

深層学習を用いた低解像度画像に対する見守りのための開・閉眼の識別 Discrimination of Eye-Opening and Closing in Low-resolution Images Using Deep Neural Networks

福永 敦[†] 金川 明弘[†] 滝本 裕則[†]
Atsushi Fukunaga Akihiro Kanagawa Hironori Takimoto

1. はじめに

現在、日本は急速な高齢化に伴い、介護職員の人手不足問題が深刻化している。この問題を解消するため、全国の介護施設では介護支援機器の導入が検討されており、見守りシステムなどの技術が注目を集めている。

介護や医療の現場で用いられる意識レベルの指標では、刺激に対して開眼するか否かで意識レベルを判別しており、被介護者に意識障害の疑いがある場合には、開・閉眼の状態が重要とされている。小松ら[1]はこれに着目し、可視光カメラで睡眠時以外の被介護者を見守り、開・閉眼の状態を識別することで、意識障害を検知する手法を提案した。研究では、カメラが検出した顔画像から開・閉眼を識別する方法として、深層学習の一種である畳み込みニューラルネットワークを用いており、低解像度の顔画像に対して開・閉眼を識別することで、1 台の広角カメラを用いた見守りの実現可能性を示した。

本稿では、介護現場への応用を可能にするため、傾いた顔に対する検出や、様々な撮影条件下における識別を想定した開・閉眼識別モデルを構築し、検証する。

2. 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク (Convolutional Neural Networks:CNN) は、畳み込み層、プーリング層、全結合層を基本構成とする順伝播型のニューラルネットワークである。畳み込み層では、畳み込みフィルタを入力画像に適用することで、フィルタが表す特徴的な濃淡構造を画像から検出する。プーリング層では、畳み込み層で抽出された特徴の位置感度を若干低下させることで、特徴に微小な位置変化があった場合でも正しく認識する。

一般に、ネットワーク構造の浅い層ほど汎用的な特徴が抽出され、深い層ほど学習データに特化した特徴が抽出されるため、浅い層には別の学習済みモデルの重みを転用するファインチューニングという手法を用いることで、少ないデータセットからでも迅速な特徴抽出が可能となる。

本研究では、VGG16 [2]とよばれる CNN 構造を用いて学習を行い、開・閉眼を識別する。また、浅い層には ImageNet とよばれる大規模データセットで訓練したモデルの重みを固定し、一部の深い層のみ学習データで訓練する。なお、活性化関数は ReLU 関数、出力関数はソフトマックス関数を用い、ミニバッチ勾配降下法を用いて損失関数であるクロスエントロピー関数を最小化し、重みの最適化を行う。本研究で用いる VGG16 の構造を図 1 に示す。



図 1 VGG16 のモデル構造

3. データセット

本研究で用いる学習用データセットは、岡山県立大学の学生 50 名の開眼状態・閉眼状態を撮影したものに加え、インターネット上より収集した開眼状態の顔画像 100 枚、およびそれらを画像処理により閉眼させた顔画像 100 枚の計 300 枚を基本構成とする。開眼画像から閉眼画像を生成することにより、容易に閉眼画像を入手できるだけでなく、同一人物の開・閉眼画像が得られるため、開・閉眼以外の特徴の差異がなくなり、識別に必要な特徴を正確に抽出できる。画像処理による閉眼画像生成の例を図 2 に示す。



図 2 開眼画像(左)から閉眼画像(右)を生成

CNN では一般に、より多様な画像を大量に学習させることで、識別モデルの汎化能力を向上させることができる。また、普遍的特徴を自動抽出する CNN は、画像内で平行移動した同一の特徴に対しては応答の不変性が保持される一方で、回転などの角度変化に対しては応答の不変性を有さない。よって本研究では、これらの特性に基づいた学習データセットの拡張を画像処理により行ない、識別モデルの汎化能力の向上を図る。

まず、学習データにおけるすべての顔画像を 128×128 ピクセルに変換する。次に、変換後の画像それぞれに対し、45 度、90 度、135 度、180 度、225 度、270 度、315 度回転した画像を生成し、1 枚につき 8 パターンの画像が生成される。また、生成された画像それぞれに対して、輝度を上げた画像、輝度を下げた画像を生成する。この処理を左右反転された画像に対しても行い、計 300 枚の学習データセットは 14,400 枚に増加した。本研究で用いた学習用データの拡張例を図 3 に示す。

[†] 岡山県立大学

また、識別モデルの性能を検証するためのテストデータセットでは、インターネット上で実際の開眼画像と閉眼画像をそれぞれ 100 枚収集し、すべて 128×128 ピクセルに変換した後、学習用データで行なった 8 パターンの回転と、左右反転の処理をそれぞれの画像に施す。この処理により、テストデータは 200 枚から 3200 枚に増加した。



