Confident Anchor-Induced Multi-Source Free Domain Adaptation

Jiahua Dong

NEURAL INFORMATION PROCESSING SYSTEMS

December 2021

Overview

1. Introduction
2. Theoretical Analysis
3. Algorithm Implementation
4. Experiments

Introduction

Given a feature space $\mathcal{X} \subset \mathbb{R}^d$ and a label space $\mathcal{Y} = [K]$, a **domain** is a joint distribution P_{XY}, where random variables $X \in \mathcal{X}$, $Y \in \mathcal{X}$.

Definition (Multi-Source Free Domain Adaptation (MSFDA))

Given n source domains $\{P^i_{XY}\}_{i=1}^n$ and a target domain P^t_{XY},

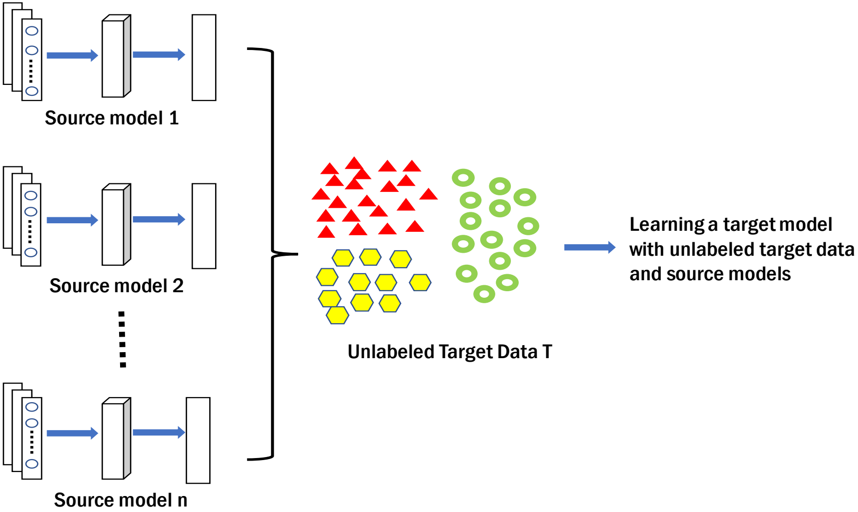
Input:

1. for any source domain P^i_{XY}, a neural network-based predictor (model) $\boldsymbol{h}^i : \mathcal{X} \to \mathbb{R}^{|\mathcal{Y}|} = \mathbb{R}^K$ is given;
2. unlabeled data $T = \{\mathbf{x}^j\}_{j=1}^m \sim P^t_X$, i.i.d is given.

Aim:

learn a target domain classifier to classify the unlabeled target data T by utilizing T and predictors $\{\boldsymbol{h}^i\}_{i=1}^n$.

Introduction



Theoretical Analysis

Basic Strategy:

- Predict **high-confident pseudo labels** using source models.
- Use high-confident pseudo labels to help learn target model.

Aim of Theoretical Analysis:

- Understand why multiple source models benefit the classification for target domain.
- Study the **effectiveness** of high-confident pseudo labels strategy.
- Investigate the basic conditions to ensure the **solvability** of multi-source free domain adaptation.

Theoretical Analysis

Meta Assumption

$P^t_{XY}, P^1_{XY}, ..., P^n_{XY}$ are drawn (i.i.d.) from a meta distribution $\mathcal{P}$, which is defined over a joint distribution space $\mathcal{P}_{XY}$.

Theorem 1

Given the Meta Assumption and some mild assumptions. Given $\eta > 0$, if $m \geq n$ and $(1-\eta)(1-\tau) > \epsilon + 2\sigma + 2\sqrt{\log(2m/\delta)/2m}$, then with probability at least $1 - \delta - g(\sigma)^n > 0$, **at least ηm target data have τ-confident pseudo labels**, where ϵ is the upper bound of the accuracies of source predictors, σ is a small constant and g is a decreasing function with $g(0) = 0$ and $g(\sigma) > 0$, if $\sigma > 0$.

Theorem 1 implies that more source models improve the probability to obtain more high-confident pseudo labels.

