



# 人工知能を用いた鼻内視鏡動画における 鼻副鼻腔乳頭腫の自動診断

由井亮輔<sup>1)2)</sup> 高橋昌寛<sup>1)</sup> 櫻井凜子<sup>1)</sup> 白木雄一郎<sup>1)</sup> 小黒亮史<sup>1)</sup>  
嶋村洋介<sup>1)</sup> 大村和弘<sup>1)</sup> 鴻信義<sup>1)</sup> 野田勝彦<sup>3)</sup> 吉田 要<sup>3)</sup>

1) 東京慈恵会医科大学 耳鼻咽喉科学教室

2) 東邦大学医療センター大森病院 耳鼻咽喉科

3) サイオテクノロジー株式会社

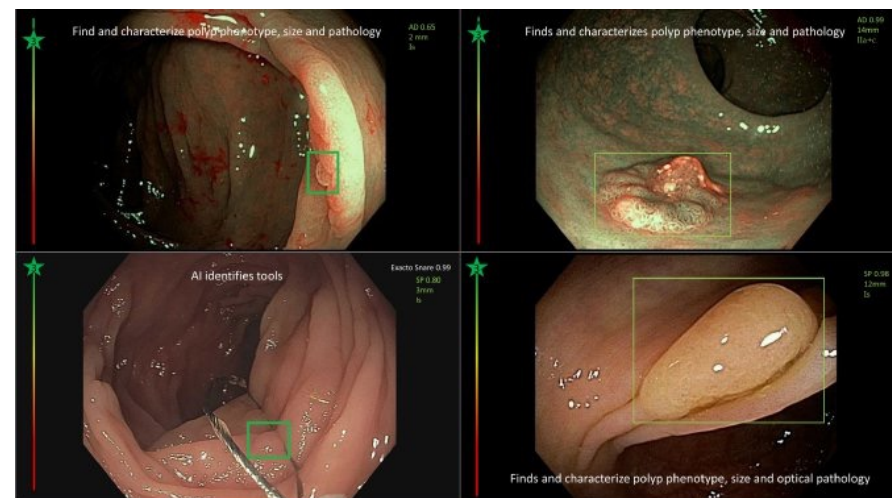
由井亮輔 高橋昌寛 櫻井凜子 白木雄一郎 小黒亮史  
嶋村洋介 大村和弘 鴻信義 野田勝彦 吉田 要  
：演題発表に関連し、開示すべき利益相反（COI）  
関係にある企業・団体等はありません。

# はじめに

大腸がんの診断サポートシステム(実用化)

→約5,000枚の内視鏡画像から作成：98%の早期癌発見率

Yamada M, Sci Rep. , 2019.

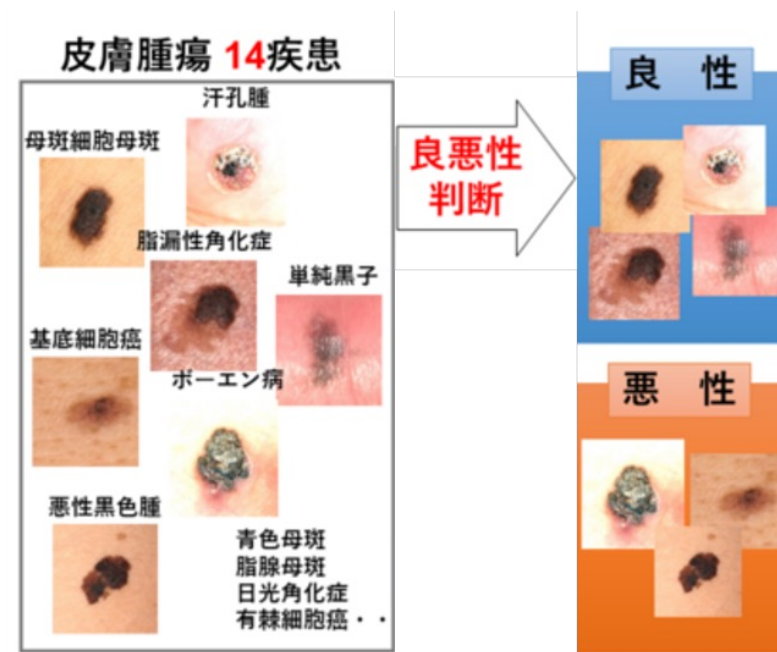


皮膚がんの診断サポートシステム(開発中)

→約5,000枚の皮膚腫瘍画像から作成：92%の悪性診断率

(1,842症例)

Esteva A, Nature, 2017.



など多くの報告がある

しかし鼻副鼻腔疾患の内視鏡所見に関する

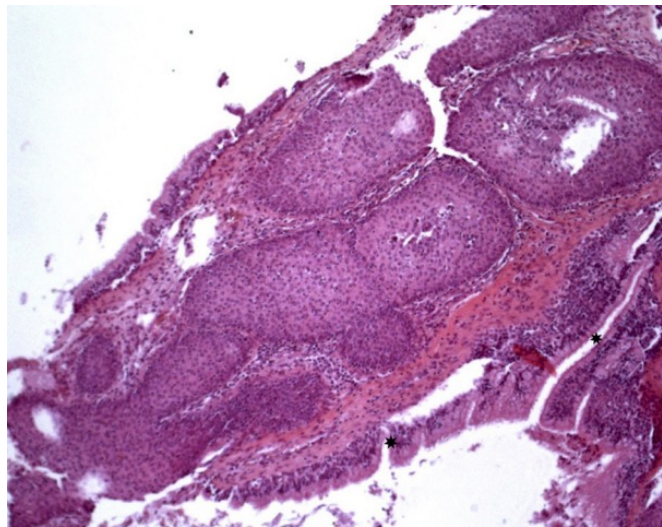
人工知能(artificial intelligence) を用いた研究報告はない



