

プライバシーが保護された訓練画像に対する 視線位置分布を用いた性別認識*

井上路子** 西山正志** 岩井儀雄**

Gender Classification using Gaze Distributions for Privacy-Protection of Training Samples

Michiko INOUE, Masashi NISHIYAMA and Yoshio IWAI

We propose a method for classifying gender using training samples after applying privacy-protection. Recently, training samples containing individuals require to protect their privacy. Head regions of training sample are usually manipulated for privacy-protection. However, the accuracy of gender classification is degraded when directly using the protected training samples. Here, we aim to use the human visual capability that people can correctly recognize males and females though the head regions are manipulated. We use gaze distributions of observers who view stimulus images for the preprocessing of gender classifier. Experimental results show that our method improved the accuracy of gender classification after manipulating the training samples by masking, pixelation and blur for privacy-protection.

Key words: Gender classification, training samples, gaze distributions, privacy-protection

1. はじめに

公共空間に設置された防犯カメラで撮影された人物画像から性別を認識する技術の開発が進んでいる。この技術に応用することで、その場に存在する人物の性別を把握し、マーケティングや防犯へ貢献することが期待されている。性別認識の既存手法¹⁾²⁾では、深層学習を適用することで高い精度を実現している。これらの手法では深層学習の効果を発揮するため訓練画像を大量に収集している。

訓練画像を収集する際に、人間である被写体を撮影することに注意しなければならない。撮影された訓練画像には、被写体が収集されることに不安を感じる情報が含まれている場合がある。特に顔は取り扱いに気を付けなければならない。さらに訓練画像は、学習モデルの精度を高めるため、長期にわたり繰り返し利用されている。したがって、訓練画像の収集にあたり、被写体のプライバシーを保護する技術が必要とされている。

画像中の被写体のプライバシーを保護するために、その頭部領域を加工することが一般的に多い。例えば、雑誌の広告ではマスク加工が施され、テレビのインタビューではモザイク加工が施され、google ストリートビューでは被写体の顔にぼかし加工が施されている。さらに、複雑な加工を施す手法が文献³⁾⁴⁾で提案されている。ただし、これらの文献では、性別認識における訓練画像へのプライバシー保護を十分に議論していなかった。既存手法を訓練画像にそのまま適用した場合、性別認識の精度に影響することが想定される。実際に文献⁵⁾では、プライバシーを保護するために訓練画像中の頭部領域にマスクやモザイクやぼかしを適用した場合、性別認識の精度が低下することが報告されている。

プライバシーが保護された訓練画像を用いて性別認識の精度を高めるために、本論文では人間の視覚特性に着目する。文

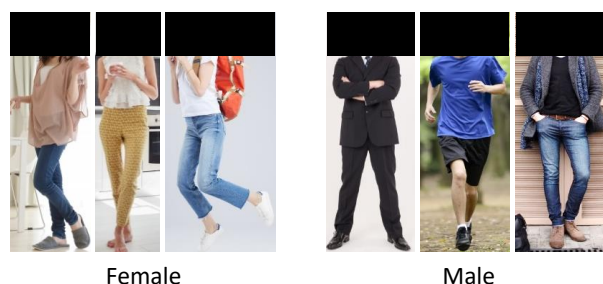


Fig.1 Examples of pedestrian images where head regions are completely masked for privacy protection

献⁶⁾では、人間は性別を見分ける際に重要な特徴を画像中の被写体から観察することが述べられている。この手法では、実験協力者が被写体を観察する際に視線位置の分布を計測し識別に利用する。具体的には、特徴抽出の前処理として、視線が集まる頭部領域を識別器が重視するように訓練画像へ重みを与えている。既存手法⁶⁾をプライバシー保護された訓練画像にそのまま適用すると、認識精度が低くなると考えられる。これは、訓練画像へプライバシー保護を施す際、前述のように頭部領域は加工されることが多いためである。ただし、被写体の頭部領域を完全に遮蔽した時でも、多くの場合において人間は見た目の性別を判断することができる。これは、体型などの頭部以外から得られる特徴を画像中の被写体から観察するためだと考えられる。例えば、図1のように、被写体の胴体が曲線の丸みを帯びている場合は女性、胴体が筋肉質な場合は男性であると人間は判断することが多い。なお、本論文における胴体とは、身体の頭、首、四肢を除いた部位を指す。

そこで本論文では、プライバシー保護のために訓練画像中の頭部領域が加工された場合でも性別認識の精度を向上させる手法を提案する。そのために、以下の2つの項目を検証する。

- 被写体の頭部領域が遮蔽された場合における実験協力者の視線位置分布

* 原稿受付 令和元年5月6日

* 掲載決定 令和元年9月4日

** 鳥取大学大学院工学研究科(鳥取市湖山町南4丁目101)

- 計測された視線位置分布を識別器に利用することの効果

1つめの検証では、実験協力者が性別を判別する際に、頭部以外で観察する領域を確認する。また、比較のために頭部領域が遮蔽されていない場合の実験も同時に行う。2つめの検証では、計測された視線位置分布をプライバシーが保護された訓練画像に適用することで、性別認識の精度が向上するかどうかを確認する。

