

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2020-525947

(P2020-525947A)

(43) 公表日 令和2年8月27日(2020.8.27)

(51) Int.Cl.		F I			テーマコード (参考)
G06T 7/00	(2017.01)	G06T 7/00	350B		5L096
G06N 3/08	(2006.01)	G06N 3/08			
G06N 20/10	(2019.01)	G06N 20/10			

審査請求 未請求 予備審査請求 未請求 (全 27 頁)

(21) 出願番号 特願2019-572463 (P2019-572463)
(86) (22) 出願日 平成30年6月29日 (2018.6.29)
(85) 翻訳文提出日 令和2年2月26日 (2020.2.26)
(86) 国際出願番号 PCT/EP2018/067692
(87) 国際公開番号 W02019/002602
(87) 国際公開日 平成31年1月3日 (2019.1.3)
(31) 優先権主張番号 1710560.2
(32) 優先日 平成29年6月30日 (2017.6.30)
(33) 優先権主張国・地域又は機関
英国 (GB)

(71) 出願人 512101187
ノルウェー ジャン ユニバーシティ オブ
サイエンス アンド テクノロジー (エ
ヌティーエヌユー)
ノルウェー 7491 トロンヘイム セ
ム サエランツ ヴァイ 14
(74) 代理人 110000040
特許業務法人池内アンドパートナーズ
(72) 発明者 ラジャ、キラン ビラパ
ノルウェー、2816 イュービク、ヨ
ハン ファルクバルゲ ヴェグ 9ビー
(72) 発明者 ラーマチャンドラ、ラガベンドラ
ノルウェー、2816 イュービク、ヨ
ハン ファルクバルゲ ヴェグ 9エフ

最終頁に続く

(54) 【発明の名称】 操作された画像の検出

(57) 【要約】

モーフィングまたは平均化された画像を検出するための装置30であって、モーフィングまたは平均化された画像が、2つまたはそれ以上の対象に対応する2つまたはそれ以上の異なる原画像からの情報を含む人工的に生成された画像である。装置は、入力画像33を受信し、該画像の特性を示す記述子特徴のセットを出力するための特徴抽出モジュールと、記述子特徴に基づいて、入力画像を、該画像がモーフィングまたは平均化されていることを示す第1のクラスまたは該画像がモーフィングまたは平均化されていないことを示す第2のクラスのいずれかに割り当てる分類部モジュール36とを備える。特徴抽出モジュールは、分類部モジュールに相補的な記述子特徴を提供する複数のニューラルネットワーク31、32を含む。装置は、各ニューラルネットワークからの記述子特徴データを組み合わせ、融合された特徴データを分類部モジュールに送信するための融合モジュール35をさらに備える。分類部モジュールは、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む訓練データセットを用いて画

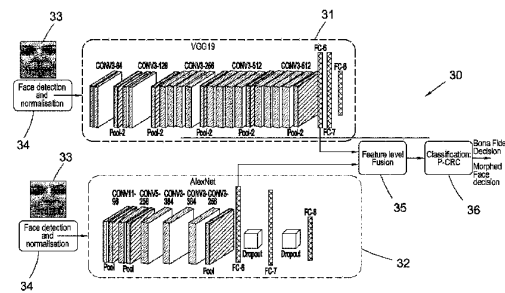


Fig. 6

【特許請求の範囲】**【請求項 1】**

モーフィングまたは平均化された顔画像を検出するための装置であって、前記モーフィングまたは平均化された画像が、2つまたはそれ以上の対象に対応する2つまたはそれ以上の異なる原画像からの情報を含む人工的に生成された画像であり、

入力画像を受信し、該画像の特性を示す記述子特徴のセットを出力するための特徴抽出モジュールと、

前記記述子特徴に基づいて、前記入力画像を、該画像がモーフィングまたは平均化されていることを示す第1のクラスまたは該画像がモーフィングまたは平均化されていないことを示す第2のクラスのいずれかに割り当てる分類部モジュールとを備え、

前記分類部モジュールが、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む訓練データセットを用いて1つの画像を分類するように訓練された機械学習システムを含むことを特徴とする、装置。

【請求項 2】

前記特徴抽出モジュールは、機械学習システムを含み、前記記述子特徴は、画像を含む訓練データセットの使用から決定されるパラメータに依存する、請求項1に記載の装置。

【請求項 3】

前記訓練データセットは、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む、請求項2に記載の装置。

【請求項 4】

前記モーフィングまたは平均化された画像は、モーフィングまたは平均化された顔画像である、請求項1～3のいずれかに記載の装置。

【請求項 5】

前記画像から目的の領域を抽出または正規化し、前処理された前記画像を前記特徴抽出モジュールに送信するように構成された画像前処理モジュールをさらに備える、請求項1～4のいずれかに記載の装置。

【請求項 6】

前記特徴抽出モジュールは、記述子特徴のセットを提供するために前記入力画像のパッチを用いて畳み込まれるフィルタのセットを含む、請求項1～5のいずれかに記載の装置。

【請求項 7】

前記記述子特徴のセットは、畳み込みによって導出される2進量のストリングを含む、請求項6に記載の装置。

【請求項 8】

前記分類部モジュールは、線形サポートベクトルマシンまたは確率的協調表現分類部を含む、請求項1～7のいずれかに記載の装置。

【請求項 9】

前記特徴抽出モジュールは、少なくとも1つの畳み込みニューラルネットワークを含む、請求項1～8のいずれかに記載の装置。

【請求項 10】

前記特徴抽出モジュールは、前記分類部モジュールに相補的な記述子特徴を提供する複数の統計的にほぼ独立したニューラルネットワークを含む、請求項9に記載の装置。

【請求項 11】

前記ニューラルネットワークは、3つまたはそれ以上の畳み込み層を有する深層畳み込みニューラルネットワークを含む、請求項9または10に記載の装置。

【請求項 12】

前記記述子特徴は、各深層畳み込みニューラルネットワークの最初の全結合層から抽出される、請求項11に記載の装置。

【請求項 13】

各ニューラルネットワークからの記述子特徴データを組み合わせ、融合された前記特徴

10

20

30

40

50

データを前記分類部モジュールに送信するための特徴レベル融合モジュールをさらに備える、請求項 10 ~ 12 のいずれかに記載の装置。

【請求項 14】

前記ニューラルネットワークは、画像がモーフィングまたは平均化されているか否かを判定するのに適した記述子特徴を提供するようにそのフィルタを訓練するために、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む画像のセットを用いて個々に訓練される、請求項 9 ~ 13 のいずれかに記載の装置。

【請求項 15】

モーフィングまたは平均化された画像を検出するための装置であって、前記モーフィングまたは平均化された画像が、2 つまたはそれ以上の対象に対応する 2 つまたはそれ以上の異なる原画像からの情報を含む人工的に生成された画像であり、

10

入力画像を受信し、該画像の特性を示す記述子特徴のセットを出力するための特徴抽出モジュールと、

前記記述子特徴に基づいて、前記入力画像を、該画像がモーフィングまたは平均化されていることを示す第 1 のクラスまたは該画像がモーフィングまたは平均化されていないことを示す第 2 のクラスのいずれかに割り当てる分類部モジュールとを備え、

前記特徴抽出モジュールが、前記分類部モジュールに相補的な記述子特徴を提供する複数のニューラルネットワークを含み、

前記装置が、各ニューラルネットワークからの記述子特徴データを組み合わせ、融合された前記特徴データを前記分類部モジュールに送信するための融合モジュールをさらに備え、

20

前記分類部モジュールが、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む訓練データセットを用いて前記画像を分類するように訓練された機械学習システムを含むことを特徴とする、装置。

【請求項 16】

前記ニューラルネットワークは、異なる深層畳み込みニューラルネットワークである、請求項 15 に記載の装置。

【請求項 17】

前記ニューラルネットワークは統計的に独立している、請求項 15 または 16 に記載の装置。

30

【請求項 18】

前記モーフィングまたは平均化された画像は、モーフィングまたは平均化された顔画像である、請求項 15 ~ 17 のいずれかに記載の装置。

【請求項 19】

請求項 2 ~ 14 のいずれかに記載された前記装置の機能をさらに備える、請求項 15 に記載の装置。

【請求項 20】

前記モーフィングまたは平均化された画像が、2 つまたはそれ以上の対象に対応する 2 つまたはそれ以上の異なる原画像からの情報を含む人工的に生成された画像であり、

40

入力画像を受信するステップと、

該画像の特性を示す記述子特徴のセットを生成するステップと、

前記記述子特徴に基づいて、前記画像を、該画像がモーフィングまたは平均化されていることを示す第 1 のクラスまたは該画像がモーフィングまたは平均化されていないことを示す第 2 のクラスのいずれかに割り当てることによって分類するステップとを含み、

前記分類ステップが、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む訓練データセットを用いて 1 つの画像を分類するように訓練された機械学習システムを使用することを特徴とする、方法。

【請求項 21】

前記記述子特徴を生成するステップは、前記分類部モジュールに相補的な記述子特徴を提供する複数の好ましくは統計的にほぼ独立したニューラルネットワークを使用すること

50

を含み、各ニューラルネットワークからの記述子特徴データは、分類に先立って組み合わせられる、請求項 20 に記載の方法。

【請求項 22】

前記分類ステップにおいて使用される分類部を訓練するステップをさらに含む、請求項 20 または 21 に記載の方法。

【請求項 23】

モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む画像のセットを用いて、前記記述子特徴を生成する特徴抽出モジュールを、画像がモーフィングまたは平均化されているか否かを判定するのに適した記述子特徴を該モジュールが提供するように訓練するステップを含む、請求項 20 ~ 22 のいずれかに記載の方法。

10

【請求項 24】

請求項 1 ~ 18 のいずれかに記載された前記装置を使用することを含む、請求項 20 ~ 23 のいずれかに記載の方法。

【請求項 25】

コンピュータに請求項 20 ~ 24 のいずれかに記載の前記方法を実行させるようにプログラミングまたは構成されたことを特徴とする、ソフトウェア製品。

【発明の詳細な説明】

【技術分野】

【0001】

20

本発明は、操作された画像の検出に関し、特に、身元確認および関連するセキュリティ対策を無効にするために操作された画像に関する。限定的ではないが、特に、本発明は、例えば、パスポートや他の身分証明書類、係る書類の申請、国境管理用途または銀行業務用途で使用される際にモーフィングまたは平均化された顔画像の検出に関する。

