2023年12月21日

顔認識技術の最新動向と犯罪捜査やテロ対策に向けた活用

澤田雅之技術士事務所（電気電子部門）所長
元警察大学校警察情報通信研究センター所長
澤田　雅之

【 目 次 】

Ⅰ 「人の目」を遥かに凌駕する顔認識技術

Ⅱ 犯罪捜査やテロ対策に向けた顔認識技術の活用方法

Ⅲ 海外における、犯罪捜査やテロ対策に向けた顔認識技術の利用動向

Ⅳ ディープラーニングによる顔認識技術

Ⅴ 米国立標準技術研究所の顔認識技術に係るベンダーテスト

Ⅵ 顔認識技術における人種バイアスの問題とその解決方法

☆ おことわり ☆

本資料の中で用いた全ての顔写真は、正当な理由・目的のもとにインターネット上に公開された顔写真を引用したものです。

Ⅰ

「人の目」を遥かに凌駕する顔認識技術

＊ 今日では、カメラ映像のリアルタイム顔認識も可能 ＊

Ⅰ－１

整形手術や長期経年変化で
別人の印象となった顔画像との照合事例

整形手術で別人の印象となった顔画像の照合事例

(出典)
https://search.yahoo.co.jp/image/search?p=□□□□&aq=1&ai=A45UppZZQ.qWtYro2JMNfA&ts=7671&ei=UTF-8&fr=top_ga1_sa

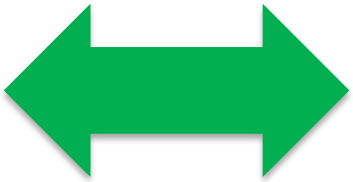
顔認識技術を用いて、左の画像で右の画像を、
数十万枚の中から類似度第１位*に*瞬時検索

経年変化で別人の印象となった顔画像の照合事例

（出典）
https://search.yahoo.co.jp/image/search;_ylt=A2RiOl7hIzleUj4ASCKU3uV7?p=◯◯◯&aq=-1&oq=&ei=UTF-8#mode%3Dsearch

顔認識技術を用いて、左の画像で右の画像を、
数十万枚の中から類似度第１位*に*瞬時検索

経年変化で別人の印象となった顔画像の照合事例



（出典）
https://search.yahoo.co.jp/image/search;_ylt=A2RiOl7hIzleUj4ASCKU3uV7?p=〇〇〇&aq=-1&oq=&ei=UTF-8#mode%3Dsearch

顔認識技術を用いて、左の画像で右の画像を、

数十万枚の中から類似度約百位に***瞬時検索***

→ ***識別精度は、顔画像の品質次第***

Ｉ－２

検索･照合速度は超高速

顔認識技術の検索・照合速度は超高速

数十万枚の顔画像の中から類似度第1位を
瞬時に検索・抽出

顔画像同士のマッチングではなく、各顔画像から生成した「顔特徴ベクトル」間の距離計算を行うことにより、類似度を算定

＊＊ ベクトル間の距離計算は超高速 ＊＊

Ⅰ－３

顔画像品質と識別精度との関係

顔画像品質と識別精度との関係

顔画像の緻密さ

目間画素数（両目の中心を結ぶ直線上の撮像素子数に相当）が数十画素あれば十分

➡ 数百画素を確保しても、識別精度の向上には繋がらない。

顔画像の鮮明さ

ブレ、ボケ、ノイズ、低コントラストが識別精度を劣化させる。

顔の撮影角度

上下方向は30度程度まで、左右方向は45度程度までOK

➡ 角度が大きいほど、他の劣化要因への余裕度が減少

顔の表情・経年変化・整形手術・メガネ等の有無

カメラ側の工夫では対処不可能な要因については、顔識別のアルゴリズムで対処

➡ 識別精度劣化の主要因ではなくなっている。

Ⅰ－４

顔認識技術の動作原理

ディープラーニングで識別の精度と速度が飛躍的に向上

顔認識技術の動作原理（ディープラーニングを用いない場合）

＊＊＊ カメラ映像のリアルタイムな顔認識は困難 ＊＊＊

顔画像の検出

顔の基本的な構造を表現した「顔パターン画像（大きさ可変）」を用いて、フレーム画像等の2次元静止画像を隅から隅までスキャン → 2次元静止画像の中から「顔パターン画像」に合致する矩形領域を「顔画像」として検出 → **動画には対応困難**

顔画像からの特徴抽出

検出して切り出した顔画像から、瞳の中心や鼻の先端等の位置を精密に求め、主成分分析等の手法で作成した「顔特徴抽出フィルタ」を適用 → 検索・照合に用いる高次元（顔特徴抽出フィルタ数に応じた次元）の「顔特徴ベクトル」を生成 → **動画には対応困難**

顔画像の検索・照合

顔画像の検索・照合は、検索・照合したい顔画像と、検索・照合の対象となる顔画像との間の類似度を算定して行う。 → 類似度は、顔画像から生成した「顔特徴ベクトル」間の距離計算により算出

顔認識技術の動作原理（ディープラーニングを用いた場合）

＊＊＊ カメラ映像のリアルタイムな顔認識が可能 ＊＊＊

顔画像の検出（R-CNNのディープラーニングを利用）

R-CNN（Ⅳ章で説明）を用いて、フレーム画像等の2次元静止画像の中に写っている「人の顔」を全て検出 → 「人の顔」が写っている領域を四角形の枠で個々に特定し、「人の顔」らしさを表す信頼度の数値を枠の上辺に表示 → **動画に対応可能**

顔画像からの特徴抽出（CNNのディープラーニングを利用）

CNN（Ⅳ章で説明）を用いて、R-CNNで検出して切り出した顔画像から様々な特徴を抽出し、それらを高次元の数値ベクトルに集約することにより「顔特徴ベクトル」を生成 → **動画に対応可能**

顔画像の検索・照合

顔画像の検索・照合は、検索・照合したい顔画像と、検索・照合の対象となる顔画像との間の類似度を算定して行う。 → 類似度は、顔画像から生成した「顔特徴ベクトル」間の距離計算により算出

Ⅱ

犯罪捜査やテロ対策に向けた顔認識技術の活用方法

＊ 顔認識技術の精度向上と顔撮像カメラの画質向上が鍵 ＊

犯罪捜査やテロ対策に向けた顔認識技術の活用方法

被疑者写真に基づく容疑者の身元割り出し
（顔認識技術の遡及的な利用）

指名手配写真に基づく指名手配犯の発見
（顔認識技術のリアルタイムな利用）

Ⅱ－１

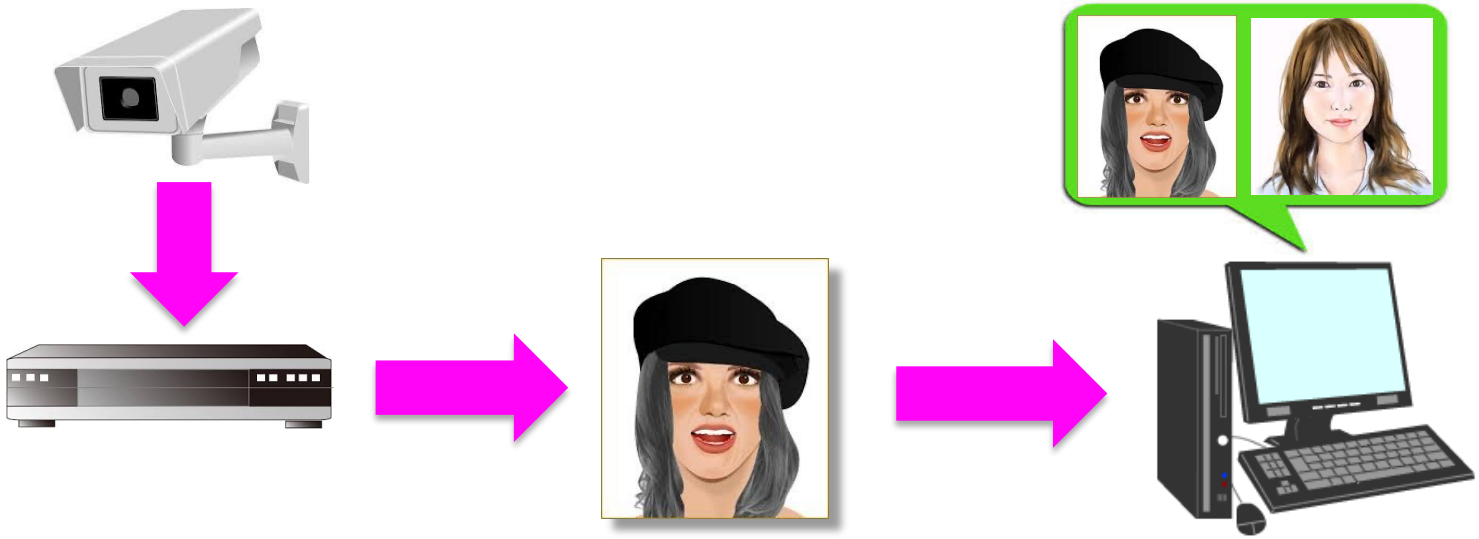
被疑者写真に基づく容疑者の身元割り出し
（顔認識技術の遡及的な利用）

被疑者写真に基づく容疑者の身元割り出し

容疑者の身元割り出しを目的として、**防犯カメラの録画映像の中から「人の目」で選定して切り出した「容疑者の遺留顔画像」を用いて、被疑者写真データベースと照合**する。 ➡ 類似度が高い順にリストアップされた被疑者写真と容疑者の遺留顔画像を「人の目」で見比べて、同一性を判定する。

近年では、

防犯カメラのデジタル化が進んで遺留顔画像の品質が格段に向上したことと、ディープラーニングにより顔認識技術の識別精度が向上したことが相俟って、**容疑者の身元確認に顔認識技術を用いるコストパフォーマンスが大幅にアップ**している。



