Adversarially Constrained Interpolation for Unsupervised Domain Adaptation

Mohamed Azzam¹; Aurele Tohokantche Gnanha¹; Hau-San Wong¹; Si Wu² ¹City University of Hong Kong, ²South China University of Technology {m.azzam, tagnanha2-c}@my.cityu.edu.hk, cshswong@cityu.edu.hk, cswusi@scut.edu.cn



Problem

We solve the problem of unsupervised domain adaptation, which aims to adapt models trained on source data (labeled) in order to perform well on a target data (unlabeled).

City University of Hong Kong

╋╋



MNIST → SVHN Source → Target

01

neighbors

C

0 0

Source domain (samples) 🥥 🛽

Previous Domain Adaptation Methods

Target domain (samples)

Decision boundary

0

1

Dom. Mixup samples (same class)

Dom. Mixup sample (diff. class)

Motivation

Common solution: Learning domain-invariant representations using domain adversarial Training.

The current domain adversarial training methods face two challenges:

- Data Inefficiency from the source and target domain
- Target-domain samples of different classes may become neighbors in the feature space



- Using Inter-domain mixup samples to train the model
- Enforcing the domain discriminator to predict the mixup ratio to facilitate the training

Inter-domain mixup samples obtain better representation through:

- 1. Binding samples of the same class together in a single cluster
- 2. Encouraging smooth transition between clusters



Proposed Method



Results

The table shows the classification accuracy (%)of models tested on the target data

Source data Target data	MNIST SVHN	<mark>SVHN</mark> MNIST	<mark>MNIST</mark> MNIST-M	CIFAR-10 STL-10	STL-10 CIFAR-10
Source-only	40.9	82.4	59.9	76.3	63.6
DANN	35.7	73.9	77.4	-	-
VADA	73.3	97.9	95.7	80.0	73.5
VMT	85.2	98.9	98.0	82.0	78.5
IIMT	-	97.3	99.5	83.1	81.6
Ours	88.7	99.0	98.1	83.7	79.7

Figures a & b shows the t-SNE plot of the last hidden layer of C in Source-only method and our method, respectively

