Domain adaptation

Павлович Владислав

Минск, 2018

Domain adaptation

- Small [un]labeled dataset from target domain.
- Huge labeled dataset in a relevant domain.

Domain adaptation

(a) GTA

(b) Cityscapes

Domain adaptation

Model	Loss	Accuracy
GTA 64x64	0.5892	0.7049
GTA 128x128	1.1422	0.6808
GTA 256x256	1.0450	0.6714
GTA 512x512	1.4202	0.6742
Cityscapes 64x64	0.3859	0.8473
Cityscapes 128x128	0.4728	0.7930
Cityscapes 256x256	0.3717	0.8325
Cityscapes 512x512	0.4498	0.8011

Notation

- Feature space X.
- Probability distribution $P(X)$.
- Domain $D = \{X, P(X)\}$.
- Objective prediction function $f(x) = P(Y|X)$.

Notation

- $D^s = \{X, P_s(X)\}$.
- $D^t = \{X, P_t(X)\}$.
- $P_s(X) \to P_t(X)$.

Domain adaptation

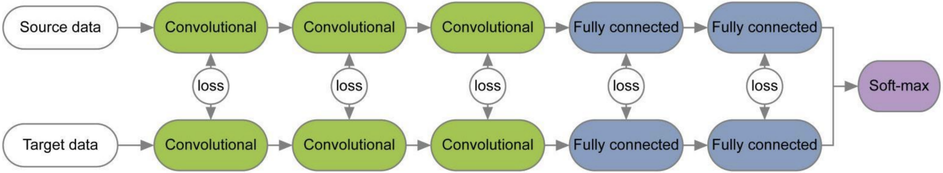
- Shallow
  - Fine-tuning
- Deep
  - SimGAN
  - CyCADA
  - GeoConGAN

Fine-tuning methods

- Class criterion.
- Architecture criterion.

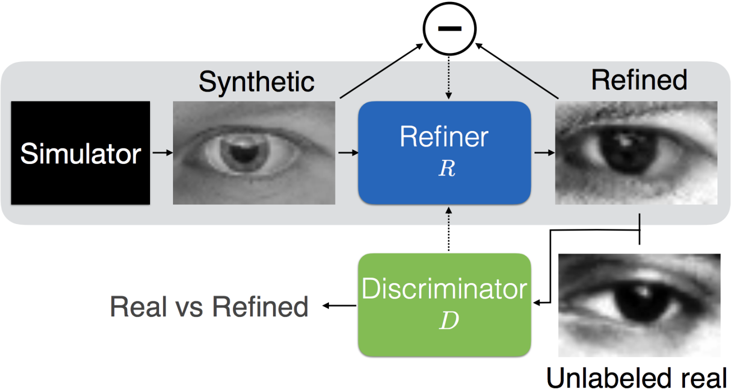
- Use pre-trained on ImageNet model as a universal feature extractor. and train a linear classifier (SVM) on extracted features.
- Fine-tune pre-trained model.

Architecture criterion



$$r_w(\theta_j^s; \theta_j^t) = \exp\left(||\theta_j^s - \theta_j^t||^2\right) - 1$$

SimGAN



Ashish S. et al. Learning from Simulated and Unsupervised Images through Adversarial Training, 2016

$$L_R(\theta) = \sum_i \ell_{real}(\theta; x_i, Y) + \lambda \ell_{reg}(\theta; x_i) =$$

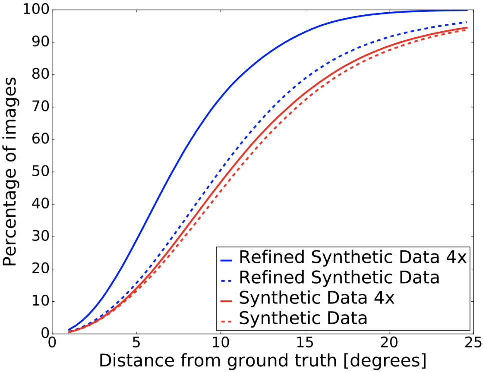
$$= -\sum_i \log(1 - D_\phi(R_\theta(x_i))) +$$

$$+ \lambda ||\psi(R_\theta(x_i)) - \psi(x_i)||$$

ψ - identity or any other transformation.

