

スマートフォンを用いた画像認識による 口腔・嚥下機能の定量的評価手法

耿世嫻¹ 平井雄太¹ 下島銀士¹ 柳田陵介² 山田大志² 小野寺宏¹ 戸原玄² 矢谷浩司¹

概要: 医師による嚥下障害の診断には、患者に定期的な通院が必要な点や診断結果が医師の経験に大きく左右される点が課題として存在する。そこで我々は、患者が在宅で簡単に嚥下障害の有無を評価できるようにするため、スマートフォンで撮影した動画から口腔・嚥下機能を評価する手法を提案する。提案手法の実現に向けて我々は、口腔・嚥下機能を分析するために必要なタスク群を決定し、147名の実験参加者から得られたそれらのタスクの動画の分析を行った。その結果、87.1%の精度、0.612のF値の識別性能を得た。

A Vision-based Oral and Swallowing Capability Quantification Approach with Smartphones

SHIXIAN GENG¹ YUTA HIRAI¹ GINSHI SHIMOJIMA¹ RYOSUKE YANAGIDA²
TAISHI YAMADA² HIROSHI ONODERA¹ HARUKA TOHARA² KOJI YATANI¹

1. はじめに

飲み込みが困難な状態は嚥下障害と呼ばれ、咳や窒息、栄養不足など、高齢者に多く見られる深刻な健康問題につながる可能性がある。脳卒中は障害の原因として非常に一般的なものであるが、嚥下障害は、その脳卒中発症後の症状としてよく見られる他、パーキンソン病や口腔がん、加齢による表情筋の衰えなど、さまざまな病気が原因で起こることもある [1]。症状が表れると、現在の評価方法では歯科医師による経験に基づく検査が必要になることが多いため、患者は頻繁に病院に通う必要がある。対面での評価では、医師が患者さんに標準ディサースリア検査 (AMSD) [2] や嚥下運動機能評価 (AMFD) [3] と呼ばれる長時間の口腔運動障害チェックを行うことが多く、その中には数時間にも及ぶ長い手順が含まれている。また、嚥下能力を調べるために、患者さんに液体を飲み込んでもらうこともある。

しかし、頻繁に病院を訪れることは患者に大きな負担を強いることになる。また、嚥下障害の評価方法の主流は経験的なものである。血圧や肺機能検査など他の健康診断と

は異なり、患者は医師による定性的なコメントを受けるだけで、数値的な測定データを得ることはできないのが普通である。一方で、定量的な測定データを取得するために、病院では舌圧測定器のような医療機器を用意している場合がある。Iowa Oral Performance Instrument (IOPI) [4], [5] を例とした舌圧測定器は、舌が口蓋を押す力を測定することで、舌の筋力を測定することができる。しかし、舌圧測定器は舌挙上時の最大舌圧を測ることができるけど、左右挺舌時の最大舌圧は測れない、嚥下障害の評価には限界があり、侵襲的でもある。また、舌圧測定器は個人で簡単に購入することができないため、在宅の患者さんには手に入らないことが多い [6]。また、舌圧は一般的に使われている健康指標ではないため、医学的知識を持たない多くの人にとって馴染みがない。

スマートフォンが普及した現在、スマートフォンによる健康状態の把握が注目されている。最近のスマートフォンは、さまざまなセンサーが搭載されている。一般的に使用されるセンサーには、カメラ [7], [8], [23], [24], [27], マイク [9], [10], [22], モーションセンサー [20], [28], 全地球測位システム (GPS) [25], [26] などがある。内蔵センサーの数が増え、オペレーティングシステムが進化したことで、スマートフォンはさまざまな種類の生体信号を検知し、身

¹ 東京大学

² 東京医科歯科大学 摂食嚥下リハビリテーション学分野

体的・精神的疾患を予測したり、必要に応じて健康介入を行ったりすることができるようになった。新しいセンシング技術により、様々な領域でのスマートフォンを用いた健康状態のセンシングが実現されている。例えば、バイタルサインの測定 [20], [21], [22], [23], 黄疸 [8] や肺機能異常 [9], [10] などの疾患の評価, メンタルヘルスのモニタリング [25], [26], 依存症のコントロール [11], [12], [13] などである。スマートフォンを利用したモバイルヘルスは、臨床でのみ行われてきたセンシングを、病院以外でも行えるようにする大きな機会となった。

本論文では、スマートフォンのカメラを用いた視覚に基づく口腔・嚥下機能の定量化手法を検討し、嚥下障害患者に定量的かつ付加的な情報を提供し、自宅で便利に自分の状態を把握・追跡できるようにする。本システムでは、スマートフォンのカメラで5回、口腔を動かす動画を撮影する。撮影した5つの動画から24の特徴量を抽出し、サポートベクトルマシン (SVM) を用いて分類する。東京医科歯科大学の訪問患者15名を含む147名の実験参加者に対してアルゴリズムをテストした結果、87.1%の精度、0.612のF値を得た。また、抽出された特徴量に対して多変量解析を行い、データの背後にある医学的な意味を考察する。