【背景技術】

【0002】

顔認識は、日常生活の一部となっている汎用バイオメトリクス方法である。顔の生体的特性を、とりわけ、制限のある状況において観察する個人自動認識により、非常に高い精度が実現される。この事実によって、顔のバイオメトリクスは、国際国境管理において主要な位置を占めることとなった。顔認識システムは、過去 40 年にわたって信号・パターン認識アルゴリズムから集められた知識に基づいており、これにより、正確かつ確実な顔認識アルゴリズムが得られている。このような性能の向上によって、顔のバイオメトリクスは、科学捜査から、監視、電子商取引および行政手続の電子申請 (e-Government applications) に対する物理的および論理的なアクセス制御までにわたる多様な用途で使用できるようになっている。

30

【0003】

バイオメトリクスによる顔基準画像は、電子パスポートおよびビザの申請の重要な一部となっており、係るパスポートは 10 年間の導入期間に 8 億通近く発行されている。顔認識が国境管理シナリオのために選択された理由の 1 つは、偽陰性システムの決定の場合、国境管理官によって視認による比較を行うことができるからである。これは、他のバイオメトリクス様式 (例えば、指紋認識) と比較して明確な利点である。これらの事実により、自動国境管理 (Automatic Border Control: ABC) の電子ゲート (e-gates) において顔認識の適用が正当化されている。典型的な ABC システムにおいて、電子機械可読旅行文書 (electronic Machine Readable Travel Document: eMRTD) とパスポート保持者 (すなわち、国境警備員に eMRTD を提示している個人) との関係は、その場で取り込まれた顔画像と eMRTD パスポートに記憶されている基準顔画像とを比較することによって自動的に確認される。これにより、非常に確実かつ正確な国境管理処理が容易となる。実際、国際民間航空機関 (International Civil Aviation Organization: ICAO) は、eMRTD 用の主要な識別子として顔画像を要求することを決定した。したがって、顔画像は、全世界のすべての電子パスポートに存在する唯一のバイオメトリクス基準

40

50

である。

【0004】

A B C システムが広く採用されるようになるにつれ、異なる種類の攻撃に対するその顔認識サブシステムの脆弱性が注目されるようになってきている。これらの攻撃は、大きく2つのタイプに分類することができる。すなわち、A B C システムに対する攻撃と、e M R T D バイオメトリクス基準に対する攻撃である。

【0005】

前者のタイプの攻撃は、典型的には、顔の偽造物を提示することにより、パスポートゲートにおいて撮像機器（またはカメラ）に対して行われる。係る攻撃は、顔スプーフィングまたは提示攻撃と呼ばれる。しかし、これらの攻撃は、顔の偽造物（すなわち、人工提示攻撃器具）を作成し、また、それをA B C 電子ゲートに提示するのに多大な労力を要する。さらに、この種の攻撃が成功するのは、e M R T D パスポートに提示される顔の偽造物に類似し得る顔写真を攻撃者が準備することを可能にする紛失または盗難 e M R T D パスポートに攻撃者がアクセスすることができる場合に限られる。

10

【0006】

後者のタイプの攻撃には、基準画像を置き換えることを目的として、（盗難）パスポートの論理データ構造（L D S）に記憶されたバイオメトリクスデータを操作することが含まれる。しかし、係る攻撃は、顔画像データ集合全体にわたって算出されるハッシュを変更することを含むことになるため、容易に見つかるはずである。したがって、より効果的な手法は、パスポート申請プロトコルおよび発行用プロトコルにおける不備を利用することである。本明細書においてさらに記載され、本発明の実施形態が関連するのは、この種の攻撃である。

20

【0007】

この種の攻撃は、実施が容易である。なぜなら、パスポート発行手続の大部分では、申請過程において（印刷された）顔写真を受け付けるからである。さらに、いくつかの国々では、パスポートの更新およびビザの申請用のウェブポータルにアップロードされたデジタル写真が受け付けられる。いずれの場合も、攻撃者が操作された顔写真をパスポート発行機関に提出し、それによって、物理的および電子的なセキュリティ機能を備えかつ操作された写真を含む正式なe M R T D パスポートを受領する機会を攻撃者に与えることになる。バイオメトリクスキオスク（biometric kiosk）の使用によってこの問題に対処することは可能であるものの、これらは、ほんの少数のパスポート申請局でしか使用されていない。

30

【0008】

係る攻撃の実行に必要な画像操作を行うためのソフトウェアは、無料で入手することができる。例えば、画像のリタッチ、特に、顔画像のリタッチを可能にするソフトウェアは、オンラインソーシャルメディアによってしばしば提供される。様々な異なる種類の顔画像の改変（例えば、幾何学形状、アスペクト比および美化）が、リタッチ画像のカテゴリに含まれ、係る攻撃に利用され得る。実際、リタッチ画像に起因する係る攻撃のリスクが認識されている。非特許文献1は、自動顔認識の性能に対するデジタル改変の影響を示す研究論文である。この論文は、リタッチされた画像を検出するアルゴリズムも紹介している。

40

【0009】

顔モーフィングは、申請者によって提供される画像に常に依拠しているe M R T D パスポートおよびビザの発行プロトコルに対する最も重大な攻撃となりつつある画像操作の別形体である。以下に説明するように、顔モーフィングは複雑な問題を引き起こすものである一方、多数のモーフィング用ソフトウェア製品が無料で入手可能であるため、モーフィング顔画像の生成の複雑性はかなり低い。例えば、G I M P ツールおよびG A P ツールは、高品質のモーフィング顔画像を最小限の労力で生成することができる。

【0010】

顔モーフィング（全体的な顔モーフィングもしくは平均化、または部分に基づく顔モー

50

フィングもしくは平均化によって行われる)の目的は、2つまたはそれ以上の異なる対象に対応する2つまたはそれ以上の異なる顔原画像から得られる固有の情報を用いて人工的に生成された新たな顔画像を生成することである。したがって、顔モーフィングは、顔画像のリタッチとは異なる。なぜなら、後者は、1つの対象に対応するただ1つの顔原画像から得られる固有の情報に依拠しているからである。生成されたモーフィング顔画像は、モーフィングされた顔に寄与した多数個のデータ対象に対応する顔の外観特徴を実質的に形成する。おそらく驚くべきことには、得られる画像は、人間の専門家(例えば、国境警備員)にとっても、また、自動バイオメトリクス顔認識システム用の類似度スコアに関しても、両方の対象に非常に類似したものとなる。

【0011】

10

これにより、任意の攻撃者(例えば、よく知られている犯罪者)が自身の顔画像を他の誰か(例えば、犯罪歴のない誰か)とモーフィングし、両者によって使用可能なeMRTDパスポートを申請するという機会が与えられる。eMRTDパスポートは国境管理用のABCシステムと共に現在広く使用可能となっているため、攻撃者は、このような攻撃をパスポート文書を偽造することなく行うことができる。したがって、国境管理処理の安全を保障するためには、この種の攻撃を減らす必要がある。

【0012】

上記登録処理が顔モーフィングによる攻撃に対して脆弱であることは、民間の顔認識アルゴリズムについて示されている。また、モーフィング画像を人間が検出することが困難であることも実験によって示されている(顔認識の専門家であっても、モーフィング顔画像の検出に失敗している)。

20

【0013】

さらなる問題は、印刷された画像が提出される国々のほとんどで使用されているパスポート申請処理に伴い、モーフィング顔画像は、ID文書の生成についてのISO/IEC規格に準じて典型的には300dpiの解像度で印刷され、その後再びスキャンされる(パスポート事務所において)可能性があるということである。モーフィング画像の再デジタル化されたバージョンでは、モーフィング画像の検出に役立ったかもしれないピクセルレベルの情報が失われている。また、デジタルモーフィングされた画像は、モーフィング画像がパスポート/ビザ発行のために任意の対象によって提出される高品質画像に類似するように、印刷および提出に先立って視覚的品質を向上させ、ゴースト外観を除去するために、リタッチを行うことが可能である。ABCシステムは、係る画像を採用するeMRTDパスポートに対して特に脆弱であることが示されている。

30

【先行技術文献】

【非特許文献】

【0014】

【非特許文献1】IEEE Transactions on Information Forensics and Security, Vol 11, No.9, September 2016, (IEEE, USA), Bharati, Singh, Vatsa & Bowyer, "Detecting Facial Retouching Using Supervised Deep Learning", pp 1903-1913

【発明の概要】

【発明が解決しようとする課題】

40

【0015】

今日までのところ、1つの画像に基づいてモーフィング顔画像を自動的かつ確実に検出するための手法は、本発明者らの研究の他には存在していない。

【課題を解決するための手段】

【0016】

本発明の第1の態様によれば、モーフィングまたは平均化された画像を検出するための装置であって、前記モーフィングまたは平均化された画像が、2つまたはそれ以上の対象に対応する2つまたはそれ以上の異なる原画像からの情報を含む人工的に生成された画像であり、入力画像を受信し、当該画像の特性を示す記述子特徴のセットを出力するための特徴抽出モジュールと、前記記述子特徴に基づいて、前記入力画像を、当該画像がモーフ

50

ィングまたは平均化されていることを示す第 1 のクラスまたは当該画像がモーフィングまたは平均化されていないことを示す第 2 のクラスのいずれかに割り当てる分類部モジュールとを備え、前記分類部モジュールが、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む訓練データセットを用いて 1 つの画像を分類するように訓練された機械学習システムを含むことを特徴とする装置が提供される。

【0017】

したがって、本発明は、分類部を訓練するための訓練データのセットを用いた機械学習によって、モーフィングまたは平均化された写真などのモーフィングまたは平均化された画像（部分または画像全体に基づく）を検出する装置を提供する。特に、訓練済み装置が、画像がモーフィングまたは平均化されているか否かを判定するために分析する 1 つの入力画像を用いてこれを行うことができる。前記訓練データセットは、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像のセットを含み得る。