警察のHPで公開された遺留顔画像（アナログカメラ）

＊ 2013年、デジタル防犯カメラの出荷台数がアナログ防犯カメラを上回った。＊

2012年に公開

2011年に公開

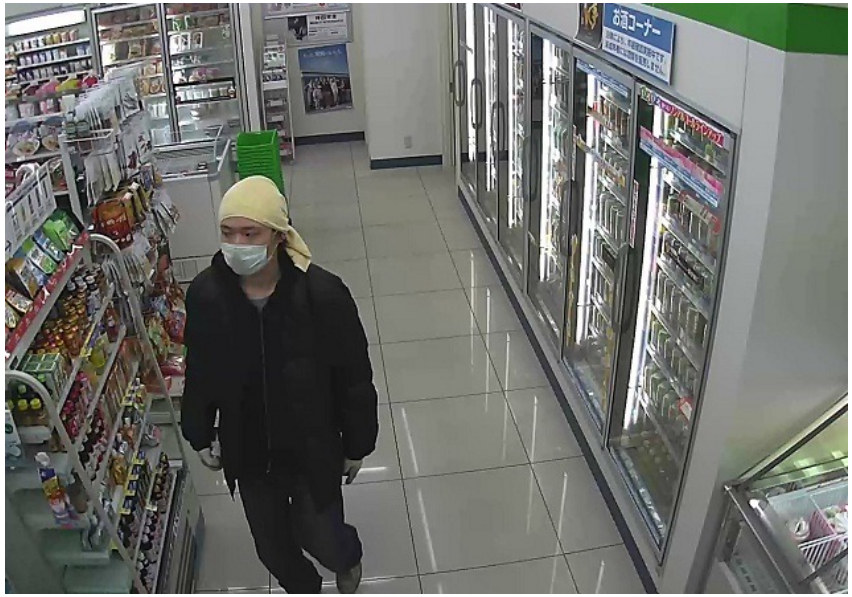
2006年に公開

出典：2023年11月現在の警視庁のHP

警察のHPで公開された遺留顔画像（デジタルカメラ）

＊ 2013年、デジタル防犯カメラの出荷台数がアナログ防犯カメラを上回った。＊

いずれも、<u>2015年に公開された画像</u>



出典：2023年11月現在の滋賀県警のHP

Ⅱ－２

指名手配写真に基づく指名手配犯の発見
（顔認識技術のリアルタイムな利用）

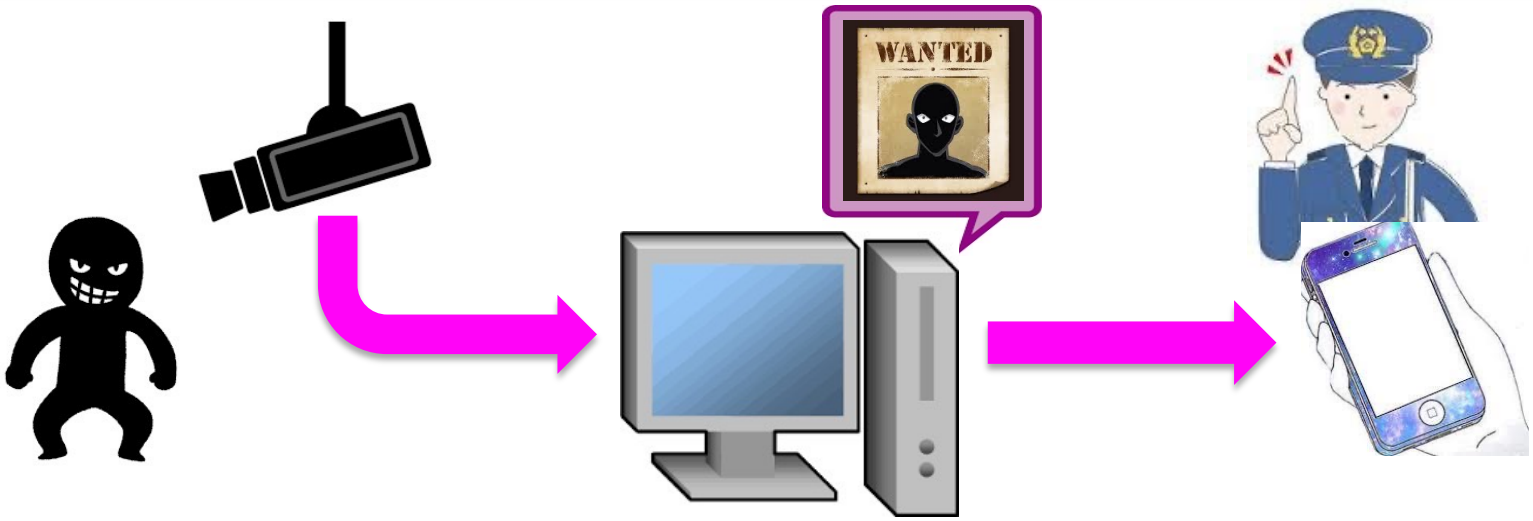
指名手配写真に基づく指名手配犯の発見

監視カメラで捉えた顔画像の中から、指名手配写真データベースに登録された顔画像と同一人物である顔画像を、リアルタイムに発見する。 ➡ 捜査員は、発見した顔画像と指名手配写真を見比べて同一人物であることを確認した上で、職務質問により最終判断する。

本人発見率と他人誤認率

本人発見率と他人誤認率はトレードオフの関係であるため、**捜査員が「指名手配犯との誤認」に振り回されないようにするには、同一性判定の閾値を最適に設定すること**が肝要である。

➡ 近年、監視カメラの高精細デジタル化が進んで顔画像の品質が格段に向上したことと、ディープラーニングにより顔認識技術の識別精度が向上したことが相俟って、**0.01%以下の他人誤認率と90%以上の本人発見率を両立させることは難しくはない。**



Ⅲ

海外における、犯罪捜査やテロ対策に向けた顔認識技術の利用動向

英国における顔認識技術の犯罪捜査等での活用と規制

出典：2018年5月13日付英国INDEPENDENT記事「ロンドン警視庁の顔認識技術は98%不正確、数字で判明」
2023年4月12日付Forbes JAPAN記事「英国警察、カメラでのリアルタイム顔認証の利用を再開へ」
2023年5月25日付日経電子版に掲載された英フィナンシャルタイムズ特約記事「英政府内に顔認証の利用拡大論」
2023年10月31日付JETROビジネス短信「英政府、AI関連施策を相次ぎ発表」

2017年以降、ロンドン警視庁やサウスウェールズ警察では、**リアルタイム顔認識（監視カメラのライブ映像の中から指名手配犯等を自動的に発見する仕組み）**について、大規模な屋外コンサートやプロサッカー試合等の機会に試験運用が繰り返し実施された。しかし、いずれも、**検知に占める誤検知の割合が9割超**であった。 ➡ ちなみに、サウスウェールズ警察では、**「顔認証カメラ作動中」を表示した車両の中で監視カメラを運用**

しかし、

2020年にかけて、リアルタイム顔認識の他人誤認率の高さやプライバシー侵害のおそれが問題視され、警察でのリアルタイム顔認識の試験運用は停止に追い込まれた。 ➡ そこで、ロンドン警視庁とサウスウェールズ警察は、**リアルタイム顔認識の性能評価を英国立物理学研究所に依頼**した。

その結果、

次のページへ

前のページから

その結果、

【英国立物理学研究所の研究報告】

本人発見率と他人誤認率を左右する閾値を適切に設定すれば、89%の本人発見率と0.017%の他人誤認率を両立させ得ることが判明した。また、人種や性別に関しても、統計上有意な偏りは見られなかったとしている。

これを受けて、

ロンドン警視庁では、「誤照合の払拭は困難であるが、リアルタイム顔認識で検知された全ての照合結果は警察官の目視で確認され、警察官が一致すると判断した場合には、その後の職務質問で最終確認される。」として、2023年からリアルタイム顔認識の使用を再開した。 → ロンドン警視庁は、**チャールズ国王の戴冠式（2023年5月）の警備でリアルタイム顔認識を使用**した。

そして、

クリス・フィリップ内務省国務相は、2023年10月29日、英国内の警察に対し、AIの利用拡大を指示。把握済みの犯罪者の追跡のため、**遡及的な顔認識技術（被疑者写真等との照合による容疑者の身元割り出し）の利用拡大**のほか、**リアルタイム顔認識技術のより広範な利用についても推奨**した。

EUにおける顔認識技術の犯罪捜査等での活用と規制

出典：2021年4月22日付朝日新聞デジタル記事「公共空間で顔認証使う捜査、原則禁止　EUがAI規制案」
2023年6月14日付Bloomberg記事「欧州議会、AI規制法案を可決ー顔認証技術など規制」
2022年5月6日付WIRED記事「欧州で賛否、顔認識の国際ネットワークは実現するか」

2021年4月、欧州委員会はAI規制案を公表。 ➡ **公共の場で警察等によるリアルタイム顔認識技術を使った捜査を原則禁止。**公共の場でのリアルタイム顔認識技術の使用は、行方不明の子どもの捜索や差し迫ったテロの脅威への対処などに限定されるべきとした。

そして、

2023年6月、欧州議会は、**公共の場で警察等によるリアルタイム顔認識技術を使った捜査の原則禁止を含むAI規制の最終交渉案を議決。**現在、EU加盟各国との協議が進められている。

他方、

2022年5月、欧州委員会は、**欧州各国の警察による被疑者写真等の顔画像の共有と、顔認識技術の遡及的な利用（被疑者写真等との照合による容疑者の身元割り出し）を認めるようにする提案**を発表。 ➡ 欧州各国の警察は、指紋、DNA、車両所有者に係る情報を、犯罪捜査目的で16年前から互いに共有していることから、犯罪捜査目的での共有対象に、被疑者写真等を新たに加えようとする提案。

米国における顔認識技術の犯罪捜査等での活用と規制

出典：2019年3月30日付CNET Japan記事「顔認識の“人種バイアス問題” なぜ解決が困難なのか」
2020年6月17日付ITmediaビジネスオンライン記事「“顔認識技術を禁止せよ” 黒人差別を受けハイテク大手の対応は？」
2022年2月1日付WIRED記事「顔認識技術は敵か味方か？ 規制と導入の狭間で揺れる米国」
2023年8月21日付MIT Technology Review記事「米国で一時高まった“顔認識規制”の動き その後を追う」

【 米国では、顔認識技術における人種バイアスを問題視 】

2018年、アメリカ自由人権協会は、**警察で用いているamazonの顔認識技術における「人種バイアス問題」を提起。** ⬅ 米連邦議会全議員（非白人が全議員に占める割合は約20%）の顔画像を、amazonの顔認識技術を用いて25,000人分のMUGSHOT（米国の被疑者写真）と照合したところ、被疑者と誤認識された議員28人の約40%は非白人であった。

➡ この結果に基づき、顔認識技術には人種バイアスがあるため、**顔認識技術の遡及的な利用（MUGSHOTや運転免許証写真等との照合による容疑者の身元割り出し）であっても、マイノリティに誤認逮捕等の不利益をもたらしかねない**とされた。

そして、

次のページへ

前のページから

そして、

2020年5月、黒人男性ジョージ・フロイド氏が白人警官に取り押さえられて窒息死した事件を契機に広がったBlack Lives Matter運動以降、**警察での顔認識技術の使用を禁止・制限する動きが加速。**

➡ 2020年6月、IBMは顔認識技術の提供を中止（顔認識技術分野から撤退）すると発表し、同年同月、amazonとMicrosoftは警察への顔認識技術の提供を1年間停止すると発表した。

