Domain Adaptation for Visual Applications
Part 3: Beyond Classical Domain Adaptation

Tatiana Tommasi
Assistant Professor, Polytechnic University of Turin, Italy
Affiliated Researcher, Italian Institute of Technology
Outline

● An Overview on several Cross-Domain Learning Settings
  ○ (annotated) source data
  ○ (annotated) target data
  ○ source / target overlap

● Self-Supervision for Cross-Domain Learning
Classical Domain Adaptation

Source (Train)  Target (Test)
Classical Domain Adaptation

Source (Train) \rightarrow Target (Test)

Labelled Source Domain \rightarrow Unlabelled Target Domain

Unsupervised DA, transductive setting

Train \rightarrow Test
Target data **not available** at training time

Target data **available** but not annotated

Annotated Source data

Annotated Target data
Annotated Source data

Multiple Source Domains

One Source Domain

Annotated Target data

Target data **not available** at training time

Target data **available** but not annotated

Only Source Model available, **(no source data)**
Transductive Unsupervised Domain Adaptation

- Annotated Source data
- Unsupervised DA, transductive setting
- Annotated Target data
Transductive Unsupervised Domain Adaptation

Annotated Source data

semi-supervised DA

Annotated Target data

Labelled Source Domain

Labelled Target Domain

Train

Test

Unlabelled

ECCV’20 ONLINE
23-28 AUGUST 2020
Transfer Learning

Model

Transductive Unsupervised Domain Adaptation

semi-supervised DA

Annotated Source data

Annotated Target data

Source Domain Model

Labelled Target Domain

Train

Test

Transfer Learning
Transfer Learning

Transductive Unsupervised Domain Adaptation

Annotated Source data

Annotated Target data

multi-source Domain Generalization

semi-supervised DA

Transfer Learning
Transfer Learning

Transductive Unsupervised Domain Adaptation

semi-supervised DA

Annotated Source data

Annotated Target data

Transfer Learning

multi-source Domain Generalization

[Deeper, Broader and Artier Domain Generalization, ICCV 2017]

Annotated Source data

[Deeper, Broader and Artier Domain Generalization, ICCV 2017]
Transfer Learning

Transductive Unsupervised Domain Adaptation

semi-supervised DA

Annotated Source data

Annotated Target data

multi-source Domain Generalization

[Generalizing to Unseen Domains via Adversarial Data Augmentation, NeurIPS 2018]

[Learning to Learn Single Domain Generalization, CVPR 2020]

single-source DG

Transfer Learning

Source data
**Transfer Learning**

- **Transductive Unsupervised Domain Adaptation**
- **semi-supervised DA**
- **Annotated Source data**
- **Annotated Target data**

**Multi-source Domain Generalization**

**Single-source DG**

- [Generalizing to Unseen Domains via Adversarial Data Augmentation, NeurIPS 2018]
- [Learning to Learn Single Domain Generalization, CVPR 2020]

**Sample in Source Domain**

**Sample in Target Domain**

**Augmented Sample**
Single Source Domain Generalization

\[ \mathcal{L}_{\text{const}} = \frac{1}{2} \| z - z^+ \|_2^2 + \infty \cdot 1 \{ y \neq y^+ \} \]
Single Source Domain Generalization

Learning to Learn Single Domain Generalization, CVPR 2020

\[ \mathcal{L}_{\text{const}} = \frac{1}{2} \| \mathbf{z} - \mathbf{z}^+ \|_2^2 + \infty \cdot 1 \{ \mathbf{y} \neq \mathbf{y}^+ \} \]

\[ \mathcal{L}_{\text{relax}} = \| \mathbf{x}^+ - V(x^+) \|_2^2 \]
Single Source Domain Generalization

\[ L_{\text{const}} = \frac{1}{2} \| z - z^+ \|_2^2 + \infty \cdot 1 \{ y \neq y^+ \} \]

\[ L_{\text{relax}} = \| x^+ - V(x^+) \|_2^2 \]

\[ L_{\text{ADA}} = L_{\text{task}}(\theta; x) - \alpha L_{\text{const}}(\theta; z) + \beta L_{\text{relax}}(\psi; x) \]