10

【0018】

前記装置は 1 つの入力画像それ自体を分析することによって判定を行うことが可能であるが、勿論、多数の係る画像を順にまたは同時にこのように分析するために使用されてもよい。

【0019】

好ましくは、前記特徴抽出モジュールは、前記記述子特徴が画像を含む訓練データセットの使用から決定されるパラメータに依存するような機械学習システムをさらに含む。

20

【0020】

本発明は、概して、モーフィングまたは平均化された画像の検出に用いることができるが、上述のパスポートおよび ID 文書の写真のモーフィングに関する公知の問題に鑑み、前記装置は、好ましくは、モーフィングまたは平均化された顔画像を検出し、前記分類部モジュールが前記 1 つの入力画像がモーフィングまたは平均化されているか否かを判定するように構成される。

【0021】

1 つの生の画像が上記装置に入力され得るが、入力画像は、可能な限り標準化されることが非常に好ましい。したがって、前記装置は、好ましくは、上記画像から目的の領域、例えば人間の顔、を抽出または正規化し、前処理された前記画像を前記特徴抽出モジュールに送信するように構成された画像前処理モジュールをさらに備える。

30

【0022】

人間の顔を抽出するために、任意の適した顔検出器アルゴリズム、例えば、Viola-Jones アルゴリズムを用いてもよい。前処理は、例えばアフィン変換によって画像を方向付けることおよび / または画像をシステムへの入力に適した適切なサイズにクロッピングすることを含み得る。

【0023】

原理的には、任意の適した特徴抽出システムを使用することができるが、前記特徴抽出モジュールは、有用には、記述子特徴のセットを提供するために前記入力画像のパッチを用いて畳み込まれるフィルタのセットを含んでもよい。パッチは、重なり合わなくてもよく、すなわち、画像をタイルのように並べてもよく、または任意の適した「ストライド」、すなわち、それらの中心間のピクセル数、の分だけ重なり合っているてもよい。画像のパッチおよびフィルタは、典型的には同じサイズの行列とみなすことができ、畳み込みは、行列の対応する要素を乗算することを含んでもよく、また、得られた結果を合計し、1 つのフィルタにつき 1 つのスカラーを与えることを含んでもよい。

40

【0024】

出力値のセットを提供するために、複数の係るフィルタが各パッチを用いて畳み込まれるのが好ましい。

【0025】

前記記述子特徴のセットは、前記畳み込み、例えば、2 進化された統計的画像特徴 (Bi

50

narised Statistical Image Features : B S I F) から導出される 2 進量のストリングを含んでいてもよい。本明細書に記載されるように、これには、上述のスカラー値のそれぞれに 1 または 0 の 2 進値を割り当てることが含まれる。

【 0 0 2 6 】

本発明は、 3×3 、 5×5 、 7×7 、 9×9 、 11×11 、 13×13 、 15×15 および 17×17 といったフィルタサイズおよび 5、6、7、8、9、10、11 および 12 といった 8 つの異なるビット長を使用し得る。例えば、ビット長が 12 である 11×11 のサイズのフィルタが、正確性の点で、他のフィルタと比べて最も優れた性能を提供し得る。

【 0 0 2 7 】

前記分類部モジュールは、任意の適したタイプであり得る。同モジュールは、例えば、線形サポートベクトルマシン (linear support vector machine) または確率的協調表現分類部 (probabilistic collaborative representation classifier) を含み得る。

【 0 0 2 8 】

本発明の好ましい形態において、前記特徴抽出モジュールは、少なくとも 1 つの畳み込みニューラルネットワークを含む。係るネットワークは、学習可能なフィルタ (「カーネル」ともいう) のセットをそれぞれ有する 1 つまたはそれ以上の畳み込み層 (段階) を含む。好ましくは、前記ニューラルネットワークは、深層畳み込みニューラルネットワーク (D-CNN) である。「深層」という用語は、複数の係るフィルタが備えられていることを意味する。3 つまたはそれ以上の畳み込み層が備えられていることが好ましいが、これらの層が、5 つまたはそれ以上存在することが最も好ましく、7 つまたはそれ以上存在することがさらに好ましい。

【 0 0 2 9 】

D-CNN の最初の畳み込み層は、その入力として、画像のパッチのセットを受信する。パッチは、重なり合っているとしてもよく、各パッチの中心間のピクセル数は「ストライド」と呼ばれる。白黒画像の場合、各パッチは、 $m \times n$ の寸法を有し、上記入力画像がカラーである場合、各パッチは、 $m \times n \times (RGB \text{ の場合 })$ の寸法を有し、 m および n は、選択されたパッチの行および列である。

【 0 0 3 0 】

フィルタは、概して、寸法においてパッチに対応しており、各畳み込みは、各行列の対応するメンバを乗算し、その結果を合計することにより算出される。換言すれば、各フィルタは、入力量 (パッチ) の幅および高さによって畳み込まれ、活性化マップ (activation map) が形成されてもよい。すべてのフィルタのマップのスタックにより、深さ寸法が形成されてもよい。したがって、層によって出力される行列の寸法は、適用されるパッチ数およびフィルタ数に対応し得る。

【 0 0 3 1 】

畳み込み層は、典型的には、所与のパッチからの出力として提供される行列が次の段階の入力を形成するように直列に配置される。これらの畳み込み層は、好ましくは各パッチを個別に (すなわち、並列に) 取り扱うため、「全結合」とは異なり、これらは単に「局所結合」される。

【 0 0 3 2 】

深層 CNN はまた、例えば、 $j \times k$ 個の値のセットがセット中の最大値に対応する 1 つの値によって置換される「最大プーリング」によって、例えば、データが「簡略化される」ように有意な特徴が抽出される「プーリング」層も有し得る。

【 0 0 3 3 】

前記装置は 1 つのニューラルネットワークを備えていてもよいが、前記特徴抽出モジュールは、前記分類部モジュールに相補的な記述子特徴を提供する複数の統計的に (少なくともほぼ) 独立したニューラルネットワークを備えることが好ましい。

【 0 0 3 4 】

以下に記載されるように、これにより、「クロスチェック」を有効に実現する独立した

10

20

30

40

50

記述子特徴を分類部が考慮することができるようになり、それによってシステムの信頼性が大きく向上するという点で、相乗効果がもたらされる。これは、ニューラルネットワークが同じデータに対して訓練される場合であっても、当該ネットワークが実質的に異なるアーキテクチャを有する場合には、特有の相補的な記述子特徴が当該ネットワークによって提供されることになるという、本発明者らによる認識に基づいている。

【0035】

この概念は、特に、また単独で、進歩性を有すると考えられる。したがって、別の態様によれば、モーフィングまたは平均化された画像を検出するための装置であって、前記モーフィングまたは平均化された画像が、2つまたはそれ以上の対象に対応する2つまたはそれ以上の異なる原画像からの情報を含む人工的に生成された画像であり、入力画像を受信し、当該画像の特性を示す記述子特徴のセットを出力するための特徴抽出モジュールと、前記記述子特徴に基づいて、前記入力画像を、当該画像がモーフィングまたは平均化されていることを示す第1のクラスまたは当該画像がモーフィングまたは平均化されていないことを示す第2のクラスのいずれかに割り当てる分類部モジュールとを備え、前記特徴抽出モジュールが、前記分類部モジュールに相補的な記述子特徴を提供する複数の（好ましくは、統計的にほぼ独立した）ニューラルネットワークを含み、前記装置が、各ニューラルネットワークからの記述子特徴データを組み合わせ、融合された前記特徴データを前記分類部モジュールに送信するための融合モジュールをさらに備え、前記分類部モジュールが、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む訓練データセットを用いて1つの画像を分類するように訓練された機械学習システムを含む装置が提供される。

10

20

【0036】

「統計的にほぼ独立している」という用語は、それらが相補的な記述子特徴を提供し、それによって、1つのニューラルネットワークを使用する場合と比べて、検出の精度が著しく改善されることを意味する。

【0037】

上述のように、前記ニューラルネットワークは、好ましくはD-CNNである。原理的には、前記記述子特徴はD-CNNの任意の層から得ることができるが、これらが各深層畳み込みニューラルネットワークの最初の全結合層から抽出されることが好ましい。本発明のこの態様は、他の態様に関連して記載される好適な機能または随意的な機能のいずれを備えていてもよい。特に、モーフィング顔画像の検出に適合させてもよい。

30

【0038】

前記装置は2つのニューラルネットワークからの出力を使用するため、前記装置は、各深層畳み込みニューラルネットワークからの記述子特徴データを組み合わせ、融合された（例えば、連結された）前記特徴データを前記分類部モジュールに送信するための特徴レベル融合（feature level fusion）モジュールをさらに備えることが好ましい。

【0039】

画像認識/物体検出および認識を目的としたニューラルネットワークは、係る目的のために正常な画像に基づいて事前訓練された状態で入手可能である。前記分類部はモーフィングまたは平均化された画像を検出するように別途訓練されるため、係るネットワークをさらに訓練することなく本発明において使用することが可能である。しかし、画像がモーフィングまたは平均化されているか否かを判断するのに適した記述子特徴を提供するようにそれらのフィルタを訓練するために、前記ニューラルネットワークは、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像（例えば、モーフィングされかつ正常な画像）を含む画像のセットを用いて個別に訓練されることが好ましい。

40

【0040】

任意の適したD-CNN、特に、適切な画像データベースに対して訓練済み（または微調整済みの）D-CNNを本発明において使用することができるが、本発明は、好ましくは、2つの一般的な事前訓練済みD-CNN、すなわち、VGG19およびAlexNet

50

tを使用する。本発明者らは、これらのD-CNNが統計的にほぼ独立していると判定した。

【0041】

VGG19は、多数（一例として、16～19）の層を有する非常に小さい（すなわち、 3×3 ）畳み込みフィルタを使用する。一方、AlexNetは、やや大きめのフィルタ（例えば、最初の層におけるストライドが4である 11×11 ）および8つの層（うち5つが畳み込み層）を使用する。これには、一部の出力がゼロに設定されているいくつかの「ドロップアウト」層も含まれる。

