

1. はじめに

統計的学習手法による人検出では、学習と異なる環境に適応させるには、追加学習を行う必要がある。文献 [2] では人体生成モデルを用いて、実環境に対応した学習サンプルを大量に生成し、追加学習することで人検出の高精度化を実現した。しかし、追加学習用のサンプル数が大量であるため、識別器の学習に時間を要するという問題がある。そこで、本研究では転移学習を利用することで少サンプルからの学習を実現する。

2. 転移学習による追加学習の効率化

提案手法では、人体生成モデルと転移学習を用いて追加学習の効率化を図る。図 1 に提案手法の流れを示す。まず事前学習として汎用的な実画像データベース T_a を用いて、AdaBoost の学習により識別器 H_a を構築する。次に、人体生成モデルを用いて実シーンに対応した少量の学習サンプル T_t を生成する。最後に生成した学習サンプルに H_a の情報を転移し CovBoost と呼ばれる学習法を用いて学習することで、実シーンに対応した識別器 H_t を構築する。

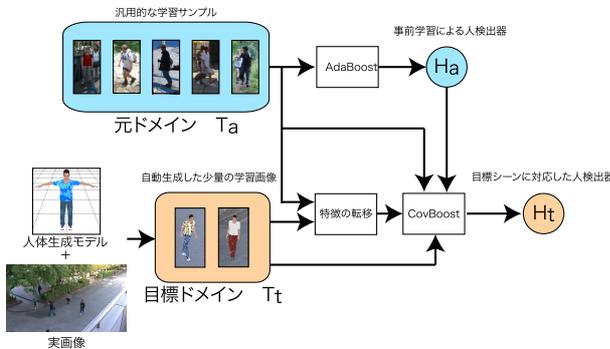


図 1：提案手法の流れ

2.1. 特徴の転移

事前学習により得られた弱識別器には、人と人以外をより良く分割できる局所領域の位置と、選択された特徴量を持つ。本手法では、この情報を転移する。特徴の転移は、図 2 のように、目標ドメイン T_t に事前学習で選択された弱識別器 (図 2(a)) の付近に候補領域を発生させ (図 2(c))、新しいシーンにおいて同様な特徴 (図 2(d)) がどこにあるかを求める。転移先の特徴の決定は、勾配方向ヒストグラムの Bhattacharyya 距離により決定する。特徴の転移は事前学習で選択された弱識別器の個数と同数行われ、選択された新しい特徴は、次項で説明する CovBoost の弱識別器候補 h_m となる。

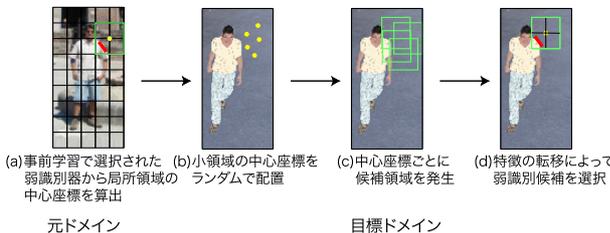


図 2：特徴の転移の流れ

2.2. 共変量を用いた Boosting 手法 (CovBoost)

Pang ら [1] によって提案された CovBoost は、事前に十分に学習された強識別器 $H_a(x)$ の知識を利用することで、少量の学習サンプルでも大量の学習サンプルを用いた時と変わらない性能を持った強識別器 $H_t(x)$ を学習する手法である。入りに I 個の目標となるシーンの学習画像 $x_i \in T_t$ とそのクラスラベル y_i と J 個の事前学習で用いた学習画像 $x_j \in T_a$ とそのクラスラベル y_j 、事前学習で構築した強識別器 $H_a(x)$ 、特徴の転移で選択した弱識別器候補 h_m を使用する。CovBoost は共変量 λ を考慮した損失関数である式 1 を最小化することが目的となる。ここで、

$$\lambda_j = \frac{1+e^{-y_j H_a(x_j)}}{1+e^{-y_j H_t(x_j)}} \text{ である}$$

$$\sum_{i=1}^I e^{-y_i H_t(x_i)} + \sum_{j=1}^J \lambda_j e^{-y_j H_t(x_j)} \quad (1)$$

以下に CovBoost の学習の流れを示す。

STEP1 弱識別器 $h^t(x)$ の決定 式 (1) の 1 次のテイラー展開を行い、最小となる $h_m^t(x)$ を求める。この STEP は事前学習の学習回数と同数繰り返される。

$$h_t = \arg \min_{h_t} \left(\sum_{i=1}^I e^{-2y_i D_t(x_i)} y_i h_t(x_i) + \sum_{j=1}^J \lambda_j e^{-2y_j D_a(x_j)} y_j h_t(x_j) \right) \quad (2)$$

STEP2 重み α_m の決定 エラー率 (式 (3)) をもとめ、式 (4) によって学習の重み α_m を計算する。

$$\epsilon_m = \frac{\sum_{h(x_i) \neq y_i} e^{-2y_i D_t(x_i)} + \sum_{h(x_j) \neq y_j} \lambda_j e^{-2y_j D_a(x_j)}}{\sum_i e^{-2y_i D_t(x_i)} + \sum_j \lambda_j e^{-2y_j D_a(x_j)}} \quad (3)$$

$$\alpha_m = \frac{1}{4} \ln \frac{1 - \epsilon_m}{\epsilon_m} \quad (4)$$

STEP3 重みの更新 式 (5)(6) により重み D_t と D_a を更新

$$D_t(x_i) = D_t(x_i) e^{-y_i \alpha^t h^t(x_i)} \quad (5)$$

$$D_a(x_j) = D_a(x_j) e^{-y_j \alpha^t h^t(x_j)} \quad (6)$$

STEP4 強識別器の構築 STEP2 から STEP4 を繰り返すことにより、最終識別器 $H_t(x) = \sum_m \alpha_m^t h_m^t(x)$ を得る。

3. 評価実験・考察

本手法の評価は AdaBoost と CovBoost に同じサンプルを入力し、人検出の精度の比較を行う。事前学習にはカメラの俯角が 21° で生成した学習画像を入力する。本研究の有用性を示すため新しいシーンとして、カメラの俯角を 41° とした学習画像を入力し検出精度の比較を行う。ただし、入力する学習画像は既存手法である AdaBoost は学習画像を 2416 枚、提案手法である CovBoost は 800 枚で実験を行う。図 3 に AdaBoost によって人検出を行った結果と、CovBoost によって人検出を行った結果の検出精度の比較を示す。

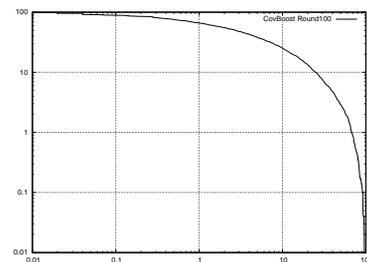


図 3：結果の比較

4. おわりに

提案手法では、学習画像の枚数を減らしたところ AdaBoost との検出精度はおよそ 100 分の 1 の検出の低下となっていました。

参考文献

[1] Le ,, Qin, Huang ,Pang , Jiang“ Treat samples differently: Object tracking with semi-supervised online CovBoost ” 2011 IEEE CVPR. pp. 627-634, 2011.
 [2] Yamauchi, Fujiyoshi “Automatic Generation of Training Samples and a Learning Method Based on Advanced MILBoost for Human Detection ”ACPR, pp. 603-607, 2011 .