様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 １８ 日現在

機関番号：２１６０２
研究種目：基盤研究(C)（一般）
研究期間：2020～2022
課題番号：２０Ｋ１１８１３

研究課題名（和文）イメージセンサノイズを用いたディープフェイク検出技術の確立

研究課題名（英文）Development of Deepfake Detection Technologies based on Image Sensor Noise

研究代表者
富岡　洋一（Tomioka, Yoichi）

会津大学・コンピュータ理工学部・上級准教授

研究者番号：１０５７４０７２

交付決定額（研究期間全体）：（直接経費）　　2,900,000円

研究成果の概要（和文）：本研究では、画像中のパターンノイズ間の距離を計測するディープメトリックラーニングモデルを提案し、画像の圧縮率の違いに頑健なパターンノイズ間の距離の推定を実現した。加えて、パターンノイズがディープフェイク検出において有効な特徴の一つとなることを確認した。更に高精度なディープフェイク検出を実現するために、顔の各パーツから複合的な特徴を抽出する畳み込みニューラルネットワークモデルを提案し、入力画像から検出できた複数部位に対応する畳み込みニューラルネットワークモデルを組み合わせたアンサンブルモデルにより、高精度なディープフェイク検出を実現できることを示した。

研究成果の学術的意義や社会的意義
近年、ディープラーニングや機械学習技術に発展に伴い，ある人物の顔を別の人物に重ねて加工する「ディープフェイク」といった改ざん技術が進歩しており、フェイクニュースの拡散が危惧されている。フェイクニュースは個人のプライバシーを侵害するだけではなく、政治経済への多大な影響を及ぼすことも考えられるため、いち早くフェイクニュースを検出し、注意喚起することが必要である。本研究の成果はディープフェイク検出の精度向上やマスク等で顔の一部が隠蔽されている場合のディープフェイク検出に有効な基盤技術であり、より安心・安全な社会の実現に貢献できると期待している。

研究成果の概要（英文）：We proposed a deep metric learning model that measures the distance between pattern noises in images which is robust to different compression ratios of the images. In addition, we confirmed that the pattern noise feature is one of the effective features for deepfake detection. To further improve the accuracy of deepfake detection, we proposed convolutional neural network models that extract composite features from each part of the face. We showed that an ensemble model combining deepfake detectors specialized for each face part that can be detected from the input image can realize highly accurate deep-fake detection.

研究分野：画像認識、画像科学捜査

キーワード：ディープフェイク　パターンノイズ　畳み込みニューラルネットワーク　アンサンブルモデル

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１。研究開始当初の背景

近年、ディープラーニングや機械学習技術による画像認識、合成の技術が急速に発展しており、このような技術を静止画や動画の改ざんに利用することが危惧されている。例えば、「ディープフェイク」は画像中の顔の領域を抽出し、他の人物の顔と置き換えることが可能である。このような技術を活用してフェイク動画を作成し、インターネット上に偽情報を拡散することが容易になってきている。偽情報は、個人のプライバシーを侵害するだけではなく、政治経済に影響を及ぼす恐れもあるため、動画の真正性をどのように保証していくかが重要な課題となっている。

ディジタル画像中には Photo Response Non-Uniformity (PRNU)ノイズ、Non Unique Artifact (NUA)と呼ばれるイメージセンサのパターンノイズが存在する。PRNU ノイズは撮像素子の感度特性のばらつきにより撮影画像に生じるノイズである。NUA はイメージセンサデザイン等に起因して生じる周期的なアーティファクトである。また、これらのパターンノイズに加えて、画像撮影時の標本化に起因するアーティファクトとして、Periodic Interpolation Artifact (PIA)が存在する。画像の拡大・縮小・回転時に補間処理を行うことで PIA も同様に変形して現れることが知られている。これらのパターンノイズはディープフェイクを検出するための重要な手がかりのひとつとなることが期待される。

２。研究の目的

本研究では、偽情報の早期発見と拡散防止に貢献することを目指し、ディープフェイク検出におけるパターンノイズの活用可能性を探ると共に、ディジタル画像・映像の改ざんの有無の自動検出を実現することを目的とする。