# はじめに

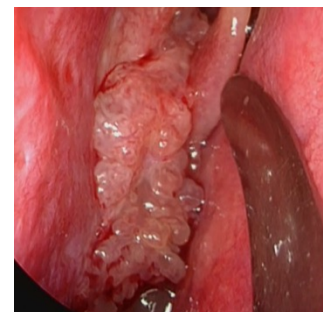
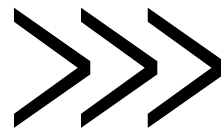
## 鼻副鼻腔腫瘍の診断

診断には生検が必要 (例：鼻副鼻腔乳頭腫)



Q. Lisan, Head and Neck diseases, 2016

病理◎



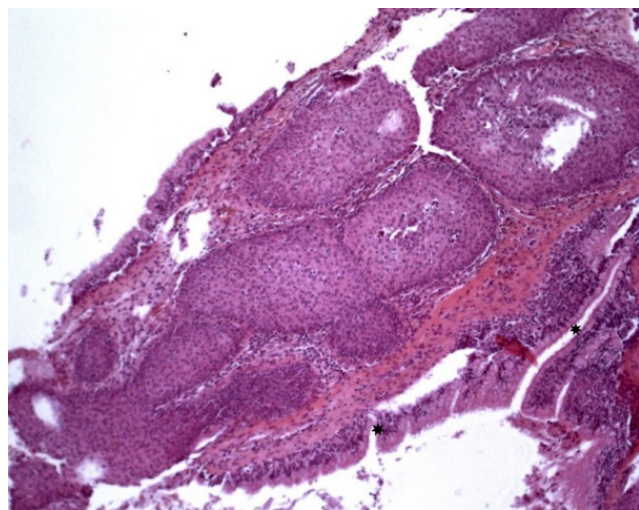
内視鏡所見△



# はじめに

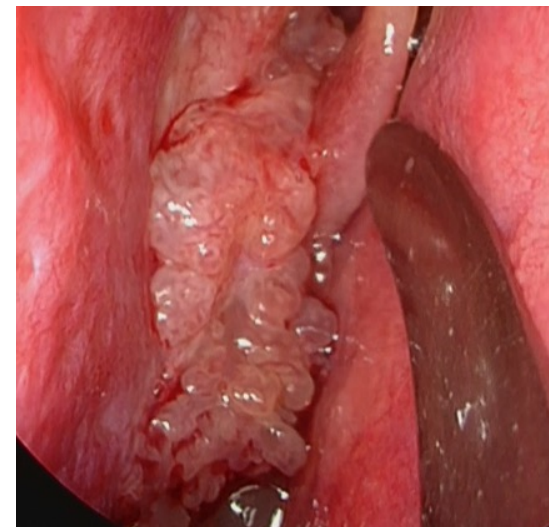
## 鼻副鼻腔腫瘍の診断

AIを用いて内視鏡の診断精度を向上させたい



Q. Lisan, Head and Neck diseases, 2016

≡



病理◎

↑ 内視鏡所見◎ ↑

# はじめに

## 鼻副鼻腔腫瘍の診断

AIを用いて内視鏡の診断精度を向上させたい

AIの学習には一般的に数千以上の症例数が必要だが

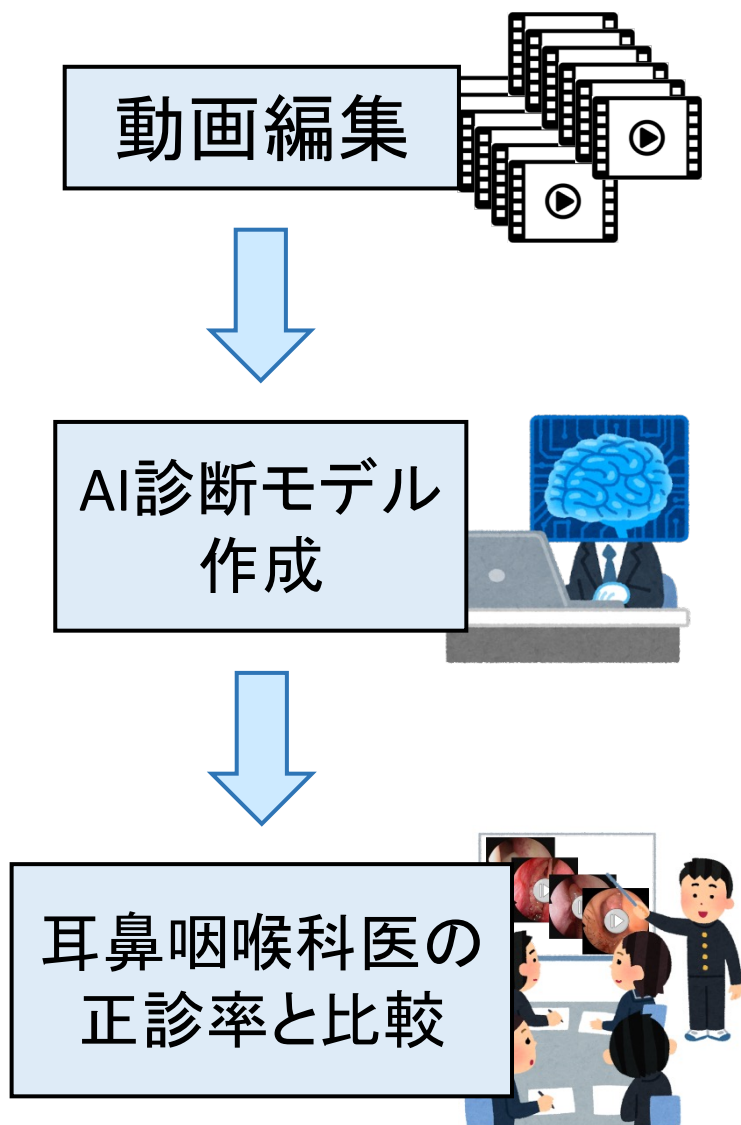
鼻副鼻腔腫瘍は症例数が少ない

### 目的

少ない症例数でも内視鏡動画を用いることで  
高い精度のAIモデルを作る



# 方法



## 対象

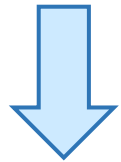
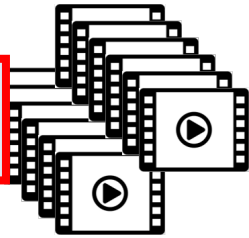
東京慈恵会医科大学附属病院において  
内視鏡下手術を施行した53症例