自治体での規制の動向

2019年、サンフランシスコ市議会は、**顔認識技術の使用を禁止する条例を可決。**以降、2021年までにカリフォルニア州やマサチューセッツ州を中心に20余りの自治体で同様の条例を可決。

➡ しかし、2022年以降は、このような条例を制定する動きは鈍化している。

➡ このような中で、2023年7月、モンタナ州議会は、顔認識技術の使用を制限する州法を可決。**警察が動画映像に対して顔認識技術を用いることを禁止**し、**警察による顔照合（運転免許証写真等も対象）には令状が必要**とした。

中国における顔認識技術の犯罪捜査等での活用と規制

出典：2018年2月8日付MIT Technology Review記事「中国警察がサイバーグラスを導入、顔認識で指名手配犯を逮捕」
2018年12月27日付日テレNEWS NNN記事「コンサート会場で容疑者の逮捕相次ぐ　中国」
2020年9月2日付CNET Japan記事「中国はいかにして顔認識技術で人々の行動を統制しているか」
2021年7月29日付DGLAB記事「中国の顔認証 “野蛮な成長”の終わり　進む法整備」
2021年8月26日付JETROビジネス短信「個人情報保護法が成立、11月1日から施行」
2023年8月9日付REUTERS記事「中国、顔認証技術の利用に関する規則案発表」

中国の「天網」は、数億台規模の固定監視カメラのネットワーク。 → 指名手配犯等の発見にも用いられているが、**固定監視カメラの近傍に警察官が配置され、発見に係る情報が当該警察官に迅速に伝達される仕組みが無ければ、発見した対象の確認と身柄の確保は難しい。**

そこで、

中国では2018年に、香港の人気歌手のコンサート（数万人規模、中国各地で開催）を見に来た約60人の指名手配犯等が、**会場に設置された固定監視カメラのリアルタイム顔認識により発見され、会場に配置された警察官により身柄を確保**された。

スマートグラスによる
リアルタイム顔認識

次のページへ

スマートグラスによるリアルタイム顔認識

前のページから

中国の警察に導入された顔認識スマートグラス
（出典は2018年2月8日付AFP BB News記事）

2018年2月時点で中国の警察は、**リアルタイム顔認識技術を用いて人混みなどで指名手配犯等を発見できるスマートグラスを導入**している。 ➡ スマートグラスに取り付けたカメラで撮影した顔画像を携帯デバイスに送信し、約1万枚の指名手配写真等を約100ミリ秒で検索可能。

中国における規制の動向

次のページへ

前のページから

中国における規制の動向

2021年1月1日、中国の**民法典**が施行された。 ➡ 民法典第111条には、「いかなる自然人の姓名や生年月日、身分証番号、生体識別情報、住所、電話番号、Eメールアドレス、健康情報、行動記録などを不法に取得、使用、加工、伝達してはならない」旨の条文。

そして、

2021年8月20日、中国の全国人民代表大会常務委員会は、**個人情報保護法**を可決し、同年11月1日から施行された。

さらに、

2023年8月8日、中国の国家インターネット情報弁公室は、**顔認識技術を使用する際の安全管理に関する規則**案を発表。 ➡ 顔認識技術の使用は、特定の目的と十分な必要性がある場合に限られ、使用には個人の同意も必要。目的を達成できる他の手段がある場合には、それを優先的に選択すべきとしている。

Ⅳ

ディープラーニングによる顔認識技術

Ⅳ－１

ディープラーニングとは？

ディープラーニングで用いるニューラルネットワーク

ディープラーニングで用いるニューラルネットワークとは、**人の頭脳内部での神経回路網の仕組みと働きを、入力層－隠れ層－出力層として、コンピュータ上で数学的に模したもの。**

具体的には、

各層には、人の脳のニューロン(神経細胞)に相当する多数のノードを配置

→ 各ノードは、人の脳のシナプス(神経細胞間を結ぶ接合部)に相当する信号回路で次の層のノードと結ばれ、ネットワークの結節点を構成

このため、

人の頭脳と同様にニューラルネットワークは、**大量の教材を用いた学習を反復(学習フェーズ)して、隠れ層内の各ノードごとの信号の伝わり具合を少しずつ変えていくことにより、ニューラルネットワーク内部に暗示的に検出・分類・識別等の能力を創出できる。**

→ **推論フェーズで活用**

ディープラーニングとは？

前ページからの再掲

人の頭脳と同様にニューラルネットワークは、大量の教材を用いた学習を反復して、隠れ層内の各ノードごとの信号の伝わり具合を変えていくことにより、ニューラルネットワーク内部に暗示的に検出・分類・識別等の能力を創出できる。

そこで、

ディープラーニングの「ディープ」とは、検出・分類・識別等の能力向上のため、隠れ層を多層化して深く学習（ディープラーニング）できるようにしたことを意味する。

また、

人には得意・不得意の分野や能力の高低があるように、ディープラーニングにも、ニューラルネットワークの構成の仕方（アーキテクチャ）の違いにより得意・不得意の分野が生じ、また、その鍛え方(学習の方法や学習用教材)の違いにより検出・分類・識別等の能力に高低が生じる。

CNN（畳み込みニューラルネットワーク）の全体構成

CNNの基本的な構成は、入力層－隠れ層（畳み込み層－プーリング層－全結合層）－出力層であり、入力された信号は、隠れ層内の各層を含めて、入力層から出力層までのネットワーク全体をフィードフォワードで伝播し、フィードバックされることはない。

【入力層】

2次元静止画像を扱うCNNの入力層は、300画素×300画素、480画素×480画素といった正方形の2次元空間であり、モノクロ画像ではチャネル数は1で、RGB画像ではチャネル数は3となる。

【隠れ層】

精度向上のために隠れ層内は、（　[複数の畳み込み層－プーリング層]－・・－[複数の畳み込み層－プーリング層]－[複数の全結合層]）、といった具合に多層化されている。

【出力層】

クラス分け（分類）の場合には、出力層にはクラス数に応じた分類結果が信頼度の数値として出力される。特徴ベクトルを求める場合には、出力層のノード数に応じた高次元（数百～数千次元）の数値ベクトルが出力される。

CNNをR-CNNに拡張

＊＊ CNNは、1枚の入力画像中の複数の事物を判別できない ＊＊

CNNは、特徴抽出器（畳み込み層とプーリング層）に判別器（全結合層）を接続して構成

➡ 1枚の入力画像中に学習済みの事物が1つだけであれば、その事物の特徴に基づき、判別器で判別（クラス分けや特徴ベクトルの作成）が可能。しかし、学習済みの事物が複数の場合には、CNNの特徴抽出器は複数同時に特徴抽出できるが、CNNの判別器は複数同時に判別できない。

そこで、R-CNNに拡張

CNNの特徴抽出器（1枚の入力画像中に学習済みの事物が複数ある場合でも、それぞれの特徴を同時に抽出可能）の出力内容を、特徴が出現した場所の手掛かりも含めて緻密に活用できるよう、判別器の機能を高度化したものがR-CNN（領域提案できる畳み込みニューラルネットワーク：Faster R-CNN、SSD、YOLO）。

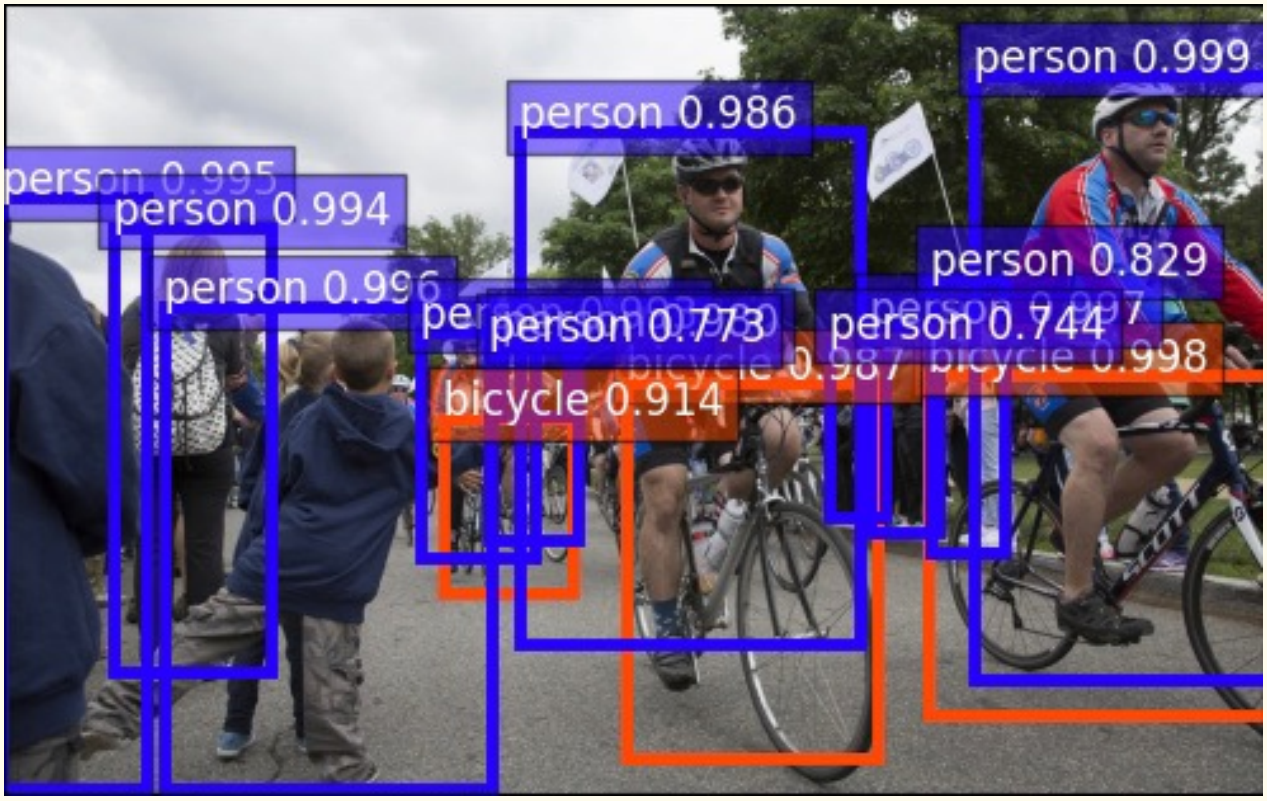
R-CNNの判別器では、

次のページへ

前のページから

R-CNNの判別器では、

1枚の入力画像中に学習済みの事物が複数ある場合でも、R-CNN（FASTER R-CNN、SSD、YOLO）の判別器では、**各事物のクラス分け(分類)を同時に行って信頼度の数値を算出**するとともに、**各事物が画像中に写っている領域を四角形の枠で特定**できる。



