

## 1. はじめに

歩行者属性認識の従来手法として、Deep Single Attribute Recognition (DeepSAR) 及び Deep Multi Attribute Recognition (DeepMAR) [3] が提案されている。これらの手法は、様々な属性を同時に認識できる。しかしながら、それらの判断根拠は不明確である。そこで本研究では、判断根拠を可視化できる Attention Branch Network (ABN) [1] を用いて歩行者属性認識を行う。さらに各属性に対応する Multi-Attention maps を手動修正し再学習することで精度向上を図る。

## 2. Attention Branch Network

ABN は、推論時にネットワークが注視する領域を Attention map として可視化するとともに認識処理に活用することで高精度な画像認識を実現している。ABN は Feature extractor, Attention branch, Perception branch の3つから構成されている。Attention branch から Attention map を獲得でき、Feature extractor で得た特徴マップに Attention map を反映させる。これにより、重要な領域に着目した特徴となり、認識精度が向上する。ABN をマルチタスクに応用する場合、Perception branch をタスク数分用意し、各属性の認識をそれぞれ行う。

## 3. 提案手法

本研究では、歩行者属性認識の精度を向上させるために、Attention map を修正する。そして、図 1 のように修正した Multi-Attention maps を用い、マルチタスク化した ABN を再学習する。

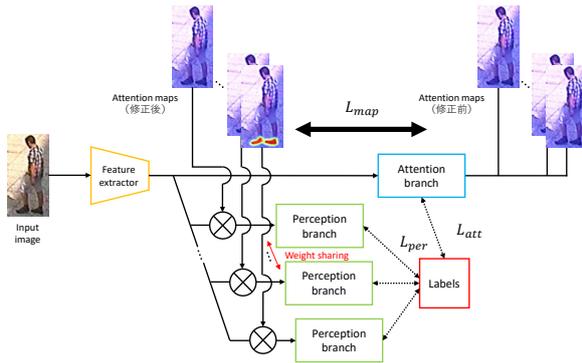


図 1: マルチタスク ABN の再学習法

### 3.1. Multi-Attention maps の手動修正

本稿では、ABN による歩行者属性認識において最も精度の低いクラスを手動修正の対象クラスとする。修正では、不要な注視領域の削除および理想的な注視領域の追加を行う。

### 3.2. ABN の再学習

ABN の再学習では、式 (1) に示すように ABN の学習誤差に Attention map の学習誤差  $L_{map}(x)$  を追加して学習する。

$$L_{all}(x) = L_{att}(x) + L_{per}(x) + L_{map}(x) \quad (1)$$

ここで  $L_{map}(x)$  は Attention map に対する誤差である。 $L_{map}(x)$  は、式 (2) に示すようにネットワークから出力される Attention map  $M(x)$  と修正した Attention map  $M'(x)$  との L2 ノルム誤差から算出する。

$$L_{map}(x) = \gamma \|M'(x) - M(x)\|_2 \quad (2)$$

ここで、 $\gamma$  は学習誤差  $L_{map}(x)$  を調整する係数である。

## 4. 評価実験

マルチタスク化した ABN による歩行者属性認識の有効性および提案手法の有効性を調査する。

### 4.1. 実験概要

本実験には、Pedestrian Attribute (PETA) dataset [2] を用いる。PETA dataset は、小規模な dataset の組み合わせで構成されている。多様なカメラ角度や解像度の異なる画像が含まれるため、汎化性能の高いモデルを学習することが可能となる。また、代表的な歩行者属性として、性別や年齢層、上半身及び下半身のスタイル、アクセサリなどがあり、合計 61 種類である。本実験では学習に 12,258 枚、評価に 7,600 枚を用いる。また、全てのデータのうちデータ数の偏りが少ない 35 クラスのみを用いて従来手法と ABN、提案手法の認識精度の比較及び、Attention map の比較を行う。

### 4.2. 手動修正

本手法では ABN の最も精度の低いタスクである footwear Shoes の学習サンプル 4,477 枚を手動で修正する。Attention map の修正例を図 2 に示す。



図 2: Attention map の修正例

### 4.3. 実験結果

従来手法とマルチタスク化した ABN および提案手法の属性クラス認識精度の平均を表 1 に示す。従来手法である DeepSAR 及び DeepMAR と比較すると、ABN は高い認識精度であることが分かる。また、提案手法である再学習後の ABN が最も高い認識精度を達成した。

表 1: 属性クラス認識精度の比較結果 [%]

DeepSAR [3]	DeepMAR [3]	ABN	提案手法
81.30	82.60	85.46	<b>85.61</b>

クラス認識精度の最も低い footwearShoes における ABN 及び提案手法の Attention map の可視化結果を図 3 に示す。図 3 より、再学習することで適切でない注視領域が抑制され、再学習前と比較して属性に関する領域を注視した Attention map を獲得できている。



図 3: Attention map の可視化結果

## 5. おわりに

本研究では、ABN を歩行者属性認識に適用した場合の有効性の調査及び手動修正した Multi-Attention maps を用いた ABN の再学習を行った。今後は提案手法の更なる高精度化を目指す。

## 参考文献

- [1] H. Fukui, *et al.*, "Attention Branch Network : Learning of Attention Mechanism for Visual Explanation", CVPR, pp.10705-10714, 2019.
- [2] Y. Deng, *et al.*, "Pedestrian Attribute Recognition At Far Distance", ACM MM, 2014.
- [3] L. Dangwei, *et al.*, "Multi-attribute Learning for Pedestrian Attribute Recognition in Surveillance Scenarios", IAPR, 2015.