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
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 \(i\)
- \(\gamma\) 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
By analysing validation data, we find that:

- \( \frac{\# \text{unknown}}{\# \text{known}} \approx 10 \)
- \( \forall 1 \leq i \leq 12, \frac{\# \text{unknown}}{\# \text{known}_i} \approx 100 \) (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
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 between 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)
Observations:

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