人物と自転車を学習させたFASTER R-CNNの推論フェーズ
（出典：https://cv.gluon.ai/build/examples_detection/demo_faster_rcnn.html）

Ⅳ－２

顔画像の検出に向けた
Ｒ－ＣＮＮのディープラーニング

R－CNNによる顔画像の検出（推論フェーズ）

顔画像の検出とは

顔画像の検出とは、フレーム画像等の2次元静止画像の中から「人の顔」を見つけ出して、その写っている画像領域を四角形の枠で囲んで特定すること。

R－CNNでは

R-CNN（領域提案できる畳み込みニューラルネットワーク ： FASTER R-CNN、SSD、YOLO）を用いれば、1枚の入力画像中に複数の「人の顔」が写っている場合でも、**それぞれの「人の顔」が写っている画像領域を四角形の枠で特定**するとともに、**それぞれの画像領域が捉えている事物が「人の顔」であることの信頼度を算出**することができる。

映像からの顔画像検出

映像（2次元静止画像であるフレーム画像が1秒間あたり数十枚連続）の中に多数の「人の顔」が写っている場合でも、**R-CNN（YOLOが最も高速）では、全ての「人の顔」をリアルタイムに検出**できる。

R－CNNによる顔画像の検出（学習フェーズ）

学習用教材の準備その１

検出漏れを防ぐため、様々な条件下（性別、年齢、人種、表情、撮影角度、髪型、髭・帽子・メガネ・マスクの有無など）と環境下（屋内、屋外、日中、夜間照明など）で、「人の顔」が撮影された画像を多数準備する。

誤検出を防ぐため、「人の顔」と誤認する恐れがある物体（人の顔がプリントされたTシャツ、お面、案山子など）が撮影された画像についても、多数準備する。

学習用教材の準備その２

*その１*で準備した画像の全てについて、アノテーションツール（Microsoft　　VoTTなど）を利用して、「人の顔」が写っている画像領域ラベルと「人の顔」であるクラスラベルを設定し、また、「人の顔と誤認する物体」が写っている画像領域ラベルと「人の顔と誤認する物体」であるクラスラベルを設定する。

誤差逆伝播による反復学習

次のページへ

前のページから

誤差逆伝播による反復学習

学習用教材をR-CNNに入力して、「人の顔」については、出力層における「人の顔」クラスでの信頼度の数値と、「人の顔」が写っている画像領域を四角形の枠で囲む精度を少しずつ高めていく。また、「人の顔と誤認する物体」についても、 「人の顔と誤認する物体」クラスでの信頼度の数値と、「人の顔と誤認する物体」が写っている画像領域を四角形の枠で囲む精度を少しずつ高めていく。

➡ 「望ましい結果との誤差」をR-CNNの隠れ層内に逆伝播させて、隠れ層（畳み込み層とプーリング層と全結合層）全体におけるノードへの信号入力時の重み付け値やノードからの信号出力時のバイアス値を少しずつ変化させる。このような学習を反復する。

➡ PyTorch、TensorFlowなどのAIフレームワークを用いれば、学習用教材を一括してセットし、反復学習を自動化することができる。

テストデータによる確認

学習用教材とは別にテストデータを多数準備しておき、反復学習を終了したR-CNNがテストデータから検出した「人の顔」の信頼度の数値や検出漏れ・誤検出の有無について、つまり、「人の顔」の検出精度が目標としたレベルに達しているか否かを確認する。

Ⅳ－３

顔画像の識別に向けた
CNNのディープラーニング

CNNによる顔画像の識別（推論フェーズ）

顔画像の識別とは

顔画像の識別とは、**2次元静止画像として捉えた「人の顔」が、誰の顔であるのかを見分けること。**
例えば、ビデオカメラ映像の中からR-CNNで検出した「人の顔」について、「識別対象者の顔画像を多数登録したデータベース」と照合して、「類似度が極めて高い顔」を顔画像データベースの中から見つけ出せば、誰の顔であるかを見分けられる。

CNNによる顔特徴ベクトルの生成

CNNの特徴抽出器（畳み込み層とプーリング層）を用いて、2次元静止画像として入力された「人の顔」から様々な特徴を抽出し、それらをCNNの判別器（全結合層）で高次元の数値ベクトルに集約することにより、**「顔特徴ベクトル」を生成する。**

「人の顔」の識別は、識別したい顔画像と識別対象となる顔画像との間の**類似度を、それぞれの顔画像から生成した「顔特徴ベクトル」間の距離計算に基づき算出**して行う。

CNNによる顔画像の識別（学習フェーズ）

＊＊ 理想的な「顔特徴ベクトル」の生成を目指して ＊＊

【 理想的な顔特徴ベクトルとは 】

顔画像の識別とは、識別対象顔画像から生成した「顔特徴ベクトル」が分布するベクトル空間において、識別したい顔画像から生成した「顔特徴ベクトル」との距離計算を行うこと。 ➡ 顔の表情や長期経年変化などが影響して「別人」に見える場合でも、**本人の顔画像の「顔特徴ベクトル」は常に近接し、他人の顔画像の「顔特徴ベクトル」とは常に離間するのが、理想的な「顔特徴ベクトル」**

そこで、

次のページへ

前のページから

そこで、

学習用教材の準備その１

多様（性別、年齢、人種など）で多数の人々の顔画像を準備するとともに、同一人物についても様々な条件下（表情、撮影角度、髭の有無、帽子・メガネ・マスクの有無など）で撮影された複数の顔画像を準備する。

学習用教材の準備その２

その1で準備した顔画像の全てについて、アノテーションツール（Microsoft　VoTTなど）を利用して、「同一人物の顔画像であること」を判別するためのIDを、ラベルとして設定する。

誤差逆伝播による反復学習

次のページへ

前のページから

誤差逆伝播による反復学習

任意の同一人物の任意の2枚の顔画像を選択するとともに、任意の別人の任意の1枚の顔画像を選択し、これらの3枚の顔画像を1つのセットとして、CNNの入力層への学習用教材として用いて、出力される3つの「顔特徴ベクトル」を比較する。 → 同一人物の2つの「顔特徴ベクトル」間の距離を少し狭めるとともに、別人の「顔特徴ベクトル」との距離を少し広げるよう、「距離の開き具合」をCNNの隠れ層内に逆伝播させて、隠れ層（畳み込み層とプーリング層と全結合層）全体におけるノードへの信号入力時の重み付け値やノードからの信号出力時のバイアス値を少しずつ変化させる。このような学習を反復する。

→ PyTorch、TensorFlowなどのAIフレームワークを用いれば、学習用教材を一括してセットし、反復学習を自動化することができる。

テストデータによる確認

学習用教材とは別にテストデータを多数準備しておき、反復学習を終了したCNNがテストデータから生成した「顔特徴ベクトル」について、同一人物の「顔特徴ベクトル」は常に近接し、別人の「顔特徴ベクトル」とは常に離間しているか否かを確認する。

V

米国立標準技術研究所の
顔認識技術に係るベンダーテスト

～ ディープラーニングの活用効果が明白 ～

【顔画像及びデータの出典】

出典１：NIST Interagency Report 8009(2014年5月公表)

出典２：NIST Interagency Report 8271(2019年9月公表)

米国立標準技術研究所の顔画像識別技術ベンダーテスト

米国立標準技術研究所（NIST：National Institute of Standards and Technology）では、顔認識技術に係る各種の評価試験を2000年以来実施

中でも、

2013年と2018年に実施した顔認識技術ベンダーテスト（FRVT：Face Recognition Vendor Test）は、**「顔を識別するアルゴリズム」の生成にディープラーニングを活用することによる識別性能の改善効果を理解する上で重要**

その理由は、

2013年のFRVTではディープラーニングを活用したベンダーは皆無。しかし、2018年のFRVTでは多くのベンダーがディープラーニングを活用し、2013年と類似したテスト環境における識別精度が桁違いに向上 → 2013年のFRVTの結果は、「NIST Interagency Report 8009」として2014年5月に公表。2018年のFRVTの結果は、「NIST Interagency Report 8271」として2019年9月に公表。 → **2018年のFRVTは、新たな参加ベンダーを含めて現在も継続実施中**

2013年のFRVTに参加したベンダー

NEC（日本）

東芝（日本）

Ayonix Inc.（日本）

3M/Cogent（米国）

HP/Virage（米国）

Decatur Industries Inc.（米国）

Safran Morpho（フランス）

Cognitec（ドイツ）

Neurotechnology（リトアニア）

清華大学(Prof. Wen)（中国）

清華大学(Prof. Su)（中国）

Beijing Ivsign Technology Co. Ltd.（中国）

中国科学院 Automation（中国）

中国科学院 Computing Technology（中国）

Zhuhai Yisheng Electronics Tech. Co. Ltd.（中国）

JunYu Technology Co. Ltd.（中国）

青字は、2013年と2018年の
両FRVTに参加したベンダー

2018年のFRVTに参加したベンダー（1/2）

NEC（日本）
東芝（日本）
グローリー（日本）
Ayonix Inc.（日本）
Cognitec（ドイツ）
Dermalog（ドイツ）
Smilart（ドイツ）
Idemia（フランス）（旧Safran Morpho）
Thales（フランス）（旧3M/Cogent）
Quantasoft（チェコ）
Eyedea Recognition（チェコ）
Neurotechnology（リトアニア）

Zhuhai Yisheng Electronics Tech. Co. Ltd.（中国）
ハイクビジョン（中国）
メグビー（中国）
センスタイム（中国）
ダーファ（中国）
依図テクノロジー（中国）
KanKan Ai（中国）
中国科学院 Shenzhen Inst. Adv. Tech.（中国）
TongYi Transportation Technology（中国）
Newland Computer Co. Ltd（中国）
Gorilla Technology（台湾）
Alchera（韓国）
Loginface Corp（韓国）

青字は、2013年と2018年の両FRVTに参加したベンダー

2018年のFRVTに参加したベンダー(2/2)

マイクロソフト(米国)
3Divi(米国)
Anke Investments（米国）
Aware（米国）
Camvi Technologies（米国）
Ever AI（米国）
Incode Technologies（米国）
Noblis（米国）
Rank One Computing（米国）
RealNetworks（米国）
Shaman Software（米国）
TigerIT Americas LLC（米国）
Vigilant Solutions（米国）

Alivia/Innovation Sys.（ロシア）
N-Tech Lab（ロシア）
Tevian（ロシア）
Vocord（ロシア）
Microfocus（英国）
Imagus Technology Pty Ltd（オーストラリア）
Innovatrics（スロバキア）
Synesis（クロアチア）
Visidon（フィンランド）
VisionLabs（オランダ）
Lookman Electroplast Industries（インド）

