

エージェントによる非定型質問への応答からの認知症検出

宇城 毅穣^{1,a)} 田中 宏季¹ 足立 浩祥² 数井 裕光³ 池田 学⁴ 工藤 喬² 中村 哲¹

概要: 簡易に認知症を検出するために、音声情報や言語情報に着目した研究が多く行われている。これらの研究で与えられる課題は、神経心理検査を元に作成されたものが多い。この場合、与えられる課題が限られてしまい、長期的に繰り返し利用する上で検出することが難しくなると考えられる。そこで我々は、各種検査の質問よりも非定型な質問を用意し、その質問への応答から認知症の検出を試みた。用意した質問は、エージェントにより研究協力者 24 名 (12 名の認知症患者と 12 名の非認知症者) に投げかけられる。我々は、高齢者の質問への応答を収録し、音声特徴量と言語特徴量を抽出した。抽出した特徴量を入力とした、機械学習による分類モデルを構築し、分類実験を行ったところ、ROC 曲線における AUC 値として 0.95 が得られた。これより、非定型質問を用いた場合においても高い認知症検出能力が得られることを示した。

キーワード: 認知症, 非定型質問, 音声特徴, 言語特徴, エージェントシステム

Detecting dementia from responses to atypically questions by agent systems

UJIRO TSUYOKI^{1,a)} TANAKA HIROKI¹ ADACHI HIROYOSHI² KAZUI HIROAKI³ IKEDA MANABU⁴
KUDO TAKASHI² NAKAMURA SATOSHI¹

Abstract: There are some studies to detect dementia utilizing acoustic and language information. Tasks used in the past studies were created on the basis of neuropsychological test. The limited variety of tasks might have a negative impact on being routinely screened for dementia, because of habituation effects. We proposed to detect dementia from the responses to more atypical questions. An agent system asked 24 participants (12 dementias and 12 non-dementia) to obtain answers of these questions. We classified two groups (dementia/non-dementia) by machine learning algorithm using the extracted features of speech and language from the recorded responses. The result showed 0.95 detection performance on the area under an ROC curve. We concluded that our system using atypical questions is also able to detect dementia.

Keywords: Dementia, atypical questions, speech feature, language feature, agent system

1. はじめに

厚生労働省の調査によると、日本の高齢化率(総人口に

占める 65 歳以上人口の割合)は、2017 年現在、27.5%となり、超高齢社会に突入している。今後も高齢化率は増加すると考えられており、高齢者の医療対策が重要な課題になっている。そのうちの一つに、認知症が挙げられる。認知症とは、記憶、思考、および行動などの認知機能が低下することで、日常生活に支障をきたす状態になることを指す。現在、認知症の有効な治療方法は、確立されておらず、患者およびその家族が将来の計画を立てるなどの診断後支援が必要となる。そのためには、認知症の早期発見が重要となる。

¹ 奈良先端科学技術大学院大学 情報科学研究科
Nara Institute of Science and Technology, Ikoma-shi, Nara, Japan

² 大阪大学 キャンパスライフ健康支援センター
Osaka University, Toyonaka-shi, Osaka, Japan

³ 高知大学 医学部 神経精神科学教室
Kochi University, Nankoku-shi, Kochi, Japan

⁴ 大阪大学大学院 医学系研究科
Osaka University, Suita-shi, Osaka, Japan

a) ujiro.tsuyoki.uq5@is.naist.jp

認知症の早期検出には、神経心理検査や血液検査、脳画像検査などを組み合わせて行われる [1]。しかし、この検査には、侵襲的な検査が含まれており、不安感やストレスを与えるため、被験者への負担が大きい。このため、非侵襲で手軽な検出法が必要となる。これまでに、非侵襲で手軽に認知症を検出する手法が、数多く提案されている。提案された手法に、言語情報を用いた研究や音声情報を用いた研究がある [2][3][4]。これらは、写真の叙述や神経心理検査中の一方的な発話を分析したものがほとんどである。また、エージェントを用いて、対話的に検出する手法も提案されている [5][6]。この手法では、神経心理検査を元で作成された質問を 3 問用意し、その応答の音声情報や言語情報から検出を試みている。しかし、神経心理検査の質問は、高齢者が覚え、あらかじめ対策を講じられる可能性がある。早期に認知症を検出するためには、定期的にかつ、長期的にモニタリングする必要がある。このため、あらかじめ対策ができないようにする必要がある。

