明治大学総合数理学部

2018 年度

卒　業　研　究

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

学位請求者　先端メディアサイエンス学科

脇　一史

目次

第 1 章

はじめに

1.1　研究背景

顔認証カメラによって取得した顧客の行動や履歴情報の防犯や商用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顧客の削除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．2014 年には，JR 大阪駅一帯にある商業ビル・公共空間の「大阪ステーションシティ」で実施する計画の「ICT 技術の利用認証実験」を予定していたが，プライバシー侵害に関する懸念が高まり，延期に追い込まれた [1]．その他では，実験のために写真撮影に同意した被験者の 3 割近くが facebook 等の SNS の情報と比較することで，氏名や住所が特定できることが判明した [2]．これに対して，山田らは，反射性のある素材により顔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 [3](以下、バイザー) を提案している．

1.2　研究目的

本研究では，一般的に顔検出を妨害する手段であるマスクや帽子などの外乱 [4] 及び，プライバシー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試験システムを実装し，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．本研究では, python の計算パッケージである numpy のみを利用するものと，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．

第 2 章

基本定義

2.1　プライバシーバイザー

プライバシーバイザーとは，デジタルカメラの顔検出をする際に主流である Viola-Jones 法 [5]，[6]，[7] で用いられる Haar-like 特徴量の算出を防ぐことで顔検出を防止するデバイスである．Haar-like 特徴量とは顔の明暗差に着目した特徴量で，2 つの異なる矩形領域で構成された Haar-like 特徴を用いて特徴量を算出する．図 2.1 に Haar-like 特徴の基本パターンを示す．プライバシーバイザーは顔全体で最も明暗差のある目元に対して，本来の明暗と反対となるように目の周りに白い反射性素材を用い，鼻筋に黒い吸収性素材を用いることで顔検出を失敗させる．本研究で使用したプライバシーバイザーを図 2.2 に表す．このプライバシーバイザーは，レンズとテンプルの接合部であるヨロイが上下に可動出来るようになっており，レンズを閉じた状態がサングラスの機能のみ持つ「通常モード」（以下，バイザー OFF），レンズを開いた状態が顔検出を不可能にする「保護モード」（以下，バイザー ON）となっている．

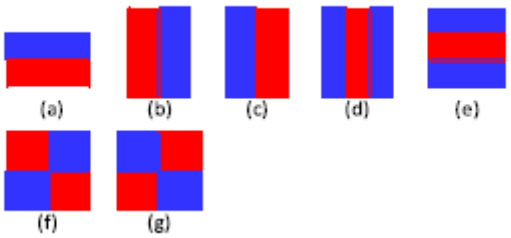


図 2.1: Haar-like 特徴の基本パターン

図 2.2: 本研究で使用したプライバシーバイザー（Technology by NII ECHIZEN Laboratory Designed by sowell design office）

2.2 Convolutional Newral Network(CNN)

Convolutional Newral Network(CNN) とは，機械学習の一つであるディープラーニングのうち，主に画像識別で用いられる手法である．Convolutional は畳み込みを表し，ニューラルネットワークの入力層に畳み込み層を追加することで，画像のピクセルの座標位置を学習に取り入れることが出来る．ニューラルネットワークは人間の脳神経系のニューロンを数値モデル化したものの組み合わせであり，学習データを入力することで，各々のニューロンが持つ重みとバイアスが逆誤差伝播法により適正化することで学習を行う．図 2.3 にニューラルネットワークの概要図を示す．

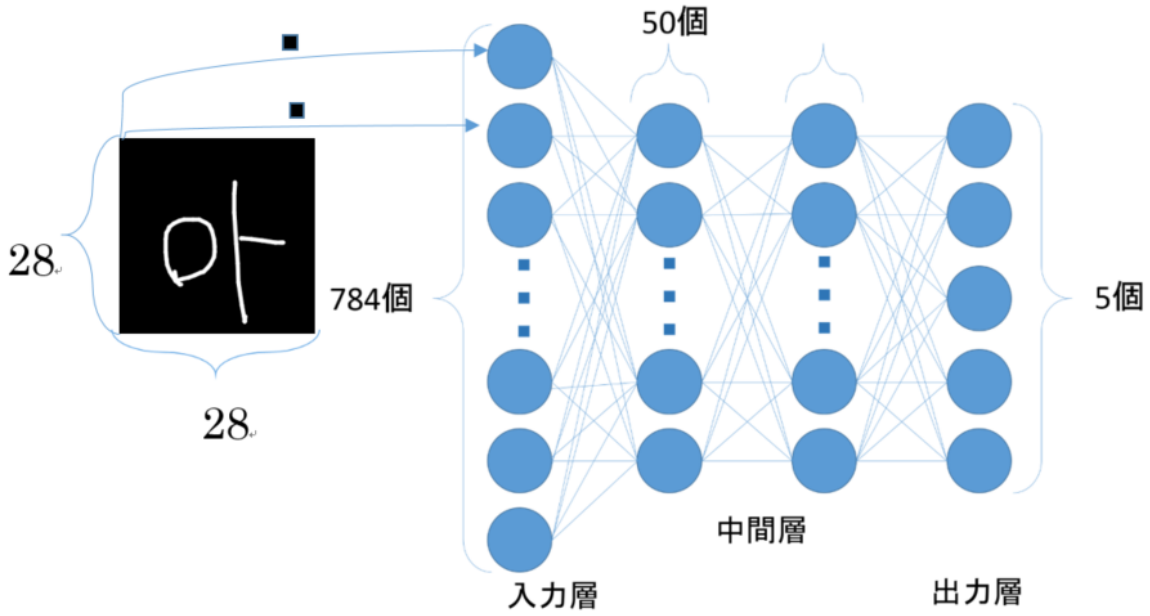


図 2.3: ニューラルネットワークの概念図 (ハングル語の母音を画素ごとに入力)