合計：49社

2013年と2018年の両FRVTで共通して用いられた2種類の顔画像

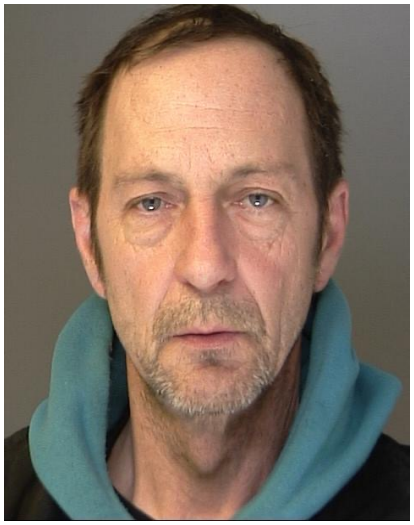
MUGSHOT画像

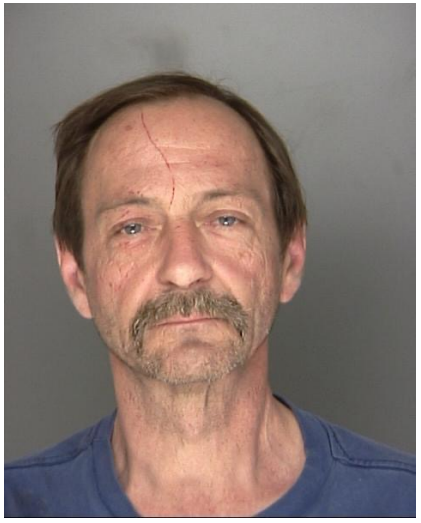
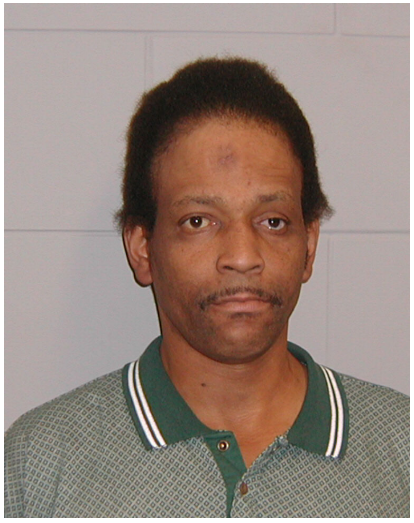
米国の警察の実際の現場で、デジタルスチルカメラにより撮影

➡ 我が国の被疑者写真に相当（高品質な顔画像）

WEBカメラ画像

米国の国境警備隊が勾留した中南米からの不法入国者を、WEBカメラにより撮影 ➡ 国境警備隊職員の指示に基づき、WEBカメラに顔を向けた瞬間を捉えている。（品質が劣る顔画像）

MUGSHOT画像




同一人物の複数のMUGSHOT画像

MUGSHOT画像の品質

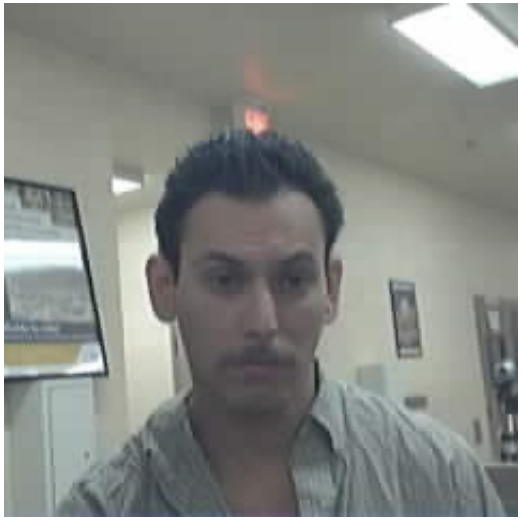
○ オリジナル画像サイズ ：
480×640、240×240、768×960

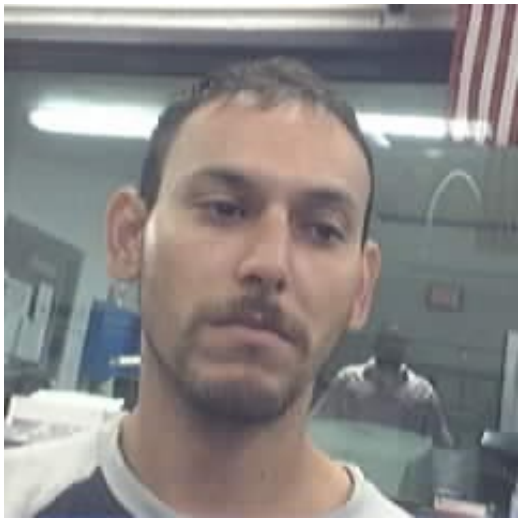
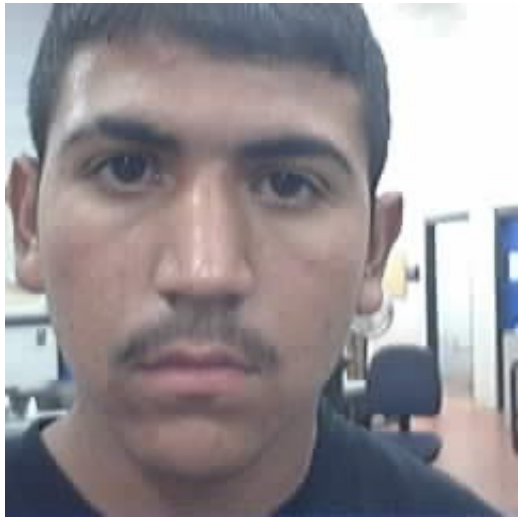
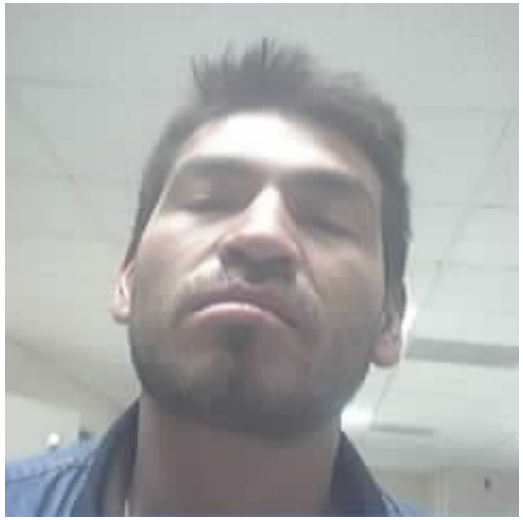
○ JPEG圧縮後の画像情報量 ：
平均48kB

○ 目間画素数 ： **平均107画素で、標準偏差は40画素**

○ ポーズ ： 顔を真正面から捉えた画像が多いが、中には、5度から10度ほど左右を向いている画像がある。また、視線がカメラに向けられていない画像が少なくない。

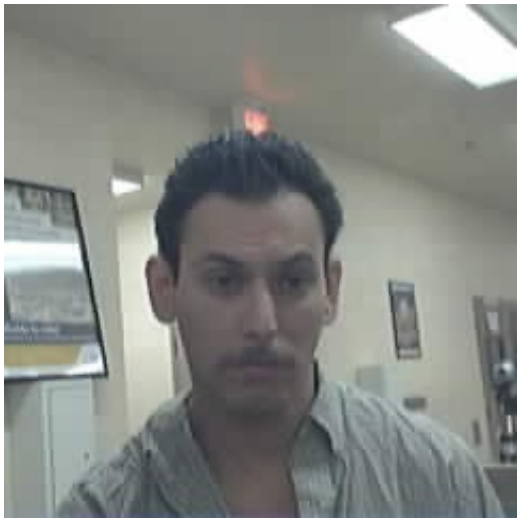
○ 照明 ： 適切な照明のもとで撮影されている画像が多い。



WEBカメラ画像

WEBカメラ画像の品質



○ オリジナル画像サイズ ：

240×240

○ JPEG圧縮後の画像情報量 ：

平均5.7kB

○ 目間画素数 ： **平均45画素で、標準偏差は12画素**

○ ポーズ ： 顔を真正面から捉えた画像は少なく、5度から10度ほど左右を向いている画像や、下から見上げるように捉えた画像が多い。

○ 照明 ： 殆どが室内照明下で撮影されているため、照明が不十分で顔に影ができている画像が大半である。

2013年と2018年の両FRVTで用いられた識別精度評価方法

2013年と2018年のFRVTでは、1対1の顔画像認証（同一人物であるか否かを顔画像により確認すること）の識別精度ではなく、**1対多数の顔画像照合（同一人物の顔画像を多数の顔画像の中から探し出すこと）の識別精度**を、各種試験により評価

そこで、

1対多数の顔画像照合では、探し出そうとする人物の顔画像（検索用顔画像）と、探す対象となる多数の顔画像（被検索顔画像）との間の「類似度」を算出 → この算出は、**検索用顔画像から生成した「顔特徴ベクトル」と、各被検索顔画像から生成した各「顔特徴ベクトル」との、ベクトル間距離を計算**して行う。

このため、

1対多数の顔画像照合の結果は、検索用顔画像との「類似度」が高い順（顔特徴ベクトル間の距離が短い順）に、被検索顔画像がスコア値（顔特徴ベクトル間距離の短さに応じた1～0の正数）とともにリストアップ → 1対多数の顔画像照合での識別精度の評価方法は、**スコア値に対する「閾値」を設けるか否か**により、2種類に大別される。

（1） スコア値に対する「閾値」を設けない識別精度評価方法

1対多数の顔画像照合を行った結果として、検索用顔画像と被検索顔画像との「類似度」のスコア値が高い順に被検索顔画像がリストアップされる。

【 スコア値に対する「閾値」を設けない場合の識別精度評価方法 】

多数の検索用顔画像を用いて1対多数の顔画像照合を行い、その結果として、検索用顔画像と同一人物の被検索顔画像が、TOP1（リストの最上位）にリストアップできなかった割合（「本人見逃し率」）を調べる。

ここで、

スコア値に対する「閾値」を設けなければ、検索用顔画像と同一人物の被検索顔画像が存在しない場合に、TOP1における「他人誤認率」が100%となることに注意を要する。

(2) スコア値に対する「閾値」を設ける識別精度評価方法

1対多数の顔画像照合を行った結果として、検索用顔画像と被検索顔画像との「類似度」のスコア値が高い順に被検索顔画像がリストアップされる。

【 スコア値に対する「閾値」を設ける場合の識別精度評価方法 】

多数の検索用顔画像を用いて1対多数の顔画像照合を行い、その結果として、検索用顔画像と同一人物の被検索顔画像が、設けた「閾値」を上回ってTOP1（リストの最上位）にリストアップできなかった割合（「本人見逃し率」）を調べる。