- 鼻副鼻腔乳頭腫(papilloma) 21例
- 好酸球性副鼻腔炎(ECRS) 32例

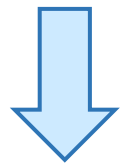
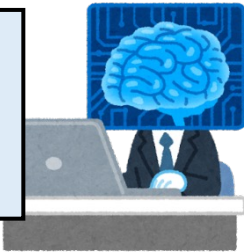


# 方法

動画編集



AI診断モデル  
作成

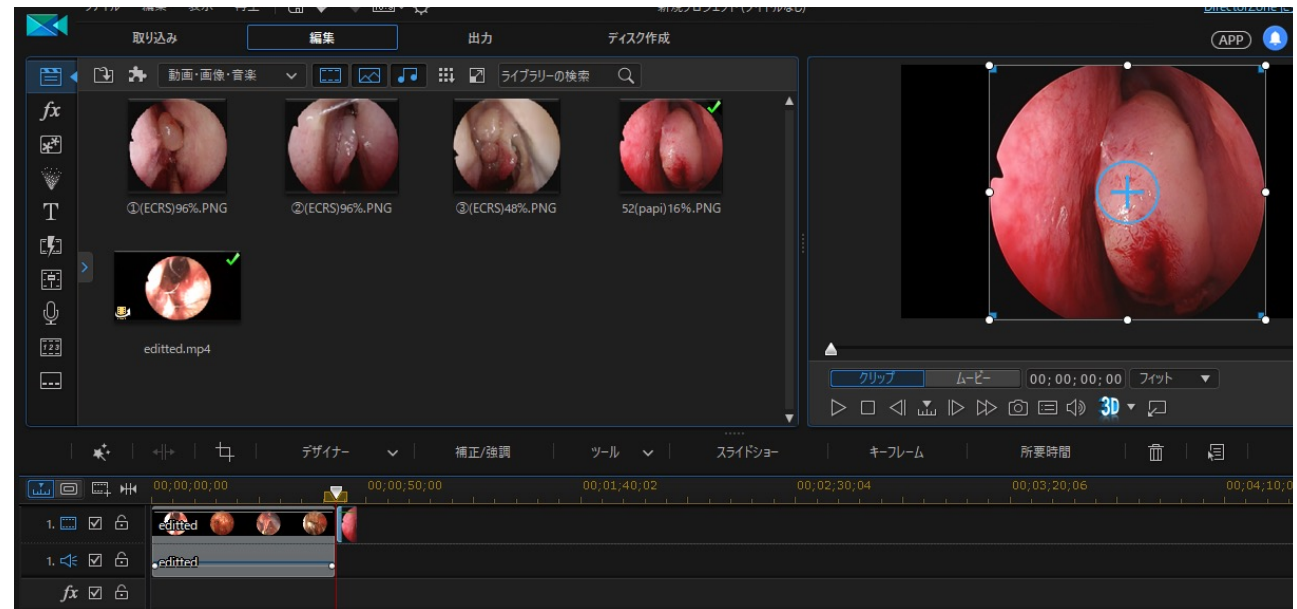


耳鼻咽喉科医の  
正診率と比較



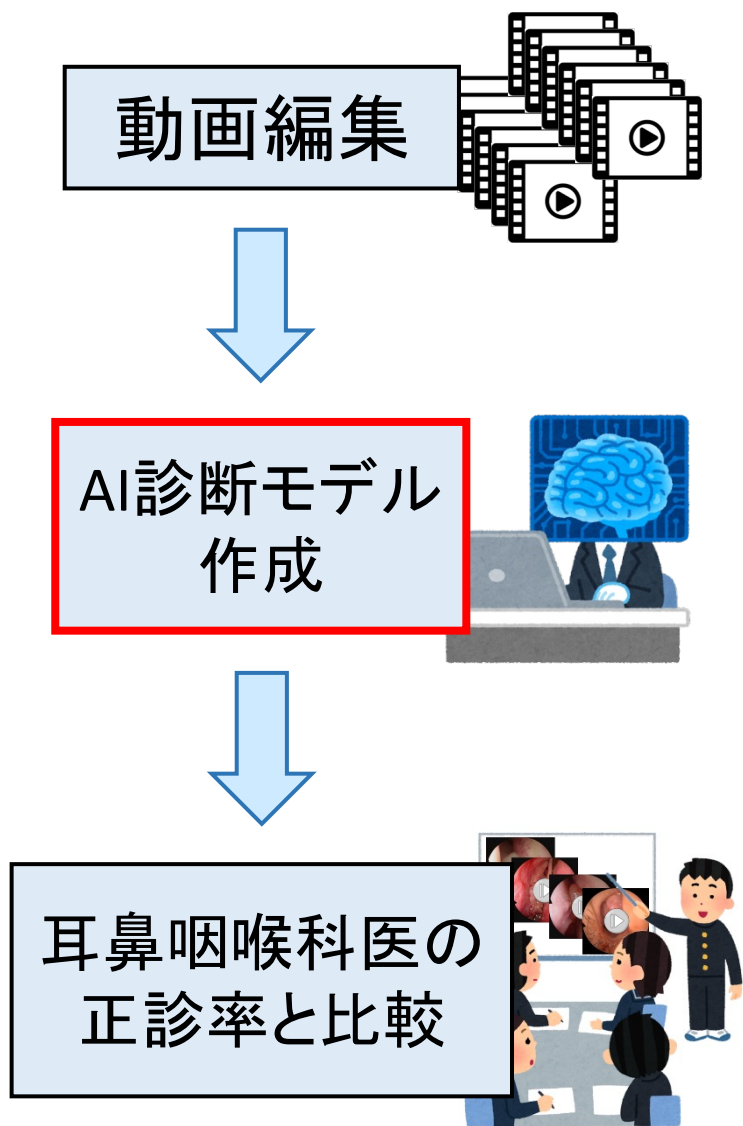
## 動画編集

- 病変操作前の手術動画を2、3分程度に
- 常に画角に病変が映っているように
- 疾患ごとの差が出ないように  
(出血量、鉗子操作、フレーム数など)