2. 医学的背景

嚥下障害の原因はさまざまである。脳卒中やパーキンソン病などの神経疾患によって、飲み込みに関わる筋肉をコントロールする神経系がダメージを受けることで起こることもある。飲み込みに重要な神経の一覧は以下の通りである [14]。

- **脳神経 VII/Cranial Nerve VII/CN VII:** 顔面神経とも呼ばれる脳神経 VII は、目を閉じたり、唇をすぼめたり、歯を露出させたりといった顔の運動を司る。嚥下過程では、CN VII は運動作用も媒介する。顔面神経に関連する障害は、顔面の一部の麻痺や非対称性または不随意運動を引き起こす。
- **脳神経 IX, X/Cranial Nerve IX, X/CN IX, X:** 脳神経 IX と X は、舌咽神経や迷走神経とも呼ばれ、飲み込むときに咽頭を活性化する。この神経はまた、発声にも関与している。CN IX および X に関連する障害は、嚥下障害および発声障害を引き起こす可能性がある。
- **脳神経 XII/Cranial Nerve XII/CN XII:** 舌下神経である脳神経 XII は、舌の下を走り、嚥下を容易にするため舌の動きを制御する。CN XII の障害は、舌の動きの異常につながる可能性がある。

このうち、脳神経 VII と脳神経 XII の障害は、スマートフォンのカメラで撮影した動画で評価できる程度に、比較的是っきりとした症状が出る。脳神経 VII の障害では顔の動きに異常が見られることが多く、脳神経 XII の障害では

舌の動きに異常が見られることがある。脳神経 IX/X の損傷は、喉の内側 (咽頭) をよく観察する必要があるため、スマートフォンのカメラで評価することは困難である。そこで、本研究では、神経学的な原因という観点から、脳神経 VII と XII の障害によって起こる嚥下障害のみを考察する。

神経障害以外の嚥下障害の原因としては、加齢による筋力低下や口腔癌などがある [1]。原因が違うにもかかわらず、しばしば似たような症状をもたらすことがある。このような場合、健康な人と同じように舌を出したり、頬を膨らませたりといった口腔内の動きをすることが困難な場合が多い。しかし、口や舌等の口腔の正常な動作は適切な嚥下行動のために必要である。そのため、口腔の動きの異常を調べることで、嚥下障害を評価することが可能となる。

病院では、医師は標準ディサースリア検査 (AMSD) [2] や嚥下運動機能評価 (AMFD) [3] に従う。この検査では、口腔運動、言語機能、呼吸、液体の嚥下など、20種類以上の作業をさまざまな条件下 (座る・寝る) で行ってもらう。このため、検査には数時間かかることもある。訓練を受けた専門家は、患者がこれらの作業をどの程度こなせるかを観察し、センサーなどにより計測された定量的データを用いず、経験に基づいて定性的に判断する。そのため、こうした複雑で経験則に基づいた評価がどのように判断されたかを説明できないことが多い。このような評価を行うためには、関連する医学的背景を持つ医師に対して、通常、広範なトレーニングが必要とされる。また、このような検査は、高齢の患者にとって負担となり、やる気を削ぐこともある。実生活では、患者は長いテストを受けるのと同様に、頻繁に病院を訪れることにも抵抗を感じる可能性がある。

AMSD や AMFD のような検査は、患者の状態について包括的な情報を提供することができるが、自宅での簡単な自己評価のために評価手順全体を再現する機械学習アルゴリズムを開発することは好ましくなく、非現実的である。よって我々の目標は、このような正式な臨床検査と組み合わせ使用できる、より少なくして簡単なタスクで済むシステムを構築し、患者が自宅でスマートフォンを使って自分の嚥下状態を把握し追跡できるように、付加的かつ定量的な情報を提供することである。

3. 関連研究

現在の嚥下障害の評価は経験的・定性的なものにとどまっているが、視覚に基づくアルゴリズムと非視覚に基づくアルゴリズムの両方を用いて、口腔・嚥下能力を自動的に評価するさまざまな方法がこれまで研究されている。

3.1 音響による口腔機能の評価

音響に基づく評価は、コミュニティによって広く探求されている。研究者らは、嚥下障害に関連する音声障害を検出するために、患者の音声を分析するアルゴリズムを開発

した [15], [16]. また, 頸部に装着するウェアラブルセンサーで嚙下音を感知する試みも行われている [17]. また, 株式会社プライムスでは, このようなハードウェアの商用プロトタイプのプロジェクトが始まっている [18]. しかし, このような実装は侵襲的とも言われ, 患者の首に密着するため, 安全性の観点で問題があると考えられる. 一方, スマートフォンのカメラで撮影する画像処理方式は, ハードウェアをカスタマイズすることなく, 身体に触れることがないため, より安全で便利である.

