

メンタルチェックを用いたフェイク動画の判別

—フェイク動画と実際の人間の特性比較—

日大生産工 塩見碧海 山下皓平 水上祐治

1. まえがき

Abbas and Taeihagh (2024)は、生成系AIの発展により、顔置換・音声合成を含むディープフェイクの品質は年々向上しているとしている[1]。また、Yang and Lyn (2019)によると、従来の検出法は、ピクセルレベルの不整合やフレーム間の位相ズレ、顔形状やポーズの破綻など視覚的アーティファクト(空間・時系列の不連続性)に着目してきたが、生成品質の向上とともに検出は困難化しているとしている[2]。そのため、それら画像等の真偽を判別する技術(ディープフェイク検出技術)が重要視されている。本研究は、ディープフェイク検出技術として、映像の「見た目」ではなく、被写体の微小振動という半生理的・運動学的特徴に着目し、非接触で画像解析を行うものである。

2. 先行研究レビュー

Ciftci and Demir (2019)によると、非接触で微細な揺れや生体由来信号(rPPGなど)を捉える手法は、偽動画では空間・時間的保存性が乏しいことを示している[3]。rPPG(リモート脈波: remote photoplethysmography)とは、心拍に伴うごく微小な色変化・揺れ(約1Hz前後)が、額・頬など顔の複数部位で同じリズムで現れることである。つまり、rPPGを測定することにより、本物画像では、例えば、額・左右頬で似た波形となり、リズムが時間的に安定する。また、位相差も自然であり、空間的にも整合的である。一方、フェイク動画では、部位ごとにリズムがズれるなど、空間的不整合がみられるなどの特徴がある。表1にrPPGによ

表1 rPPGによる画像の判定の目安

本物動画	心拍帯域のピークが明瞭、領域間の相関が高く、位相差の分散が小さい。高周波比は低め。
フェイク動画	ピークが不明瞭／揺れる、領域間相関が低め、位相差がバラつく。高周波比が相対的に高い。

Ciftci and Demir (2019)を元に筆者作成

る画像の判定の目安を示す。

3. 分析フレームワーク

3.1. 分析手順

- Step1: ディープフェイクを作成可能なプログラムを入手する。
- Step2: オリジナルとなる映像を取得する。
- Step3: 合成に用いる偽物の顔写真を入手する。
- Step4: Step 2で取得したオリジナル映像を、「Mental-Checker」を用いて解析する。
- Step5: Step1のプログラムとStep3の顔写真を用いて、ディープフェイク映像を作成する。
- Step6: Step4と同様に、作成したディープフェイク映像(Step5)を「Mental-Checker」で解析する。
- Step7: Step4とStep6の解析結果を比較し、どの項目に違いが見られるかを確認する。
- Step8: この工程を複数回繰り返して法則性を見つけ出し、フェイク動画を判別する手法を確立する。

3.2. 分析環境

本物動画： 実際の映像。(30名分)

フェイク動画： 本物の映像にフェイク動画の編集をしたもの(30名分)

実験装置： Mental-Checker

3.3. 分析結果

メンタルチェックを用いてフェイク動画



図1 実験装置の使用例

Discriminating Deepfake Videos with a “Mental Checker”
— A Comparison of Deepfake and Real-Human Characteristics —

Aoi SHIOMI Kohei YAMASHITA and Yuji MIZUKAMI

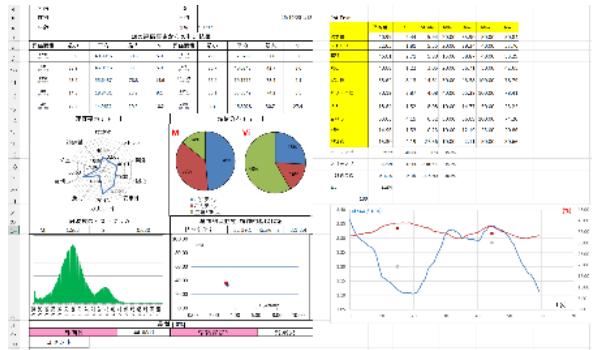


図2 本物動画の分析結果

とリアルな動画を分析し比較した。その結果、今回比較したものすべてに同じ違いがみられた。今回のリアルの映像の分析結果（画像2）とフェイク動画（画像3）を比較した際にフェイク動画の周波数のヒストグラムがすべてリアルの映像のものより右にずれていたことが分かった。このことからフェイク動画の映像の方が頭の振動が大きく精神的に興奮状態になっていることがわかった。

4. まとめ

今回のメンタルチェッカーを用いてフェイク動画とリアルな動画を分析し比較した。その分析結果から、すべてのフェイク動画の映像の方が頭の振動が大きく精神的に興奮状態になっていることがわかった。

この結果は、ディープフェイク動画が単に「生理学的データがない」のではなく、「不自然に過剰なデータ」を含んでいることを示唆しており、以下の2つの重要な可能性が導き出される。1つ目はディープフェイク特有の「合成ノイズ」を興奮と誤認である。フェイク動画は、元の映像に別の顔画像を張り合わせる際、完全に静止させることができないため、微細な映像ノイズ（ちらつき、ピクセルの歪み、画像フレームの不規則なズレ）が常に発生している。

そのため「メンタルチェッカー」がこの不自然な「ノイズ」や「合成処理による微細な揺れ」を、人間が興奮しているときに発生させる「生理的な振動」と誤認し、興奮状態と判定した可能性が最も高いと判断できる。よって、検出された「大きな振動」は、偽造過程で発生した「合成されたデータ」の証拠である。

2つ目は元の映像の「動き」の強調である。ディープフェイクは、元の人物の動画（リアル）にターゲットの顔を貼り付ける。このリアルの動画が、たまたま極端にブレたり、動きの激しいものであった場合、その不自然な動きのパターンがそのまま出力される。



図3 フェイク動画の分析結果

そのため、処理の過程で、不自然な動きを隠すために、あえてフレームレートや動きの振幅を過剰に「滑らかに」したり、「大きく」見せたりする処理がされている可能性がある。この過剰な「滑らかさ」や「ブレ」の補正が、振動分析において興奮状態と判定されるパターンに合致した。このことから、「大きな振動」は、フェイク動画が持つ「不必要的動的な情報（不自然なブレや動きの補正）」を露呈しており、実際の動画で見られるような「感情と一致した振動の抑制」が機能していない証拠である。

このことから、フェイク動画の「大きな振動と興奮状態」のデータは、「偽造処理によって導入されたデジタルノイズや不自然な動きの強調」を、メンタルチェッカーが「生きている人間の生理的な興奮」と誤認した結果である可能性が高い。この「視覚情報と生理学的振動データの決定的な不一致」は、動画が人間ではなく合成されたデータであることを示す重要な科学的証拠となることがわかる。

参考文献

- 1) Abbas, F., & Taeihagh, A. (2024). Unmasking deepfakes: A systematic review of deepfake detection and generation techniques using artificial intelligence. *Expert Systems with Applications*, 252(Part B), 124260.
- 2) Yang, X., Li, Y., & Lyu, S. (2019). Exposing deep fakes using inconsistent head poses. In 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8261–8265). IEEE
- 3) Ciftci, U. A., & Demir, I. (2019). FakeCatcher: Detection of synthetic portrait videos using biological signals.