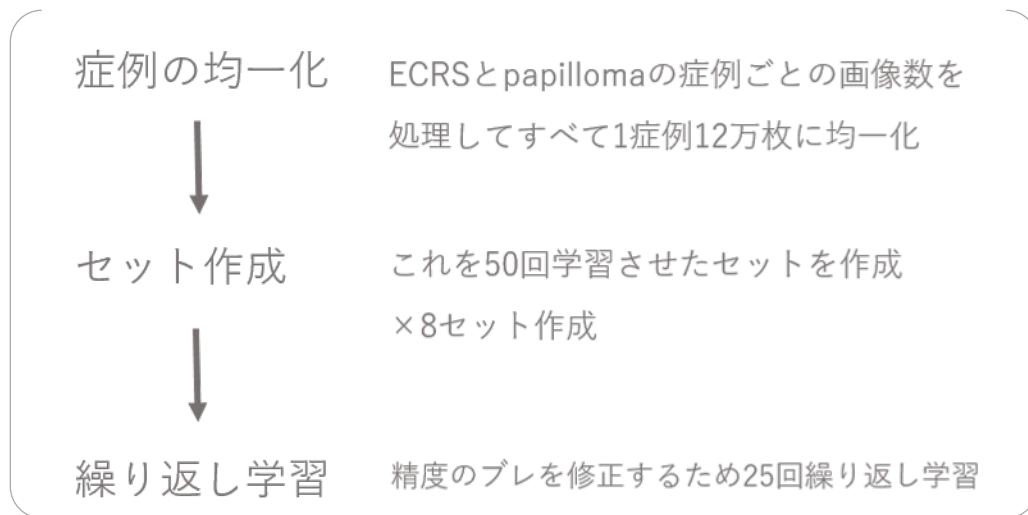


# 方法

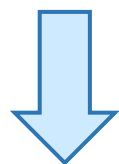
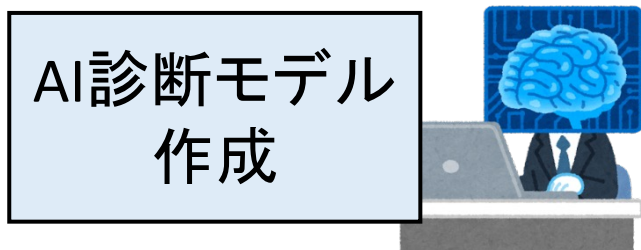
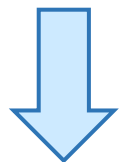
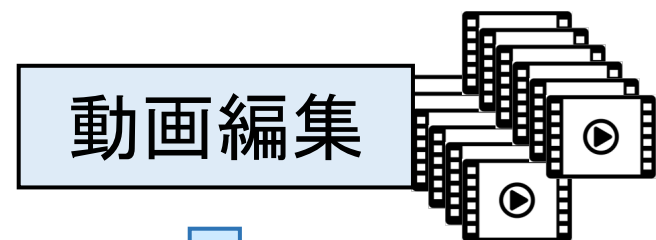


## AI診断モデルの作成

- AI専門チームに依頼した
- 少ない症例数でも高い精度の診断モデルを作成できるように、独自の技術で繰り返しデータを学習させた



# 方法



耳鼻咽喉科医の  
正診率と比較

## 耳鼻咽喉科医の視診による診断

- ・耳鼻咽喉科学教室員25名を対象とした
- ・すべての動画を見て診断を予想してもらう
- ・Googleフォームで回答、計算した



質問 回答 合計点: 56

« 1 »papilloma or not ?

papilloma

not papilloma

« 2 »papilloma or not ?

papilloma

not papilloma

« 3 »papilloma or not ?

papilloma

# 結果

## AI診断モデル

$$\text{精度} = (\text{感度} + \text{特異度}) \div 2$$

	モデル	感度	特異度	精度
連続数	シングル	0.779	0.860	0.820
	アンサンブル	0.810	0.873	0.841
5秒間 スコア	シングル	0.775	0.869	0.822
	アンサンブル	0.810	0.876	0.843

判定手法  
今回は4通りの手法で  
最良の診断モデルを模索



# 結果

## AI診断モデル

	モデル	感度	特異度	精度
連続数	シングル	0.779	0.860	0.820
	アンサンブル	0.810	0.873	0.841
5秒間 スコア	シングル	0.775	0.869	0.822
	アンサンブル	0.810	0.876	<u>0.843</u>

