Adversarial Multiple Source Domain Adaptation

Han Zhao†, Shanghang Zhang†, Guanhang Wu†, João Costeira†, José Moura† & Geoffrey J. Gordon†
†Carnegie Mellon University, ‡Instituto Superior Técnico
{hzhao1, shanghaz, guanhanw, moura, ggordon}@andrew.cmu.edu, jpc@isr.ist.utl.pt

Summary
Unsupervised Domain adaptation: Source ≠ Target

- We theoretically analyze the multiple source domain adaptation problem under both classification and regression settings.
- We propose two models using adversarial neural networks for multiple source domain adaptation.
- We conduct extensive experiments on sentiment analysis, digit recognition and vehicle counting problems, and we achieve superior adaptation performances on all the tasks.

Preliminary

Given hypothesis class $\mathcal{H}$ and $A_H := \{h^{-1}(1) \mid h \in \mathcal{H}\}$, $\mathcal{H}$-divergence is: $d_H(D, D') := 2 \sup_{h \in A_H} | \Pr(D) - \Pr(A) |$.

Generalization bound for single-source-single-target binary classification (Blitzer et al. NIPS’ 08), using $m$ instances, with probability $\geq 1 - \delta$, $\forall h \in \mathcal{H}$:

$$
\epsilon_T(h) \leq \bar{\epsilon}_S(h) + \frac{1}{2}d_{\Delta H}(\hat{D}_S; \hat{D}_T) + \lambda + O\left(\frac{d \log (1/\delta)}{m}\right)
$$

$\bar{\epsilon}_S(h) = \frac{1}{2}d_{\Delta H}(\hat{D}_S; \hat{D}_T)$: empirical population source/target binary classification error.

$\lambda := \min_{h \in \mathcal{H}} \epsilon(h) + \epsilon(h')$.

A naive extension to $k$ source domains with union bound:

$$
\epsilon_T(h) \leq \max_{i \in [k]} \left\{ \bar{\epsilon}_S(h) + \frac{1}{2}d_{\Delta H}(\hat{D}_S; \hat{D}_T) + \lambda \right\} + O\left(\frac{1}{m} \log \frac{k}{\delta} + d \log \frac{m}{d}\right)
$$

Models and Algorithms

Desired Task

Domain label $(S_1, T)$

Domain label $(S_2, T)$

Domain label $(S_k, T)$

Source cameras: with label

Target camera: without label

Theorem (informal): $\mathcal{H}$ is a hypothesis class and $\forall c \subseteq 2^d$, $\bar{\epsilon}_T$ is the classification error. Using $H$-divergence, we prove a generalization bound for single-source-single-target binary classification.

$$
\epsilon_T(h) \leq \sum_{i \in [k]} \alpha_i (\bar{\epsilon}_S(h) + \frac{1}{2}d_{\Delta H}(\hat{D}_T; \hat{D}_S)) + \lambda + O\left(\frac{d \log (1/\delta)}{km}\right)
$$

Datasets:

- WebCamT (Zhang et al., CVPR’ 17), public dataset for vehicle counting. Image resolution: $352 \times 240$.

Experiments

- 8 cameras. 6 as sources and each of the rest two as target. 2,000 images for each domain.

Methods:

- FCN: Fully-convolutional NN, without domain adaptation.
- DANN: Combine all sources into one, with adversarial learning.

Table: Counting error statistics. $S$ is the number of source cameras; $T$ is the target camera id.

<table>
<thead>
<tr>
<th>S</th>
<th>T</th>
<th>Ours</th>
<th>DANN</th>
<th>FCN</th>
<th>T</th>
<th>Ours</th>
<th>DANN</th>
<th>FCN</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1</td>
<td>1.8101</td>
<td>1.7140</td>
<td>1.9490</td>
<td>1.9094</td>
<td>B</td>
<td>2.5059</td>
<td>2.1435</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1.3776</td>
<td>1.2363</td>
<td>1.3683</td>
<td>1.5545</td>
<td>B</td>
<td>1.9092</td>
<td>1.8680</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1.3868</td>
<td>1.1965</td>
<td>1.5520</td>
<td>1.5499</td>
<td>B</td>
<td>1.7375</td>
<td>1.8487</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>1.4021</td>
<td>1.1942</td>
<td>1.4156</td>
<td>1.7925</td>
<td>B</td>
<td>1.7758</td>
<td>1.6016</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>1.4359</td>
<td>1.2877</td>
<td>2.0298</td>
<td>1.7505</td>
<td>B</td>
<td>1.5912</td>
<td>1.4644</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>1.4381</td>
<td>1.2984</td>
<td>1.5426</td>
<td>1.7646</td>
<td>B</td>
<td>1.5989</td>
<td>1.5126</td>
</tr>
</tbody>
</table>

Reference

- Blitzer et al., Learning bounds for domain adaptation, NIPS 2010.
- Zhang et al., Understanding traffic density from large-scale web camera data, CVPR 2017.