Università di Genova | MaLGa

Domain adaptation

26/03/2025

Vito Paolo Pastore

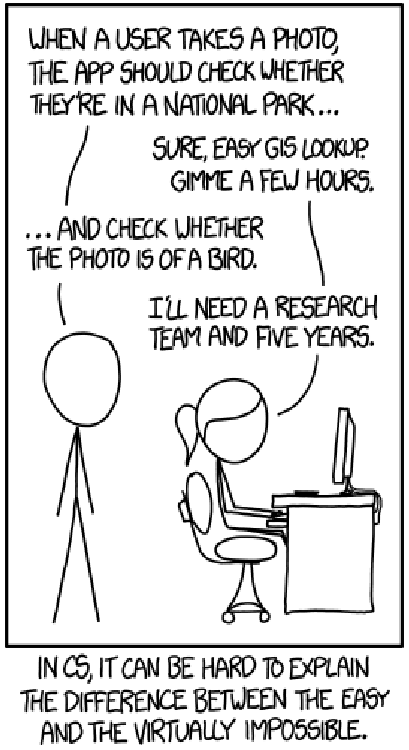
Deep learning a.y. 2024/2025

Domain Adaptation

Credits

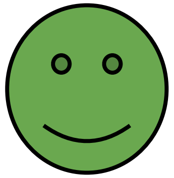
- Tutorial by Pietro Morerio and Massimiliano Mancini
- Some slides courtesy of Prof. Elisa Ricci and Vittorio Murino

Is there a bird?



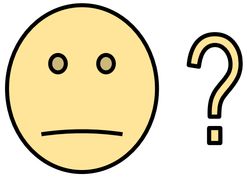
https://xkcd.com/1425/

Is there a bird?



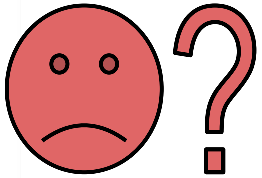
Is there a bird?

Is there a bird?



https://bam-dataset.org/

Is there a bird?



https://bam-dataset.org/

A Real-World Example

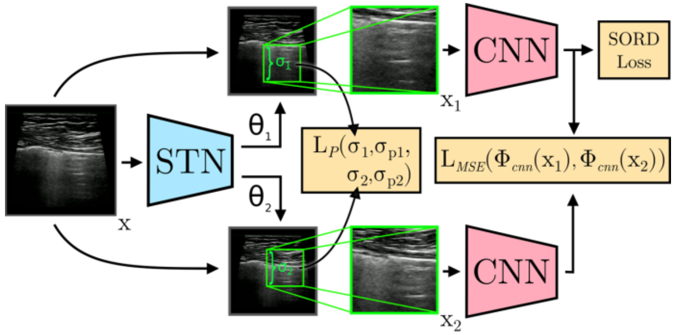
Prediction of COVID-19 markers in lung ultrasonography images.

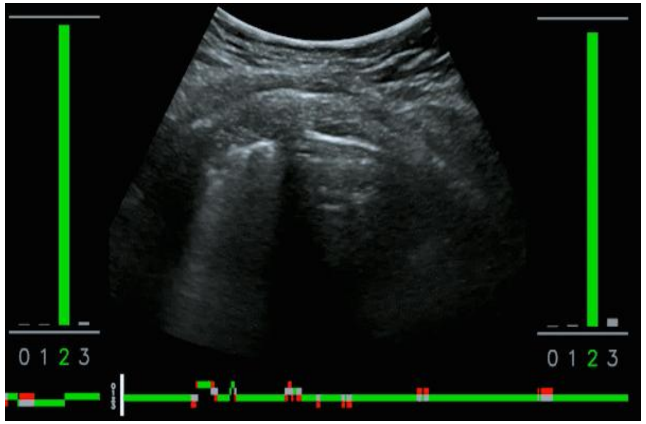
Data from different hospitals (and different sensors) in Italy

Train/test data on same hospital ≥85% accuracy.

Performance drops of even 20% considering data from different hospitals.

Issues in having annotated data from all the hospitals.





S. Roy et al., *Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound*, IEEE Trans. on Medical Imaging 2020.

Domain Shift

- A domain shift is a change in the statistical distribution of training and test data.

- *Generally* affects the performance of a model.

- ***In the vision context, it consists in visual appearance changes degrade the performance of visual recognition systems.***

Accuracy: 85%

B. Kulis, K. Saenko, and T. Darrell, "What You Saw is Not What You Get: Domain Adaptation Using Asymmetric Kernel Transforms" CVPR, 2011.

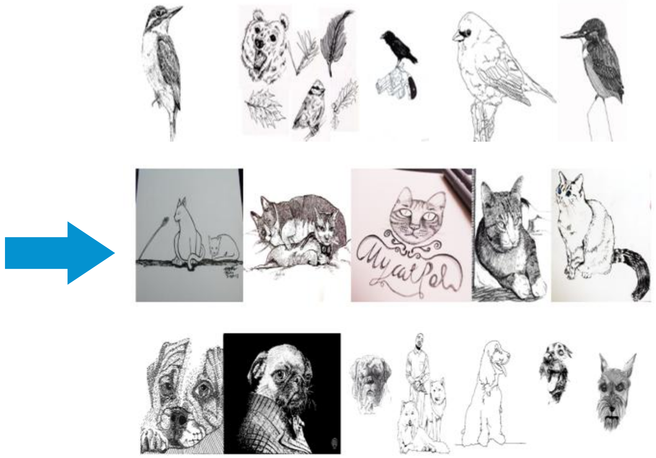
Domain Shift

- A domain shift is a change in the statistical distribution of training and test data.

- *Generally* affects the performance of a model.

- ***In the vision context, it consists in visual appearance changes degrade the performance of visual recognition systems.***

Accuracy: 85%



Accuracy: 55%

B. Kulis, K. Saenko, and T. Darrell, "What You Saw is Not What You Get: Domain Adaptation Using Asymmetric Kernel Transforms" CVPR, 2011.

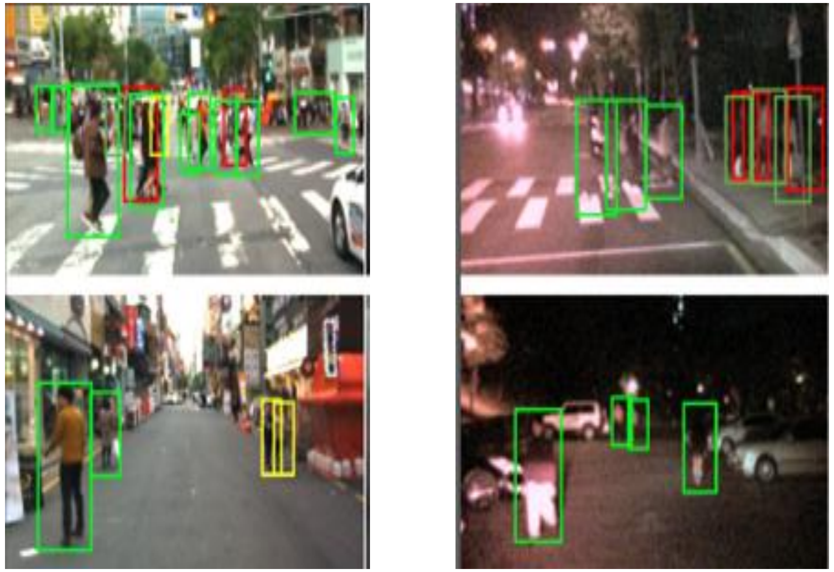
Domain Shift: why do we care about?

Domain Shift: why do we care about?

Z. Chen et al., "Deep Learning Features at Scale for Visual Place Recognition". ICRA 2017.

Appearance changes due to seasonal and time changes.

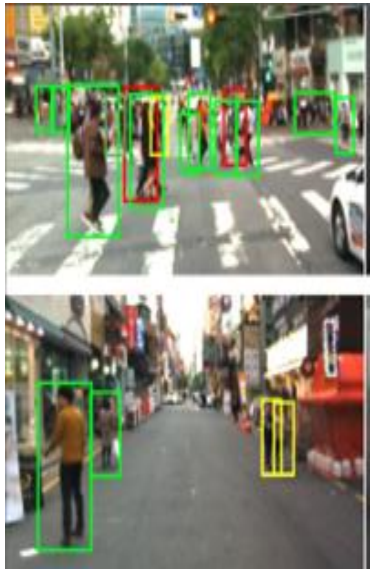
Domain Shift: why do we care about?

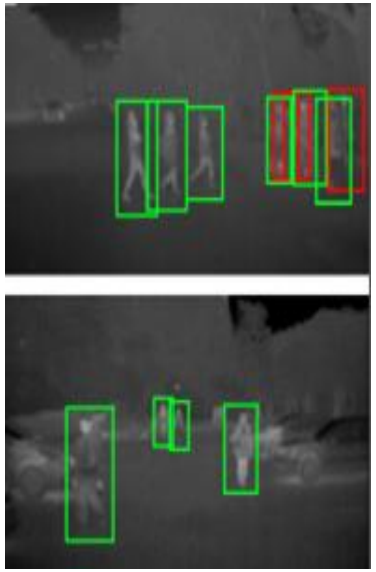


S. Hwang et al., “Multispectral Pedestrian Detection: Benchmark Dataset and Baseline”, CVPR, 2015.

Appearance changes due to illumination conditions

Domain Shift: why do we care about?





S. Hwang et al., “Multispectral Pedestrian Detection: Benchmark Dataset and Baseline”, CVPR, 2015.

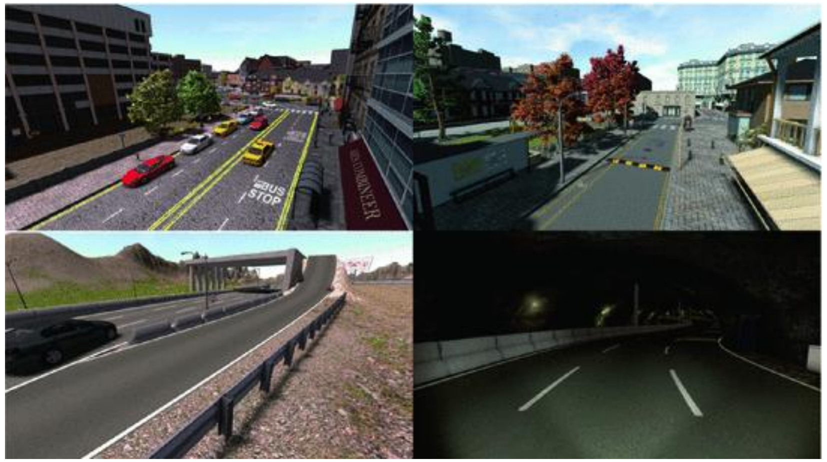
Appearance changes due to different sensors.

Domain Shift: why do we care about?

https://www.zdnet.com/article/robots-to-the-rescue-searching-for-survivors-checking-on-structural-damage-in-japan/

Overcoming costly/unfeasible data collection

Domain Shift: why do we care about?

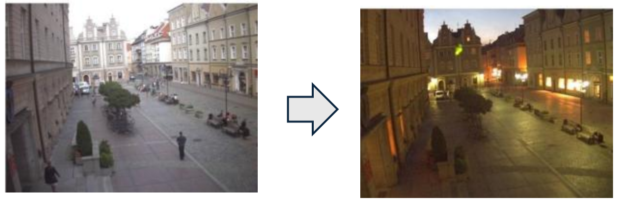


http://synthia-dataset.net/

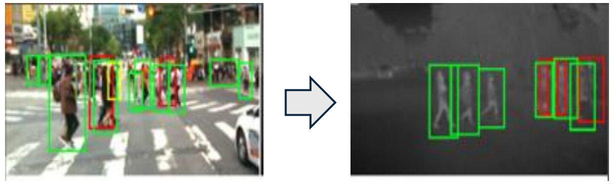
Use of synthetic data

Domain Shift is *ubiquitous* ...

Time & Environmental Changes



Different Modalities



Synthetic to Real Images

CAD to Real Images

Domain shift and dataset bias

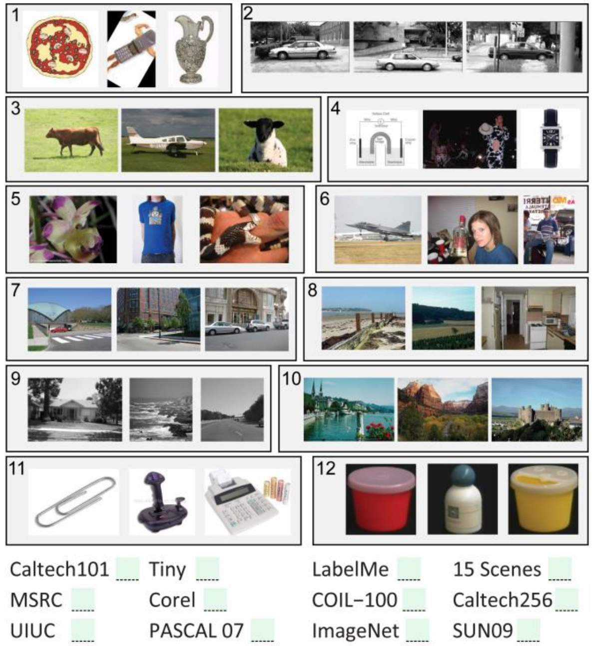
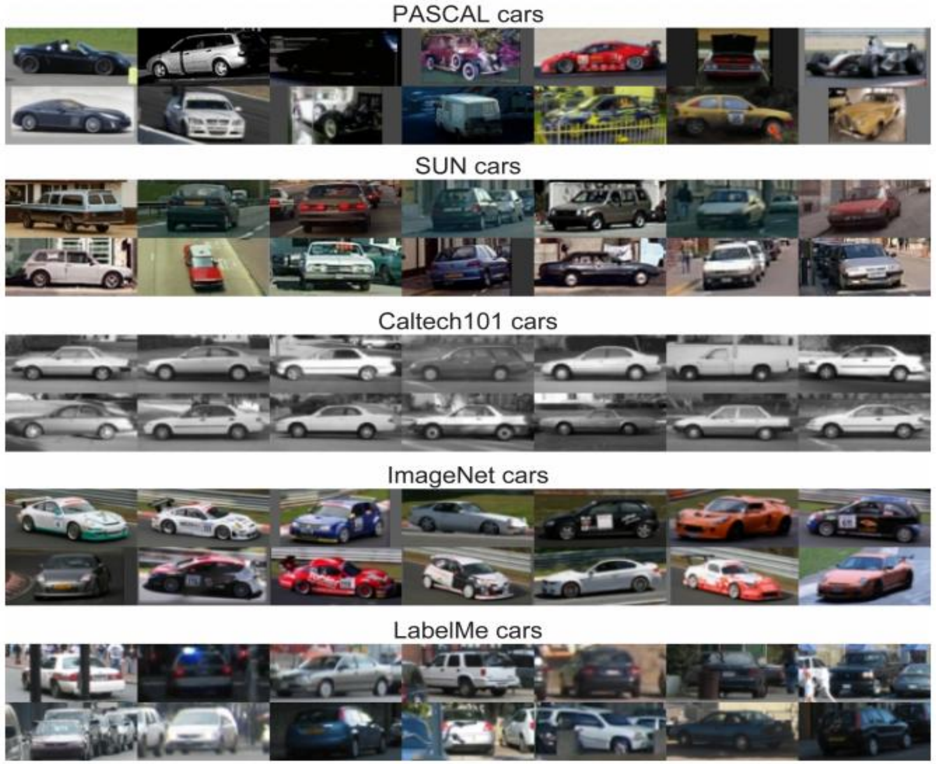
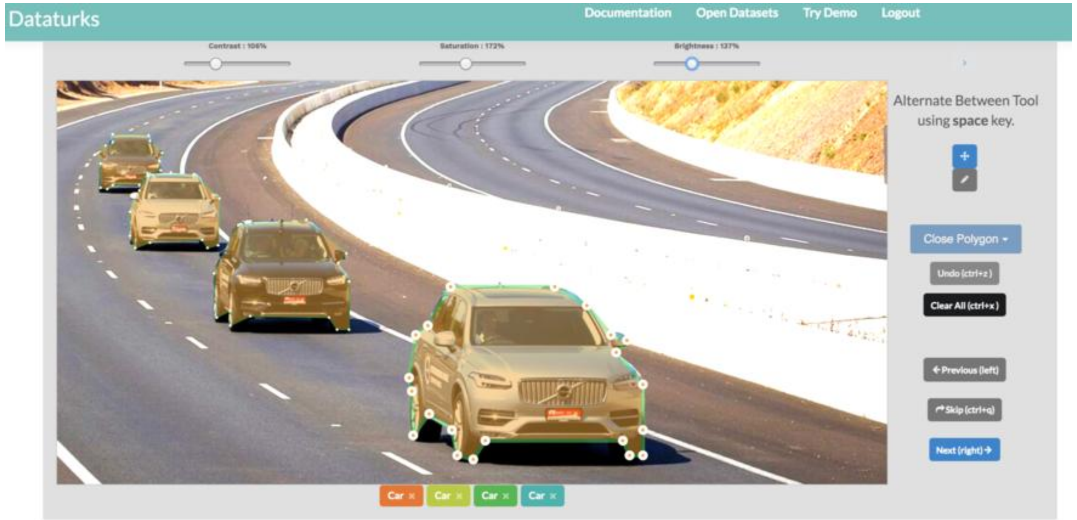


Figure 1. Name That Dataset: Given three images from twelve popular object recognition datasets, can you match the images with the dataset? (answer key below)



Torralba and Efros .,"Unbiased look at dataset bias, CVPR, 2011.

