

# 表情情報を用いた「せん妄」患者の識別モデル生成と特徴分析

生田 光輝（東京大学／産総研），○松本 吉央（東京理科大／産総研），  
小川 朝生，大谷 清子（国立がん研究センター）

## Development of Classification Model and Feature Analysis Using Facial Feature Information for "Delirium" Diagnosis Support

Mitsuki Ikuta (The University of Tokyo / AIST), ○Yoshio Matsumoto (Tokyo University of Science / AIST), Asao Ogawa, Kiyoko Otani (National Cancer Center)

Abstract: Delirium is a temporary and reversible mental disorder related to attention and cognition. Its onset is common in healthcare settings, and despite the importance of early detection and response, many are overlooked. Therefore, this study aims to create a classification model and extract features using video images of dialogue during delirium. In addition, an image data collection system using a neck-mounted camera device that is easy to use in clinical practice is investigated.

Key words: Delirium, Psychiatry, Machine learning, Explainable AI

### 1. はじめに

せん妄は、注意力障害を中核症状とし、種々の精神症状を伴う中枢神経系の機能障害の一形態である。入院患者の約 20%に出現するとされており、患者の予後や QOL にも影響することから、発症・重症化の予防が重要である[1,2]。

せん妄のスクリーニングに関する研究としては、電子カルテの情報からの発症予測やリスクの高い入院患者の特定を行うものが盛んである。RandomForest を用いた研究としては、術前の電子カルテを使用して術後のせん妄を予測し 80%の精度でせん妄を予測したもの[3]、また患者の入院中の電子カルテを学習し入院時のデータからせん妄のリスクが高い入院患者を 74.1% の感度と 82.2% の特異性で予測したもの[4]がある。また、約 5 万人の入院に関するデータを利用し、多重ロジスティック回帰を用いて、AUC 0.83 のスコアでせん妄を予測した研究もある[5]。しかし、医療機関においてせん妄の診断や管理を行う専門医はまだ少なく、実際の医療現場での診断や重症度の評価は難しい。

患者の顔画像から疾患をスクリーニングする研究として、顔画像から認知症患者を検出するもの[6]や、動画像からパーキンソン病を検出する研究[7]などがある。我々はこれまでに、せん妄患者を顔画像（表情特徴）をもとにスクリーニングする手法を提案し、予備的検討を行った[8]。せん妄患者は、表情筋の変化が遅いことや反応潜時が延長することなどが知られており、動画像を用いて表情からせん妄の診断モデルの作成とその特徴の可視化ができる可能性がある。本稿では、顔画像を用いた識別モデル作成とその特徴分析の結果について述べる。

### 2. 表情情報を用いた「せん妄」患者の識別手法

#### 2.1 特徴量の抽出

顔画像から表情情報を抽出するために、本研究ではオープンソースの画像処理ソフトウェアである OpenFace[9]を用いる。OpenFace は、1 枚の顔画像からランドマークと呼ばれる目、鼻、口、眉などの特徴点（合計 68 箇所）の 2 次元座標、FACS（Facial Action Coding System）における表情運動要素である Action Unit (AU) [10]の動き、視線の方向、頭部の姿勢等を検出することが可能である。FACS とは、人の表情を解剖学的に分解した顔筋肉の動きの特徴の定義で、主に心理学や表情・感情認識で使われている。OpenFace を動画像に対して用いると、これらの情報をリアルタイムに取得することができる。

#### 2.2 学習モデル

OpenFace から取得した AU と視線の動きの時系列データを学習し、せん妄の分類モデルを作成する。本研究では、機械学習アルゴリズムである XGBoost(eXtreme Gradient Boosting) [11] を用いた。このアルゴリズムは、勾配ブースティング決定木（GBDT:Gradient Boosting Regression Tree）の一種で、決定木を大量に生成し、各決定木の結果を集計して予測する手法である。この手法には、非線形性や相互作用を考慮しやすい、数値を扱うテーブルデータの分析に優れていて精度が高い、可読性に優れており特徴量について分析しやすいなどの特徴があることから、本研究に適していると判断した。

## 2.3 特徴分析

学習モデルに対して、SHAP (SHapley Additive exPlanations) [12]を用いてせん妄に関連のある特徴量を解析する。この手法は、機械学習モデルの解釈手法の一つで、Fig. 1 のように各特徴量が予測値にどのように影響を与えたのかを可視化できる。具体的には、分類モデルが正解ラベルに対して、入力データの各変数がプラスに働いたのか、あるいはマイナスに働いたのかを明らかにできる。そのため、予測時に用いることでどのような特徴からその予測結果に至ったのかを説明でき、サンプル全体で各特徴量の SHAP 値の分布や平均を見ることが予測値と各特徴量の関係性が分かる。この特性を利用して、各特徴量がどのような大きさだとせん妄と判断できるのかについて分析する。