図3 学習用データの拡張例

本研究では、広角カメラによる見守りを実現するため、低解像度の画像で学習・テストを行い、識別モデルの有用性を検証する。学習・テストそれぞれにおいては、顔画像データを任意のサイズまで縮小し、VGG16 の推奨サイズまで拡大してから入力するものとする。

4. 画素補間

画像に対して縮小や拡大、回転などの処理を施す場合、新規の画素を周辺の画素で近似して補う画素補間が必要である。変換処理の内容に応じて最適な補間手法を適用することにより、変換後の画像のジャギー軽減や、滑らかで自然となるような内挿を実現できる。

本研究では、学習・テストデータの顔画像に対する縮小処理にはバイリニア補間法を適用し、回転・拡大処理にはバイキュービック補間法を適用する。

バイリニア補間法は、補間画素の 2×2 近傍の画素を参照し、その輝度値の中間の値を新規画素の輝度値とする手法である。画像座標系の縦方向、横方向それぞれにおいて補間画素から参照画素までの距離を d とおくと、補間画素の輝度値は参照画素の輝度値と重み W の加重平均で決定され、重み関数 W は式 (1) のように表される。

$$W = \begin{cases} 1-d & (0 \leq d < 1) \\ 0 & (1 \leq d) \end{cases} \quad (1)$$

バイキュービック補間法は、補間画素の 4×4 近傍の画素を参照し、その輝度値の中間の値を新規画素の輝度値とする手法である。重み W は距離 d を用いて式 (2) のように表され、 sinc 関数をテイラー展開により 3 次の項まで近似した関数である。ここで、 a は鮮鋭度のパラメータであり、一般に -1.0 から -0.5 の値をとる。

$$W = \begin{cases} 1 - (a+3)d^2 + (a+2)d^3 & (0 \leq d \leq 1) \\ -4a + 8ad - 5ad^2 + d^3 & (1 \leq d < 2) \\ 0 & (2 \leq d) \end{cases} \quad (2)$$

5. 実験

実験では、あらかじめ解像度を落とした画像を学習させることで、実際に低解像度の顔画像が検出された場合でも、開・閉眼識別が可能となるように識別モデルを構築した。

顔画像のサイズを 64×64、32×32、16×16 ピクセルに変換後、VGG16 モデルの推奨サイズ(224×224 ピクセル)まで拡大し、それぞれの解像度で学習・テストを行った。テスト結果を表 1 に示す。

表1 様々な解像度でのテスト結果

画像サイズ[pix]	64×64	32×32	16×16	8×8
テスト精度[%]	87.09	80.63	66.22	57.84

また、データ拡張における回転を行わず、傾いている顔を取り除いたデータセットでも学習・テストを行ったが、精度の向上は見られなかった。さらに、輝度調整の度合いを弱めたデータセットで構築したモデルでは、テスト結果の精度は 1~2%ほど低下した。

6. まとめ

本稿では、傾いた顔などの顔検出が困難な画像に対し、様々な角度の顔画像を学習させた CNN を用いることで、見守りシステムのための開・閉眼の識別が可能であることを示した。また、輝度調整や人工的な閉眼画像生成など、多様性に富んだ画像処理を学習データに施し拡張することにより、様々な撮影条件下における実際の画像に対しても有効な識別モデルを構築した。

また、ImageNet で学習したモデルの重みを浅い層に適用するファインチューニングでは、少ないデータセットからでも十分な特徴抽出が可能であり、顔画像などの特徴量が異なるタスクにおいても効率のよい学習が可能であった。

様々な解像度における学習・テストでは、あらかじめ低解像度の画像を学習させることにより、実際の低解像度画像に対する識別が期待できるが、介護現場へ応用できる精度に達していないため、あらゆる顔の傾きに対する顔検出、および開・閉眼の識別手法において検討の余地がある。

今後の課題としては、CNN を用いた顔検出器の作成により、傾いている顔や正面を向いていない顔を含む、様々なパターンの顔を画像から検出し、開・閉眼を識別するモデルを構築すると同時に、介護現場への応用に向けて精度向上や処理時間短縮などのパフォーマンス向上を実現する必要がある。

謝辞

本研究の一部は科学研究費助成事業 基盤研究(C)(課題番号: 17K01573)の補助を受けたことを記し、ここに謝意を表します。

参考文献

- [1] 小松他, “ディープラーニングを用いた見守りのための閉眼状態検出”, 日本 OR 学会秋期アブストラクト集, pp.134-135, 2017.
- [2] K.Simonyan, A.Zisserman, “Very Deep Convolutional Networks For Large-Scale Image Recognition”, Visual Geometry Group, Dpt. of Eng. Sci, University of Oxford, 2015