Domain Shift: how do we solve it?

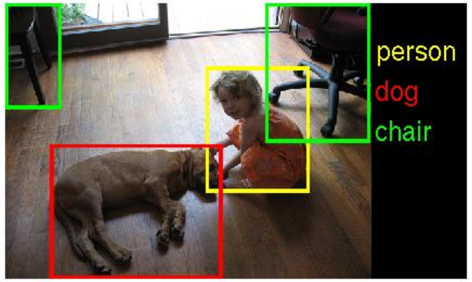


Source image

Annotate target data....

UniGe | MaLGa

Domain Shift & Data Annotation

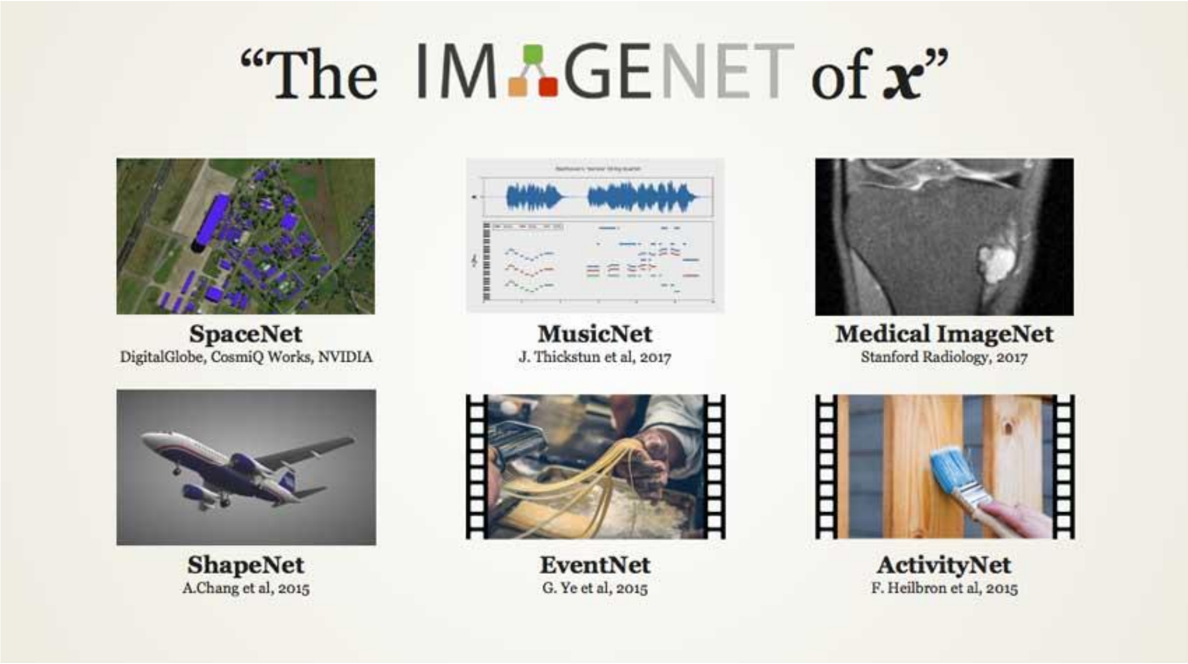


- about hand-annotated 14M images
- about 1M with bounding-boxes
- more than 21,000 categories

J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. CVPR) 2009.

Data collection is costly

Domain Shift & Data Annotation



Slide credit: [Fei-Fei Li]

We cannot annotate everything.

Domain Shift & Data Annotation

https://www.zdnet.com/article/robots-to-the-rescue-searching-for-survivors-checking-on-structural-damage-in-japan/

Sometimes collecting data is impossible

A bit of notation

$\mathcal{X}$	Input space
$\mathcal{Y}$	Output space
$X \in \mathcal{X}$	Input variable (image)
$Y \in \mathcal{Y}$	Output variable (label)
$D = \{\mathcal{X}, P(X)\}$	Domain
$T = \{\mathcal{Y}, P(Y\|X)\}$	Task

F. Orabona & T. Tommasi. Tutorial on Domain Adaptation and Transfer Learning. *ECCV 2014*.

The Domain Adaptation (DA) Problem

Source Domain

$$D^s = \{\mathcal{X}^s, P(X^s)\}$$
$$T^s = \{\mathcal{Y}^s, P(Y^s|X^s)\}$$

Target Domain

$$D^t = \{\mathcal{X}^t, P(X^t)\}$$
$$T^t = \{\mathcal{Y}^t, P(Y^t|X^t)\}$$

DA problem

$$D^s \neq D^t \qquad T^s = T^t$$

F. Orabona & T. Tommasi. Tutorial on Domain Adaptation and Transfer Learning. In ECCV 2014.

The Domain Adaptation (DA) Problem

Source Domain

$$D^s = \{\mathcal{X}^s, P(X^s)\}$$
$$T^s = \{\mathcal{Y}^s, P(Y^s|X^s)\}$$

Target Domain

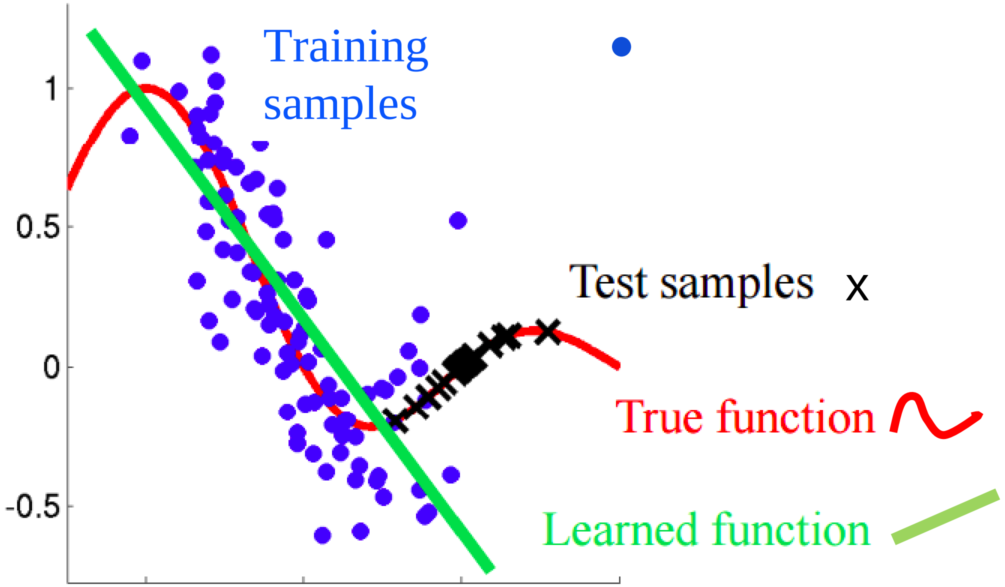
$$D^t = \{\mathcal{X}^t, P(X^t)\}$$
$$T^t = \{\mathcal{Y}^t, P(Y^t|X^t)\}$$

DA problem

$$\mathcal{X}^s \neq \mathcal{X}^t$$
$$P(X^s) \neq P(X^t)$$

$$\mathcal{Y}^s = \mathcal{Y}^t$$

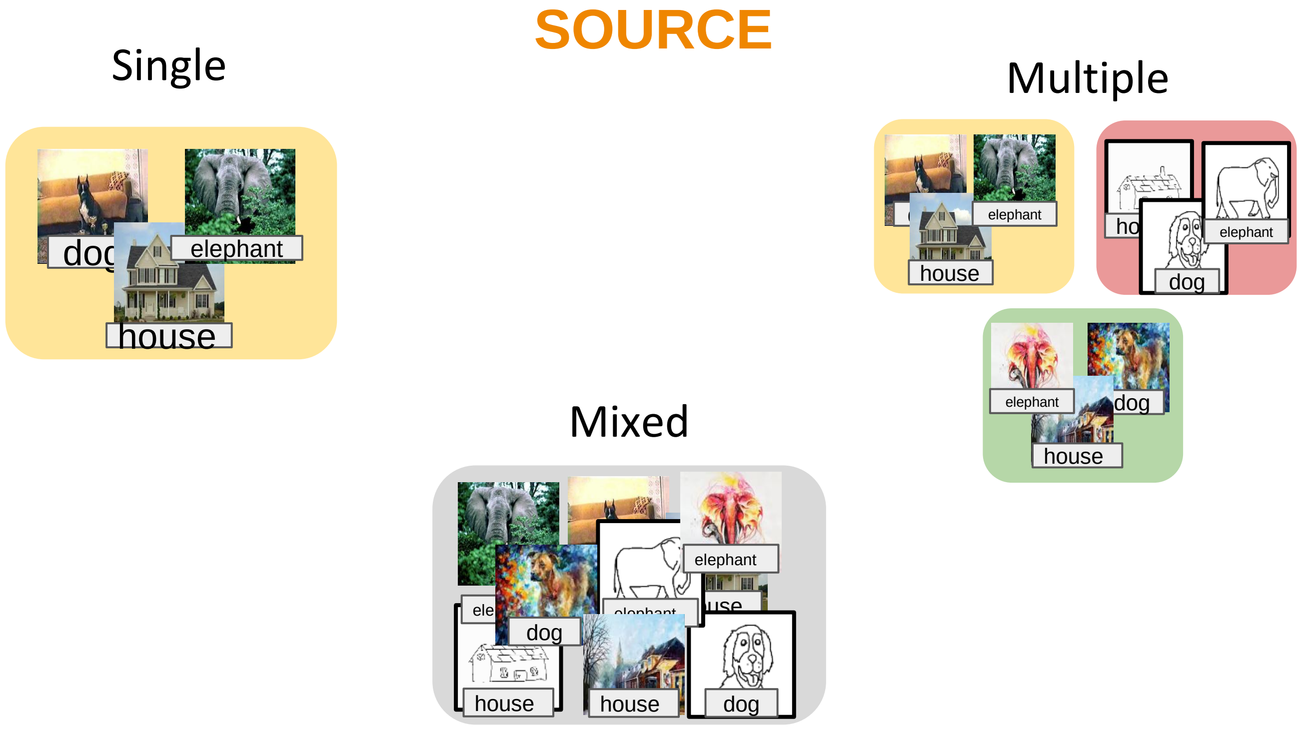
Covariate Shift Assumption



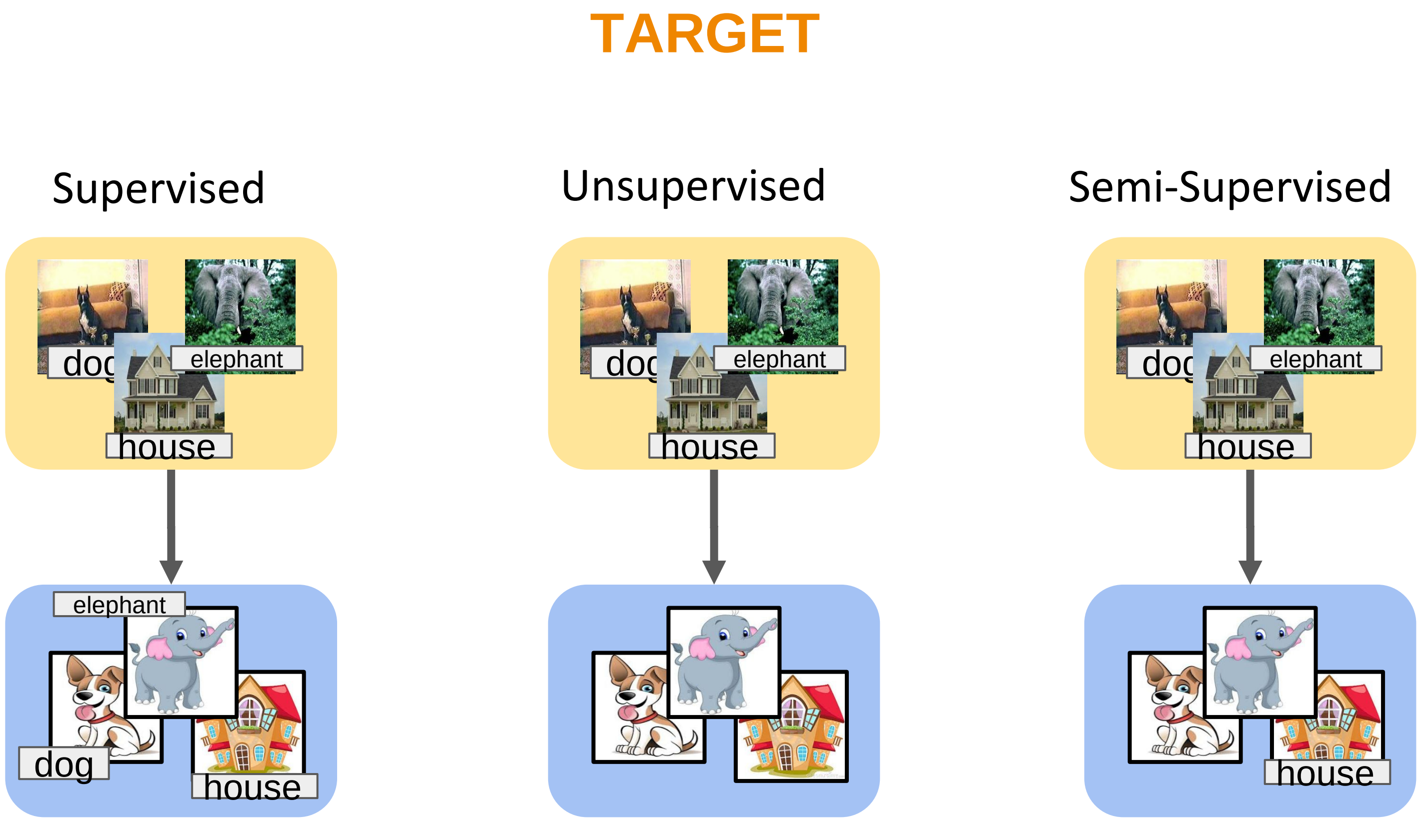
UniGe | MaLGa

DA Scenarios

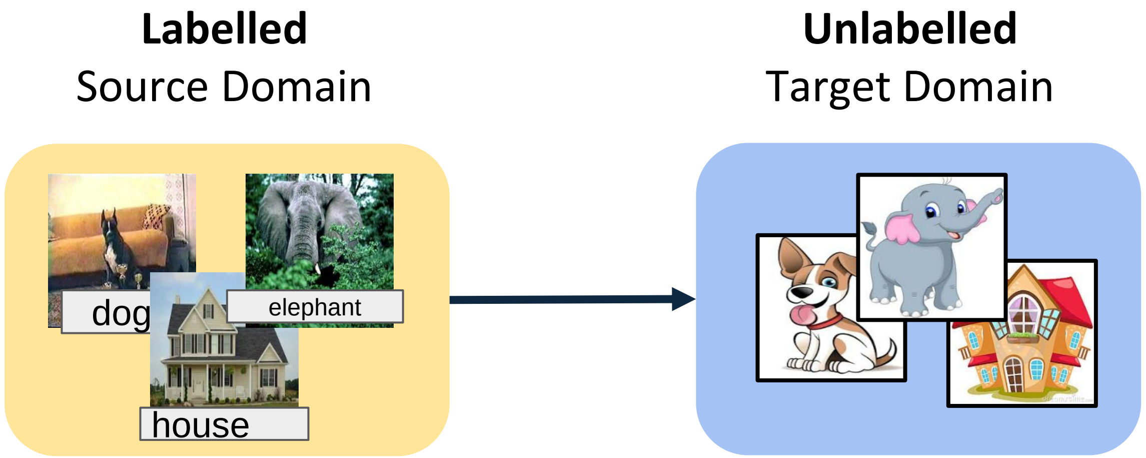
DA Scenarios: number of (source) domains



