

## 疾患サブタイプを考慮したアンサンブル学習に基づく 少音読課題音声からのパーキンソン病検出

伊藤 有生†1 加藤 昇平†2 佐久間 拓人†2  
大嶽 れい子†3 榎田 道人†4 渡辺 宏久†3

†1 名古屋工業大学 工学部 情報工学科 †2 名古屋工業大学 大学院工学研究科 工学専攻  
†3 藤田医科大学 医学部 脳神経内科学 †4 名古屋大学 大学院医学系研究科 神経内科

### 1 はじめに

パーキンソン病 (Parkinson's Disease : PD) はアルツハイマー病に次いで 2 番目に多い神経変性疾患である。PD の診断方法は既に確立されているが、専門医による診断が必要であるため、検査への抵抗感や医療従事者にとって検査の時間的負担が大きい問題がある。こうした理由から、PD の診断を補助するスクリーニングシステムが必要である。近年では、発声特徴を用いた機械学習によるスクリーニング手法が提案されている。Suppa ら [1] は PD の構音障害に着目して、単母音の連続発声から機械学習モデルを構築し、PD 検出の有効性を示唆した。坂口ら [2] は認知負荷が小さく発話を促しやすいとされる音読課題の遂行音声から認知症を判別するシステムを提案した。単語および文章を読み上げる音声から認知症希少疾患を判別し、認知症検出の可能性を示唆しているが、PD では未検証である。本研究では、坂口らと同様の音読課題を用いて、課題遂行音声による PD の簡易検出手法を提案する。

### 2 方法

#### 2.1 実験協力者

実験には、大脳皮質基底核変性症 (CBS) 8 名、多系統萎縮症 (MSA) 17 名、パーキンソン病 (PD) 81 名、健常者 (HC) 94 名の計 200 名が参加した。CBS および MSA は PD の疾患サブタイプであり、PD と似た症状を持つ疾患である。HC の対象者は、認知機能検査により定め、ミニメンタルステート検査  $\geq 26$  点かつ修正版 Addenbrooke の認知検査 (ACE-R)  $\geq 89$  点を満たす者とした。また発話に関する特徴は性別や年齢によって影響されるため、HC の年齢および性別を疾患群に合うように統制した。

#### Parkinson's disease detection from a few oral reading task speech based on ensemble learning considering disease subtypes

Yuki ITO†1, Shohei KATO†2, Takuto SAKUMA†2, Reiko OHDAKE†3, Michihito MASUDA†4, Hirohisa WATANABE†3  
†1 Dept. of Computer Science, Nagoya Institute of Technology  
†2 Dept. of Engineering, Graduate School of Engineering, Nagoya Institute of Technology  
†3 Dept. of Neurology, Fujita Health University School of Medicine  
†4 Dept. of Neurology, Nagoya University Graduate School of Medicine  
†1†2 Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan  
{ito, shohey}@katolab.nitech.ac.jp

表 1: 音読課題の一覧

課題番号	課題
課題 1	まど
課題 2	パイプ
課題 3	電話
課題 4	なす
課題 5	ばなな
課題 6	33
課題 7	ゆきだるま
課題 8	25 パーセント
課題 9	92 分の 1
課題 10	とけいえんぴつ
課題 11	電話が鳴っています
課題 12	魚屋は元気でした
課題 13	兄はまだ戻りません
課題 14	日本高校野球連盟
課題 15	だけどやっぱりでもはだめ
課題 16	新しい甘酒を 5 本のひょうたんに入れなさい

#### 2.2 実験環境及び収録機器

音声のデータは名古屋大学医学部附属病院、名古屋大学医学部保健学科大幸キャンパス並びに大阪大学医学部附属病院にて専門の言語聴覚士により収集された。録音機器にはオーディオテクニカ製の単一指向性の卓上マイク AT9921 を用いた。

#### 2.3 実施手順

ACE-R および WAB 失語症検査で実施される検査項目から、発話特徴量の抽出に適切と思われる項目を選択し、それらに対する回答音声を録音した。選択された課題のうち、本研究では音読課題を対象とした。音読課題とは WAB 失語症検査の復唱課題に「とけいえんぴつ」を追加した計 16 課題を順番にモニターに提示し、音読する課題である。表 1 に課題の一覧を示す。各課題ごとに  $n=200$  のサンプルを取得した。

#### 2.4 データ解析

##### 2.4.1 前処理

収録環境によるノイズによる影響を最小化するため、前処理としてノイズ除去を施した。ノイズ除去には周波数領域で雑音を除去する Spectral Subtraction 法を用いた。

##### 2.4.2 特徴量の定義

得られた音声データから、課題ごとに音響・言語・時間に関する特徴量をそれぞれ抽出する。特徴量について以下に詳述する。

##### ・音響特徴量 (384 種)

音響特徴量として INTERSPEECH 2009 Emotion

Challenge[3] で定義された 384 種の音響特徴量セットを用いる。これらの特徴量は音声解析ソフト openSMILE を使用して抽出する。

• 言語特徴量 (15 種)

発話文を Microsoft 社製の音声認識システムによりテキスト化し、MeCab を用いて形態素解析した。言語特徴量として各品詞 (動詞, 形容詞, 形容動詞, 名詞, 副詞, 連体詞, 接続詞, 感動詞, 助動詞, 助詞, 記号, フィラー) の割合と, 総語数, 異なり語数, 音声認識システムの出力である確信度の計 15 種を定義する。

