

## 嚥下障害診断補助ツールの開発

## Development of a diagnostic support tool for dysphagia

牛島諒羽<sup>1</sup> 古川大輔<sup>2</sup> 黒岩眞吾<sup>3</sup> 西田昌史<sup>1</sup> 西村雅史<sup>1,4</sup>

Ryo Ushijima Daisuke Furukawa Shingo Kuroiwa Masafumi Nishida Masafumi Nishimura

## 1. はじめに

誤嚥性肺炎は日本人の死因第 6 位であり[1]、その主な原因である嚥下障害の早期発見と適切な対応が求められている。嚥下障害の主流な診断方法には、人員や専門医の不足、検査機器・設備の不足に加え、侵襲性が高いといった課題がある。医師や言語聴覚士によって実施される頸部聴診は高価な機器を必要とせず、非侵襲であるが主観評価に依存しているという問題点がある。近年、嚥下音に深層学習を利用することで嚥下障害を診断するスクリーニング手法が提案されている。無線型ウェアラブルシステムを用いた不顕性誤嚥の自動診断に関する研究[3]では、臨床応用にあたって高性能な PC や高価なウェアラブル装置を必要とし、現場導入には課題がある。一方で我々はこれまでに嚥下音に CNN を用いて嚥下障害者と健常高齢者と若年者の 3 クラス分類において 82% の分類精度を達成し、XAI による可視化にも成功している[2]。そこで本研究では、文献[2]での研究を踏まえて、臨床現場での使用を想定し、結果の表示までを半自動化した嚥下障害診断補助ツールを開発した。

## 2. 嚥下障害診断補助ツール

図 1 に本ツール使用時の画面を示す。t-SNE の結果が表示され、データをクリックすることで分類結果や可視化結果の確認が行える。画面左には表示モードの選択メニューと 2 次元プロットが表示され、「嚥下障害者・健常高齢者」

「嚥下障害者・健常者(若年者・高齢者)」「嚥下障害者・健常高齢者・若年者」のいずれかを選択して表示内容を切り替えられる。学習データと診断データは個別または同時に表示可能であり、各データ点をクリックするとファイル名、分類結果、信頼度、元ラベルが表示される。さらに、スカログラム画像や Score-CAM 画像に加え、嚥下音の再生も可能で、視覚・聴覚の両面から診断を補助する。図 2 に本ツールの全体構成と処理フローを示す。嚥下音の収集・切り取りを手動で行い、その後の結果の表示までの処理を自動化している。ツール利用時には、診断対象となる患者の嚥下音データを図 2 左上に示す咽喉部に直接装着するマイクを使用して収録する。収録した嚥下音を視察により 1 秒間で切り取り、フォルダに保存後、ツールを実行することで自動分類および 2 種類の可視化処理が行われる。

## 2.1 嚥下音に基づく自動分類

本節では、図 2 に示す処理フローのうち、嚥下音からの特徴抽出および自動分類処理を説明する。本ツールは、既存の健常高齢者と若年者および嚥下障害者データで事前学習した ConvNeXt を使い、新たに収録された診断対象データ(以下、診断データ)を分類する構成である。システムを実行すると、ゼリー嚥下時に収録した音声に連続ウェーブレット変換を適用し、スカログラム画像を生成する。生成画像を ConvNeXt に入力し、対象データのクラスを推定するとともに、その分類に対する信頼度が算出される。

