

# 重み付き誤差関数を導入した DCNN による Heterogeneous Learning の高精度化

EP12036 加藤優

指導教授：藤吉弘亘，山下隆義

## 1.はじめに

複数のタスクを单一の識別器で学習・識別するアプローチとして、Heterogeneous Learning が提案されている [1]。複数のタスクを同時に学習する際、各タスクにおける学習誤差の収束値の差が、学習に悪影響を与える場合がある。そこで、本研究では重み付き誤差関数を導入した Deep Convolutional Neural Network(DCNN) による Heterogeneous Learning を提案する。これにより、複数のタスク間の学習誤差を均一化し、高精度な顔属性認識を実現することが期待できる。

## 2. DCNN による Heterogeneous Learning

DCNN による Heterogeneous Learning は、図 1 のように複数のタスクを单一の DCNN で学習する [2]。これにより、タスクの数に比例して学習、識別コストを増やすことなく、複数のタスクを单一の DCNN で識別することが可能となる。本研究では、顔画像を対象とし、回帰推定タスクでは顔器官点検出、年齢推定、笑顔度推定を行い、認識タスクでは性別認識、人種認識を行う。

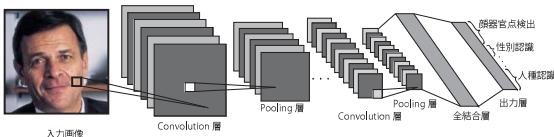


図 1 : DCNN による Heterogeneous Learning

図 2 に各タスクを单一の DCNN で学習した際の学習誤差を示す。図 2 より、各タスクの収束値に差があることがわかる。顔器官点検出では、各顔器官点の位置の変化が少ないため、難易度が低い。一方、年齢推定や笑顔度推定では見えが大きく変化するため、難易度が高い。このように、タスクの難易度の差により、収束値には大きな差が生じる。また、回帰推定タスクのラベル値は 0 から 1 の連続値であるのに対して、認識タスクのラベル値は 0 か 1 の離散値である。そのため、認識タスクでは回帰推定タスクに比べ、学習誤差が大きくなることから、収束値に大きな差が生じる。これら 2 つの要因による収束値の差は、DCNN による Heterogeneous Learning の学習に悪影響を与えることがある。

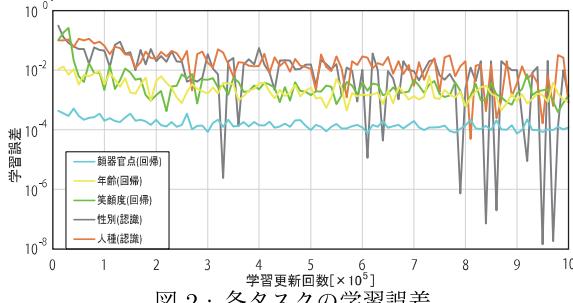


図 2 : 各タスクの学習誤差

## 3. 提案手法

提案手法では、顔器官点検出以外のタスクに対して、学習誤差関数  $E$  に重みを付与する。各タスクに対して重みを与えることで、各タスクの学習誤差の収束値の差が低減し、安定した学習が可能となる。学習誤差関数に対する重みは、式 (1) のように与える。

$$E = \frac{1}{M} \sum_{m=1}^M \left( \| \mathbf{L}_{f,m} - \mathbf{O}_{f,m} \|_2^2 + \sum_{t \neq f}^T w_t \| \mathbf{L}_{t,m} - \mathbf{O}_{t,m} \|_2^2 \right) \quad (1)$$

