

ウェアラブルデバイスのマイクを用いた マスク装着状態の検知に向けて

小野 翔多^{1,a)} 西山 勇毅^{1,b)} 瀬崎 薫^{2,1,c)}

概要: 感染症予防において、マスクの装着は飛沫による感染症への感染リスクを低下させる有効な手段の一つである。日常生活中におけるマスク装着の有無やその種類を自動的に検出できれば、感染リスクの判定や Just-in-Time での注意喚起、行動記録など様々な応用サービスが実現可能になる。しかし、映像処理や専用機器を用いず、日常生活中において自動的にマスクの装着状態を検知する手法はまだ提案されていない。本研究では、市販のスマートウォッチの内蔵マイクのみを用いてマスクの装着状態を検出する。マスク装着時の音声特性調査とマスク装着状態判定モデルの評価実験から、マスク装着時・未装着時の音声データと機械学習を用いてマスク装着状態を検知できる可能性が示唆された。

キーワード: 感染症予防, マスク装着情報, スマートウォッチ, 機械学習, 音声データ

1. はじめに

新型コロナウイルス感染症 (COVID-19) の日本国内の感染者数は 2022 年 1 月 9 日時点で 240 万人以上、死亡者数は累計 1 万 8 千人以上となるなど、感染爆発から約 2 年経過してもなお猛威を奮っている [1]。COVID-19 は新型コロナウイルス (SARS-CoV-2) に感染することで発症する感染症のことであり、主に飛沫・接触感染によって感染する [2]。ウイルスの潜伏期間は 1 から 14 日 (多くは 5 日程度) と言われており、感染に無症状の状態でも他者に感染させる可能性がある。そのため、感染爆発を抑えるためには、全ての人が飛沫・接触感染の機会を減らすことが重要になる。

主な感染経路である飛沫感染の可能性を低下させるために、マスクの装着が推奨されている [3]。例えば、COVID-19 の感染拡大の抑制を目的に、公共交通機関や公共スペース、職場などにおけるマスクの装着の義務付けなどが行われている。また感染時に行われる積極的疫学調査において、マスク装着の有無は感染リスクの一つとして調査項目になっている。マスクには、一般家庭で利用される家庭用の「布マスク」や「不織布マスク」、または医療用の「N95 マス

ク」などが挙げられる。マスクの基本性能は、フィルタ部のろ過効率で評価される。一般的に、N95 マスクが最もろ過効率が高く、不織布、布マスクと続く。逆に呼吸のしやすさなどの快適性に関しては、布マスクが高く、N95 マスクが低くなる。飛沫のサイズは 3.0 から 5.0 μm と言われており、ろ過効率が十分高いマスクの装着により、飛沫の排出と吸引量を減少できる [4]。

感染症予防行動を促進するシステム [5] を構築する上で、特別な操作なしに「マスク装着の有無」を自動的に検出できれば、マスク装着を促す様々なアプリケーションの構築が容易になる。例えば、Just-in-Time での感染症予防行動支援や感染リスクの評価、マスク装着情報に応じた特典の付与など、様々な応用アプリケーションが実現できる。既存研究では、画像処理を用いたマスク装着状態の検知手法 [6], [7], [8] が提案されているが、日常生活中に常に動画を記録することは現実的に難しい。また、マスク装着時と非装着時の音声の違いから、マスク装着状態の検知を行う研究 [9] も行われているが、データセットと計測環境、計測手法に偏りがあり、現実環境で利用できるかは疑問が残る。

本稿では「スマートウォッチを用いたマスク装着状態検知の可能性」について調査する。具体的には、スマートウォッチに搭載されたマイクのみを用いて、マスク装着時と非装着時の音声の違いから、マスク装着状態の検知が可能であるかを調査する。

まず 2 章にて本稿の関連研究について整理する。3 章にて関連研究の課題を整理した上で、本稿の目的について述

¹ 東京大学生産技術研究所

IIS, The University of Tokyo

² 東京大学空間情報科学研究センター

CSIS, The University of Tokyo

a) shota@mcl.iis.u-tokyo.ac.jp

b) yuukin@iis.u-tokyo.ac.jp

c) sezaki@iis.u-tokyo.ac.jp

べる。マスク装着時の音声特性調査を4章にて行い、マスク素材の違いが音声の周波数に与える影響とスマートウォッチ内蔵マイクの高周波成分の録音性能を明らかにする。5章では、これまでの調査結果を踏まえてマスク装着状態を検知する機械学習モデルのプロトタイプを実装し、その性能を評価する。評価結果をもとに本稿のアプローチの可能性と限界点にて6章で整理する。最後に7章にて本稿をまとめる。

本稿の貢献は以下の通りである。

- スマートウォッチのマイクを用いてマスク装着時と非装着時の音声を記録し周波数成分の違いを調査したこと
- マスクの素材の違いが音声の周波数成分に与える影響を調査したこと
- スマートウォッチの内蔵マイクを用いたマスク装着状態検知の可能性と課題について整理したこと

2. 関連研究

本章では、まず既存の画像と音声データを用いたマスク装着状態の検出手法について述べ、各手法の利点と課題について整理する。ついでウェアラブルデバイス、特にスマートウォッチを用いた、自動的な人の状態検知に関する取り組みと、その応用可能性について整理する。