2.3　適合率及び平均適合率

本研究では，顔認証の精度の評価指標として適合率を算出する．適合率とは正と予測したデータのうち，実際に正であるものの割合である．A さんの適合率は

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

と定義する．平均適合率は全ユーザについての適合率の平均とする．

2.4　再現率及び平均再現率

本研究では，顔認証の精度の評価指標として再現率を算出する．再現率とは実際に正であるもののうち，正であると予測されたものの割合である．A さんの再現率は

$$R_A = \frac{\text{A と正しく判定した数}}{\text{本物の A さんの画像数}}$$

と定義する．平均再現率は全ユーザについての再現率の平均とする．

第 3 章

プライバシーバイザー評価実験

3.1　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．

3.2　実験内容

3.2.1　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, CNN で用いる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顔を撮影し，顔の検出は，openCV を用いて取得する範囲を画面上に表示させ，図 3.1 に示すように，被験者の顔の範囲である青枠に基準となる赤枠を重ね合わせるように撮影した．使用するカメラは mac に標準で搭載されている web カメラを用いる．取得間隔については，パソコンの前にいる被験者を 3 フレーム毎に顔を上下左右に動かしながら撮影し, 2 日間で計 200 枚の学習データを取得した（図 3.2）．さらに，表 2 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．

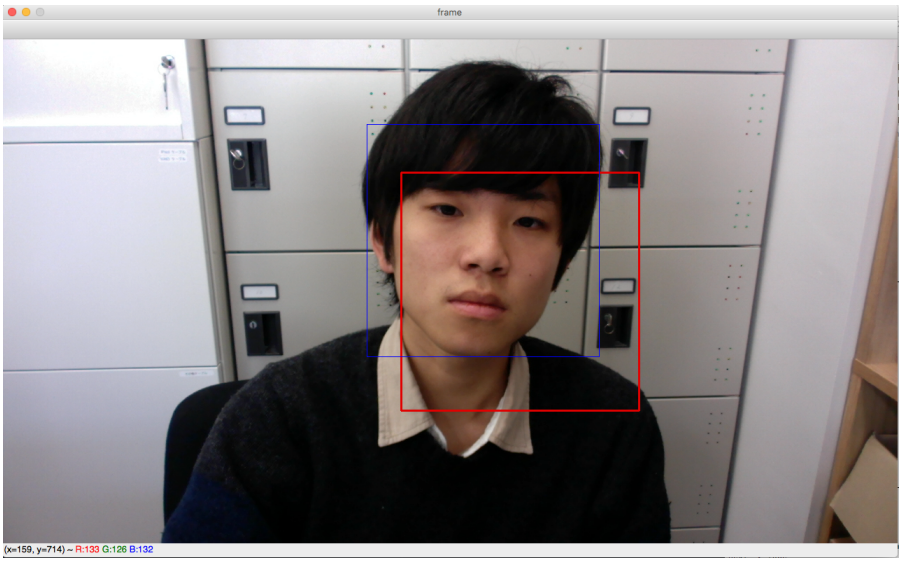


図 3.1: 顔認証システム

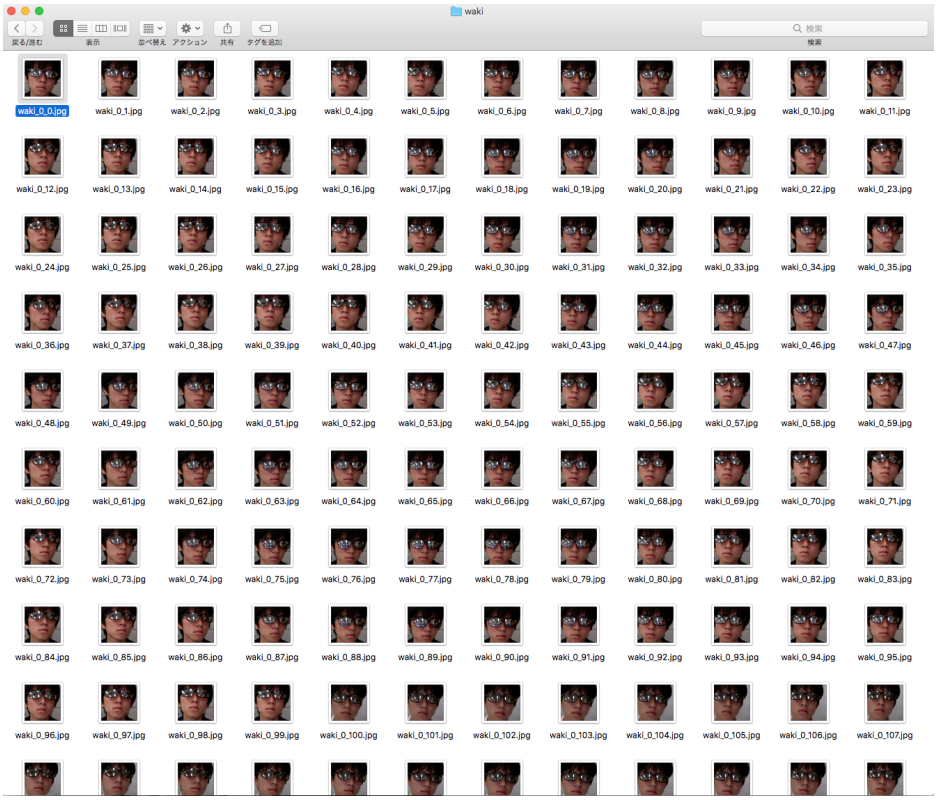


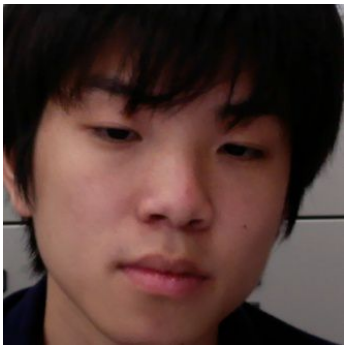
図 3.2: 学習データの一覧

3.2.2 顔画像データの拡張

CNN を用いる場合にデータセットの数は重要な要素であり，学習データの枚数を増やすことにより CNN の精度が向上する．本研究では，Data Augmentation と呼ばれるデータ拡張を行う．Data Augmentation は直訳で「水増し」と呼ばれ，元の学習データに変換を加えてデータ量を増やす手法である．取得した顔画像に対し, keras の機能を用いて 224 × 224 にリサイズを行い, openCV を用いて表 4.1 に示す各パラメータについて 9 種類のデータに拡張した．コントラスト調整は一定以下の低輝度の画素を 0，一定以上の高輝度の画素を 255 にし，中間の輝度を調整する．本研究では調整の割合を輝度の幅で区分している．ガウシアンノイズは各画素にガウス分布に基づく生成値を足してノイズを付加する．本研究では標準偏差 σ を表 4.1 に示す値に設定する．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．図 3.3 に各々のパラメータで拡張した結果を示す．