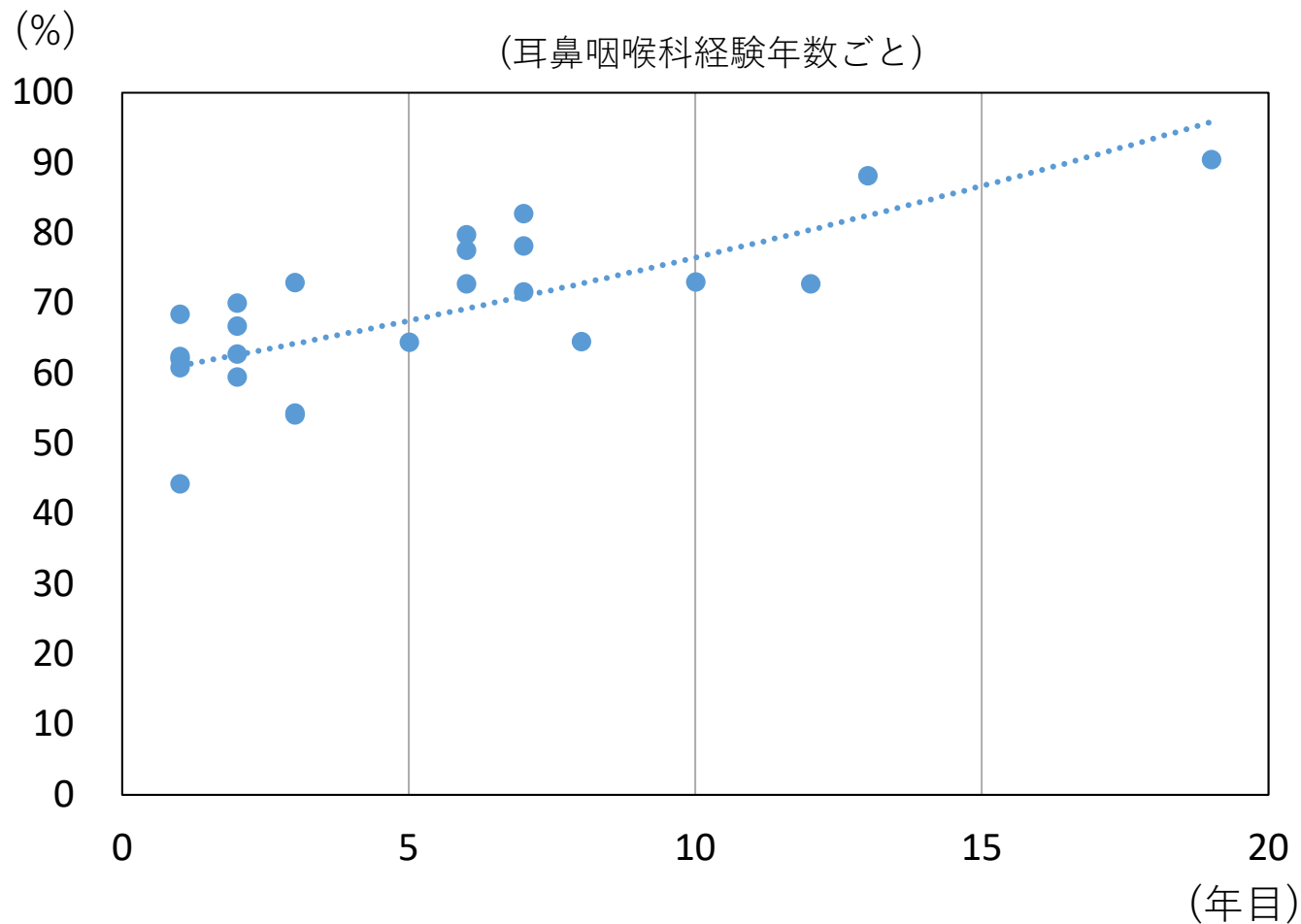
少ない症例数であったが  
84.3%の精度のAI診断モデルを作成できた



# 結果

## 耳鼻咽喉科医の視診による診断

(耳鼻咽喉科経験年数ごと)



耳鼻咽喉科医25名

-5年目 …61.8 ± 7.9%

6年目- …77.6 ± 7.3%

全体平均…69.4 ± 10.9%

(\* : <0.01)