本論文では、認知症を検出するための新たなアプローチを提案する。我々は、音声対話システムのエージェントを利用し、非定型な質問セットを用意した。その質問がエージェントから投げかけられ、その質問に対する応答から高い認知症検出能力を備えた手法を提案する。

2. 関連研究

多くの先行研究では、音声情報や言語情報から認知症の早期検出を目指し、その有効性が示されている。以下にそれらをまとめる。

Roark ら [4] は、ウェクスラーの記憶検査 [7] 中の音声を取録し、非認知症と軽度認知障害 (MCI) の方との分類を行い、神経心理検査検査よりも高い精度の結果となった。構文木の複雑さや音声特徴を利用して、分類を行い、認知症の早期検出に言語情報や音声情報が有用である可能性を示した。また、Aramaki ら [2] は、軽度認知障害 (MCI) の疑いがある方の認知能力と言語能力の関係を調べた。軽度認知障害の判定には、HDS-R [8] を用いた。言語能力の測定のために、作文とスピーチの課題を与えた。作文を解析した結果、非認知症と軽度認知障害の方との間に大きな違いは見られなかった。しかし、スピーチを解析した結果、非認知症よりも軽度認知障害の疑いがある人の方が、語彙数が多くなるという結果となった。これより、発話の分析によって、認知機能低下の初期段階を検出できる可能性を示した。コンピュータエージェントとの対話を通じて、認知症を検出する試みも行われている。Tanaka ら [5] は、MMSE [9]などを参考に質問を用意し、それらの質問への応答から認知症の検出を試みた。質問への応答を取録し、言語特徴や音声特徴、表情を抽出して、非認知症者と認知症患者の分類を行った。その結果、線形カーネルを用いた SVM において、ROC 曲線の AUC が 0.94 という検出能力

となった。Bahman ら [6] の研究も、精神科医の問診と同様の形式の質問を用意し、エージェントとの会話から認知症の検出を試みた。音声特徴や言語特徴を用いて、約 91 % の精度で分類することができた。これより、質問への応答から認知症を検出する方法を示した。

我々は音声対話のエージェントシステムを用いて、非定型質問に対する高齢者の応答から認知症を検出する手法を提案する。

3. データ取得

3.1 エージェントシステム

データ取得のために Tanaka ら [5] のエージェントシステムを利用した。このシステムは、エージェントとして、MMDAgent を使用しており、高齢者 (ユーザ) が理解しやすいように、エージェントの発話速度を下げ、発話内容に字幕を付与している。また、認知症を検出するために 6 種類の課題を用意されている (図 1)。以下にその概要を示す。

- 自己紹介: エージェントが自己紹介し、ユーザの名前と年齢を尋ねる。
- 視線計測: ディスプレイに小さな黒い点を表示し、移動させる。ユーザはその点を目で追うように指示する。
- 読み課題: 画面に文章を表示させ、ユーザに声を出して読むように指示する。文章はウェクスラーの記憶検査から引用している。
- 固定質問: MMSE などの神経心理検査の質問を用意し、ユーザに尋ねる。
- ランダム質問: 計 13 問の質問を用意し、ランダムで 5 問をユーザに尋ねる。
- リテリング: ウェクスラーの記憶検査中の一文をシステムが読み上げ、ユーザに記憶させる。読み上げた後、ユーザに記憶した分を復唱させる。