[Learning to Learn Single Domain Generalization, CVPR 2020]
Zero-Shot Domain Adaptation

[Zero-Shot Deep Domain Adaptation, ECCV 2018]
Zero-Shot Domain Adaptation

[Zero-Shot Deep Domain Adaptation, ECCV 2018]

Step 1: simulate target representation

Step 2: domain adaptation

source domain: gray scale
target domain: RGB

- available at training
- unavailable at training
Zero-Shot Domain Adaptation

[Zero-Shot Deep Domain Adaptation, ECCV 2018]
Zero-Shot Domain Adaptation

[Zero-Shot Deep Domain Adaptation, ECCV 2018]
Zero-Shot Domain Generalization

[Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020]
Zero-Shot Domain Generalization

[Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020]
Zero-Shot Domain Generalization

[Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020]
Zero-Shot Domain Generalization

[Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020]
Zero-Shot Domain Generalization

[Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020]
Zero-Shot Domain Generalization

[Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020]
Zero-Shot Domain Generalization

\[ \phi(a_i, a_j, a_k) = \lambda a_i + (1 - \lambda)(\alpha a_j + (1 - \alpha) a_k) \]

Domain Picking
Zero-Shot Domain Generalization

Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020

\[ \phi(a_i, a_j, a_k) = \lambda a_i + (1 - \lambda)(\alpha a_j + (1 - \alpha)a_k) \]
Zero-Shot Domain Generalization

[Towards Recognizing Unseen Categories in Unseen Domains, ECCV 2020]
Transfer Learning

Transductive Unsupervised Domain Adaptation

Annotated Source data

Annotated Target data

zero-shot DA

zero-shot DG

multi-source Domain Generalization

predictive DA

Annotated Source data

multi-source DA

single-source DG

zero-shot DG

single-source DG
Predictive DA

[Multivariate Regression on the Grassmannian for Predicting Novel Domains, CVPR 2016]
[AdaGraph: Unifying Predictive and Continuous Domain Adaptation through Graphs, CVPR 2019]
Predictive DA

[Multivariate Regression on the Grassmannian for Predicting Novel Domains, CVPR 2016]
[AdaGraph: Unifying Predictive and Continuous Domain Adaptation through Graphs, CVPR 2019]
Predictive DA

[Multivariate Regression on the Grassmannian for Predicting Novel Domains, CVPR 2016]
[AdaGraph: Unifying Predictive and Continuous Domain Adaptation through Graphs, CVPR 2019]
Predictive DA

[Multivariate Regression on the Grassmannian for Predicting Novel Domains, CVPR 2016]
[AdaGraph: Unifying Predictive and Continuous Domain Adaptation through Graphs, CVPR 2019]
Transfer Learning

Transductive Unsupervised Domain Adaptation

Annotated Source data

Annotated Target data

multi-source Domain Generalization

predictive DA

zero-shot DG

single-source DG

zero-shot DA

semi-supervised DA

Transfer Learning

zero-shot DG

single-source DG
Transfer Learning
Transductive Unsupervised Domain Adaptation
Annotated Source data
Annotated Target data
multi-source Domain Generalization
predictive DA
zero-shot DA
zero-shot DG
single-source DG
online DA
semi-supervised DA
zero-shot DG
multi-source DG
online DA
[Kitting in the Wild through Online Domain Adaptation, ICRA 2018]
Transfer Learning

Transductive Unsupervised Domain Adaptation

Annotated Source data

Annotated Target data

zero-shot DA

multi-source Domain Generalization

predictive DA

single-source DG

online DA

domain generalization

zero-shot DG

semi-supervised DA

Transfer Learning

multi-source DG
Source-Free DA

[Universal Source-Free Domain Adaptation, CVPR 2020]
[Do We Really Need to Access the Source Data? Source Hypothesis Transfer for Unsupervised Domain Adaptation, ICML 2020]
Source-Free DA