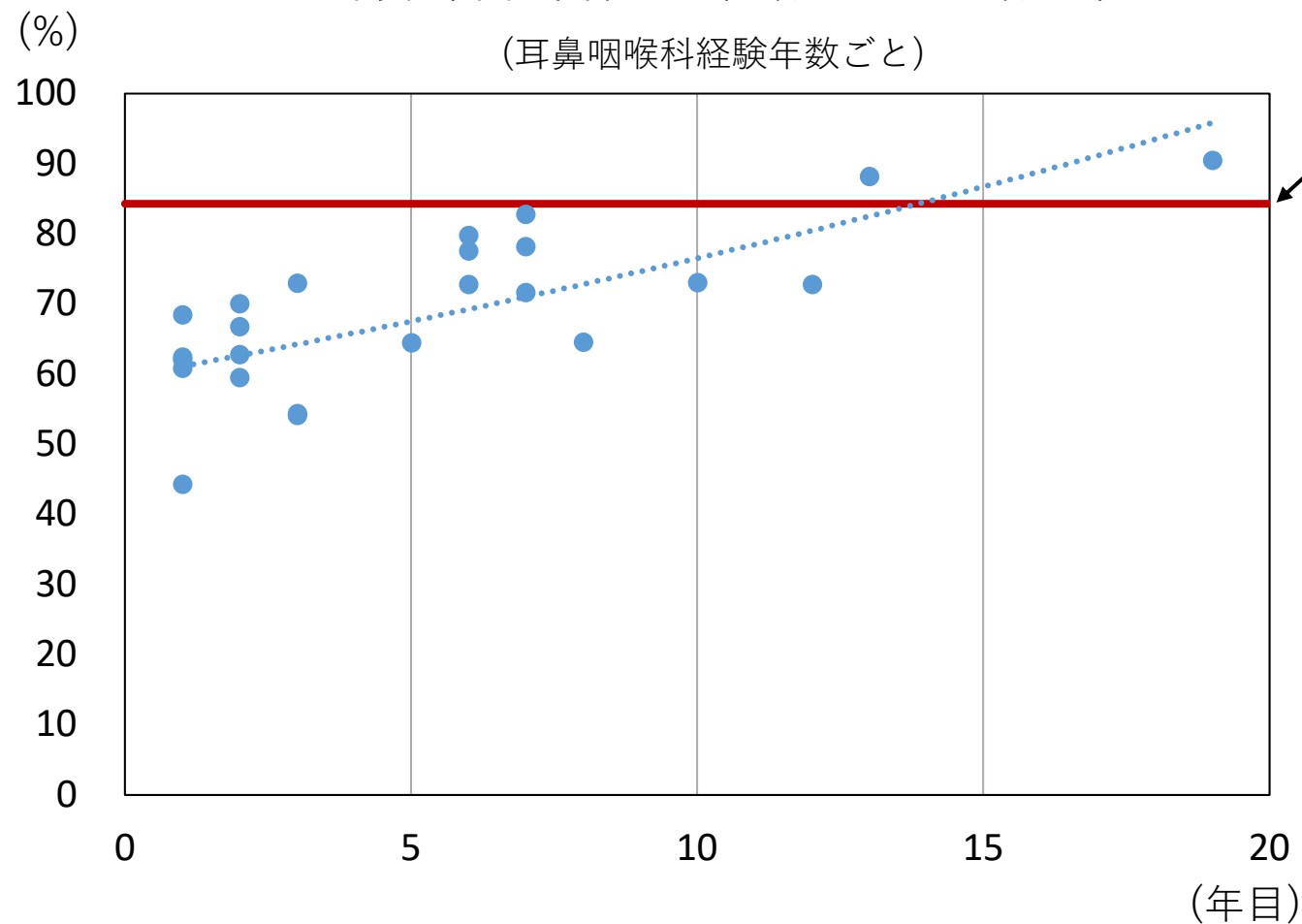
経験年数が増えるにつれて、成績が上昇する傾向



# 結果

## 耳鼻咽喉科医の視診による診断

(耳鼻咽喉科経験年数ごと)



AIの診断率 84.3%

耳鼻咽喉科医25名

-5年目 …61.8 ± 7.9%

6年目- …77.6 ± 7.3%