【0042】

本発明は、対応する方法にも適用され、したがって、別の態様によれば、本発明は、前記モーフィングまたは平均化された画像が、2つまたはそれ以上の対象に対応する2つまたはそれ以上の異なる原画像からの情報を含む人工的に生成された画像であり、入力画像を受信するステップと、当該画像の特性を示す記述子特徴のセットを生成するステップと、前記記述子特徴に基づいて、前記画像を、当該画像がモーフィングまたは平均化されていることを示す第1のクラスまたは当該画像がモーフィングまたは平均化されていないことを示す第2のクラスのいずれかに割り当てることによって分類するステップとを含み、前記分類ステップが、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む訓練データセットを用いて1つの画像を分類するように訓練された機械学習システムを使用することを含むことを特徴とする方法を提供する。

【0043】

上述したように、前記記述子特徴を生成するステップは、好ましくは、前記分類部モジュールに相補的な記述子特徴を提供する複数の統計的にほぼ独立したニューラルネットワーク、好ましくは、D-CNNを使用することを含み、各ニューラルネットワークからの記述子特徴データは、分類に先立って組み合わせられる。

【0044】

同様に、前記方法は、前記分類ステップにおいて使用される分類部を使用に先立って訓練するステップをさらに含み得る。同方法はまた、モーフィングまたは平均化された画像およびモーフィングまたは平均化されていない画像を含む画像のセットを用いて、前記記述子特徴を生成する特徴抽出モジュールを、画像がモーフィングまたは平均化されているか否かを判定するのに適した記述子特徴を当該モジュールが提供するように訓練するステップを含み得る。

【0045】

実際、前記方法は、本発明の装置に関連して上述された好適な機能または随意的な機能のいずれかまたはすべてに対応するステップをさらに含み得る。

【0046】

具体的な例としては、本発明の特定の用途は、パスポート/ビザ発行所に提示され得るか、あるいは、実際、犯罪または詐欺目的で自動パスポートゲートに提示され得るモーフィングまたは平均化された画像の検出にある。したがって、本発明は、パスポートまたは他の身分証明書類に関連するモーフィング画像を検出することによって係る行為を防止する方法、および係る目的で自動化された装置（係る装置を備える自動パスポートゲートを含む）に好適に適用される。

【0047】

本発明は、典型的にはメモリ、プロセッサおよびデータ記憶装置を含むコンピュータまたは他の適したデータ処理装置を用いて概して実現される。したがって、本発明はまた、係る装置を本発明の装置および/または方法を実現するようにプログラミングまたは構成するためのソフトウェア製品にも適用される。したがって、本発明の別の態様によれば、係るソフトウェア製品が提供される。

【0048】

本発明の上記態様は、モーフィングまたは平均化された画像の検出に特に着目したものであるが、上述の装置および方法は、他の方法で操作された（例えば、リタッチされた）

10

20

30

40

50

画像を検出するために同様に使用され得ることが理解される。したがって、本発明は、操作された画像の検出においてより広く使用することができる。

【0049】

したがって、本発明の別の態様よれば、操作された画像を検出するための装置であって、入力画像を受信し、当該画像の特性を示す記述子特徴のセットを出力するための特徴抽出モジュールと、前記記述子特徴に基づいて、前記入力画像を、当該画像が操作されていることを示す第1のクラスまたは当該画像が操作されていないことを示す第2のクラスのいずれかに割り当てる分類部モジュールとを備え、前記特徴抽出モジュールが、前記分類部モジュールに相補的な記述子特徴を提供する複数のニューラルネットワークを含み、前記装置が、各ニューラルネットワークからの記述子特徴データを組み合わせ、融合された前記特徴データを前記分類部モジュールに送信するための融合モジュールをさらに備え、前記分類部モジュールが、操作された画像および操作されていない画像を含む訓練データセットを用いて1つの画像を分類するように訓練された機械学習システムを含む装置が提供される。また、本発明のさらに別の態様において、対応する方法も提供される。本発明のこれらの態様は、本発明の上述の態様に関連して上述された随意的な機能および/またはステップの恩恵を同様に受けることが理解される。

【図面の簡単な説明】

【0050】

【図1】図1は、本発明の第1の実施形態に係るモーフィング顔画像検出システムのブロック図である。

【図2】図2は、図1の実施形態のBSIF段階の定性的な結果を示す図である。図2(a)は正規化された顔画像の集合、図2(b)はBSIF特徴を示す対応する画像の集合、図2(c)はそれらのBSIF特徴のヒストグラム、図2(d)は同ヒストグラムの拡大部分をそれぞれ示す。

【図3】図3は、モーフィング画像の生成に用いる対象の一对の写真である。

【図4】図4は、モーフィング画像データベースの生成に用いる対象対の正規化された画像および対応するモーフィング画像のセットである。

【図5】図5は、データベースからのモーフィング画像をモーフィング画像の生成元である対象と比較することによって得られた散乱比較スコアを示すグラフである。

【図6】図6は、本発明の第2の実施形態に係るモーフィング顔画像検出システムを示すブロック図である。

【図7】図7は、第2の実施形態で使用された3×3のフィルタサイズを有する(a)公知のAlexNetネットワークおよび(b)公知のVGG19ネットワークでそれぞれ使用された深層畳み込みニューラルネットワークの畳み込み層において使用されたフィルタの重みを示す。

【図8】図8は、第2の実施形態で使用されたモーフィング画像データベースの生成に用いる対象対の正規化された画像および対応するモーフィング画像のセットである。

【図9】図9は、IG-SVM法の性能(a)を第2の実施形態の性能(b)と比較する一对のグラフである。

【発明を実施するための形態】

【0051】

以下に、添付の図面を参照し、本発明についての好適な実施形態を、あくまで例として説明する。

【0052】

まず図1を参照すると、ローバストなモーフィング顔画像検出のためのシステムの第1の実施形態10のブロック図が示されている。同システムは、2つの主要な段階を有する。すなわち、前処理11およびモーフィング顔検出12である。

【0053】

第1のステップは、検出される画像13をシステムに入力するためのものである。

【0054】

次に、前処理段階 11 は、正規化された顔領域を抽出する。これには、まず、画像内の顔を検出することが含まれる。これは、ローバストであり実際のシナリオにおいて良好に機能する周知の物体検出フレームワークである Viola - Jones アルゴリズムを用いて行われる。なお、他の任意の顔検出器、例えば、D - CNN に基づく顔検出器を使用してもよい。次のステップにおいて、画像 13 を正規化し、アフィン変換を用いて回転を補正し、最後に、正規化された画像を $x \times y$ 個のピクセル（ここでは、 120×120 個のピクセル）にリサイズする。これにより、標準サイズの正規化された画像 14 の出力が得られる。

【0055】

概略的には、モーフィング顔検出段階 12 は、正規化された顔画像 14 から抽出された微小テクスチャ（micro-texture）特徴の解析に基づくものである。これは、2 進化された統計的画像特徴（Binarised Statistical Image Features）（BSIF）フィルタ 15 を用いて行われ、その出力（検出された特徴のセット）は、線形サポートベクトルマシン（SVM）16 に渡されて分類される。これにより、画像は、「正常な顔」17 または「モーフィング顔」18（すなわち、拒否される）のいずれかに分類される。

【0056】

BSIF フィルタリングは、画像記述子を提供するための公知の技術である（J. Kanna la and E. Rahtu. "BSIF: Binarized Statistical Image Features" 21st International Conference on Pattern Recognition (ICPR), 2012 pages 1363-1366 を参照）。まず、グレースケール画像を、0 が平均強度を示すように、すなわち、正規化された画像を表す行列が正と負の値を有するように正規化し、次いで、 $I \times I$ 個のピクセルの正方形パッチに分割する。すなわち、各パッチは $I \times I$ 個の実数行列で表される。

【0057】

次いで、各パッチに対して、それぞれが同様に $I \times I$ 個の実数行列である一連の n 個の線形フィルタで畳み込みを行い、フィルタ応答を得る。各フィルタについてのフィルタ応答は、各行列の対応するメンバを乗算し、その結果を合計することにより算出される。次いで、0 より大きい場合には 1 の出力、そうでない場合には 0 の出力を与えることにより、合計を 2 進化する。 n 個のフィルタのセットが適用されるので、 n 個の 2 進ビットのストリングが得られ、その値は、画像の BSIF 記述子（特徴）を構成し、それを分類するために用いることができる。

【0058】

図 1 を参照すると、段階 15 において、上記の特徴が画像から得られる。使用される特定のフィルタセットは、オープンソースのセットである。これは元々、13 個の異なる自然風景画像からランダムにサンプルされた 50,000 個の画像パッチを用いた機械学習訓練段階によって得られたものである（A. Hyvriinen, J. Hurri, and P. O. Hoyer. Natural Image Statistics, volume 39. Springer, 2009 を参照）。この機械学習は、教師なし学習であり、独立成分分析法を用いたものであった（J. H. van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265 (1394):359-366, 1998 を参照）。得られたフィルタは、統計的に高度に独立している。したがって、上記フィルタのセットは、画像認識用に最適化されているが、本実施形態においてはモーフィング画像検出に関して特別に訓練されていない。

【0059】

BSIF フィルタは教師なし学習に基づいて生成されるため、異なるサイズを有する任意の数のフィルタが使用され得る。本発明者らは、 3×3 、 5×5 、 7×7 、 9×9 、 11×11 、 13×13 、 15×15 および 17×17 といった 8 つの異なるフィルタサイズを、5、6、7、8、9、10、11 および 12 といった 8 つの異なるビット長を用いて評価した。開発データセット（下記を参照）に対する実験に基づき、正確性の点で最も好ましいものとして、ビット長が 12 である 11×11 のフィルタサイズを選択した。

【0060】

図2は、正常な顔画像およびモーフィング顔画像の両方について上記好ましいフィルタサイズを用いて得られたBSIF特徴の定性的な出力を示している。第1および第2の対象20、21の正規化された画像を、これら2つの対象の正規化されたモーフィング画像22と共に図2(a)に示す。

【0061】

フィルタを適用した結果が図2(b)に示されており、同図において、BSIF特徴(すなわち、各パッチの記述子)は、これらが関連しているパッチの位置においてグレースケールに変換されて示されている。加えて、図2(c)に示すヒストグラムは、各パッチについての記述子の値をピンの線形シーケンスで示している。12個のフィルタが使用されているため、各記述子のビット長は12であり、したがって、寸法は、 1×4096 (すなわち、 2^{12})である。