DA Scenarios: the target labels



DA Scenarios: unsupervised DA



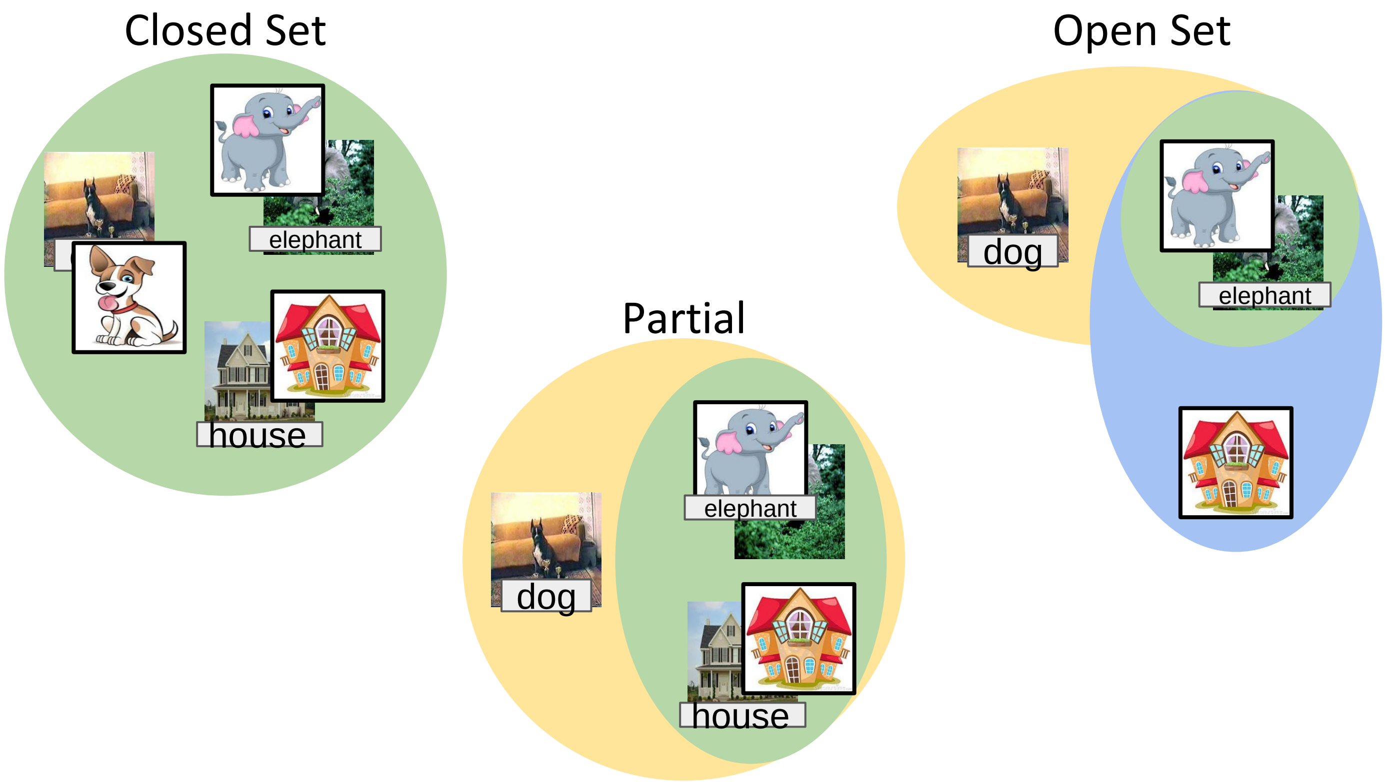
$$P(X^s) \neq P(X^t)$$
$$\mathcal{Y}^s = \mathcal{Y}^t$$

Real images vs Cartoons

{elephant, dog, house}

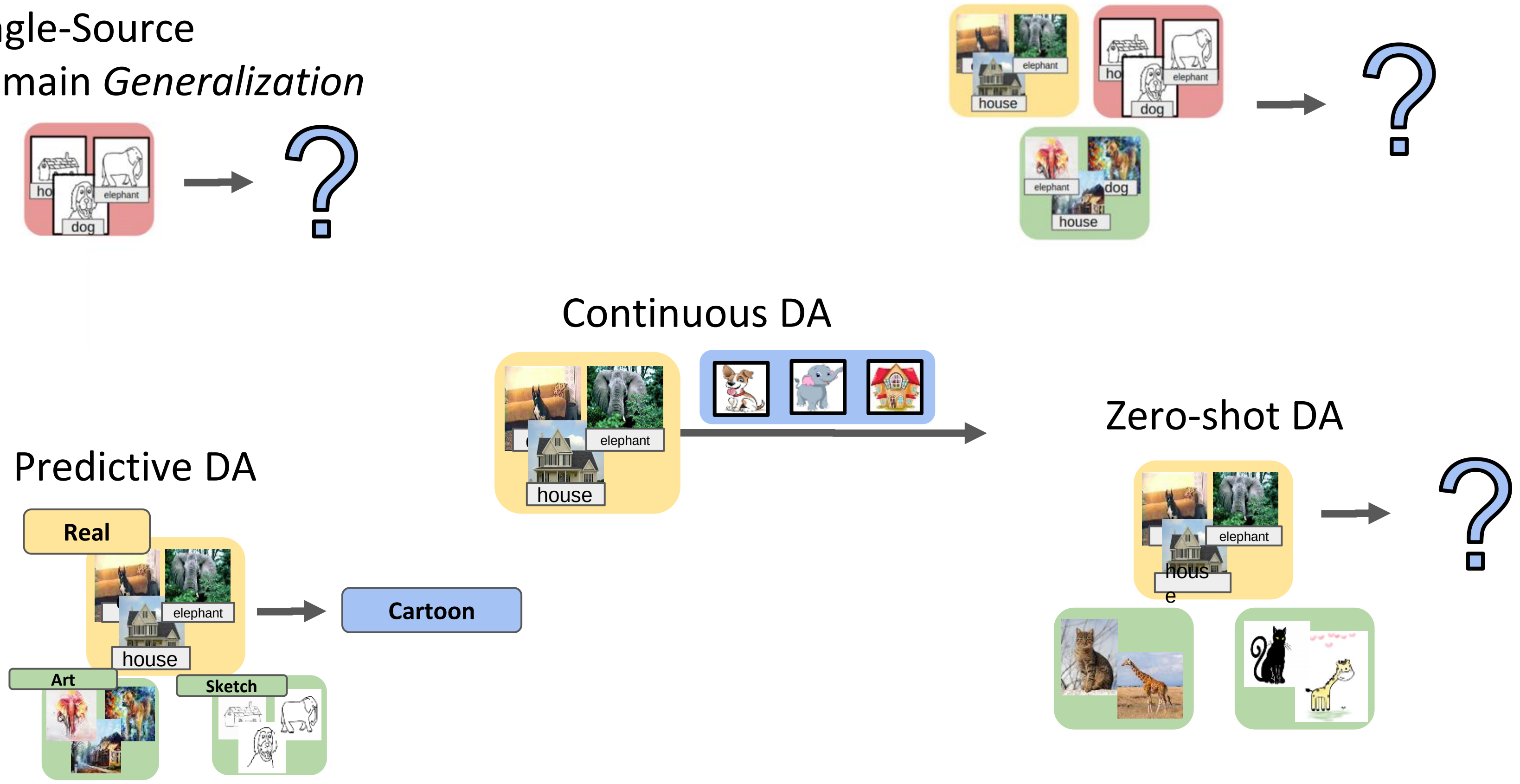
http://www.eecs.qmul.ac.uk/~dl307/project_iccv2017

DA Scenarios: the label space



DA Scenarios: without target data!

(Multi-Source)
Domain *Generalization*



- **Predictive DA:** scenario where no target data is available and the system has to learn to generalize from annotated source images plus unlabeled samples with associated metadata from auxiliary domains

Mancini, Massimiliano, et al. "Adagraph: Unifying predictive and continuous domain adaptation through graphs." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.

Deep Domain Adaptation

Three main categories:

- **<u>Discrepancy-based</u>**: fine-tuning the deep network with labeled or unlabeled target data to diminish the domain shift
- **<u>Adversarial-based</u>**: using domain discriminators to encourage domain confusion through an adversarial objective
- **<u>Reconstruction-based</u>**: using the data reconstruction as an auxiliary task to ensure feature invariance

G. Csurka, "Domain adaptation for visual applications: A comprehensive survey". arXiv preprint arXiv:1702.05374, 2017

Feature distribution discrepancy-based methods

Setting and Introduction (shared across next methods)

$$\mathcal{D}_s = \{(\mathbf{x}_i^s, y_i^s)\}_{i=1}^{n_s} \qquad \mathcal{D}_t = \{\mathbf{x}_j^t\}_{j=1}^{n_t}$$

- Deep neural network able to learn transferable features that bridge the cross-domain discrepancy, and build a classifier $y = \theta(\mathbf{x})$ which can minimize target risk $\epsilon_t(\theta) = \Pr_{(\mathbf{x},y)\sim q}[\theta(\mathbf{x}) \neq y]$ using source supervision

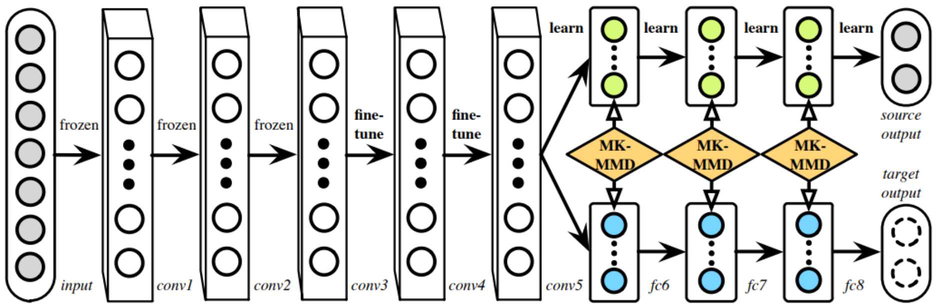
- To approach this problem, many existing methods aim to bound the target error by the source error plus a discrepancy metric between the source and the target, e.g., maximum mean discrepancies (MMD)

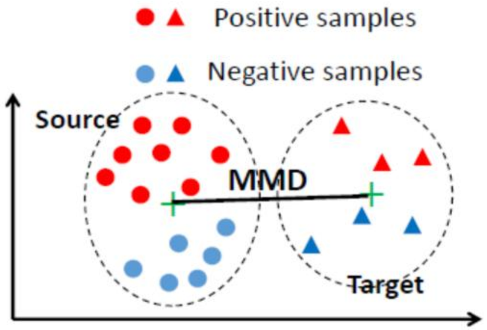
- MMD was introduced in Gretton et al., to decide whether two observations X and Y independently and identically distributed from two probability distributions p and q, decide whether p and q are the same or not.

F. Orabona & T. Tommasi. . In ECCV 201Gretton, A. et al. . “A kernel method for the two-sample-problem”. In NIPS. 2007.
4.

Domain adaptation network (1)

Long, M., Cao, Y., Wang, J., & Jordan, M. (2015, June). Learning transferable features with deep adaptation networks. In *International conference on machine learning* (pp. 97-105). PMLR.





Mean embedding of distribution p

$$d_k^2(p,q) \triangleq \left\| \mathbf{E}_p[\phi(\mathbf{x}^s)] - \mathbf{E}_q[\phi(\mathbf{x}^t)] \right\|_{\mathcal{H}_k}^2$$

$$\mathbf{E}_{\mathbf{x}\sim p} f(\mathbf{x}) = \langle f(\mathbf{x}), \mu_k(p) \rangle_{\mathcal{H}_k}$$

Feature map associated to kernel:

$$k(\mathbf{x}^s, \mathbf{x}^t) = \langle \phi(\mathbf{x}^s), \phi(\mathbf{x}^t) \rangle$$

Convex combination of L positive defined kernel

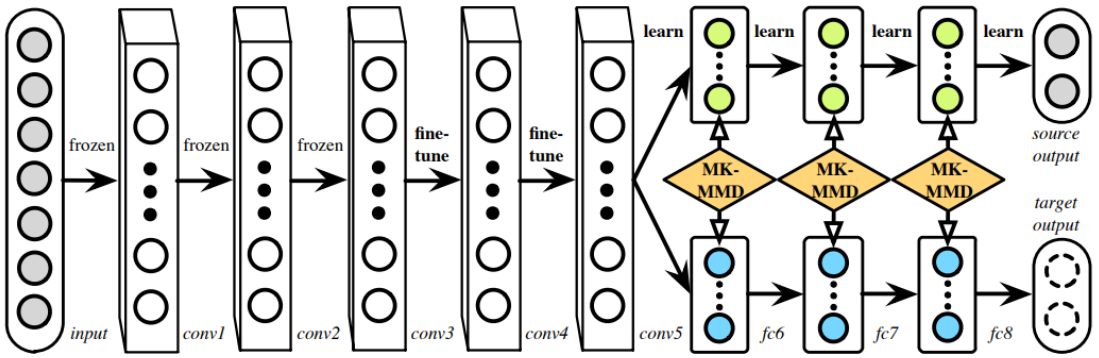
$$k(\mathbf{x}^s, \mathbf{x}^t) = \sum_{l=1}^{L} \beta_l k_l(\mathbf{x}^s, \mathbf{x}^t), \text{s.t.} \beta_l \geq 0, \sum_{l=1}^{L} \beta_l = 1.$$

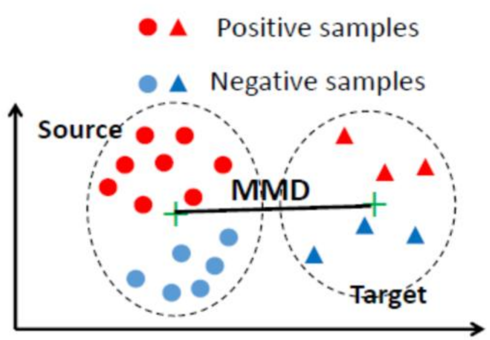
- Distance between embeddings of the source and target probability distributions in a **reproducing kernel Hilbert space**.

- In a deep network, bottom layers capture general patterns of the data, whereas top ones are more tailored to the classification task *and thus biased towards the source/target domain*.

- **Idea**: perform a re-mapping of the feature representation in a common embedding so that such bias is removed. Such re-mapping is made at different layers and is induced by a MMD distance over a kernel function (introduced the Multiple Kernel MMD OR MK-MMD)

Domain adaptation network (2)





- The CNN is fine-tuned on the source labeled examples requiring that source and target to become similar under the hidden representations of fully connected layers fc6-fc8.