3.2 画像による口腔機能の評価

また, 画像を用いた嚙下・口腔機能の評価法もある程度研究されている. 例えば, サルコペニックは画像認識でサルコペニアの嚙下障害を評価するため, 頸部の静止画像の画像処理に着目した [29]. しかし, 口腔・嚙下機能の低下は, 喉の機能低下だけでなく, 口を開けることができない, 舌を出すことができないなどにもつながる. これは, これらの器官が嚙下の過程で重要な役割を果たすからである. また, これらの器官の動的な動きも評価時の重要な特徴である. そのため, 画像データだけでなく動画データも充実させ, 自宅で簡単に行えるような評価方法を模索する必要があると考えている.

3.3 センサーとしてのスマートフォン

近年のスマートフォンは, カメラ, マイク, 加速度センサーなど, 優れたセンシング能力を備えている. HCI 分野の研究者たちは, モバイルヘルスのためのセンサーとしてスマートフォンを研究している. 例えば, Seismo はスマートフォンのカメラと加速度センサーを利用して血圧を測定するアプリケーションを開発した [19]. OralCam はスマートフォンのカメラと機械学習アルゴリズムを用いて, 歯や歯茎のトラブルなど口腔内の病気を分析する [24]. また, SpiroSmart ではスマートフォンを用いて, 音響信号処理による低コストのスパイロメトリーシステムを構築した [10]. スマートフォンをセンサーとして利用することで, 自宅で健康診断ができるようになり, 臨床検査を外部でできるようになる可能性が大きくなる. また, スマートフォンの使用データは, タバコやアルコールなどの依存症対策 [11], [12], [13], パーキンソン病の薬物摂取のモニタリング [28] など, 様々なトピックで活用されている. スマートフォンで健康管理を行うメリットの一つは, ユーザーが自分の状態を簡便かつ頻繁に把握できることである. スマートフォンで収集されたデータは, ユーザーが自分の健康について把握・認識できるだけでなく, 自宅にいるときにも追加の情報を医師に提供することができる.

4. データ収集

現在の評価方法は経験則に基づくものが多いため, 実験

全体を設計する上で, 医師の方々の意見を考慮に入れることが不可欠である. そこで, 神経内科医に何度かヒアリングを行い, 実験やデータ収集の方法を共同で議論を行った.

4.1 医師インタビュー

システムの最大の特徴は使いやすさであり, 複雑な説明や長時間のテストは好ましくない. そこで, 作業の複雑さを最小限に抑えつつ, 分類に必要な特徴量を確保するために, 神経内科医に複数回ヒアリングを行った.

そこで, 嚙下時に有用な舌の機能と表情筋を評価するための課題を検討し, 最終的に以下の 5 つの口腔を動かすタスクに決定した. なお, プライバシー保護のため, 参加者の顔は下半分のみ撮影した.

- 舌をできるだけ左に突出させ, 3 秒以上保持する.
- 舌をできるだけ右に突き出し, 3 秒以上保持する.
- 舌をできるだけ前方に突き出し, 3 秒以上保持する.
- 口をすぼめ, 3 秒以上保持する
- 頬を膨らませ, 3 秒以上保持する.

4.2 実験参加者の募集

参加者の募集は 3 つのグループで行った. 最初のグループは, クラウドソーシングで参加者を募集し, 5 つのタスクの説明動画に従って自分のビデオをアップロードしてもらった. このグループの参加者は健康であることが期待される. 年齢層は 20 歳代から 60 歳代までである. 2 つめのグループは, 文京区シルバー人材センターで参加者を募集した. このグループの参加者は 60 歳以上であり, 健康であることが期待される. 最後のグループでは, 東京医科歯科大学の医師の協力のもと, 病院を訪れる 15 人の患者のデータを集めた. 動画はすべて, さまざまな機種 of スマートフォンで撮影したものである. フレームレートは, 30fps に設定した.

さらに, 性別・年齢と, 嚙下障害の症状別転帰尺度として広く使われている自記式の Eating Assessment Tool (EAT-10) [30] への回答をお願いした. EAT-10 のスコアが 3 以上であれば, 嚙下障害が疑われ, 医師への診断が推奨される. 本研究では, どのグループに所属しているかに関わらず, EAT-10 のスコアが 5 点以上であれば嚙下障害であるとみなした. EAT-10 のスコアが 5 以上であった参加者は 19 名であり, 嚙下障害陽性とした.

表 1 に, 3 つのグループの参加者から, 照明条件などの理由で不適格となったデータを削除した結果を示す.