2. 関連研究

2.1 プライバシー保護

文献⁷⁾⁸⁾⁹⁾で述べられているように、画像中に存在する被写体のプライバシー保護は重要な課題であり、近年注目が高まっている。被写体の画像を撮影する前の段階や、画像を用いて認識処理を行う段階など、処理の全体を通して様々な既存手法が提案されている。文献¹⁰⁾では、被写体自身が保護メガネをかけることで、一般的な検出アルゴリズムから顔が見つからないようにする手法が提案されている。また、文献¹¹⁾では、カメラの撮像系に特殊な機構を組込むことで、被写体の頭部を自動で遮蔽する手法が提案されている。これらの既存手法を用いることでプライバシーが保護された画像のみが記録される利点がある。ただし、撮影時に特別な機材を準備する必要があるため、本論文で対象とする訓練画像の収集時に、これらの既存手法が適しているとは言えない。文献³⁾では、被写体の顔を架空の人物の顔と取り換えることで、プライバシーを保護する手法が提案されている。新しいプライバシー保護の形と言えるが、社会的に十分に受け入れられているとは言えないため、本論文では一般的な画像加工を対象とする。文献⁴⁾では、架空人物の画像に本人の特徴量を視覚的に分からないよう埋め込むことで、認識処理を行う手法が提案されている。認識アルゴリズムが今後変更されないのであれば有効であるが、新たに開発された認識アルゴリズムを適用すると埋め込まれた特徴量は復元されなくなる。このため本論文では、様々な認識アルゴリズムに対応できるように前処理の手法を取り扱うことを考える。

2.2 視覚特性

画像認識の分野において、人間の視覚特性の活用が期待されている。文献¹²⁾では、視線位置分布を用いて顕著性マップの推定精度を向上させている。文献¹³⁾では、ウェアラブルカメラを用いて注目領域推定と行動認識を同時に行っている。文献¹⁴⁾では、視線位置と画像特徴を用いて選好推定を行っている。文献¹⁵⁾では、商品リコメンデーションのために視線の移動パターンを推定している。文献¹⁶⁾では、視線位置分布のみから物体認識を行っている。さらに文献¹⁷⁾では、視線による注目領域マスキングで顔画像の属性を認識している。また文献¹⁸⁾では、視線位置分布を深層学習に組み込むことでファッション画像の属性を認識している。これらの既存手法は様々な視線の応用を議論しているものの、プライバシー保護を取り扱っていないため、本論文は性別認識を対象とし新たに取り組むこととする。

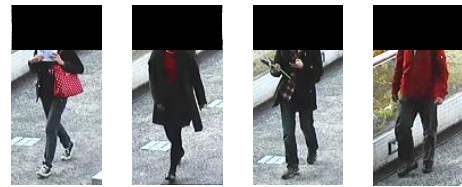
3. 被写体の頭部領域が遮蔽された場合における 実験協力者の視線位置分布

3.1 刺激画像

実験協力者が被写体の性別を見分ける際、頭部以外で観察する領域を明らかにする検証を行った。そのため本論文では、画



Fig.2 Examples of pedestrian images after applying privacy protection



(a) S_1



(b) S_2

Fig.3 Examples of stimulus images for measuring gaze distributions

像中の被写体の頭部領域を実験協力者が観察できないように、頭部およびその周辺領域を完全に遮蔽した。その例を図2(a)に示す。他の方法として、例えば図2(b)のように、頭部領域のみを遮蔽することが考えられる。ただしこの場合、頭部領域の周囲に矩形の強いエッジが発生するため、実験協力者の視線が人物領域だけでなくエッジ付近にも集まることが予想される。図2(c)や(d)のように、頭部及びその周辺領域にモザイクやぼかしを施すことも考えられる。ただし、被写体の頭部を観察できるため、実験協力者の視線が頭部領域にも集まることが予想される。従って本実験では、頭部領域及びその周辺領域を完全に遮蔽する方法を採用した。以下では、実験協力者が観察する人物画像を刺激画像と呼ぶ。

実験協力者の視線位置分布を計測する際、刺激画像に適用する画像加工の条件を以下に挙げる。

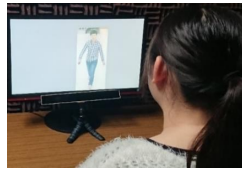
S_1 : 頭部領域を遮蔽した場合

S_2 : 頭部領域を遮蔽しない場合

比較のために、条件 S_1 に加え S_2 の計測も行うこととした。条件 S_1 で使用する刺激画像の例を図3(a)に、条件 S_2 で使用する刺激画像の例を図3(b)に示す。刺激画像としてPETAデータセット¹⁹⁾に含まれるCUHKデータセットを使用した。このデータセットに含まれる画像の人物領域は、人手によってある程度の位置合わせがなされている。その程度を調査するために、CUHKデータセットに含まれる全ての人物画像から条件 S_1 と S_2 の平均画像を生成した。条件 S_1 の平均画像を図4(a)に、 S_2 の平均画像を(b)に示す。先に図中の(b)について考察する。画像上部に現れた黒い円が人物の頭部領域、画像の中央付近に現れた黒い楕円が人物の胴体領域、画像下部に現れた淡い黒色の部分が人物の足領域に対応していると言える。図4(a)では、(b)の画像上部に現れた黒い円が観察されて



Fig.4 Average images with and without masking head regions



(a)