- First, the CNN (AlexNet in the paper) **is trained on the source dataset.** Then, the adaptation consists in fine-tuning only layers from conv4 included (the more specifically tailored to the source dataset)

- Now, in the fine-tuning, we use unannotated target data, and we modify the objective function to include a variant of the Maximum Mean Discrepancy

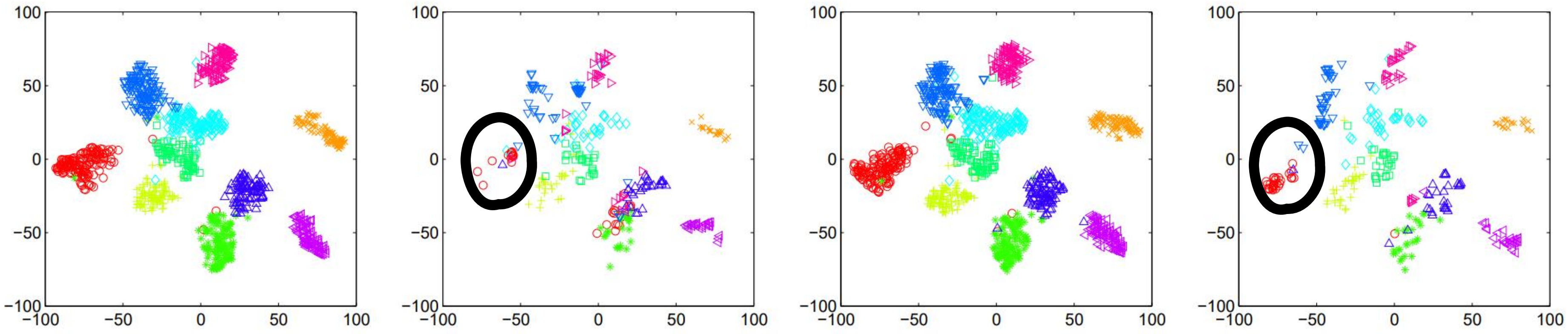
$$\min_{\Theta} \frac{1}{n_a} \sum_{i=1}^{n_a} J\left(\theta\left(\mathbf{x}_i^a\right), y_i^a\right) + \lambda \sum_{\ell=l_1}^{l_2} d_k^2\left(\mathcal{D}_s^\ell, \mathcal{D}_t^\ell\right)$$

layer

Domain adaptation network (3)

Office 31. 4,652 images within 31 categories collected from three distinct domains: Amazon (A), which contains images downloaded from amazon.com, Webcam (W) and DSLR (SLR camera from office)

Target features: More discriminative and more aligned with the source domain



(a) DDC Features on Source (b) DDC Features on Target (c) DAN Features on Source (d) DAN Features on Target

Long, M., Cao, Y., Wang, J., & Jordan, M. (2015, June). Learning transferable features with deep adaptation networks. In *International conference on machine learning* (pp. 97-105). PMLR.

Weighted Domain adaptation network (WDAN) (1)

The method aims to solve the issue of class prior distributions change between source and target domains, and the correspondent degrading performances.

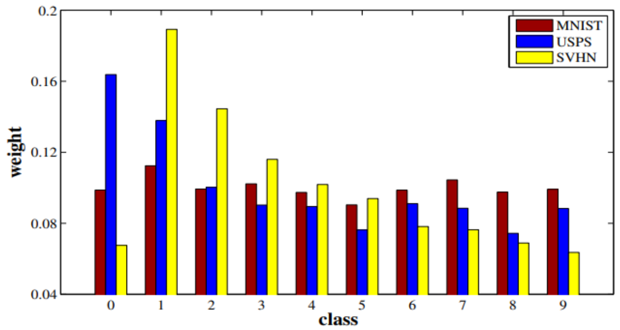


Figure 1. Class prior distributions of three domains for digit recognition. As is shown, class bias exists across domains. It is natural to see that the class weight of 0 and 1 are relatively high in postal service (*USPS*), and the class weight of 1 and 2 are relatively high in house numbers (*SVHN*).

The authors introduce an auxiliary weight for each class, obtaining a weighted discrepancy metrics that is used for adaptation

Yan, H., Ding, Y., Li, P., Wang, Q., Xu, Y., & Zuo, W. (2017). Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2272-2281).

Weighted Domain adaptation network (WDAN) (2)

MMD is modified in this way:

$$\mathrm{MMD}_w^2(\mathcal{D}_s, \mathcal{D}_t) = \left\| \frac{1}{\sum_{i=1}^{M} \alpha_{y_i^s}} \sum_{i=1}^{M} \alpha_{y_i^s} \phi(\mathbf{x}_i^s) - \frac{1}{N} \sum_{j=1}^{N} \phi(\mathbf{x}_j^t) \right\|_{\mathcal{H}}^2 \quad (8)$$

Source class weight bias

- A CNN is first trained on source dataset, then the last layers are fine-tuned using the following objective function

$$\min_{\mathbf{W}, \{\hat{y}_j\}_{j=1}^{N}, \boldsymbol{\alpha}} \frac{1}{M} \sum_{i=1}^{M} \ell(\mathbf{x}_i^s, y_i^s; \mathbf{W}) + \gamma \frac{1}{N} \sum_{j=1}^{N} \ell(\mathbf{x}_j^t, \hat{y}_j^t; \mathbf{W})$$

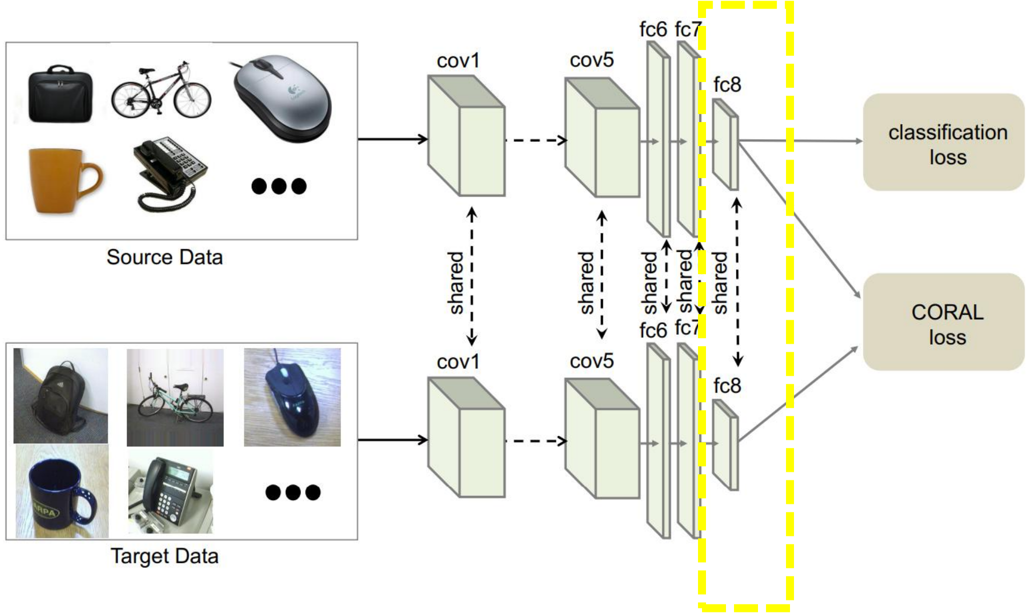
$$+ \lambda \sum_{l=l_1}^{l_2} \mathrm{MMD}_{l,w}(\mathcal{D}_s^l, \mathcal{D}_t^l), \quad (11)$$

Pseudo-labeling

Weighted MMD

Yan, H., Ding, Y., Li, P., Wang, Q., Xu, Y., & Zuo, W. (2017). Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2272-2281).

Deep DA – CORrelation ALignment, CORAL



$$\ell_{CORAL} = \frac{1}{4d^2}\|C_S - C_T\|_F^2$$

$$C_S = \frac{1}{n_S - 1}(D_S^\top D_S - \frac{1}{n_S}(\mathbf{1}^\top D_S)^\top(\mathbf{1}^\top D_S))$$

$$C_T = \frac{1}{n_T - 1}(D_T^\top D_T - \frac{1}{n_T}(\mathbf{1}^\top D_T)^\top(\mathbf{1}^\top D_T))$$

Covariances

$\|\cdot\|_F^2$ Squared matrix frobenius norm

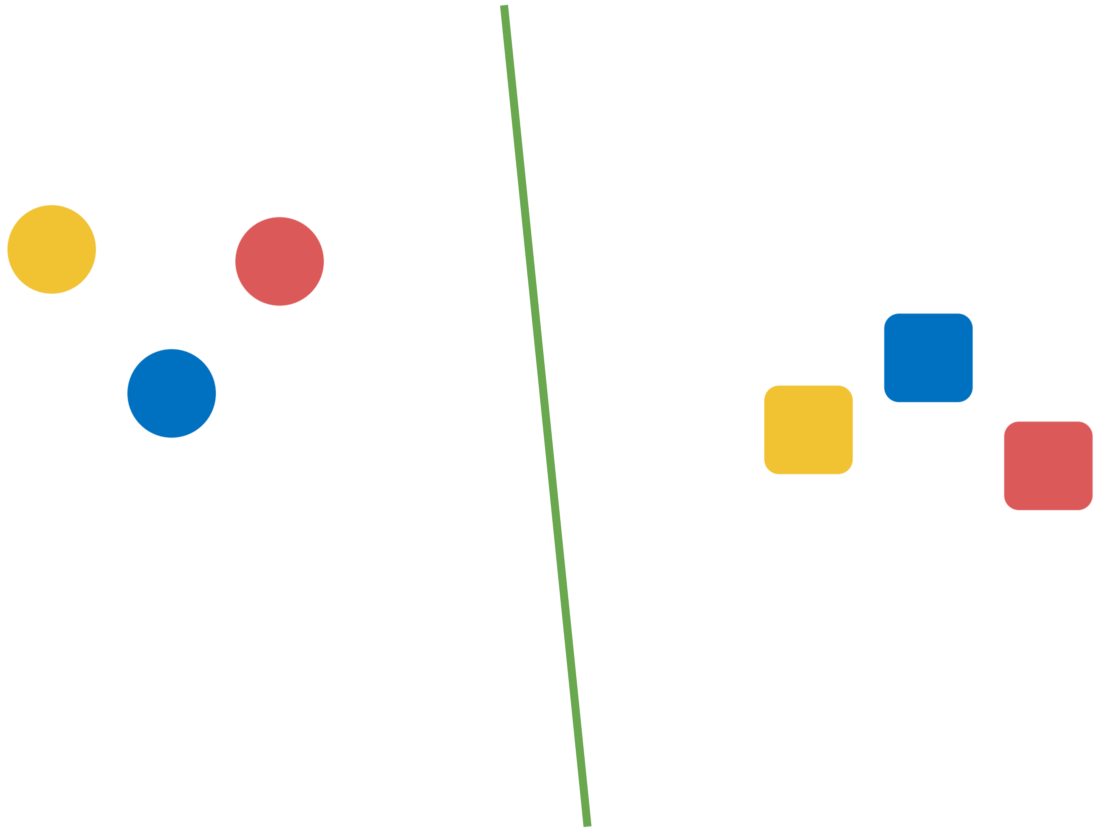
Yan, H., Ding, Y., Li, P., Wang, Q., Xu, Y., & Zuo, W. (2017). Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2272-2281).

Adversarial-based methods

Domain Classification

Domain Classification

Domain Classification

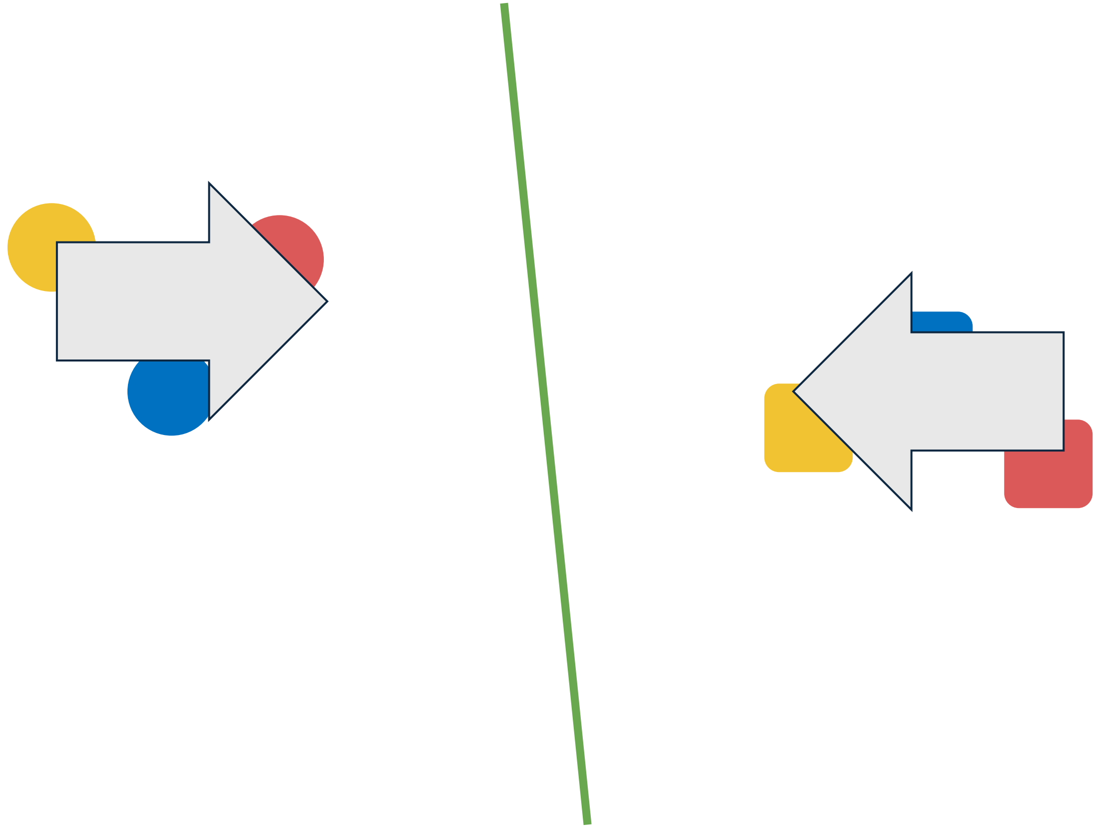


Adversarial approaches to domain adaptation

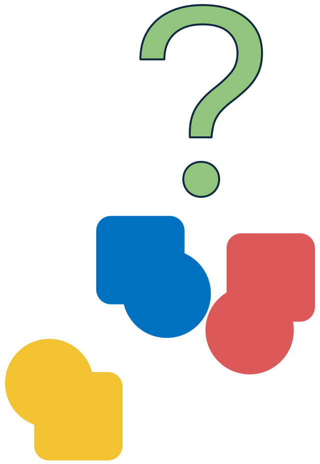
- Theory on domain adaptation suggests predictions must be made based on features that cannot discriminate between the training (source) and test (target) domains, for effective domain transfer;

- Thus, these approaches consist in promoting the emergence (during training) of features that are discriminative towards the source classes classification problem, and indiscriminative for a target/source domain classification task.