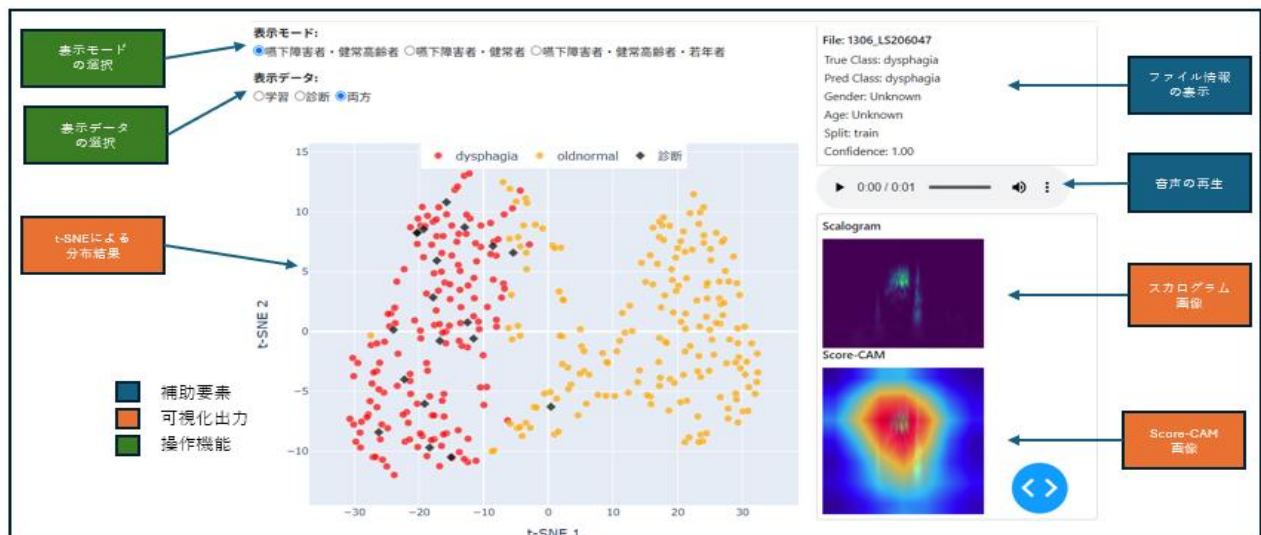


図 1 嚥下障害診断補助ツールの表示例

- 1 静岡大学 Shizuoka University  
 2 君津中央病院 Kimitsu Chuo Hospital  
 3 千葉大学 Chiba University  
 4 愛知産業大学 Aichi Sangyo University

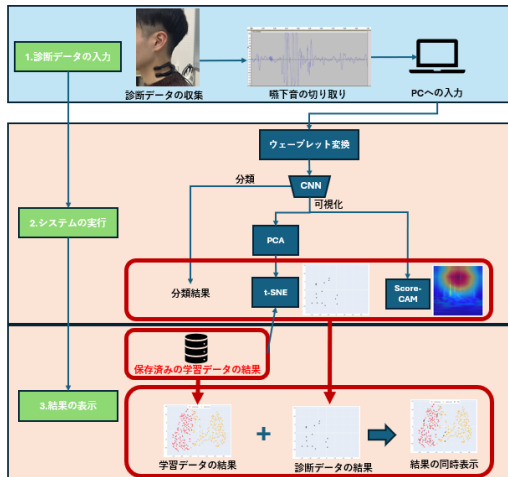


図2 嚥下障害診断補助ツールの処理フロー

## 2.2 t-SNEによる可視化

図2に示す処理フローのうち、t-SNEを用いた可視化処理を説明する。前項で学習や分類に使用した既存データのスカログラム画像に対し、処理を高速化するため全学習データをPCAにより30次元に圧縮したのち、t-SNEでさらに2次元に圧縮した。診断データに対しても同様にOpen t-SNEによって、この学習済みの2次元空間上にマッピングし、学習済データと位置関係を比較できるようにした。この手法により、診断データが健常者群または障害者群のいずれに近い分布を示すかを直感的に把握できる。

## 2.3 Score-CAMによる可視化

図2に示す処理フローのうち、Score-CAMは、分類結果に対してモデルがどの領域に注目して判断を下しているかを示す手法であり、本ツールでは各スカログラム画像に対して適用される。図3にScore-CAMによる注目領域の違いを示す。嚥下障害者では嚥下音中央部の低周波成分が強調され、健常高齢者では高周波成分への広範な注目が確認される。若年者はその中間的傾向を示しており、分類根拠の妥当性を視覚的に確認できる。

