

# Heterogeneous Learning と 重み付き誤差関数の導入による顔画像解析

加藤 優<sup>1,a)</sup> 福井 宏<sup>1,b)</sup> 山下 隆義<sup>1,c)</sup> 山内 悠嗣<sup>1,d)</sup> 藤吉 弘亘<sup>1,e)</sup>

## 1. はじめに

画像認識分野において、Deep Convolutional Neural Network(DCNN) が高精度な識別が可能であることから注目されている。DCNN の学習法に、複数の識別タスクを単体の DCNN で識別できる Heterogeneous Learning がある。しかし、Heterogeneous Learning は識別タスクに認識のタスクと回帰推定のタスクが共存した場合、ラベル値が連続値と離散値の違いが原因で、算出される学習誤差に大きな差が発生し、学習に影響を与えることがある。そこで、本研究では、Heterogeneous Learning に重み付き誤差関数を導入する。認識のタスクで発生する突発的な学習誤差を、重み付き誤差関数により軽減することで、Heterogeneous Learning の学習を安定させ、高精度な認識を実現する。

## 2. Heterogeneous Learning

Heterogeneous Learning は、図 1 に示すように複数のタスクを単体の DCNN で学習する学習法である。

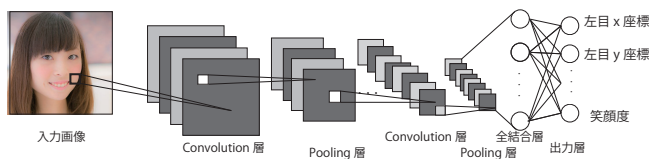


図 1 Heterogeneous Learning を導入した DCNN の構造

一般に複数の識別タスクを扱う場合、タスクの数に比例して識別器を構築する必要がある。それにより、識別タスクの数に比例して計算コストが増加するため、実用化の面で非現実的である。Heterogeneous Learning を用いて、一つの識別器を構築することで、複数のタスクを同時に学習、識別することが可能である。

## 3. 提案手法

提案手法では、顔器官点検出、性別認識、年齢推定、人種認識、笑顔度推定の 5 つのタスクを対象とする。そして、性別認識と人種認識を認識のタスク、顔器官点検出と年齢推定、笑顔度推定を回帰推定のタスクとする。また、学習の誤差関数には 2 乗誤差関数を用いる。単一タスクの学習誤差は図 2 のように出力される。

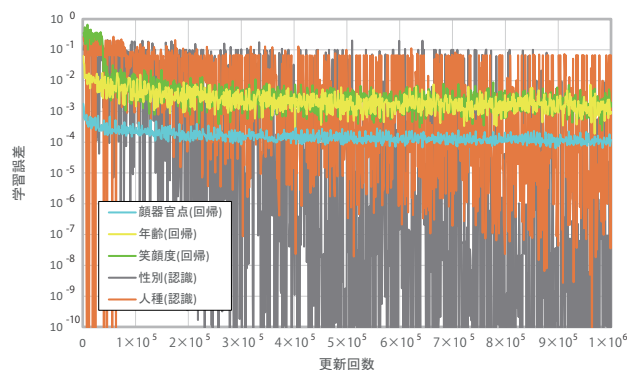


図 2 各識別タスクの学習誤差

図 2 から、認識のタスクと回帰推定のタスクを比較した場合、学習誤差に大きな振幅の差が生じていることが確認できる。回帰推定のタスクのラベル値は、連続値であるため中間的な表現をラベル値に反映できる反面、認識のタスクでは学習誤差が突発的に増減する現象が発生する。突発的な学習誤差の増減は、学習誤差が常に小さい回帰推定のタスクの学習に大きな影響を与えるため、精度が低下しやすい。

この突発的な学習誤差の増減の影響を軽減するため、本稿では式 (1) のように重み付き誤差関数  $E_M$  を導入する。認識のタスクと回帰推定のタスクに異なる重みを与えることで認識のタスクで発生する突発的な誤差を減少し、回帰推定のタスクに与える影響を低減する。ここで、 $M$  はバッチサイズ、 $T_c$ 、 $T_r$  は認識と回帰推定のタスク数、 $w_{t_c}$ 、 $w_{t_r}$  は認識と回帰推定のタスクに与える重み、 $L$  はラベル値、 $O$  は DCNN により出力される値を示している。

<sup>1</sup> 中部大学 〒487-0027 愛知県春日井市松本町 1200

a) u\_kato@vision.cs.chubu.ac.jp

b) fhiro@vision.cs.chubu.ac.jp

c) yamashita@cs.chubu.ac.jp

d) yuu@vision.cs.chubu.ac.jp

e) hf@cs.chubu.ac.jp

$$E_M = \frac{1}{M} \sum_{m=1}^M \left( \sum_{t_c=1}^{T_c} w_{t_c} \|L_{t_c} - O_{t_c}\|_2^2 + \sum_{t_r=1}^{T_r} w_{t_r} \|L_{t_r} - O_{t_r}\|_2^2 \right) \quad (1)$$