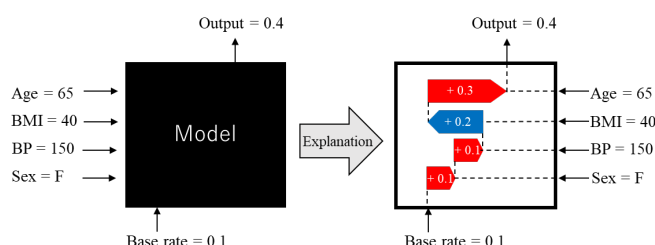


Fig.1: SHAP (SHapley Additive exPlanations)

## 2.4 データ収集とデータセットの作成

本研究では、国立がん研究センターの倫理委員会の承認のもと、同東病院にてせん妄患者の顔動画を収集した。具体的には入院中のせん妄患者 6 人、非せん妄患者 15 人を対象に、病室での診察場面の医師との対話の様子を撮影した (Fig.2)。撮影には、WEB カメラを接続したタブレット (Microsoft Surface 3) を利用した。各患者の動画の長さは 1 分から 5 分程度であった。

本実験では、収集している表情データが診察場面のものであるため、すべての患者に共通する表情変化として、医師の問いかけに対する患者の応答 (頷き) や視線の動きが含まれる。そこで、患者の応答時の 3 秒間の AU と視線データを抽出し、その分散値を特徴量とした。なお、ここでは OpenFace によって分析可能な 17 の AU (Table 1) をすべて利用した。最終的に、データの大きさは 559 個となったため、一人当たりの診断動画から平均約 27 個のデータを作成したことになる。

また、せん妄有無の正解データとして、電子カルテのデータを利用した。ここには、せん妄の有無 (専門医による診断結果) や重症度、処方薬、他の病状 (認知症等) についての情報が含まれており、動画撮影時のせん妄の発症状況を用いてラベル付けを行った。

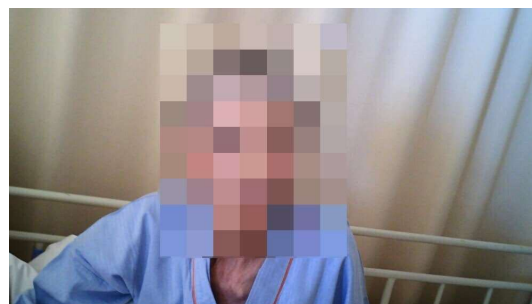


Fig.2: Example of Recorded Patient Image

Table 1: Action Units used for Analysis

AU	Description	12	Lip Corner Puller
1	Inner Brow Raiser	14	Dimpler
2	Outer Brow Raiser	15	Lip Corner Depressor
4	Brow Lowerer	17	Chin Raiser
5	Upper Lid Raiser	20	Lip stretcher
6	Cheek Raiser	23	Lip Tightener
7	Lid Tightener	25	Lips part
9	Nose Wrinkler	26	Jaw Drop
10	Upper Lip Raiser	45	Blink

## 3. 分析結果と考察

### 3.1 分類モデルの分類精度

作成したデータセットをもとに XGBoost にて分類モデルを作成した結果、5 分割の交差検証 (クロスバリデーション) による精度は 91% となった。この値は比較的高く、表情や視線の情報からせん妄をスクリーニングできる可能性が十分にあることが示唆される。

### 3.2 分類に重要な特徴

Fig. 3 は、SHAP を用いてせん妄患者と非患者を分類する上で重要な特徴を可視化したものである。青で表されている Class 0 はせん妄でない人への予測貢献度であり、赤で表されている Class 1 はせん妄を持った患者への予測貢献度である。今回の分析では、AU05 (上眼瞼挙筋)、AU02 (前頭筋、外側筋)、AU07 (眼輪筋、眼瞼下垂部)、AU23 (口輪筋)、Gaze angle (視線) が大きく寄与する特徴となった。AU05 は上瞼を持ち上げる動作で、AU02 は眉の外側を持ち上げる動作、AU07 は瞼の緊張で、AU23 は唇を閉じる動作である。結果として、せん妄であるかを識別するのに寄与した表情特徴は、瞼や眉、視線など、目元に集まる形となった。