• 時間特徴量 (3 種)

課題を提示してから回答までの反応時間, 発話時間, および総語数を発話時間で除した発話速度の計 3 種を時間特徴量として抽出する。

2.4.3 特徴選択

抽出した特徴量数はサンプル数よりも多いため, 訓練データを過学習するおそれがある。そこで, 特徴量数を削減し汎化性能を向上させるため, 赤池情報基準を用いた変数増加法により特徴選択した。特徴量数の上限を 100 とする。

2.5 モデル構築

PD と HC の判別モデルは, 音読課題ごとに作成した学習器を弱学習器とする重み付きアンサンブルモデルによって構築する。

はじめに各課題ごとに PD と HC を判別する学習器を作成する。学習器には RBF カーネルを用いた Support Vector Machine (SVM) を使用する。SVM のパラメータである  $C$  と  $\gamma$  は,  $[10^{-3}, 10^{-2}, \dots, 10^3]$  の範囲で grid search によりハイパーパラメータをチューニングした。

本研究では音読課題が 16 種あるが, スクリーニングの場面では時間短縮のため, より少ない課題で判別することが望ましい。また疾患によって, 判別に寄与する課題は異なる可能性が高いことから, 疾患サブタイプを考慮したアンサンブル学習モデルを構築する。課題の重みは「疾患ごとの F 値」によって決定する。各課題の判別モデルから, 疾患ごとに F 値を算出する。各疾患ごとに分類性能上位  $m (1 \leq m \leq 16)$  位に含まれる課題の数を学習器の重みとする。例えば  $m=1$  の場合, 各疾患で F 値が最も高い課題を選択する。

実験では,  $m$  の値を 1~16 まで変化させ判別性能が最良となる  $m$  を求める。汎化性能の評価には 5-fold Cross Validation (CV) を用いて, 評価指標として F 値, 感度, 特異度を算出する。

3 結果と考察

表 2 に各音読課題で構築された判別モデルの判別性能の平均値を示す。F 値, 感度, 特異度について, 平均値は 0.79, 0.78, 0.79 であった。各疾患ごとに F 値を算出し, 1 位から 16 位までのアンサンブル学習モデルを作成した結果,  $m=3$  (上位 3 位) で最良となり, F 値 0.91, 感度 0.91, 特異度 0.89 であった。上位 3 位では課題 3, 4, 6, 7, 10 の 5 つの課題が選択された。図 1 に最も F 値が高い課題 4 「なす」と上位 3 位のアンサンブル学習モデルの CV から算出された混同行列を

表 2: 課題ごとの判別性能の平均と標準偏差

評価指標	値
F 値	0.79 ± 0.04
感度	0.78 ± 0.06
特異度	0.79 ± 0.04

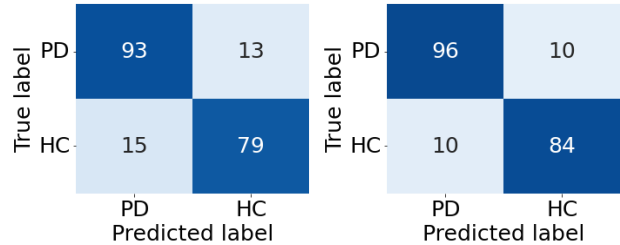


図 1: 混同行列 (n=200)

左: 最良課題: 「なす」

右: 上位 3 位のアンサンブル学習モデル

示す。図 1 より, アンサンブル学習モデルによって感度, 特異度ともに改善していることが分かる。

実験の結果, PD 検出モデルの構築に必要な課題数は 5 つであることが分かった。これらの課題は全て単語課題であり, 全 16 課題のうち 5 課題で実施可能なため, スクリーニング時間の短縮が期待される。従って「疾患ごとの F 値」上位 3 位までの学習器によるアンサンブル学習モデルが PD の簡易検出モデルとして適していると考えられる。今後サンプル数を増加させ, より検討を進めたい。

4 おわりに

本研究では少音読課題での PD スクリーニング手法を提案した。200 名の音読課題の回答音声から, 音響・言語・時間特徴量を抽出し課題ごとに学習器を作成し, 疾患サブタイプを考慮したアンサンブル学習モデルを検証した。その結果, 上位 3 位のアンサンブル学習モデルにおいて, 判別性能が最大となった。今後は, さらに実験協力者を増やし結果の再現性を確認するとともに, PD の検出に有効な特徴量について調査し, PD の疾患検出率の向上を目指す。

謝辞

本研究は, 一部, 文部科学省科学研究費補助金 (JP19H01137, JP21K07428) の助成により行われた。

参考文献

- [1] Suppa A, Costantini G, Ascì F, Di Leo P, Al-Wardat MS, Di Lazzaro G, Scalise S, Pisani A, Saggio G: Voice in Parkinson's Disease: A Machine Learning Study, Front Neurol(2022).
- [2] 坂口巧一, 加藤昇平, 花井俊哉, 佐久間拓人, 大嶽れい子, 榎田道人, 渡辺宏久: 音読課題音声からの FTLD・ALS 簡易検出モデル, 情報処理学会第 83 回全国大会, Vol. 1(2021).
- [3] Kockmann, Marcel and Burget, Lukás and Cernocký, Jan: Brno University of Technology system for Interspeech 2009 emotion challenge, Interspeech, pp. 348-351(2009).