VisDA Classification Challenge: Honorable Mention Talk

*Kaichao, You *Zhangjie, Cao Hong, Liu Mingsheng, Long

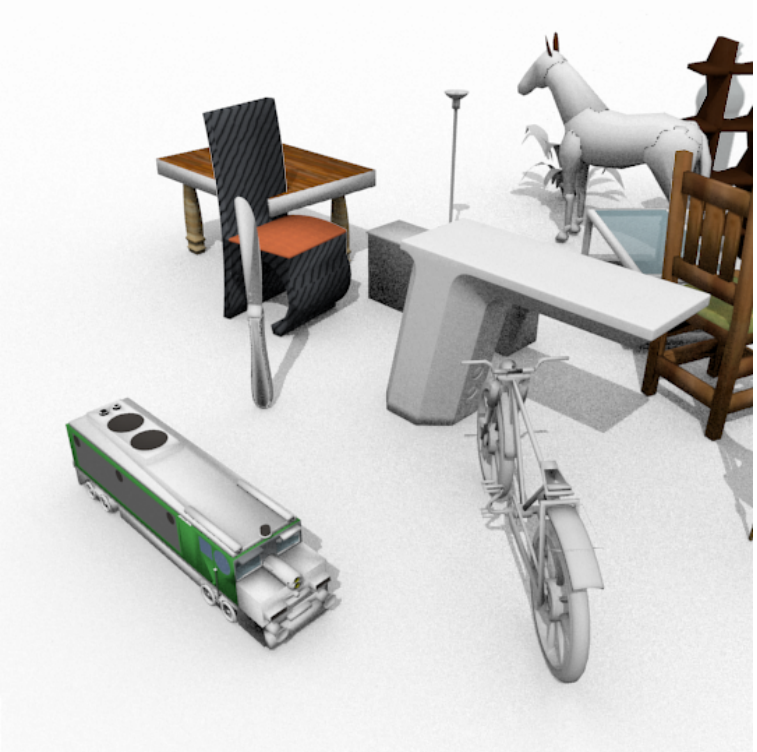
Tsinghua University, China

September 14, 2018

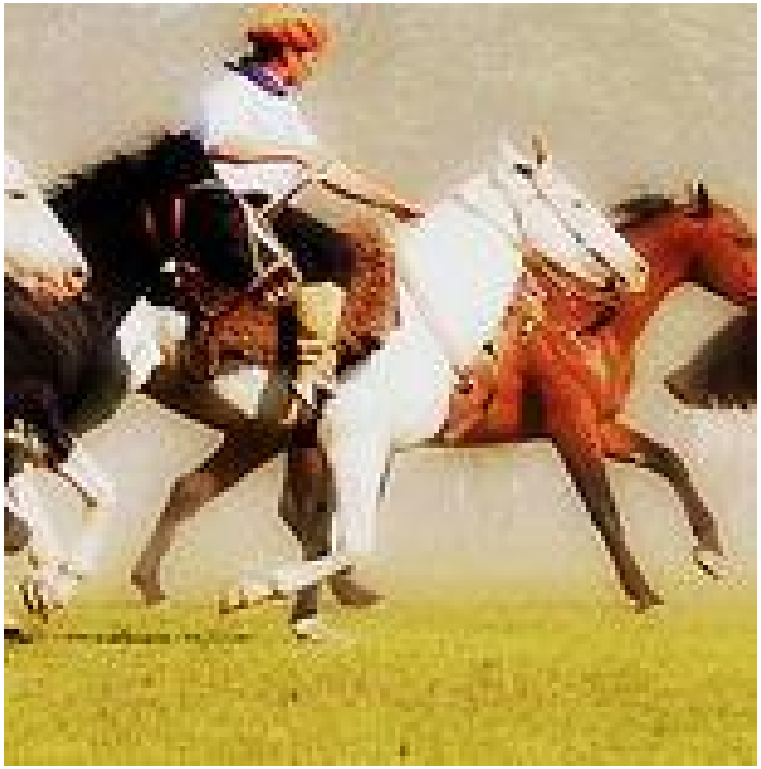
Logic:

- Challenges
- Solutions

Relation Shift



(a) horse in training set



(b) horse in validation set

Figure: examples showing relation-shift problem in VisDA2018

Relation Shift

- Training data is generated by render engines of games
- Each single object looks real
- Object relation is not the same as that in reality
- Domain adaptation model may suffer from such relation shift

Relation Shift

- Refine training data in an automatic way
- Images with multiple objects often have low confidence on each class.
- Train a 13-way classifier on source with denoise cross entropy loss

$$L = \frac{1}{n}\sum_{i=1}^{n} max\{L_i, \gamma\} \tag{1}$$

  - n is the mini-batch size
  - L_i is the original cross entropy loss for example i
  - γ is progressively adjusted
  - noisy examples are ignored
- Images with a single object would have high confidence scores
- Keep those images with only a single object by controlling confidence threshold

Overwhelming Target Unkown Examples

By analysing validation data, we find that:

- $\frac{\#unknown}{\#known} \approx 10$
- $\forall 1 \leq i \leq 12, \frac{\#unknown}{\#known_i} \approx 100$ (There are 12 classes which are known)
- Extreme class unbalance which is hard to tackle

High risk of negative transfer:

- Standard domain adaptation methods will matching the overwhelming unknown target class data with source data
- Images in common label space will be ignored due to their small proportion

