CNN を用いた顔認証システムはプライバシーバイザーに効果があるのか

脇 一史 †

† 明治大学 総合数理学部 先端メディアサイエンス学科 菊池研究室

1 はじめに

顔認証カメラによって取得した映像データを顧客の行動特定に活用することが進んでいる．その一方, 追跡を停止して欲しい要求や自分の登録データの開示請求などの課題が生じている．これに対して，山田らは，反射性のある素材により検出を防止するメガネ型デバイスのプライバシーバイザー [1](以下、バイザー) を提案している．

そこで，本研究では，バイザーに追跡を停止出来る効果があるのかを検証するため，画像識別として主流であるディープラーニングを用いてシステムを試験実装し，様々な条件下のもとで被験者の顔を識別する実験を行う．本実験では, TensorFlow 上で実行可能なフレームワークである Keras で Convolutional Neural Network(CNN) を実装し，様々な条件に対してバイザーごと学習させて，検出機能を検査する．実験結果を基に，セキュリティと生活者のプライバシーについて考察する．

2 実験目的

1. CNN による顔認証の精度を明らかにする．
2. 素顔を学習した場合に，素顔及び, バイザー OFF（一般的なサングラス状態）, バイザー ON（顔検出を防ぐ保護状態）の計 3 種類に対し識別精度の違いを明らかにする．
3. バイザーごと学習することで，被験者を追跡できるのかを明らかにする．

3 実験内容

3.1 顔画像データの取得

CNN は明るさや表情, 髪形などの変化に対して汎用性を持つ必要性がある. そのため, 顔画像データを被験者 20 名に対し 1 日毎に 100 枚ずつ異なる時間に撮影する．openCV を用いて顔を検出し，顔がその範囲内に収まるように調整する．使用するカメラは iMac に標準で搭載されている web カメラである．取得間隔については，被験者を 3 フレーム毎の感覚で顔を上下左右に動かしながら撮影する．さらに，表 1 に示す異なる条件下で認証精度を評価するために, 顔を固定させた状態で，別日にそれぞれ 100 フレーム分の評価データを取得した．

表 1　データ拡張

手法	コントラスト調整	左右反転	ガウシアンノイズ
種類	12% 23% 35% 47% 59% 70%	1	2 4
計	6	1	2

3.2 顔画像データの拡張

取得した顔画像に対し, keras を用いて 224 × 224 にリサイズを行い, 表 1 に示す各パラメータについて 9 種類のデータに拡張する. 元画像を含め, 一人一種類当たり 2,100 枚, 合計 42,000 枚の顔画像データを作成した．

3.3 CNN の構成

本研究では VGG-16[2] と呼ばれる ImageNet Large Scale Visual Recognition 2014(ILSVRC2014) の 1000 クラス識別タスクにおいて, 2 位となった識別手法を改良したモデルを用いる．VGG は構造がシンプルであり，応用性が高いため多用されている．さらに，[3] を参考に，ニューロンをランダムに消去しながら学習する Dropout をすることで，過学習を抑制できる手法を追加する．認証精度を向上するため，学習済みのモデルを転用して新たなモデルを生成する fine tuning を使用する．上記のモデルに対し，学習データを 44000 枚，テストデータを 2000 枚に分け, 学習回数である epoch は，10 回とする．学習した CNN に対し, web カメラで取得した顔画像データをリアルタイムに識別する顔認証システムを実装した．実行画面を図 1 に示す．

3.4 学習と評価が同じ組み合わせの評価手法

3 種類のデータに対し，学習データと評価データが同一のもので実験を行う．学習データを学習させた CNN について，同一条件の評価データでテストする．

3.5 学習と評価が異なる組み合わせの評価手法

3 種類のデータに対し，種類別の汎用性を調べるため，学習データと評価データが異なるもので実験を行う．学習データで学習させた CNN について，異なる条件の評価データでテストする．

†Kazushi Waki，Department of Frontier Media Science，School of Interdisciplinary Mathmatical Science，Meiji University，Kikuchi Laboratory.