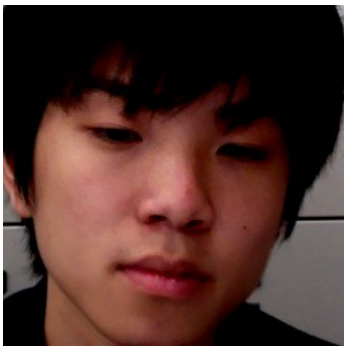
表 3.1: データ拡張

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

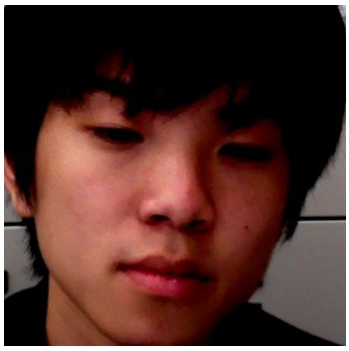


元画像

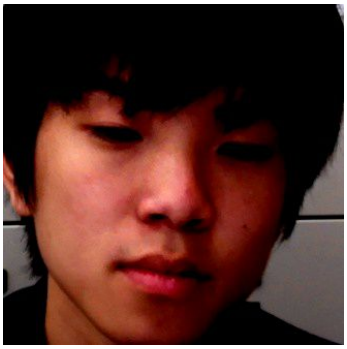
コントラスト 12%



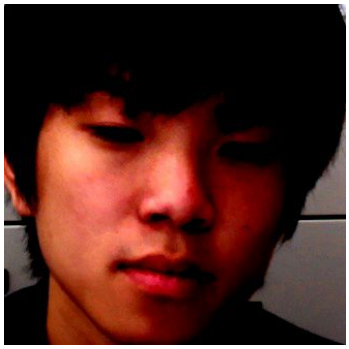
コントラスト 23%



コントラスト 35%



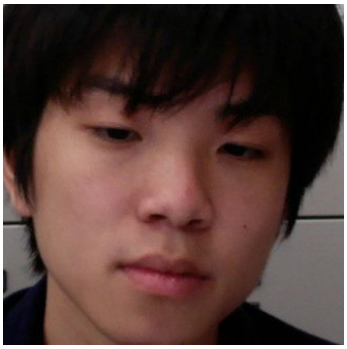
コントラスト 47%



コントラスト 59%

コントラスト 70%

左右反転



ガウシアンノイズ σ=2

ガウシアンノイズ σ=4

図 3.3: 画像の変換例

3.2.3 CNN の構成

本節では Keras を用いて，VGG-16[8] 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参考に作成した．ImageNet とは，2009 年にプリンストン大学によって発足された機械学習のためのデータセットであり，2 万を超えるカテゴリと 1400 万枚以上の画像を有している．VGG とは畳み込み層と全結合層を組み合わせたシンプルな構造かつ，高い精度を発揮できるため様々な機械学習のモデルとして多用されている手法である．さらに，本研究では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 [9] を追加した. また, 認証精度を向上するため，ImageNet の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 を使用した．これにより，学習データの数が少ない状態においても高い精度の識別が可能である．本研究で実装した CNN の階層図を図 7 に示す．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, 学習パラメータは表 4.2 に示す設定とする．

表 3.2: CNN の構成表 (活性化関数及び損失関数は vgg16 に基づく)

パラメータ名	値
モデル	vgg16+Dropout
入力画素	224 × 224
バッチサイズ	128
エポック数	10
出力層	20

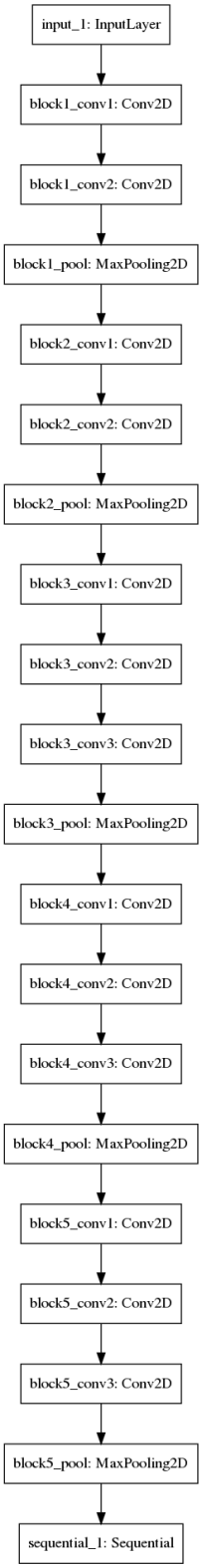

input_1: InputLayer
block1_conv1: Conv2D
block1_conv2: Conv2D
block1_pool: MaxPooling2D
block2_conv1: Conv2D
block2_conv2: Conv2D
block2_pool: MaxPooling2D
block3_conv1: Conv2D
block3_conv2: Conv2D
block3_conv3: Conv2D
block3_pool: MaxPooling2D
block4_conv1: Conv2D
block4_conv2: Conv2D
block4_conv3: Conv2D
block4_pool: MaxPooling2D
block5_conv1: Conv2D
block5_conv2: Conv2D
block5_conv3: Conv2D
block5_pool: MaxPooling2D
sequential_1: Sequential

図 3.4: CNN の全体図

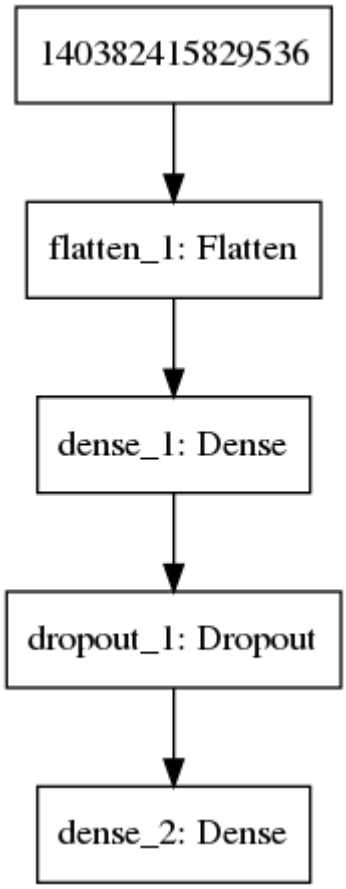


図 3.5: Sequential 層 (図 3.4 の最下層の内容)