### 4. 評価実験

提案手法の有効性を示すために、各識別タスクを単一タスクの DCNN と従来の Heterogeneous Learning, 重み付き誤差関数を導入した Heterogeneous Learning の識別精度を比較する。本実験では、顔器官点検出, 性別認識, 年齢推定, 人種認識, 笑顔度推定を行う。顔器官点検出では右目, 左目, 鼻, 口の右端, 口の左端の計 5 点の位置を回帰により推定し, 笑顔度推定は 0 から 99 までの値を回帰により推定する。また, 人種認識では, アジア人, 白人, 黒人の 3 種類を判別する。回帰推定の重みは, 器官点検出が他のタスクと比べ非常に誤差のレンジが低いため, 年齢推定と笑顔度推定にかけるものとする。データセットには, Web から収集した 53,663 枚のデータセットを使用する。ここで, 学習サンプル 42,663 枚, 評価サンプル 11,000 枚とする。入力するサンプルは 100 × 100 画素のグレースケール画像である。

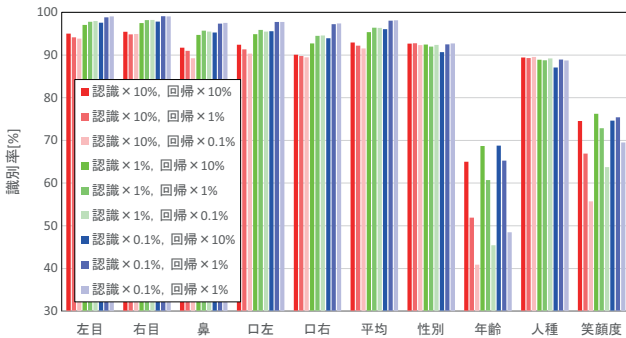


図 3 重みの変化による識別精度の比較

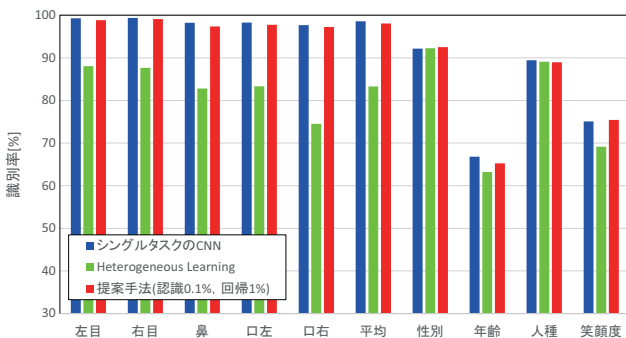


図 4 従来法と提案手法の識別精度の比較

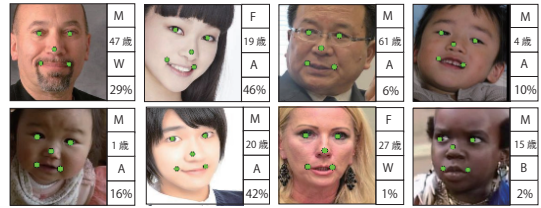


図 5 識別結果の一例

#### 4.1 重み付き誤差関数の重みを変更した際の精度の比較

図 3 より, 回帰推定のタスクは認識のタスクの重み  $w_{t_c}$  を小さくすることで精度が向上している。これは, 認識のタスクで突発的に発生する学習誤差の増減を重み付き誤差関数により抑制したため, 回帰推定のタスクの精度が向上したと考えられる。また, 認識のタスクは重みを小さくした場合, 精度に大きな変化は見られなかった。

#### 4.2 従来法と提案手法の比較

図 4 より, 従来の Heterogeneous Learning と比較したとき, 顔器官点の検出精度の平均が約 15% 向上している。また, 年齢推定, 笑顔度推定においても推定精度がそれぞれ約 2%, 約 6% 向上している。一方, 性別認識と人種認識においては従来の Heterogeneous Learning と同等の識別精度であり, 単一タスクで識別した DCNN と比較した場合には, 提案手法の検出精度は同程度の識別精度であった。

図 5 に, 重み付き誤差関数を導入した Heterogeneous Learning の識別例を示す。図 5 では, 図中の緑の点が顔器官点を示している。また, 右枠の上から順に性別 (M: 男性, F: 女性), 年齢, 人種 (A: アジア人, W: 白人, B: 黒人), 笑顔度を示している。

また, 1 枚あたりの処理時間は, Intel Core i7-4790(3.60GHz), メモリ 8GB の環境で行った場合, 単一タスクの DCNN で 5 つのタスクを実行するには約 110ms の処理時間を有する。しかし, 重み付き誤差関数を導入した Heterogeneous Learning では, 5 つのタスクの識別時間は約 22ms であった。よって, 単一タスクの DCNN では識別するタスクに比例して処理時間が増加していくが, Heterogeneous Learning では識別するタスクが増加しても処理時間は変化しないことが確認できる。

### 5. おわりに

本稿では, Heterogeneous Learning に重み付き誤差関数の導入を提案した。Heterogeneous Learning における各タスクに適した重み付き誤差関数を導入することで, 従来の Heterogeneous Learning より高精度な識別を実現した。今後は, 高精度かつ効率的な顔画像解析を行う。

#### 参考文献

[1] Z. Zhang, *et al.*, “Facial Landmark Detection by Deep Multi-task Learning”, ECCV, 2014.  
 [2] M. Kimura, *et al.*, “Facial Point Detection Based on Deep Convolutional Neural Network with Optimal Minibatch”, PRMU, 2015.