2.1 マスク装着状態の検知

画像内に映った人々のマスク装着・非装着状態を機械学習を用いて検出する手法が提案されている [6], [7]。Lodhらの手法 [6] では、マスク検出モデルと人物識別モデルを組み合わせることで、1枚の画像内から複数人のマスク装着者を検出できるほか、動画内からリアルタイムにマスクの装着・非装着を検出できる。また、同様に画像から「マスクを装着している」、「マスクを装着しているが正しい付け方ではない」、「マスクを装着していない」の三段階で、マスク装着状態を識別する手法も提案されている [7]。これらの手法は、医療現場での誤ったマスク装着状態の早期発見やマスク装着時のみ利用できる公共サービスの提供において利用できる。しかし、マスク装着状態の検知には画像が必要なため、日常生活中において個人々のマスク装着状態を常時記録する用途には適していない。

マイクで収集した音声を用いて、マスクの装着・非装着を検出する手法も提案されている [9], [10]。マスク装着時の音声は通常時と比較すると変化が生じており [11], [12]、特に1kHz以上の高い周波数を減衰させることが明らかになっている [13]。Mallol-Ragoltaらの研究 [9] では、マスクが持つ音のフィルタ特性を活用し、1kHz以上の周波数を含む音声データを用いることで、マスク装着・非装着の Unweighted Average Recall (UAR) が66.42%から70.70%に向上することを明らかにした。しかし、Mallol-Ragolta

が利用したデータセット [14] は、N95マスクのみが利用されており、一般向けの布マスクや不織布マスクを用いた場合の結果は未知数である。また音声の計測も据え置きのコデンサマイクで実施されており、人々が日頃携帯するスマートフォンやスマートウォッチに搭載されたマイクを用いて、マスク装着の検知に必要な音の特徴量を収集できるかは、まだ分かっていない。

2.2 スマートウォッチを用いた人の状態検知

スマートウォッチに代表されるウェアラブルデバイスの出荷台数は年々増加している [15]。特にスマートウォッチには、スマートフォンと同様に多種多様なセンサ類が搭載されており、それらを活用した人の状態検知手法が多く提案されている [16], [17], [18]。例えば、Samyounらの研究 [16] では、スマートウォッチに搭載されたモーションセンサを用いることで、手洗い時間の計測とその質の判定を実現している。また多くのスマートウォッチには、内蔵マイクも搭載されており、内蔵マイクを用いて収集した音から会話時間の計測 [18] や尿流量を測定する手法 [17] も提案されている。このようにスマートウォッチは、豊富なセンサ類と計算能力を内蔵しており、人の状態を検知する上で有用な基盤であると言える。しかしこれまでのところ、スマートウォッチを用いてマスク装着状態を検出する研究は行われていない。

3. 本研究の目的

日常生活の自然な行動内から自動的にマスク装着状態を検出できれば、Just-in-Timeでの注意喚起や感染リスクの判定、マスク装着情報のライフログ化など様々な応用が実現できる。近年、高性能なセンサと計算機構を搭載したスマートウォッチの普及が進んでいる [15]。スマートウォッチでは、端末内のセンサと計算機構を用いて歩数や手洗いの状態 [16]、会話時間 [18] など様々な人の行動を常時検出・蓄積できる。スマートウォッチは、行動認識基盤として既に日常生活中に利用されており、スマートウォッチを用いてマスク装着状態を自動的に検出できれば、幅広いユーザー層や場面でマスク装着情報を利用可能になる。

関連研究では、画像 [6] や音声データ [9] からマスク装着状態を取得する手法が提案されている。画像処理による手法 [6] は、顔が写った画像のみから高い精度でマスク装着状態を検知できる。しかし、マスク装着状態の検知には、常に自身の顔を撮影する必要があるため、日常生活中に常時利用することは困難である。また、音声を用いた手法 [9] では、N95マスク装着時と非装着時の音声データを用いて機械学習モデルを構築し、マスク装着状態を判定している。しかし、学習データとして利用されている音声は、据え置きの高性能なコンデンサマイクを用いて、医療用のN95マスクの装着時と非装着時の音声のみ集音している。そのた

