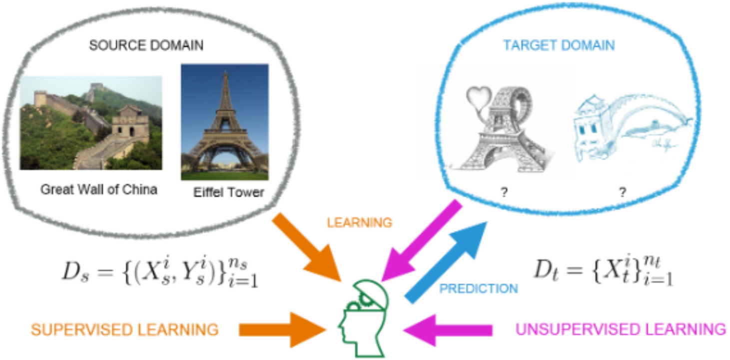
Domain adaptation (DA)



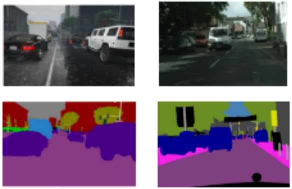
Leveraging labeled **source domain**, to learn a model for the **target domain**.

Example scenarios

Recognition

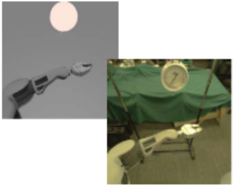
Detection

Segmentation



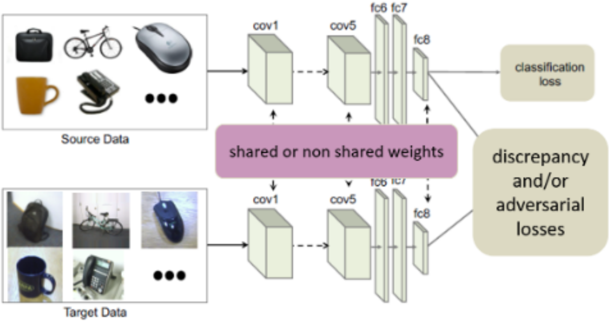
Re-identification

Control



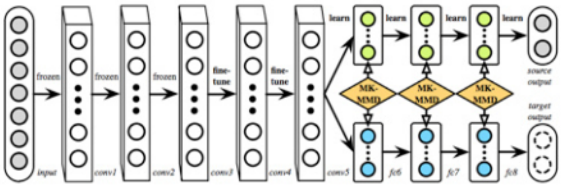
Visual localization

Discriminative models



- Siamese network, one source and one target stream
  - Both stream initialised with the pretrained-model on the source
- Classification (cross-entropy) loss on the source
- Domain alignment:
  - minimizing the distribution discrepancy
  - adversarial domain confusion

Minimizing feature distribution discrepancy



- Kernelized MMD loss, DAN (Long+@ICML'15)

$$MMD(S,T) = \sum_{l=1}^{L} \| \mathbb{E}(\phi(M_S^l)) - \mathbb{E}(\phi(M_T^l)) \|_2$$

where ϕ is a kernel projection and $\mathbb{E}(X) = \frac{1}{|X|}\sum_{x \in X}$ is the empirical expectation.

- Weighted discrepency, WDAN (Yan+@CVPR'17)

Adversarial learning

Principles of GAN

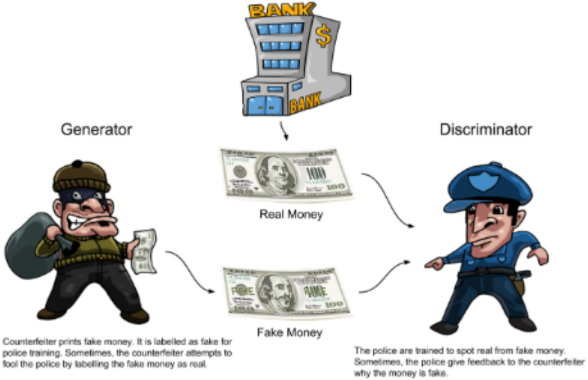


Image: Courtesy to Richard Gall.

- Generative adversarial nets (GAN), Goodfellow^{+}@NIPS'14

Increase domain confusion

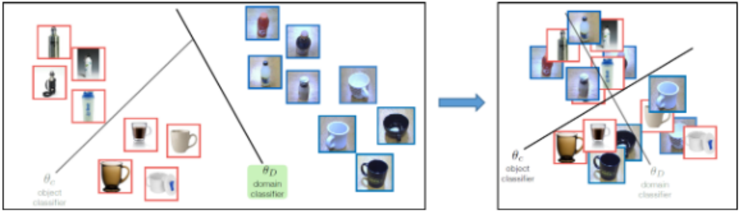


Image: Courtesy to Judy Hoffman.