【0062】

なお、BSIFヒストグラムの特徴(図2(c))は、正常な顔画像とモーフィング顔画像との間のヒストグラムプロファイルにおけるバラツキを示している。この違いは、図2(d)に示されるように、3つの顔画像すべてについての対応するピンにおけるヒストグラムプロファイルを拡大することによってさらに明らかとなる。

【0063】

図1を再び参照すると、記述子データ(すなわち、図2(c)に示すデータ)は、次いで、線形SVM分類部16に渡される。これは、正の(正常な顔)サンプルおよび負の(モーフィング)サンプル(例えば、以下の実験を参照)の大きいセットを用いて、これらを高い信頼度で区別できるように周知の方法で事前に訓練されている。したがって、SVM分類部は、記述子データを用いて、提示された顔画像が正常な(参照符号17)クラスまたはモーフィング(参照符号18)クラスに属しているかを判定し、対応する出力をユーザに提供する。このシステムが使用される場所、例えば、パスポート/ビザ発行所では、画像にモーフィングクラスを割り当てることにより、当該画像は拒否されることになる。

【0064】

第1の実施形態の変形例において、BSIF特徴抽出システムに代えて、後述する第2の実施形態において用いられるような単一の深層畳み込みニューラルネットワーク(D-CNN)が使用される(記載されるどちらのD-CNNを使用してもよい)。第2の実施形態(第2の実施形態の特徴レベル融合は必要ではない)との関連でも説明されるように、D-CNNの最初の全結合層から入力を受信し得る、第2の実施形態の分類部システムを使用してもよい。

【0065】

[実験]

本発明者らは、110個のデータ対象から得られた異なる組み合わせの顔画像を用いて生成された450個のモーフィング画像で構成される新規な大型モーフィング顔データベースを構築した。データ収集における最初のステップは、eMRTDパスポート仕様に規定されるICAOの取り込み規格に従って顔画像を取り込むことであった。この点で、本発明者らは、まず、スタジオ環境において、均一な照明、均一な背景、ニュートラルな姿勢および通常の顔の表情で正面顔画像を収集した。画像は、三脚に載置され、対象から2メートルの距離に配置されたキャノンEOS 550D DSLRカメラを用いて取り込んだ。図3は、2つのデータ対象の取り込まれた高品質顔画像の例を示す。

【0066】

次いで、無料で入手可能なGNU画像操作プログラム(GNU Image Manipulation Program)v2.8(GIMP)ツールおよびGIMPアニメーションパッケージ(GIMP Animation Package)(GAP)ツールを用いて、モーフィング顔画像を生成した。モーフィングされる2つの顔画像を手動で位置合わせし、GAPツールに対する入力として供した。次いで、GAPツールは、ある対象から別の対象への移行を示す画像フレームのシーケンス

10

20

30

40

50

を生成した。最終のモーフィング画像を、モーフィング処理への貢献対象の顔との類似性を確認することによって手動で選択した。図4は、それぞれに対して2つの異なる対象を用いて得られたモーフィング顔画像の3つの例を示しているが、モーフィング画像はまた、3つの対象を用いて同様に生成した。左の列は第1の対象の画像を示し、第2の列は第2の対象の画像を示し、第3の列はそれぞれのモーフィング画像を示している。

【0067】

モーフィング顔データベースを適切に評価し、モーフィング顔検出アルゴリズムをベンチマークテストするために、450個のモーフィング画像のデータベース全体を3つの独立したサブセットに分割した。すなわち、訓練セット、開発セットおよび試験セットである。訓練セットは200個のモーフィング画像で構成され、これらは、SVM分類部を訓練するためだけに使用した。開発セットは50個のモーフィング画像で構成され、これらは、特にBSIFフィルタの大きさおよび長さの選択において、提案されるスキームのパラメータを調整するために使用した。試験セットは200個のモーフィング画像で構成され、これらは、モーフィング顔検出アルゴリズムの結果を報告するためだけに使用した。

【0068】

モーフィング顔画像データベースに対し、Neurotechnology (<http://www.neurotechnology.com/verilook.html>)を参照)から入手可能な市販のVerilook顔認識SDKを用いて脆弱性分析を行った。分析は、モーフィング顔画像をVerilookに登録し、モーフィング画像の生成に使用したデータ対象のうちの1つに対応するプロブサンプルを使用することによって行った。したがって、登録されたモーフィング顔のそれぞれについて、モーフィング顔画像の生成に使用された対象の数に応じて、2つまたは3つの異なる比較スコアが得られた。

【0069】

図5は、Verilook顔SDKを用いて450個のモーフィング画像で構成されるデータベース全体に対して得られた散乱比較スコアを示す。これらのスコアの有効性は、FRONTEX (欧州対外国境管理協力機関)のガイドラインに従って評価される。FRONTEXは、検証モードで動作するABCシステムにおける顔検証アルゴリズムは他人受入率(FAR)が0.001(0.1%)またはそれよりも良好となる性能を実現するものであるのがよいと推奨している。本構成において、FRRは、0.05(5%)を上回らないほうがよい。実験においては、SDKによって得られた閾値スコアを用いたが、これは、上記特定のFAR=0.1%に対して36である。図5の赤色の垂直線は、目標FAR=0.1%に対応する検証閾値を示している。したがって、36を超える比較スコアは、検証の成功とみなされる。図5から分かるように、すべてのモーフィング画像がこの閾値と良好に一致しており、これにより、実際の用途についてのモーフィング顔画像の攻撃ポテンシャル(attack potential)が示されている。

【0070】

第1の実施形態を自動モーフィング顔検出に用いた定量的な結果について、4つの異なる最新の特徴抽出スキーム、すなわち、画質解析(Image Quality Analysis: IQA)、局所バイナリパターン(Local Binary Patterns: LBP)、局所位相量子化(Local Phase Quantisation: LPQ)および2D高速フーリエ変換(2DFFT)の比較評価と共に以下に説明する。これらの比較器は、問題に対するその関連性および近年の入力データ攻撃検出(Presentation Attack Detection)研究におけるその精度を考慮して選択された。これらの特徴抽出スキームの分類は、提案されるスキームに合わせて、線形SVMを用いて行った。

【0071】

結果を定量化するため、2つの異なる性能評価メトリクスを定義した。すなわち、(i)正常な顔画像として分類されるモーフィング顔画像(NFCM)(換言すれば、正常な顔画像がモーフィング顔画像として分類される比率)および(ii)正常な顔画像として分類されるモーフィング顔画像(MFCN)(換言すれば、正常な顔として分類されるモーフィング顔画像の比率)である。全体的な精度は、 $ACER = (NFCM + MFCN)$

/ 2 として定義される平均分類エラー率 (A C E R) を用いて測定することができる。

【 0 0 7 2 】

表 1 は、第 1 の実施形態 (表中の「提案される方法」) の定量的な結果を、本研究において使用した 4 つの異なるベースラインアルゴリズムと共に示している。

【 0 0 7 3 】

【表 1】

アルゴリズム	MFCN (%)	NFCM (%)	ACER (%)
画質 [7] - SVM	1.73	73.37	37.55
LBP [16] - SVM	37.66	13.20	25.43
LPQ [2] - SVM	29.00	11.47	20.23
2DEFT [15] - SVM	61.03	37.22	49.12
提案される方法	3.46	0	1.73

表 1

【 0 0 7 4 】

得られた結果に基づき、以下のことが観察される。

- ・最も優れた性能は、 A C E R が 1 . 7 3 % である第 1 の実施形態 (「提案される方法」) の使用について認められる。
- ・最も優れた M F C N は、 M F C N が 1 . 7 3 % である画質解析特徴について認められるが、この設定の N F C M 値は 7 3 . 3 7 % とかなり高く、実際のシナリオにおいては適用不可能である。

【 0 0 7 5 】

したがって、 B S I F フィルタに基づく統計的な画像の使用は、従来の特徴抽出技術と比較して最も優れた性能を示した。得られた結果は、自動モーフィング顔検出についての第 1 の実施形態の適用可能性を示している。

【 0 0 7 6 】

以下に、本発明の第 2 の実施形態について、他の図面を参照して説明する。本実施形態は、印刷・スキャン処理を経ており、したがって「デジタル」モーフィング画像よりも検出が困難であるモーフィング画像の認識に特に適している。印刷・スキャン処理は、最も広く採用されているパスポート申請処理に対応するものである。

【 0 0 7 7 】

図 6 は、第 2 の実施形態のシステム 3 0 を示すブロック図である。より詳細に後述されるように、同システムは、モーフィング顔画像を検出するための、2 つの事前訓練済み深層畳み込みニューラルネットワーク (D - C N N) 3 1、3 2 の特徴レベル融合に基づいている。使用されるニューラルネットワークは、画像認識において使用されるものとして公知であり、当該目的で事前訓練済みである。

【 0 0 7 8 】

畳み込みニューラルネットワークは、1 つまたはそれ以上の畳み込み層 (段階) を含み、これらはそれぞれ、上述の実施形態において使用されたものに類似する学習可能なフィルタ (「カーネル」ともいう) のセットを有する。「深層」という用語は、複数の係るフィルタが備えられていることを意味する。

【 0 0 7 9 】

第 1 の実施形態において用いられた B S I F 特徴抽出システムと同様、 D - C N N の第 1 の畳み込み層は、上記画像のパッチのセットをその入力として受信する。パッチは重なり合っているように、各パッチの中心間のピクセル数は「ストライド」と呼ばれる。白黒画像の場合、各パッチは、 $m \times n$ の寸法を有し、入力画像がカラーである場合、各パッチは、 $m \times n \times 3$ (R G B の場合) の寸法を有し、 m 、 n は、画像パッチの行および列である。フィルタは、それぞれ、寸法においてパッチに対応しており、各フィルタについてのフィルタ応答 (すなわち、畳み込み) は、各行列の対応するメンバを乗算し、その結果を合計することにより算出される。換言すれば、各フィルタは、入力量 (パッチ) の幅およ

