CNN を用いた顔認証システムの開発と追跡停止に対する評価

○脇一史・森駿文・菊池浩明

明治大学総合数理学部

背景：顔認証カメラ

- 個人の追跡や防犯、商用に活用されている
- 追跡回避や削除要求（オプトアウト）などの課題
  - マスクやサングラスを着用することで追跡が停止できると言われている

朝日新聞　２０１４年１月６日　朝刊　１ページ　大阪本

大阪駅ビル 顔で追跡

4月から実験

JR大阪駅の駅ビル「大阪ステーションシティ」（大阪市北区）で通行人の顔をカメラ約90台で撮影し、その特徴を登録して同一人物を自動的に追跡する実験が4月から始まる。顔認証技術＝■＝の精度を確かめるのが狙いで、データは個人が識別できない処理をしたうえで、JR西日本に提供されるという。不特定多数の人を撮影しデータを収集する行為に、専門家はプライバシー侵害への懸念を示している。▼31面＝こっそり収集

大阪ステーションシティの地下通路にはすでに顔認証用カメラが設置されている

カメラ90台 行動把握

総務省所管の独立行政法人「情報通信研究機構」（東京都小金井市）がJR西日本とステーションシティを運営する「大阪ターミナルビル」の協力を得て、2年間実施する。

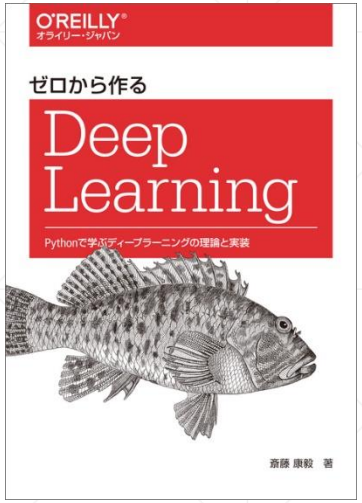
実験では、各カメラで3㍍四方にいる数十人の顔を瞬時に撮影する。両目の幅など100カ所程度の人の顔の特徴を抽出し、人の顔の特徴を抽出定のIDを与えて登別のカメラが同じ特つ顔を識別すると、物と判断して追跡。内での約1週間の行握できる仕組みだ。

研究目的

- マスクやサングラス、帽子などの外乱によって追跡が停止できるのか？
  - では、マスクごと学習して識別されたらどうか？
  - 目や口などの部位を計測する専用認識器では外乱に弱い
    - » 顔部位検出法（Viola-Jones法）では難しい

研究方法

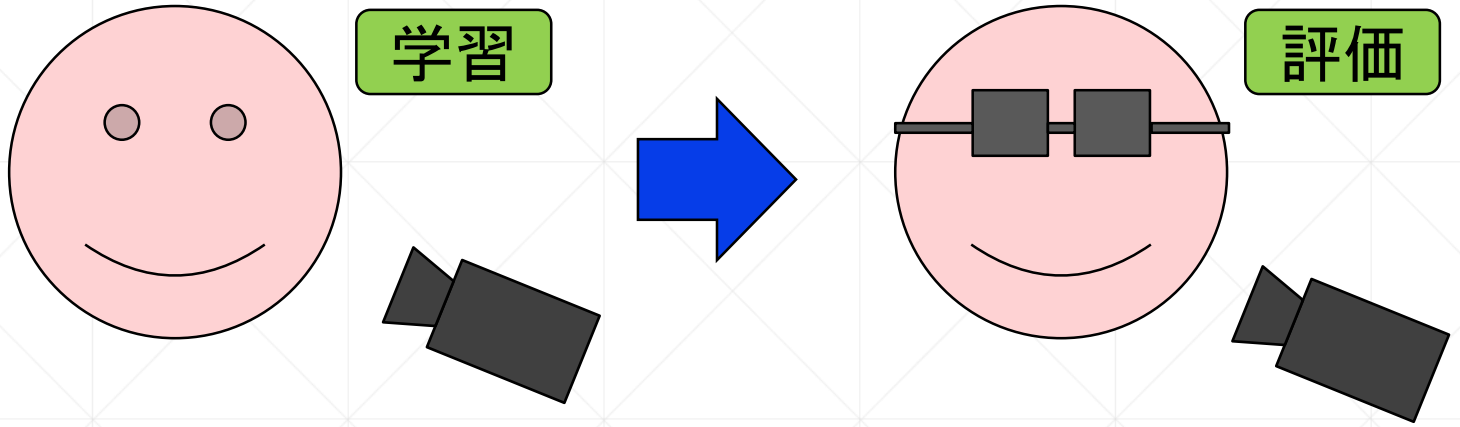
- 「ゼロから作るDeep Learning」[1]を基に実装した
  - 本書に記載されているプログラムを改変し実装
- 実装したConvolutional Neural network(CNN)に対して外乱ごとに学習させた場合追跡ができるのかを明らかにした



