

Convolutional Neural Network による顔の向きに頑健な顔器官検出

木村 真稔[†] 福井 宏[†] 山下 隆義^{††} 山内 悠嗣[†] 藤吉 弘亘^{††}

[†] 中部大学 〒487-0027 愛知県春日井市松本町 1200

E-mail: [†]{kimura001,fhiro,yuu}@vision.cs.chubu.ac.jp, ^{††}{yamashita,hf}@cs.chubu.ac.jp

あらまし 本稿では, Convolutional Neural Network (CNN) による顔向きに頑健な顔器官検出手法を提案する. CNN の高精度な性能や頑健性が様々な分野において注目されているが, 学習におけるハイパーパラメータ等の設定や学習サンプルの与え方が複雑であり学習の難しさに課題が残っている. そこで, 我々は学習サンプルの与え方についてどのような与え方が適しているかを明らかにする.

キーワード Convolutional Neural Network, 顔の器官点検出, ミニバッチ学習

Facial Point Detection Based on Deep Convolutional Neural Network with Optimal Minibatch

Masatoshi KIMURA[†], Hiroshi FUKUI[†], Takayoshi YAMASHITA ^{††}, Yuji YAMAUCHI[†], and
Hironobu FUJIYOSHI ^{††}

[†] College of Engineering, Chubu University 1200, Matsumoto-cho, Kasugai, AICHI

E-mail: [†]{kimura001,fhiro,yuu}@vision.cs.chubu.ac.jp, ^{††}{yamashita,hf}@cs.chubu.ac.jp

Abstract We propose a Convolutional Neural Network (CNN)-based method to ensure both robustness to variations in facial pose. Although the robustness of CNN has attracted attention in various fields, the training process suffers from difficulties in parameter setting and the manner in which training samples are provided. We demonstrate a manner of providing samples that results in a better network. Experimental results indicate that the subset with augmentation technique has sufficient variations and quantity to obtain the best performance.

Key words Convolutional Neural Network, Facial Points Detection, minibatch

1. ま え が き

顔画像を用いた個人認識や表情推定, 視線推定の前処理として顔の器官点検出は重要な技術である. 器官点は, 個人によって器官点の配置パターンが異なっているため, 個人認識に有効な特徴量とされており, 器官点の配置パターンを用いることで表情の推定も可能である. このような機能を実現するために, 顔画像が様々な表情, 照明条件, オクルージョンで撮影された状態においても, 正確に器官点を検出することが求められている.

顔の向きに頑健な検出法として, 顔の向き毎に識別器を作成する Conditional Regression Forests (CRF) が挙げられる [2]. CRF は顔の向きごとに Regression Forests を作成し, 顔の向きを認識した後, その顔向きに応じた Regression Forests を利用して, 各器官点を回帰により検出している. CRF の前提条件として顔の向きを認識が必要であるため, 高い向きを認識精度が不可欠であり, また器官点検出の計算コストが高いという

デメリットがある.

本研究では, Convolutional Neural Network (CNN) [1] を用いた顔の向きに頑健な顔の器官点検出を提案する. また, 器官検出の精度を向上させるために, 学習におけるサンプルの与え方について最適な方法を検討する.

2. 提 案 手 法

CNN は, 確率的勾配降下法により逐次学習サンプルを与え, ネットワークのパラメータを学習する. その際, 学習サンプルを一定数ごとに与えるミニバッチ学習が一般的である. ミニバッチ学習におけるサブセットの作成に関して, これまでランダムに学習サンプルを選択して作成する方法が多く用いられているが, どのような与え方が適しているか検討が十分なされていない. そこで提案手法では, 2つのミニバッチの作成方法を提案することで, CNN の学習と識別の変化を調査する.

2.1 Random minibatch

Random minibatch によるミニバッチ作成方法を図 1 に示

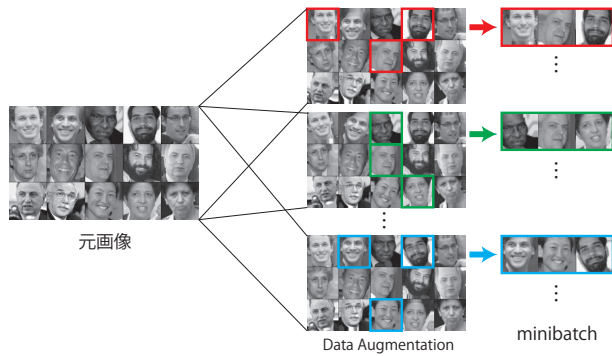


図 1 Random minibatch のアルゴリズム

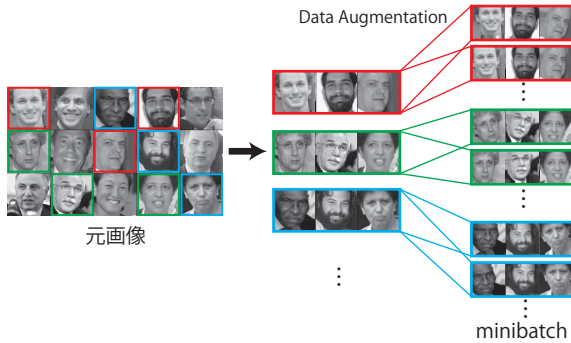


図 2 Fixed-person minibatch のアルゴリズム

す。まず、Data Augmentation により学習サンプルに幾何変化を与えることでサンプル数を増幅させる。そして、増幅した学習サンプル群からランダムに学習サンプルを選択してミニバッチを作成する。Random minibatch は、ミニバッチ内の学習サンプルの組み合わせが多いため、汎化性能が向上しやすくなっている。