び高さにわたって畳み込まれ、活性化マップが形成される。すべてのフィルタについてのマップのスタックにより、深さ寸法が形成される。したがって、上記層によって出力される行列の寸法は、適用されるパッチ数およびフィルタ数に対応している。

【0080】

畳み込み層は、所与のパッチからの出力として提供される行列が次の段階の入力を形成するように直列に配置される。これらの畳み込み層は各パッチを個別に（すなわち、並列に）取り扱うため、「全結合」とは異なり、これらは単に「局所結合」される。

【0081】

深層CNNはまた、例えば、 $j \times k$ 個の値のセットがセット中の最大値に対応する1つの値によって置換される「最大プーリング」によって、例えば、データが「簡略化される」ように有意な特徴が抽出される「プーリング」層も有する。

10

【0082】

複数の畳み込み層（および、通常、プーリング層）に続いて、画像全体の特性に基づいてデータを受信し、入力画像を多数の分類のうち1つに割り当てる出力を提供する分類段階を含む1つまたはそれ以上の全結合層が存在する。

【0083】

本実施形態において用いる2つの公知の深層CNN（画像認識用に設計された）は、いくつかの全結合層が後に続く、多数の畳み込み層およびプーリング層をそれぞれ有する。これらは、BSIFシステムに関連して記載したものと概して同様の方法で画像認識のために事前に訓練されている（すなわち、訓練データセットが機械学習処理において用いられ、それにより、画像認識のための最適なフィルタが学習される）。

20

【0084】

図6を再び参照すると、第2の実施形態のシステムは、(i)ブロック34における画像33の前処理、(ii)ブロック35における2つのD-CNN31、32の特徴レベル融合および(iii)ブロック36における分類を含む3つの主要な機能ブロックを用いて構築される。各ブロックの機能については以下に述べる。

【0085】

前処理ブロック34は、それぞれ、第1の実施形態に関連して記載した対応するブロック11と概して類似している。したがって、Viola-Jonesアルゴリズムを用いて顔画像33から顔検出を行い、検出された顔領域を正規化して、アフィン変換を用いて回転を補正し、画像を一例として 227×227 個のピクセルにリサイズする（このサイズは、当該サイズの画像を用いて同じく事前に訓練されたD-CNNの入力層のサイズに適合可能であるという理由で使用される）。

30

【0086】

次いで、2つの異なるD-CNNのそれぞれに対して同様の前処理済み出力を与える。

【0087】

システムは、2つの一般的な事前訓練済みD-CNN、具体的には、VGG19（参照符号31）およびAlexNet（参照符号32）を使用する（(i) K. Simonyan and A. Zisserman. "Very deep convolutional networks for large-scale image recognition". arXiv preprint arXiv:1409.1556, 2014 and A. Krizhevsky, I. Sutskever; and (ii) G. E. Hinton. "Imagenet classification with deep convolutional neural networks". In Advances in neural information processing systems, pages 1097-1105, 2012をそれぞれ参照）。

40

【0088】

VGG19は、多数（一例として16~19）の層を有する非常に小さい（すなわち、 3×3 ）畳み込みフィルタを使用する。一方、AlexNetは、やや大きいフィルタ（例えば、最初の層におけるストライドが4である 11×11 ）および8つの層（うち5つが畳み込み層）を使用する（一部の出力がゼロに設定されたいくつかの「ドロップアウト」層も有する）。各D-CNNの畳み込み層およびプーリング層を図面において概略的に示す。

50

【 0 0 8 9 】

いずれの D - C N N も同じ大型の I m a g e N e t データベースに対して訓練される。しかし、これらは異なる構成（すなわち、アーキテクチャ）に基づいているため、同じデータベースに対して訓練されているにもかかわらず、モーフィング顔画像の検出に適した相補的な特徴を提供する。

【 0 0 9 0 】

D - C N N は、モーフィング顔データベース（以下を参照）を用いて使用される前に微調整される。

【 0 0 9 1 】

2つのネットワークの組み合わせは特徴レベルで行われ、特徴レベルでは、畳み込み層のセットを用いて各ネットワークから抽出された特徴が例えば連結によって融合される。図 6 から分かるように、出力は、両方のネットワーク 3 1、3 2 の最初の全結合層（F C - 6）から抽出される。この層は、特に A l e x N e t においてドロップアウト層が使用される前に完全な特徴を利用するために選択される。次いで、特徴レベル融合ブロック 3 5 において、これらの特徴を連結して 1 つのベクトルを形成し、次いで、ベクトルを分類ブロック 3 6 に入力する。

【 0 0 9 2 】

第 1 の実施形態の分類部と同様、公知の方法で事前に訓練された分類ブロック 3 6 は、この特徴ベクトルを用いて画像が正常であるかモーフィングされているか（したがって、拒否されるべきか）を判定する。使用される分類部は、分類を行うための他のクラスと共に試験サンプルの尤度比を最大化する、確率的協調表現分類部（P - C R C）である。

【 0 0 9 3 】

第 1 の実施形態と同様、第 2 の実施形態の分類部 3 6 は、正常な画像およびモーフィング画像の両方を含むモーフィング顔画像データベースからの訓練データセットを用いて訓練される。このデータベースの詳細を表 2 に記載する。

【 0 0 9 4 】

【表 2】

種類	サンプル数		
	訓練セット	試験セット	合計
真正画像	2 0 6	1 4 6	3 5 2
モーフィング画像	2 2 5	2 0 6	4 3 1

表 2

【 0 0 9 5 】

ただし、さらに、この訓練データセットは、画像認識において事前に訓練済みであるが、モーフィング画像に関しては事前に訓練されていない V G G 1 9 ネットワークおよび A l e x N e t ネットワークを個別に微調整するために使用される。この微調整には、個々の D - C N N を、画像をそれらが上述のように組み合わせられる前に、正常な画像またはモーフィング画像に分類するように訓練することが含まれる。

【 0 0 9 6 】

ネットワークを微調整する間、最後の層に対し、ネットワークのその他の層と比べて高い値の学習率が適用される。使用された学習パラメータは、重み学習率係数が 1 0、バイアス学習率係数が 2 0 であった。

【 0 0 9 7 】

〔分析〕

微調整済みの V G G 1 9 および A l e x N e t のネットワーク特徴を可視化するために、両方の D - C N N からの 3 番目の畳み込み層（C o n v 3）からの特徴を図 7 に示す。

【 0 0 9 8 】

10

20

30

40

50

図7(a)は、畳み込みフィルタのサイズが 3×3 で長さが384の微調整済みAlexNetからのConv3の特徴を示している。興味深いことに、各畳み込みフィルタが顔の特徴およびテクスチャの特徴を示していることが観察される。さらに、強調表示領域（異なる色で示す）は、微調整済みのAlexNetによって学習された顔特有の特徴を示している。

【0099】

同様のことが、図7(b)からも観察される。同図は、畳み込みフィルタのサイズが 3×3 で長さが256のConv3層を有する微調整済みVGG19についての対応する図である。図7(b)の強調表示領域は、VGG19の微調整によって学習された顔の特性からの要素を示している。

【0100】

2つの微調整済みD-CNNネットワークからの特徴が相補的な情報（したがって、相乗効果）をもたらす度合いを、該当する相互相関関数（CC）を以下のように計算することによって調べた。

【0101】

【数1】

$$CC = \frac{\sum_m \sum_n (F_{Alex} - \bar{F}_{Alex})(F_{VGG} - \bar{F}_{VGG})}{\sqrt{\left(\sum_m \sum_n (F_{Alex} - \bar{F}_{Alex})^2\right) \left(\sum_m \sum_n (F_{VGG} - \bar{F}_{VGG})^2\right)}} \quad (1)$$

式中、 F_{VGG} および F_{Alex} は、それぞれのD-CNNからの記述子特徴の値を表し、 \bar{F}_{VGG} および \bar{F}_{Alex} は、それらの平均値である。これに応じて、平均値は、行および列（ m および n ）の全体にわたるピクセル強度に対応している。CCの値が低いほど、データがより相補的であることを示す。

【0102】

以下の表3に結果を示す。

【0103】

【表3】

データタイプ	特徴	CCC
デジタル	訓練Alex-訓練VGG	0.047
	試験Alex-試験VGG	0.045
印刷・スキャン (HP)	訓練Alex-訓練VGG	0.024
	試験Alex-試験VGG	0.009
印刷・スキャン (リコー)	訓練Alex-訓練VGG	0.007
	試験Alex-試験VGG	0.012

表3

【0104】

なお、CC値（表中では「CCC」と称する）はいずれも非常に低く、これにより、2つのD-CNNから得られた特徴が本質的に相補的である、すなわち、これらが組み合わせにおいて相乗効果を有するということが示されている。

【0105】

分類部のモーフィング検出性能を調べるため、AlexNetから抽出されたFC-6特徴を F_A 、VGG19から抽出されたFC-6特徴を F_V とする。次いで、抽出された特徴を、これらを連結することによって組み合わせ、P-CRC（SVMやRandom Forestといった他の同等の分類部を使用することもできる）を訓練するために使用

する1つの特徴ベクトル $Tr_F = [F_A \parallel F_V]$ を形成した。次いで、試験顔画像 F_{Te} を AlexNet ネットワークおよび VGG19 ネットワークの FC - 6 層に対して個別に投影し、対応する特徴を得た。これらを F_{Te_A} および F_{Te_V} とする。次いで、特徴連結を用いてこれらの特徴を組み合わせ、特に P - CRC を用いた試験特徴ベクトルとして使用する1つのベクトル $Te_F = [F_{Te_A} \parallel F_{Te_V}]$ を形成し、モーフィング検出スコアを得た。本研究で用いた P - CRC は、以下の式で表される、学習された特徴ベクトル対プローブ特徴ベクトルに対する正則化最小二乗 (Regularised Least Square Regression: LSR) を使用する。

【0106】

【数2】

$$\hat{F} = \operatorname{argmin}_{\alpha} \|Te_F - \mathcal{D}\alpha\|_2^2 + \lambda \|\alpha\|_2^2 \quad (2)$$

式中、 Te_F は試験画像の特徴ベクトルであり、 \mathcal{D} は Tr_F を用いた学習済み協調部分空間辞書 (collaborative subspace dictionary)、 α は係数ベクトル、 λ は正則化パラメータである。得られた距離は、モーフィング顔検出性能 (以下の実験に基づく図9の記載を参照) を得るためのモーフィング検出スコアとして使用される。