Ashish S. et al. Learning from Simulated and Unsupervised Images through Adversarial Training, 2016

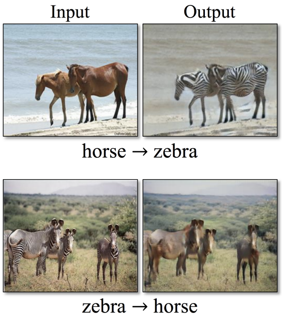
SimGAN



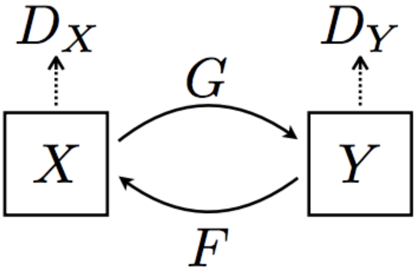
Ashish S. et al. Learning from Simulated and Unsupervised Images through Adversarial Training, 2016

Method	R/S	Error
Support Vector Regression (SVR) [33]	R	16.5
Adaptive Linear Regression ALR) [23]	R	16.4
Random Forest (RF) [36]	R	15.4
kNN with UT Multiview [47]	R	16.2
CNN with UT Multiview [47]	R	13.9
k-NN with UnityEyes [43]	S	9.9
CNN with UnityEyes Synthetic Images	S	11.2
CNN with UnityEyes Refined Images	S	**7.8**

Ashish S. et al. Learning from Simulated and Unsupervised Images through Adversarial Training, 2016

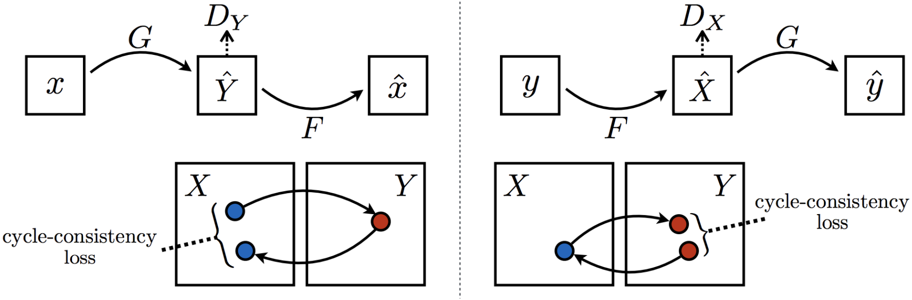


Jun-Yan Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, 2017



Jun-Yan Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, 2017

CycleGAN



Jun-Yan Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, 2017

CycleGAN

$$
\begin{aligned}
L(G, F, D_X, D_Y) &= L_{GAN}(G, D_Y, X, Y) + \\
&+ L_{GAN}(F, D_X, Y, X) + \\
&+ \lambda L_{cyc}(G, F) \\
L_{GAN}(G, D_Y, X, Y) &= E_{y \sim p_{data}(Y)}[\log D_Y(y)] + \\
&+ E_{x \sim p_{data}(X)}[\log(1 - D_Y(G(x))] \\
L_{cyc}(G, F) &= E_{x \sim p_{data}(X)}[||F(G(x)) - x||_1] + \\
&+ E_{y \sim p_{data}(Y)}[||G(F(y)) - y||_1]
\end{aligned}
$$

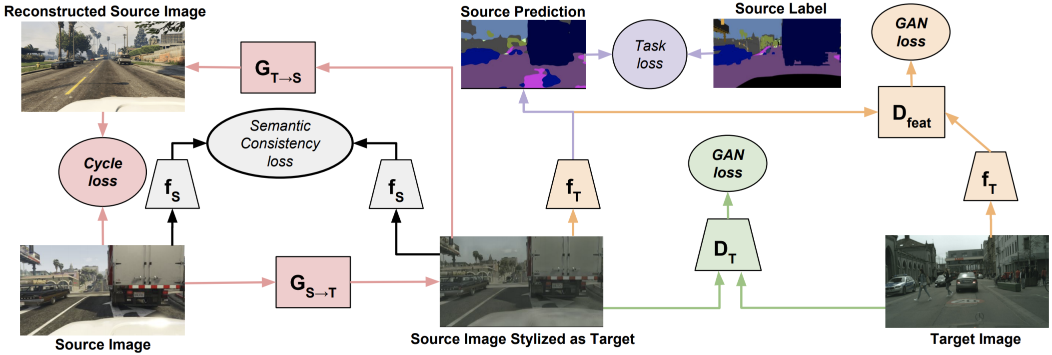
Jun-Yan Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent