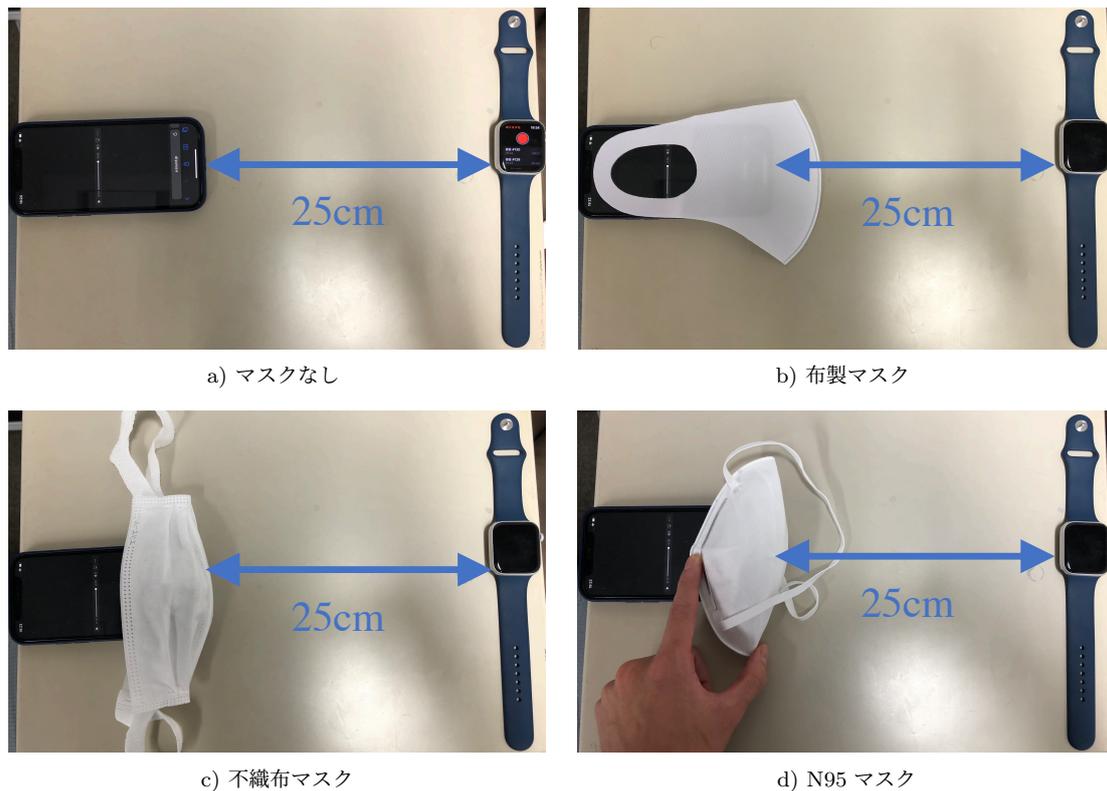


図 1: 実験の様子

め、音声を用いた手法が、市販のウェアラブルデバイス上で動作し、かつ布マスクや不織布マスクなどの一般向けマスクでも動作するかは不明である。

本研究は、市販のスマートウォッチに搭載されているセンサを用いてマスクの装着状態を検知することを目的とする。特に多くのスマートウォッチに内蔵されている「マイク」を用いて、マスク装着時と非装着時の音声の変化からマスク装着状態を検出する。既存研究 [9] から、マスクによって高周波の音声成分が吸収されることと、機械学習を用いて N95 マスクの装着・非装着をある程度判定できる。しかし、前述したように既存研究で利用されているデータセットは偏りが大きく、収集されているデータセットの内容と環境も限定的であるため、本研究の想定環境において利用できるかは明らかでは無い。

そこで本研究では目的達成に向けて、本研究が想定する環境下におけるマスク装着時の音声特性調査を実施する。具体的には、(調査項目 1) 布・不織布・N95 マスク装着時の周波数成分への影響調査と、(調査項目 2) スマートウォッチ内蔵マイクの高周波成分の録音性能の調査である。その上で、スマートウォッチより収集したマスク装着・非装着時の音声データを用いて、マスク装着状態を判定する機械学習モデルのプロトタイプを作成し、その可能性について議論する。

4. マスク装着時の音声特性調査

既存研究 [9] より、マスクの装着は音声、特に高周波成分に影響を与えることが明らかになっている。しかし既存研究の調査手法は本研究の想定環境とは大きく異なるため、本研究の想定環境において同様の現象が観測されるかを調査する必要がある。本節では「布・不織布・N95 マスク装着時の周波数成分への影響調査」と「スマートウォッチ内蔵マイクの高周波成分の録音性能」の調査結果について述べる。なお本稿では、スマートウォッチとして Apple Watch SE を使用し、音声の録音も Apple Watch の内蔵マイクを用いた。マスクは、「布マスク」としてポリエステル製マスク、「不織布マスク」としてメルトブローン不織布（ウイルス飛沫ろ過試験：99%カット）製の三層構造マスク、「N95 マスク」として米国国立労働衛生研究所認定の N95 対応規格マスクをそれぞれ利用した。

4.1 布・不織布・N95 マスク装着時の周波数成分への影響調査

マスク装着時と非装着時の周波数成分の変化を、異なる種類のマスクを通して測定した。音源は、周波数成分への影響を明らかにするために、スマートフォンのアプリを用いて生成した機械音を用いた。周波数は 1000 Hz と 10000 Hz の 2 種類を用いた。実験の様子を図 1 に示す。同じ音量、音程の音声データを用いるために、事前に録音しておいた

