

平成29・30年度厚生労働科学研究費補助金
(政策科学総合研究事業(臨床研究等 ICT 基盤構築・人工知能実装研究事業))
分担研究報告書

表情、音声などのセンシングデータの人工知能解析

研究分担者 岸本泰士郎 慶應義塾大学医学部精神・神経科学教室 専任講師

研究要旨

うつ病と認知症はそれぞれへの治療アプローチが異なるため、鑑別が極めて重要である一方で、両者の鑑別に有効かつ簡便なバイオマーカーは存在しない。このため、両者の鑑別は、記憶障害の程度に対する機能障害の程度、自身の記憶障害への自覚の程度など臨床的な判断に頼らざるを得ないのが現状である。人工知能技術を利用して両者の鑑別が可能になれば有用であるし、また人工知能に判断を委ねるのではなくとも、どのような指標が両者の分別に(機械学習の特徴量として)有用なのかが明らかになることで、病態の理解につながり、臨床的に有用な示唆となる。

本分担研究課題においては、学習モデル構築および精度検証に用いるデータセットとして、日本医療研究開発機構(AMED)の委託研究「表情・音声・日常生活活動の定量化から精神症状の客観的評価をリアルタイムで届けるデバイスの開発(研究代表者:岸本泰士郎)」のデータを用いた。この研究では、診療場面で医師と患者が行う通常の会話の様子を録音・録画している。

初年度は、うつ病や認知症の症状が明確に存在するデータセットをそれぞれ51、56データセット抽出し、音響学的特徴を解析対象とした。音声をFormant frequency等の指標に変化した後、特徴エンジニアリングを行い、support vector machineによる分類器の生成を行った。結果、うつ病と認知症の分類精度は82.2%であった(一個抜き交差検証による)。次年度では、年齢を揃えたうつ病あるいは認知症患者それぞれ64名、74名分のデータセットを用い、音響学的特徴および表情特徴を解析対象とした。support vector machine with cubic kernelによる分類器の生成を行った結果、うつ病と認知症を分類する精度は89.9%であった(10-fold cross validationによる)。本検証結果は前述の研究に基づくデータセットを対象にしており、外部データでの検証を行う必要があるが、高い精度での分別が可能だったことから、今後、実用化に向けての開発が望まれる。

A. 研究目的

うつ病と認知症の鑑別は臨床上、しばしば困難である。両者の合併例も存在するが「仮面認知症」のように臨床表現型は認知症様であってもうつ病の治療に反応する一群の患者も存在する。うつ病と認知症では治療アプローチが異なるため、両者の鑑別は極めて重要である。一方で、両者の鑑別に有効かつ簡便なバイオマーカーは存在せず、画像検査においても特に初期の認知症の判断は困難である。その結果、両者の鑑別は、ライフイベントの関与の程度、記憶障害の程度に対する機能障害

の程度、自身の記憶障害への自覚の程度、さらには抗うつ薬への反応性など臨床的な判断に頼らざるを得ないのが現状である。人工知能技術を利用して両者の鑑別が可能になれば有用である。あるいは人工に判断を委ねなくとも、こういった指標が両者の分別に(機械学習の特徴量として)有用なのかが明らかになることで、病態の理解につながり、臨床的に有用な示唆となる。

B. 研究方法

学習および判別能力の精度検証に用いたデータ

セットとして、日本医療研究開発機構 (AMED) の委託研究で慶應義塾大学が代表機関を務める「表情・音声・日常生活活動の定量化から精神症状の客観的評価をリアルタイムで届けるデバイスの開発(研究代表者:岸本泰士郎)」のデータを用いた。この研究では、診療場面で医師と患者が行う通常の会話の様子を 10-30 分間、録音、録画し (いわば診療の様子を機械に観察させ)、音声や画像データを種々のソフトウェアを用いて数値化し、機械学習に用い、重症度分類等を試みている。現在までに、種々の重症度のうつ病・双極性障害患者 172 名、834 データセット、認知症 60 名、218 データセットが蓄積されている。

初年度の研究開発においては、前述の目的に向けての初めのステップとして、典型的なうつ病患者および認知症患者の分類を行う学習モデルを生成し、その精度を検証した (うつ病 51 データセット、認知症 56 データセットを利用)。初年度は診療観察データのうち、音声音響学的要素のみを対象にした。患者音声に解析には Praat を用いた。Praat は音声音響学的分析にしばしば用いられるソフトウェアで、話者の Formant、Pitch、Intensity、Pulse 等を解析することが可能である^{1,2)}。