Domain Classification



Domain Classification

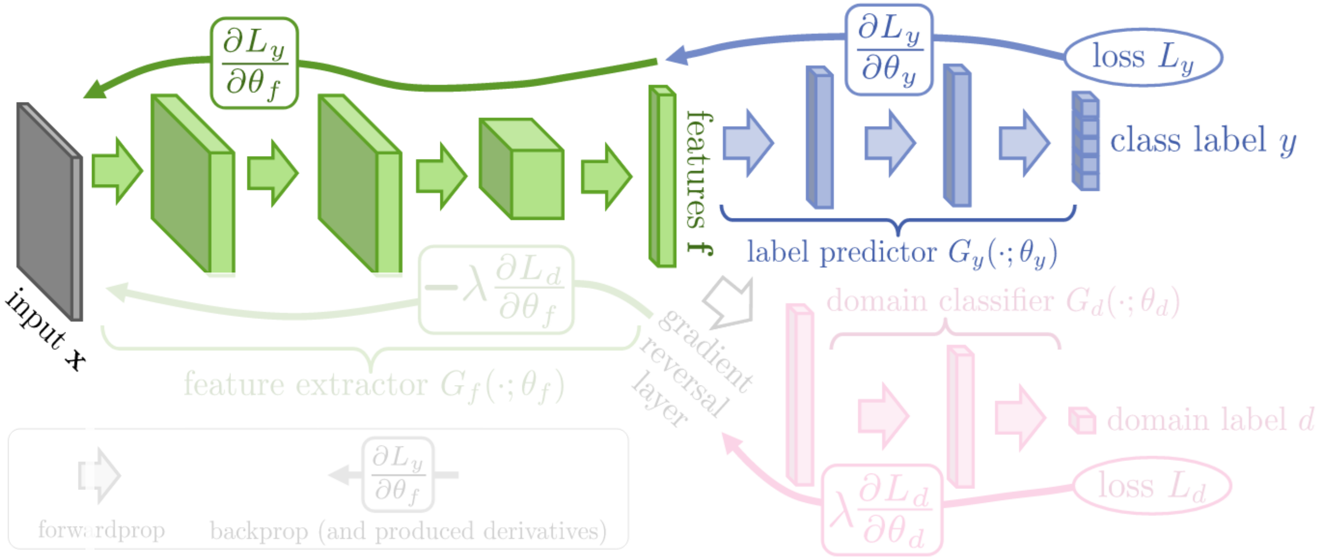


UniGe | MaLGa

Domain Adversarial Neural Network (DANN)

Domain Adversarial Neural Network (DANN) (1)

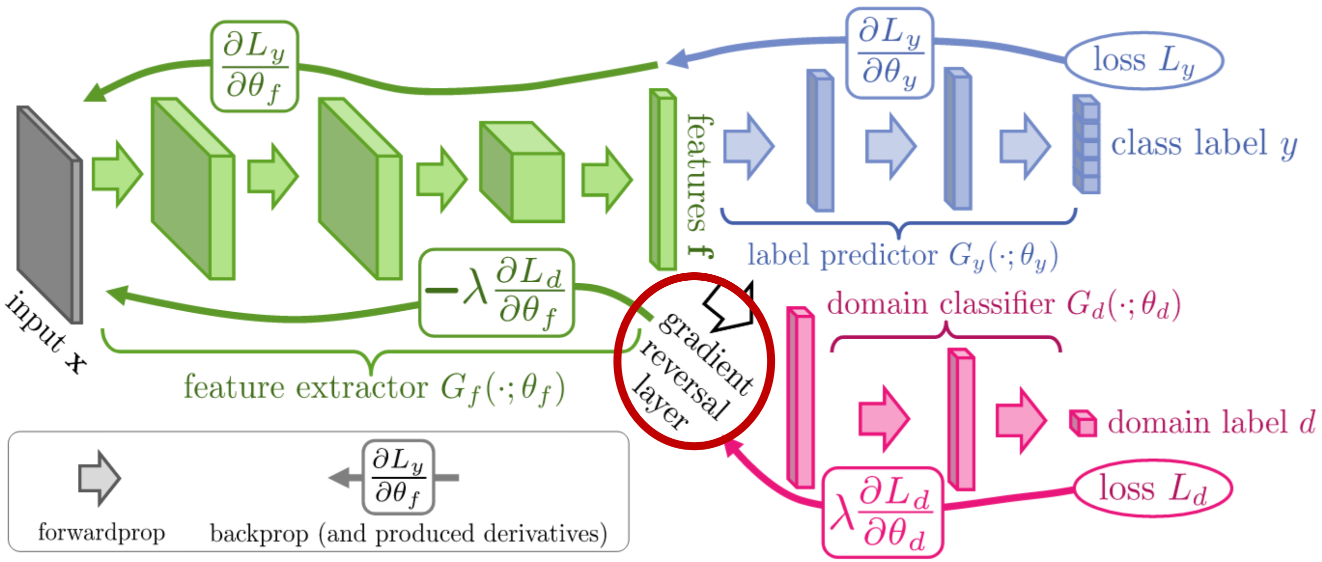
- The approach is generic as a DANN version can be created for almost any existing feed-forward architecture that is trainable by backpropagation;
- The first component is a regular feed-forward CNN, that predicts source labels



Ganin, Y., et al. "Domain-adversarial training of neural networks". The Journal of Machine Learning Research, 17(1), 2096-2030, 2016..

Domain Adversarial Neural Network (DANN) (2)

- Then, a domain classifier is connected to such network, via a gradient reversal layer, ensuring that target and source features are as not distinguishable as possible

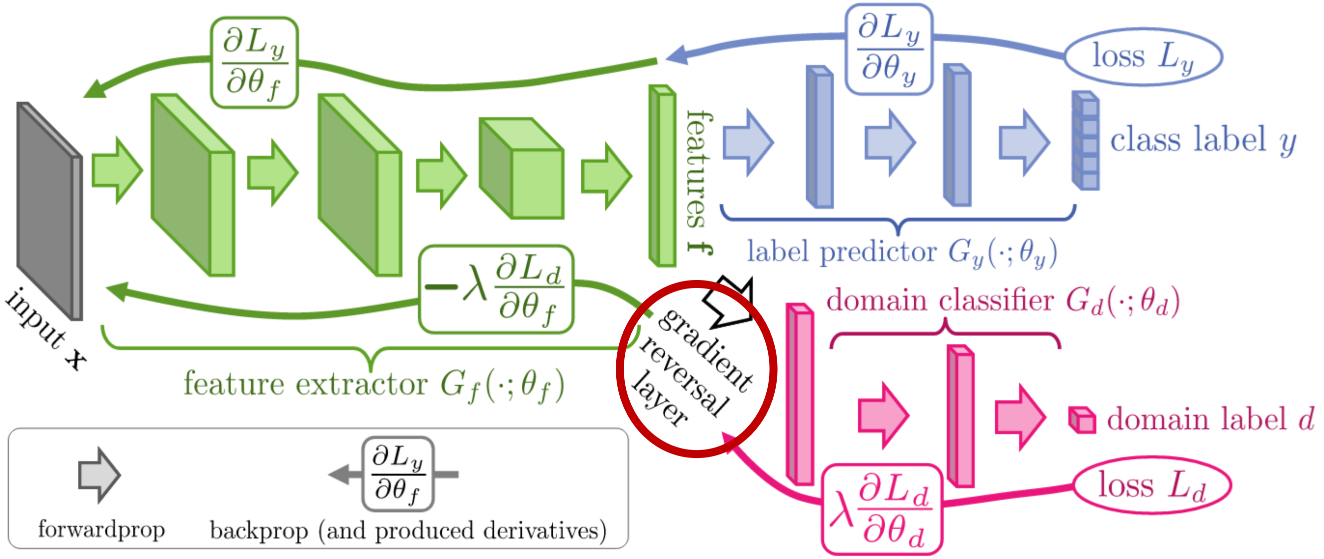


Ganin, Y., et al. "Domain-adversarial training of neural networks". The Journal of Machine Learning Research, 17(1), 2096-2030, 2016.

Domain Adversarial Neural Network (DANN) (2)

- Then, a domain classifier is connected to such network, via a gradient reversal layer, ensuring that target and source features are as not distinguishable as possible;

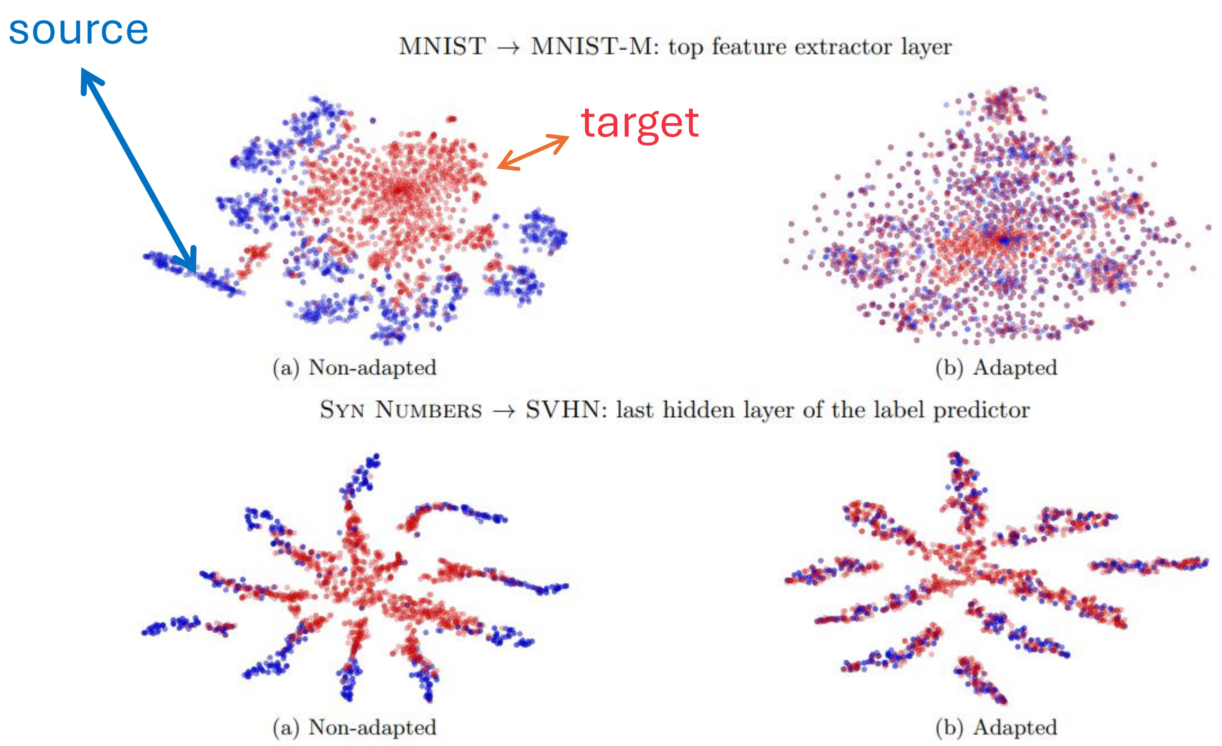
- Such layer multiplies the gradient by a certain negative constant during the backpropagation-based training.



Ganin, Y., et al. "Domain-adversarial training of neural networks". The Journal of Machine Learning Research, 17(1), 2096-2030, 2016.

Domain Adversarial Training

- Target and source features alignment with DANN

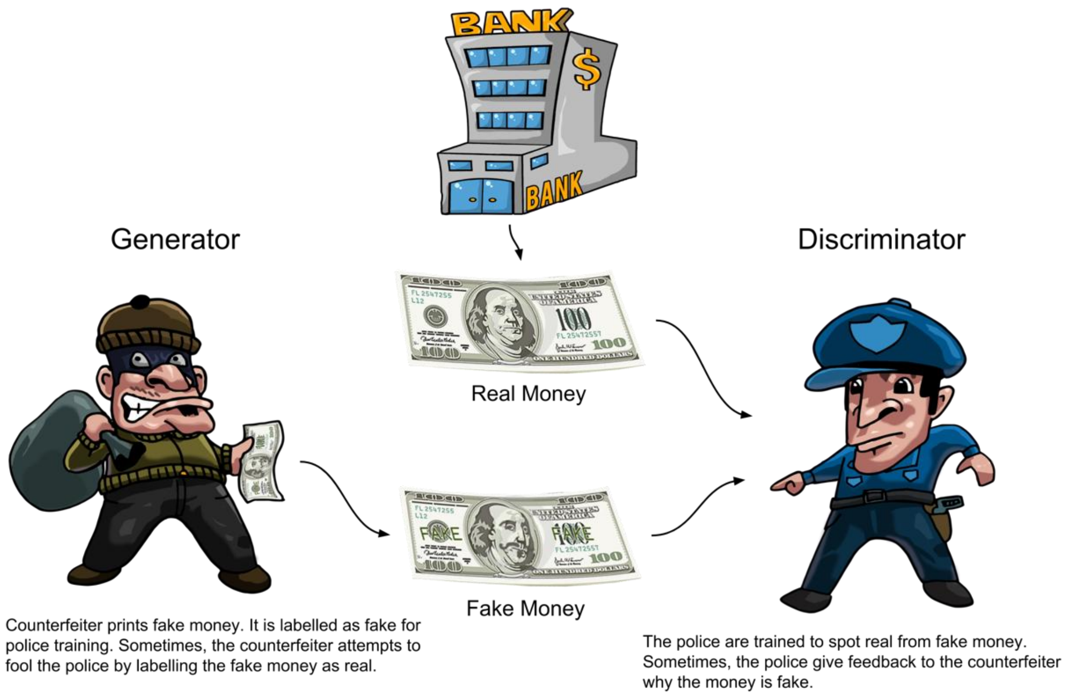


Ganin, Y., et al. “Domain-adversarial training of neural networks”. The Journal of Machine Learning Research, 17(1), 2096-2030, 2016..

UniGe | MaLGa

Adversarial Discriminative Domain Adaptation (ADDA)

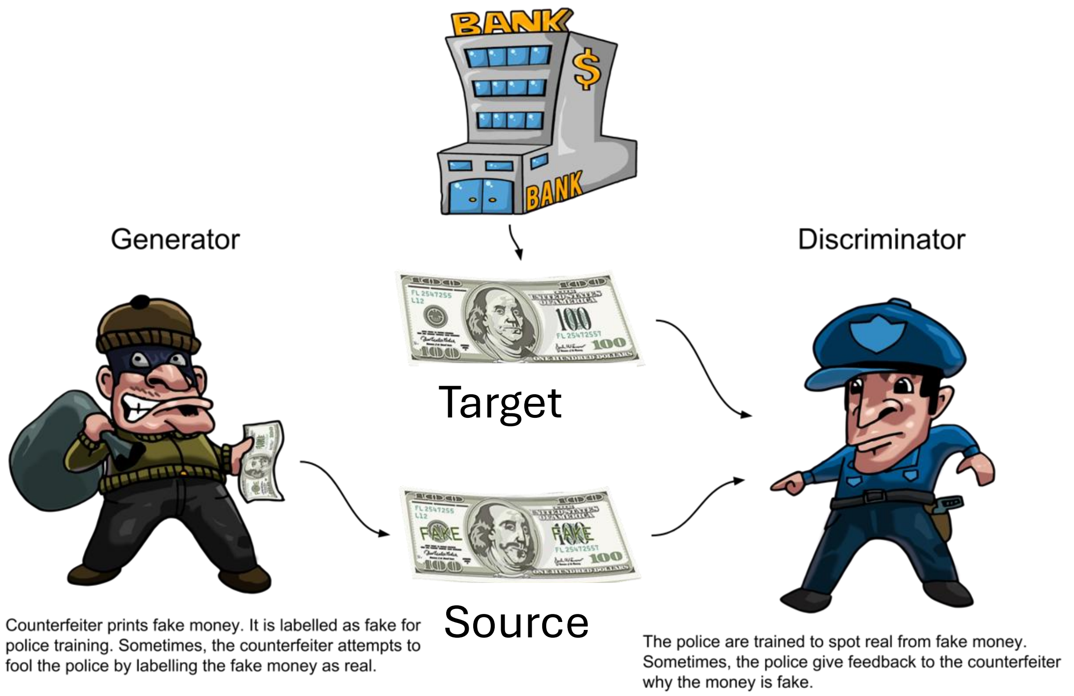
GANs and Domain adaptation (1)



Long, M., et al. “Learning Transferable Features with Deep Adaptation Networks”. In ICML 2015..

Image: Courtesy to Richard Gall.

