

報道発表資料

2023年8月3日
東京慈恵会医科大学
サイオステクノロジー株式会社

希少疾患の鼻副鼻腔乳頭腫を動画から判定する AI 診断モデルの作成に世界で初めて成功 —ベテラン耳鼻科医の診断精度 77.6%を上回る 84.3%を実現—

東京慈恵会医科大学耳鼻咽喉科学講座 由井亮輔助教、高橋昌寛講師、鴻信義教授、小島博己講座担当教授らとサイオステクノロジー株式会社 野田勝彦、吉田要らの研究グループは、鼻腔内視鏡動画を用いて有病率の低い鼻副鼻腔乳頭腫に関する高い精度の AI 診断モデルの作成に成功しました。

本研究の精度向上などの手法を応用することにより鼻副鼻腔乳頭腫だけでなく他の希少疾患に対する人工知能研究にもつながることが期待されます。

<ポイント>

- ・ 希少疾患のため AI 学習用のデータ数の確保が難しい鼻副鼻腔乳頭腫について、鼻腔内視鏡動画から診断する AI 診断モデルを作成しました
- ・ 作成した診断モデルの精度は 84.3%と、6 年以上経験のある耳鼻科医の平均正答率 77.6%を上回りました
- ・ 今回の診断モデルが得意・不得意とした条件の検証を進めることで、様々な希少疾患への AI 診断の活用につながるとともに、医師による病変発見についても新たな基準が見つかる可能性が期待されます

本研究の成果は 8 月 2 日に英国科学誌「Scientific Reports」に掲載されました。

【本研究内容についてのお問い合わせ先】

東京慈恵会医科大学 耳鼻咽喉科学講座

助教 由井亮輔

講師 高橋昌寛

電話 03-3433-1111 (代)

【報道機関からのお問い合わせ窓口】

学校法人慈恵大学 経営企画部 広報課 電話 03-5400-1280 メール koho@jikei.ac.jp

サイオス株式会社 広報 電話 03-6401-5120 メール mktg@sios.com

1. 背景

鼻副鼻腔乳頭腫は良性腫瘍ですが、再発や悪性化する可能性があり、早期診断と鼻腔内視鏡下での外科的切除が望まれます。耳鼻咽喉科では外来診察に鼻腔内視鏡を用いていますが、鼻副鼻腔乳頭腫と鼻腔ポリープ（鼻茸）との鑑別が困難な症例も多くあります。



確定診断には病理検査が必要ですが、1週間程度の時間がかかり、鼻腔内視鏡検査によって精度の高い補助診断ができれば臨床的に有用です。そこで我々は、鼻腔内視鏡動画を用いて鼻副鼻腔乳頭腫を診断するコンピュータ支援診断システムを開発しました。

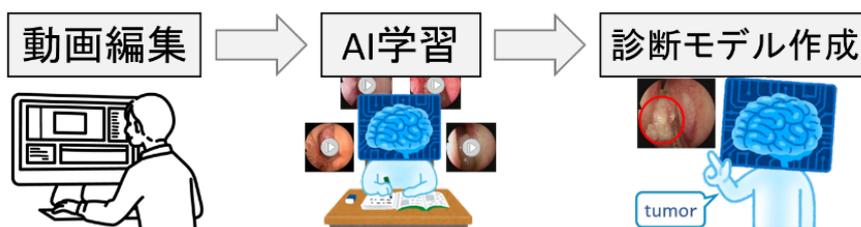
近年の人工知能（AI）および機械学習技術の進歩は著しく、医療分野においても重要になってきています。予測アルゴリズムの作成は長い間、医療における意思決定を助けるために使用されてきましたが、機械学習手法の実用化は2000年に始まりました。その後、コンピュータのハードウェア性能の大幅な向上により、2010年にはDeep Neural Network（DNN）が導入され、2012年にはImageNet Large Scale Visual Recognition ChallengeにおいてDNNの精度が従来の画像処理手法を上回り、最終的には2015年に人間の画像認識精度を上回りました。しかし、DNNモデルの学習には一般的に大量のデータが必要であり、鼻副鼻腔乳頭腫のような希少疾患の診断への応用は困難と考えられていました。

過去に鼻腔内視鏡動画を用いて鼻副鼻腔乳頭腫を診断するDNNを利用した研究は報告がありません。そこで本研究では、DNNモデルを用いて鼻腔内視鏡診断の精度を向上させることができるかどうかを調査することを目的としました。さらに、DNNモデルを様々な耳鼻咽喉科医による評価と比較し、その実用性を検討しました。

2. 手法

2018年から2021年に東京慈恵会医科大学附属病院耳鼻咽喉科で内視鏡下鼻副鼻腔手術を受けた患者の中で、病理検査で鼻副鼻腔乳頭腫と診断された患者21例、鼻腔ポリープを伴う慢性副鼻腔炎患者32例の計53例（男性=33、女性=30、平均年齢 51.2 ± 12.6 歳）を対象としました。

鼻副鼻腔乳頭腫、鼻腔ポリープそれぞれの病変が画面に映っている場面のみを手術動画（鼻腔内視鏡動画）を編集しました。次に患者を無作為に8グループに分け、学習用と評価用を分けて交差検証を行いました。