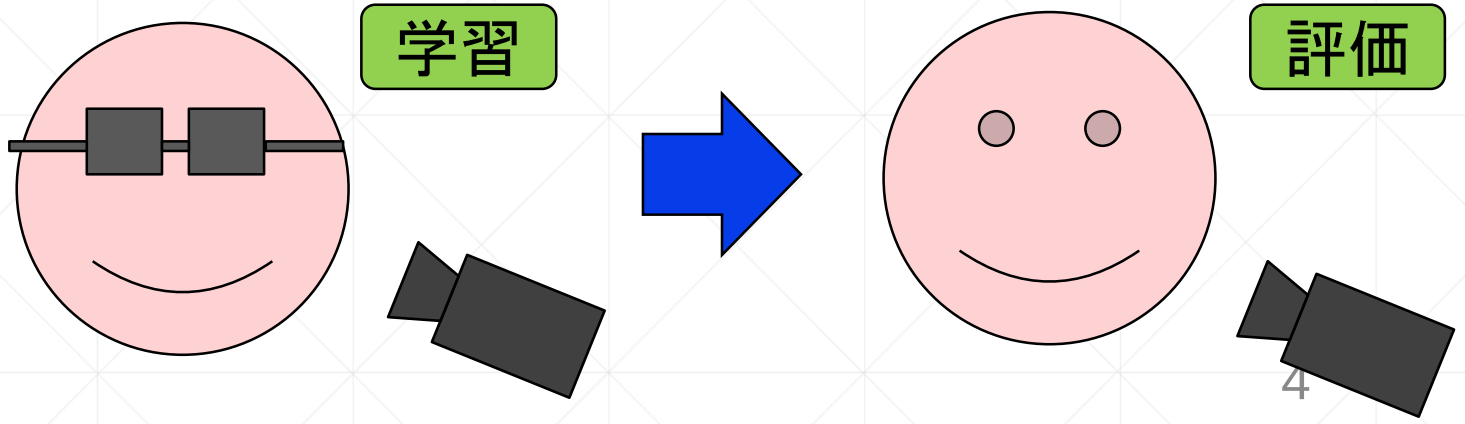
[1]齋藤康毅, “ゼロから作るDeep Learning python で学ぶディープラーニングの理論と実装”, OREILLY, 2016.

実験概要

1. 素顔を学習したときに識別精度を下げる外乱を明らかにする
   - 素顔
   - 帽子
   - マスク
   - サングラス
   - マスク+サングラス



2. 外乱ごと学習したときの外乱の精度を明らかにする



顔画像データの取得

- iMacのwebカメラを使用
- 画面上に枠を出し、その範囲内に収まるように撮影
- CNN作成用データ
  - 被験者　：6名
  - 撮影枚数/人: 100枚×5日＝500枚
- 評価用データ
  - 被験者　：5名
  - 撮影枚数: 100枚

顔画像データの拡張

- 取得画像を112×112にリサイズ
- 右図の15パターンの調整を行った
  - 元画像含め$500 \times 16 = 8{,}000$枚
- ランダムな位置で96×96に切り出しを行った
- $8{,}000$枚$\times 6$人$= 48{,}000$枚
  - 38,400枚を学習データ
  - 9,600枚をテストデータ

手法	コントラスト調整	輝度変換	ガウシアンノイズ
種類	12% 23% 35% 47% 59% 70%	0.5 0.7 0.9 1.1 1.3 1.5 1.7	2 4
計	6	7	2

CNNの構成

- VGG-11[1]という識別手法を参考にした
  - Dropout[2]を全結合層に追加
    - » ニューロンをランダムに消去し過学習を抑制
- 学習回数: 2epoch

素顔データ

[1] Karen Simonyan and AndrewZisserman(2014):Very Deep Convolutional Networks for Large-Scale Image Recognition, ICLR,2014.
[2]齋藤康毅, “ゼロから作るDeep Learning python で学ぶディープラーニングの理論と実装”, OREILLY, 2016.