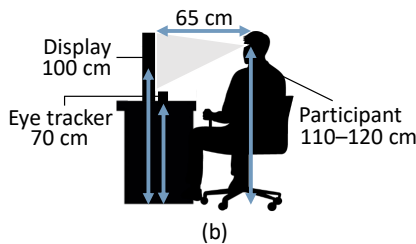


Fig.5 Experimental setting

いないことから、人物の頭部領域が完全に遮蔽されていると言える。なお、条件 S_1 では画像上部から高さ 25% の領域を遮蔽した。この決め方として、被写体の頭部が完全に遮蔽され、かつ、肩以下の胴体が大きく遮蔽されないこととした。遮蔽された刺激画像がこの決め方を全て満たすことを目視で確認した。刺激画像中の遮蔽領域の画素値はディスプレイ背景の画素値と同じとした。刺激画像のサイズは条件 S_1 と S_2 ともに 80×160 画素とした。実験条件を統制するため、刺激画像に含まれる被写体の男女の割合を同じにした。また、画像中の被写体の体向きの割合を前後左右で同じにした。被写体に同一人物は出現しないようにした。刺激画像の枚数は条件 S_1 で 32 枚、 S_2 で 32 枚とした。

実験協力者数を 16 名 (男性 10 名、女性 6 名) とした。平均年齢は 22.4 ± 1.0 歳であった。視線計測の実験風景の例を図 5 (a) に、その配置を (b) に示す。ディスプレイから 65 cm の位置に実験協力者を座らせた。目の高さが 110 から 120 cm になるよう椅子の高さを実験協力者に調節させた。ディスプレイの大きさは 24 インチ (解像度: 1920×1080 画素)、視線計測装置は Gaze Point GP3 とした。視線計測装置のサンプリングレートは 60 fps とした。角度分解能は 0.5 度から 1.0 度の間であると仕様書に記載されていた。刺激画像を 480×960 画素に拡大して提示した。また、センターバイアスを避けるためにディスプレイ上のランダムな位置に刺激画像を提示した。条件 S_1 の刺激画像が提示されたディスプレイの例を図 6 (a) に、条件 S_2 の刺激画像が提示されたディスプレイの例を (b) に示す。実験協力者 16 名のうち 8 名には条件 S_1 、残りの 8 名には条件 S_2 の刺激画像を提示した。条件 S_1 では男性 4 名と女性 4 名の計 8 名、条件 S_2 では男性 6 名と女性 2 名の計 8 名とした。

3.2 実験手順

実験協力者の視線位置分布を計測する手順は以下とした。



Fig.6 Examples of stimulus images on the display

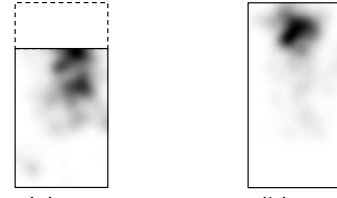


Fig.7 Gaze maps generated under the conditions S_1 and S_2

- P_1 : 実験協力者を 16 名とした。
- P_2 : 実験協力者を 1 名ランダムに選び出した。
- P_3 : 視線位置分布を計測する時の条件を S_1 または S_2 のどちらかにランダムに設定した。
- P_4 : 刺激画像中の被写体の性別を見分けるタスクを実行するように実験協力者に説明した。
- P_5 : 灰色画像を 2 秒間提示した。
- P_6 : ランダムに選択された刺激画像を 2 秒間提示した。
- P_7 : 黒色画像を 3 秒間提示し、実験協力者に口頭で被写体の性別を回答させた。
- P_8 : 手順 P_5 から P_7 を刺激画像の枚数分繰り返した。
- P_9 : 全ての実験協力者が終了するまで手順 P_2 から P_8 を繰り返した。

次に、視線マップの生成手法について述べる。ここでは、手順 P_6 で計測された視線位置分布から、文献⁶⁾と同様の手法で視線マップを生成した。視線計測装置が停留していると判別した実験協力者の視線位置のみを用いた。条件 S_1, S_2 毎に、停留点の出現位置の頻度分布を全ての実験協力者、および、全ての刺激画像から求めた。視線マップの大きさは刺激画像と同じとなるようリサイズした。

3.3 計測された視線位置分布の解析

条件毎の視線マップの生成結果を図 7 に示す。以下では条件 $i \in \{S_1, S_2\}$ の視線マップを g_i で表す。図中において、視線マップの黒色領域は、実験協力者からの視線位置が集まることを表している。逆に白色領域は視線位置が集まらないことを表している。図 7 (a) の視線マップ g_{S_1} を図 4 (a) の平均画像中の人物シルエットと照らし合わせると、遮蔽がある場合に実験協力者の視線位置は被写体の胴体領域に集まることが分かった。各刺激画像中の人物領域を観察すると、被写体の胴体の体型や服装や持ち物が出現していた。一方、図 7 (b) の視線マップ g_{S_2} から、遮蔽がない場合に実験協力者は刺激画像中の被写体の頭部領域を重点的に観察することが分かった。これは既存手法⁶⁾と同様の傾向であった。条件 S_1 と S_2 ともに、被写体の足付近には実験協力者の視線が集まらなかった。

次に、個々の刺激画像に対する視線マップを確認した。図 8 に、図 3 のそれぞれの刺激画像に対する視線マップを示す。例えば、図 3 の右上の刺激画像を観察した時の全実験協力者の視