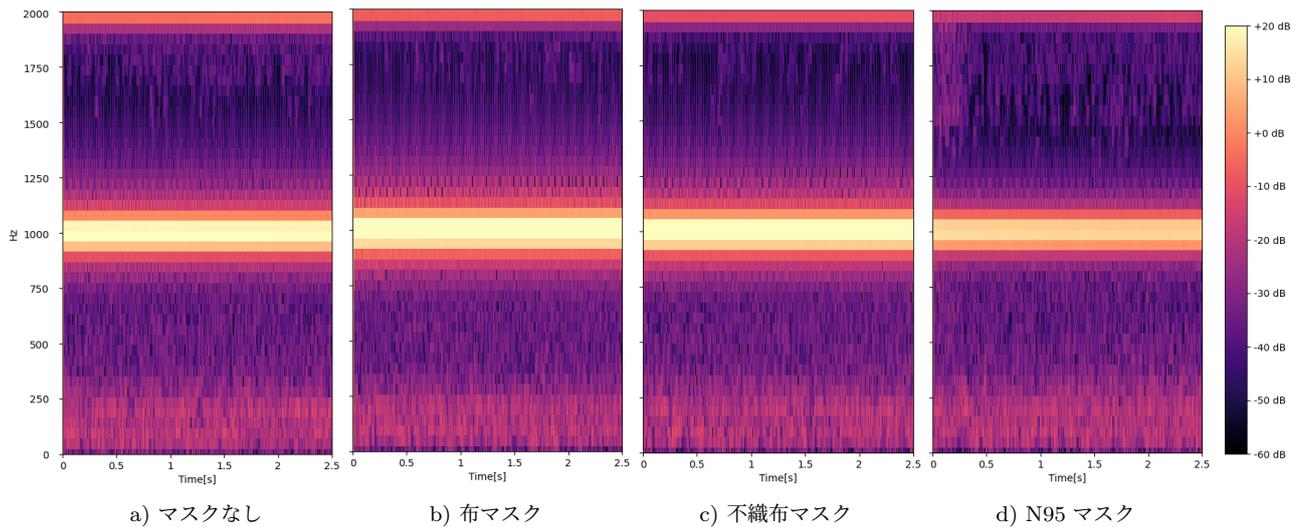


図 2: 1000Hz

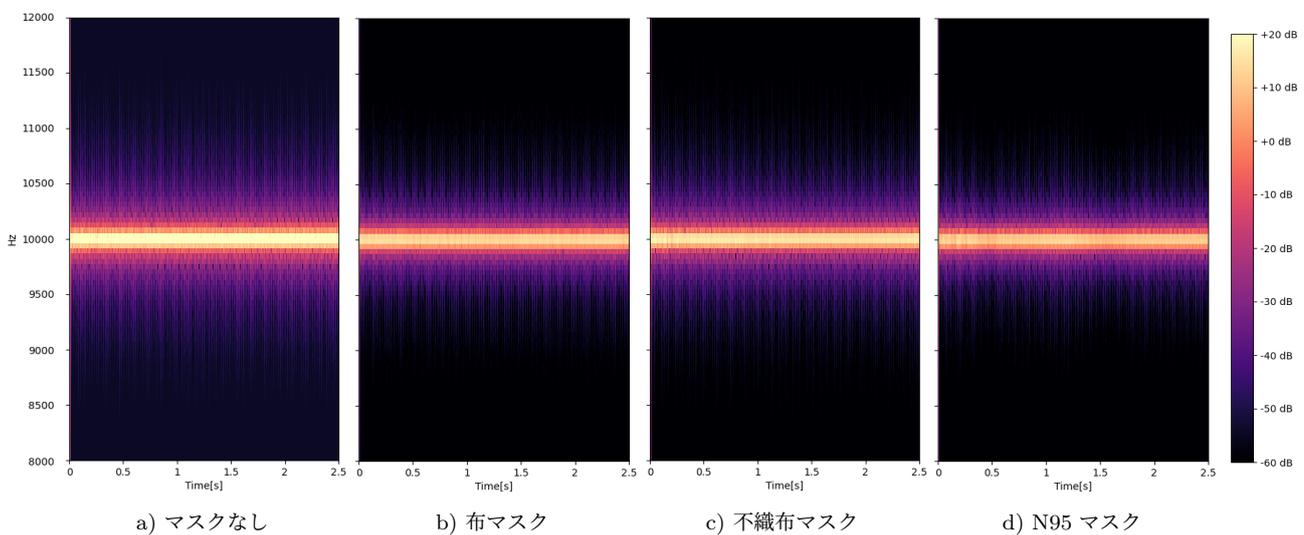


図 3: 10000Hz

音声データを iPhone のスピーカから流して、25 cm 離れた場所に置いてあるスマートウォッチのマイクで計測する。この計測をそれぞれの周波数において、スピーカに対してマスク装着なし、不織布マスク装着、布マスク装着、N95 マスク装着の 4 種類で実行した。3 種類のマスクを用いた実験の際は、マスクによる周波数成分の変化を測定するために、スピーカにそれぞれマスクを覆い被せて録音した。

図 2 に周波数 1000 Hz で測定した音声データにおける 2.5 秒間のスペクトラムを示す。横軸が時間の経過を示し、縦軸は周波数成分を示す。1000 Hz 帯に着目すると、マスクなし、布マスク、不織布マスクでは周波数成分の変化はあまりみられない。一方 N95 マスクを用いた実験では他の 3 種類の測定データと比較して 1000 Hz 帯の周波数成分が少ない。結果から、周波数 1000 Hz 帯では N95 マスクを用いた時にのみ、周波数成分が変化していると考えられる。

従来手法ではコンデンサマイクを用いた計測で同様の結果を得られていたが、今回の実験を通してスマートウォッチで取得した音声データでもそれが確かめられた。また、周波数 1000 Hz 帯では、布マスクや不織布マスクを用いた場合、マスク装着状態の検知は難しいと考えられる。