3.2 非定型質問

本研究では、エージェントシステムのランダム質問に着目した。ランダム質問では、エージェントの発話内容として、神経心理検査の質問よりも非定型な質問を 13 問用意した。質問は、精神科医師との相談の上で決定した。その質問を表 1 に示す。エージェントは、13 問の質問からランダムで 5 問を選び、高齢者に質問する。エージェントの会話中において、システムはユーザの音声と動画をそれぞれマイクと内蔵カメラで記録した。システムは、15 秒の無回答時間を検出すると、自動的に次の質問へ移行する。

4. 認知症の検出手法と評価

4.1 研究協力者

我々は、計 24 名の研究協力者を募集した。研究協力者の情報を表 2 に示す。各研究協力者は、事前に十分な説明を受けた上で本実験の参加に同意した。このうち、12 名は

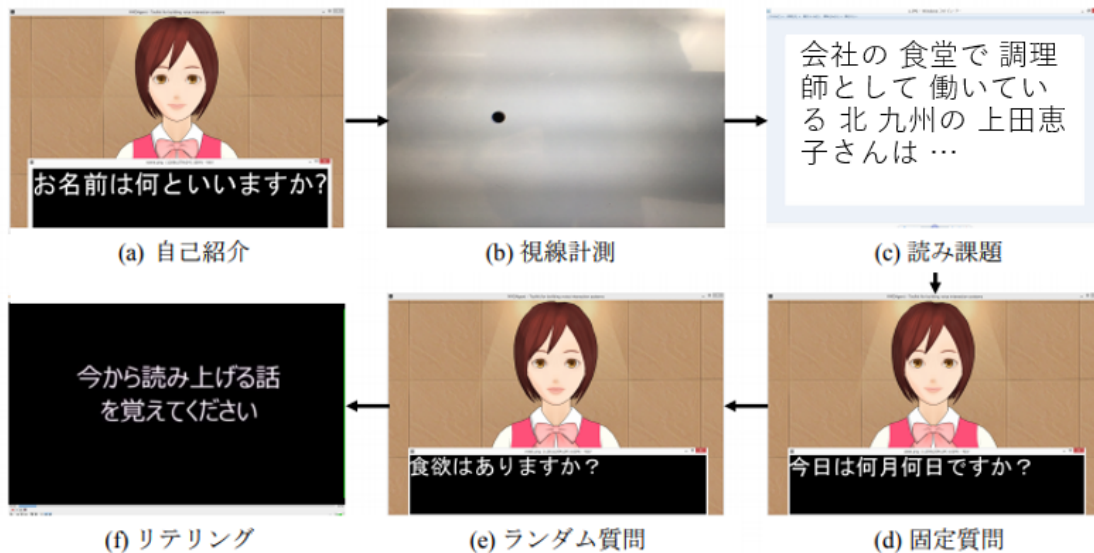


図 1 データ取得のためのエージェントシステム

表 1 用意した質問セット

質問	内容
Q1	ご家族のことを教えてください
Q2	日頃ストレスを感じていることを話してください
Q3	趣味は何ですか
Q4	好きな歌は何ですか
Q5	石原裕次郎さんについて知っていることを教えてください
Q6	長嶋茂雄さんについて知っていることを教えてください
Q7	美空ひばりさんについて知っていることを教えてください
Q8	今の日本の総理大臣は誰ですか
Q9	今の季節は何ですか
Q10	今年は何年ですか
Q11	あなたは左利きですか、右利きですか
Q12	夜は眠れていますか
Q13	食欲はありますか

表 2 研究協力者の年齢や MMSE スコア、教育歴の平均と標準偏差.