Cross Entropy Loss

(1) Information Maximization Loss
   It makes the target outputs
   ● individually certain (as entropy minimization)
   ● globally diverse

(2) Target Clustering & Nearest Centroid Classifier
● the models on different nodes have different convergence rates
the models on different nodes have different convergence rates

increase the weight of those nodes whose gradients are beneficial to the target domain

measure how well the target features can be clustered (cluster gap statistics gain)
the models on different nodes have different convergence rates

- increase the weight of those nodes whose gradients are **beneficial** to the target domain

- measure how well the target features can be **clustered** (cluster gap statistics gain)
- the models on different nodes have different convergence rates
- increase the weight of those nodes whose gradients are **beneficial** to the target domain
- measure how well the target features can be **clustered** (cluster gap statistics gain)
Federated DA

- the models on different nodes have different convergence rates
- increase the weight of those nodes whose gradients are **beneficial** to the target domain
- measure how well the target features can be **clustered** (cluster gap statistics gain)
• the models on different nodes have different convergence rates
• increase the weight of those nodes whose gradients are **beneficial** to the target domain
• measure how well the target features can be **clustered** (cluster gap statistics gain)
Federated DA

- the models on different nodes have different convergence rates
- increase the weight of those nodes whose gradients are beneficial to the target domain
- measure how well the target features can be clustered (cluster gap statistics gain)
Federated DA

- the models on different nodes have different convergence rates
- increase the weight of those nodes whose gradients are **beneficial** to the target domain
- measure how well the target features can be **clustered** (cluster gap statistics gain)
[Universal Domain Adaptation, CVPR 2019]
Partial DA

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
Partial DA

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
Partial DA

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
Partial DA

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
[A Balanced and Uncertainty-aware Approach for Partial Domain Adaptation, ECCV 2020]
Partial DA

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
[A Balanced and Uncertainty-aware Approach for Partial Domain Adaptation, ECCV 2020]
Partial DA

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
[A Balanced and Uncertainty-aware Approach for Partial Domain Adaptation, ECCV 2020]
Entirely Digital.

Partial DA

- Entropy-weighted adversarial domain discriminator (e-DANN)

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
[A Balanced and Uncertainty-aware Approach for Partial Domain Adaptation, ECCV 2020]
Partial DA

- Entropy-weighted adversarial domain discriminator (e-DANN)
- Borrow a fraction of the source sample per class and consider them as target
Partial DA

- Entropy-weighted adversarial domain discriminator (e-DANN)
- Borrow a fraction of the source sample per class and consider them as target
- Encourage uniform and low prediction scores for incorrect classes of the source

[Partial Adversarial Domain Adaptation, ECCV 2018]
[Learning to transfer examples for partial domain adaptation, CVPR 2019]
Open-Set DA

[Open Set Domain Adaptation by Backpropagation, ECCV 2018]
[Separate to Adapt: Open Set Domain Adaptation via Progressive Separation, CVPR 2019]
[Open Set Domain Adaptation by Backpropagation, ECCV 2018]
[Separate to Adapt: Open Set Domain Adaptation via Progressive Separation, CVPR 2019]
Open-Set DA

[Open Set Domain Adaptation by Backpropagation, ECCV 2018]
[Separate to Adapt: Open Set Domain Adaptation via Progressive Separation, CVPR 2019]
Open-Set DA

[Open Set Domain Adaptation by Backpropagation, ECCV 2018]
[Separate to Adapt: Open Set Domain Adaptation via Progressive Separation, CVPR 2019]
Open-Set DA

[Open Set Domain Adaptation by Backpropagation, ECCV 2018]
[Separate to Adapt: Open Set Domain Adaptation via Progressive Separation, CVPR 2019]
Open-Set DA

