

発話音声を用いた軽度認知症スクリーニング - DNN を用いた判別モデルの検討 -

清水 絢子[†] 加藤 昇平^{†‡} 佐久間 拓人^{†‡}

[†]名古屋工業大学大学院工学研究科 [‡]名古屋工業大学情報科学フロンティア研究院

1 はじめに

現在、日本社会は著しく高齢化が進んでおり、早期に認知症を発見するための認知症スクリーニング技術の発展は強く期待されている。現在主流である医療機関での医師や臨床心理士によるスクリーニングは、簡易であっても5~20分程度の時間がかかる。そのため、医師などの負担が大きく、他の患者の診療に影響が出る可能性もある。そこで、負担を軽減する代替手段としてコンピュータによるデータマイニング技術を用いた認知症スクリーニングが研究されてきた。

大掛かりな装置や高価な機材を使用せず手軽に行える認知症スクリーニングとして音声を用いた認知症スクリーニングがある。本稿では、近年さまざまな分野で活発に研究されているディープラーニングの一種であるディープニューラルネットワーク (DNN) を用いたスクリーニングシステムを構築し、その有効性を検証した。

2 先行研究

加藤ら [1] は認知課題に対する高齢者の返答音声から音声韻律特徴を特徴量として抽出し、スクリーニングを行うシステムを提案している。ここで用いられている判別アルゴリズムは、音声韻律特徴の一つとして課題実行時の質問終了から返答までの反応時間 T を用いている。 T は「あー」「えー」などの無意味な音声を除き、課題に対する返答音声の開始時点を見つける必要があるため自動的な抽出が難しく、これまで人力で抽出されていた。認知症スクリーニングシステムの実用場面を考慮すると、特徴量は全て自動的に抽出できることが望ましい。しかしながら、認知課題によっては T を用いないことによる判別性能の低下が著しく、判別アルゴリズムの改善が求められてきた。

本稿では、人力で抽出されていた特徴量の代わりとして自動的に抽出可能な特徴量を用いて DNN による判別を行い、その性能評価および考察を行う。

3 スクリーニング手法

発話音声から特徴量を抽出し、特徴量選択を行い、二群判別を行う。発話音声には CDR [2] により評価された正解ラベルが付与されている。健常者 (CN) と CDR1 相当のアルツハイマー型軽度認知症患者 (AD) の二群、および健常者 (CN) と CDR0.5 相当の軽度認知機能障害 (MCI) の二群をそれぞれ判別する。

3.1 特徴量

本研究では先行研究で用いられている 129 種の特徴量のうち、時間に関する特徴量 T を発話までの反応時間 T^* に代えた 129 種の音声特徴を用いる。具体的にはスペクトルとピッチに関する特徴量を 13 種、フォルマントに関する特徴量を 33 種、エネルギーに関する特

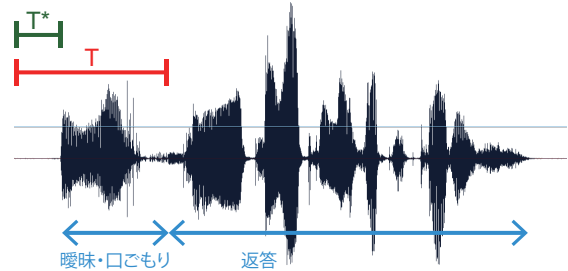


図 1: 時間に関する特徴量 T と T^*

徴量を 22 種、メルケプストラムに関する特徴量を 60 種、自動的に抽出可能な時間に関する特徴量 (T^*) を用いた。時間に関する特徴量 T と T^* の関係は図 1 のようになっている。

3.2 使用音声

日時見当識課題に回答した 243 名の高齢者 (年齢 65~96 歳, 男性 72 名, 女性 171 名) の音声を用いた。回答者は CN, AD, MCI のいずれも 81 名である。

4 DNN

DNN により CN/AD 判別と CN/MCI 判別を行う。

DNN についてパラメータ調整を行った。パラメータとして活性化関数、層数、ユニット数、繰り返し回数がある。これらのチューニングとしてグリッドサーチを行うには計算量が膨大で現実的ではないため、比較するパラメータはヒューリスティックを用いて決定した。用いた活性化関数は Ian J. Goodfellow らが発表した Maxout [3], 正規化線形関数とも呼ばれるランブ関数 (Rectifier), シグモイド関数の一種であるハイパボリックタンジェント (Tanh), これらの関数に過学習を抑制する効果がある Dropout [4] を併用したものの 6 種類である。