また、2 年度の研究開発においては、以下のように年齢を揃えたデータセットを用意した。

- ・ 57-81 歳 (データセット数が確保でき、2 群に年齢差が生じない範囲として任意に設定)
 - ・ 10 分以上の自由会話の動画データがあるもの
 - ・ 認知症患者の場合：Mini Mental State Examination のスコアが 23 点以下、かつ、Geriatric Depression Scale (Long Form) において 10 点未満のもの
 - ・ うつ病患者の場合：Hamilton Rating Scale for Depression 17 項目版のうち 8 点以上のもの
- 結果、上述を満たした、うつ病、認知症それぞれ 64 名、74 名分のデータを抽出し、研究用のデータセットとした。また、2 年度の研究開発においては、初年度同様、Praat を用いた音響学的な特徴量を利用しただけでなく、表情データから得られる

特徴量を利用した。表情の解析に際しては、オムロン社の OKAO Vision を用いた。OKAO Vision は顔検出・顔器官検出を自動で行い、表情 (喜び・驚き・怒り・悲しみ・無表情)、瞬目などを定量することが可能である^{3,4)}。これらのうち、当研究グループによる種々の施行から、うつ病重症度等の解析に有用であった要素を利用し、特徴エンジニアリングを行い、次に Lasso (least absolute shrinkage and selection operator) を用いて特徴選択を行った。うつ病あるいは認知症と標識したデータで support vector machine with cubic kernel による分類器の生成を行った。

(倫理面への配慮)

本研究の施行に際しては、人を対象とする医学系研究の倫理指針に準拠した倫理委員会に諮り、承認を得て行った。また、全ての被験者から同意を取得して行った。

C. 研究成果

初年度の音声のみを用いたうつ病・認知症の判別 (認知症を同定する) の精度、感度、特異度は一個抜き交差検証 (Leave one out cross validation) で以下の様であった。

	精度	感度	特異度
音声データのみ	0.822	0.824	0.821

また、2 年度に取り組んだ音声および表情データを用いた精度、感度、特異度を以下に示す (10-fold cross validation による) 音声、表情それぞれ単独のモダリティを用いた検証は以下の様であった。

	精度	感度	特異度
音声データのみ	0.884	0.919	0.844
表情データのみ	0.826	0.813	0.838

音声・表情データ両者を用いた分類器の精度、感度、特異度を以下の様であった

	精度	感度	特異度
音声・表情データの組み合わせ	0.899	0.919	0.875

D. 考察

診療場面から得られる音声および表情データを用いたうつ病と認知症の分類器の精度は 89.9%に達した。なお、本検証結果は前述の研究に基づくデータセットのみを対象にしており、外部データによる検証を行う必要がある。

今後、さらなるデータ収集や学習モデルの改善を通じて、分類の精度をあげることでより臨床上の有用性は増すものと考えられる。また、現在は 10 分程度の撮影データを元に解析を行っているが、より短時間でも可能かどうかの検証も有用と考える。将来は治療反応を予測するような臨床的により有用な予測技術の開発が求められる。

E. 結論

診療場面で医師と患者が行う通常の会話の様子を録音・録画したデータを用いて 90%程度の精度でうつ病と認知症の分類が可能であった。さらなるデータ収集や学習モデルの改善を通じて、分類の精度をあげることでより臨床上の有用性は増すものと考えられる。また、現在は 10 分程度の撮影データを元に解析を行っているが、より短時間でも分別が可能かどうかの検証も有用と考える。さらに、本研究では、うつ病で認知機能の低下を認めない症例と、認知症でうつ症状を認めない症例とで構成されるデータセットを用いて分類器の生成を行ったが、今後は、鑑別が実際に困難だったケースなどでの検討が必要である。

上述のような課題はあるものの、90%程度での分別能力は実臨床での使用を考えた際にも有用であり、今後は実用化に向けての開発が望まれる。

F. 健康危険情報

なし

G. 研究発表

1. 学会発表・招致講演等

岸本泰士郎. 情報通信技術や人工知能を用いた精神科領域における新たな試み,第38回多摩精神科

臨床研究会, 東京, 2017.11.8

岸本泰士郎. ICTや機械学習を用いた精神科領域における新しい試み,慶應義塾大学ウェルビーイング研究センターリサーチコンプレックス推進事業設置講座, 神奈川県, 2017.11.30