CNNでの追跡停止の評価

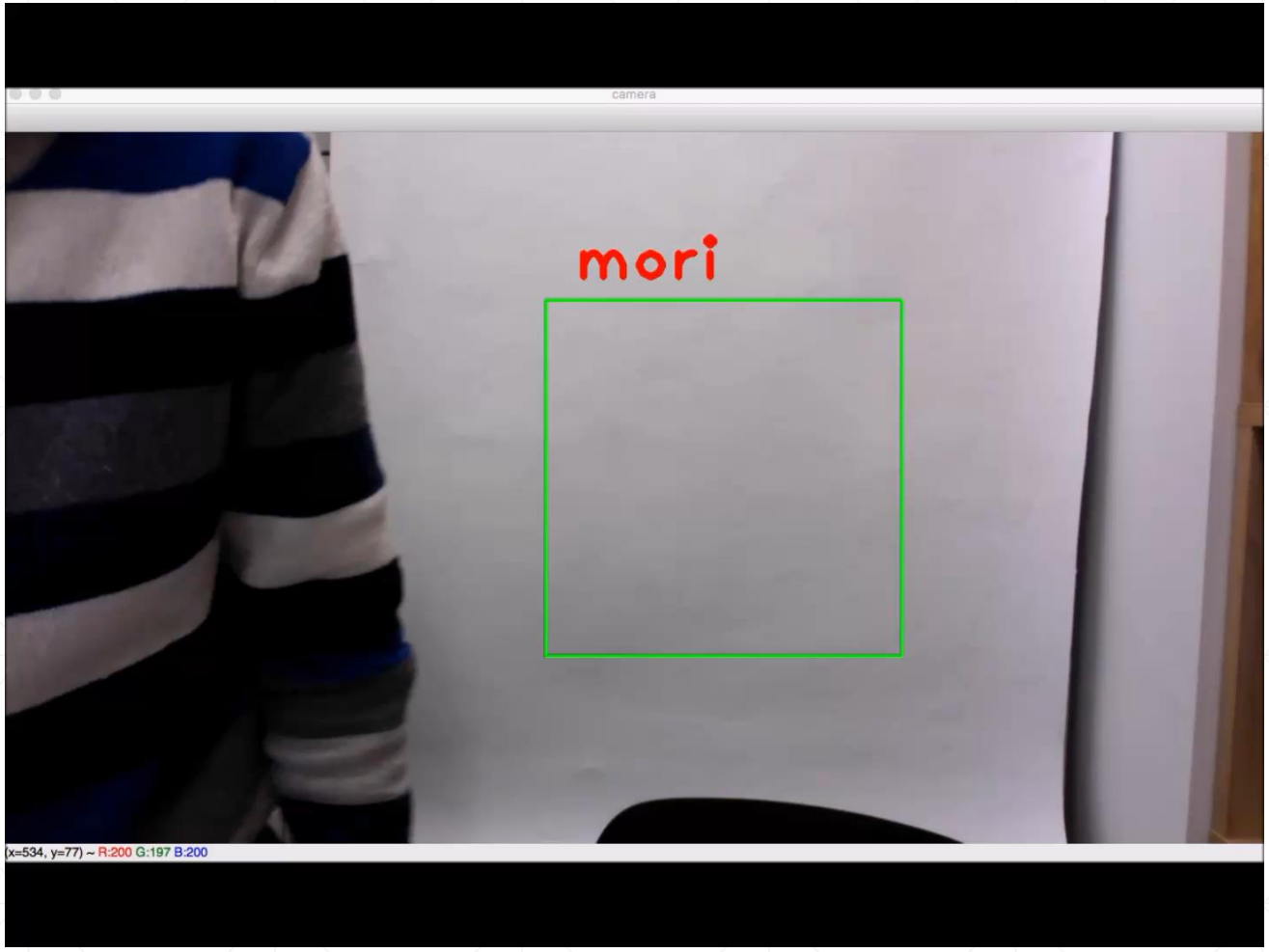
- 素顔&外乱画像それぞれのパラメータでCNNを構築
  - 5つのパラメータを持ったCNNを作成
  - それぞれの画像で再現率を計算
    - 追跡停止の可否を検証するために再現率を求めた
  - Aさんの再現率$R_A = \dfrac{A\text{と正しく判定した数}}{\text{本物の}A\text{さんの画像数}}$
  - 全員の再現率の平均を比較