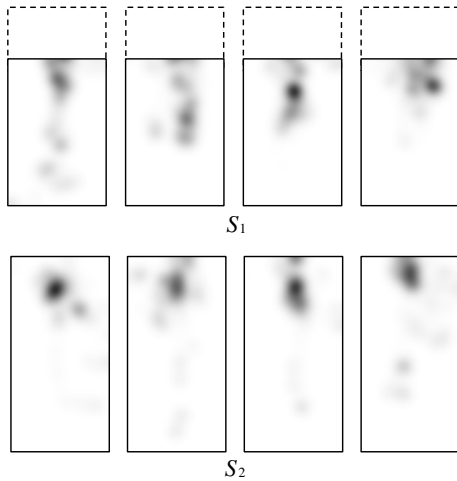
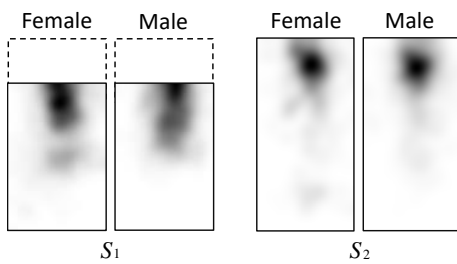
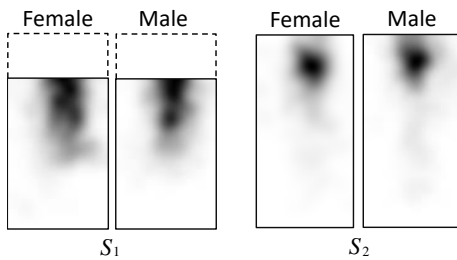


Fig.8 Examples of the gaze maps of the stimulus images



(a) Participants



(b) Subjects

Fig.9 Gaze maps measured from males and females of participants or subjects

視線位置分布が、図 8 の右上の視線マップとなる。条件 S_1 では、多くの刺激画像において実験協力者の視線は被写体の胴体へ主に集まっていたが、一部の刺激画像において足や腰にも視線が集まっていた。条件 S_2 では、全ての刺激画像において実験協力者の視線は頭部へ集まっていた。

実験協力者による性別認識の正解率について述べる。実験協力者の回答が刺激画像の性別ラベルと一致していれば正解とした。実験協力者の正解率は、条件 S_1 で $87.0 \pm 5.5\%$ 、条件 S_2 で $95.9 \pm 3.1\%$ であった。

計測時の実験協力者、及び、画像中の被写体の性差による視線位置分布の違いを検証した。実験協力者の性別毎に生成した視線マップを図 9 (a) に、被写体の性別毎に生成した視線マップを (b) に示す。図中の (a) では、条件 S_1 と S_2 とともに、視線が集まる領域には実験協力者の性別による顕著な差は見られなかった。図中の (b) では、(a) と同様に、被写体の性別による顕著な差は見られなかった。条件 S_1 と S_2 における実験協力者、及び、被写体の性別間の相関係数を、ピアソンの積率相関係数により算出した。その結果を表 1 に示す。全ての条件で性別間の相関係数は高かった。

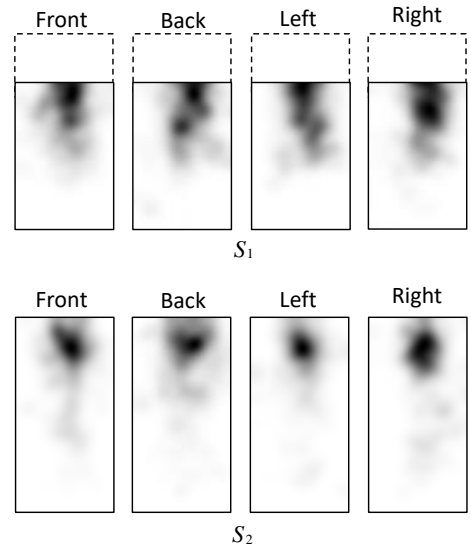


Fig.10 Gaze maps measured from the subjects of each direction

Table1 Correlation coefficients of gaze maps measured from males and females of participants or subjects.

Participants		Subjects	
S_1	S_2	S_1	S_2
0.88	0.97	0.97	0.92

刺激画像中の被写体の体向きの影響で、視線マップに差が見られるかどうかを調査した。体向き毎の視線マップを図 10 に示す。ピアソンの積率相関係数により、体向きの間で相関係数を算出した。その結果、条件 S_1 では 0.91 ± 0.03 、 S_2 では 0.93 ± 0.03 となった。どちらの条件でも相関係数が高いため、被写体の体向きの影響で、視線が集まる領域に顕著な差は見られないと考えられる。

4. 計測された視線位置分布を識別器に利用することの効果

4.1 実験条件

本実験では視線位置分布を用いることで、被写体の頭部領域が遮蔽された訓練画像で性別認識の精度が向上するかどうかを検証した。訓練画像には刺激画像と同様に PETA データセットに含まれる CUHK データセットを用いた。ただし視線計測時の刺激画像は、訓練画像とテスト画像に含まれないようにした。性別認識の精度を求める際、総数 2540 枚の画像に 10-fold 交差検証を適用した。交差検証の各回では、訓練画像の枚数を 2286 枚 (男性 1143 枚, 女性 1143 枚)、テスト画像の枚数を 254 枚 (男性 127 枚, 女性 127 枚) とした。訓練画像に対する加工の条件を以下とした。