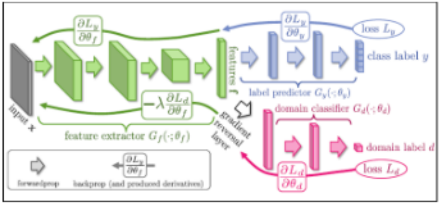
- Adversarial (GAN) loss, ADDA (Tzeng+@CVPR'17)

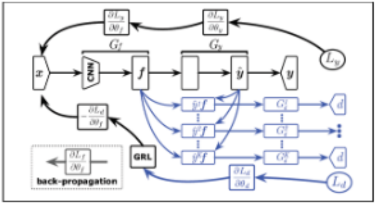
$$\max_{D} \{\mathbb{E}_{\mathbf{x}\sim p_S(\mathbf{x})}[\log D(M_S(\mathbf{x}))] + \mathbb{E}_{\mathbf{x}\sim p_T(\mathbf{x})}[\log(1 - D(M_T(\mathbf{x}))]\}$$

$$\max_{M_T} \{\mathbb{E}_{\mathbf{x}\sim p_T(\mathbf{x})}[\log D(M_T(\mathbf{x}))]\}$$

- Deep domain confusion, DDC (Tzeng+@ARXIV'14)
- Jensen-Shannon divergence (by GAN), GAM (Huang+@ECCV'18)

Gradient reversal layers

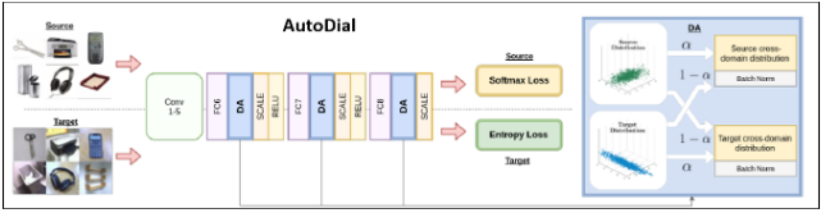




- RevGrad (Ganin$^+$@JMLR'16), MADA (Pei$^+$@AAAI'18), SimNet (Pinhero$^+$@CVPR'18)

$$\min_{M_S,M_T} \max_{D} V(D, M_S, M_T) = \mathbb{E}_{\mathbf{x}\sim p_S(\mathbf{x})}[\log D(M_S(\mathbf{x}))]\mathbb{E}_{\mathbf{x}\sim p_T(\mathbf{x})}[\log(1 - D(M_T(\mathbf{x})))]$$

Adapting the batch



- Domain specific batch normalization, AutoDial (Carlucci^{+}@ICCV'17), AdaBN (Li^{+}@PR'18), DSBN (Chang^{+}@CVPR'19)
- Batch Nuclear-norm Maximization, BNM (Cui^{+}@CVPR'20)
- Batch Whitening, DWT (Roy^{+}@CVPR'19)
- Learning batch re-weighting with mass shift, JD-BW (Binkowski^{+}@ICCV'19)

Transfer domain style

Paired I2I

Un-paired I2I

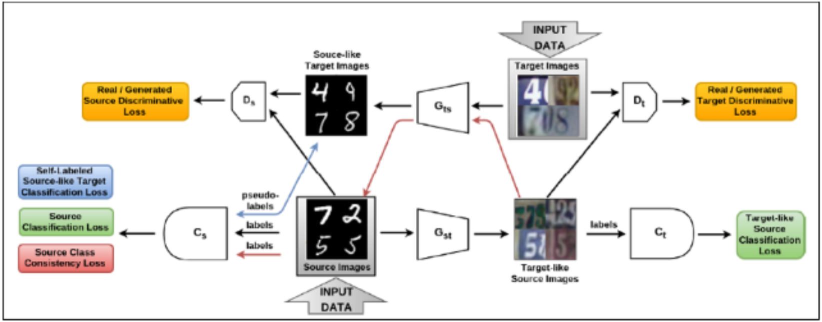
Paired image-to-image style transfer as preprocessing

- Csurka^{+}@TASKCV'17, Thomas^{+}@ACCV'19, Jackson^{+}@CVPR-WS'19

Unpaired image-to-image style transfer learning

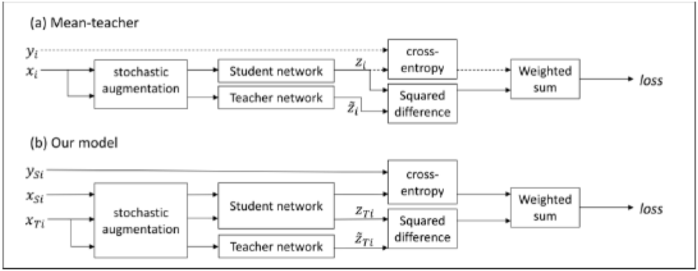
- I2I (Zhu^{+}@ICCV'17), I2IAd (Murez^{+}@CVPR'18)