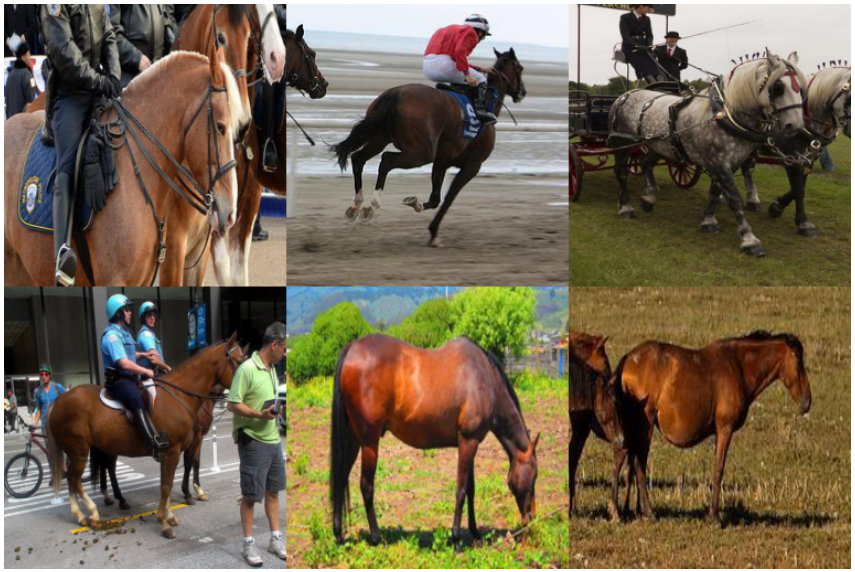
Overwhelming Target Unkown Examples

Exclude target unknown class in training process.

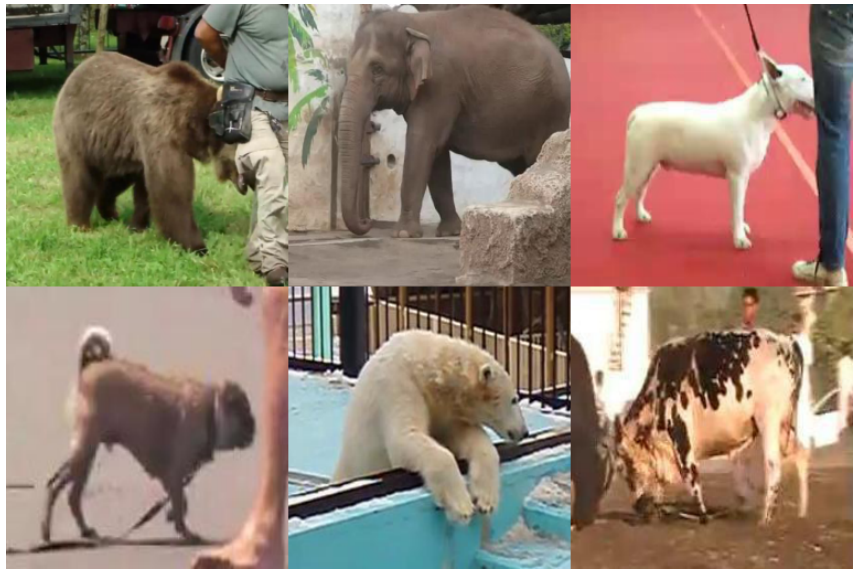
1. Train a 12-way classification model on refined source data and apply it to target domain
2. Select out those images with highest confidence
3. Train a 12-way classification model with selected images and refined images and apply it to target domain
4. Go to step 1 and repeat several times
5. Obtain target images with high confidence from known class
6. Label these target images with sudo-label predicted by our classifier
7. Semi-supervised domain adaptation using source images of known classes and selected target images with sudo-label of high confidence
8. Images with low confidence score are classified as unknown class

Foo/Bar-alike Images

- Can't tell horses from dogs when there are only horses in training set
- An intrinsic problem when models trained on closed set are applied to open set classification
- Treat horse-alike images as horse.



(a) horse in validation set



(b) horse-alike in validation set(labeled with unknown class)

Leaderboard

#	User	Entries	Date of Last Entry	Per Category Accuracy												Known ▲	Mean ▲
				plane ▲	bcycl ▲	bus ▲	car ▲	horse ▲	horse ▲	mcycl ▲	person ▲	plant ▲	sktbd ▲	train ▲	truck ▲	unknown ▲		
1	Yingwei.Pan	13	08/27/18	95.8 (1)	93.5 (1)	94.3 (2)	98.6 (1)	93.5 (1)	98.5 (1)	91.5 (1)	82.3 (1)	97.2 (1)	91.7 (1)	93.3 (1)	92.3 (1)	77.2 (1)	93.5 (1)	92.3 (1)
2	lianqing	23	08/28/18	91.0 (3)	76.6 (4)	86.5 (7)	94.5 (3)	83.1 (2)	36.9 (4)	83.4 (4)	69.8 (2)	90.0 (4)	40.8 (4)	69.1 (4)	65.4 (2)	9.8 (12)	73.9 (2)	69.0 (2)
3	youkaichao	15	08/28/18	94.0 (2)	79.2 (3)	90.5 (4)	97.4 (2)	63.2 (4)	36.0 (5)	81.1 (7)	45.6 (4)	93.0 (3)	35.4 (7)	84.5 (2)	46.2 (6)	42.2 (8)	70.5 (3)	68.3 (3)

Observations:

- Our "unknown" accuracy is higher due to the special disposal of "unknown" category.
- Accuracies for "horse" abd "person" are still not satisfying due to Foo/Bar-alike images