- T_{S_1} : 訓練画像中の頭部領域を遮蔽した場合
- T_{S_2} : 訓練画像中の頭部領域を遮蔽しない場合

条件 T_{S_1} の訓練画像の例を図 11 (a) に、条件 T_{S_2} の訓練画像の例を (b) に示す。

訓練画像に対する視線マップの効果を確認するために、以下の条件で精度を比較した。

- G1 : 遮蔽あり視線マップ g_{S_1} を用いた提案手法



Fig.11 Examples of training samples with and without masking head regions

- G2 : 遮蔽なし視線マップ g_{S_2} を用いた既存手法
 G3 : 視線マップを用いないベースライン手法

精度を比較する条件の数は、 T_{S_1} と T_{S_2} に対し、G1 と G2 と G3 を組み合わせた 6 通りとした。訓練画像とテスト画像から視線を毎回計測することは避け、3. の刺激画像から計測された視線マップのみを用いることとした。視線マップを識別器の前処理として利用する手法は、文献⁶⁾と同様とした。以下でその概要を説明する。実験協力者の視線位置が集まる位置の画素値に、訓練画像上で相対的に大きくなるように重みを与えた。一方、実験協力者の視線位置が集まらない位置の画素値に、小さな重みを与えた。具体的には以下の手続きとした。視線マップの位置 (x, y) における値を $g_i(x, y)$ で表す。 $g_i(x, y)$ が取り得る値の範囲は $[0, 1]$ とした。訓練画像の画素値を $t(x, y)$ とすると、前処理後の画素値 $t'(x, y)$ を式 (1) で算出した。

$$t'(x, y) = c(g_i(x, y))t(x, y) \quad (1)$$

ここで $c(\cdot)$ は視線マップの値を強調する補正関数を表し $c(z) = z^a + b$ とした。a, b はハイパーパラメータである。a>1 の時は視線マップの値が強調され、a<1 の時は値が緩和される。b は視線マップのオフセットである。なお、RGB 色空間を HSV 色空間に変換し、V 値のみに重みを与えた。

視線マップにより重み付けされた訓練画像に対し、以下の識別器を適用した。

- **CNN** : Convolutional Neural Network (CNN) を用いた。ネットワークモデルの前段には畳み込み層を 2 層としそれぞれの間にプーリング層を設けた。後段には全結合層を 2 層を設けた。このモデルは文献²⁰⁾の Mini-CNN と同様である。
- **SVM** : 線形の Support vector machine²¹⁾ を用いた。ペナルティパラメータ C は 1 とした。
- **LR** : 文献²²⁾のロジスティクス回帰を用いた。正則化パラメータは 1 とした。

なお、テスト画像には遮蔽を適用しないこととし、訓練画像と同じ手続きで視線マップによる重み付けを行った。遮蔽なしの理由として、性別認識を実行する端末内で、テスト画像のプライバシーが問題になることが少ないことが挙げられる。これは、端末内でテスト画像を識別器に入力し、その場で画像を破

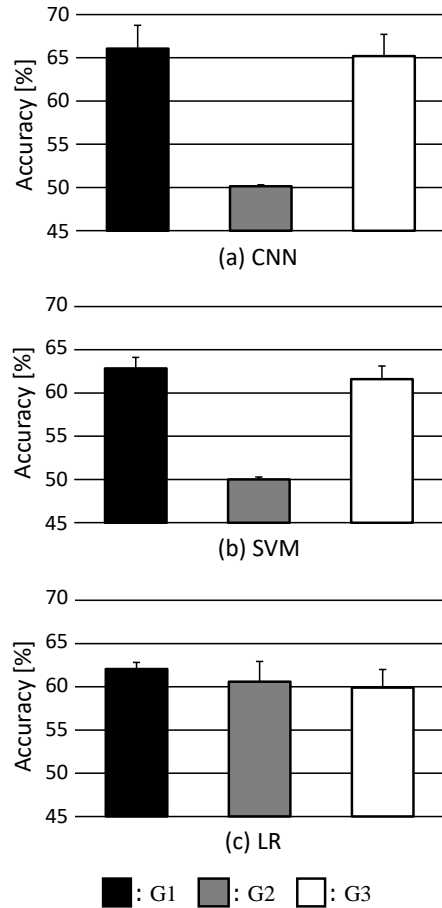


Fig.12 Accuracy of gender classification under masking head regions of the condition T_{S_1}

棄する仕様が多いためである。よって本実験では、テスト画像としてプライバシー保護の画像加工が施される前の状態である遮蔽なし人物画像を用いて精度検証を行った。

4.2 訓練画像が遮蔽された場合の実験結果

頭部を遮蔽した条件 T_{S_1} の性別認識の結果を図 12 に示す。図中の (a) は CNN, (b) は SVM, (c) は LR を用いた時の性別認識の精度である。図中の (a) では、条件 G1 は G3 より精度がわずかに高かった。また、条件 G2 は G1 と G3 より精度が大幅に低下した。(b) では (a) と同様の傾向がみられた。(c) では、条件 G2 は G3 より精度が高く、条件 G1 は G2 と G3 より精度が高かった。これらの結果より、訓練画像の頭部領域を遮蔽した場合、視線マップは遮蔽なしと比べて遮蔽ありの方が有効であることが分かった。

次に頭部を遮蔽しない条件 T_{S_2} の性別認識の結果を図 13 に示す。(a) では、条件 G1 と G3 の精度に顕著な差はなかった。また、条件 G2 は G1 と G3 より精度が高かった。(b) では、条件 G1 は G2 と G3 より精度が低かった。条件 G2 は G3 より精度が高かった。(c) では、(b) と同様の傾向がみられた。これらの結果より、訓練画像の頭部領域を遮蔽しない場合、視線マップは遮蔽ありと比べて遮蔽なしの方が有効であることが分かった。訓練画像への遮蔽の有無によって性別認識に有効な視線マップが異なることから、それぞれ適した視線マップを切り替えて用いることが必要と考えられる。