３。研究の方法

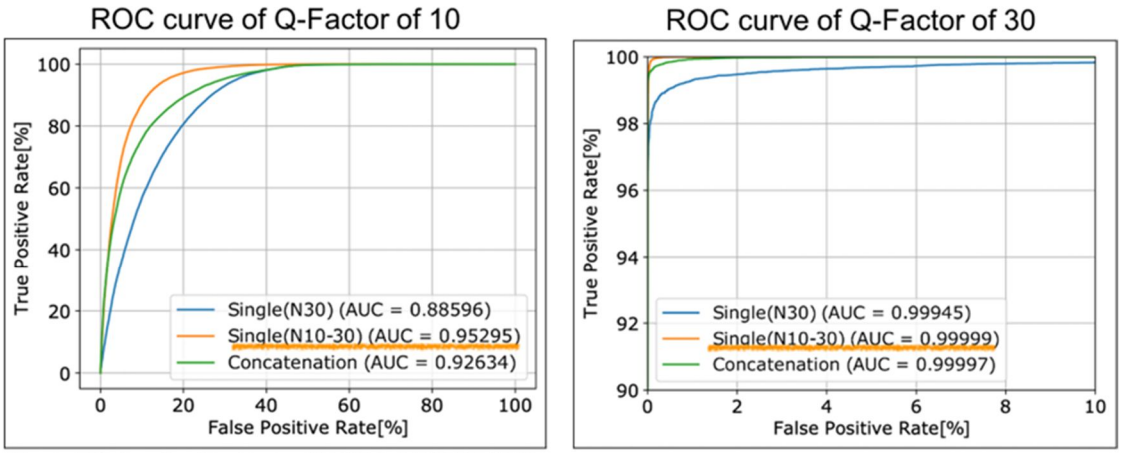
パターンノイズをディープフェイク検出に活用するためには、画像中からパターンノイズ特徴を適切に抽出することが必要となる。イメージセンサのパターンノイズは微弱な信号であり、画像圧縮によりさらにノイズが汚染されることでパターンノイズ特徴の抽出が難しくなることが知られている。良く用いられている画像圧縮の一種である JPEG 圧縮は画像の圧縮率を Quality Factor で調整することができ、Quality Factor の低下とともにパターンノイズ情報も劣化してしまう。同機種のカメラで撮影した画像であれば、同じ NUA パターンノイズ特徴を抽出できるようにすべきであるが、既存の学習方法では同機種の画像であっても異なる Quality Factor で JPEG 圧縮された 2 対の画像からは、異なる特徴を抽出してしまう傾向がある。本研究ではこの問題を緩和するため、異なる Quality Factor であっても同機種であれば同じ特徴を、異なる機種であれば別の特徴が抽出できるように畳み込みニューラルネットワークの学習方法を改良した。Quality Factor の違いに頑健な畳み込みニューラルネットワークを学習するためには、様々な Quality Factor の JPEG 画像から抽出した大量のノイズが必要となるが、機種ごとに Quality Factor の異なる大量のノイズを準備することは多大な時間と労力を要する。そこで、各ノイズからより低い Quality Factor の JPEG 圧縮に対応する疑似ノイズを作成し、畳み込みニューラルネットワークの学習に用いた。また、トリプレットロスによるディープメトリックラーニング技術に基づき、同機種間であれば特徴間の距離が近く、異なる機種間であれば特徴間の距離が遠くなるように畳み込みニューラルネットワークを学習し、パターンノイズ間の相違度を測定する手法を実現した。

次に、パターンノイズ間の相違度を測定する手法をディープフェイク動画と実動画に適用し、ディープフェイク検出器を構築、評価した。この結果、パターンノイズはディープフェイク検出のための特徴として有効ではあるが、パターンノイズだけでは十分な精度でディープフェイクを検出することが困難であることが判明した。そこで、ディープフェイク検出の特徴として、パターンノイズだけではなく、ディープフェイク画像の不自然さなどの見た目の特徴も合わせて用いることにした。また、新型コロナウィルスの拡大により、マスクを着用することも多くなり、顔の一部が隠れていても本物の画像と偽物の画像を正しく認識できることも重要と考えられる。そこで、本研究では、目、角膜、鼻、口といった顔パーツに着目し、顔パーツ毎にディープフェイク検出の可能性について調査を行った。また、入力画像から検出できた顔パーツに特化したディープフェイク検出モデルで構成されるアンサンブルモデルにより高精度にディープフェイクを検出する技術を確立した。

４。研究成果

提案のパターンノイズ間の距離を測定するディープメトリックラーニングモデルを用いて、与えられた一組のノイズ画像を同機種と異なる機種の 2 クラスに分類したときの Receiver