3.2.4　リアルタイム顔認証システム

web カメラの映像をフレーム毎に取得し，opwnCV を用いて顔領域のみトリミングしたものを学習済みの CNN に入力させ，リアルタイムに被験者を判定するシステムを実装した．実行画面を図 3.6 に示す．

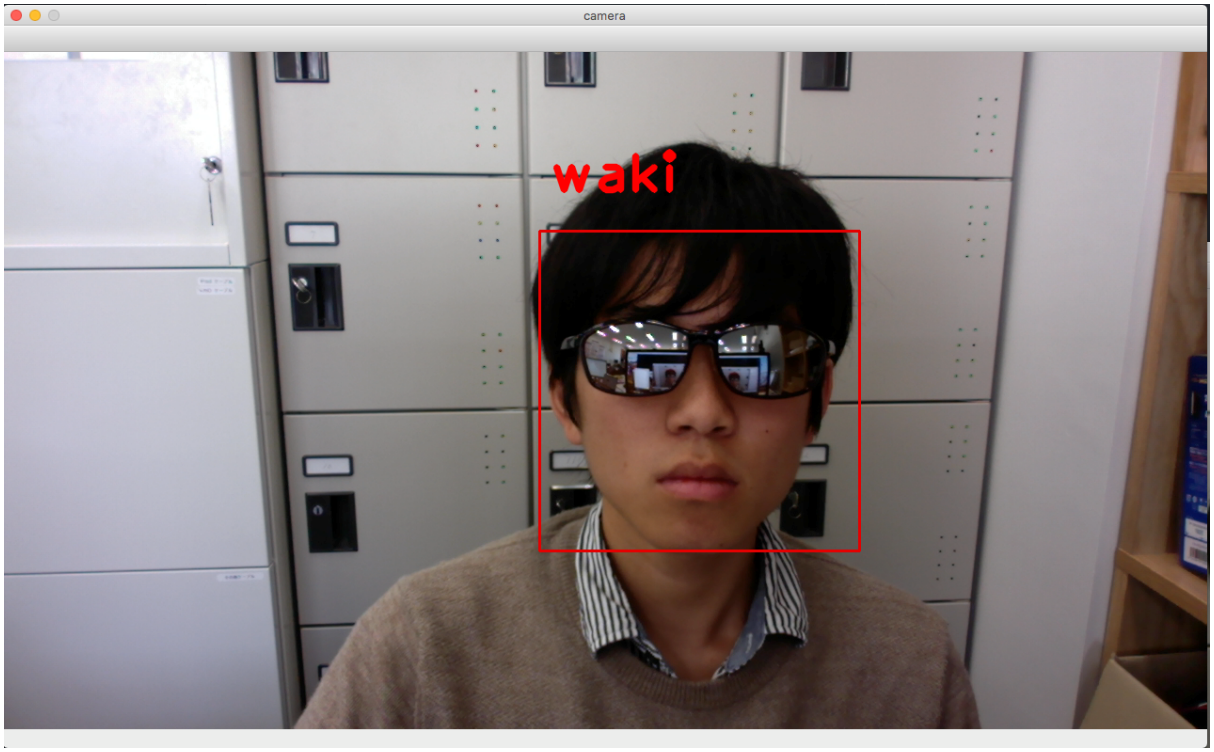


図 3.6: リアルタイム顔認証システムの実行例（バイザー ON）

3.2.5　学習と評価が同じ組み合わせの評価手法

素顔，バイザー OFF，バイザー ON の計 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．学習データで学習させた CNN について，同一条件の評価データでテストする．

3.2.6　学習と評価が異なる組み合わせの評価手法

素顔，バイザー OFF，バイザー ON の 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．

3.3　実験結果

3.3.1　素顔で学習した CNN に対する追跡停止の評価

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 3.3 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 3.7 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．各評価データのばらつきを表す標準偏差を表 3.4 に示す．各々の外乱の組み合わせ毎のばらつきは一様であった．最後に，平均再現率を 3.5 に示す．平均適合率とほぼ変わらない数値となった．

3.3.2　外乱を加えた画像で学習した時の評価

表 3 より, それぞれの外乱について以下に述べる.

- バイザー OFF を学習：素顔に対しては 27.36% と低かったものの，バイザー ON 時では 44.78% と高い結果となった.
- バイザー ON を学習：3 種類すべての評価データに対して約 40% と平均的な適合率が得られた.

表 3 に示す値を CNN のエポックごとの学習推移で表したグラフを表 3.8 に示す.

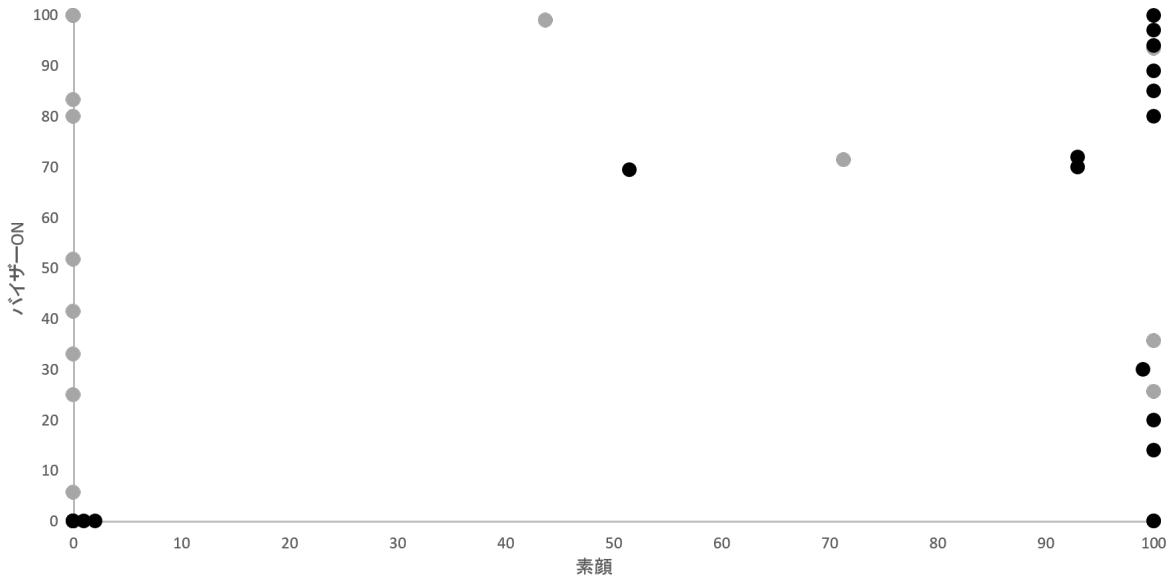


図 3.7: 各ユーザの認証精度の変動 (黒点：素顔，灰点：バイザー ON)