岸本泰士郎. 情報通信技術によって開かれる精神科診療の展望と課題. 第55回兵庫県精神医療学術講演会, 神戸, 2018.1.20

岸本泰士郎. ICTやAIを活用した今後の精神科医療の展望& 3つの試験デザインに基づくLAIの有効性の検討. 第129回福井県神経科精神科医会, 福井, 2018.2.2

岸本泰士郎. AI を用いた精神疾患臨床症状定量化の試み, 平成 29 年度 医薬品評価委員会 臨床評価部会総会, 東京, 2018.2.21

岸本泰士郎. 精神疾患・認知症診断におけるテクノロジー利活用の課題と展望, 東海大学認知症研修会プログラム, 神奈川, 2018.2.27

岸本泰士郎.情報通信技術 (ICT) や人工知能 (AI) を活用した診断支援の試み. 第30回東京都医師会医療とITシンポジウム, 2018.3.10

岸本泰士郎. AI はこころの病を診ることができるか? -人工知能を用いたうつ症状定量化の試み-. ぐんま・脳とこころのアカデミー, 群馬, 2018.03.24

岸本 泰士郎. 機械学習による精神運動制止の評価の試み, 第114回日本精神神経学会学術総会, 神戸, 2018.6.21

岸本 泰士郎. 情報通信技術や機械学習を活用した臨床症状評価, 第114回日本精神神経学会学術

総会, 神戸, 2018.6.21

岸本 泰士郎. 情報通信技術 (ICT) や人工知能 (AI) の活用という観点で30年後の精神科医療を展望する, 第114回日本精神神経学会学術総会, 神戸, 2018.6.23

岸本 泰士郎. 情報通信技術や人工知能がもたらす近未来の精神医学の展望, 第114回日本精神神経学会学術総会, 神戸, 2018.6.23

澤田 恭助, 高宮 彰紘, 岸本 泰士郎, 三村 将. うつ病患者に対する音声データの臨床的有用性の検討: システムティックレビューとメタ解析, 第114回日本精神神経学会学術総会, 神戸, 2018.6.22

岸本 泰士郎. 遠隔医療は国民に寄り添った医療になるか? 本邦における遠隔精神科医療の展望と課題, 日本精神神経科診療所協会 第24回学術研究会, 兵庫, 2018.6.24

岸本 泰士郎. 情報通信技術 (ICT) や機械学習を用いたうつ病診療の展望, 第3回CNSサミット, 東京, 2018.8.5

岸本泰士郎. 情報通信技術や機械学習を用いた精神症状定量化の試み, 日本線維筋痛症学会第10回学術集会, 東京, 2018.9.29

岸本 泰士郎. 遠隔医療や機械学習を活用した認知症診療の展望, 第37回日本認知症学会学術集会, 札幌, 2018.10.13

工藤弘毅, 岸本泰士郎. 「メンタルヘルスリテラシーと人工知能 (AI)」, 第38回日本精神科診断学会, 埼玉, 2018.10.19

岸本泰士郎. 情報通信技術や人工知能を活用した

近未来の精神科医療の展望, 第3回 Venlafaxine研究会, 神奈川, 2018.10.24

岸本泰士郎. 情報通信技術や人工知能を活用した近未来の医療の展望, 平成30年度女性医師の勤務環境に関する病院長等との懇談会, 沖縄, 2018.10.30

岸本泰士郎. 精神科医療の遠隔化は診療所、精神病院、総合・大学病院に普及し得るか, 第38回医療情報学連合大会 (第19回日本医療情報学会学術大会), 福岡, 2018.11.23

岸本泰士郎. 情報通信技術や機械学習を活用した精神科領域の展望, 第2回デジタルヘルス学会, 東京, 2018.12.23

岸本泰士郎. 人工知能技術を用いた精神疾患症状定量化の試み, 第1回日本メディカルAI学会, 東京, 2019.1.26

H. 知的財産権の出願・登録状況
出願準備中

参考文献

1. Boersma P and Weenink D. Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from <http://www.praat.org/>
2. Boersma P and Weenink D. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10): 341-345. 200114; 33(2): 219-37.
3. Konishi Y, Kinoshita K, Kawade M. Sparse bayesian regression for head pose estimation. 18th International Conference on Pattern Recognition (ICPR'06). Vol 0, Hong Kong, China, 20-24, August 2006:507-510.
4. Kitazawa M, Yoshimura M, Kuo-Ching

L, Wada S, Mimura M, Tsubota K, Kishimoto T.
Utilization of Facial Image Analysis Technology
for Blink Detection: A Validation Study. Eye
Contact Lens. 2018. Nov;44 Suppl 2:S297-S301.