ニューラルネットワークと学習

学習時には、224×224 ピクセルのサイズに切り出した画像を用いて、画像の病変部を範囲内に収めながら DNN モデルを学習させました。1つの DNN モデルの1回の学習サイクルでは、50回の反復学習を繰り返し実行しました。この学習サイクルを8つのデータセットで行い、1つの学習セットで8つのモデルを生成しました（学習セット：評価セット=7：1）。各 DNN モデルの学習は少数の患者からオーグメントで生成した大量のデータを用いるため、学習するたびに能力・精度に差が出ます。その能力・精度の変動を検証するために、24の学習セットを作成しました。その結果、8データセット×24=192個の診断モデルが生成されました。

耳鼻咽喉科医による視診による診断

経験年数のさまざまな耳鼻咽喉科医 25名が、AIが評価した鼻腔内視鏡動画と全く同じものを見てもらい、鼻副鼻腔乳頭腫症例か、鼻腔ポリープ症例かを問い、正答率を評価しました。

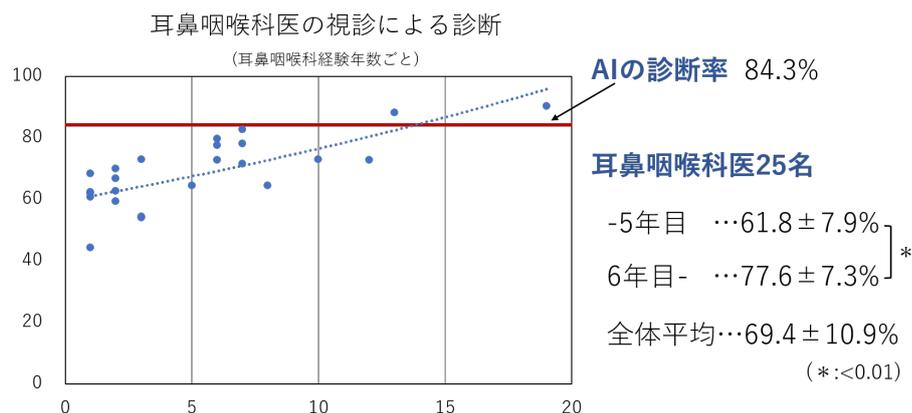
3. 成果

すべての診断モデルの中で最も優れた性能のものは、5秒間スコア分析を用いたアンサンブル予測による精度 84.3%（感度 81.0%、特異度 87.6%）でした。

耳鼻咽喉科医の診断精度・AIとの比較

耳鼻科医の診断精度は平均 69.4%でした。また、学年が上がるごとに成績が上がる傾向にあり、6年目以上は、それ以下と比べて有意差に高い正答率でした（6年目以上の平均正答率 77.6%、5年目以下の平均正答率 61.8%）。

以下のグラフの通り、AIの精度は、耳鼻咽喉科医全体の平均だけでなく、6年目以上の平均正答率を上回る結果でした。



AIの精度が耳鼻咽喉科医の診断率よりも高かった原因を探るため、AIと耳鼻科医の診断精度が異なる症例を検討しました。まず、耳鼻咽喉科医にとっては明らかな鼻副鼻腔乳頭腫の所見であるにも関わらず、AIが正しく診断できなかった症例がありました。これは機械学習された症例に偏りが生じていて、典型的な所見の一つを学習できていなかったためと考えられます。次に反対に、AIがほぼ完璧に診断したにもかかわらず、耳鼻咽喉科医による正しい診断率が低い症例もありました。これは、AIが人間と異なる部位を認識しているためと考えられますが、AIの具体的な判断基準は不明です。このような症例に

は病変発見の鍵が隠されている可能性があると考えました。

本研究には症例数が少なく、診断率にも一部ばらつきがありました。この問題を解決するためには、前向き研究を行い、より多くの症例を、同様の方法で撮影し、精度を向上していくことが不可欠です。今後は多施設での臨床研究や、対象疾患を増やしていく予定であり、鼻腔内視鏡で様々な疾患のスクリーニングができるようになれば、健診や非専門医の診察にも AI を活用できるようになることが期待できます。

4. 今後の応用、展開

鼻副鼻腔乳頭腫の鼻腔内視鏡動画を用いた AI 予測の精度を検討し、症例数が少ないにもかかわらず、精度の高い AI を作成できました。このことは鼻副鼻腔乳頭腫の自動診断の重要な第一歩となると考えています。そして他の希少疾患に対する AI 研究に光を与えるものであると考えています。今回の経験を活かし他の有病率の低い疾患に対する解析も行って参ります。

発表雑誌

雑誌名： *Scientific Reports*

論文タイトル： Preoperative prediction of sinonasal papilloma by artificial intelligence using nasal video endoscopy: A retrospective study

著者： Ryosuke Yui, Masahiro Takahashi, Katsuhiko Noda, Kaname Yoshida, Rinko Sakurai, Shinya Ohira, Kazuhiro Omura, Nobuyoshi Otori, Kota Wada, Hiromi Kojima

DOI 番号： 10.1038/s41598-023-38913-0

5. 脚注、用語説明

深層学習（ディープラーニング）：

深層学習とは、機械学習の手法の 1 つであり、多層のニューラルネットワークまたは複数の機械学習アルゴリズムを組み合わせた手法の事を指します。

ネットワークモデル：

深層ニューラルネットワークの構造は、多数提案されており、代表的な構造には名称が付与され、総じてネットワークモデルと呼びます。

オーグメント：

オリジナル画像の明度やコントラストの変更、回転・反転、拡大・縮小などにより、擬似的に画像の枚数を増やす行為を指します。

以上