2.2 Fixed-person minibatch

図 2 に Fixed-person minibatch によるミニバッチ作成方法を示す。Fixed-person minibatch では、はじめに学習サンプル群からミニバッチで使用する学習サンプルを選択する。このとき、選択した学習サンプルはミニバッチごとにサブセットとして保存する。そして、保存したサブセットに対して Data Augmentation をすることでサブセット単位で学習サンプルを増幅させている。これにより、各ミニバッチ内の学習サンプルに対して、顔の器官点検出の場合に特定の人物が固定された状態でサンプルを生成することができる。

3. 評価実験

提案手法の有効性を示すために、評価実験により器官点の検出精度を比較する。評価実験では、10 点の器官点を検出し、教師信号と検出結果を用いて評価する。検出する顔の器官点は、両目の目尻と目頭、鼻翼、上下唇の中心、口尻の 10 点である。データセットには、人の顔画像から構成されている Labeled Faces in the Wild (LFW) データセットを使用し、学習サンプル 1,500 枚、評価サンプル 927 枚を用いる。各サンプルのサイズは 100×100 のグレースケール画像である。Data Augmentation により学習サンプルに与える幾何変換は -15 度から 15 度までの回転と -10 pixel から 10 pixel までの平行移動を与え、サンプル数を 14 倍程度に生成する。

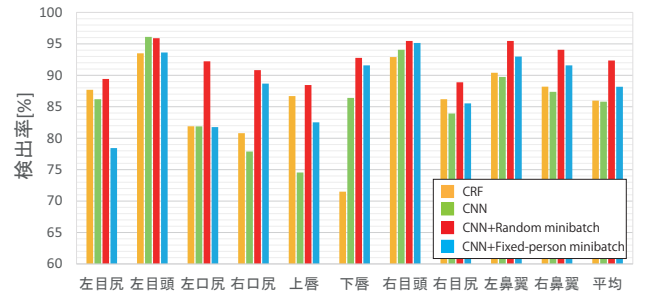


図 3 検出率の比較

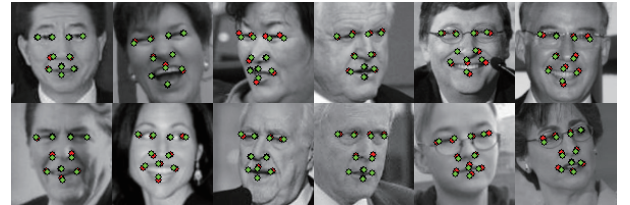


図 4 検出結果の一例

3.1 各ミニバッチ作成方法による精度の比較

図 3 の実験結果から従来の Data Augmentation を行わない CNN と比較して、Data Augmentation を行った CNN の方が検出精度が高いことが分かる。また、各器官点の検出率の平均から Random minibatch が Fixed-person minibatch よりも約 4% 検出精度が向上している。Random minibatch は平均だけでなく、すべての器官点において Fixed-person minibatch よりも精度がよい。Random minibatch はミニバッチ内で人物が固定されておらず様々な組み合わせがある。これによりサブセットのバリエーションが多くなっており、CNN の学習においてはバリエーションの豊富さが不可欠であることが分かる。

3.2 従来法と比較

図 3 より、従来の CNN の各器官点の平均検出率は CRF とほぼ同等である。一方、Random minibatch を用いた場合、検出精度が約 6% 向上していることが分かる。図 4 に検出結果を示す。ここで、赤い点が教師信号、緑の点が検出結果を示している。図 4 から、表情や顔の向きの変化に対しても頑健に検出していることが分かる。これにより、CNN の学習に Random minibatch を導入することで、顔向き変化に頑健な器官検出を実現することができている。

4. おわりに

本稿では、CNN による顔器官検出および、ミニバッチ作成方法について検討した。CNN のミニバッチ作成方法は、ミニバッチ内のバリエーションが多くあることが学習には不可欠であり、提案する Random minibatch により、従来手法と比較して顔の器官点検出精度を約 6% の向上させることができた。今後は、CNN による顔器官検出の更なる高精度化を目指してネットワークの最適な初期値の設定方法等を検討していく。

文 献

- [1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. In IEEE, 1998.
- [2] M. Dantone, J. Gall, G. Fanelli, and L. V. Gool. Real-time facial feature detection using conditional regression forests. In CVPR, 2012.