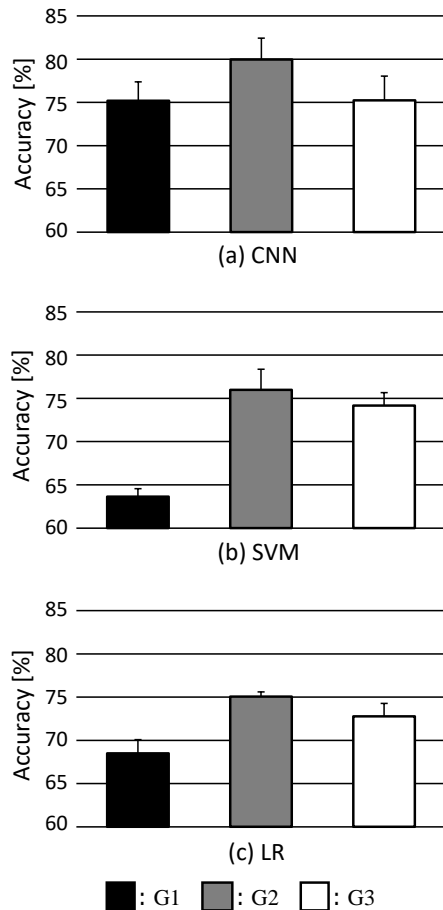


Fig.13 Accuracy of gender classification under *not* masking head regions of the condition T_{S_2}

4.3 被写体の頭部領域にモザイクやぼかしを施した場合

訓練画像中の被写体のプライバシーを保護するために、遮蔽以外に用いられる加工方法を適用した際の性別認識の精度を検証した。ここではモザイクとぼかしを採用した。加工前の訓練画像の例を図 14 (a) に、モザイクを施した訓練画像の例を図 (b) に、ぼかしを施した訓練画像の例を図 (c) に示す。遮蔽と異なり、モザイクやぼかしはいくつかの加工程度を持つ。検証では、訓練画像に 4 つの加工程度 $k \in \{1, 2, 3, 4\}$ を施した。モザイクの加工程度を条件 T_{p_k} 、ぼかしの加工程度を条件 T_{b_k} で表す。それぞれの条件における加工程度はモザイクで $\{16 \times 32, 12 \times 24, 8 \times 16, 4 \times 8\}$ ブロックとした。ぼかしの加工程度はガウシアンブラーを用い、その σ を $\{3, 7, 11, 15\}$ とした。識別器には 4.2 節で最も精度の高かった CNN を用いた。視線マップには 4.1 節の G1 と G2 と G3 を用いた。

プライバシー保護としてモザイクを施した性別認識の結果を図 15 (a) に示す。条件 T_{p_1} から T_{p_3} では、G2, G3, G1 の順に精度が高かった。一方、条件 T_{p_4} では G1, G3, G2 の順に精度が高かった。これらの結果より、モザイクの加工程度が大きい訓練画像を用いた場合、遮蔽あり視線マップを利用する手法が最も精度が良くなるのが分かった。

次に、プライバシー保護としてぼかしを施した性別認識の結果を図 15 (b) に示す。条件 T_{b_1} から T_{b_4} において、G1, G2, G3 の順に精度が高かった。これらの結果より、ぼかしの加工程度が小さい訓練画像を用いた場合でも、遮蔽あり視線マップを利用することで性別認識の精度が高まる場合があることが分

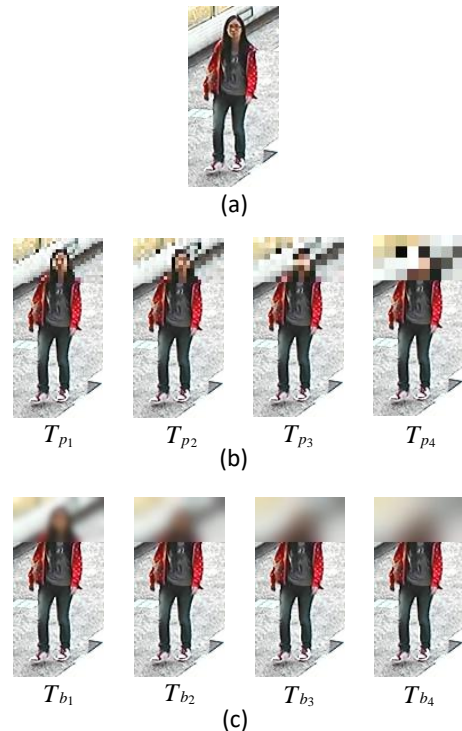


Fig.14 Examples of training samples after protecting privacy using pixelation and blur while changing parameters

かった。

4.4 人物領域を用いた場合との比較

視線マップによる重み付けではなく、人物領域そのものを用いた場合と性別認識の精度を比較した。人物領域の抽出には Mask R-CNN²³⁾ を用いた。その訓練画像の例を図 16 に示す。遮蔽なしの訓練画像に対し Mask R-CNN を適用し、その上部を遮蔽した。性別認識の精度を求める際、総数 2540 枚の画像に 10-fold 交差検証を適用した。交差検証の各回では、訓練画像の枚数を 2286 枚 (男性 1143 枚, 女性 1143 枚), テスト画像の枚数を 254 枚 (男性 127 枚, 女性 127 枚) とした。識別器には 4.2 節で最も精度の高かった CNN を用いた。