表 3.3: 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

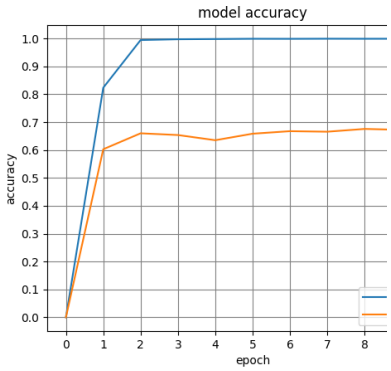
学習＼評価	素顔	バイザー OFF	バイザー ON	平均
素顔	62.59	36.67	20.75	40.00
バイザー OFF	27.36	59.37	44.78	43.80
バイザー ON	40.34	42.58	42.28	41.73
平均	43.40	46.21	35.90	41.84

表 3.4: 平均適合率時の標準偏差

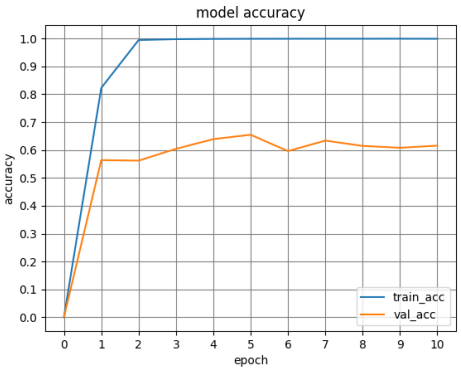
学習＼評価	素顔	バイザー OFF	バイザー ON	平均
素顔	0.40	0.43	0.38	0.40
バイザー OFF	0.36	0.37	0.39	0.37
バイザー ON	0.41	0.34	0.38	0.38
平均	0.39	0.38	0.38	0.38

表 3.5: 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再現率 [%]

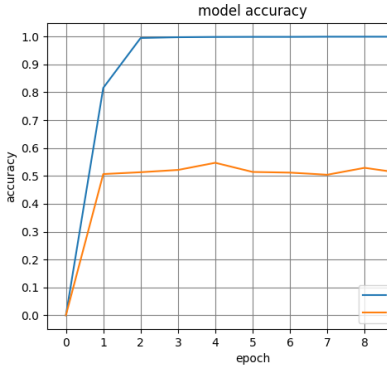
学習＼評価	素顔	バイザー OFF	バイザー ON	平均
素顔	66.98	25.70	11.93	34.87
バイザー OFF	39.15	62.68	39.75	47.19
バイザー ON	41.03	50.10	46.25	45.79
平均	49.05	46.16	32.66	42.62