グループ	N	年齢		MMSE		教育歴	
		平均	SD	平均	SD	平均	SD
非認知症	12	74.5	4.3	27.5	1.8	8.8	2.6
認知症	12	75.9	7.6	21.2	5.1	13.9	3.8

大阪大学医学部附属病院における認知症グループとして収録に参加し、12名は奈良先端科学技術大学院大学における非認知症グループとして収録に参加した。認知症グループは既存の基準に従って、大阪大学医学部附属病院の医師により、DSM-IVに基づいて、早期認知症であると診断を受けている。我々は、全研究協力者から年齢と MMSE スコア、および、教育歴を取得した。音声対話エージェントを使用するために、Surface Pro 3 を使用し、全研究協力者が課題を完了することが確認できた。

4.2 応答からの特徴量セット

収録データに対し、発話書き起こしを行った。音声特徴

量と言語特徴量を 5 つの質問に対するそれぞれの応答について抽出し、全質問での平均値を算出した。我々は、先行研究 [5] に基づいて 21 種類の特徴量を選択した。

4.2.1 音声特徴量

音声特徴量の抽出に関して、我々は Snack sound toolkit*1 を使用した。ここでは、Gap や Pause、基本周波数、パワー、発話速度に関する特徴を抽出した。Gap は、エージェントの質問終了時から研究協力者の応答開始までの時間差のことである。Pause は、研究協力者の発話の中で時間差が 1 秒以上の回数をカウントしている。また、時間差が最長のものも算出した。基本周波数に関しては、変動係数、平均値、最大値、中央値、最小値、レンジを算出した。また、パワーに関しては、平均値、最大値、最小値を算出した。

4.2.2 言語特徴量

我々は、MeCab*2 を用いて日本語の形態素解析を行った。MeCab の出力より、トークン (形態素) 数、フィラー、名詞や動詞、形容詞、副詞の数を算出した。

4.3 識別モデル

我々は、抽出した特徴量から、2 種類の機械学習による識別モデルを構築した。識別モデルとして、ロジスティック回帰と線形サポートベクタマシンを使用した。識別モデルの入力として、それぞれの特徴量が最小値 0、最大値 1 になるよう正規化し、認知症グループと非認知症グループを分類するモデルを学習した。分類精度の評価に関しては、Roark ら [4] や Tanaka ら [5] の研究を基に、Leave-one-participant-out 交差検証から ROC 曲線を描き、ROC 曲線の AUC を算出した。