5 判別結果と考察

5.1 DNN のパラメータと判別結果の考察

DNN のパラメータの違いによる判別性能を検証した。結果を一部抜粋して表 1 に CN/AD 判別における結果、表 2 に CN/MCI 判別における結果を一部抜粋して示す。活性化関数によって正診率の高いパラメータが異なり、活性化関数ごとにパラメータを調整しなければならぬことがわかる。

表 1 より、DNN による CN/AD 判別において最も正診率が高かったのは活性化関数 TanhWithDropout, 層数 1, ユニット数 13, 繰り返し回数 100 のとき、活性化関数 TanhWithDropout, 層数 3, ユニット数 52, 繰り返し回数 500 のときおよび活性化関数 TanhWithDropout, 層数 3, ユニット数 130, 繰り返し回数 500 のときの 83.95%であった。表 2 より、CN/MCI 判別においては活性化関数 TanhWithDropout, 層数 4, ユニット数 13, 繰り返し回数 1000 のとき 75.31%であった。

CN/AD 判別, CN/MCI 判別ともに最も高い正診率を出した活性化関数は TanhWithDropout であった。このことからこのスクリーニング手法に DNN を用いる場合に適した活性化関数は TanhWithDropout である可能性が高いといえることができる。

*Screening of Mild Alzheimer's Disease and Mild Cognitive Impairment from Speech: binary discrimination using deep neural networks, Ayako SHIMIZU[†], and Shohei KATO^{†‡} Takuto SAKUMA^{†‡}

[†]Graduate School of Engineering, Nagoya Institute of Technology

Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

[‡]Frontier Research Institute for Information Science, Nagoya Institute of Technology

{shimizu, shohey, sakuma}@katolab.nitech.ac.jp

表 1: 各パラメータにおける CN/AD 判別性能 (一部抜粋)

活性化関数	層	ユニット	繰り返し	正診率 (%)
Maxout	3	13	500	79.01
MaxoutWithDropout	3	13	500	80.25
Rectifier	3	13	500	78.40
RectifierWithDropout	3	13	500	82.10
Tanh	3	13	500	82.10
Tanh	4	13	500	82.72
TanhWithDropout	1	13	100	83.95
TanhWithDropout	3	13	500	81.48
TanhWithDropout	3	52	500	83.95
TanhWithDropout	3	130	500	83.95

表 2: 各パラメータにおける CN/MCI 判別性能 (一部抜粋)

活性化関数	層	ユニット	繰り返し	正診率 (%)
Maxout	3	13	500	64.81
MaxoutWithDropout	3	13	500	67.90
Rectifier	3	13	500	62.96
RectifierWithDropout	3	13	500	67.90
Tanh	3	13	500	65.43
TanhWithDropout	3	13	500	73.46
TanhWithDropout	4	13	1000	75.31

また表 1, 表 2 より, 層数を 3, ユニット数を 13, 繰り返し回数を 500 にしたとき, Maxout, Rectifier, Tanh のいずれにおいても Dropout を併用した活性化関数を用いた場合のほうが正診率が高くなっていることがわかる. このことから CN/AD 判別および CN/MCI 判別においては Dropout により過学習を抑制できていると思われる.

5.2 判別結果の比較と考察

DNN の性能を検証するため, サポートベクターマシン (SVM), ランダムフォレスト, ロジスティック回帰分析による判別, AUC による判別 [1] を用いた比較実験を行った. 表 3 に CN/AD 判別および CN/MCI 判別における各手法の正診率を示す. また, 表 4 に返答までの反応時間 T を用いた先行研究の正診率を示す. SVM はカーネル関数を ANOVA RBF, コストを 5, 分類と回帰のタイプを C-svc, n 重交差確認法の n を 5 とした. ランダムフォレストはツリー数を 8000, CN/AD 判別の特徴量選択数を 6, CN/MCI 判別では 31 で判別を行った.