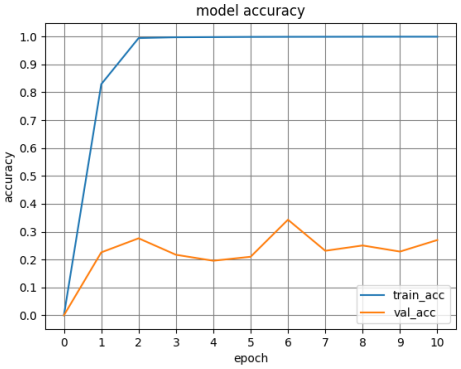
学習:素顔
評価:素顔



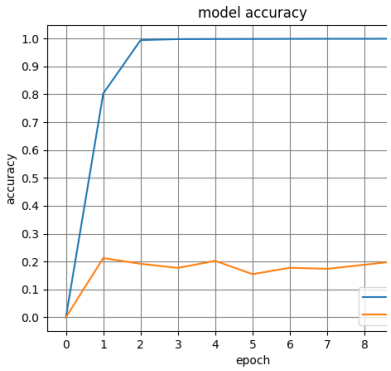
学習:バイザー OFF
評価:バイザー OFF



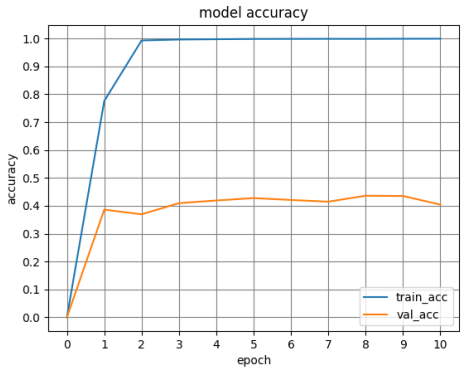
学習:バイザー ON
評価:バイザー ON



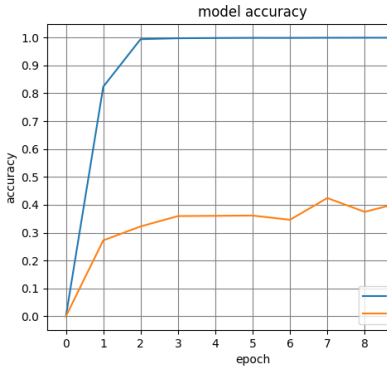
学習:素顔
評価:バイザー OFF



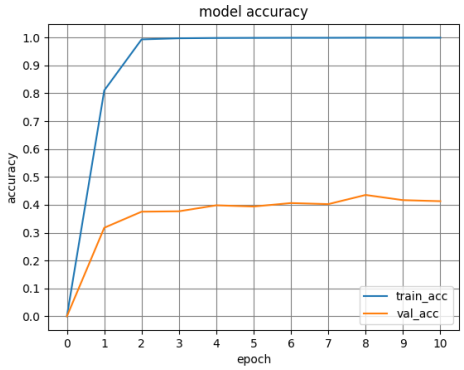
学習:素顔
評価:バイザー ON



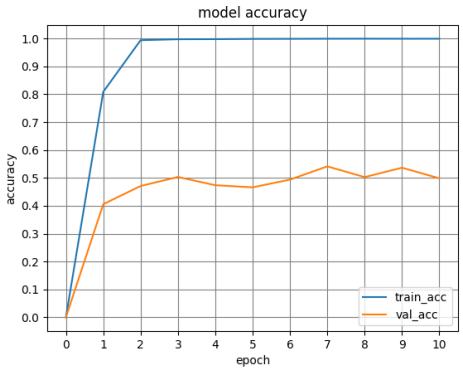
学習:バイザー OFF
評価:素顔



学習:バイザー OFF
評価:バイザー ON



学習:バイザー ON
評価:素顔



学習:バイザー ON
評価:バイザー OFF

図 3.8: エポック毎の学習遷移（青線：学習データ　橙線：評価データ）

第 4 章

日用品を用いた評価実験

4.1　実験目的

1. CNN による顔認証の精度を明らかにする．
2. 素顔を学習させた場合に，素顔及び, マスク, サングラス, 帽子, マスク＋サングラスの計 5 種類のうち最も識別精度を下げる外乱はなにかを明らかにする．
3. 外乱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．

4.2　実験内容

4.2.1　顔画像データの取得

CNN で用いる顔画像データを被験者 6 名に対し 1 日毎に 100 枚ずつ異なる時間に顔を撮影し，顔の検出は，openCV を用いて 取得する範囲を画面上に表示させ， 顔がその範囲内に収まるように撮影した．使用するカメラは mac に標準で搭載されている web カメラを用いる．取得間隔については, パソコンの前にいる被験者を 3 フレーム毎に顔を上下左右に動かしながら撮影し, 5 日間で計 500 枚取得した. さらに被験者 5 名に対し, 表 3 に示す異なる条件下で認証精度を評価するために, 上記 5 種類について別日にそれぞれ 100 フレーム分の画像を取得した.

4.2.2　顔画像データの拡張

取得した顔画像に対し， openCV を用いて 112 × 112 にリサイズを行い， 表 4.1 に示す各パラメータについて $6 + 7 + 2 = 15$ 種類のデータに拡張した． 元画像を含め， 一人一種類当たり 8,000 枚， 合計 48,000 枚の顔画像データを作成した．さらに， 画像の顔の位置によって識別さることを防ぐため， 全ての画像をランダムな位置から 96 × 96 に切り出しを行った．

表 4.1: データ拡張

手法	コントラスト調整	輝度変換	ガウシアンノイズ
種類	12% 23% 35% 47% 59% 70%	0.5 0.7 0.9 1.1 1.3 1.5 1.7	2 4
計	6	7	2

4.2.3 CNN の構成

本節は numpy を用いて，VGG-11 を参考に作成し，Dropout を追加した．また，ランダムに学習データを 38400 枚，テストデータを 9600 枚に分け，各種パラメータは表 4.2 に示す値とする．学習した CNN に対し，web カメラでリアルタイムに取得した顔画像データを渡すことで結果を画面に表示する顔認証システムを実装した．実装画面を図 4.1 に示す．

表 4.2: CNN の構成表

パラメータ名	値
モデル	vgg11+Dropout
入力画素	96 × 96
バッチサイズ	300
学習係数	0.001
Optimizer	Adam
エポック数	2
出力層	5

4.2.4 学習と評価が同じ組み合わせの評価手法

上記の計 5 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．学習データで学習させた CNN について，同一条件の評価データでテストする．

4.2.5 学習と評価が異なる組み合わせの評価手法

上記の計 5 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．

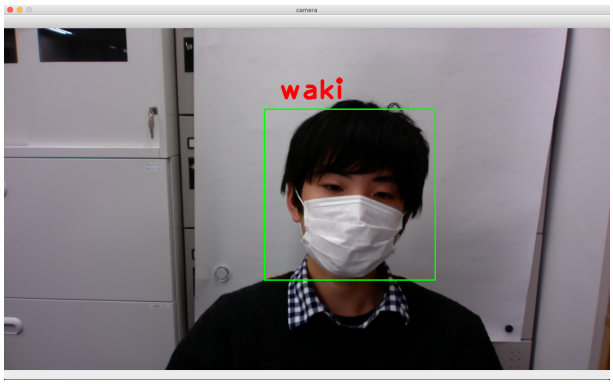


図 4.1: 顔認証システム

4.3 実験結果

4.3.1 素顔で学習した CNN に対する追跡停止の評価

学習させた CNN に対し異なる条件下で評価した精度を表 4.3 に示す．素顔を学習させた場合，素顔の評価については 60.6%，マスクなどの外乱は 20%〜40% の再現率となった．また，被験者 5 名の 100 フレームの画像に対し，横軸を素顔，縦軸をマスクにした際に被験者毎の出力値の平均をプロットした結果を図 4.2 に示す．被験者 C を除いて，負の相関関係がみられた．

4.3.2 外乱を加えた画像で学習した時の評価

表 4.3 より，それぞれの外乱について以下に述べる，

- 帽子を学習... 素顔とマスク＋サングラスに関して 50% 前後の再現率となった．
- マスクを学習... 同じマスクでの評価は 99.4% となったが，一方でそれ以外については 30% 以下の再現率となった．また，図 1 より，被験者 B を除いて，素顔の出力値がほぼ 0% に対し，マスクはほぼ正しく出力されている．
- サングラスを学習... 同じサングラスでの評価は 94.4% となった．さらに，素顔，マスク，サングラスについても 50% 以上の再現率となった．
- マスク＋サングラスを学習... 同じマスク＋サングラスでの評価は 78.6% となり，それ以上にマスクが 84.2% の再現率となった．一方で，サングラスは 22.0% の再現率であった．

表 4.3: 学習させた CNN に対し異なる条件下で評価した場合の再現率

train \ test	素顔	帽子	マスク	サングラス	マスク + サングラス	平均
素顔	60.6	21.4	28.8	37.0	21.2	33.8
帽子	51.4	74.8	20.0	48.8	20.0	43.0
マスク	20.2	20.0	99.4	20.0	30.0	37.9
サングラス	53.0	20.6	50.4	94.4	53.4	54.8
マスク + サングラス	25.2	20.0	84.2	22.0	78.6	46.0
平均	42.1	31.8	56.6	44.4	40.6	43.1