5. 収集したデータの分析

ビデオデータは非常に高次元であるため, 訓練用のデータ数が限られる嚙下障害陽性のデータに対して機械学習のアルゴリズムを直接適用することは困難である. そこで, 高精度な分類の実現に向けて, 我々はまず有用だと予想さ

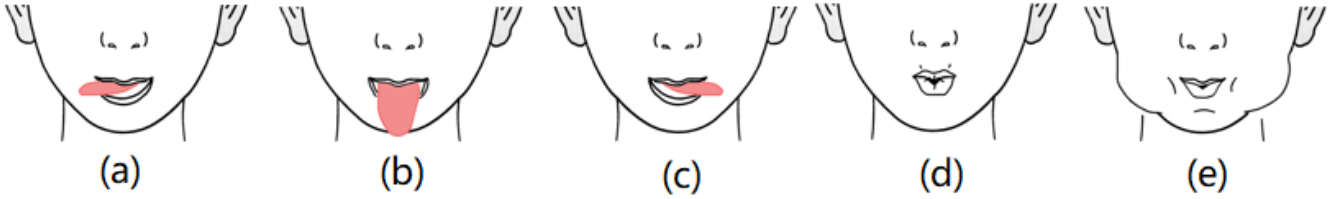


図 1: 撮影した口・舌の動きの種類. (a) 左挺舌 (b) 前挺舌 (c) 右挺舌 (d) 口すぼめ (e) 頬膨らませ

表 1: 参加者情報

	参加者数	グループ概要
グループ 1	74 (嚙下障害 2)	クラウドソーシングで募集, 若者, 健康
グループ 2	58 (嚙下障害 3)	シルバー人材センター経由で募集, 高齢者, 健康
グループ 3	15 (嚙下障害 14)	病院, 患者

れる特徴量の抽出を行い, 特徴量の次元を削減することにした. 我々は表 2 に示す 25 個の口・舌の動きに対する特徴量を分析の対象とした.

5.1 キーポイントのラベル付け

舌や頬などの位置を正確にラベル付けするために, 我々は顔の部位の位置 (ランドマーク) の自動検出を用いず, 手動でラベル付けすることにした. このようにした他の理由として, プライバーを考慮して我々のデータは顔の下半分しか含まれていないため, 顔全体の映像を必要とするランドマーク検出器を用いることができなかったということが挙げられる. ランドマークの検出を正確に行うためには, 一般的に顔全体が必要であることが多い.

表 2 の特徴量 #1-14 においては, 最も遠く舌を出した, 最も大きく口を窄めた等, 最も良かった動きをしたフレームを 1 つ選択し, 長さや角度をそれぞれ計算している.

表 2 の特徴量 #15~25 においては, 舌・唇・頬にランダムに 5 点のキーポイントをラベル付けする. これらのキーポイントの特徴量の計算手法は次節で詳述する.

また, カメラと実験参加者の間の距離が特徴量に与える影響を排除するため, 基準長として鼻の幅のピクセル数を利用することにした. Covid-19 の流行を考慮して, データ収集時に実験参加者の鼻幅を直接測定することは避けた. 産総研デジタルヒューマン研究所の統計によると, 日本人の平均鼻幅は表 3 の通りである [31]. この値を実験参加者の鼻の長さとし, これと鼻の幅のピクセル数を比較して, 1 ピクセルが何ミリメートルに対応するか計算し, 特徴量の計算を行った. この手法に起因する誤差の検討に関しては, 今後の課題としたい.

$$F = L_{vid}/L_{aist} \quad (1)$$

ここで, L_{vid} は収集した映像データの鼻幅 (ピクセル), L_{aist} は AIST のデータベースによる鼻幅である.

5.2 口・舌のトラッキング

口・舌の運動時の特徴を評価するために, オプティカルフロートラッキングを行う Kanade-Lucas-Tomas feature tracker (KLT) [32] を用いた. KLT はオブジェクトの先端に対するラベル付けを得意とするため, ラベル付けするポイントは舌, 唇, 頬の先端とした.

各タスクの平均動作特徴量は, 以下の式で算出される.

$$m_{avg} = \sum_{n=f_{start}}^{f_{end}-1} \sum_{i=1}^5 \frac{F \sqrt{(x_{i,n} - x_{i,n+1})^2 + (y_{i,n} - y_{i,n+1})^2}}{5(f_{end} - f_{start})} \quad (2)$$

ここで f_{start} 及び f_{end} は, 各タスクの動画データにおいて, タスクを開始及び終了したフレーム番号を表す. 表 2 に示すように, i は 5 つのラベル付き点のインデックスである.

また, 平均的な動きだけでなく, FFT (Fast Fourier Transform) 処理後に最大値となった時系列の動きの周波数も追加している (特徴量 #20-24). これにより, 舌の動きの中で最も顕著な周波数がわかる可能性がある.

6. 結果

25 個の特徴量の計算後, 各特徴量を, 放射状基底関数 (Radial Basis Function, RBF) をカーネルとするサポートベクターマシン (Support Vector Machine, SVM) への入力とした. 嚙下障害陽性の実験参加者は 19 名しかいなかったため, データセットを有効活用するため 5 分割交差検証を行った. また, ハイパーパラメータの最適化のために, グリッドサーチを用いた. 次節以降で, F1 スコア, ROC (Receiver Operating Characteristic Curve), PR カーブ (Precision-Recall curve) などを用いた分析結果を示す.

