

# 覚醒低下検出モデルにおける Variational Autoencoder の潜在空間の次元数最適化

松井 恵莉香† 高野 聖人† 南雲 健人†  
大岩 孝輔† 野澤 昭雄† (青山学院大学)

## 1. はじめに

近年、運転中のドライバーの覚醒度維持を目的とした眠気検知技術の開発が進んでいる。これまでの眠気検知技術は時系列の変化を捉えて眠気を検出する方法が提案されてきた<sup>(1)</sup>。しかし、これらの時間的な情報を用いた手法では眠気検出に時間を要することが課題である。

そこで、我々は時間的な情報を用いず顔面皮膚温の情報である顔面熱画像の空間特徴量のみを利用することで、短時間で眠気検出をできると考えた。顔面熱画像は非接触・非侵襲で計測可能であり、ドライバーの精神的・物理的負担を軽減することができる。これまでに、CNNを用いた覚醒度段階推定を行ってきた<sup>(2)</sup>。CNNを用いた手法は、モデル構築時に多くの覚醒状態と覚醒低下状態の顔面熱画像が必要であるが、覚醒低下状態の顔面熱画像は十分な量を収集することが困難である。

この問題を解決する方法の1つとして、Variational Autoencoder (VAE) を用いた異常検知手法の検討が行われている<sup>(3)</sup>。VAEは「正常」なデータのみで異常検知モデルの構築が可能である。したがって、VAEを用いた手法は「正常」なデータに相当する覚醒状態の顔面熱画像のみで、「異常」なデータに相当する覚醒低下状態の検出が期待できる。一方で、VAEは同一の学習条件下でも異常検知精度がばらついてしまうという問題がある。先行研究<sup>(4)</sup>では、顔面熱画像を用いたVAEの異常検知精度向上の試みとして、中核温度変動の検出におけるVAEのパラメータ最適化が行われた。その結果、VAEのパラメータ最適化により異常検知精度のばらつきが抑えられた。

本研究では、VAEを用いた顔面熱画像に基づく一過性覚醒低下検出を目指すとともに、覚醒低下検出におけるVAEのパラメータの1つである潜在空間の次元数最適化を検討し、一過性覚醒低下の検出精度を評価した。

## 2. VAE の基本構造

教師なし学習の1つであるVAEは、生成モデルの代表的な方法であり、入力データと出力データが一致するように学習する。入力データを  $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$ 、出力データを  $\hat{\mathbf{x}} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N\}$  とする。VAEは $\mathbf{x}$ が与えられたとき、潜在変数を確率分布に仮定することで、 $\hat{\mathbf{x}}$ を生成する。

図1にVAEの学習プロセスを示す。VAEは入力空間を低次の空間に射影するEncoderと、低次の空間から元の空間に戻すDecoderから構成される<sup>(5)</sup>。Encoderに学習データ $\mathbf{x}$ が入力されると、潜在変数 $\mathbf{z}$ の変分事後分布 $q_\theta(\mathbf{z}|\mathbf{x})$ の

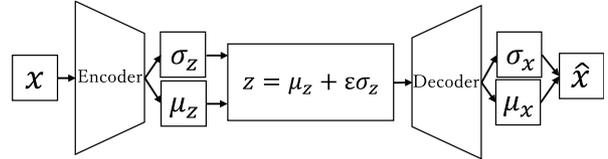


図1 VAEの学習プロセス

パラメータである平均ベクトル $\mu_z$ と標準偏差ベクトル $\sigma_z$ を出力する。そして、これらをパラメータとした正規分布から潜在変数 $\mathbf{z}$ が確率的にサンプリングされる。なお、潜在変数 $\mathbf{z}$ は、平均0、分散1あるいは共分散行列 $\mathbf{I}$ の正規分布からサンプリングされる $\epsilon$ を用いて表現される<sup>(6)</sup>。潜在変数 $\mathbf{z}$ をDecoderに入力すると、 $\mathbf{x}$ の条件付き確率 $p_\theta(\mathbf{x}|\mathbf{z})$ のパラメータである平均ベクトル $\mu_x$ と標準偏差ベクトル $\sigma_x$ が出力され、 $\mathbf{x}$ を復元する。この再構成された $\hat{\mathbf{x}}$ 、つまりDecoderの出力 $\hat{\mathbf{x}}$ が最終的なVAEの出力となる。