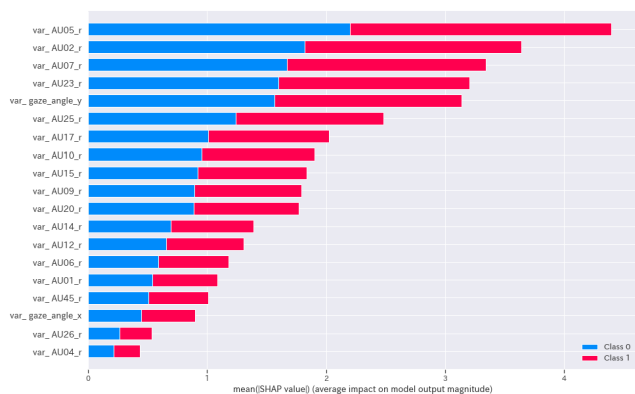


Fig.3: Features Contributing to Classification of Delirium

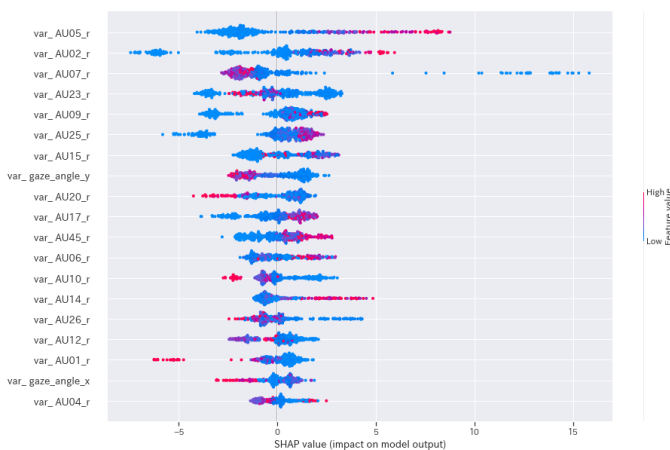


Fig.4: Visualization of the feature contribution (Patients with delirium)

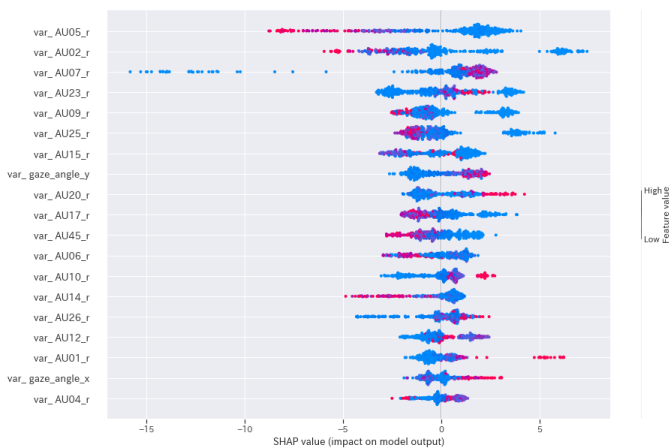


Fig.5: Visualization of the feature contribution (Non-patients)

Fig.4 と Fig.5 はせん妄患者と非患者のそれぞれの分類に寄与した重要な特徴の分布を示したものである。

Fig. 4 から、せん妄患者は AU05 と AU02 からの上瞼と眉の動きが大きくなることが分かった。一方で、AU07 と AU23 から、瞼の緊張と唇を閉じる動作が小さくなることが分かった。また、全体的な傾向として、口元

の動きが小さくなり、瞼付近は動きが大きくなることがあり、視線の動きも小さくなることが分かった。この特徴から、従来言われていた表情筋の弛緩は、主に口元に発生している可能性が考えられる。瞼の動きが大きくなる点についてはあまり予想はしていなかったが、眼瞼痙攣などの痙攣は、せん妄と同様に中枢神経系の異常によって発生するため、せん妄と同様に神経伝達異常が原因で、症状が表層に現れてきている可能性もある。また、認知症などで、表情の動きが減衰した場合でも、瞼の反射などの動きは最後まで減衰しないといわれている[13]。そのため、そのような反射活動が意識障害中でも残留している可能性もある。この分析結果を踏まえて、実際にせん妄患者の動画を確認したところ、視線があまり合わず、ぼんやりとしているような雰囲気を感じたため、視線や目元以外の表情の動きの減衰が、特徴として表れていると考えられる。