表 3: AIST データベースによる鼻幅

	若年層	年長者
男性	36.2 mm	38.7 mm
女性	33.1 mm	35.9 mm


6.1 識別結果

評価は 5 重交差検証により行った。最終的な SVM モ

デルの精度は 87.1% であり、F1 スコアは 0.61 (Precision: 0.50, Recall: 0.79) であった。表 4 は 5 重交差検証による予測結果混同行列である。

不均衡なデータセットでは、ROC 曲線 (図 2) は偏ることがあるが、PR 曲線 (図 3) は偏らない。図に示すように、我々のモデルは妥当な ROC 曲線 (AUC = 0.8174, PR 曲線 (AUC = 0.4025) を示し、我々のモデルがある程

表 2: タスクの一覧

	キーポイント (赤丸の部分)	特徴量
タスク 1		左舌突出量最大時の舌長さ (特徴量 #1) と顔面垂直中心線との角度 (特徴量 #2)
タスク 2		最大右側舌突出部の舌長さ (特徴量 #3) と顔面垂直中心線との角度 (特徴量 #4)
タスク 3		最大前方舌突出部の舌長さ (特徴量 #5) と顔面垂直中心線との角度 (特徴量 #6)
タスク 4		安定した口笛の水平・垂直方向の口唇長 (特徴量 #7 and #8) 顔の垂直中心線に対する水平・垂直方向の口唇角 (特徴量 #9 and #10)
タスク 5		安定した頬の膨らみの水平・垂直方向の口唇長 (特徴量 #11 and #12) 顔の垂直中心線に対する水平・垂直方向の口唇角 (特徴量 #13 and #14)
タスク 1		保持時の舌上 5 点の平均的な動き (特徴量 #15), FFT 後の最もパワーの大きい周波数 (特徴量 #20)
タスク 2		保持時の舌上 5 点の平均的な動き (特徴量 #16), FFT 後の最もパワーの大きい周波数 (特徴量 #21)
タスク 3		保持時の舌上 5 点の平均的な動き (特徴量 #17), FFT 後の最もパワーの大きい周波数 (特徴量 #22)
タスク 4		唇の 5 点保持時の平均的な動き (特徴量 #18), FFT 後の最もパワーの大きい周波数 (特徴量 #23)
タスク 5		唇の 5 点保持時の平均的な動き (特徴量 #19), FFT 後の最もパワーの大きい周波数 (特徴量 #24)
その他	-	年齢 (特徴量 #25)

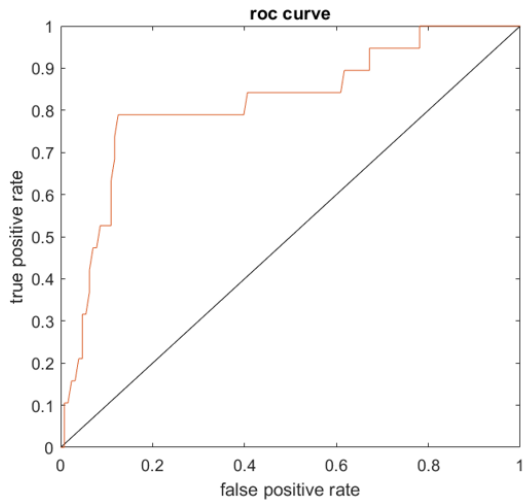


図 2: ROC 曲線 (AUC = 0.8174) AUC がベースラインの 0.5 超え, ランダムなモデルより良く分類していると思われるが, 不均衡なデータセットでは, ROC 曲線は偏ることがある.

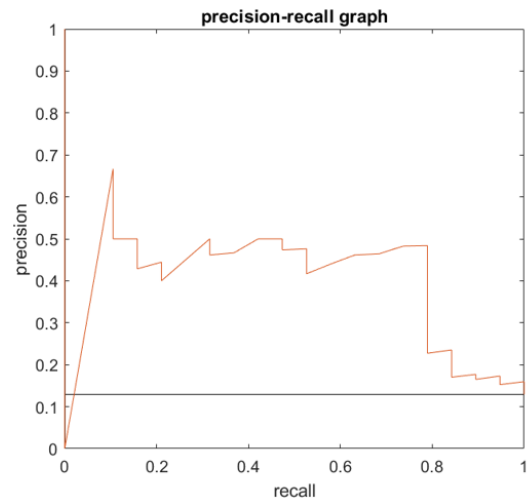


図 3: PR 曲線 (AUC = 0.4025) AUC がベースラインの 0.13 を超えているので, ある程度正負のケースを分類していると思われる. 不均衡なデータセットでも, RP 曲線は偏らない.

表 4: 混同行列

	予測 N	予測 P
実際 N	113	15
実際 P	4	15

度問題を区別して分類できることを示している. PR 曲線では, 0.8 リコールあたりで精度が急激に低下し, モデルがいくつかの正負のケースを区別することが困難であることを示している.