## 3. 実験

本研究では、覚醒低下状態を喚起すべく、意図的に覚醒低下の誘発及び抑制を繰り返す実験を行った。

### 3.1 実験環境

実験は室温 $25.4 \pm 0.8^\circ\text{C}$ の対流のないシールドルームで行い、覚醒低下の誘発効果を高めるため、暗室で実施した。被験者は安静状態で着座した。被験者の前方1.0mに赤外線サーモグラフィ(A615, FLIR)を設置し、撮影速度1fpsで熱画像を記録した。なお、本装置の熱画像の解像度は $640 \times 480$ pixel、温度分解能は $0.1^\circ\text{C}$ 以下、皮膚の熱放射率は $\epsilon = 0.98$ とした。また、顔表情による覚醒度を評価するため、顔表情計測用の赤外線カメラを被験者の前方0.6mに設置した。

### 3.2 実験手順

被験者は健康成人の男性5名(年齢:20~21歳)とした。被験者には予め実験内容・目的・調査対象を口頭および書面にて十分説明し、署名により実験協力に対する同意を確認した。室温への順化を図るため、実験は被験者入室後から30分以上経過後に実施した。実験は、15分間の眠気誘発区間、5分間の覚醒低下の抑制区間、30秒間の休憩区間、各区間前後3分間の安静状態区間により構成される。以上の一連の流れを1セットとして、3セット連続で行った。安静状態区間は1分ごとに開眼、閉眼、開眼を繰り返した。眠気誘発区間は、被験者に目の前に設置された液晶モニターに表示された眠気誘発移動表的を注視することを指示した。眠気誘発移動表的は、円軌道を描くように円形の移動目標が1周期あたり2秒として移動する。覚醒低下の抑制区間は、約20秒間のアラーム音再生後、繰り上がりありの2桁の足し算を3秒間隔

で課した。暗算課題は液晶モニタに表示され、被験者はテンキーを使って回答した。休憩区間に被験者は、疲労を軽減するためにストレッチを行った。

## 4. 解析方法

### 4.1 覚醒度の評価

北島らが提案した下記に示す 5 段階の眠気評価法<sup>(6)</sup>を用いて被験者以外の 5 名の評価者が評価した。評価者は被験者の顔表情動画から 20 秒おきに覚醒度を評価した。各評価者が評定した値の平均を四捨五入し、眠気レベルを分類した。

- (1) 眠気レベル 1: 全く眠くなさそう (視線の移動が早く頻繁である, 瞬きの周期は安定している, 動きが活発で身体動きを伴う)
- (2) 眠気レベル 2: やや眠そう (唇が開いている, 視線移動の動きが遅い)
- (3) 眠気レベル 3: 眠そう (瞬きはゆっくりと頻発, 口の動きがある, 座り直しがあがる, 顔に手をやる)
- (4) 眠気レベル 4: かなり眠そう (意図的と思われる瞬きがある, 頭振る, 肩の上下動など無用な体全体の動きがある, あくびは頻発し深呼吸も見られる, 瞬きも視線の動きも遅い)
- (5) 眠気レベル 5: 非常に眠そう (瞼を閉じる, 頭が前後に傾く)

### 4.2 VAE による覚醒低下検出モデルの構築

実験で得られた熱画像を顔面領域が入るように 139×88pixel にトリミングし, 最大値 1, 最小値 0 となるように規格化した。本研究では, 覚醒低下抑制区間における顔面熱画像を正常状態, 眠気レベル 5 と分類された顔面熱画像を異常状態とし, それぞれの状態における顔面熱画像を Normal, Anomaly と定義した。全 5 名の被験者から採取したサンプルのうち, 4 名分の Normal にあたる 3600 枚の熱画像から, 熱画像の一部をランダムに 10 万枚切り出した。これらのパッチを学習データとし, VAE を用いてモデルを構築した。最適なパラメータを探索するため, 学習条件は潜在空間の次元数  $N_z = \{2, 8, 16, 32, 64\}$ , 熱画像を切り出すパッチサイズ  $S_p = \{8^2, 16^2, 32^2, 64^2\}$  から選択した。パラメータの組み合わせは 20 通りである。なお, 構築されるモデルのばらつきを考慮し, 各パラメータの組み合わせに対して 10 個のモデルを構築した。