全体平均…69.4 ± 10.9%

(\* : <0.01)

AI診断モデルの精度は、全体平均だけでなく専門医(6年目以上)を上回る

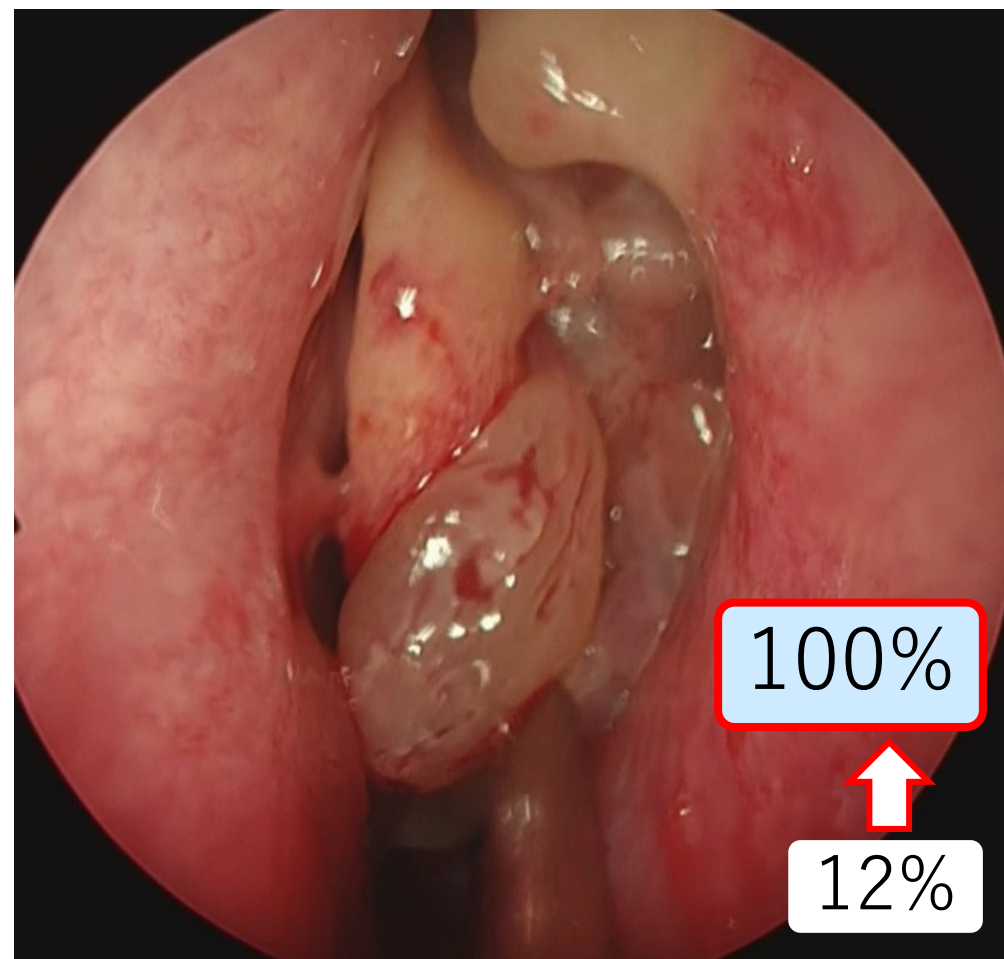
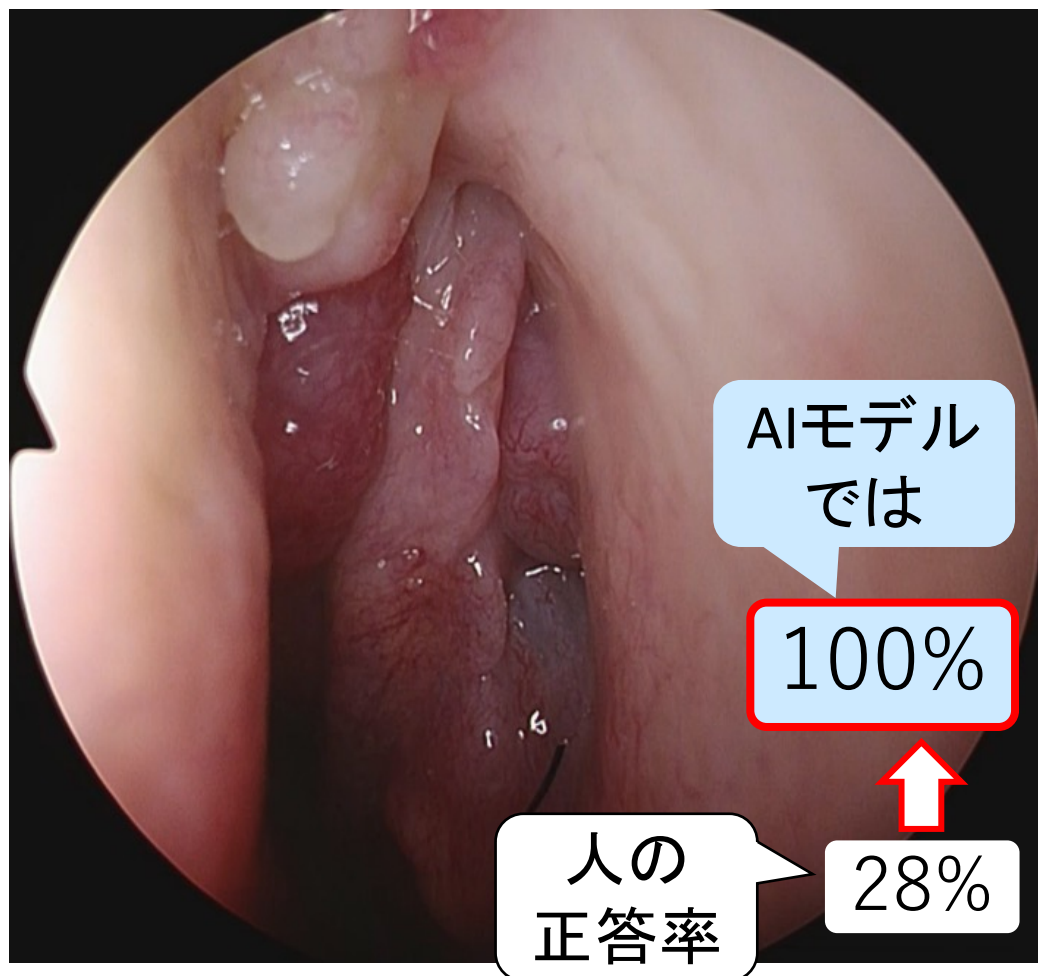