[Open Set Domain Adaptation by Backpropagation, ECCV 2018]
[Separate to Adapt: Open Set Domain Adaptation via Progressive Separation, CVPR 2019]
Self-Supervised Learning
Solve Jigsaw Puzzles

- Decompose an image in patches
- Shuffle them = remove their spatial co-location
- Ask a network to recompose the original image

[Unsupervised learning of visual representations by solving jigsaw puzzles, CVPR 2016]
Recognize Image Orientation

- Rotate the image = remove the original orientation
- Ask a network to predict the rotation angle

Self-Supervised Learning

[Unsupervised Representation Learning by Predicting Image Rotations, ICLR 2018]
Self-Supervised Learning

Ex. 1: **Inpainting** (remove patch and then predict it)

\[
\begin{pmatrix}
\text{Image} \quad \text{Patch} \quad \text{Predicted Image}
\end{pmatrix}, \ldots
\]

Ex. 2: **Context** (given two patches, predict their spatial relation)

\[
\left\{ \begin{array}{c}
\{\text{Flamingo}, \text{“south east”}\}, \\
\{\text{Fur}, \text{“west”}\},
\end{array} \right\}, \ldots
\]

Ex. 3: **Colorization** (predict color given intensity)

\[
\begin{pmatrix}
\text{Color Image} \quad \text{Intensity Image} \quad \text{Predicted Color Image}
\end{pmatrix}, \ldots
\]

[Steering Self-Supervised Feature Learning Beyond Local Pixel Statistics, CVPR 2020]
[Unsupervised learning of visual representations by solving jigsaw puzzles, CVPR 2016]
[Unsupervised representation learning by predicting image rotations, ICLR 2018]
[Colorization as a proxy task for visual understanding, CVPR 2017]
[Self-supervised feature learning by learning to spot artifacts, CVPR 2018]
[Colorization as a Proxy Task for Visual Understanding, CVPR 2017]
[Self-Supervised Feature Learning by Learning to Spot Artifacts, CVPR 2018]
Self-Supervised Learning

[Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019]
[Self-Supervised Learning Across Domains, ArXiv 2020]
[On the Effectiveness of Image Rotation for Open Set Domain Adaptation, ECCV 2020]

[Tackling Partial Domain Adaptation with Self-Supervision, ICIAP 2019]
Self-Supervised Learning

- Zero-shot Domain Generalization (DG)
- Multi-source Domain Generalization
- Predictive Domain Adaptation (DA)
- Single-source Domain Generalization

Transductive Unsupervised Domain Adaptation

- Zero-shot Domain Adaptation
- Online Domain Adaptation (DA)
- Semi-supervised Domain Adaptation

One-Shot Unsupervised Cross-Domain Detection, ECCV 2020

Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019

Self-Supervised Learning Across Domains, ArXiv 2020

Source data

Target data

Transfer Learning

Source-free DA

Federated DA
Self-Supervision + Domain Generalization

\[
\arg\min_{\theta_f, \theta_c, \theta_p} \frac{1}{n^s} \sum_{i=1}^{n^s} \mathcal{L}_c(G_c(G_f(x_i^s)), y_i^s) + \alpha_s \frac{1}{K^s} \sum_{k=1}^{K^s} \mathcal{L}_p(G_p(G_f(z_k^s)), p_k^s)
\]

[Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019]
[Self-Supervised Learning Across Domains, ArXiv 2020]
Self-Supervision + Domain Generalization

- **[Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019]**
- **[Self-Supervised Learning Across Domains, ArXiv 2020]**
- **[Deeper, Broader and Artier Domain Generalization, ICCV 2017]**