*1 <http://www.speech.kth.se/snack/>

*2 <http://taku910.github.io/mecab/>

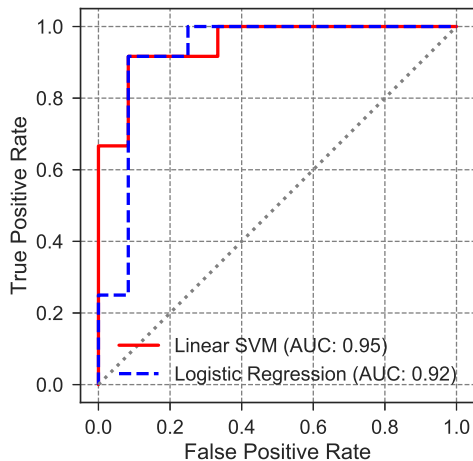


図 2 各モデルに対する ROC 曲線。赤色実線は、線形サポートベクタマシンの ROC 曲線。青色点線は、ロジスティック回帰の ROC 曲線を示している。

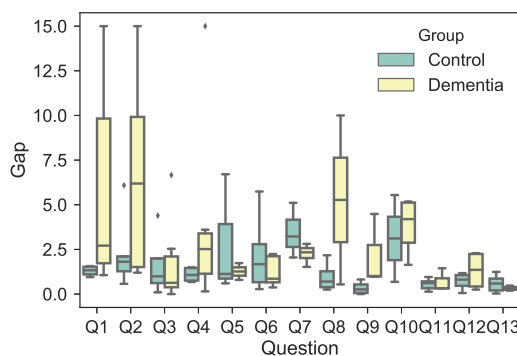


図 3 質問ごとの非認知症と認知症患者の Gap の箱ひげ図。

4.4 結果

構築した 2 種類のモデルの ROC 曲線を描いた結果を図 2 に示す。2 つの識別モデルから、それぞれ 0.95 (線形サポートベクタマシン), 0.92 (ロジスティック回帰) の AUC を得た。これらの結果は、先行研究 [5] と同様、高い検出能力であることがわかる。また、ロジスティック回帰において、特徴量の重みの大きさが最も大きかったのは Gap である。先行研究 [5] においても、最も効果があったと報告されている。ここで、研究協力者ごとに Gap の平均を算出した。マン・ホイットニーの U 検定で p 値を算出すると、非認知症者と認知症患者のグループ間に有意差がみられた ($p < 0.05$)。さらに、Gap を質問ごとに算出した結果を図 3 に示す。図 3 より、基本的には、非認知症者よりも認知症患者の方が Gap が大きく、差がみられる。特に、Q8 の総理大臣は誰かという問いに関しては、大きな差がみられることがわかる。しかし、質問の種類によっては、差がみられないものがあることがわかる。Q5, Q6, Q7 に関しては、過去の有名人について尋ねる質問であり、研究協力者が青年期の頃の記憶に関する質問に関しては、Gap の差がみられない可能性があると考えられる。

5. まとめ

本研究では、音声対話エージェントを用いて、非定型質問と高齢者の応答から認知症を検出する手法を提案した。計 24 名の認知症患者と非認知症者により、本手法の効果を実験的に検証し、結果として、ROC 曲線の AUC において 0.95 の検出精度を達成することができた。これより、非定型質問を用いた場合においても、認知症を検出することができる可能性が示された。今後は、質問による認知症患者と非認知症者の違いを分析する予定である。

謝辞 本研究の一部は JSPS 科研費 JP17H06101, JP16K16172 の助成を受けて行われた。

参考文献

- [1] Laske, C., Sohrabi, H. R., Frost, S. M., López-de Ipiña, K., Garrard, P., Buscema, M., Dauwels, J., Soekadar, S. R., Mueller, S., Linnemann, C. et al.: Innovative diagnostic tools for early detection of Alzheimer's disease, *Alzheimer's & dementia: the journal of the Alzheimer's Association*, Vol. 11, No. 5, pp. 561–578 (2015).
- [2] Aramaki, E., Shikata, S., Miyabe, M. and Kinoshita, A.: Vocabulary size in speech may be an early indicator of cognitive impairment, *PloS one*, Vol. 11, No. 5, p. e0155195 (2016).
- [3] McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J. and Mayeux, R.: The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, *Alzheimer's & dementia: the journal of the Alzheimer's Association*, Vol. 7, No. 3, pp. 263–269 (2011).
- [4] Roark, B., Mitchell, M., Hosom, J.-P., Hollingshead, K. and Kaye, J.: Spoken language derived measures for detecting mild cognitive impairment, *IEEE transactions on audio, speech, and language processing*, Vol. 19, No. 7, pp. 2081–2090 (2011).
- [5] Tanaka, H., Adachi, H., Ukita, N., Ikeda, M., Kazui, H., Kudo, T. and Nakamura, S.: Detecting Dementia Through Interactive Computer Avatars, *IEEE journal of translational engineering in health and medicine*, Vol. 5, pp. 1–11 (2017).
- [6] Mirheidari, B., Blackburn, D., Harkness, K., Walker, T., Venneri, A., Reuber, M. and Christensen, H.: An avatar-based system for identifying individuals likely to develop dementia, *Proc. Interspeech 2017*, pp. 3147–3151 (2017).
- [7] Wechsler, D.: *WAIS-III, Wechsler adult intelligence scale: Administration and scoring manual*, Psychological Corporation (1997).
- [8] Hosokawa, T., Yamada, Y., Isagoda, A. and Nakamura, R.: Psychometric equivalence of the Hasegawa Dementia Scale—revised with the Mini-Mental State Examination in stroke patients, *Perceptual and motor skills*, Vol. 79, No. 1, pp. 664–666 (1994).
- [9] Folstein, M. F., Folstein, S. E. and McHugh, P. R.: “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician, *Journal of psychiatric research*, Vol. 12, No. 3, pp. 189–198 (1975).