# 考察 人の正答率 vs. AIの精度

ECRS

papilloma

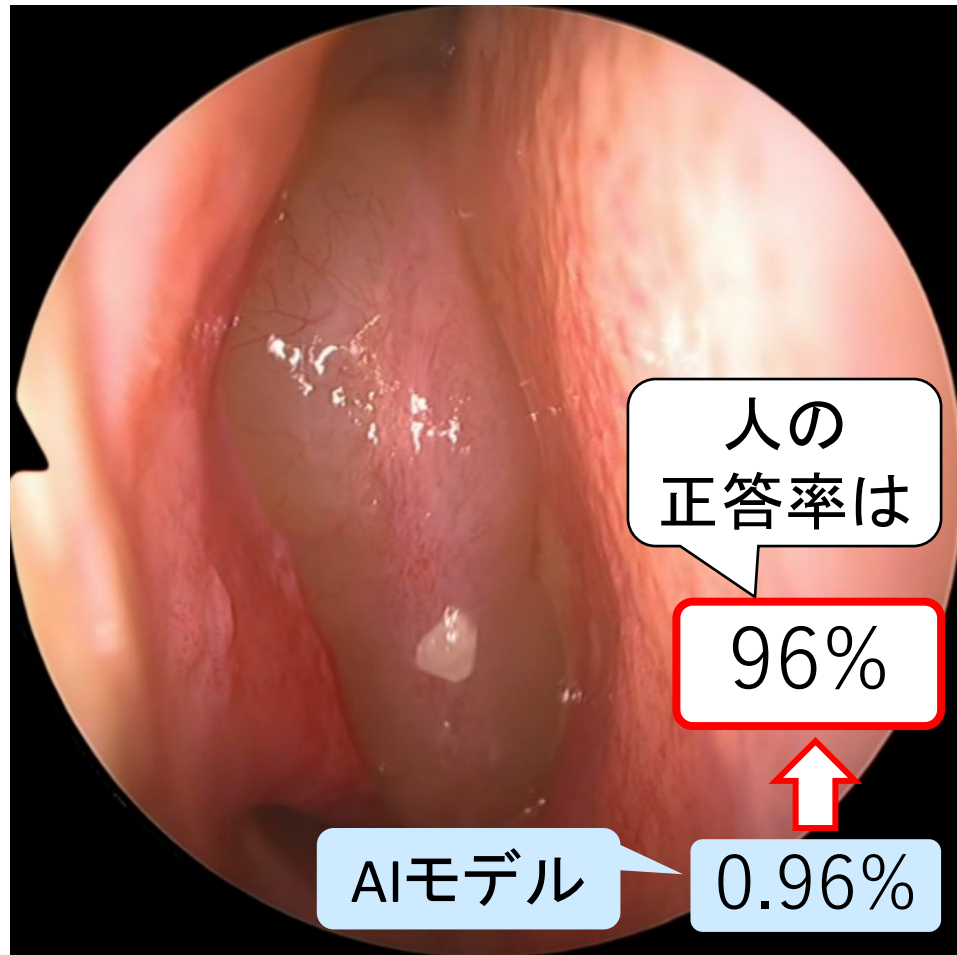


⇒人にとって難しい症例でも、AIの精度は高い

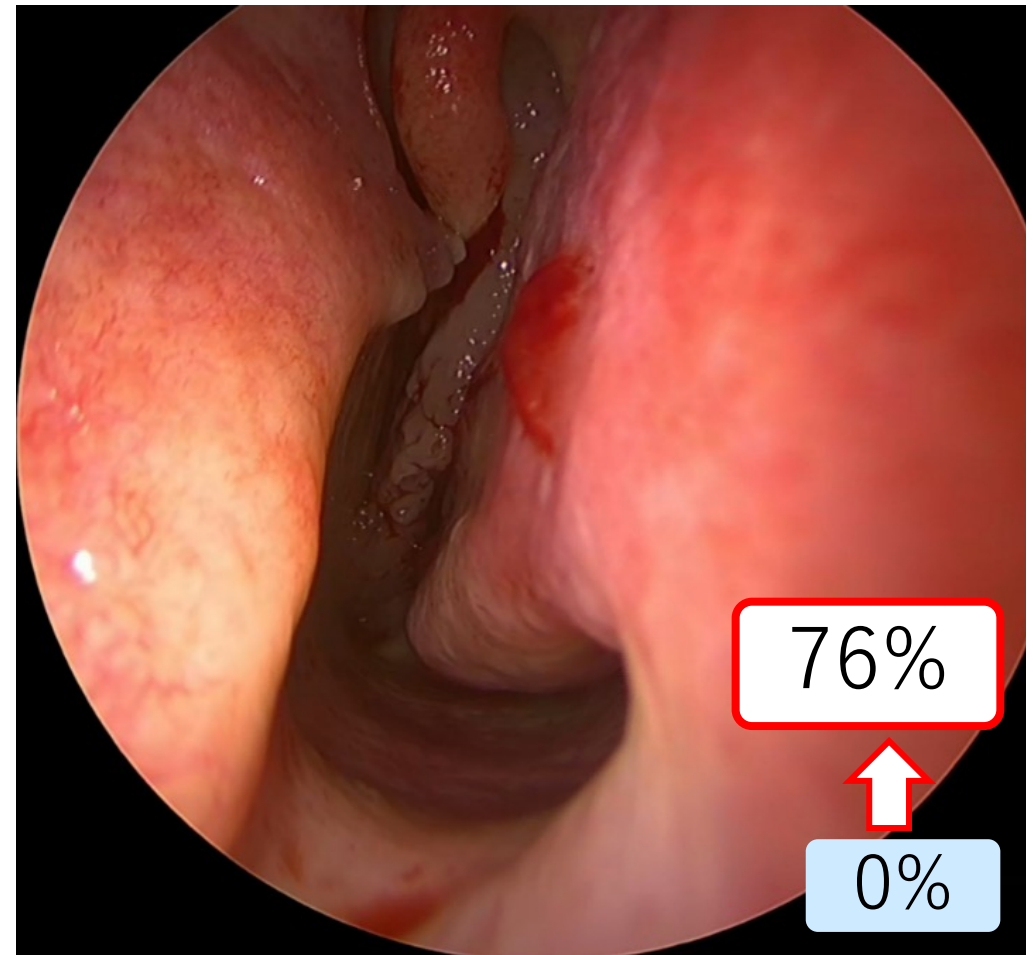


# 考察 AIの精度 vs. 人の正答率

ECRS



papilloma



⇒AIと人では正答率が高い症例は異なる

# 考察

## 今後AIの精度をあげるために

- ・認識している部分が人とAIでは異なる

AIシステムにおいて、画像中の遠方にある病変の検出は弱い

Yamada M, Sci Rep, 2019.

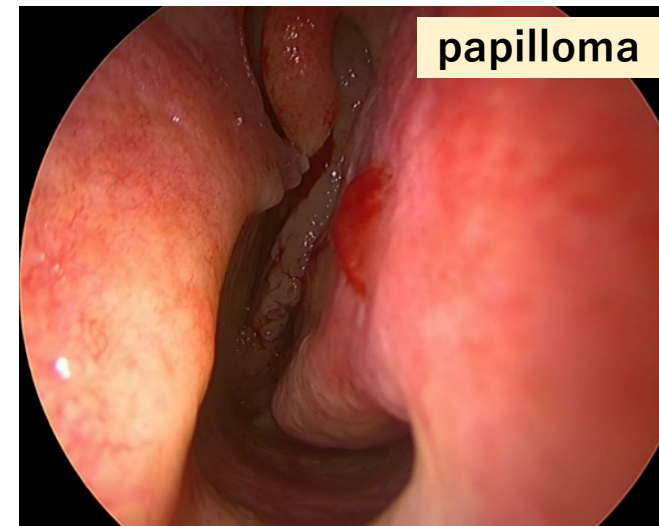
⇒近接した画像を学習させる

- ・人が当然と考える著しい所見でも類似例がないとAIの精度が上がらない

同系統の画像の学習が多いと、新規画像での性能評価において十分な性能が担保できない

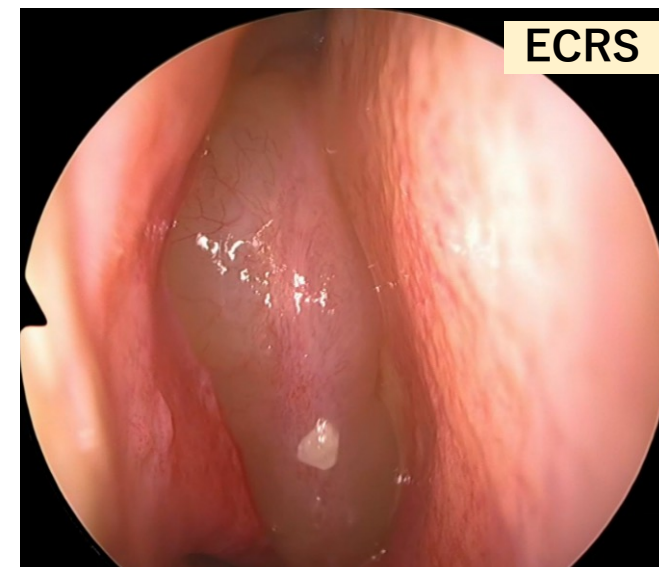
Chen PC, Nat Mater, 2019.

⇒症例を増やすことが必要



papilloma

人76% AI 0%



ECRS

人96% AI 0.96%

# まとめ

- ① 非常に少ない症例数でも、動画を用いることで耳鼻咽喉科専門医以上に相当する精度84.3%のAI診断モデルを作成できた
- ② 人の正答率が低い症例でもAIの精度は非常に高く、今後の自動診断に期待できる
- ③ AIの精度が低い症例も存在し、今後は撮像方法の標準化と症例数増加によりさらなる精度向上を目指す
- ④ 診断に難渋する耳鼻咽喉科疾患は、多くの症例を学習させることが難しいため、今回の手法を他疾患に広げていきたい