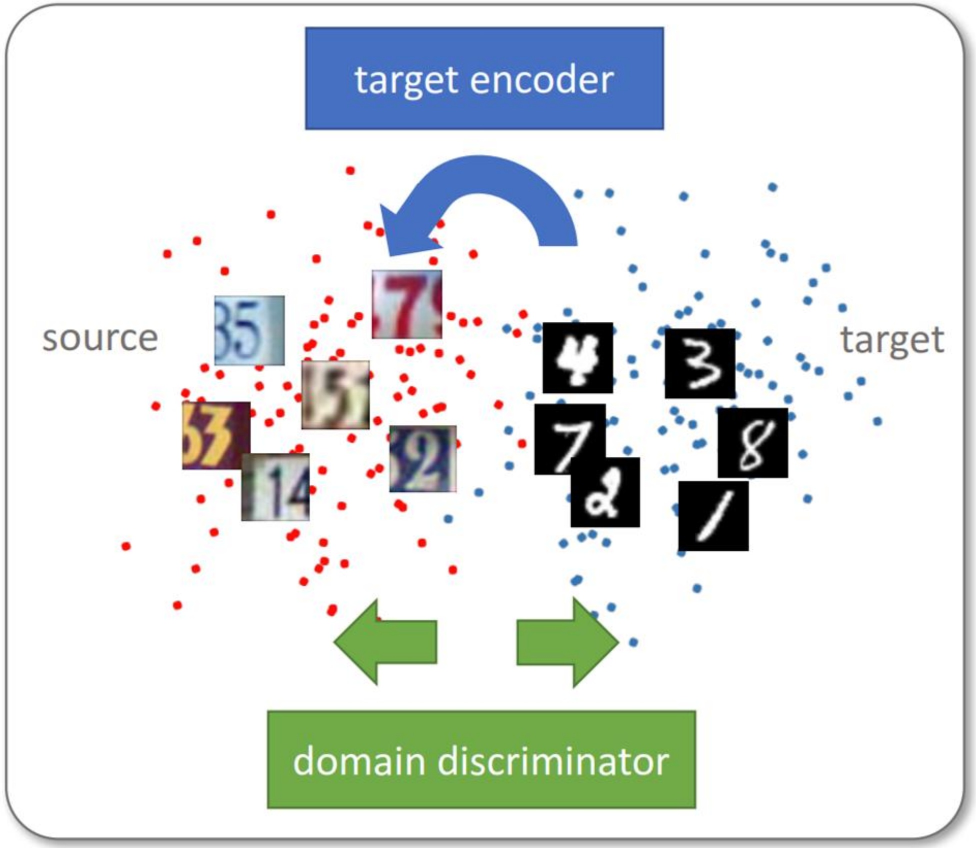
GANs and Domain adaptation (2)



Long, M., et al. "Learning Transferable Features with Deep Adaptation Networks". In ICML 2015..

Image: Courtesy to Richard Gall.

GANs and Domain adaptation (3)



Generative Adversarial Neural networks

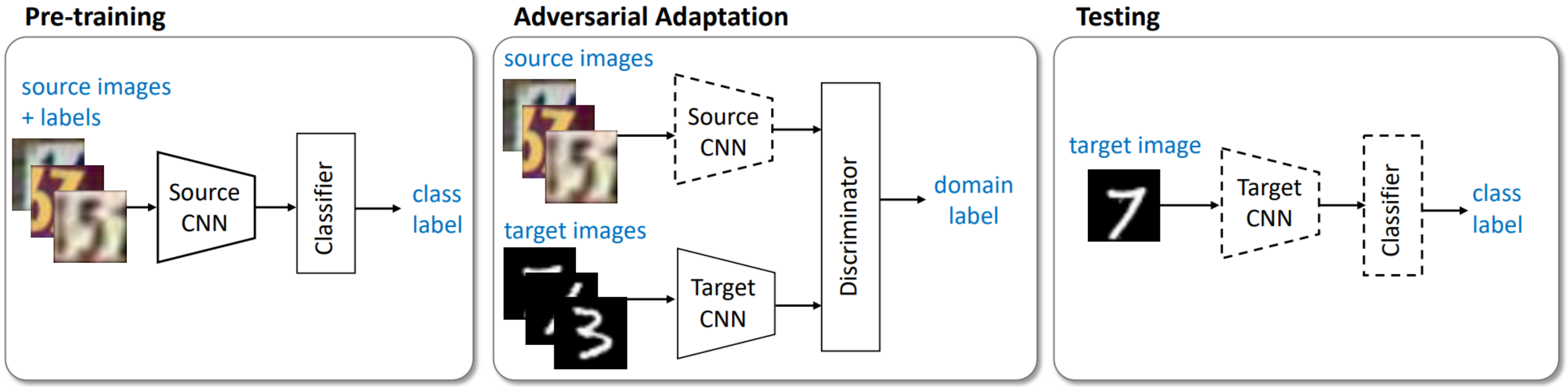
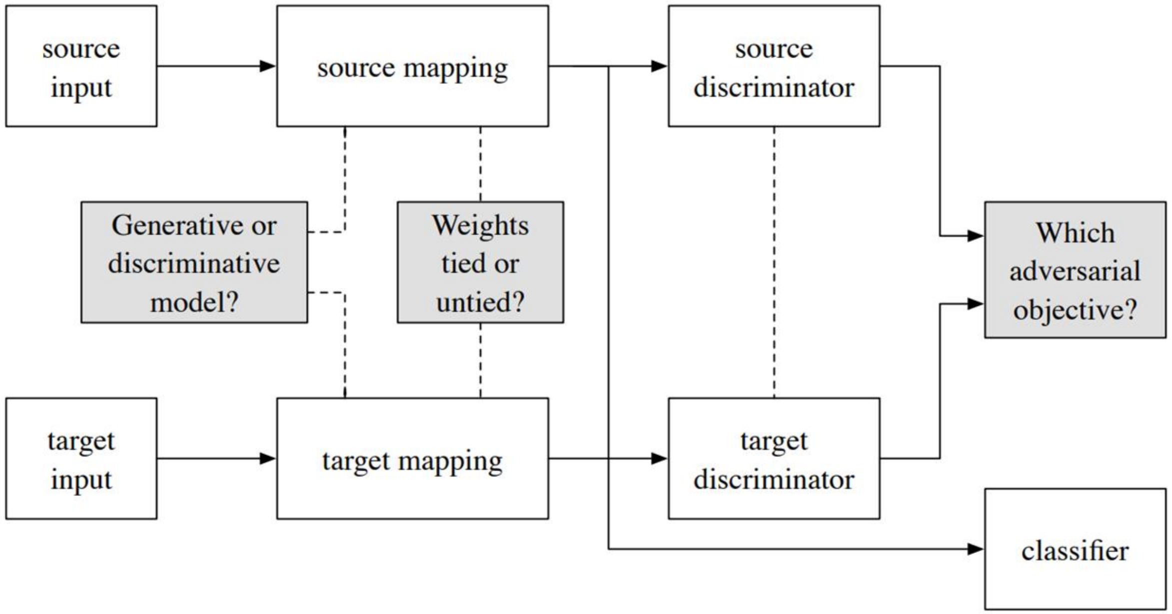


Figure 3: An overview of our proposed Adversarial Discriminative Domain Adaptation (ADDA) approach. We first pre-train a source encoder CNN using labeled source image examples. Next, we perform adversarial adaptation by learning a target encoder CNN such that a discriminator that sees encoded source and target examples cannot reliably predict their domain label. During testing, target images are mapped with the target encoder to the shared feature space and classified by the source classifier. Dashed lines indicate fixed network parameters.

Long, M., et al. “Learning Transferable Features with Deep Adaptation Networks”. In ICML 2015..

Tzeng, E., Hoffman, J., Saenko, K., & Darrell, T., ” Adversarial discriminative domain adaptation”. In CVPR 2017.

In general, ADDA-like approaches



Tzeng, E., Hoffman, J., Saenko, K., & Darrell, T., " Adversarial discriminative domain adaptation". In CVPR 2017.

Reconstruction-based domain adaptation (with Image Translation)

Deep Reconstruction-Classification Network (DRCN) (1)

- Auxiliary reconstruction task on target data is used to support the adaptation of target labels

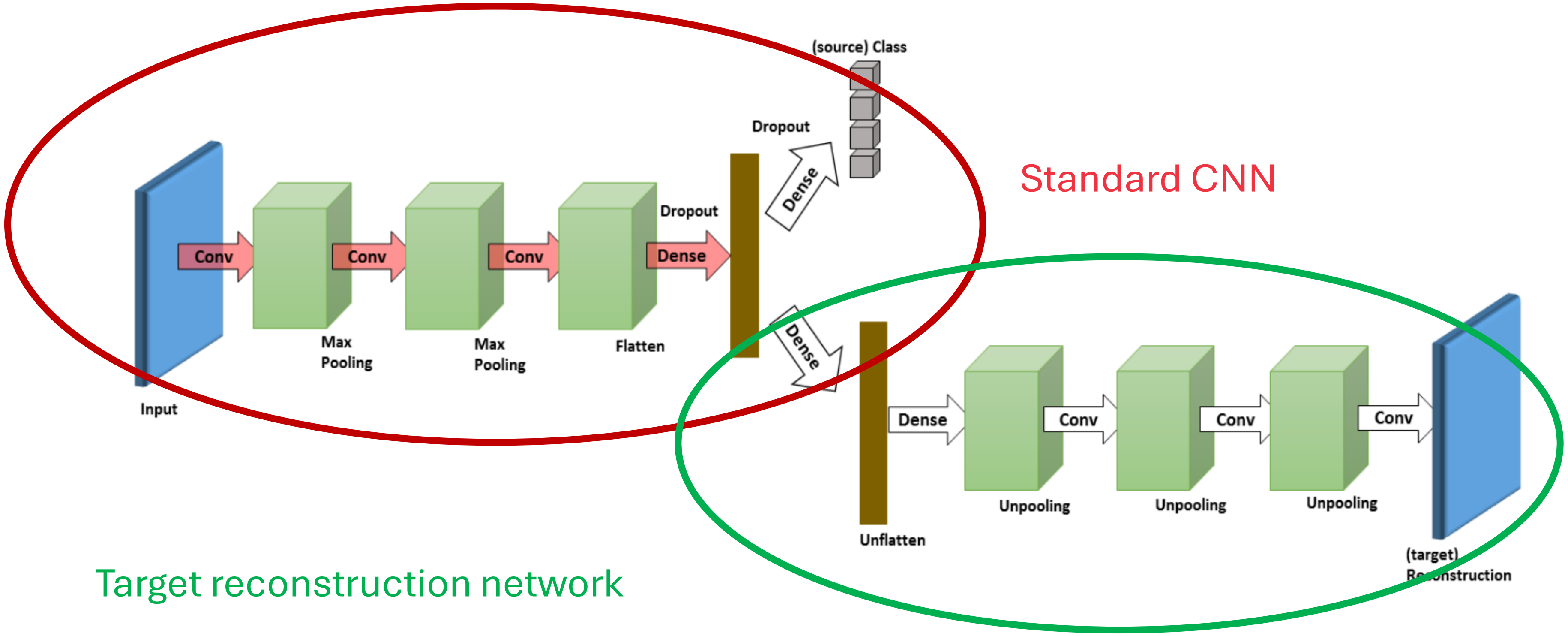


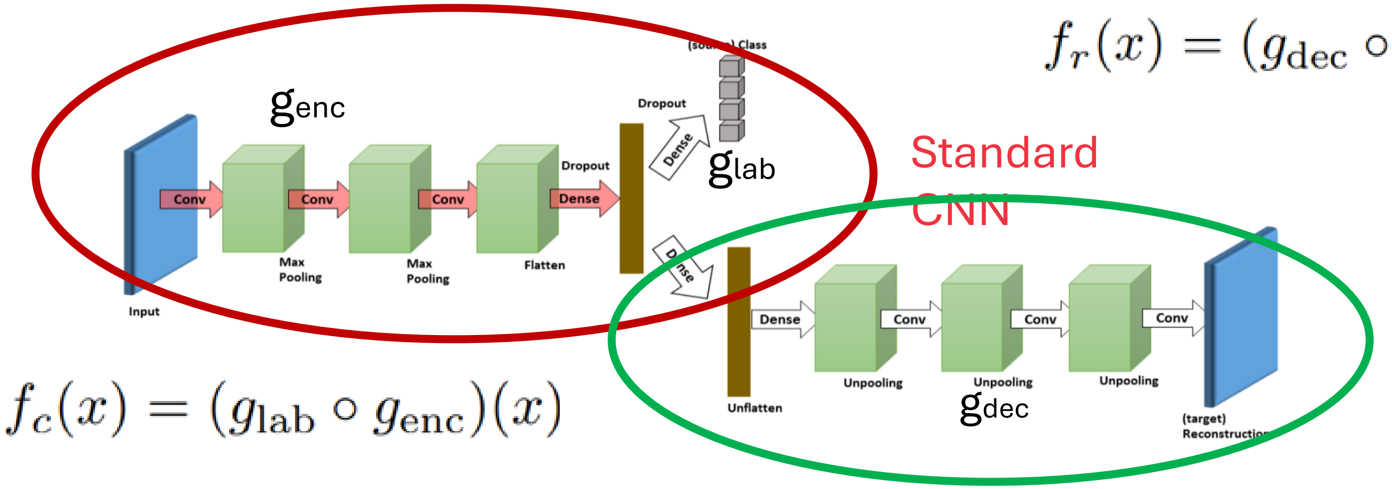
Fig. 1. Illustration of the DRCN's architecture. It consists of two pipelines: i) label prediction and ii) data reconstruction pipelines. The shared parameters between those two pipelines are indicated by the red color.

Ganin, Y., et al. "Domain-adversarial training of neural networks". The Journal of Machine Learning Research, 17(1), 2096-2030, 2016.

Deep Reconstruction-Classification Network (DRCN) (2)

- Auxiliary reconstruction task on target data is used to support the adaptation of target labels

$$f_r(x) = (g_{\mathrm{dec}} \circ g_{\mathrm{enc}})(x)$$



Standard CNN

$$f_c(x) = (g_{\mathrm{lab}} \circ g_{\mathrm{enc}})(x)$$

Target reconstruction network

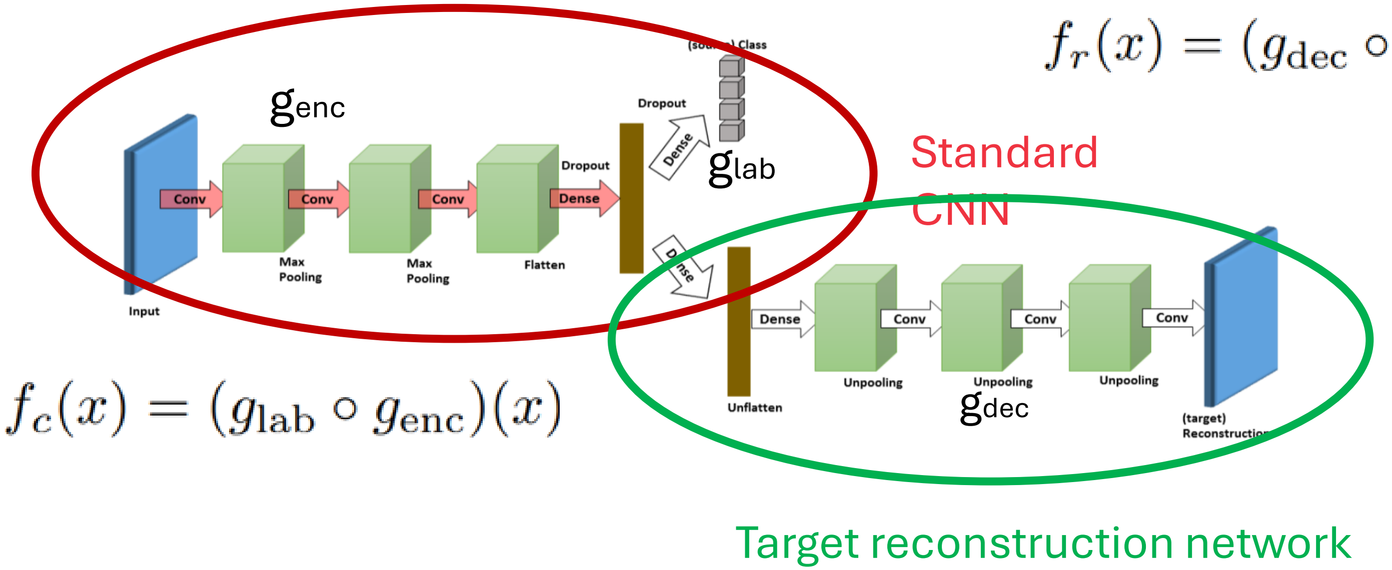
Let fc : X → Y be the (supervised) label prediction pipeline
and fr : X → X be the (unsupervised) data reconstruction pipeline. Define three additional functions:

- an encoder / feature mapping g_{enc} : X → F, 2)
- decoder g_{dec} : F → X
- feature labeling g_{lab} : F → Y.

