

共変量シフトを用いた Random Forest による転移学習

土屋成光^{1,a)} 藤吉弘亘^{1,b)}

1. はじめに

Random Forest (RF) [1] はランダムネスにより高い汎化性能をもつマルチクラス識別器である。コンピュータビジョンの分野では、画像分類、セマンティックセグメンテーション、キーポイント検出、人検出などに利用されている。しかし、物体検出においては学習時と実際に使用する環境でのデータの分布が一般に異なり、その性能低下を招く。そのため、実際に使用する対象のシーンでサンプルを再採集する必要があるが、これは非常に人的コストが高い。サンプル収集では、サンプル画像内の物体の位置やスケールを揃えることが求められるためである。このような、異なる分布を持つ複数ドメイン（目標ドメイン、事前ドメイン）のサンプルに対し、一方から一方へ「転移」を行うことで効率的にシステムを構築する、転移学習手法が提案されている [2][3]。これらは、物体検出問題において、目的とするシーンからの新規データの再収集をより省力化し、少数の目標シーンサンプルと多数の事前シーンサンプルによって、シーン間のカメラ俯角や照明の差異に対応している。そこで本稿では、RF への、共変量シフトによる転移学習 [3] の導入を提案する。

2. 共変量シフトを用いた Random Forest による転移学習

提案手法では、学習前に事前ドメインとそれを学習した RF を既得とし、事前ドメインと目標ドメイン両方のサンプルを学習する。その際、両ドメインのサンプルは木毎に同数、ランダムに選択され、サブセットを構成する。学習過程において、目標ドメイン $X_t \ni x_i$ と事前ドメイン $X_a \ni x_j$ の分布の差に応じて事前ドメインに λ による重み付けを行い、事前ドメインサンプルに重み付けを行う。これにより、目標ドメインの学習に有効な事前ドメインサンプルを取り込み、齟齬のあるサンプルの影響を低減する。学習の流れを図 1 に示す。

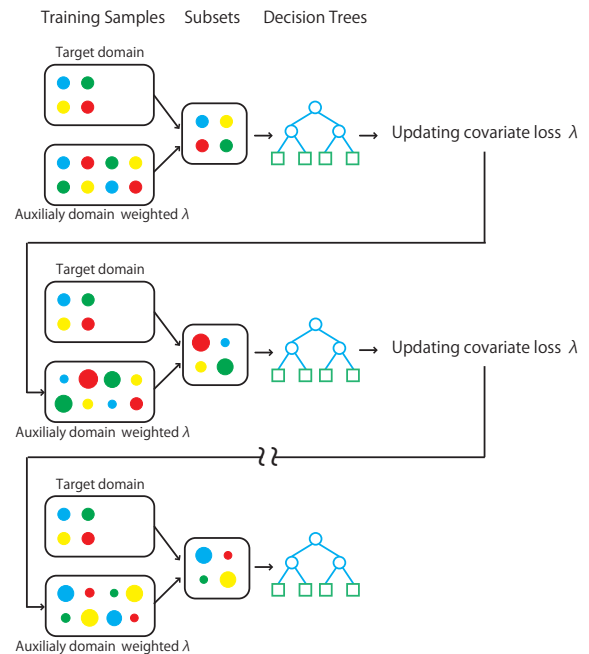


図 1 転移学習を導入した Random Forest の学習

2.1 Random Forest への共変量損失 λ の導入

Forest の構築は RF 同様ランダムに特徴を選択し、末端ノードに到達した学習サンプル群から $c \in C$, $p(c|x)$ によってクラス確率を生成する。重み λ は Phang らの提案する Covariate shift Boost[3] で用いられているものをマルチクラス識別器である RF に合わせ次式で求める。

$$\lambda_j = \frac{1 + e^{P_a(c_k|L, x_j)}}{1 + e^{P_t(c_k|L, x_j)}} \quad (1)$$

ここで、 c_k は事前サンプル x_j の正解クラスラベル、 L は到達したリーフノード、 $P_a(c_k|L, x_j)$ は既得の事前ドメインサンプルによって学習された RF の正解ラベル確率、 $P_t(c_k|L, x_j)$ は現在構築中の Forest の正解ラベル確率である。これにより、学習中、目標サンプルに不足している事前サンプルを重視し、目標サンプルと齟齬のあるサンプルをリジェクトする。