図 3 に周波数 10000 Hz で測定した音声データにおける 2.5 秒間のスペクトラムを示す。図 2 と同様に横軸が時間の経過を示し、縦軸は周波数成分を示す。周波数 10000 Hz 帯では、マスクなしと比較して、布マスク、不織布マスク、N95 マスクの 3 種類とも周波数成分が変化していることが読み取れた。結果から、高周波数の場合は、N95 マスクのような高性能マスク以外でも周波数に変化が見られた。

この結果から、マスクの素材に関わらず、低周波数よりも高周波数の方がマスクによる周波数成分の変化が大きく現れるため、マスク装着状態の検知には高周波の音に着目

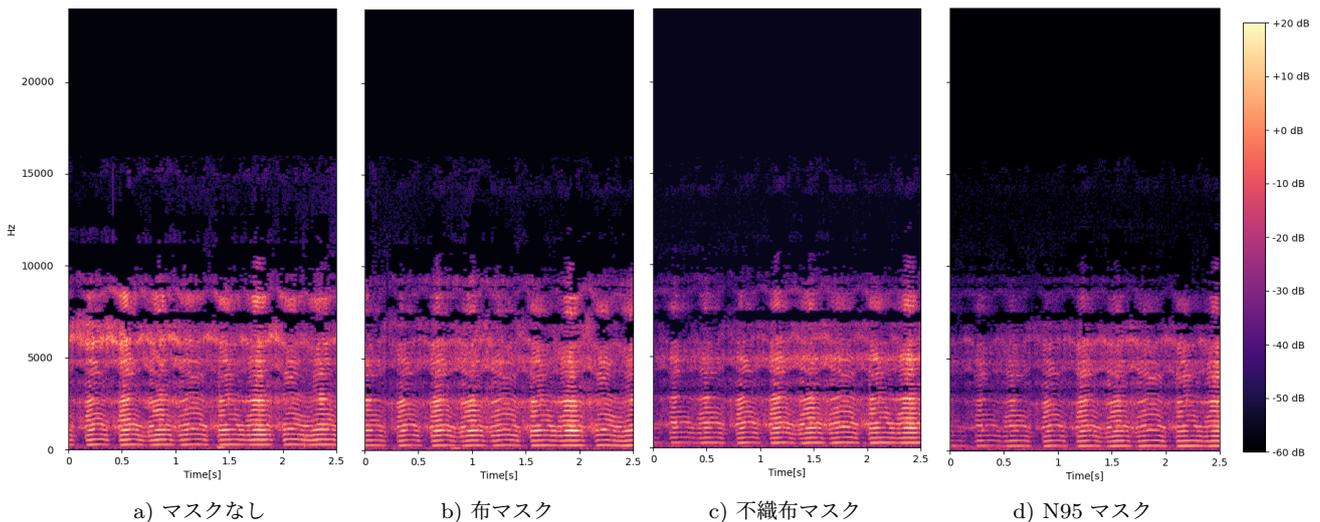


図 4: High freq. の音声データ用いた際のスペクトラム

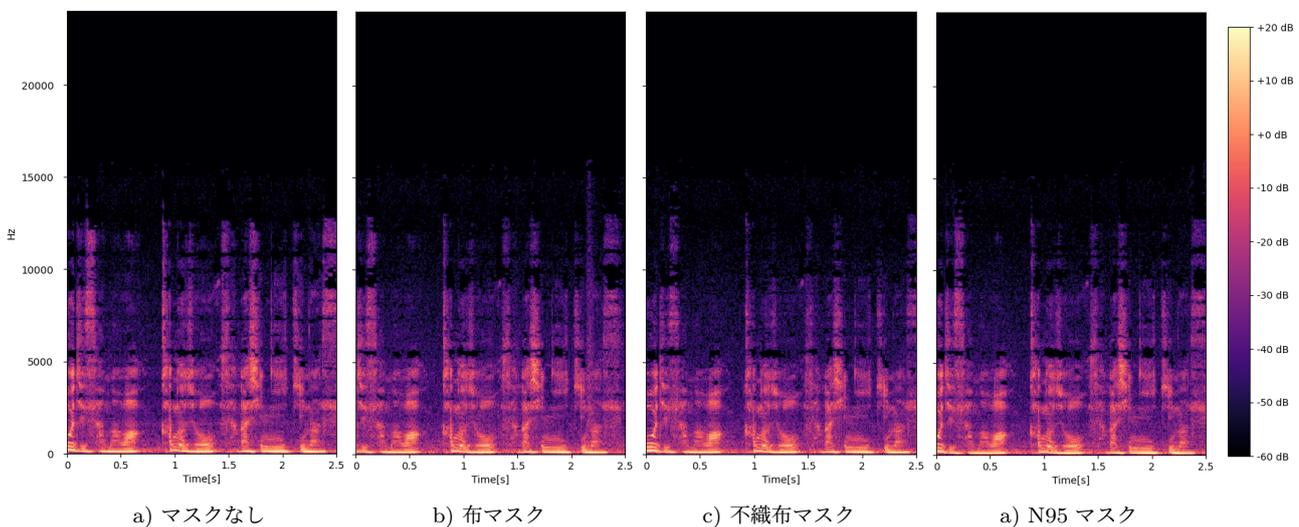


図 5: Mix freq. の音声データ用いた際のスペクトラム

すべきだという知見が得られた。従来研究では、N95 マスクが周波数成分に与える影響のみが調査されており、他の種類のマスクが与える影響までは明らかになっていなかった。本調査により、他の種類のマスク（不織布や布）でも N95 マスクと同様に、高周波数成分が変化することが明らかになった。