Cyclic consistency



- Predict source from predicted target, LTR (Sener[+]@NIPS'16)
- Predict from traget-like source image, SBADA-GAN (Russo[+]@ARXIV'17)

Teacher-student paradigm



Mean-teacher of data augmented ensemble classifier

- SelfEns (French+@ICLR'18), DWT (Roy+@CVPR'19)

Refine student classifier's decision-boundary with a teacher

- DIRT-T (Shu+@ICLR'18)

Cluster alignment with a teacher

- CAT (Deng, +@ICCV'19)

To summarize

Winning strategies:

- Adversarial adaptation vs discriminative (CDAN, GAM)
- GAN (CyCADA, DRIT) better vs encoder-decoder
- Exploit score distributions to guide feature alignment (MCD, RWOT, DWT)
- Curriculum/Self-learning using pseudo-labels (PFAN, iCAN)

The results are to be taken cautiously as

- The results come from various papers
- Not clear how the hyperparameters for each model were selected
- Not always clear how comparable the models (*e.g.* diff architecture)

DeepDA becoming extremely popular in CV

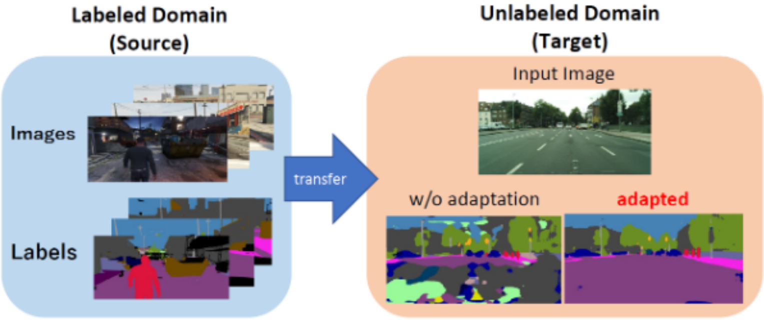
Many method proposed for:

- Semantic segmentation
- Person Re-ID
- Object detection

But recent DA methods were also proposed for:

- Pose/action recognition
- Depth estimation
- Low level image enhancement
- Control in robotics
- 3D/Visual localization
- Medical imaging

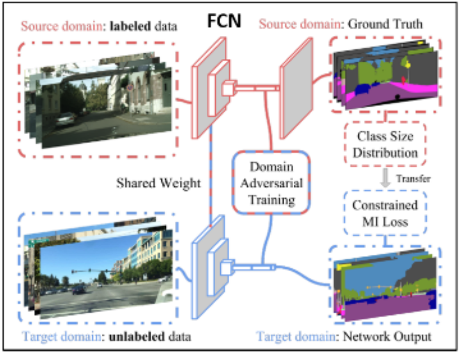
Image Segmentation

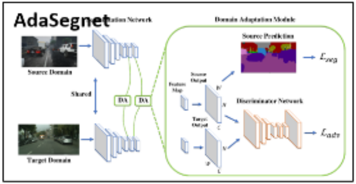


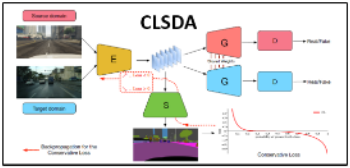
From Synthetic to real data

- Easy to obtain pixel level annotation
- Poor labeling due to domain shift

Segmentation model adaptation

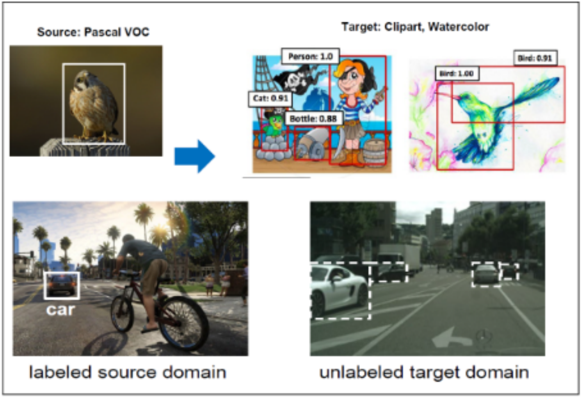






- Transferring label statistics, FCN-WLD (Hoffman^{+}@CoRR'16)
- Backpropagating contrastive loss, CLSDA (Zhu^{+}@ECCV'18)
- Multilevel Adversarial Learning, AdaSegNet (Tsai^{+}@CVPR'18)

Object detection



- Adapting Faster R-CNN, Chen+@CVPR'18, Zhu+@CVPR'19, Saito+@CVPR'19, Xu+@CVPR'20
- Self-training, RoyChowdhury+@CVPR'19, Inoue+@CVPR'18, Kim+@ICCV'19