Ghifary, M., Kleijn, W. B., Zhang, M., Balduzzi, D., & Li, W. (2016). Deep reconstruction-classification networks for unsupervised domain adaptation. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14 (pp. 597-613). Springer International Publishing.

Deep Reconstruction-Classification Network (DRCN) (3)

$$f_r(x) = (g_{\mathrm{dec}} \circ g_{\mathrm{enc}})(x)$$



$$f_c(x) = (g_{\mathrm{lab}} \circ g_{\mathrm{enc}})(x)$$

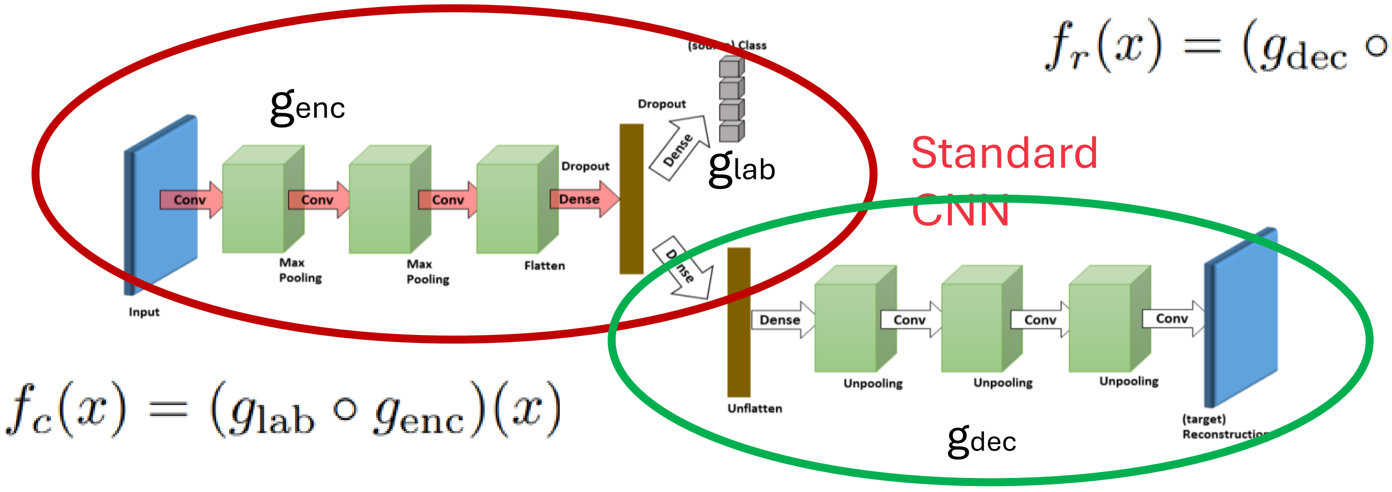
$\Theta_c = \{\Theta_{\mathrm{enc}}, \Theta_{\mathrm{lab}}\}$ Params of supervised model

$\Theta_r = \{\Theta_{\mathrm{enc}}, \Theta_{\mathrm{dec}}\}$ Params of unsupervised model

Ghifary, M., Kleijn, W. B., Zhang, M., Balduzzi, D., & Li, W. (2016). Deep reconstruction-classification networks for unsupervised domain adaptation. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14 (pp. 597-613). Springer International Publishing.

Deep Reconstruction-Classification Network (DRCN) (3)

$$f_r(x) = (g_{\text{dec}} \circ g_{\text{enc}})(x)$$



$$f_c(x) = (g_{\text{lab}} \circ g_{\text{enc}})(x)$$

Target reconstruction network

Single objective functions

$$\mathcal{L}_c^{n_s}(\{\Theta_{\text{enc}}, \Theta_{\text{lab}}\}) := \sum_{i=1}^{n_s} \ell_c\left(f_c(\mathbf{x}_i^s; \{\Theta_{\text{enc}}, \Theta_{\text{lab}}\}), \mathbf{y}_i^s\right),$$

$$\mathcal{L}_r^{n_t}(\{\Theta_{\text{enc}}, \Theta_{\text{dec}}\}) := \sum_{j=1}^{n_t} \ell_r\left(f_r(\mathbf{x}_j^t; \{\Theta_{\text{enc}}, \Theta_{\text{dec}}\}), \mathbf{x}_j^t\right).$$

Objective to minimize

$$\min \lambda \mathcal{L}_c^{n_s}(\{\Theta_{\text{enc}}, \Theta_{\text{lab}}\}) + (1-\lambda)\mathcal{L}_r^{n_t}(\{\Theta_{\text{enc}}, \Theta_{\text{dec}}\})$$

Ghifary, M., Kleijn, W. B., Zhang, M., Balduzzi, D., & Li, W. (2016). Deep reconstruction-classification networks for unsupervised domain adaptation. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14 (pp. 597-613). Springer International Publishing.

Deep Reconstruction-Classification Network (DRCN) (3)

Algorithm 1 The Deep Reconstruction-Classification Network (DRCN) learning algorithm.

Input:
- Labeled source data: $S^s = \{(\mathbf{x}_i^s, y_i^s)\}_{i=1}^{n_s}$;
- Unlabeled target data: $S_u^t = \{\mathbf{x}_j^t\}_{i=j}^{n_t}$;
- Learning rates: α_c and α_r;

```
Initialize parameters Θ_enc, Θ_dec, Θ_lab
while not stop do
    for each source batch of size m_s do
        Do a forward pass according to (1);
        Let Θ_c = {Θ_enc, Θ_lab}. Update Θ_c:
```

$$\Theta_c \leftarrow \Theta_c - \alpha_c \lambda \nabla_{\Theta_c} \mathcal{L}_c^{m_s}(\Theta_c);$$

```
    end for
    for each target batch of size m_t do
        Do a forward pass according to (2);
        Let Θ_r = {Θ_enc, Θ_dec}. Update Θ_r:
```

$$\Theta_r \leftarrow \Theta_r - \alpha_r (1-\lambda) \nabla_{\Theta_r} \mathcal{L}_r^{m_t}(\Theta_r).$$

```
    end for
end while
```

Output:
- DRCN learnt parameters: $\hat{\Theta} = \{\hat{\Theta}_{\text{enc}}, \hat{\Theta}_{\text{dec}}, \hat{\Theta}_{\text{lab}}\}$;

Supervised source label classification

Unsupervised target reconstruction

Ghifary, M., Kleijn, W. B., Zhang, M., Balduzzi, D., & Li, W. (2016). Deep reconstruction-classification networks for unsupervised domain adaptation. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14 (pp. 597-613). Springer International Publishing.