一方、非患者については、Fig. 5 から視線の動きや口元の動きが大きいことがわかる。非患者の動画を確認したところ、場合によっては笑顔を浮かべたり、診察の場面で流暢に会話が進んでいた人が多かった。非患者の電子カルテを参照すると、過去に発症したせん妄の治療が完了して、退院調整をしていたりと比較的元気な人が多かった。したがって、その点が表情の動き方の違いとして現れた可能性も考えられる。また、分析結果を現場でせん妄の対応にあたる医師等に共有したところ、せん妄患者の特徴として共感を得られた。特に、視線については違和感を感じていたようで、せん妄患者とは視線が合いづらく、患者の視界に入りこむように診断を心掛けることがあることが分かった。そのため、このような表情特徴や視線情報から、せん妄を評価したり、診断支援に役立てられる可能性があると考えられる。

## 5. おわりに

本研究では、せん妄患者と非患者の対話中の動画像を用いて、表情や視線の情報から分類モデルを作成し、91%の分類精度が得られた。さらにせん妄患者の特徴を分析したところ、せん妄患者には表情や視線に特徴があることが示唆された。今後、これらの情報を用いた診断支援システムへの応用が期待される。ただし、現在の分析ではまだ利用できたデータの患者数が少ないため、より多くのデータを用いた分析が必要である。

## 参考文献

- [1] Esther S Oh, Tamara G Fong, Tammy T Hsieh, Sharon K Inouye. "Delirium in Older Persons : Advances in Diagnosis and Treatment" JAMA,

- pp.1161-1174 (2017)
- [2] Tammy T Hsieh , Sharon K Inouye, Esther S Oh  
“Delirium in the Elderly” *Psychiatr Clin North Am*, pp.1-17 (2018)
  - [3] Anis Davoudi, Ashkan Ebadi, Parisa Rashidi, Tazcan Ozrazgat-Baslanti, Azra Bihorac, and Alberto C. Bursian. “Delirium Prediction using Machine Learning Models on Preoperative Electronic Health Records Data” *IEEE Int Symp Bioinformatics Bioeng* , pp.568–573 (2018).
  - [4] Stefanie Jauk, Diether Kramer, Birgit Großauer , Susanne Rienmüller , Alexander Avian , Andrea Berghold, Werner Leodolter , Stefan Schulz. “Risk prediction of delirium in hospitalized patients using machine learning: An implementation and prospective evaluation study” *Journal of the American Medical Informatics Association* Volume 27, pp.1383–1392 (2020).
  - [5] Caitlin E. Coombes, Kevin R. Coombes ,Naleef Fareed ”A novel model to label delirium in an intensive care unit from clinician actions” *BMC Medical Informatics and Decision Making*, Vol.21, Article number: 97(2021).
  - [6] Yumi Umeda-Kameyama, Masashi Kameyama , Tomoki Tanaka , Bo-Kyung Son, Taro Kojima, Makoto Fukasawa , Tomomichi Iizuka, Sumito Ogawa, Katsuya Iijima , Masahiro Akishita “Screening of Alzheimer’s disease by facial complexion using artificial intelligence” *Aging* (Albany, NY) Volume 13, pp.1765—1772 (2021).
  - [7] Mohammad Rafayet Ali, Taylor Myers, Ellen Wagner, Harshil Ratnu, E. Ray Dorsey & Ehsan Hoque. “Facial expressions can detect Parkinson’s disease: preliminary evidence from videos collected online” *npj Digital Medicine* volume 4, (2021).
  - [8] 生田光輝,松本吉央,小川朝生, 大谷清子, “画像を用いた「せん妄」診断支援に向けたモデル作成・特徴分析およびデータ収集システムの検討”, 第 23 回計測自動制御学会システムインテグレーション部門講演会, 2A2-E10 (2022).
  - [9] Baltrusaitis, T., Zadeh, A., Lim, Y. C. & Morency, L.-P. “Openface 2.0: facial behavior analysis toolkit.” *IEEE International Conference on Automatic Face & Gesture Recognition*, pp.59–66 (2018).
  - [10] Ekman, P., Friesen, W. V. & Hager, J. C. “Facial action coding system (FACS) : A Technique for the Measurement of Facial Action.” *Consulting Psychologists Press*, Palo Alto, CA (1978).
  - [11] Tianqi Chen and Carlos Guestrin. “XGBoost: A Scalable Tree Boosting System.” *SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.785–794 (2016).
  - [12] Scott M. Lundberg, Su-In Lee. “A Unified Approach to Interpreting Model Predictions.” *NIPS*, pp.4768–4777 (2017).
  - [13] Jonsson C-O, Mälhammar G, Waldton S. “Reflex elicitation thresholds in senile dementia.” *Acta Psychiatr Scand*, 55(2):81-96 (1977).