6.2 相関分析

また, データの背後にある医学的な意味を見出すために, 25 個の特徴量について Spearman の相関分析を行った. 図 4 は, Spearman の相関係数とその p 値を示したものである. このプロットから, 特徴量 #1-6, 特徴量 #8, 特徴量 #12 は陽性例と負の相関を持つことがわかる. このことから, 嚥下障害陽性例では, 舌の突出長が比較的小さく, 顔

の垂直線に対して舌が作る角度が小さいことがわかる. また, タスク 4 とタスク 5 では, 唇の垂直方向の長さが小さいことから, 参加者がこれらのタスクを実行できない可能性があることが示された. 一方, 特徴量 #18-22 と特徴量 #25 は嚥下障害と正の相関があり, タスク 4 と 5 で保持時の動きがより速いことが嚥下障害と相関している可能性があることが示唆される.

7. 本結果を用いた想定されるシステム

本結果を用いて, 今後もバルアプリを構築する予定である. システム全体では, ユーザからデータを取得した後, ソフトウェアプロトタイプの前処理技術として顔ランドマーク検出器を用いることで, 本実験で行ったようにユーザがキーポイントにラベル付けする必要がないようにする予定である.

図 5 に, 提案システムの全体的なユーザシナリオを表す. まず, ユーザは指示に従い, スマートフォンのフロントカ

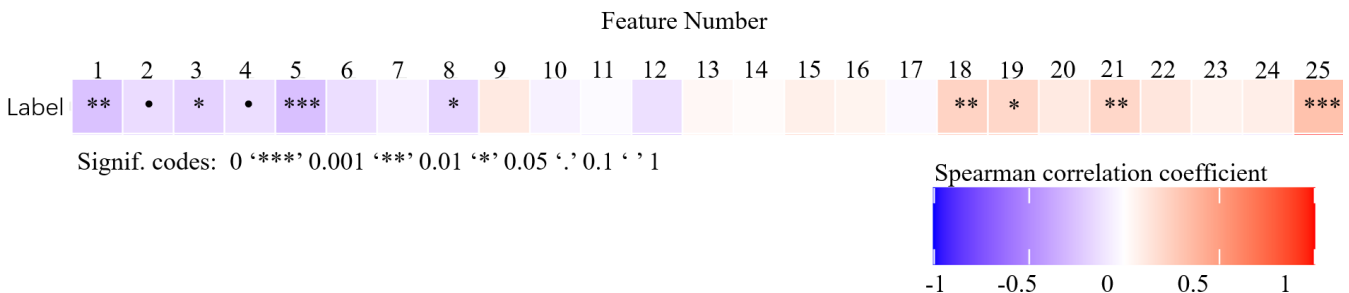


図 4: Spearman の相関分析: 特徴量 #1-6, 特徴量 #8, 特徴量 #12 は陽性例と負の相関を持つことがわかる. 一方, 特徴量 #18-22 と特徴量 #25 は嚥下障害と正の相関がある.

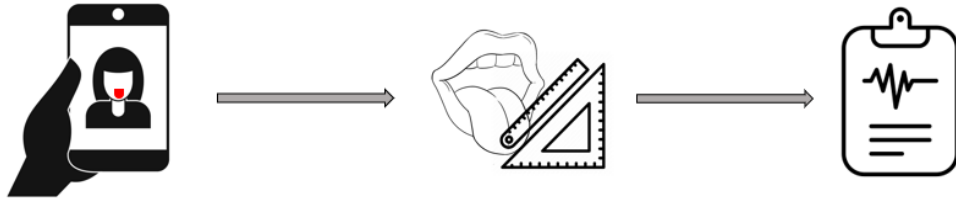


図 5: 本研究の結果を用いて構築すること想定しているモバイルシステム。ユーザはまず、スマートフォンのフロントカメラで決められた手順で動画を撮影する。その後、システムはさまざまな特徴量を定量化する。その特徴量を SVM に入力し、予測結果に基づいて評価レポートを作成する。

メラで 5 つの口腔内を動かすタスクを記録する。次に、顔ランドマーク検出器により特徴量抽出を行う。そして、その結果を機械学習アルゴリズムに入力し、モデルの予測に基づき、評価レポートを作成する。

ユーザが自分でキーポイントをラベリングすることは面倒である可能性がある。そこで、一般的な顔のランドマーク検出器をスマホ上で動作させる [33], [34]。これにより少なくとも鼻、唇、頬を検出することができる。そして、これらのキーポイントのみをクラウドに送信して解析することで、ユーザのプライバシーを確保することができる。しかし、舌の検出については、まだ十分に開発されたモデルがないので、可能であれば、舌のキーポイントを追跡するための舌検出アルゴリズムを独自に構築したいと考えている。できなかった場合、ユーザが手動で舌のキーポイントにラベルを付ける必要がある。

8. まとめと今後の課題

我々はスマートフォンを用いた視覚に基づく口腔・嚥下能力を定量化するアプローチを提案する。我々のモデルは、5 つの動作タスクから抽出された 25 の特徴量に基づき、妥当な性能で嚥下障害を予測することができる。これにより、自宅での嚥下障害の追跡が可能となり、医師に追加情報を提供できる可能性がある。