性別認識の精度は、視線マップ gs_1 を用いる提案手法では $66.1 \pm 2.7\%$ 、人物領域を用いる比較手法では $52.4 \pm 2.5\%$ であった。提案手法は比較手法より精度が高かった。遮蔽あり視線マップで訓練画像に重みを与える提案手法は、人物領域を用いる比較手法より有効であると言える。

4.5 訓練画像およびテスト画像の頭部領域を遮蔽した場合

訓練画像とテスト画像において被写体の頭部領域を遮蔽した場合について精度を評価した。総数 2540 枚の画像に 10-fold 交差検証を適用した。交差検証の各回では、訓練画像の枚数を 2286 枚 (男性 1143 枚, 女性 1143 枚), テスト画像の枚数を 254 枚 (男性 127 枚, 女性 127 枚) とした。視線マップには 4.1 節の G1 と G2 と G3 を用いた。識別器には CNN と SVM と LR を用いた。

訓練画像およびテスト画像の頭部領域を遮蔽した場合の精度を図 17 に示す。図中の (a) は CNN, (b) は SVM, (c) は LR を用いた時の精度である。(a) では、条件 G1 は G2 より精度が高かった。条件 G1 は G3 より精度が僅かに高かった。一方、条件 G2 は G3 より精度が僅かに低かった。(b) では、条件 G1 は G2 より精度が僅かに高かった。条件 G1 は G3 より精度が高

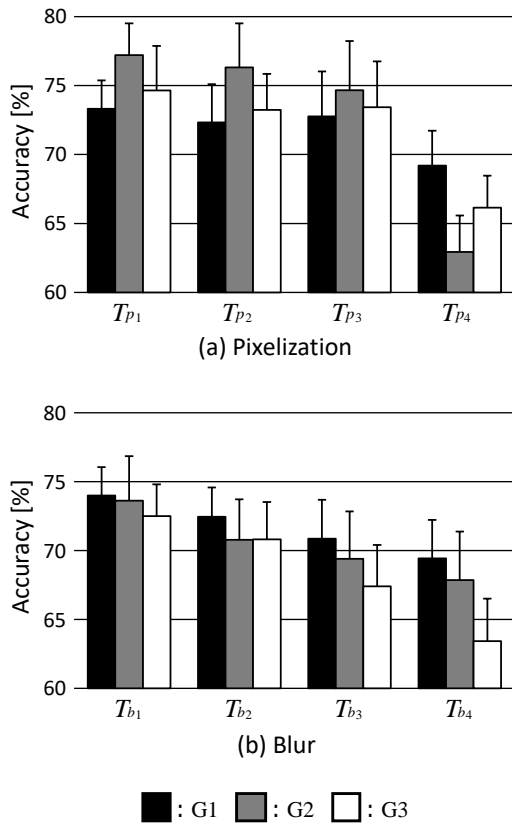


Fig.15 Accuracy of gender classification using training samples of pixelation and blur



Fig.16 Examples of training samples after applying body region segmentation

かった. 条件 G2 は G3 より精度が僅かに高かった. (c) では (b) と同様の傾向がみられた. これらの結果より, 提案手法 G1 では既存手法 G2 と比べて, 僅かではあるが精度向上が期待できると考えられる. また, 提案手法 G1 はベースライン G3 と比べて精度向上に有効であると言える.

5. おわりに

被写体のプライバシーが保護された訓練画像を用いた性別認識において, 実験協力者の視線位置分布を利用することで精度を向上させる手法について議論した. 視線位置分布の計測では, 被写体の頭部領域が遮蔽された刺激画像を実験協力者に提示した. 視線計測の結果, 被写体の胴体領域を実験協力者は観察することが明らかになった. 得られた視線位置分布を利用し, 遮蔽やモザイクやぼかしの加工でプライバシーが保護された訓練画像を用いて性別認識の実験を行い, 提案手法の有効性を確認した.

今後の課題として, 提案手法と既存手法の視線位置分布を組み合わせた性別認識の手法の開発や, 性別認識以外の属性タスクにおける遮蔽あり視線マップの有効性に関する検証, 訓練画像が遮蔽なしでテスト画像が遮蔽ありの場合の検討, 遮蔽あり

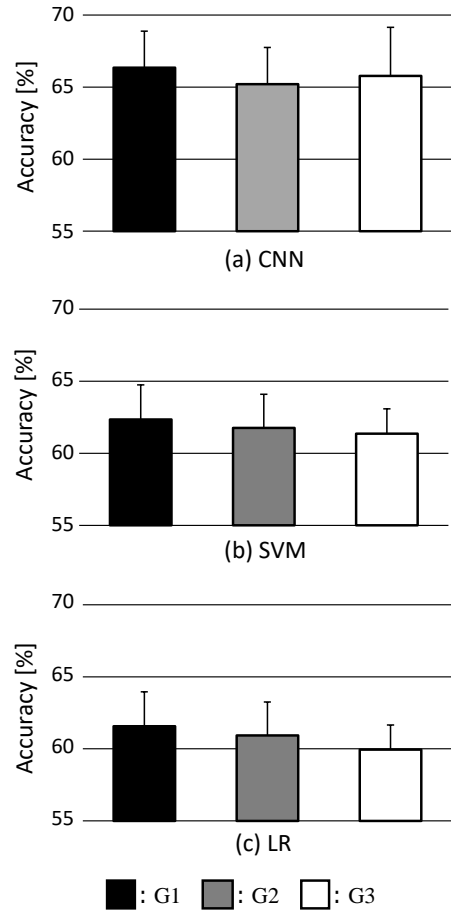


Fig.17 Accuracy of gender classification under masking head regions of training samples and test samples

視線マップの更なる有効活用の検討, また, 停留点の決め方の影響調査が挙げられる.

