（様式１１）

博士学位論文審査結果要旨

西暦 2023年 3月 1日

研究科，専攻名　　バイオ・情報メディア研究科　コンピュータサイエンス専攻

学位申請者氏名　　張　重陽

論　文　題　目　　Personal Photo Privacy Protection by Attacking Face Detectors

審査結果の要旨

2023年3月1日に東京工科大学内において学位申請者 張 重陽 氏の学位審査公開発表会が開催され，博士学位論文に関する発表とその質疑応答が行われた．

本論文では，悪意のある顔画像の収集や加工(本人に無断でWeb上の顔画像を取得し，それを用いてコラージュ画像を作り出して拡散させるなど)を防止する手法について検討を行っている．現在，代表的な顔検出手法(顔検出器)としてMTCNN(Multi-Task Convolutional Neural Network)，SSD(Single Shot multibox Detector)，S3FD(Single Shot Scale-invariant Face Detector)の3手法が広く知られているが，これらの手法すべてで顔検出が行えないようにするための手法(つまり，顔検出器への攻撃方法)を確立することが本研究の目的である．

以下に本論文の内容を述べる．

第１章は序論である．本章ではまず悪質な顔画像加工の被害事例を紹介している．そして，このような行為の防止がいかに重要であるのかについて言及し，本研究の目的を明確化している．続いて，代表的な顔面交換手法(顔部分だけを別の写真で交換したコラージュ画像を生成する手法)を概説し，このような顔面交換手法がうまく機能しないようにするためには顔検出自体を機能させないことが最も効率的であることについて言及している．

続く第2章では，本研究に関連する様々な技術を紹介している．具体的には，代表的な顔検出手法であるMTCNN，SSD，S3FDについて詳細に述べるとともに，敵対的攻撃の原理についても言及している．

第3章，第4章，第5章は本論文の中心をなす章である．

まず第3章では，顔検出器の攻撃手法に関する先行研究を踏まえ，画像中の有効な攻撃範囲を明確化することを目的としている．この課題に取り組むにあたり，第一ステップとしてStyle Transferと呼ばれる技術により顔の特徴を抽出し，背景との合成を行う．続いて第二ステップとして得られた合成画像に対して顔検出器(MTCNN)により顔領域の検出を行う．この結果，Style Transferを用いて背景の検出確率を高めても，顔検出器の顔検出結果には影響しないことを明らかにした．このことは，画像中の顔を検出できないようにするためには背景に処理を加えるだけでは不十分で，顔領域自体に処理を加える必要があることを示している．

続く第4章では，顔検出器の中で最も性能が優れているMTCNNを対象とした攻撃方法について検討している．MTCNNでは，SSDやS3FDとは異なり画像ピラミッド処理が組み込まれている．この処理は，様々な画像サイズの顔を検出するために有効な処理である．画像ピラミッド処理では画像のサイズを比例的に縮小させるため，この過程で画素の数と値の両方が変化する．つまり攻撃側から見れば，摂動情報が変化してしまうために敵対サンプルの攻撃力を低下

させることになってしまう.そこで,これに対応するために，摂動情報のみを補間し，異なるサイズの摂動情報を融合させることにより，摂動の劣化を抑え，MTCNNに対する攻撃を実現している.

さらに，第5章では，MTCNN，SSD，S3FDの顔検出器すべてを攻撃するために，顔に黒線を追加する方法を提案している．黒線構造は，MTCNNで検出した目，口角，境界ボックスの頂点の座標を結んで実現する．評価実験の結果，黒線構造は顔検出器の攻撃に有効であることが判明している．しかしながら，黒線構造は黒線で顔の一部を覆うことになるため画像の使い勝手が著しく悪くなってしまう．そこで，いかに細い黒線を少なく入れるのかについて詳細に検討するとともに，黒線が引かれるべき画像領域に黒線ではなくランダム摂動を挿入する手法を提案している．さらに，ランダム摂動の挿入方法について遺伝的アルゴリズムに基づく最適化手法を提案し，評価実験を実施している.

最後の第6章は本論文のまとめである．以上述べてきた成果をまとめるとともに今後の課題について言及している.

以上まとめると，本博士学位論文は，悪意のある顔画像の収集や加工を防止する手法について検討を行ったものであり,これを実現するために主要な3種の顔検出器に対する攻撃方法について詳細検討を行ったものである.特に,画像中の有効な攻撃範囲の明確化(第3章)，顔検出器の中でも最も性能が優れているMTCNNを対象とした攻撃方法の提案(第4章),3種の顔検出器すべてを攻撃するための黒線構造ならびに黒線構造領域へのランダム摂動の埋め込み方法の提案(第5章)については顕著な学術的成果を含む内容であり,即座の実用化にはまだ課題が残るとは言え,当該分野での学術的価値は高いものである.

以上のように，本学位論文は，コンピュータサイエンス分野の学術発展に寄与する部分が少なくない．また，学位審査公開発表会後に実施された筆記試験では，専門基礎，専門，外国語(英語)に関する基本的な知識が問われたが，この試験に対する回答も高い得点が得られていることから，博士として十分な学力を有しているものと判定される.

以上の理由から，審査委員会は 張 重陽 氏 に対して博士(コンピュータサイエンス)の学位を授与するのに十分な学識と能力を有しているものと認める.

審査委員　主査

東京工科大学　教授　青木　輝勝