本研究にはいくつか課題がある。まず、5 つの動作タスクのみで嚥下障害を予測するデータで良好な性能を得ることができたが、嚥下障害は非常に複雑な疾患であり、医師が様々な観点から慎重に検討する必要がある。我々のアプローチよりも複雑な評価方法を用いることで、より信頼性の高い結果が得られる可能性がある。また、前述のように、今回の研究では、脳神経 IX と X の評価は含まれていない。現実には、嚥下障害の原因として考えられるものは、すべて医師によって検討される必要がある。したがって、我々のアプローチは、医師による嚥下障害の臨床評価に取って代わるものではない。第二に、我々のデータはすべて 2 次元の動画であるが、実際、画像における深度マップは動きや長さを計測するのに役立つので考慮する価値がある。スマートフォンの最新モデルには専用の深度カメラが搭載されているものもあるが、ほとんどのスマートフォンはそ

うではない。このような制約があること、また、スマートフォンを利用したアプローチを実現することが目的であることから、本研究では深度マップを考慮しないこととした。また、嚥下障害陽性例のサンプル数は比較的少ない。今後の研究では、より多くの嚥下障害陽性例を対象とした大規模な研究を行う必要がある。これにより、機械学習モデルの性能が向上するだけでなく、医学的な意味を見出す際に、より多くの変数の相互作用を伴う多重ロジスティック回帰などの統計分析を行うことができるようになる。

謝辞

システム設計やデータ収集の過程でアドバイスをいただいた浜松医科大学の長島優先生に厚く御礼申し上げます。

参考文献

- [1] U.S. National Library of Medicine NIH (National Institute of Health): Dysphagia: Swallowing Disorders, 2016. (<https://medlineplus.gov/swallowingdisorders.html>).
- [2] 西尾正輝: AMSD 標準ディサースリア検査 (Assessment of Motor Speech for Dysarthria), インテルナ出版, 2004.
- [3] 西尾正輝, 阿部尚子, 岡本卓也, 福永真哉: 標準ディサースリア検査の嚥下障害への臨床的応用の試み: AMFD の開発, Japan Journal of Clinical Research in Dysarthria, Vol.6 No.1 pp4-10, 2016.
- [4] Yuka Kamijo, Eiichiro Kanda, Keisuke Ono, Keizo Maeda, Akane Yanai, Kazuya Honda, Ryuji Tsujimoto, Mai Yanagi, Yoshitaka Ishibashi, and Masayuki Yoshida. 2018. Low tongue pressure in peritoneal dialysis patients as a risk factor for malnutrition and sarcopenia: a cross-sectional study. Renal Replacement Therapy 4, Dec 2018.
- [5] IOPI Medical. Iowa Oral Performance Instrument (IOPI), 2001. (<https://iopimedical.com/>).
- [6] 株式会社ジーシー: JMS 舌圧測定器 TPM-02. (<https://www.gcdental.co.jp/sys/data/item/1197/>).
- [7] Shichao Yue and Dina Katabi: Liquid Testing with Your Smartphone. In Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys 2019.
- [8] Alex Mariakakis, Megan A. Banks, Lauren Phillipi, Lei Yu, James Taylor, and Shwetak N. Patel: BiliScreen: Smartphone-Based Scleral Jaundice Monitoring for Liver and Pancreatic Disorders. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies (IMWUT). 2017.
- [9] Mayank Goel, Elliot Saba, Maia Stiber, Eric Whitmire, Josh Fromm, Eric C. Larson, Gaetano Borriello, and Shwetak N. Patel: SpiroCall: Measuring Lung Function