【0107】

[実験]

104個の対象で構成された一般に入手可能な顔データベースを用いて生成されたモーフィング顔画像データベースについて、第2の実施形態を評価した。モーフィング顔画像は、GNU画像操作プログラムv2.8 (GIMP) および GIMP アニメーションパッケージ (GAP) ツールなどの無料で入手可能なソフトウェアパッケージを用い、標識点の位置合わせを手動で行い、高品質モーフィング画像を得た。次いで、パスポート発行手続の実際のシナリオをシミュレーションするため、デジタルモーフィング画像を印刷・スキャンしたバージョンを生成した。

【0108】

したがって、高品質写真用紙 (300グラム) および印刷解像度 1200 dpi の HP Photosmart 5520 を用い、それぞれのデジタルモーフィング画像を印刷した。次のステップにおいて、2つの異なる種類のスキャナ、すなわち (i) HP の Photo Scanner および (ii) リコーのオフィススキャナを用い、印刷された画像をスキャンした。ePass における顔画像参照に関する ICAO 仕様に従い、印刷された写真を 300 dpi となるようにスキャンした。

【0109】

図8は、デジタルモーフィング画像40 (上の行)、HPのスキャナ41を用いて印刷・スキャンした画像 (中央の行) およびリコーのスキャナ42を用いて印刷・スキャンした画像 (下の行) の例を示す。図8の各行の中央の画像44は、モーフィング顔画像の両側に示す2つの異なる対象43、45を用いて生成されたモーフィング顔画像に対応している。

【0110】

図8から分かるように、知覚的な画質は、印刷・スキャン処理に伴って劣化する。この影響により、検出がより困難なものとなる。本研究で使用するデータベースは、訓練および試験に用いられる2つのばらばらの (あるいは重なり合わない) サブセットに分割された352個の真正顔画像および431個のモーフィング顔画像 (デジタル、HPを用いた印刷・スキャン、リコーを用いた印刷・スキャンなどの3つのバージョン全てについて) を有する。訓練セットおよび試験セットについて、サンプルの分布は上記表2に示されている。

【0111】

10

20

30

40

50

モーフィング顔画像を生成する処理は、真正画像の質と比べて異なる質をもたらす画像に対する一連の前処理操作および後処理操作を含む。分類部がモーフィングの影響ではなく、質ノ圧縮差を検出する場合があるため、このような差分的な質指標は、モーフィング検出性能を偏らせ得る。したがって、本発明者らは、真正顔画像およびモーフィング顔画像の両方に対して前処理操作および後処理操作の同等のシーケンスを適用することにより、モーフィング画像と真正顔画像が同じ質を有するよう、さらに留意した。

【0112】

I S O / I E C 3 0 1 0 7 - 3 に従って以下のように規定される攻撃提示分類エラー率 (Attack Presentation Classification Error Rate : A P C E R) と真正提示分類エラー率 (Bona Fide Presentation Classification Error Rate : B P C E R) との相関を示す検出誤りトレードオフ (Detection Error Trade-off : D E T) プロットを用い、モーフィング顔検出アルゴリズムの性能を報告する。

- ・ A P C E R : 特定のシナリオにおいて、攻撃提示 (すなわち、モーフィングされた顔サンプル) が誤って真正提示として分類される割合。

- ・ B P C E R : 特定のシナリオにおいて、真正な提示が誤って提示攻撃 (すなわち、モーフィング画像) として分類される割合。

加えて、5 % および 1 0 % での固定 A P C E を設定し、対応する B P C E R を報告することにより、結果を報告する。モーフィング顔検出アルゴリズムによって得られる A P C E R および B P C E R の両方の値が低いほど、最も優れた検出性能を示す。

【0113】

最新のモーフィング顔検出アルゴリズムを用いた本発明者らのベンチマークにおける方法のロバスト性を評価するため、2 つの異なる実験を行った。

【0114】

実験 1 : 本実験においては、デジタル顔画像データベースおよび印刷・スキャン済みモーフィング顔画像データベースについて、提案される方法の検出性能を最新の方法と共に個別に評価した。本実験では、訓練データおよび試験データが同じソース (デジタル、印刷・スキャン済み (H P) または印刷・スキャン済み (リコー)) からのものとなるように、検出スキームの性能を分析した。したがって、検出方法では、検出される印刷・スキャン済みモーフィング顔画像の生成に使用するプリンタ / スキャナの種類の種類が予めわかっている。

【0115】

実験 2 : 本実験は、本実施形態のモーフィング顔検出アルゴリズムを含むモーフィング顔検出アルゴリズムのロバスト性を試験するために設計されたものである。モーフィング検出方法は、異なるデータソースに対して訓練および試験した。例えば、モーフィング顔検出アルゴリズムをデータベースのデジタルバージョンに対して訓練し、印刷スキャン済みバージョン (H P またはリコー) に対して試験した。本実験により、モーフィング検出アルゴリズムの汎化能力が示された。さらに、世界中の異なるパスポート発行当局で異なるスキャナが使用されているので、実験 2 はまた、モーフィング顔画像の検出の現実的なシナリオにも対処している。

【0116】

以下の表 4 は、デジタルサブセット、H P 印刷・スキャン済みサブセットおよびリコー印刷・スキャン済みサブセットに対して個別に評価された 4 つの異なる最新アルゴリズムを用いたベンチマークにおける提案されるスキームの定量的な結果を示している。

【0117】

10

20

30

40

【表 4】

データセット	アルゴリズム	D-EER (%)	BPCER (%) a t	
			APCER = 10%	APCER = 5%
デジタル	LBP-SVM [10]	29.28	44.52	56.84
	LPQ-SVM [10]	26.12	42.46	52.73
	IG-SVM [10]	21.88	30.13	41.78
	BSIF-SVM [10]	22.70	38.25	49.31
	提案されるスキーム	8.23	7.53	14.38
HP 印刷・スキャン	LBP-SVM [10]	34.35	61.64	69.17
	LPQ-SVM [10]	49.82	92.46	94.52
	IG-SVM [10]	38.35	73.28	80.13
	BSIF-SVM [10]	26.12	45.89	55.47
	提案されるスキーム	17.64	32.87	41.78
リコー 印刷・スキャン	LBP-SVM [10]	22.70	58.21	42.46
	LPQ-SVM [10]	39.18	74.65	82.87
	IG-SVM [10]	34.35	52.05	63.69
	BSIF-SVM [10]	23.29	43.83	54.79
	提案されるスキーム	12.47	16.43	28.76

表 4

【 0 1 1 8 】

図 9 は、3 つのモーフィング顔データタイプすべてに対する画像勾配 (IG-SVM) に基づく提案される方法および最新の方法の性能を示す DET 曲線を示している。簡略化のため、2 本の DET 曲線のみを含んでいるが、すべての結果の詳細は上記表 4 に示されている。以下は、実験 1 から得られる主な所見である。

・第 2 の実施形態のシステムは、モーフィング顔データベースのデジタルバージョンおよび印刷・スキャン済みバージョンの両方に対する 3 つの個別評価すべてにおいて、最も優れた検出性能を示した。同システムは、デジタル画像データベースに対して最小の検出等価エラー率 (D-EER (%)) 8.23%、印刷・スキャン済み画像データベース (HP) に対して D-EER 17.64%、印刷・スキャン済み画像データベース (リコー) に対して D-EER 12.47% を示した。2 番目に優れた検出性能は、BSIF-SVM スキームについて認められる。

・全体として、微小テクスチャスキームに基づく最新スキームの検出性能は、データベースの 3 つのバージョンすべてに対して劣化した性能を示している。

・第 2 の実施形態のシステムを含むモーフィング検出アルゴリズムの検出性能は、データベースのデジタルバージョンに対して、印刷・スキャン済みバージョンと比べて劣化していることが示されている。しかし、リコーを用いて生成された印刷・スキャン済みデータベースに対する本実施形態の性能は、D-EER が 12.47%、APCER = 5% での BPCER = 16.43% という妥当な性能を示していることは興味深い。このことは、デジタルモーフィング顔画像および印刷・スキャン済みモーフィング顔画像の両方の検出に対する提案されるスキームのロバスト性を示している。

・スキャナの種類もまた、モーフィング顔検出アルゴリズムの性能に影響を及ぼす。上記実験により、高品質の印刷写真をフォトスキャナと共に用いることによってモーフィング検出がより困難となることが明らかとなった。このことは上記実験において証明されてい

る。なぜなら、本実施形態を含むモーフィング顔検出アルゴリズムの性能は、リコー製オフィススキャナを用いて生成された印刷・スキャン済みデータベースと比べて、HP製フォトスキャナを用いて生成された印刷・スキャン済みデータベースに対して高いエラー率を示しているからである。

【0119】

実験2に従ったデータベース交差評価に対する提案されるスキームおよび最新のスキームの性能を以下の表5に示す。

【0120】

【表5】

訓練セット	試験セット	アルゴリズム	D-EER (%)	BPCER (%) at	
				APCER = 10%	APCER = 5%
デジタル	印刷・スキャン (HP)	BSIF-SVM	26.70	48.63	56.16
		提案されるスキーム	15.05	24.65	39.72
デジタル	印刷・スキャン (リコー)	BSIF-SVM	27.53	57.53	65.75
		提案されるスキーム	15.05	17.80	28.08
印刷・スキャン (HP)	デジタル	BSIF-SVM	21.29	40.41	52.05
		提案されるスキーム	19.88	34.24	45.89
印刷・スキャン (HP)	印刷・スキャン (リコー)	BSIF-SVM	28.11	58.90	67.80
		提案されるスキーム	13.06	21.23	28.76
印刷・スキャン (リコー)	印刷・スキャン (HP)	BSIF-SVM	27.53	60.95	65.75
		提案されるスキーム	19.05	28.08	40.41
印刷・スキャン (リコー)	デジタル	BSIF-SVM	23.29	34.93	52.05
		提案されるスキーム	20.71	30.13	37.67