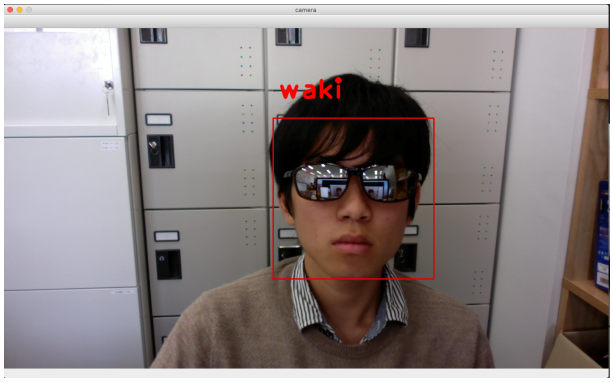

図 1　顔認証システム

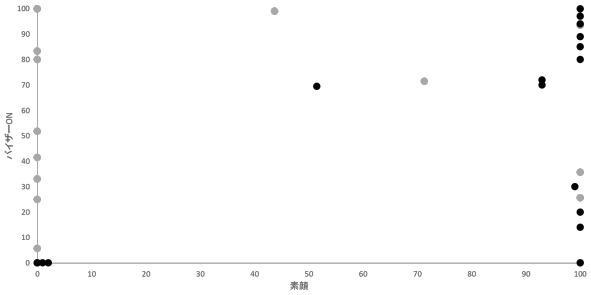

図 2　各ユーザの認証精度の変動 (黒点：素顔，灰点：バイザー ON)

表 2　学習させた CNN に対し異なる条件下で評価した場合の平均適合率

学習＼評価	素顔	バイザー OFF	バイザー ON	平均
素顔	62.59	36.67	20.75	40.00
バイザー OFF	27.36	59.37	44.78	43.80
バイザー ON	40.34	42.58	42.28	41.73
平均	43.40	46.21	35.90	41.84

4　実験結果

学習させた CNN に対し，異なる条件下で評価した平均適合率を表 2 に示す．素顔を学習させた場合，素顔の評価データについての適合率は 62.59%, と最も高い結果となった．次に，外乱に対する個人差を見るために，素顔とバイザー ON 時の被験者毎の平均適合率の散布図を図 2 に示す．被験者によって，素顔の評価にバイザーの汎用性の違いがみられた．さらに，バイザーの評価データは素顔の認証能力がないことも明らかとなった．

A 被験者の再現率は

$$P_A = \frac{\text{A と正しく判定した数}}{\text{A 被験者と判定した全ての画像数}}$$

と定義する．と定義する．全ユーザについての適合率の平均を平均適合率と呼ぶ．

4.1　考察

表 2 より，バイザーを学習し，バイザーで評価した際に平均適合率が 42% 程であった原因として，目元が隠れているため特徴量が減り，適合率が低下したためと考える．これに対し，[4] では一般的なサングラスを用いて評価したが，94.4% と非常に高い精度であった．これは，一般的なサングラスはグレーに着色された色付きレンズを使用しており，レンズ部の光反射が無いため CNN への影響が少なく，サングラス以外の露見している部分に対して細かく学習したことで精度を高めたと考える．

一方，バイザーはミラー型のレンズを使用しているため，バイザーに反射する景色が CNN に影響を与えたと考えられる．反射する景色は，被験者の顔の傾き等の個人の特徴が影響する場合と，研究室内の人の移動等の外部の環境変化が影響する二通りがある．従って，被験者毎の特徴量に違いがなかったため，適合率に結びつかなかったと考える．本研究と [4] の比較を表に示す．

表 3　本研究と [4] の比較

	2017	2018
手法	独自 CNN(vgg-11)	keras(vgg-16)
外乱	帽子，マスク，サングラス	プライバシーバイザー
精度	54.80	41.73
外乱頑強性	低い	高い

5　おわりに

バイザーを装着することにより，顔認証による追跡を素顔の場合と比べ 20 %～35% 防止できることが示された．しかし，バイザーを装着した画像を学習すると，本人と識別される割合が高くなることが分かった．

画像処理によってレンズ部の反射の景色を消去することで，一般的なサングラスと同様の精度になるかは今後の課題である．

参考文献

[1] 山田隆行，合志清一，越前功 “光の反射・吸収特性を利用した撮影画像からの顔検出防止手法”，情報処理学会論文誌 Vol.55，No 9，pp.2104-2119，Sep. 2014.

[2] Karen Simonyan and AndrewZissserman(2014):Very Deep Convolutional Networks for Large-Scale Image Recognition. ICLR,2014.

[3] 齋藤康毅, “ゼロから作る Deep Learning python で学ぶディープラーニングの理論と実装”, OREILLY 2016.

[4] 脇一史, 菊池浩明, “CNN を用いた顔認証システムの開発と追跡停止に対する評価”, 情報処理学会第 80 回全国大会，7W-03，pp.3_543-3-544，2018.