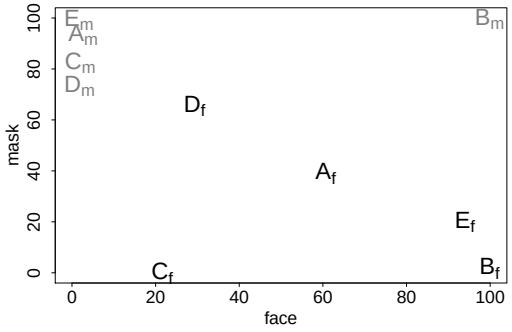


図 4.2: 認証精度の評価 (f-素顔,m-マスク)

第 5 章

頭髪による顔認証精度の影響

5.1　実験目的

顔取得システムでは，Viola-Jones 法による顔検出手法を利用して実装した．取得した顔画像を図 5.1 に示す．被験者により個体差はあるが，取得画像に占める頭髪の割合は大きいものであり，画像から学習する CNN に対して，頭髪は学習結果にどれほど影響があるのか検証する．

5.2　実験内容

被験者 20 名，データセットは学習データ 200 枚，評価データ 100 枚とし，以下の実験を行う．

1. 頭髪を含めた学習データで CNN の学習を行い，顔認証の結果が一致しない被験者において，学習データと評価データの頭髪の違いを調査する．次に，前髪を手で上げた状態でリアルタイム顔認証システムを利用して評価を行う．
2. 学習データの画像の上部 80 ピクセル分トリミングを行い，1 と同様に前髪を挙げた状態で評価を行う．

図 5.1: 顔画像の取得

5.3　実験結果

1. 学習データの顔画像とリアルタイム顔認証システムの実行画面を図 5.2 に示す．頭髪を含めた学習データの CNN で評価した場合，学習データ時の髪形と評価時の髪形が大きく変化することで，結果が異なることが明らかになった．次に前髪を上げた状態で評価した結果を図 5.3 に示す．前髪を上げることで，正しく認証していた被験者に対しても異なる被験者を示す結果となった．
2. 学習データをトリミングした場合の学習データを図 5.4 に示す．また，トリミング対応したリアルタイム顔認証システムの実行画面を図 5.5 に示す．被験者によって正誤に違いがでる結果となったが，頭髪を含む場合とは異なる被験者を指した．

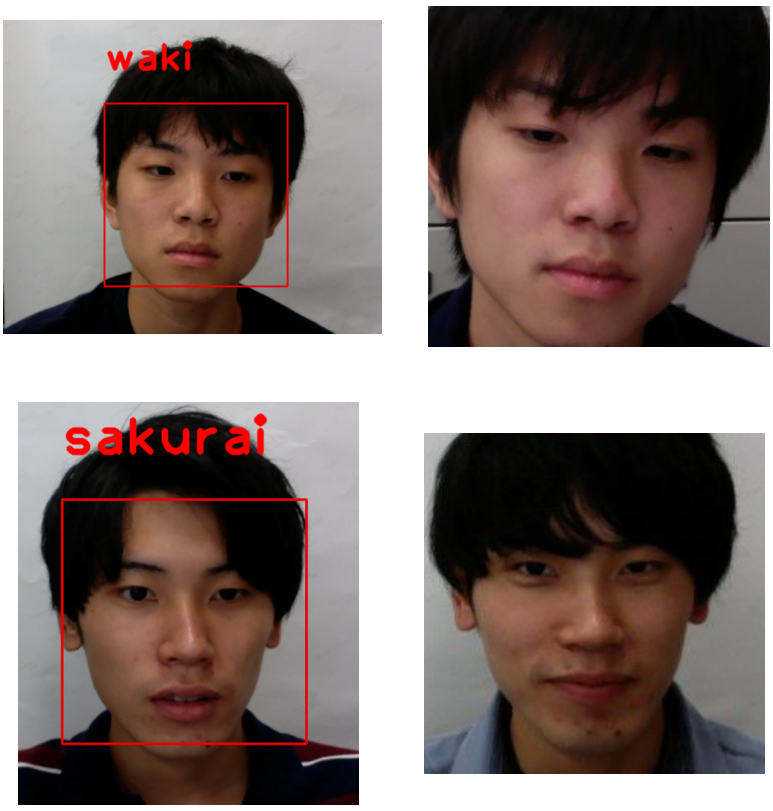


図 5.2: 学習データの顔画像 (右) と顔認証システム (左)