Introduction to Source-Free domain adaptation and domain generalization

Unsupervised Domain Adaptation – Recap

TRAINING

Target Domain training set

No labels

+

Source Domain training set

Labels

TEST

Target Domain test set

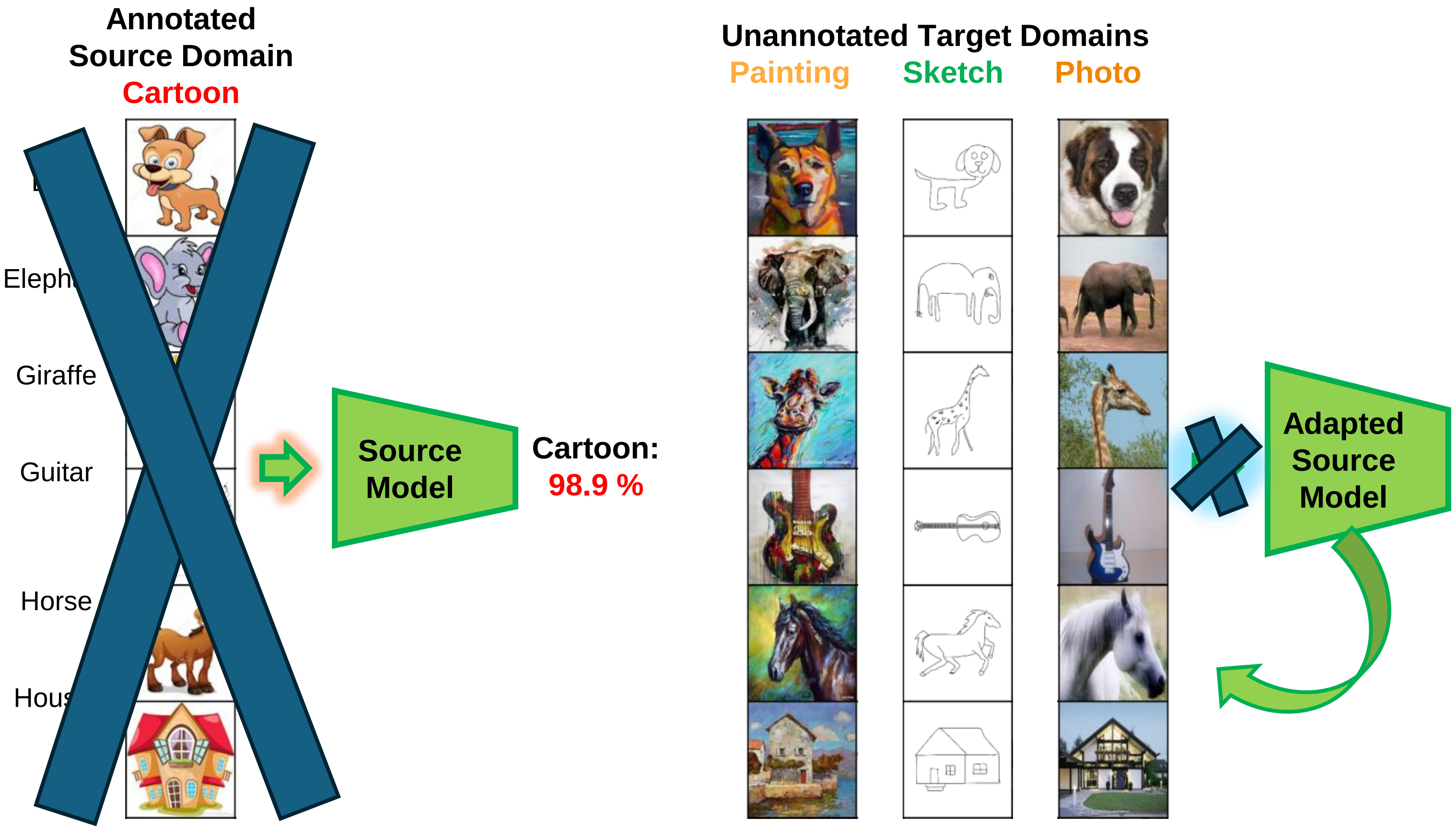
Labels

+

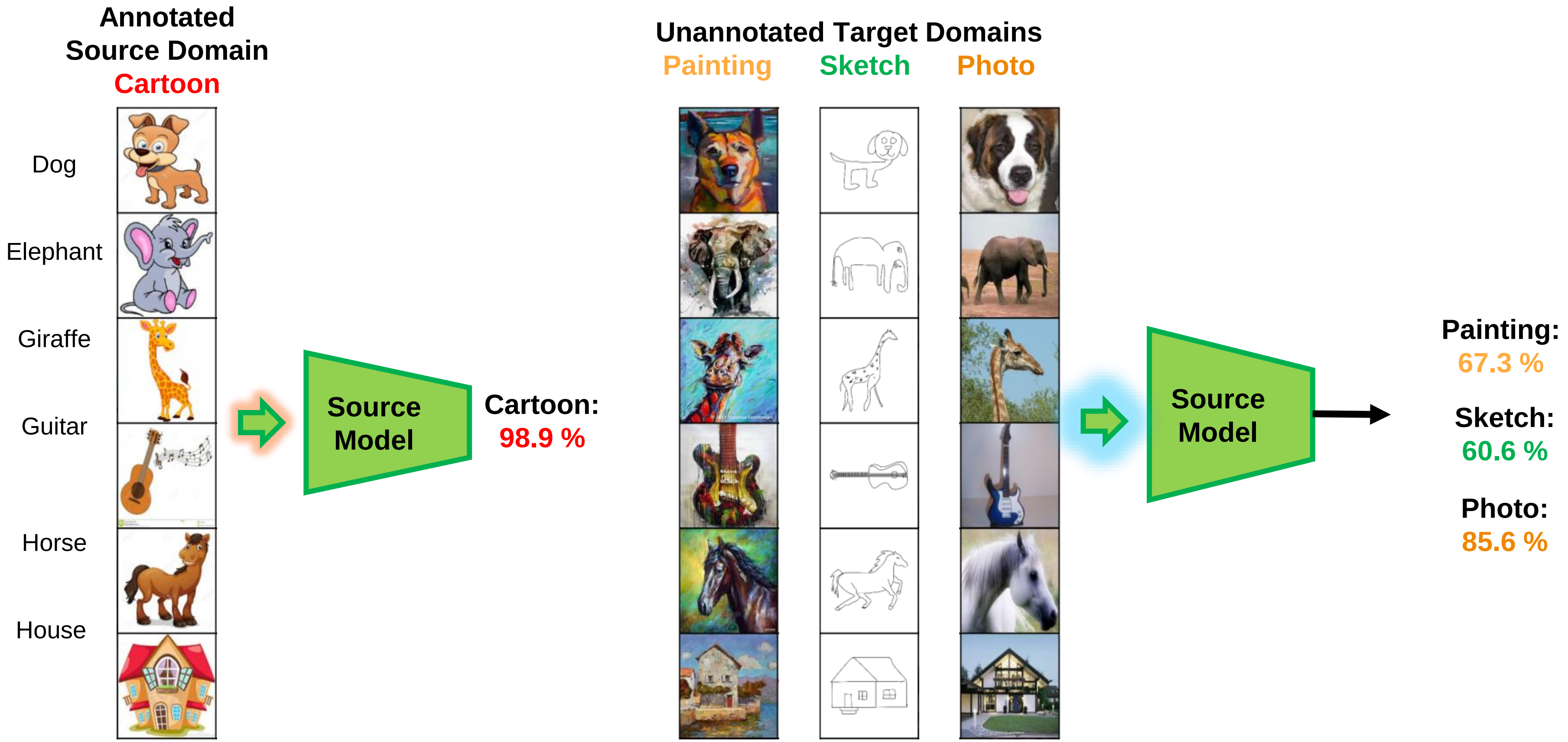
Source Domain test set

Labels

Unsupervised Domain Adaptation

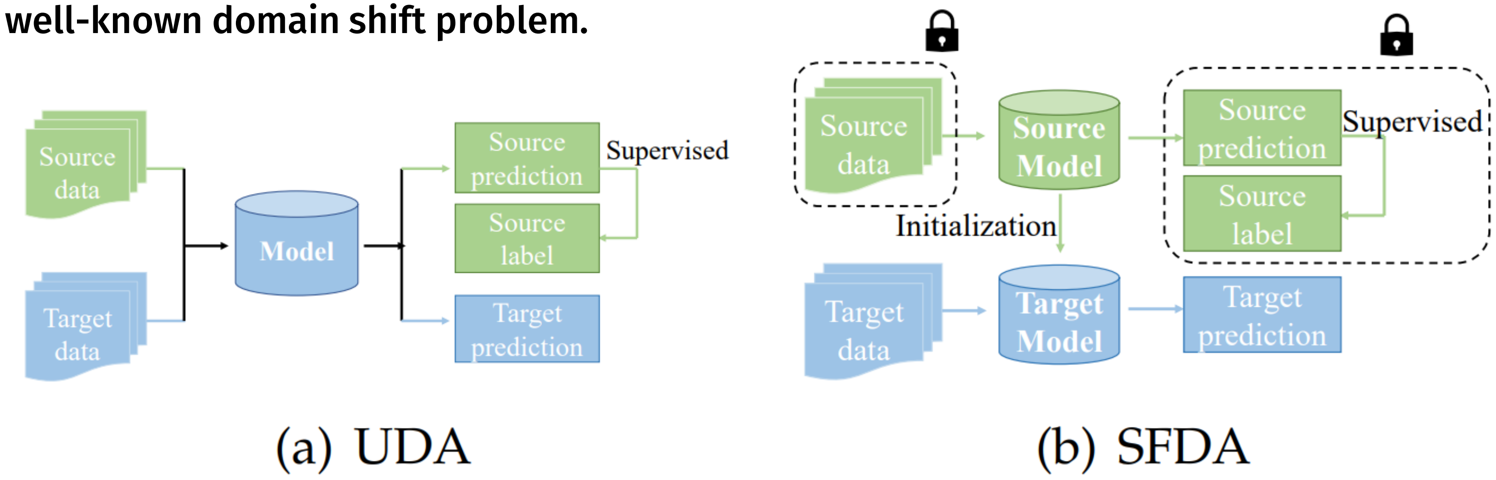


Unsupervised Domain Adaptation



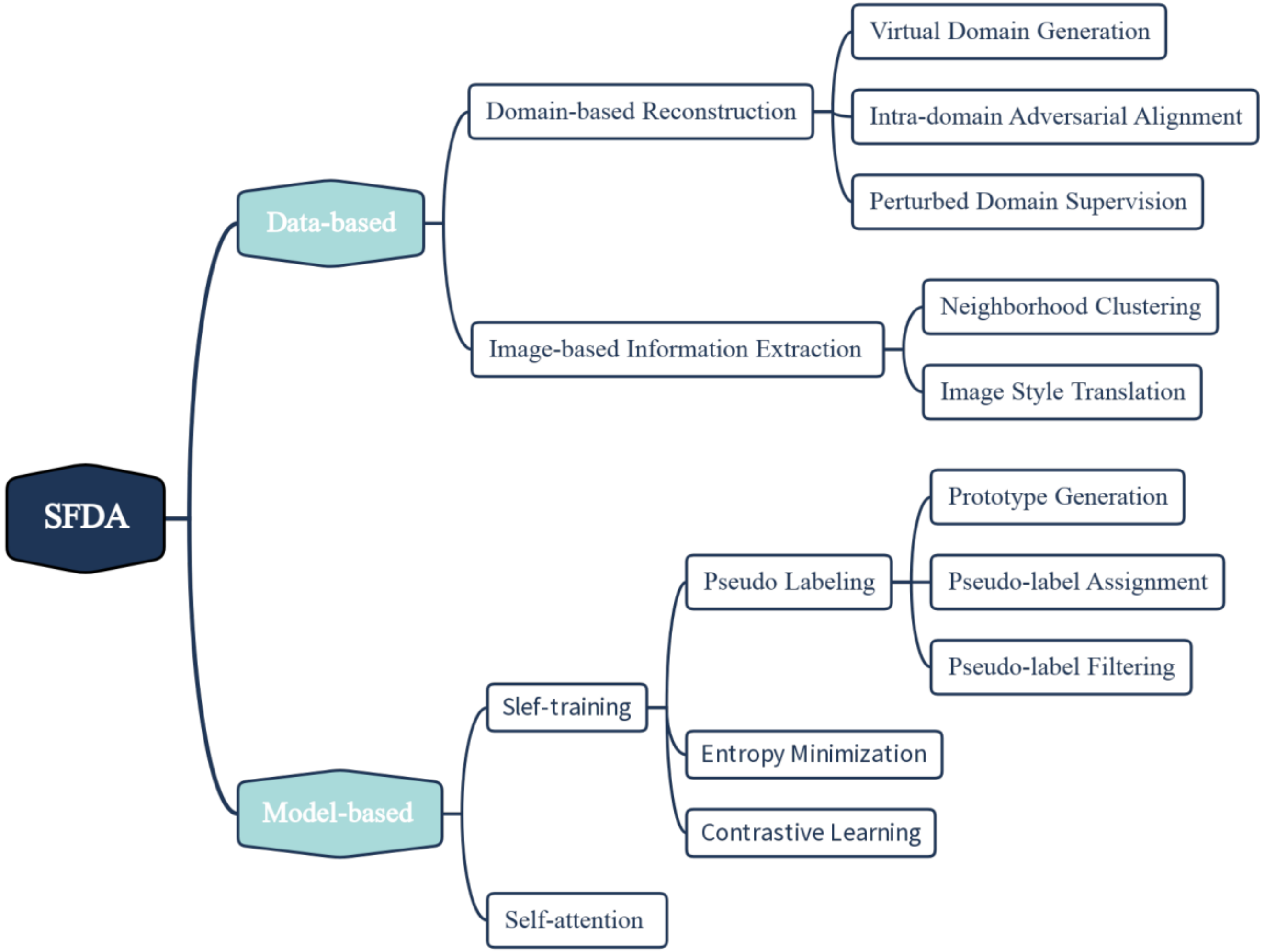
Source-free Unsupervised Domain Adaptation (SFDA)

- Conventional UDA methods assume **source and target data to be simultaneously available** during training.
- They either try to align target and source domain distributions by certain metrics (discriminative methods), or to obtain domain-invariant features (Adversarial and reconstruction)
- Such an assumption may not hold in practice, as source data is often inaccessible (e.g., due to privacy reasons).
- UDA model can be trained using both the source and target domain data, **while SFDA can only utilize the source model** to initialize the target model and then update it with the unlabeled target data
- On the contrary, a pre-trained source model is usually available, which performs poorly on target due to the well-known domain shift problem.



Li, J., Yu, Z., Du, Z., Zhu, L., & Shen, H. T. (2024). A comprehensive survey on source-free domain adaptation. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Source-free Unsupervised Domain Adaptation (SFDA) (2)



Li, J., Yu, Z., Du, Z., Zhu, L., & Shen, H. T. (2024). A comprehensive survey on source-free domain adaptation. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Examples of Data-centric and model-centric approaches and intuitions

Data-centric

- **Domain-based reconstructions** methods may aim to reconstruct a domain or make further divisions within the target domain to compensate for missing source domain data, so that the UDA approaches can be extended to the SFDA setting;

- **Neighborhood clustering** is based on the observation that the underlying intrinsic structure in the target data is embedded in the neighbor relationship, even though the target domain data distribution can not explicitly align with the source classifier (in each category of the target domain will exist neighbors located within the source model's boundary, so those hard samples can be mapped to the source domain distribution by maintaining consistency among neighbors).

Model-centric

- Source pre-trained model has a certain degree of generalization over the target domain due to the similarity of the source and target domains Thus, pseudolabeling may be used, specifically by first pseudo-labeling the high-confidence samples in the target domain, and optimize the model thereafter by the obtained pseudo-labels

Li, J., Yu, Z., Du, Z., Zhu, L., & Shen, H. T. (2024). A comprehensive survey on source-free domain adaptation. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Domain Generalization

Problem formulation

Each dataset carries its own bias [1], and models trained on it result biased, too.

[1] A. Torralba and A. Efros. **Unbiased Look at Dataset Bias**. *CVPR11*

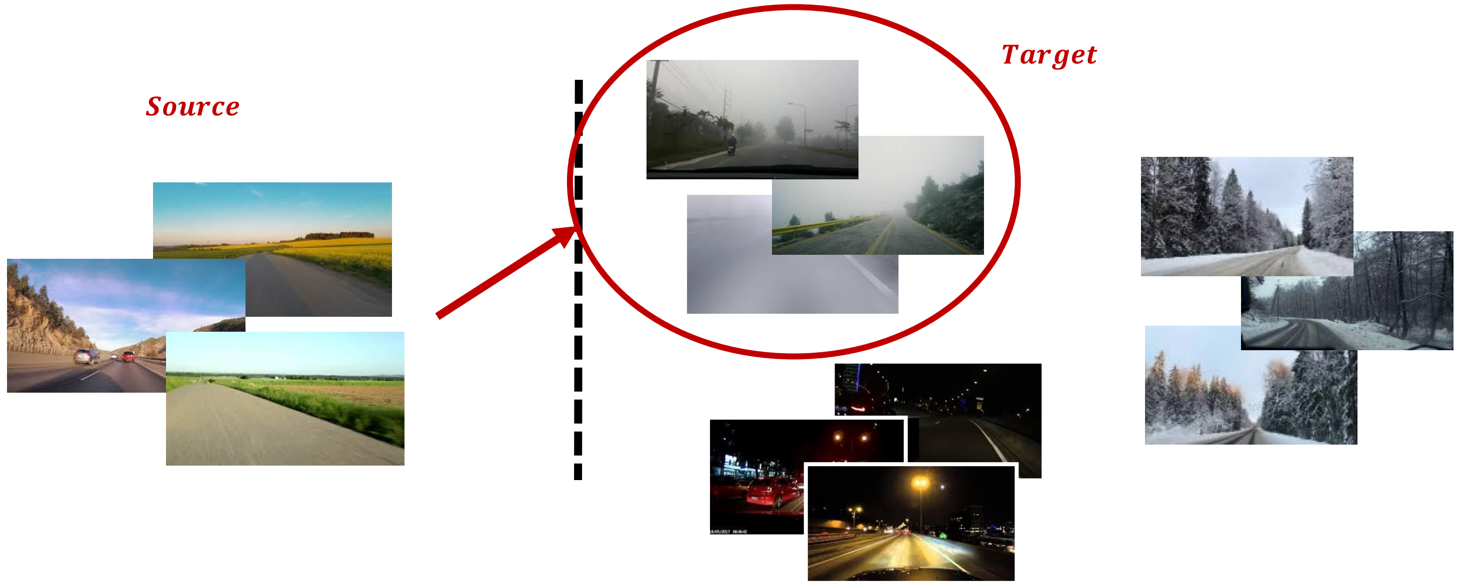
Domain adaptation

Source

Domain adaptation has been the main strategy to bridge the gap between source and target distributions.

Assumptions: *we can fix a priori a target distribution and we are able to sample from it.*

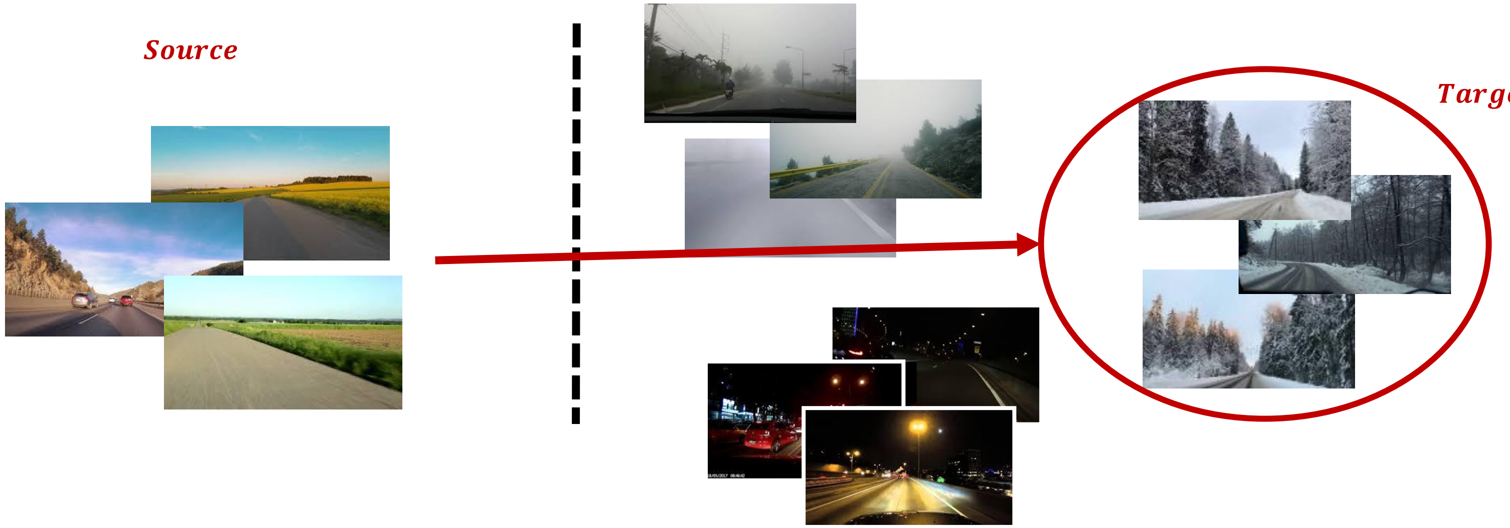
Domain adaptation



Domain adaptation has been the main strategy to bridge the gap between source and target distributions.

Assumptions: *we can fix a priori a target distribution and we are able to sample from it.*

Domain adaptation



Domain adaptation has been the main strategy to bridge the gap between source and target distributions.

Assumptions: *we can fix a priori a target distribution and we are able to sample from it.*

Domain adaptation

Domain adaptation has been the main strategy to bridge the gap between source and target distributions.

Assumptions: *we can fix a priori a target distribution and we are able to sample from it.*

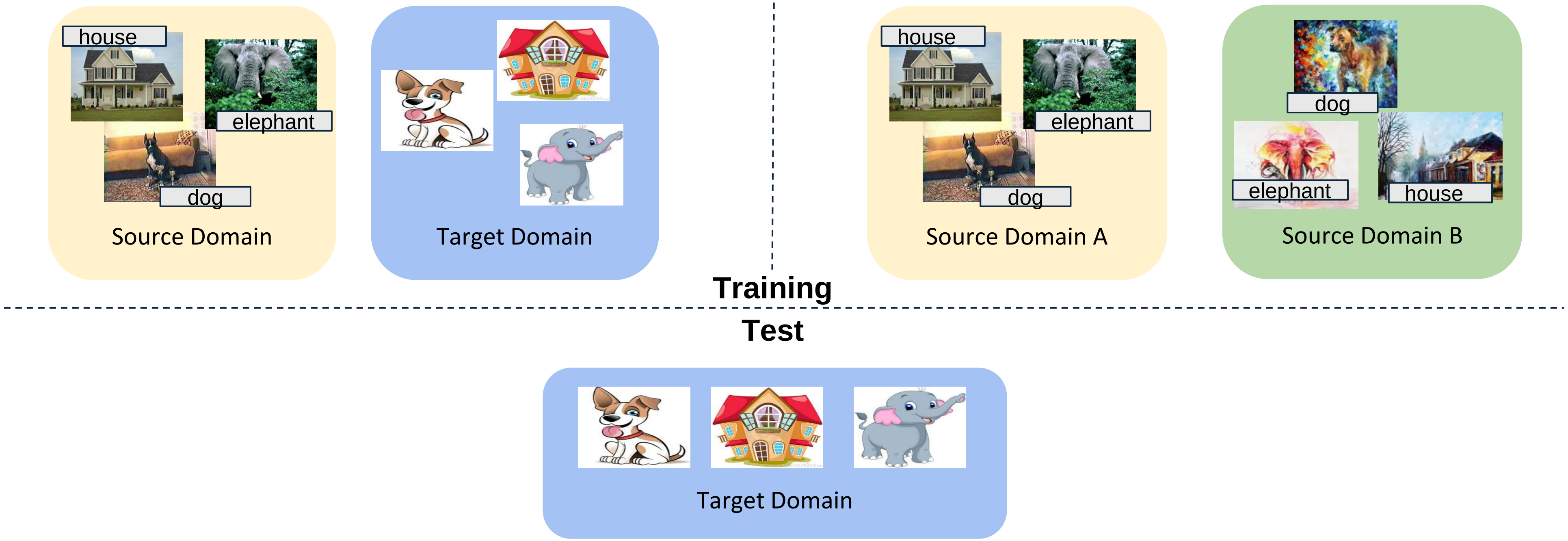
Generalising to unseen domains …

Goal: generalizing to unseen domains using data from a single source.

Domain Generalization (DG)

Domain Generalization:

Given a set of multiple labeled source domains, we want to find a model able to generalize to any target domain for which no data are available during training:



Domain Adaptation:
Given a one or multiple source domains for which we have labeled data, we want to find a model able to generalize to a target domain for which few or no labeled data are available during training.

UniGe | MaLGa