4.2 スマートウォッチ内蔵マイクの高周波成分の録音性能

スマートウォッチ内蔵マイクの録音性能、特に「人の音声」の録音性能を調査した。4.1 節で行った結果より、高周波の音が機械学習の判定精度に大きく影響すると予想できる。そのため、スマートウォッチ内蔵マイクを用いて収集した実際の人の音声に、どの程度高周波の音が含まれるかを調査する。

調査では、まず「意識的に高い音で話した音声 (High

freq.)」と「通常の音声 (Mix freq.)」の 2 種類の音声を iPhone で録音した。次に録音した音声データをスピーカから流し、スピーカに対してマスク装着なし、不織布マスク装着、布マスク装着、N95 マスク装着の 4 種類において、スマートウォッチのマイクを用いて音声を取得した。マスク装着の有無によって、2 種類のデータセットで周波数成分にどのような変化が現れているのかについて調査する。

図 4 に高周波数成分を録音したデータセットの中から一つを選出し、マスク装着状態別に 4 種類の音声データ 2.5 秒間のスペクトラムを示す。マスクなしと比較して、布マスク、不織布マスク、N95 マスクの周波数が変化していることが確認できる。特に 10 000 Hz から 15 000 Hz にかけての周波数が大きく変化している。マスク同士を比較すると N95 マスクが最も周波数の変化が顕著に表れており、次点で不織布マスク、最後に布マスクの順番となっている。

結果から、マスクの種類によって性能は異なるものの、マスクなしと比較して高周波数成分を中心にマスク装着状態の違いが読み取れることがわかった。

図 5 に低周波数成分を含む音声データを録音したデータセットの中から一つを選出し、マスク装着状態別に 4 種類の音声データ 2.5 秒間のスペクトラムを示す。マスクなしと比較して、マスクを装着している場合ではわずかながら、周波数が変化していることが確認できる。しかし、この結果は、図 4 と比較すると微小なものであり、これらからマスク装着状態を検出することは困難であると考えられる。

5. 機械学習を用いたマスク装着状態の分類

4 章の調査結果を踏まえて、「スマートウォッチの内蔵マイクを用いたマスク装着状態の検知モデル」のプロトタイプを実装する。学習データには、4.2 節で収集したマスク装着時・非装着時の音声データを利用した。検知モデルの実装には macOS の CreateML 3.0 を使い、特徴量抽出には VGGish を用いた。高周波数の音声データセット、低周波数を含む音声データセットとして、共に 5 秒程度の音声ファイルを 10 個用意し、8:2 の割合でトレーニングデータとテストデータにして交差検証した。

表 1 に 4 種類のマスク装着状態の分類結果を示す。マスクなし、布マスク、不織布マスク、N95 マスクをそれぞれ機械学習分類のための class として、その適合率 (Prec.) と再現率 (Recall) を機械学習によって算出した。ここで、適合率とはマスク装着状態が正であると予測したものがどれだけ正しかったかを見る指標である。つまり、適合率が高いほど検出したマスク装着状態検知の正確性が高いと言える。また、再現率とはマスク装着状態が実際に正であったもののうち、どれだけ正と予測できたかを見る指標である。つまり、再現率が高いほど検出したいマスク装着状態検知の損失が少ないと言える。

マスクなし、布マスク、不織布マスクにおいては、Mix freq. と比較して High freq. を用いた結果が適合率、再現率共に高くなっている。これは、高周波数成分ではどのマスクを用いた場合でも、周波数成分を変化させる性質があるという実験結果と一致している。一方、N95 マスクを用いた場合にのみ、High freq. を用いた場合よりも Mix freq. を用いた時の適合率及び再現率が高くなっている。これは、低周波数では、N95 マスクを用いた場合にのみ周波数の変化が見られたため、他 3 種類の測定データとの差異が現れたことが関係していると考えられる。

表 2 にマスクなしと布マスク分類結果を示す。表から、High freq. を用いた分類の方が、Mix freq. を用いた分類より、適合率・再現率共に高い結果を示していることがわかる。特に Mix freq. を用いた場合のマスクなしの分類結果が低くなっている。これは、Mix freq. では布マスクによる周波数成分の変化がほとんど見られず、周波数による分類

表 1: マスク装着状態の分類 (4 種類)

class	High freq.		Mix freq.	
	Prec.[%]	Recall[%]	Prec.[%]	Recall[%]
マスクなし	56.0	56.4	43.0	51.6
布	61.0	56.6	33.0	25.4
不織布	51.2	62.2	51.0	46.6
N95	56.6	61.8	68.2	67.4

表 2: マスク装着状態の分類 (マスクなし vs 布)

class	High freq.		Mix freq.	
	Prec.[%]	Recall[%]	Prec.[%]	Recall[%]
マスクなし	74.8	76.2	43.6	42.4
布	80.8	73.6	52.8	59.8