評価について

- (1) 素顔で学習したとき、追跡停止に最も効果的な外乱はどれか？
- (2) 外乱ごと学習したとき、追跡停止を妨害できるのはどれか？

顔認証システム実行画面

- webカメラで撮影した画像をリアルタイムに判定可能

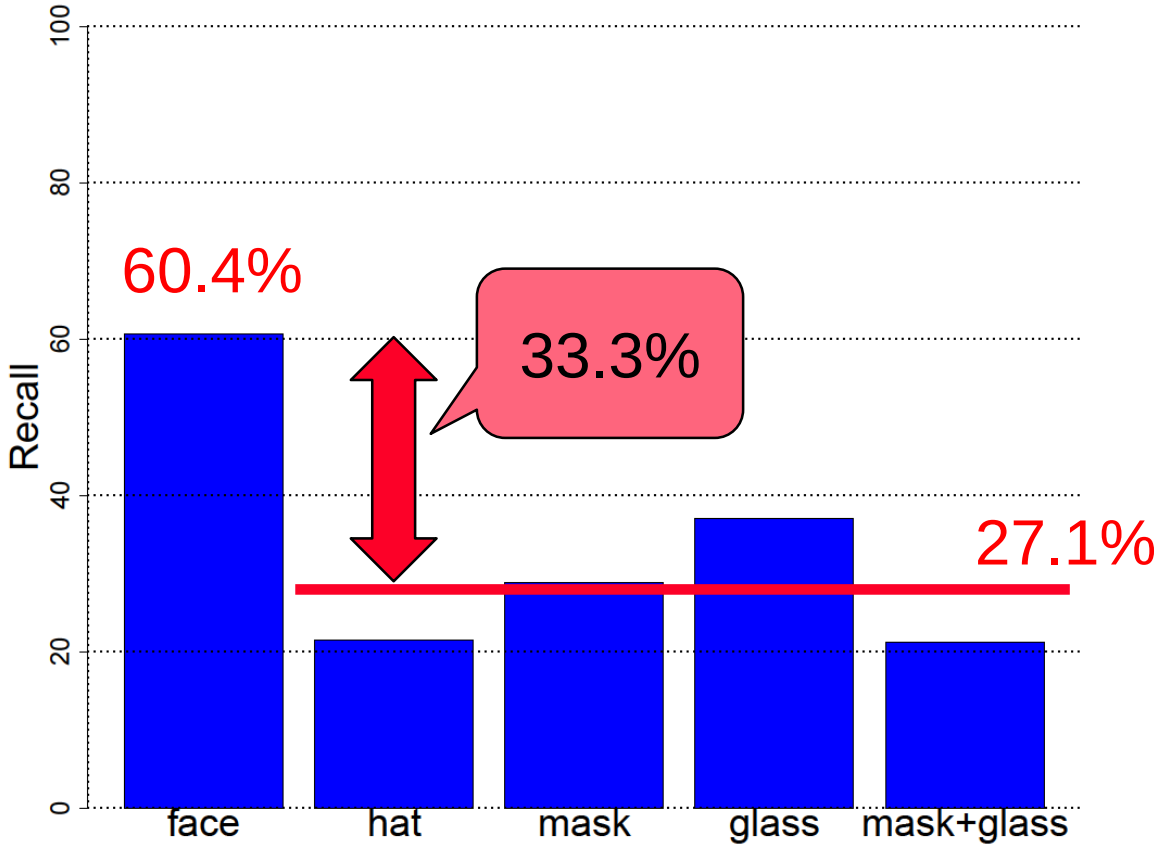


脇

森

(1)素顔で学習したときの結果

- 素顔が60.4%と低い結果となった
  - 素顔は表情の変化などによる個人内での変動が大きくなったため
- 外乱を加えることで識別率は平均33.3%低下した



(2)外乱画像で学習した平均再現率

- 学習と評価が同じ組み合わせの中では素顔が最も悪かった
- マスク+サングラス→マスクでも84.2%となった
  - マスクやサングラスの掛け方も特徴となった
- サングラスで学習したときが平均が一番高かった
  - まばたきや眼球の動きが一番個人内の変動が激しいと考えられる
  - そこを隠すことで静的な特徴を学習できた

学習＼評価	素顔	帽子	マスク	サングラス	マスク+サングス	平均
素顔	60.6	21.4	28.8	37.0	21.2	33.8
帽子	51.4	74.8	20.0	48.8	20.0	43.0
マスク	20.2	20.0	99.4	20.0	30.0	37.9
サングラス	53.0	20.6	50.4	94.4	53.4	54.8
マスク+サングラス	25.2	20.0	84.2	22.0	78.6	46.0
平均	42.1	31.8	56.6	44.4	40.6	43.1

おわりに

- (1)素顔で学習したとき、追跡停止に最も効果的な外乱はどれか？
  - 帽子
    - » 撮影角度によって精度に変化がある
  - マスクやサングラスなどにより、顔認証の追跡を33%防止できることを示した
- (2) 外乱ごと学習したとき、追跡停止を妨害できるのはどれか？
  - **マスク**
    - » ただし、汎用性ではサングラスの方が性能が良かった

今後の課題

- 外乱の組み合わせを増やす
  - マスク＋帽子など
- Keras等のフレームワークを用いて実装し、検証する
  - 全体の認識精度向上を目指す