Operating Characteristic (ROC)曲線を図 1 に示す。Single(N10-N30)は Quality Factor が 10%から 30%のときのノイズを学習したモデルであり、Quality Factor が 30%のノイズだけを学習したモデルに対して、Area Under Curve (AUC)を改善できており、提案手法により Quality Factor の違いに対してより頑健な特徴が抽出できていることがわかる。



Single(N30)：Quality Factorが30%のノイズだけを学習
Single(N10-30)：Quality Factor10%から30%のノイズを組み合わせて学習
Concatenation：各Quality Factorのノイズを学習したモデルを作成し、抽出した特徴を接続した特徴を使用

図 1 与えられた一組のノイズ画像を同機種と異なる機種の 2 クラスに分類する問題に対する ROC 曲線

図 2 に、各顔パーツに特化したディープフェイク検出器とそれらを組み合わせたアンサンブルモデルの F1-score と実行時間の評価を示す。各顔パーツの検出器として、ResNet-18 を採用した。図 2 に示すように、目(E)、角膜(C)、鼻(N)、口(M)だけでもディープフェイク検出に有効な特徴を含んでいることを確認できた。また、各顔パーツに特化したディープフェイク検出器を組み合わせることで顔全体を使用した場合と同程度の高精度なディープフェイク検出を実現できた。

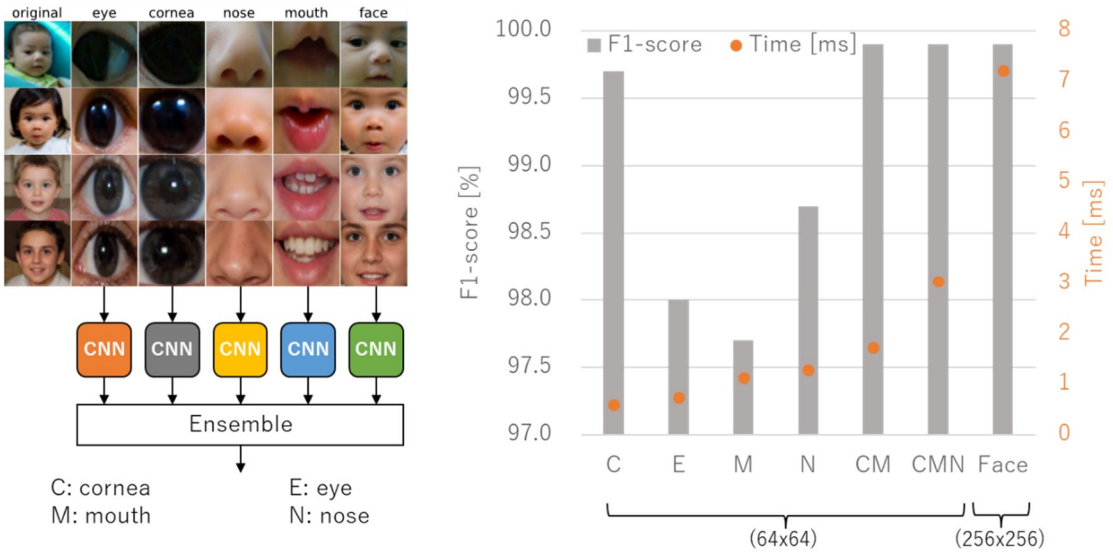


図 2 各顔パーツに特化したディープフェイク検出器とそれらを組み合わせたアンサンブルモデルの F1-score と実行時間の評価 (図中の画像は FFHQ データセットと StyleGAN2 の画像を使用)

５．主な発表論文等

〔雑誌論文〕　計0件

〔学会発表〕　計3件（うち招待講演　0件／うち国際学会　2件）

１．発表者名 Mai Uchida and Yoichi Tomioka
２．発表標題 CNN-based Camera Model Classification and Metric Learning Robust to JPEG Noise
３．学会等名 International Conference on Awareness Science and Technology（国際学会）
４．発表年 2020年

１．発表者名 内田麻衣，富岡洋一
２．発表標題 畳み込みニューラルネットワークを用いたカメラモデル分類におけるJPEG圧縮の影響の評価及び解析
３．学会等名 電子情報通信学会パターン認識・メディア理解研究会
４．発表年 2020年

１．発表者名 AkihisKawabe, Ryuto Haga, Yoichi Tomioka, Yuichi Okuyama and Jungpil Shin
２．発表標題 Fake Image Detection Using An Ensemble of CNN Models Specialized For Individual Face Parts
３．学会等名 IEEE International Symposium on Embedded Multicore/Many-core Systems-on-Chip（国際学会）
４．発表年 2022年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関