ここで、 $M$  はバッチサイズ、 $T$  はタスク数、 $w_t$  は各タスクに与える重み係数、 $\mathbf{L}_{t,m}$  は各タスクのラベル値、 $\mathbf{O}_{t,m}$  は各タスクの出力値、 $\mathbf{L}_{f,m}$  は顔器官点のラベル値、 $\mathbf{O}_{f,m}$  は顔器官点検出の出力値を示している。各タスクの重み係数  $w_t$  の算出には、正規分布を用いて算出する。全学習誤差の 99.7% が、平均  $\mu$  と標準偏差  $\sigma$  の 3 倍の区間に含まれている。この区間外は突発的な学習誤差となっており、これを無視して重みを付与する。これにより、突発的な学習誤差の影響を受けず、重みを付与することができる。この値を顔器官点検出と各タスクで算出し、それぞれ  $N_f$ 、 $N_t$  とする。算出された  $N_f$ 、 $N_t$  から、式 (2) により重み係数  $w_t$  を算出する。

$$w_t = \frac{N_f}{N_t} \quad (2)$$

## 4. 評価実験

評価実験では、単一タスクの DCNN、従来法である重み無しの DCNN による Heterogeneous Learning、提案手法の 3 つの学習法を比較する。データセットには、Web から収集したサンプルを使用する。また、学習サンプル 42,663 枚、評価サンプル 11,000 枚とし、100 × 100 画素のグレースケール画像を使用する。

### 従来法との比較

図 3 に、単一タスクの DCNN、従来法、提案手法におけるタスク毎の識別率を示す。図 3 より、従来法は単一タスクの DCNN と比較した場合、回帰推定タスクの精度が大幅に低下していることがわかる。また、提案手法は従来法と比較した場合、全てのタスクにおいて約 5% 識別率が向上した。特に、顔器官点検出において、約 14% 精度が向上している。そして、提案手法は単一タスクの DCNN と比較した場合、全てのタスクにおいて精度が向上、もしくは同等の精度で識別できることを確認した。

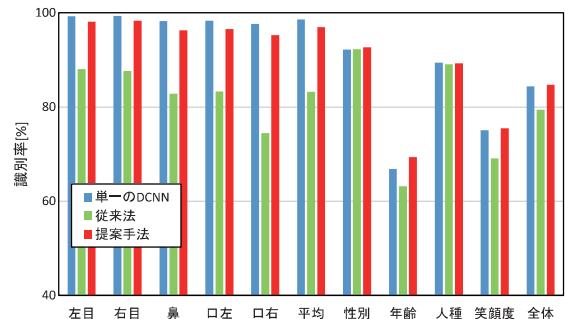


図 3 : 従来法と提案手法の識別精度の比較

図 4 に、顔の複数属性認識の例を示す。従来法では特に顔器官点の検出が困難であるが、提案手法では重み付き誤差関数を導入することで安定した学習が可能となり、より正確に顔器官点を検出することができた。

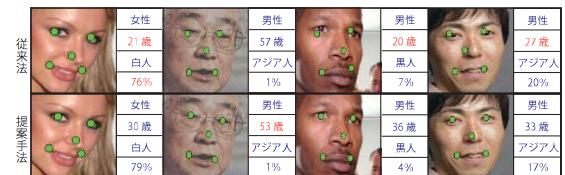


図 4 : 顔の複数属性認識の例

### 処理時間の比較

Intel Core i7-4790、メモリ 8GB の環境において、単一の DCNN で 5 つのタスクを処理するには約 110[ms] の処理時間を必要とする。一方、Heterogeneous Learning では約 22[ms] であり、1 タスク分の処理時間で複数のタスクを処理できる。

### 5. おわりに

本研究では、重み付き誤差関数を導入した DCNN による Heterogeneous Learning を提案した。提案手法を導入することで、高精度かつ高速に複数のタスクを識別できた。今後は、さらに高精度かつ効率的な DCNN による Heterogeneous Learning の学習を行う。

### 参考文献

- [1] X.Yang, et al., "Heterogeneous Multitask Learning with Joint Sparsity Constraints", NIPS, 2009.
- [2] Z. Zhang, et al., "Facial Landmark Detection by Deep Multi-task Learning", ECCV, 2014.