表 3 より, CN/AD 判別における正診率はロジスティック回帰分析による判別が一番高かった. 次に AUC による判別, DNN, ランダムフォレスト, SVM と続く. DNN とロジスティック回帰分析による判別の正診率の差は 3.70% であった.

CN/MCI 判別における正診率は AUC による判別の正診率が一番高かった. 次にロジスティック回帰分析による判別, DNN, ランダムフォレスト, SVM と続く. DNN と AUC による判別の正診率の差は 5.55% であった.

DNN についてはグリッドサーチなどによりチューニングをするには現実的でない計算量となるため現時点では十分にパラメータを調整できていないが, より適切なパラメータにすることで結果が向上する可能性もある. また, 現在用いている特徴量は線形判別手法に相応しいものであるため, より DNN で学習するに相応しい入力特徴量へ変更することにより, 大きく性能が向上する可能性が考えられる.

一方で表 4 より, 返答までの反応時間 T を用いた先行研究の正診率は CN/AD 判別で 89.51%, CN/MCI

表 3: 各手法による正診率 (%) (T^* 使用)

	DNN	SVM	ランダムフォレスト
CN/AD	83.95	79.01	80.25
CN/MCI	75.31	69.14	67.90
	ロジスティック	AUC	
CN/AD	87.65	85.80	
CN/MCI	79.63	80.86	

表 4: 先行研究による正診率 (%) (T 使用)

CN/AD	89.51
CN/MCI	75.93

判別で 75.93% であった. また, 特徴量 T に代えて発話までの反応時間 T^* を用いた判別手法の内最も高い CN/AD 判別の正診率はロジスティック回帰分析による判別の 87.65%, CN/MCI 判別では AUC による判別の 80.86% であった. CN/AD 判別において先行研究の正診率 89.51% のほうが 1.86% 高いものの, ともに 87% 以上と十分高い値であった. CN/MCI 判別においては先行研究の 75.93% よりも特徴量 T^* を用いた AUC による判別の正診率 80.86% は 4.93% も高かった. 自動的な抽出が容易である特徴量 T^* は特徴量 T の代替として十分に有用であると考えられる.

6 おわりに

本稿では, 音声による認知症スクリーニングシステムにおける DNN を用いた判別について検討した. 複数の判別アルゴリズムと, DNN の性能を比較した結果, CN/AD 判別, CN/MCI 判別ともに, 今回用いた 5 種類の判別アルゴリズムの中で 3 番目の正診率であった. 最も高かった正診率と比較すると DNN の正診率は CN/AD 判別において 3.70%, CN/MCI 判別においては, 5.55% 低い結果となった.

また, 返答までの反応時間 T に代えて発話までの反応時間 T^* を用いた場合, CN/AD 判別においてロジスティック回帰分析による判別が, CN/MCI 判別においては AUC による判別が最も高い正診率であることを確認した. 今研究で CN/MCI 判別において先行研究の正診率よりも優れた正診率を示した. CN/AD 判別においても先行研究には及ばないものの 87% 以上と十分に高い正診率を示した. これらより, 自動的に抽出可能な特徴量のみを用いた判別でも十分な判別結果を得られることが示唆された.

今後, DNN のパラメータとしてより適切なパラメータが存在すると思われるのでそれについても求めたい. また, 音声情報をそのまま入力情報とすることも考えている. 上記のような工夫を重ねて, ディープラーニングを用いた判別性能の向上を目指したい.

参考文献

- [1] Shohei Kato, Akira Homma, Takuto Sakuma, Munehiro Nakamura. Detection of Mild Alzheimer's Disease and Mild Cognitive Impairment from Elderly Speech: binary discrimination using logistic regression. The 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '15), Milano, pp.5569-5572, August 25-29, 2015.
- [2] Morris, J. C. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, Vol.43, No.11, pp.2412-2414, 1993.
- [3] Ian J. Goodfellow, David Warde-Farley, Mehdi Mirza, Aaron Courville, Yoshua Bengio. Maxout Networks. *ICML (3)*, 28, pp.1319-1327, 2013.
- [4] P. Baldi, P. Sadowski. Understanding Dropout. *NIPS*, pp.2814-2822, 2013.