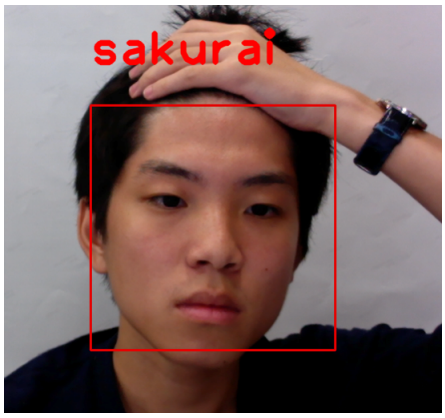


図 5.3: 前髪を上げた状態（認証結果は誤り）

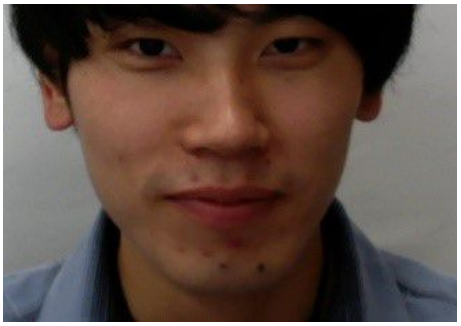


図 5.4: 頭髪をトリミングした学習データ

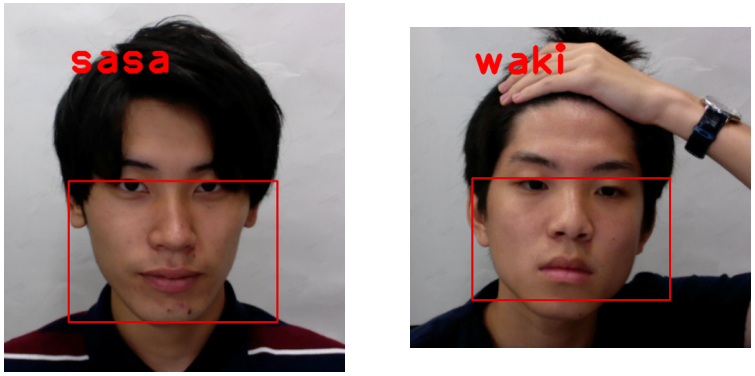


図 5.5: トリミング時の顔認証システム（左図の被験者は誤りだが，右図の被験者は正しい）

第 6 章

考察

6.1　プライバシーバイザーについて

6.1.1　素顔を学習した場合

表 3.3 より，素顔を学習し，素顔で評価した場合の適合率が最も高くなった原因として，顔全体の露見している部分に対して各自の細かな特徴を学習したのではないだろうか．一方，バイザーで評価した場合は，後述で述べるレンズの特性により，適合率が低下したと考える．

6.1.2　バイザーを学習した場合

表 3.3 より，同じバイザーで評価した場合，認証精度が 42% 程となったのに対し，[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 6.1: 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

6.1.3　ユーザ毎の変動評価

図 3.7 によると，バイザー ON を評価データとした場合．素顔の精度は 0% に対し，バイザーの精度は向上した．これは，CNN が素顔を学習した場合に目元の特徴で被験者を判断している一方で，バイザーを学習すると目元以外の特徴で判断しているためと分析できる．さらに，被験者によって素顔の評価データに，バイ

ザーへの汎用性の違いがみられた原因として，目元の特徴量に依存しない被験者はバイザーを装備しても識別されてしまうと考える．すなわち，CNN は被験者それぞれに対して，その被験者を識別する顔の特徴的な部位を学習し，判断していると考える．

6.1.4　エポック毎の学習遷移

図 10 の全ての学習遷移において，学習データの学習率が 2eoch 目で 100% となり，過学習の可能性が指摘される．これは，本研究ではデータ拡張を行っていることから，図 6 に示すように類似の画像が 1 エポック内で繰り返し学習されるため，学習率が指数関数的に増加したと考える．

6.2　日用品について

表 4.3 より，外乱を与えた場合，素顔時よりも精度が向上した原因として，バイザー時と同様に顔の一部が隠れているため，表情などの変化に頑強であり，露見している部分に対して各自の細かな特徴を学習したのではないかと考えられる．一方で，帽子であれば被り方，マスクやサングラスであれば付け方といったように外乱によって被験者を識別しているとも考える．

顔がほぼ隠れているマスク＋サングラスについて，マスクに関しても 84.2% の再現率が得られたことから，マスクを学習した場合と同じ特徴を学習していると考えられる．同じマスク＋サングラスにおいても 78.6% の再現率を得ていることから，マスクが識別する上で重要な特徴を生み出していると考える．

表 4.3 より，サングラスで学習した場合に 5 種類の評価の平均が 54.8% と最も高い数値を得た理由として，まばたきや眼球の動きなど顔の表情の中で最も変化が激しい部位である目をサングラスで隠すことによって，被験者毎の本質的な特徴を学習したため，帽子を除く外乱に対して 50% 程度の再現率が得られたと考える．

6.3　頭髪の影響について

頭髪が顔認証システムの結果に影響を与えた原因として，顔画像において最も流動的かつ特徴を持った部位であるからといえるだろう．頭髪は前髪以外にも，髪色，髪の量，耳がどれほど隠れているか，もみあげの有無など多くの情報を含んでいる．そして，これらは日毎に変わりやすい部分でもあるため，学習データの精度は 100% であっても正しく評価出来ない結果になったと考える．一方で，トリミングしたにも関わらず正しい結果とならない被験者がいた原因として，学習データを撮影した時間帯による明るさなどの外的要因によるものだと考える．

第 7 章

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．

本研究では，静止画像を評価データとして用いた．これは，2019 年 1 月現在で利用されている顔認証システムの認証方式が，固定されたカメラに目線を合わせ，数秒顔を固定するものであったため，この方式に倣い本研究においても採用した．しかし，評価の段階で適合率が 0% に近い被験者と 100% に近い被験者の二極化されてしまい，本研究で実装したリアルタイム顔認証システムの精度が個人間で大きな差となった．そのため，静止の顔画像を取得する際は 1 枚毎に一定の期間を空ける分散撮影をすることが求められるだろう．

頭髪は顔認証システムに影響を与えることが明らかとなった．これに対し，頭髪のトリミングはその影響を抑えるものとして一定の効果があることが示された．しかし，外部環境によってはトリミングの効果が低減されることになる．従って，流動的な部位や状況を固定化させることで，CNN の汎用性をより生かした顔認証システムの実装が可能であるだろう．

謝辞

本研究に際して, 様々なご指導をいただきました菊池浩明教授に深く感謝します. また，顔認証システムの実装に関して，顔画像の提供にご協力いただいた菊池研究室の皆様に厚く御礼申し上げます．最後に, 本論文を著するにあたり，参考画像として掲載に同意していただいた池上和輝様，清水崇喜様に感謝の意を表すると共に, 謝辞にかえさせていただきます.

参考文献

[1] 朝日新聞，“大阪駅ビル顔で追跡”，2014 年 1 月 6 日 朝刊 pp.1”

[2] Face Recognition Study – FAQ，available from (http://www.heinz.cmu.edu/ acquisti/face-recognition-study-FAQ/).

[3] 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.

[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.

[5] Viola，P. and Jones，M.: Rapid object detection using a boosted cascade of simple features，Proc. Computer vision and Pattern Recognition 2001 (CVPR 2001)，pp.I-511-I-518(2001).

[6] Shakhnarovich，G.，Viola，P．and Moghaddam，B.: A unified learning framework for real time face detection and classification，Proc．Automatic Face and Gesture Recognition 2002 (FG2002) (2002).

[7] Viola，P. and Jones，M.: Robust Real-Time Face Detection，International Journal of Computer Vision(IJCV)，Vol.57，No.2，pp.134-157(2004).

[8] Karen Simonyan and AndrewZissserman(2014)，“Very Deep Convolutional Networks for Large-Scale Image Recognition”，pp.1409-1556， ICLR，2014.

[9] 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.

研究業績

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.