ここで、

【 「本人見逃し率」と「他人誤認率」は、トレードオフの関係であることに注意 】

「閾値」を高く設定して「他人誤認率」を小さくするほど「本人見逃し率」は大きくなり、「閾値」を低く設定して「他人誤認率」を大きくするほど「本人見逃し率」は小さくなる。 → そこで、「他人誤認率」が1/10、1/100、1/1000となる「閾値」をパラメータとして設定して、各「閾値」での「本人見逃し率」を調べて識別精度を評価する。

Ⅳ－１

2013年FRVT → 2018年FRVT

ディープラーニングが識別精度を飛躍的に向上

2013年のFRVTではディープラーニングの活用は未だ見られないが、2018年のFRVTでは多くのベンダーがディープラーニングを活用して、識別精度を飛躍的に向上させている。

そこで、2013年と2018年の両FRVTに参加した6社【　　NEC、Thales（旧3M/Cogent）、東芝、Neurotech.、Cognitec、Idemia（旧Safran　Morpho）】について、 同様な試験条件下での結果を各社ごとに比較すれば、ディープラーニングの導入効果を調べることができる。

なお、ここでの同様な試験条件とは、160万人分（160万枚）の顔画像データベースに対して、高品質なMUGSHOT画像により検索した場合の「本人見逃し率」と、品質の劣るWEBカメラ画像により検索した場合の「本人見逃し率」を、それぞれ求めるものである。

2013年FRVTと2018年FRVTの検索精度（試験方法）

検索用MUGSHOT画像

2013年のFRVT

検索対象は、160万人分（160万枚）の顔画像データベース（大半はMUGSHOT画像であるが、96,885人分のWEBカメラ画像を含む。）

① 5万人分（5万枚）のMUGSHOT画像（データベース内の同一人物画像とは別画像）で検索し、「閾値を設けないTOP1の本人見逃し率」を算出

② 10,660人分（10,660枚）のWEBカメラ画像（データベース内の同一人物画像とは別画像）で検索し、「閾値を設けないTOP1の本人見逃し率」を算出

検索用WEBカメラ画像

2018年のFRVT

検索対象は、160万人分（160万枚）の顔画像データベース（全てMUGSHOT画像）

① 154,549人分（154,549枚）のMUGSHOT画像（データベース内の同一人物画像とは別画像）で検索し、「閾値を設けないTOP1の本人見逃し率」を算出

② 82,106人分（82,106枚）のWEBカメラ画像（データベース内の同一人物画像とは別画像）で検索し、「閾値を設けないTOP1の本人見逃し率」を算出

2013年FRVTと2018年FRVTの検索精度（比較上の注意点）

2013年と2018年のFRVTで用いられた顔画像データベースはいずれも160万人分（160万枚）であるが、2018年の顔画像データベースは全て高品質なMUGSHOT画像であるのに対して、2013年の顔画像データベースには品質の劣るWEBカメラ画像が全体の６％ほど（160万枚中の96,885枚）含まれている。

しかし、

このような品質の劣る顔画像の混在は、「本人見逃し率」の劣化に繋がる。

そこで、

【 劣化の度合いを推測する目安として 】

2013年のFRVTでは、2万人分（2万枚）の顔画像データベース（全てMUGSHOT画像）を検索対象として、2万人分（2万枚）のMUGSHOT画像（データベース内の同一人物画像とは別画像）で検索し、「閾値を設けないTOP1の本人見逃し率」を算出しているので、その結果を次ページに青字の括弧書きで示す。

2013年FRVTと2018年FRVTの検索精度（試験結果）

閾値を設けない場合のTOP1（類似度ランク1位）の本人見逃し率

NEC	MUGSHOT : 0.041(0.028)	WEBCAM : 0.113
	MUGSHOT : 0.003	WEBCAM : 0.010
Thales *（旧3M/Cogent）*	MUGSHOT : 0.172(0.105)	WEBCAM : 0.364
	MUGSHOT : 0.006	WEBCAM : 0.020
東芝	MUGSHOT : 0.107(0.060)	WEBCAM : 0.237
	MUGSHOT : 0.007	WEBCAM : 0.022
Neurotech.	MUGSHOT : 0.205(0.142)	WEBCAM : 0.702
	MUGSHOT : 0.007	WEBCAM : 0.024
Cognitec	MUGSHOT : 0.136(0.085)	WEBCAM : 0.576
	MUGSHOT : 0.008	WEBCAM : 0.025
Idemia *（旧Safran Morpho）*	MUGSHOT : 0.091(0.068)	WEBCAM : 0.307
	MUGSHOT : 0.009	WEBCAM : 0.032

黒字 : 2013年FRVT
赤字 : 2018年FRVT

青字の括弧内の数字は、2013年FRVTにおいて、2万人分（2万枚）のMUGSHOT画像に対して、本人の別のMUGSHOT画像で、検索を2万回実施した結果（閾値を設けないTOP1の本人見逃し率）を示す。

2013年FRVTと2018年FRVTの検索精度（考察）

2013年と2018年の両FRVTに参加した6社【　　NEC、Thales（旧3M/Cogent）、東芝、Neurotech.、Cognitec、Idemia（旧Safran　Morpho）　】は全て、高品質なMUGSHOT画像による検索時と、品質の劣るWEBカメラ画像による検索時のいずれも、2013年のFRVT（**青字で括弧内に示した数値を含めて**）と較べて2018年のFRVTでは、**「本人見逃し率」の数値に桁違いの改善**が見られる。

ディープラーニングは、2013年のFRVTでは未だ活用されず、2018年のFRVTでは活用が一気に進んだことから、**識別性能が飛躍的に向上した主因はディープラーニングの活用**にあると言える。

Ⅳ－２

2018年FRVT
品質の劣る顔画像に対する識別特性

防犯カメラで捉えた顔画像は、MUGSHOT画像のような高品質な顔画像（つまり、緻密かつ鮮明な無表情顔をほぼ正面から捉えた顔画像）と較べて、緻密さや鮮明さ、撮影角度などの点で「品質」が劣る場合が多い。

このため、高品質なMUGSHOT画像による識別精度と、緻密さや鮮明さ、撮影角度の点で品質が劣るWEBカメラ画像による識別精度を詳しく比較検討すれば、防犯カメラで捉えた顔画像に顔認識技術を適用していく上での有用な知見を得ることができる。

MUGSHOT画像とWEBカメラ画像の検索精度（試験方法その1）

＊＊＊ MUGSHOT画像で検索する場合 ＊＊＊

検索用MUGSHOT画像

「他人誤認率が1/10、1/100、1/1000となる閾値」の求め方

160万人分（160万枚）の顔画像データベース（全てMUGSHOT画像）を検索対象として、331,254人分（331,254枚）のMUGSHOT画像（いずれもデータベース内に同一人物の顔画像が存在しない。）で検索し、「他人誤認率が1/10、1/100、1/1000となる閾値」を求める。

「TOP1における本人見逃し率」の求め方

160万人分（160万枚）の顔画像データベース（全てMUGSHOT画像）を検索対象として、154,549人分（154,549枚）のMUGSHOT画像（データベース内の同一人物画像とは別画像）で検索し、「閾値を設けないTOP1の本人見逃し率」と、「他人誤認率が1/10、1/100、1/1000となる閾値を設けたTOP1の本人見逃し率」を算出する。

MUGSHOT画像とWEBカメラ画像の検索精度（試験方法その2）

＊＊＊ WEBカメラ画像で検索する場合 ＊＊＊

検索用WEBカメラ画像

「他人誤認率が1/10、1/100、1/1000となる閾値」の求め方

160万人分（160万枚）の顔画像データベース（全てMUGSHOT画像）を検索対象として、331,254人分（331,254枚）のWEBカメラ画像（いずれもデータベース内に同一人物の顔画像が存在しない。）で検索し、「他人誤認率が1/10、1/100、1/1000となる閾値」を求める。

「TOP1における本人見逃し率」の求め方

160万人分（160万枚）の顔画像データベース（全てMUGSHOT画像）を検索対象として、82,106人分（82,106枚）のWEBカメラ画像（データベース内の同一人物画像とは別画像）を用いて検索し、「閾値を設けないTOP1の本人見逃し率」と、「他人誤認率が1/10、1/100、1/1000となる閾値を設けたTOP1の本人見逃し率」を算出する。

MUGSHOT画像とWEBカメラ画像の検索精度（試験結果その1）

閾値を設けない場合のTOP1（類似度ランク1位）の本人見逃し率

NEC	MUGSHOT画像	0.003	WEBカメラ画像	0.010
佐図	MUGSHOT画像	0.004	WEBカメラ画像	0.008
マイクロソフト	MUGSHOT画像	0.003	WEBカメラ画像	0.011
センスタイム	MUGSHOT画像	0.005	WEBカメラ画像	0.016
VisionLabs	MUGSHOT画像	0.003	WEBカメラ画像	0.015
N-Tech Lab	MUGSHOT画像	0.006	WEBカメラ画像	0.017
Lookman	MUGSHOT画像	0.012	WEBカメラ画像	0.039
Alivia	MUGSHOT画像	0.007	WEBカメラ画像	0.023
Neurotech.	MUGSHOT画像	0.007	WEBカメラ画像	0.024
東芝	MUGSHOT画像	0.007	WEBカメラ画像	0.022

MUGSHOT画像とWEBカメラ画像の検索精度（試験結果その2）

他人誤認率を一定(緑字1/10、赤字1/100、青字1/1000)としたTOP1の本人見逃し率

NEC	MUGSHOT 0.003、0.004、0.004	WEBカメラ 0.011、0.013、0.017
依図	MUGSHOT 0.004、0.007、0.012	WEBカメラ 0.011、0.017、0.027
マイクロソフト	MUGSHOT 0.004、0.008、0.014	WEBカメラ 0.016、0.024、0.037
センスタイム	MUGSHOT 0.007、0.012、0.023	WEBカメラ 0.025、0.040、0.063
VisionLabs	MUGSHOT 0.005、0.012、0.029	WEBカメラ 0.025、0.051、0.090
N-Tech Lab	MUGSHOT 0.010、0.021、0.039	WEBカメラ 0.032、0.059、0.094
Lookman	MUGSHOT 0.016、0.027、0.047	WEBカメラ 0.052、0.075、0.105
Alivia	MUGSHOT 0.012、0.027、0.062	WEBカメラ 0.039、0.068、0.107
Neurotech.	MUGSHOT 0.012、0.025、0.056	WEBカメラ 0.042、0.074、0.130
東芝	MUGSHOT 0.013、0.029、0.065	WEBカメラ 0.041、0.074、0.118