### 4.3 覚醒低下検出モデルの評価

学習データから除外した 1 名分のデータセットの Normal, Anomaly それぞれ 100 枚をテストデータとして, 学習済みモデルの汎化性能を評価した。なお, 被験者ごとに 5 分割交差検証を行った。評価指標として Receiver Operating Characteristic (ROC) 曲線の下部面積である Area Under Curve (AUC) を用いた。AUC は 0 から 1 までの値をとり, 値が大きいほど, モデルの精度が高いことを示す。予測モデルの AUC の目安は 0.9 が high accuracy, 0.7 以上 0.9 未満が moderate accuracy, 0.5 以上 0.7 未満が low accuracy とされている<sup>(7)</sup>。

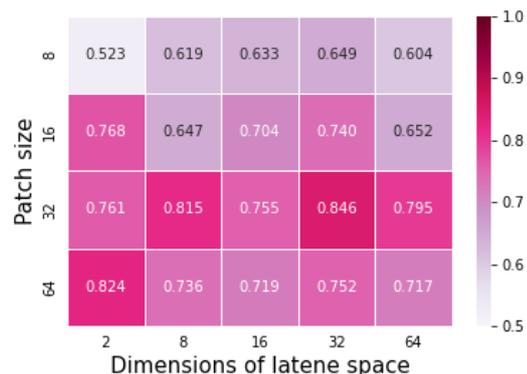


図2 各パラメータにおける平均 AUC

## 5. 結果及び考察

図 2 に各パラメータ下で構築した 10 個のモデルの平均 AUC を示す。ただし, 横軸は潜在空間の次元数, 縦軸はパッチの一边の長さである。得られた結果より, パッチサイズが 32<sup>2</sup>以上になると AUC が 0.7 以上になった。しかし, 各パッチサイズにおいて, 潜在空間の次元数と AUC の値に共通の関係は確認できず, 最適な潜在空間の次元数の決定には至らなかった。

## 6. まとめ

本研究では, VAE を用いた顔面熱画像に基づく一過性覚醒低下検出を目指すとともに, 検出精度向上に向けた検討として, 覚醒低下検出における VAE のパラメータの 1 つである潜在空間の次元数最適化を検討した。その結果, パッチサイズが 32<sup>2</sup>以上になると AUC が 0.7 以上になったが, 最適な潜在空間の次元数の決定には至らなかった。

### 参考文献

- (1) Saito, Y, Makoto, I, Inagaki, T: “Driver assistance system with a dual control scheme: Effectiveness of identifying driver drowsiness and preventing lane departure accidents”, IEEE Transactions on Human-Machine Systems, Vol.46, No.5, pp.660-671 (2016)
- (2) Adachi, H, Oiwa, K, Nozawa, A: “Drowsiness level modeling based on facial skin temperature distribution using a Convolutional Neural Network”, IEEJ Transactions on Electrical and Electronic Engineering (TEEE C), Vol.14, No.6, pp.870-876 (2019)
- (3) Masaki, A, Nagumo, K, Lamsal, B, Oiwa, K, Nozawa, A: “Anomaly detection in facial skin temperature using Variational Autoencoder”, Artificial Life and Robotics, Vol.26, Iss.1, pp.122-128 (2021)
- (4) 高野聖仁, 南雲健人, 大岩孝輔, 野澤昭雄: 「VAE を用いた顔面皮膚温度の異常検知精度向上の試み」, 2021 年電気学会電子・情報・システム部門大会, GS11-3 (2021)
- (5) 野村泰稔: 「深層学習を用いた異常検知技術」, 材料, Vol.69, No.9, pp.650-656 (2020)
- (6) 北島洋樹, 沼田仲穂, 山本恵一, 五井美博: 「自動車運転時の眠気の予測方法についての研究(第一報, 眠気表情の評定法と眠気変動の予測に有効な指標について)」, 日本機械学会論文集 C 編, Vol.63, No.613, pp.3059-3066 (1997)
- (7) John A. Swets: “Measuring the Accuracy of Diagnostic System”, American Association for the Advancement of Science, Vol.240, No.4857, pp.1285-1293 (1988)