謝 辞

本研究の一部は, 総務省 SCOPE(No. 172308003) および JSPS 科研費 JP17K00238 の助成を受けたものである.

参 考 文 献

- 1) P. Sudowe, H. Spitzer, and B. LeibeSudowe. Person attribute recognition with a jointly-trained holistic cnn model. In *Proceedings of IEEE International Conference on Computer Vision Workshop*, (2015), 329.
- 2) Y. Lu, A. Kumar, S. Zhai, Y. Cheng, T. Javidi, and R. Feris. Fully-adaptive feature sharing in multi-task networks with applications in person attribute classification. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2017), 1131.
- 3) S. Ribaric, A. Ariyaeinia, and N. Pavesic. De-identification for privacy protection in multimedia content. *Signal Processing: Image Communication*, **47**, (2016), 131.
- 4) S. J. Oh, M. Fritz, and B. Schiele. Adversarial image perturbation for privacy protection a game theory perspective. In *Proceedings of IEEE International Conference on Computer Vision*, (2017), 1491.
- 5) N. Ruchaud, G. Antipov, P. Korshunov, J. L. Dugelay, T. Ebrahimi, and S. A. Berrani. The impact of privacy protection filters on gender recognition. *Applications of Digital Image Processing XXXVIII*, (2015), 959906.
- 6) M. Nishiyama, R. Matsumoto, H. Yoshimura, and Y. Iwai. Extracting discriminative features using task-oriented gaze maps measured from observers for personal attribute classification. *Pattern Recognition Letters*, **112**, (2018), 241.
- 7) O. S. Joon, B. Rodrigo, F. Mario, and S. Bernt. Faceless person recognition: Privacy implications in social media. In *Proceedings of European Conference on Computer Vision*, (2016), 19.
- 8) F. Flammini, R. Setola, and G. Franceschetti. Effective surveillance

- for homeland security: Balancing technology and social issues. *CRC Press*, (2013).
- 9) P. Campisi. Security and privacy in biometrics: Towards a holistic approach. *Security and Privacy in Biometrics*, Springer, (2013), 1.
 - 10) T. Yamada, S. Gohshi, and I. Echizen. Privacy visor: Method for preventing face image detection by using differences in human and device sensitivity. In *Proceedings of International Conference on Communications and Multimedia Security*, (2013), 152.
 - 11) Y. Zhang, Y. Lu, H. Nagahara, and R. Taniguchi. Anonymous camera for privacy protection. In *Proceedings of 22nd International Conference on Pattern Recognition*, (2014), 4170.
 - 12) M. Xu, Y. Ren, and Z. Wang. Learning to predict saliency on face images. In *Proceedings of IEEE International Conference on Computer Vision*, (2015), 3907.
 - 13) A. Fathi, Y. Li, and J.M. Rehg. Learning to recognize daily actions using gaze. In *Proceedings of 12th European Conference on Computer Vision*, (2012), 314.
 - 14) Y. Sugano, Y. Ozaki, H. Kasai, K. Ogaki, and Y. Sato. Image preference estimation with a data-driven approach: A comparative study between gaze and image features. *Eye Movement Research*, **7**, 3, (2014), 862.
 - 15) Q. Zhao, S. Chang, F. M. Harper, and Joseph J. A. Konstan. Gaze prediction for recommender systems. In *Proceedings of 10th ACM Conference on Recommender Systems*, (2016), 131.
 - 16) N. Kaessli, Z. Akata, B. Schiele, and A. Bulling. Gaze embeddings for zero-shot image classification. In *Proceedings of IEEE conference on computer vision and pattern recognition*, (2017), 4525.
 - 17) N. Murrugarra-Llerena and A. Kovashka. Learning attributes from human gaze. In *Proceedings of IEEE Winter Conference on Applications of Computer Vision*, (2017), 510.
 - 18) H. Sattar, A. Bulling, and M. Fritz. Predicting the category and attributes of visual search targets using deep gaze pooling. In *Proceedings of IEEE International Conference on Computer Vision Workshops*, (2017), 2740.
 - 19) D. Yubin, L. Ping, L. C. Change, and T. Xiaoou. Pedestrian attribute recognition at far distance. In *Proceedings of 22nd ACM International Conference on Multimedia*, (2014), 789.
 - 20) A. Grigory, B. Sid-Ahmed, R. Natacha, and D. Jean-Luc. Learned vs. hand-crafted features for pedestrian gender recognition. In *Proceedings of 23rd ACM International Conference on Multimedia*, (2015), 1263.
 - 21) C. Corinna and V. Vladimir. Support-vector networks. *Machine Learning*, **20**, 3, (1995), 273.
 - 22) D. R. Cox. The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)*, **20**, 2, (1958), 215.
 - 23) H. Kaiming, G. Georgia, D. Piotr, and G. Ross. Mask r-cnn. In *Proceedings of IEEE International Conference on Computer Vision*, (2017), 2980.