MUGSHOT画像とWEBカメラ画像の検索精度（考察）

閾値を設けない場合のTOP1における本人見逃し率

試験結果上位10社の「閾値を設けないTOP1の本人見逃し率」は、MUGSHOT画像による検索では0.003～0.012で、WEBカメラ画像による検索では0.008～0.039 ➡ 高品質画像による検索時と較べて、品質が劣る画像による検索時には、10社の全てで「本人見逃し率」が数倍に劣化 ➡ 顔認識技術を用いる場合には、検索用及び被検索用の顔画像のいずれも、ほぼ正面から捉えた緻密かつ鮮明な顔画像が望ましい。

閾値を設けて他人誤認率を1/1000とした場合の本人見逃し率

試験結果上位10社の「他人誤認率が1/1000となる閾値を設けた本人見逃し率」は、MUGSHOT画像による検索では0.004～0.065で、WEBカメラ画像による検索では0.017～0.130 ➡ 閾値を設けない場合と較べて、閾値を設けた場合には、10社の全てで「本人見逃し率」が劣化 ➡ 他人誤認率を極力低くする閾値を設ける場合には、検索用顔画像及び被検索用顔画像のいずれも、ほぼ正面から捉えた緻密かつ鮮明な顔画像であることが欠かせない。

Ⅳ－３

2018年FRVT
顔の長期経年変化に対する識別特性

被疑者写真データベースには、20年以上前に撮影された顔画像も多く含まれている。このため、防犯カメラ等に遺留された犯人の顔画像を元に被疑者写真データベースを検索する場合には、長期にわたる顔の経年変化で容貌が大きく変わった場合でも、同一人物か否かを高精度に判別できることが望まれる。

そこで、2018年のFRVTでは、高品質なMUGSHOT画像を用いて、最大で18年に及ぶ「顔の経年変化」が識別精度に及ぼす影響を調べている。

顔の経年変化に対する識別特性（試験方法）

「他人誤認率が1/1000となる閾値」の求め方

300万人分（300万枚）の顔画像データベース（全てMUGSHOT画像）を検索対象として、331,254人分（331,254枚）のMUGSHOT画像（いずれもデータベース内に同一人物の顔画像が存在しない。）で検索し、「他人誤認率が1/1000となる閾値」を求める。

「TOP1における本人見逃し率」の求め方

１人につき複数枚のMUGSHOT画像がある3,068,801人の累犯者について、最も古いMUGSHOT画像のみを3,068,801枚集めて検索対象の顔画像データベースとし、それ以外のMUGSHOT画像（2,853,221人分の10,951,064枚）で10,951,064回検索する。

そして、最も古いMUGSHOT画像と検索に用いたMUGSHOT画像との間の経過年数（例えば、【2～4年の隔たり】や【10～12年の隔たり】などをパラメータとして、「閾値を設けないTOP１の本人見逃し率」と、「他人誤認率が1/1000となる閾値を設けたTOP１の本人見逃し率」を算出する。

顔の経年変化に対する識別特性（試験方法）

検索用MUGSHOT画像

検索用MUGSHOT画像

検索用MUGSHOT画像

最も古いMUGSHOT画像

１人につき複数枚のMUGSHOT画像がある場合に、最も古いMUGSHOT画像のみを3,068,801枚集めたものを検索対象として、それ以外のMUGSHOT画像(2,853,221人分の10,951,064枚)で検索

顔の経年変化に対する識別特性（試験結果その1）

閾値を設けない場合のTOP1（類似度ランク1位）の本人見逃し率

【0～2年前】、【2～4年前】、【4～6年前】、【6～8年前】、【8～10年前】、【10～12年前】、【12～14年前】、【14～18年前】

NEC	0.003	0.004	0.004	0.004	0.004	0.005	0.006	0.004
依図	0.006	0.008	0.008	0.008	0.009	0.011	0.015	0.021
マイクロソフト	0.003	0.005	0.006	0.007	0.009	0.010	0.013	0.016
VisionLabs	0.004	0.006	0.008	0.009	0.011	0.013	0.015	0.019
N-Tech Lab	0.005	0.009	0.015	0.022	0.030	0.040	0.057	0.080
Lookman	0.012	0.014	0.016	0.017	0.017	0.018	0.021	0.026
Alivia	0.009	0.012	0.014	0.016	0.018	0.020	0.023	0.031
Neurotech.	0.009	0.012	0.014	0.015	0.017	0.019	0.023	0.031

顔の経年変化に対する識別特性（試験結果その2）

他人誤認率が1/1000となる閾値を設けたTOP1の本人見逃し率

【0～2年前】、【2～4年前】、【4～6年前】、【6～8年前】、【8～10年前】、【10～12年前】、【12～14年前】、【14～18年前】

NEC	0.007	0.009	0.011	0.013	0.015	0.017	0.021	0.027
依図	0.012	0.020	0.031	0.047	0.067	0.096	0.14	0.20
マイクロソフト	0.027	0.047	0.072	0.10	0.13	0.16	0.20	0.26
VisionLabs	0.048	0.080	0.13	0.17	0.21	0.24	0.30	0.36
N-Tech Lab	0.035	0.063	0.10	0.15	0.20	0.26	0.34	0.42
Lookman	0.043	0.069	0.098	0.13	0.16	0.19	0.23	0.28
Alivia	0.064	0.11	0.14	0.19	0.24	0.28	0.34	0.40
Neurotech.	0.065	0.10	0.14	0.18	0.23	0.27	0.33	0.40

顔の経年変化に対する識別特性（考察）

NEC社の【14～18年の隔たり】における「閾値を設けないTOP1の本人見逃し率」の数値（0.004）は、他の上位7社の数値（0.016～0.080）と較べて桁違いに優秀。同社の【14～18年の隔たり】における「他人誤認率が1/1000となる閾値を設けたTOP1の本人見逃し率」の数値（0.027）も、他社の数値（0.20～0.42）と較べて桁違いに優秀 ⇒ 数十年前に撮影した被疑者写真も多く含まれる被疑者写真データベースの検索・照合において、【14～18年の隔たり】における識別性能の桁違いの優秀さは大きな意義がある。

【 顔の長期経年変化に対する識別精度を更に向上させるには 】

「顔を識別するアルゴリズム」をディープラーニングにより生成する「学習フェーズ」で、より多くの「同一人物の昔の顔写真と今の顔写真のセット」を教材として用いて、反復学習することが効果的

Ⅳ－４

2018年FRVT
真横顔に対する識別特性

2013年の時点では、真横顔画像を正面顔画像と照合して同一人物であるか否かを見分けられる顔認識技術は存在せず、顔の正面から見て30度から45度ほどの斜め横方向から撮影した顔画像が、正面顔画像との照合により同一人物であるか否かを見分けられる限界であった。

しかし、2018年の時点では、真横顔画像を正面顔画像と照合して同一人物であるか否かを見分けられる顔認識技術が出現している。

真横顔に対する識別特性（試験方法）

「他人誤認率が1/10、1/100、1/1000となる閾値」の求め方

160万人分（160万枚）の顔画像データベース（全て正面顔のMUGSHOT画像）を検索対象として、10万人分（10万枚）の**真横顔のMUGSHOT画像**（いずれもデータベース内に同一人物の顔画像が存在しない。）で検索し、**「他人誤認率が1/10、1/100、1/1000となる閾値」**を求める。

「TOP1における本人見逃し率」の求め方

160万人分（160万枚）の顔画像データベース（全て正面顔のMUGSHOT画像）を検索対象として、10万人分（10万枚）の**真横顔のMUGSHOT画像**（いずれもデータベース内に撮影時点が異なる同一人物の正面顔のMUGSHOT画像が存在する。）で検索し、**「閾値を設けないTOP1の本人見逃し率」**と、**「他人誤認率が1/10、1/100、1/1000となる閾値を設けたTOP1の本人見逃し率」**を算出する。

真横顔に対する識別特性（試験方法）

160万人分（160万枚）のMUGSHOT画像（正面顔）に対して、10万人分のMUGSHOT画像（真横顔）で10万回検索

最新のMUGSHOT画像（検索用）

最新のMUGSHOT画像（真横顔）で、最新から2番目のMUGSHOT画像（正面）を検索

最新から2番目のMUGSHOT画像（被検索用）

真横顔に対する識別特性（試験結果）

他人誤認率一定（緑字1/10、赤字1/100、青字1/1000）としたTOP1の本人見逃し率
紫字は、閾値を設定しないTOP1の本人見逃し率

マイクロソフト	0.089、0,109、0.148、0.203
VisionLabs	0.130、0,198、0.322、0.461
N-Tech Lab	0.208、0,317、0.443、0.566
NEC	0.272、0,340、0.479、0.664
Tevian	0.329、0,483、0.661、0.910
Dermalog	0.517、0,642、0.856、0.948
依図	0.831、0,845、0.875、0.902

真横顔に対する識別特性（考察）

最上位4社の「閾値を設けないTOP1の本人見逃し率」の数値（0.089～0.272）は、160万人分（160万枚）の顔画像データベース（全て高品質な正面顔のMUGSHOT画像）を検索対象として、高品質な真横顔のMUGSHOT画像で検索を行った結果 ➡ 被疑者写真データベースの検索・照合において大きな意義 ➡ 犯人の遺留顔画像の撮影角度が真横に近かったとしても高品質（顔画像が緻密で鮮明）であれば、被疑者写真データベースを検索して得られるTOP5000（類似度の上位5000位以内）に、犯人の被疑者写真をリストアップし得ることを示している。

【 真横顔に対する識別精度を更に向上させるには 】

「顔を識別するアルゴリズム」をディープラーニングにより生成する「学習フェーズ」で、より多くの「真横顔と正面顔のセット」を教材として用いて、反復学習することが効果的

Ⅳ－５

2018年FRVT
同一人物の複数ショット照合による識別精度の向上

被疑者写真データベースでは、累犯者には複数の被疑者写真が存在する。このため、防犯カメラの録画映像に遺留された犯人の顔画像に基づき、被疑者写真データベースを検索する際には、累犯者の被疑者写真の扱い方について2通りの方法がある。つまり、最新の被疑者写真のみを検索対象とする方法と、全ての被疑者写真を検索対象とする方法である。どちらの方法が識別精度に優れるのか、興味深いところである。

同一人物の複数ショット照合による識別精度の向上（試験方法）

「他人誤認率」が1/1000となる「閾値」の求め方

160万人分（160万枚）の顔画像データベース（全てMUGSHOT画像）を検索対象として、331,254人分（331,254枚）のMUGSHOT画像（いずれもデータベース内に同一人物の顔画像が存在しない。）で検索し、「他人誤認率が1/1000となる閾値」を求める。