表5

【0121】

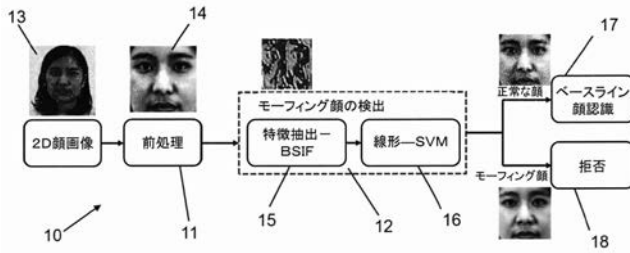
簡略を期すため、提案されるスキームの結果は、BSIF-SVMに基づく2番目に優れた方法と共に示されている。以下は、実験2から推論される主要な所見である。

・モーフィング検出アルゴリズムの性能は、実験1で得られた結果のそれと比較して劣化している。

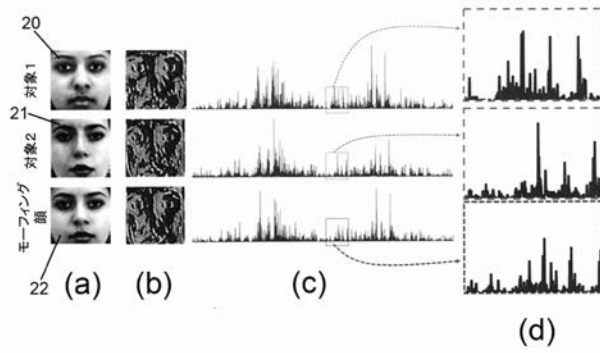
・最も優れた検出性能は、最新のスキームと比べて向上した性能を顕著に示す本実施形態について認められる。

・本実施形態の最も優れた検出性能は、訓練データが、モーフィング顔データベースのデジタルバージョンに対応している場合およびデータベースの印刷・スキャン済みバージョンを試験する場合に観察される。このことは、最新スキームと比較した場合の本実施形態のロバスト性を示している。したがって、広範にわたる実験に基づき、第2の実施形態のシステムは、異なる種類のスキャナから生成されたデジタルバージョンおよび印刷・スキャン済みバージョンの両方におけるモーフィング顔画像の検出の最も優れた方法であることが明らかとなった。同システムはまた、実験2で示されたプロトコルに従ったデータベース交差評価に対して最も優れた性能を示しており、他の最新の方法と比較した場合の提案される方法の適用可能性を示すものである。

【図 1】



【図 2】



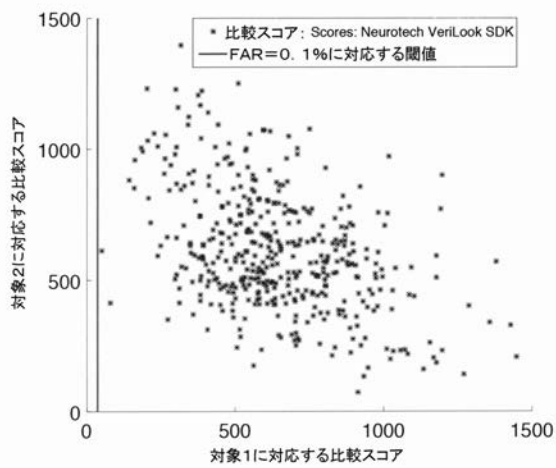
【図 3】



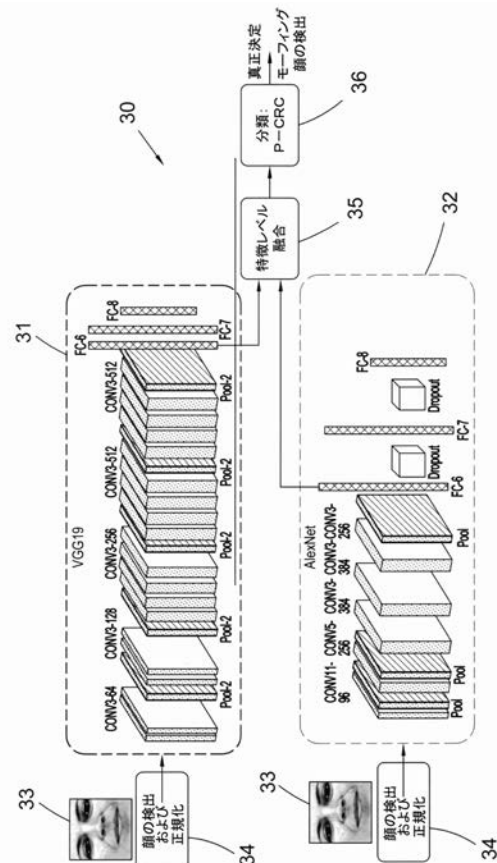
【図 4】



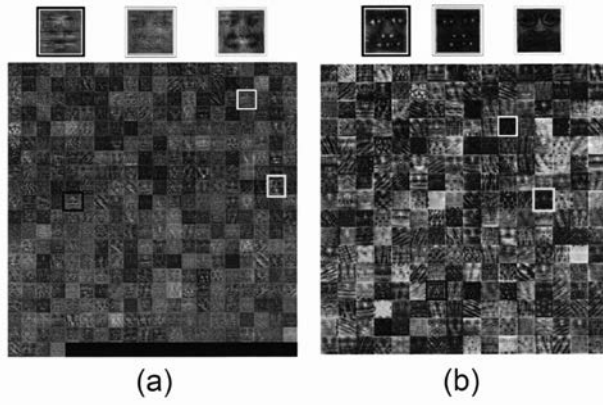
【図 5】



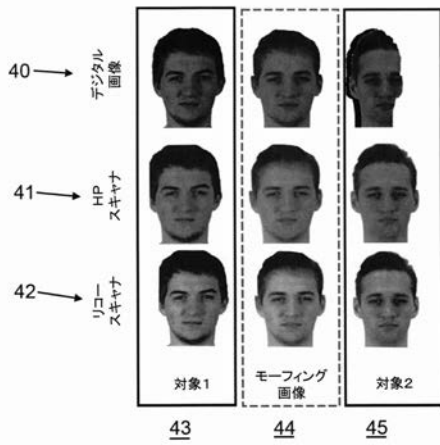
【図 6】



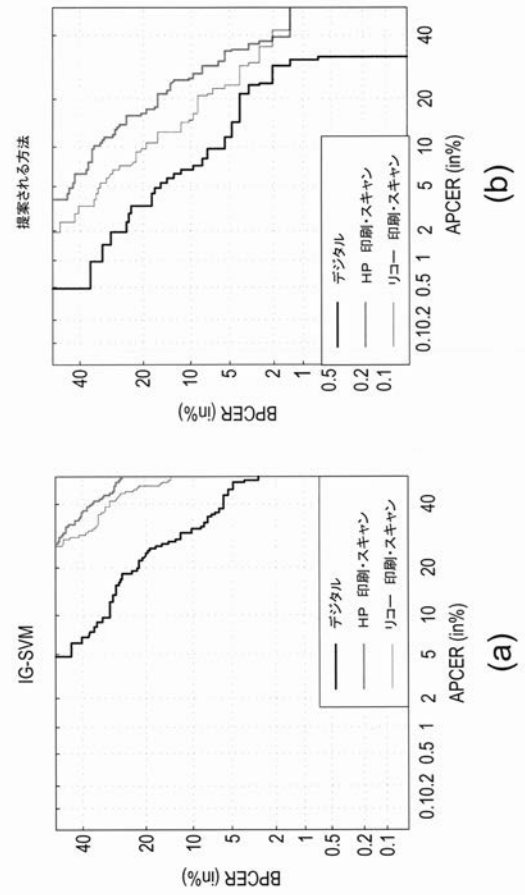
【図 7】



【図 8】



【図 9】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2018/067692

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06K9/00 G06K9/46 G06K9/62
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAGHAVENDRA R ET AL: "Detecting morphed face images", 2016 IEEE 8TH INTERNATIONAL CONFERENCE ON BIOMETRICS THEORY, APPLICATIONS AND SYSTEMS (BTAS), IEEE, 6 September 2016 (2016-09-06), pages 1-7, XP033035149, DOI: 10.1109/BTAS.2016.7791169 [retrieved on 2016-12-19] the whole document ----- -/--	1-25

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

28 September 2018

Date of mailing of the international search report

08/10/2018

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Bouganis, Alexandros

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2018/067692

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BHARATI APARNA ET AL: "Detecting Facial Retouching Using Supervised Deep Learning", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE, PISCATAWAY, NJ, US, vol. 11, no. 9, 1 September 2016 (2016-09-01), pages 1903-1913, XP011613776, ISSN: 1556-6013, DOI: 10.1109/TIFS.2016.2561898 [retrieved on 2016-06-13] abstract page 1903, right-hand column section IV	1-25
X,P	----- RAGHAVENDRA R ET AL: "Transferable Deep-CNN Features for Detecting Digital and Print-Scanned Morphed Face Images", 2017 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), IEEE, 21 July 2017 (2017-07-21), pages 1822-1830, XP033145970, DOI: 10.1109/CVPRW.2017.228 [retrieved on 2017-08-22] the whole document	1-25
A	----- ERGUN HILAL ET AL: "Fusing Deep Convolutional Networks for Large Scale Visual Concept Classification", 2016 IEEE SECOND INTERNATIONAL CONFERENCE ON MULTIMEDIA BIG DATA (BIGMM), IEEE, 20 April 2016 (2016-04-20), pages 210-213, XP032946001, DOI: 10.1109/BIGMM.2016.86 [retrieved on 2016-08-16] abstract section III.B -----	9-19,21

フロントページの続き

(81)指定国・地域 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT

(72)発明者 ヴェンカテッシュ、スシュマ

ノルウェー、 2 8 1 6 イェービク、 ヨハン ファルクバルゲ ヴェグ 9 エフ

(72)発明者 ブッシュ、クリストフ

ノルウェー、 2 8 2 1 イェービク、 タルンヴェイエン 7

Fターム(参考) 5L096 AA02 AA06 EA13 FA32 FA34 FA35 FA69 GA30 GA32 GA55
HA11 KA04 MA07

【要約の続き】

像を分類するように訓練された機械学習システムを含む。

【選択図】 図6