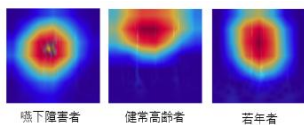


図3 Score-CAMの注目領域

## 3. 実データによる評価

本ツールの有用性を検証するため、まず既存の学習データを用いて分類モデルを構築した。この学習データとしては、管理環境で収録された若年者57名の193嚥下、EAT-10の調査で嚥下機能に問題ないと考えられた健常高齢者47名の184嚥下に加え、病院ベッドサイドで収録された嚥下障害者54名の155嚥下である。いずれも一口の嚥下リードと呼ばれるゼリーを食した際の嚥下音である。

一方、今回新たに嚥下障害の疑いで入院した患者を対象に同様のデータを収集し評価した。その内訳は、言語聴覚士による嚥下音の判断で、嚥下障害の疑いありとされた10嚥下と、正常な嚥下ができているとされた9嚥下である。

結果として、言語聴覚士が嚥下障害者の疑い有りとした10件のうち、システムは7件を嚥下障害者と分類し3件を健常高齢者と誤分類した。また、言語聴覚士が正常な嚥下ができていると判断した9件についても、7件を嚥下障害者と判断してしまった。これらのデータのt-SNEによる可視化結果でも、すべて嚥下障害者側に分布されるという結果となった。一方、Score-CAMによる可視化では、嚥下障害者の疑いのある患者データの注目領域は概ね図3に示す例と類似していたが、正常な嚥下ができている患者の注目領域も類似の傾向を示しており、t-SNE同様、これらの区別ができていないことがわかった。

今回期待した分類性能は得られなかった理由としては、学習データに用いた健常高齢者のデータと病院で収集した入院患者のデータでは嚥下能力に差があったことが原因だと考えられる。特に、入院患者内の健常高齢者は常食のような固形食を摂取しており、大学で収集した健常高齢者とは同じ健常でも、身体状態などに違いがあり、嚥下機能にばらつきが存在する。

## 4. おわりに

本研究では、言語聴覚士による臨床利用を想定し、嚥下障害の診断を支援する補助ツールを開発した。本ツールは、直感的なユーザーインターフェースと簡便な操作性を備えている。入院患者に対する評価実験では、分類精度や視覚的妥当性を十分に確認することはできなかった。この結果を踏まえ、さらなる性能向上が今後の課題である。

また、本ツールでは分析対象となる1秒間の嚥下音を視察により手動で切り出す必要があるが、この工程の自動化についても並行して検討を進めている。今後は、自動切り出しを含む処理の一括実行が可能なシステムとして統合し、臨床現場での実用化に向けた開発を進めていく予定である。

## 謝辞

本研究の一部はJSPS科研費18H03260、21K18305の助成を受けました。

## 参考文献

- [1] 政策統括官付参事官付人口動態・保健社会統計室, 令和5年(2023)人口動態統計月報年計(概数)の概況, 厚生労働省. <https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai23/index.html>, (2025-6-5)
- [2] R. Ushijima, D. Furukawa, S. Kuroiwa, M. Nishida, and M. Nishimura, "Visualization and classification of swallowing difficulties using scalogram-based CNN," in Proceedings of NCSP, pp.356-359, 2025.
- [3] B. Shin, S. H. Lee, K. Kwon, Y. J. Lee, N. Crispe, S.-Y. Ahn, S. Shelly, N. Sundholm, A. Tkaczk, M.-K. Yeo, H. J. Choo, and W.-H. Yeo, "Automatic clinical assessment of swallowing behavior and diagnosis of silent aspiration using wireless multimodal wearable electronics," *Advanced Science*, vol. 11, no. 34, p. 2404211, Sep. 2024.