表 3: マスク装着状態の分類 (マスクなし vs 不織布)

class	High freq.		Mix freq.	
	Prec.[%]	Recall[%]	Prec.[%]	Recall[%]
マスクなし	73.4	76.8	77.6	77.8
不織布	78.0	66.2	77.2	66.4

表 4: マスク装着状態の分類 (マスクなし vs N95 マスク)

class	High freq.		Mix freq.	
	Prec.[%]	Recall[%]	Prec.[%]	Recall[%]
マスクなし	81.6	78.8	67.8	67.6
N95	82.2	77.2	66.2	55.2

が困難になっていることが関係していると考えられる。

表 3 にマスクなしと不織布マスクの分類結果を示す。表から、High freq. を用いた分類と Mix freq. を用いた分類共に同程度の分類結果を示していることがわかる。どちらの分類においても不織布マスクを用いた場合の再現率が最も低く見られる。これから、不織布マスクではマスクを装着しているのにも関わらず、装着していると判定されにくい特性を持っている可能性があることが読み取れる。

表 4 にマスクなしと N95 マスクの分類結果を示す。表から、High freq. を用いた場合 Mix freq. を用いた場合と比較して、高い適合率と再現率を示している。これは、N95 マスクを用いると高周波数成分の変化が顕著に現れ、マスクなしの場合の周波数成分との差異が大きくなっていることが関係していると考えられる。また、この結果は表 2 や表 3 と比較して高い精度のマスク装着状態の検知である。

以上の結果から、高周波数成分に着目することで、マスクの装着状態を検出することができることが示唆された。3 種類のマスクを用いた実験を行ったが、中でも N95 マスクを用いた場合に高い精度でマスク装着状態を検出できた。これは、周波数の変化がより顕著に現れるため、その影響は分類精度にも影響しているからだと考えられる。

6. 議論

6.1 内蔵マイクを用いたマスク装着状態検知の可能性

マスク装着の有無が音声の周波数成分に与える影響を調査した結果、既存研究と同様にマスクの材質に関わらず高周波の音を吸収する特性は見られた。これは高周波の音声であれば、スマートウォッチ内蔵のマイクを用いても、音声のみからマスク装着状態を検出できることを示唆している。さらに、音声には高周波の音も含まれており、高周波の音のみでマスク装着状態の判定モデルを実装すると、判定精度が向上することが分かった。これらの結果より、条件を整えばスマートウォッチな内蔵マイクでもマスク装着状態を検出できる可能性が高いと考えられる。

6.2 複数センサを用いたマスク装着状態検知

スマートウォッチには、内蔵マイクの他に加速度やジャイロ、心拍センサなど様々なセンサが搭載されている。内蔵マイクの音声だけでなく他のセンサデータも横断的に活用することで、マスク装着状態の検知精度を高めることができる可能性がある。例えば、モーションセンサやマイクを用いて、マスク装着時の「耳にマスクの紐を引っ掛ける動作」や「マスクの位置を調整する動作」、「一時的にマスクの位置を下げる動作」など、特徴的な動作を検知できれば、内蔵マイク単体よりも高い精度でマスク装着状態を検知できると考えられる。

6.3 本稿の限界点

本稿では高周波数の音に着目することで、マスク装着状態をスマートウォッチの内蔵マイクを用いて検出できる可能性があることを示唆した。しかし、今回の実験では実際の生活環境内での評価は行っておらず、本実験結果がそれらの環境において利用できるかは追加の評価実験が必要である。また、本稿で実装したマスク装着状態の判定モデルでは、1名の音声データのみを用いてマスクの判定モデルを実装しているため、異なる被験者に対して本モデルが有効であるかは明らかではない。今後は被験者数や計測場所を増やすことで、異なる性別や場所においてマスク装着状態を検出できるかを調査する。

7. おわりに

本稿では、ウェアラブルデバイスの内蔵マスクを用いたマスク装着状態の検知に向けて、マスク装着時の音声特性調査およびプロトタイプシステムの実装・評価を行った。調査結果より、スマートウォッチの内蔵マイクを用いた場合でも、従来手法と同様にマスクによる高周波成分のフィルタ効果を観測できた。また、マスクの素材に関わらず、高周波成分をフィルタしていることも明らかになった。調

査結果を踏まえて、スマートウォッチの内蔵マイクを用いてマスク装着状態を判定するプロトタイプシステムを実装した。その結果、高周波成分(8000Hz以上)の音声を多く含む場合には、約70%の精度でマスク装着状態を検出できた。本手法では、会話中から高周波数部分のみを抜き出す必要があり、それらの実装は今後の課題である。またスマートウォッチの内蔵マイクだけでなく、内蔵のモーションセンサなど複数センサと組み合わせることで、検知精度の向上が期待できる。