- over a Phone Call. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016.
- [10] Eric C. Larson, Mayank Goel, Gaetano Borriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N. Patel: SpiroSmart: using a microphone to measure lung function on a mobile phone. In The 2012 ACM Conference on Ubiquitous Computing, UbiComp '12, 2012.
- [11] Soujanya Chatterjee, Alexander Moreno, Steven Lloyd Lizotte, Sayma Akther, Emre Ertin, Christopher P. Fagundes, Cho Lam, James M. Rehg, Neng Wan, David W. Wetter, and Santosh Kumar: SmokingOpp: Detecting the Smoking ' Opportunity' Context Using Mobile Sensors. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies, 2020.
- [12] Sangwon Bae, Denzil Ferreira, Brian Suffoletto, Juan C. Puyana, Ryan Kurtz, Tammy Chung, and Anind K. Dey: Detecting Drinking Episodes in Young Adults Using Smartphone-based Sensors. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies, 2017.
- [13] Alex Mariakakis, Sayna Parsi, Shwetak N. Patel, and Jacob O. Wobbrock: Drunk User Interfaces: Determining Blood Alcohol Level through Everyday Smartphone Tasks. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 2018.
- [14] Paul Brazis Jose Biller, Gregory Gruener: DeMyer's The Neurologic Examination: A Programmed Text, Seventh Edition, 2017.
- [15] Juan Ignacio Godino-Llorente, Pedro Gomez-Vilda, and Manuel Blanco-Velasco: Dimensionality reduction of a pathological voice quality assessment system based on Gaussian mixture models and short-term cepstral parameters. IEEE transactions on biomedical engineering 53, 10, 2006.
- [16] Juan Ignacio Godino-Llorente and Pedro Gómez-Vilda: Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. IEEE Transactions on Biomedical Engineering 51, 2, 2004.
- [17] Younggeun Choi, Minjae Kim, Baekhee Lee, Xiaopeng Yang, Jinwon Kim, Dohoon Kwon, Sang-Eok Lee, HyangHee Kim, Seok In Nam, Saewon Hong, Giltae Yang, Duk L. Na, and Heecheon You: Development of an Ultrasonic Doppler Sensor-Based Swallowing Monitoring and Assessment System. Sensors, 2020.
- [18] PLIMES Inc: GOKURI, 人工知能が嚥下を測る, 2018. (<https://www.plimes.com/gokur>)
- [19] Edward Jay Wang, Junyi Zhu, Mohit Jain, Tien-Jui Lee, Elliot Saba, Lama Nachman, and Shwetak N. Patel: Seismo: Blood Pressure Monitoring Using Built-in Smartphone Accelerometer and Camera. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), 2018.
- [20] Reham Mohamed and Moustafa Youssef: HeartSense: Ubiquitous Accurate Multi-Modal Fusion-based Heart Rate Estimation Using Smartphones. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies, 2017.
- [21] WeXi Gu, Yuxun Zhou, Zimu Zhou, Xi Liu, Han Zou, Pei Zhang, Costas J. Spanos, and Lin Zhang: SugarMate: Non-intrusive Blood Glucose Monitoring with Smartphones. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies, 2017.
- [22] Chen-Hsuan (Iris) Shih, Naofumi Tomita, Yanick X. Lukic, Álvaro Hernández Reguera, Elgar Fleisch, and Tobias Kowatsch: Breeze: Smartphone-based Acoustic Real-time Detection of Breathing Phases for a Gamified Biofeedback Breathing Training. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies, 2019.
- [23] Edward Jay Wang, William Li, Doug Hawkins, Terry Gernsheimer, Colette Norby-Slycord, and Shwetak N. Patel: HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp, 2016.
- [24] Yuan Liang, Hsuan-Wei Fan, Zhujun Fang, Leiyang Miao, Wen Li, Xuan Zhang, Weibin Sun, Kun Wang, Lei He, and Xiang Anthony Chen: OralCam: Enabling Self-Examination and Awareness of Oral Health Using a Smartphone Camera. CHI Conference on Human Factors in Computing Systems (CHI '20), 2020.
- [25] Haroon Rashid, Sanjana Mendu, Katharine E. Daniel, Miranda L. Beltzer, Bethany A. Teachman, Mehdi Boukhechba, and Laura E. Barnes: Predicting Subjective Measures of Social Anxiety from Sparsely Collected Mobile Sensor Data. Proceedings of the ACM on Interactive Mobile Wearable and Ubiquitous Technologies, 2020.
- [26] Rui Wang, Weichen Wang, Min S. H. Aung, Dror Ben-Zeev, Rachel Brian, Andrew T. Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Emily A. Scherer, and Megan Walsh: Predicting Symptom Trajectories of Schizophrenia using Mobile Sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies, 2017.
- [27] Tsuneo KAWANO Yusuke TANAKA, Yutaka FUKUI: タブレット PC のモバイル性を活かした現場作業者用「腰痛判定カメラ」の開発. モバイル学会誌, 2012.
- [28] Hanbin Zhang, Chenhan Xu, Huining Li, Aditya Singh Rathore, Chen Song, Zhisheng Yan, Dongmei Li, Feng Lin, Kun Wang, and Wenyao Xu: PDMove: Towards Passive Medication Adherence Monitoring of Parkinson's Disease Using Smartphone-based Gait Assessment. Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies (IMWUT), 2019.
- [29] Sakai K, Gilmour S, Hoshino E, Nakayama E, Momosaki R, and Yoneoka D Sakata N: A Machine Learning-Based Screening Test for Sarcopenic Dysphagia Using Image Recognition. Nutrients 13, 11, 2021.
- [30] Belafsky PC, Mouadeb DA, Rees CJ, Pryor JC, Postma GN, Allen J, and Leonard RJ: Validity and reliability of the Eating Assessment Tool (EAT-10). Annals of Otolaryngology, Rhinology & Laryngology 117, 12 (2008), 2008.
- [31] AIST Digital Human Laboratory: AIST 人体寸法データベース 1991-92, 1992. (<https://www.airc.aist.go.jp/dhrt/91-92/>).
- [32] Carlo Tomasi and Takeo Kanade: Detection and Tracking of Point Features. International Journal of Computer Vision, 1991.
- [33] Google Inc.: Detect faces with ML Kit on Android. (<https://developers.google.com/>).
- [34] Apple Inc.: Tracking the User's Face in Real Time. (<https://developer.apple.com/>).