Theoretical Analysis

Theorem 2

Given some mild assumptions and suppose that high-confident pseudo labels are true labels, with probability at least $1 - \delta - g(\sigma)^n > 0$, for any $\boldsymbol{h}$ from hypothesis space $\mathcal{H}$ with finite Natarajan dimension

$$|\mathrm{err}(\boldsymbol{h}) - \widehat{\mathrm{err}}^{\tau}(\boldsymbol{h})| \leq C\sqrt{1/m} + C(\sigma, \delta, \epsilon, \tau),$$

where $\mathrm{err}(\boldsymbol{h})$ is the error for $\boldsymbol{h}$, $\widehat{\mathrm{err}}^{\tau}(\boldsymbol{h})$ is the empirical error for $\boldsymbol{h}$ predicted by high-confident pseudo labels with threshold τ and $C(\sigma, \delta, \epsilon)$ is a small constant related to σ, δ, threshold τ and source model's error ϵ.

Above theorem indicates that under some mild assumptions, if predicted high-confident pseudo labels are true labels, then the error of target domain can be approximated by the empirical error predicted by pseudo labels.

Contributions of Our Theory

- Our theory provides the first theorem to estimate the low bound for the number of high-confident pseudo labels.
- Our theory shows that multiple source models benefit the number of high-confident pseudo labels.
- Our theory shows that under the meta assumption, multi-source free domain adaptation is solvable, if the high-confident pseudo labels are true labels.

Algorithm Implementation

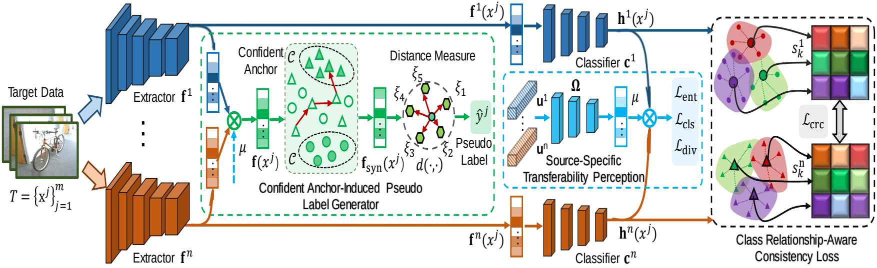
Three components:

- **Source-Specific Transferability Perception.**
  Aim to **quantify the transferability contributions** of complementary knowledge from source domains.

- **Confident Anchor-Induced Pseudo Label Generator.**
  Aim to **mine confident pseudo labels** for target data.

- **Class Relationship-Aware Consistency Loss.**
  Ensure the semantic consistency of underlying inter-class relationships across domains to promote **more shared transferable knowledge** from source domains towards target adaptation.

Based on the three components, we develop a **Confident Anchor-induced multi-source free Domain Adaptation** (CAiDA) model.



Experiments

We conduct experiments on datasets: **Office-31**, **Office-Caltech**, **Office-Home** and **Digits-Five**.

Table: Comparisons between our model and other competing methods on Digits dataset.

Methods	R → MM	R → MT	R → UP	R → SV	R → SY	Avg.
Source only	63.4	90.5	88.7	63.5	82.4	77.7
MDAN	69.5	98.0	92.4	69.2	87.4	83.3
DCTN	70.5	96.2	92.8	77.6	86.8	84.8
M^3SDA	72.8	98.4	96.1	81.3	89.6	87.7
MDDA	78.6	98.8	93.9	79.3	89.7	88.1
LtC-MSDA	85.6	99.0	98.3	83.2	93.0	91.8
Source model only	25.2	90.0	93.3	42.8	77.8	65.8
BAIT	87.6	96.2	96.7	60.6	90.5	86.3
PrDA	86.2	95.4	95.8	57.4	84.8	83.9
SHOT	90.4	98.9	97.7	58.3	83.9	85.8
MA	90.8	98.4	98.0	59.1	84.5	86.2
DECISION	93.0	**99.2**	97.8	82.6	97.5	94.0
Ours	**93.7**	99.1	**98.6**	**83.3**	**98.1**	**94.6**

Thank You !

dongjiahua@sia.cn