よりも高い精度であり、特に全体の正答率の低い症例に対して有効であることがわかりました。

Q2. なぜこのような研究を行ったのですか？

2022 年より生成 AI の開発が進み、その活用が急激にすすんでいます。2023 年には ChatGPT によってアメリカの医師国家試験(USMLE)や司法試験で合格レベルに達したという報告がされて以降、ChatGPT の医療・ヘルスケア分野での活用可能性が世界的に大きな注目を浴びるようになりました。さらに、2023 年 9 月より GPT-4Vision が発表され、画像認識もできることが報告されました。

その中で、実際の耳鼻科診療における活用方法については十分にわかっていないため、使ってみることで検討しようと考えました。AI というとよくわからないイメージのある方もおられると思いますが、日常生活にも普及し始めており、上手に活用していくことが求められていると感じています。

どんなふうに使えて、何が課題なのかがわかると、具体的な活用方法も明確になってきます。

Q3. この研究が教えてくれることと今後の課題について教えてください。

今回の研究では AI の医学分野における応用の研究として、耳鼻咽喉科領域で一定の精度を実現することがわかりました。さらに、中耳疾患における活用方法の有用性を示すことができ、適切な設定を行えば、AI を安全に使用できる可能性があります。

また、課題としては今回使用したデータが限られることや、情報の取り扱い、使用者がその回答を理解して診療につなげることができるかが挙げられます。

今後、さらに検証を進めていくことで、実診療の効率や質向上につなげることができます。その第一歩としての成果を発信させていただきます。

GPT-4VisionによるマルチモーダルAIによって中耳疾患を分類できるか？

