Transfer learning through self-learning

Course project for *Machine Learning for Computer Vision*, under *Professor Vinay Namboodiri*

- Chirag Gupta
- Akshay Aggarwal

Transfer learning through self-learning

- Target class contains very few labeled samples
- Auxiliary, related class available with large number of samples
- Cats and dogs, cars and bikes
- **Build a robust classifier for Class 1, <u>adapt</u> it to Class 2**

- Learn a classifier that gives scores to data-points
- Predict scores for out-of-sample points
- Tag high-confidence positive (high positive score) and negative (high negative score) data-points
- Use this data for re-learning

'Self-learn' the classifier of domain 1 on domain 2.

Original Paper: Online Domain Adaptation of a Pre-Trained Cascade of Classifiers

Problem : Face Detection

Idea : Image specific domain adaptation

- Assume high confidence faces are true faces.
- Use them to detect other faces.

Online domain adaptation of a pre-trained cascade of classifiers, Vidit Jain and Erik Learned-Miller

Original Paper: Online Domain Adaptation of a Pre-Trained Cascade of Classifiers

Details :

- Standard face detection model - given bounding boxes, identify those that contain faces.
- Typically, large number of candidate bounding boxes.
- **Pre-Trained Cascade (Viola Jones classifier) :**
  - Boost simple linear classifiers. Use in a cascade.
  - Use fewer features in early stages, and optimize to minimize false negatives (lower the threshold).
  - Thus, prune down the possible number of bounding boxes.

- Once candidate bounding boxes are low, learn a GPR model.

Our work : Explore this idea in a standard transfer learning, non-online setting

Cats → Dogs

Dataset : PASCAL VOC 2007

Training : 15 random images each of only-cats(OC) and only-dogs (OD) + 30 random non-cat, non-dog images (NCND)

Testing : 30 random images each of only-cats(OC) and only-dogs (OD) + 60 random non-cat, non-dog images (NCND)

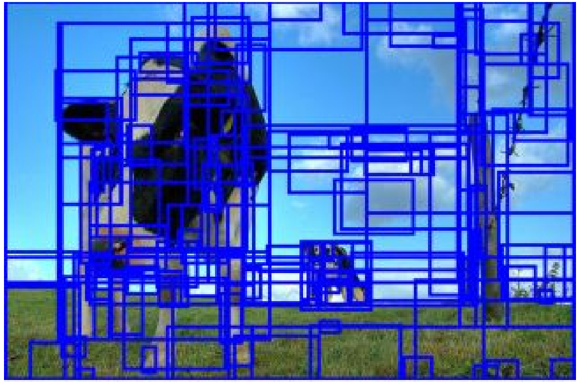
Model : Latent SVM for learning

1. Train a model for object detection of cats on 15 OC + 15 NCND
2. Predict using learnt model on 30 OD + 30 NCND
3. Again train a fresh model on the the top 20 predictions (by confidence)
4. Final model outputted and tested on the same test

Bounding boxes for object detection

Standard approach :

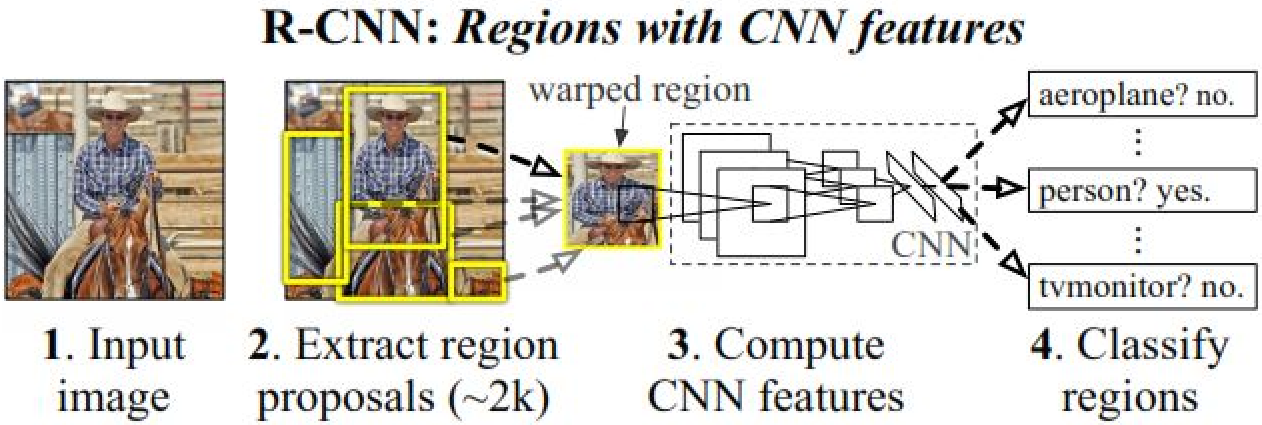
1. Break image into candidate bounding boxes.
2. Extract features for each bounding box.
3. Take a max on the bounding boxes.



Features could be SIFT, HOG, CNN features.

Bounding boxes for object detection

R-CNN features:



Segmentation As Selective Search for Object Recognition
K. E. A. van de Sande, J. R. R. Uijlings, T. Gevers, A. W. M. Smeulders

Latent SVM for object detection

Dataset does not provide boxes, only 0-1 tag.

→ Bounding boxes act as latent variables.

$$\min_{\boldsymbol{w}} \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}\left(\max_{(\hat{y},\hat{h})\in\mathcal{Y}\times\mathcal{H}}[\boldsymbol{w}\cdot\Phi(x_i,\hat{y},\hat{h}) + \Delta(y_i,\hat{y},\hat{h})]\right)$$

$$- C\sum_{i=1}^{n}\left(\max_{h\in\mathcal{H}} \boldsymbol{w}\cdot\Phi(x_i,y_i,h)\right).$$

Latent SVMs: Learning Structural SVMs with Latent Variables, Chun-Nam John Yu and Thorsten Joachims

For Object Detection: Object and Action Classification with Latent Variables, Hakan Bilen, Vinay P. Namboodiri, Luc J. Van Gool

Latent SVM for cat detection

Initial Model trained using 15 OC and 15 NCND. (CM)

Accuracy of CM on cat-test (30OC, 30NCND) : 91.6 %

Accuracy of CM on dog-test (30OD, 30NCND) : 70 %

New model trained on top 20 predictions of above^ (NDM)

Accuracy of DM on dog-test (30OD, 30NCND) : 73 %

Another model trained using 15 OD and 15 NCND. (DM)

Accuracy of DM on dog-test (30OD, 30NCND) : 87 %

Dog classified as cat with high confidence

Dog classified as cat with low confidence

Dog classified as non-cat with low confidence

Dog classified as non-cat with high confidence

Classified as Cat with some confidence !

Classified as non-cat with high confidence

Image classified negative by CM but positive by NDM

References

- Jain, Vidit, and Erik Learned-Miller. "Online domain adaptation of a pre-trained cascade of classifiers." *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011.
- Yu, Chun-Nam John, and Thorsten Joachims. "Learning structural SVMs with latent variables." *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 2009.
- Bilen, Hakan, Vinay P. Namboodiri, and Luc J. Van Gool. "Object and Action Classification with Latent Variables." *BMVC*. Vol. 2. 2011.
- Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014.