CycleGAN

(c) Source

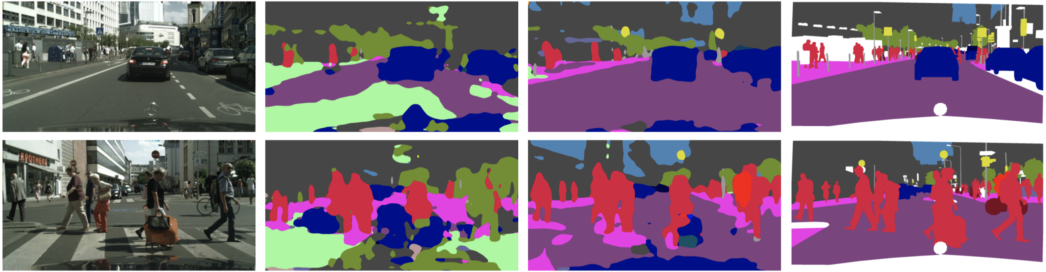
(d) Generated

CyCADA



Judy Hoffman et al. CYCADA: cycle-consistent adversarial domain adaptation, 2017

CyCADA



(a) Test Image (b) Source Prediction (c) CyCADA Prediction (d) Ground Truth

Judy Hoffman et al. CYCADA: cycle-consistent adversarial domain adaptation, 2017

CyCADA

$$L_{task}(f_S, X_S, Y_S) = -E_{(x_s,y_s)\sim(X_S,Y_S)} \sum_{k=1}^{K} \log(\sigma(f_S^{(k)}(x_s)))$$

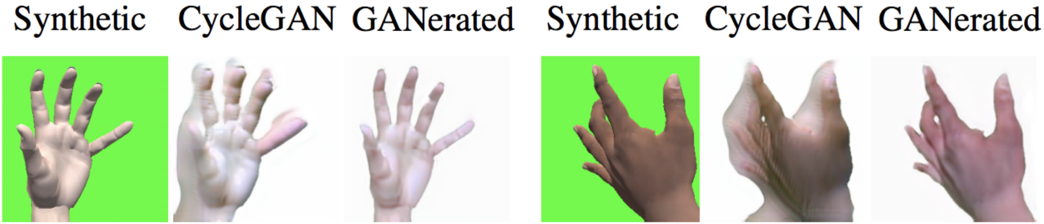
$$\begin{aligned} L_{sem}(G_{S\to T}, G_{T\to S}, X_S, X_T, f_S) =& L_{task}(f_S, G_{T\to S}(X_T), p(f_S, X_T)) + \\ &+ L_{task}(f_S, G_{S\to T}(X_S), p(f_S, X_S)) \end{aligned}$$

Judy Hoffman et al. CYCADA: cycle-consistent adversarial domain adaptation, 2017

CyCADA

Method	mIOU
Source only	21.7
CyCADA feat-only	31.7
CyCADA pixel-only	37
CyCADA pixel+feat	39.5

GeoConGAN



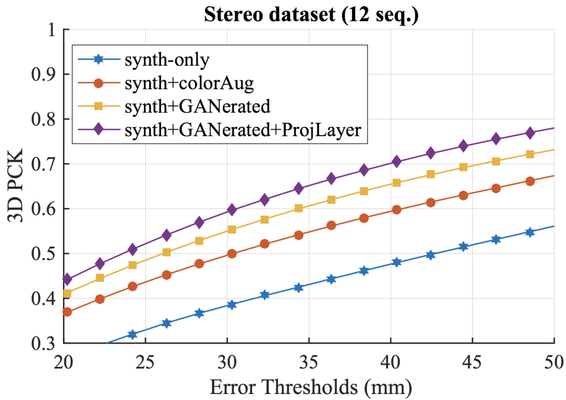
Franziska Mueller et al. GANerated Hands for Real-Time 3D Hand Tracking from Monocular RGB, 2018

GeoConGAN

- Based on CycleGAN.
- Extra cross-entropy loss for segmented images to preserve annotations.

Franziska Mueller et al. GANerated Hands for Real-Time 3D Hand Tracking from Monocular RGB, 2018

GeoConGAN



Franziska Mueller et al. GANerated Hands for Real-Time 3D Hand Tracking from Monocular RGB, 2018

Thank You!