複数ショットを含まない場合の「TOP1における本人見逃し率」の求め方

160万人分（160万枚）の顔画像データベース（全てMUGSHOT画像で、同一人物の複数ショットを含まない。）を検索対象として、154,549人分（154,549枚）のMUGSHOT画像（データベース内の同一人物画像とは別画像）で検索し、「他人誤認率が1/1000となる閾値を設けたTOP1の本人見逃し率」を算出する。

複数ショットを含む場合の「TOP1における本人見逃し率」の求め方

160万人分（3,351,206枚）の顔画像データベース（全てMUGSHOT画像で、同一人物の複数ショットを全て含む。）を検索対象として、154,549人分（154,549枚）のMUGSHOT画像（データベース内の同一人物画像とは別画像）で検索し、「他人誤認率が1/1000となる閾値を設けたTOP1の本人見逃し率」を算出する。

同一人物の複数ショット照合による識別精度の向上（試験方法）

160万人分(160万枚)のMUGSHOT画像（最新から2番目の画像）に対して、最新のMUGSHOT画像（154,549人分）で154,549回検索

160万人分(3,351,206枚)のMUGSHOT画像（最新から2番目以前の全ての画像）に対して、最新のMUGSHOT画像（154,549人分）で154,549回検索

最新のMUGSHOT画像（検索用）

全てのMUGSHOT画像を検索

最新から2番目のMUGSHOT画像のみを検索

同一人物の最新以外の複数の MUGSHOT画像（被検索用）

同一人物の複数ショット照合による識別精度の向上（試験結果）

他人誤認率が1/1000となる閾値を設けたTOP1の本人見逃し率

赤字は、検索対象が160万人分(160万枚)のMUGSHOT画像の場合

青字は、検索対象が160万人分(3,351,206枚)のMUGSHOT画像の場合 ← 同一人物の複数ショット照合

NEC	0.0044	0.0021
依図	0.0123	0.0074
マイクロソフト	0.0141	0.0080
センスタイム	0.0234	0.0165
VisionLabs	0.0289	0.0185
N-Tech Lab	0.0391	0.0301
Lookman	0.0463	0.0425
Alivia	0.0620	0.0402
Neurotech.	0.0564	0.0527
東芝	0.0648	0.0529

160万人分(3,351,206枚)の内訳
1人1枚　80.1%
1人2枚　13.4%
1人3枚　3.7%
1人4枚　1.4%
1人5枚　0.6%
1人6枚　0.3%
1人7枚以上 0.2%（最大33枚）

同一人物の複数ショット照合による識別精度の向上（考察）

160万人分(3,351,206枚)のMUGSHOT画像では、160万人中の約32万人(全体の約2割)が、撮影期日の異なる複数枚のMUGSHOT画像を有するため、160万人分(160万枚)のMUGSHOT画像と較べて検索対象枚数が甚だしく増加（160万枚から約335万枚に倍増）している。しかし、検索結果については、検索対象枚数の倍増にも関わらず、いずれのベンダーついても識別精度に顕著な改善が見られる。

このことから、

顔画像データベース内に1人につき複数枚の顔画像が存在する場合には、最新の1枚のみに対して照合するよりも、最古も含めた複数枚全てに対して照合する方が、一定の閾値の下で他人誤認率を悪化させることなく、本人見逃し率を低減できる。

視点を変えれば、

ライブ映像の中から目的とする人物（ターゲット）を見つけ出す場合には、映像内の1ショットのみの顔画像でターゲットデータベースを照合するよりも、映像内の複数ショットの顔画像を用いてターゲットデータベースを照合する方が、一定の閾値の下で他人誤認率を悪化させることなく、本人見逃し率を低減できる。

Ⅵ－６

2018年FRVTは今日でも継続中

トップランナーに躍り出た中国企業

2018年FRVTは今日でも継続中

2018年FRVTは、参加した49社の結果が「NIST Interagency Report 8271」として纏められ、2019年9月に公表された。 ➡ しかし、2018年FRVTは、今日でも継続して実施されている。つまり、**新たなテストの申し込みを随時に受け付けて、2018年FRVTと同様の手順に基づくテストを実施し、その結果を「NISTIR 8271 DRAFT SUPPLEMENT」にまとめて公表**している。

【 これまでに公表された「NISTIR 8271 DRAFT SUPPLEMENT」 】

◯ 2021年11月に公表されたDRAFT SUPPLEMENT ➡ 我が国のCANONや富士通、Cloudwalk Technology（中国）、Samsung S1（韓国）などが新たに参加（参加企業の総数は80社）

◯ 2023年4月に公表されたDRAFT SUPPLEMENT ➡ 21社が新たに参加（参加企業の総数は101社）

ここから分かることは、

次のページへ

前のページから

ここから分かることは、

2013年FRVT（NISTIR 8009として2014年5月に公表）と2018年FRVT（NISTIR 8271として2019年9月に公表）では、**NEC（日本）が大半のテスト項目でトップの成績を残し、センスタイム等の中国企業がNECを凌駕したテスト項目は皆無**

しかし、

2018FRVTの継続として、2021年11月に公表されたNISTIR 8271 DRAFT SUPPLEMENTでは、**センスタイムやCloudwalk Technologyといった中国企業がNECにキャッチアップ**

→ さらに、2023年4月に公表されたNISTIR 8271 DRAFT SUPPLEMENTでは、**センスタイムやCloudwalk TechnologyがNECを凌駕する傾向も見られる。**

具体的には、

次のページへ

160万人分のMUGSHOT画像を被検索対象として、
他人誤認率を一定(1/1000)とした場合の、TOP1における本人見逃し率
緑字は、2019年9月公表時のデータ
赤字は、2021年11月公表時のデータ
青字は、2023年4月公表時のデータ

【 MUGSHOT画像（正面顔）で検索した結果 】

NEC	0.004（2019年9月）	0.002（2021年11月）	0.004（2023年4月）
センスタイム	0.023（2019年9月）	0.002（2021年11月）	0.001（2023年4月）
Cloudwalk	― （2019年9月）	0.002（2021年11月）	0.002（2023年4月）

【 WEBカメラ画像（正面顔）で検索した結果 】

NEC	0.017（2019年9月）	0.013（2021年11月）	0.009（2023年4月）
センスタイム	0.063（2019年9月）	0.014（2021年11月）	0.007（2023年4月）
Cloudwalk	― （2019年9月）	0.013（2021年11月）	0.011（2023年4月）

【 MUGSHOT画像（真横顔）で検索した結果 】

NEC	0.664（2019年9月）	0.622（2021年11月）	0.147（2023年4月）
センスタイム	1.000（2019年9月）	0.173（2021年11月）	0.868（2023年4月）
Cloudwalk	― （2019年9月）	0.133（2021年11月）	0.063（2023年4月）

顔認識技術におけるディープラーニングの効能・効果

ディープラーニングが用いられていない2013年FRVTでは、**真横顔画像を正面顔画像と照合して同一人物か否かを見分けられる顔認識技術は皆無** ➡ 顔の正面から見て30度から45度ほどの斜め横方向から撮影した顔画像が、正面顔画像との照合により同一人物であるか否かを見分けられる限界 ⬅ **真横顔画像と正面顔画像との照合は、「人の目」でも困難**

しかし、

ディープラーニングが用いられるようになった2018年FRVT（2019年9月に公表）では、**真横顔画像を正面顔画像と照合して、同一人物か否かを見分けられる顔認識技術が出現**

➡ その後の短期間（2021年11月公表と2023年4月公表）に、**前記照合の精度が格段に向上** ⬅ **ディープラーニングの効能・効果を如実に示している。**

このことから、

次のページへ

前のページから

このことから、

【 顔認識技術の精度向上は、ディープラーニングの工夫次第 】

真横顔画像と正面顔画像との照合時の識別精度を向上させるには、「顔を識別するアルゴリズム」をディープラーニングにより生成する「学習フェーズ」において、より多くの「真横顔と正面顔のセット」を学習用教材として用いて、反復学習を適切に行うことに尽きる。

一般論として、

ディープラーニングによる顔認識技術において、顔画像の検出精度、識別精度、分類精度をさらに向上させるには、

「顔を検出、識別、分類するそれぞれのアルゴリズム」をディープラーニングにより生成する「学習フェーズ」において、より多くの適切な学習用教材を用いて、反復学習を適切に行うこと

VI

顔認識技術における
人種バイアスの問題とその解決方法

顔認識技術における人種バイアスの問題

米国では、犯罪捜査に顔認識技術を用いた際の人種バイアス、つまり、白人の顔に比べてアフリカ系やアジア系の有色人種の顔に対する他人誤認率が高くなることが、人種的偏見の助長に繋がりかねないとして、2010年代後半から大きな社会問題化

ところで、

顔認識技術に用いられるディープラーニングは、人の頭脳内部での神経回路網の仕組みと働きを、コンピュータ上で数学的に模したもの　→　人種バイアスの原因や対策を検討するには、「人の目」における人種バイアスについて検討することが有益

「人の目」では、

人は、相手の顔を見て美醜などの特徴を瞬時に見分けている。この見分ける基となっているものは、誕生以来、見たり接したりしてきた無数の顔から生み出された「頭の中の平均顔」と考えられる。見たり接したりしてきた顔の大半が日本人であれば、日本人の平均顔が頭の中に生み出されるであろう。日本人には、白人の顔、あるいは黒人の顔が、日本人の顔ほどには見分けが付けづらいが、日本人の平均顔を基に顔の特徴を見分けているとすれば、白人の顔、あるいは黒人の顔が、日本人には皆同じように見えてしまうのも無理はない。

【警察政策第17巻(2015)「顔画像識別における人の目の特性と機械の目の特性」からの引用】

ここから類推すれば、

次のページへ

前のページから

ここから類推すれば、

大半を日本人の顔で学習した顔認識技術では、白人や黒人の顔に対する識別精度が日本人の顔ほどには上がらない。 ➡ 米国で犯罪捜査に顔認識技術を用いた場合における**人種バイアスの問題は、白人の顔に比べてアフリカ系やアジア系の有色人種の顔に対する学習不足が最大の原因**

解決策は、

「顔を識別するアルゴリズム」をディープラーニングで生成する「学習フェーズ」での学習用教材として、アフリカ系やアジア系の有色人種の顔画像を更に多く用いて、白人の顔画像と同等の識別精度が達成できるまで、効果的かつ効率的な反復学習を行うこと

つまり、

ディープラーニングを活用した顔認識技術における識別精度の向上には、反復学習の充実強化が最も重要なファクター ⬅ 真横顔画像と正面顔画像との照合時の識別精度向上方策と同じ

2023年12月21日

終

顔認識技術の最新動向と犯罪捜査やテロ対策に向けた活用

澤田雅之技術士事務所（電気電子部門）所長
元警察大学校警察情報通信研究センター所長
澤田　雅之