2.2 λ を用いた木構造の分岐関数

木構造の分岐関数は、Covariate loss λ を考慮して事前サンプルについては重み付きで評価を行う。Forest の分岐関数の評価式は次式で示される。

¹ 中部大学, 〒 487-8501 愛知県春日井市松本町 1200

^{a)} tsuchiya@vision.cs.chubu.ac.jp

^{b)} hf@cs.chubu.ac.jp

$$\Delta E = \frac{|I_l|}{|I|} E(I_l) \frac{|I_r|}{|I|} E(I_r) \quad (2)$$

$E(I_l)$, $E(I_r)$ は選択された分岐関数候補におけるシャノンのエントロピーである。シャノンエントロピーは次式で示される。

$$E(I) = \sum_{k=1}^C P(c_k) \log P(c_k) \quad (3)$$

$P(c_k)$ は各クラスに属するサンプルから計算されるクラス c_k のクラス確率であり、次式で計算される。

$$P(c_k) = \frac{|I_{c_k}|_t + |I_{c_k}|_a}{|I|}. \quad (4)$$

ここで、サブセット内の目標ドメインに関する評価 $|I_{c_k}|_t$ は通常の RF 同様分岐されたサブセット内に存在する全目標サンプル数と、 c_k に属するサンプルの比に基づいて計算されるが、事前ドメインに対する評価 $|I_{c_k}|_a$ は全事前サンプル、 c_k に属するサンプルに対し、それぞれの λ の総和の比を求める。末端ノードが持つクラス確率についても同様に、末端ノードに到達したサンプル群から式 (4) を用いて各クラスのクラス確率が算出される。各木の出力の結合は RF と同様である。また、 λ は逐次的に適応を行うため、 T 本の木で構成された Forest を構築する場合、 $2T$ 本の木を構築し後半の T 本を使用する。これは、同様に逐次的な転移学習である [2] で用いられている方法と同様である。

3. 実験

実験として、2つのドメインから提案手法、マルチタスク学習、再学習をそれぞれ行い性能を比較する。マルチタスク学習は両ドメインを足しあわせた集合を学習サンプルとし、重み付けなどは行わず RF を構築する。再学習は目標ドメインのサンプルのみで RF を構築するものである。物体検出問題としては HOG 特徴量による人検出を用いる [4]。表 1 に、実験に用いたツリーパラメータを示す。また、パラメータは各手法で同一である。学習用画像として、

Number of tree	50
Max depth of tree	5
Number of feature	3780
Number of feature selection trial	63($\cong \sqrt{3780}$)

目標ドメインに INRIA Person Dataset[4] から人画像 2,416 枚、非人画像 1,238 枚、事前ドメインに DaimlerChrysler Mono Pedestrian Detection Benchmark Dataset[5] から人画像 15,600 枚、非人画像 67,000 枚を使用した。また、評価用画像として INRIA Person Dataset から、1,135 枚の人画像と 453 枚の背景画像を用いた。

3.1 実験結果

各手法に対し、目標ドメインの人画像を 2,416 枚から 800

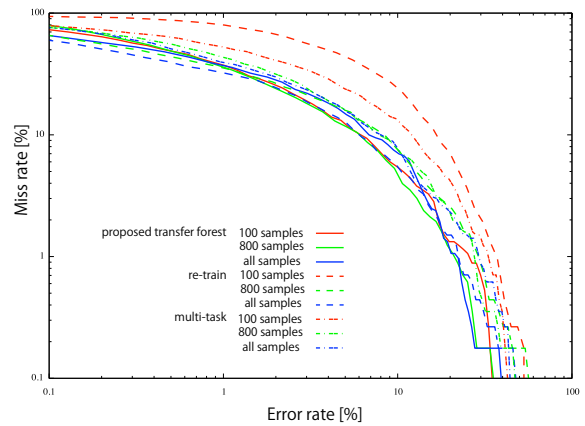


図 2 識別精度の比較

枚、100 枚へと減少させた際の識別精度を図 2 に示す。100 枚まで減少させた際、再学習では識別性能が大きく低下しており、マルチタスク学習ではその低下を軽減できている。それに対し、提案手法は全データを使用したものとほぼ変わらない性能である。また、全データを使用するとマルチタスク学習では事前ドメインが障害となり、再学習より性能が低いことに対し、提案手法は性能を安定して発揮できている。

4. おわりに

本稿では、共変量シフトによる転移学習を導入した RF フレームワークを提案した。提案手法は複数ドメインのサンプルに対する効率的な学習を可能とし、実験において、再学習やマルチタスク学習に比べ、より少数のサンプルにおいて安定的に高精度を示した。今後は、マルチカテゴリの画像分類など、マルチクラス、マルチドメイン問題における性能評価を行う。また、サンプルだけでなく、事前に構築した RF の木構造識別器群を再利用する更なる省力化を検討する。

参考文献

- [1] L. Breiman. “Random forests.” *Machine Learning*, 45(1):5-32, October 2001
- [2] W. Dai, Q. Yang, G. Xue and Y. Yu, “Boosting for Transfer Learning”. *International Conference on Machine Learning*. 2007.
- [3] J. Pang, Q. Huang, S. Yan, S. Jiang and L. Qin, “Transferring Boosted Detectors Towards Viewpoint and Scene Adaptiveness”, *IEEE Trans. on Image Processing* 20(5): pp.1388-1400, 2011.
- [4] N. Dalal and B. Triggs: “Histograms of oriented gradients for human detection”, *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 886-893, 2005.
- [5] M. Enzweiler and D. M. Gavrila. *Monocular Pedestrian Detection: Survey and Experiments* *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.31, no.12, pp.2179-2195, 2009.