謝辞 本研究の一部は国立研究開発法人情報通信研究機構に支援いただいた。

参考文献

- [1] 厚生労働省: 新型コロナウイルス感染症について | 国内の発生状況など, <https://www.mhlw.go.jp/stf/covid-19/kokunainohasseijoukyou.html>. Accessed on 2022/01/30.
- [2] for Disease Control, C. and (CDC), P.: COVID-19 — How COVID-19 Spreads, <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html>. Accessed on 2022/01/30.
- [3] (WHO), W. H. O.: COVID19 — When and how to use masks, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks>. Accessed on 2022/01/30.
- [4] Fischer, E. P., Fischer, M. C., Grass, D., Henrion, I., Warren, W. S. and Westman, E.: Low-cost measurement of face mask efficacy for filtering expelled droplets during speech, *Science Advances*, Vol. 6, No. 36, p. eabd3083 (online), DOI: 10.1126/sciadv.abd3083 (2020).
- [5] Nishiyama, Y., Yonezawa, T. and Sezaki, K.: Self-Guard: Semi-Automated Activity Tracking for Enhancing Self-Protection against the COVID-19 Pandemic, *Proceedings of the 18th Conference on Embedded Networked Sensor Systems, Virtual Event, Japan, SenSys '20*, No. 2, New York, NY, USA, Association for Computing Machinery, p. 780–781 (online), DOI: 10.1145/3384419.3430592 (2020).
- [6] Lodh, A., Saxena, U., Khan, A., Motwani, A., Shakkera, L. and Sharmasth, V. Y.: Prototype for Integration of Face Mask Detection and Person Identification Model – COVID-19, *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1361–1367 (online), DOI: 10.1109/ICECA49313.2020.9297399 (2020).
- [7] Jiang, M. and Fan, X.: RetinaMask: A Face Mask detector, *CoRR*, Vol. abs/2005.03950 (online), available from <https://arxiv.org/abs/2005.03950> (2020).
- [8] Kumar, A., Kaur, A. and Kumar, M.: Face Detection Techniques: A Review, *Artif. Intell. Rev.*, Vol. 52, No. 2, p. 927–948 (online), DOI: 10.1007/s10462-018-9650-2 (2019).
- [9] Mallol-Ragolta, A., Liu, S. and Schuller, B. W.: The Filtering Effect of Face Masks in their Detection from Speech, *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 2079–2082 (online), DOI: 10.1109/EMBC46164.2021.9630634 (2021).
- [10] Mohamed, M. M., Nessiem, M. A., Batliner, A.,

- Bergler, C., Hantke, S., Schmitt, M., Baird, A., Mallol-Ragolta, A., Karas, V., Amiriparian, S. and Schuller, B. W.: Face mask recognition from audio: The MASC database and an overview on the mask challenge, *Pattern Recognition*, Vol. 122, p. 108361 (online), DOI: <https://doi.org/10.1016/j.patcog.2021.108361> (2022).
- [11] Everest, F. A., Pohlmann, K. C. et al.: *Master handbook of acoustics*, Vol. 4, McGraw-Hill New York (2015).
- [12] Montacé, C. and Caraty, M.-J.: Phonetic, Frame Clustering and Intelligibility Analyses for the INTERSPEECH 2020 ComParE Challenge, *Proc. Interspeech 2020*, pp. 2062–2066 (online), DOI: [10.21437/Interspeech.2020-2243](https://doi.org/10.21437/Interspeech.2020-2243) (2020).
- [13] Corey, R. M., Jones, U. and Singer, A. C.: Acoustic effects of medical, cloth, and transparent face masks on speech signals), *The Journal of the Acoustical Society of America*, Vol. 148, pp. 2371 – 2375 (2020).
- [14] Schuller, B. W., Batliner, A., Bergler, C., Messner, E.-M., Hamilton, A., Amiriparian, S., Baird, A., Rizos, G., Schmitt, M., Stappen, L., Baumeister, H., MacIntyre, A. D. and Hantke, S.: The INTERSPEECH 2020 Computational Paralinguistics Challenge: Elderly Emotion, Breathing Masks, *Proc. Interspeech 2020*, pp. 2042–2046 (online), DOI: [10.21437/Interspeech.2020-0032](https://doi.org/10.21437/Interspeech.2020-0032) (2020).
- [15] Counterpoint Technology Market Research: Global Smartwatch Shipments Jump 35% YoY in Q1 2021, <https://www.counterpointresearch.com/global-smartwatch-shipments-q1-2021/>. Accessed on 2022/01/30.
- [16] Samyoun, S., Shubha, S. S., Sayeed Mondol, M. A. and Stankovic, J. A.: iWash: A smartwatch hand-washing quality assessment and reminder system with real-time feedback in the context of infectious disease, *Smart Health*, Vol. 19, p. 100171 (online), DOI: <https://doi.org/10.1016/j.smhl.2020.100171> (2021).
- [17] Arjona, L., Enrique Diez, L., Bahillo Martinez, A. and Arruza Echevarria, A.: UroSound: A Smartwatch-based Platform to Perform Non-Intrusive Sound-based Uroflowmetry, *IEEE Journal of Biomedical and Health Informatics*, pp. 1–1 (online), DOI: [10.1109/JBHI.2022.3140590](https://doi.org/10.1109/JBHI.2022.3140590) (2022).
- [18] 小松勇輝, 下条和暉, 西山勇毅, 瀬崎薫: 腕時計型ウェアラブルデバイスを用いた会話時間計測手法の構築に向けて (To appear), 情報処理学会第 84 回全国大会 (愛媛大学), 情報処理学会, (オンライン), 入手先 (<https://www.ipsj.or.jp/event/taikai/84/index.html>) (2022).