### Table: PACS Results

<table>
<thead>
<tr>
<th>Method</th>
<th>art_paint.</th>
<th>cartoon</th>
<th>sketches</th>
<th>photo</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alexnet</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeepAll</td>
<td>63.30</td>
<td>63.13</td>
<td>54.07</td>
<td>87.70</td>
<td>67.05</td>
</tr>
<tr>
<td>TF</td>
<td>62.86</td>
<td>66.97</td>
<td>57.51</td>
<td>89.50</td>
<td>69.21</td>
</tr>
<tr>
<td>DeepAll</td>
<td>64.44</td>
<td>72.07</td>
<td>58.07</td>
<td>87.50</td>
<td>70.52</td>
</tr>
<tr>
<td>D-SAM</td>
<td>63.87</td>
<td>70.70</td>
<td>64.66</td>
<td>85.55</td>
<td>71.20</td>
</tr>
<tr>
<td>DeepAll</td>
<td>63.40</td>
<td>66.10</td>
<td>56.60</td>
<td>88.50</td>
<td>68.70</td>
</tr>
<tr>
<td>Epi-FCR</td>
<td>64.70</td>
<td>72.30</td>
<td>65.00</td>
<td>86.10</td>
<td>72.00</td>
</tr>
<tr>
<td>DeepAll</td>
<td>64.91</td>
<td>64.28</td>
<td>53.08</td>
<td>86.67</td>
<td>67.24</td>
</tr>
<tr>
<td>MLDG</td>
<td>66.23</td>
<td>66.88</td>
<td>58.96</td>
<td>88.00</td>
<td>70.01</td>
</tr>
<tr>
<td>DeepAll</td>
<td>67.21</td>
<td>66.12</td>
<td>55.32</td>
<td>88.47</td>
<td>69.28</td>
</tr>
<tr>
<td>MetaReg</td>
<td>69.82</td>
<td>70.35</td>
<td>59.26</td>
<td>91.07</td>
<td>72.62</td>
</tr>
<tr>
<td>DeepAll</td>
<td>63.30</td>
<td>63.10</td>
<td>54.00</td>
<td>87.70</td>
<td>67.03</td>
</tr>
<tr>
<td>PAR</td>
<td>68.70</td>
<td>70.50</td>
<td>64.60</td>
<td>90.40</td>
<td>73.54</td>
</tr>
<tr>
<td>DeepAll</td>
<td>68.09</td>
<td>70.23</td>
<td>61.80</td>
<td>88.86</td>
<td>72.25</td>
</tr>
<tr>
<td>MMLD</td>
<td>66.99</td>
<td>70.64</td>
<td>67.78</td>
<td>89.35</td>
<td>73.69</td>
</tr>
<tr>
<td>DeepAll</td>
<td>66.50</td>
<td>69.65</td>
<td>61.42</td>
<td>89.68</td>
<td>71.81±0.26</td>
</tr>
<tr>
<td><strong>Jigsaw</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotation</td>
<td>67.76</td>
<td>70.79</td>
<td>64.01</td>
<td>89.64</td>
<td>73.05±0.20</td>
</tr>
<tr>
<td>DeepAll</td>
<td>69.43</td>
<td>69.40</td>
<td>65.20</td>
<td>89.17</td>
<td>73.30±0.47</td>
</tr>
<tr>
<td><strong>Jigsaw+Rotation</strong></td>
<td>69.70</td>
<td>71.00</td>
<td>66.00</td>
<td>89.60</td>
<td>74.08±0.32</td>
</tr>
</tbody>
</table>
Self-Supervision + Domain Generalization

<table>
<thead>
<tr>
<th>PACS</th>
<th>art_paint.</th>
<th>cartoon</th>
<th>sketches</th>
<th>photo</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepAll</td>
<td>63.30</td>
<td>63.13</td>
<td>54.07</td>
<td>87.70</td>
<td>67.05</td>
</tr>
<tr>
<td>TF</td>
<td>62.86</td>
<td>66.97</td>
<td>57.51</td>
<td>89.50</td>
<td>69.21</td>
</tr>
<tr>
<td>DeepAll</td>
<td>64.44</td>
<td>72.07</td>
<td>58.07</td>
<td>87.50</td>
<td>70.52</td>
</tr>
<tr>
<td>D-SAM</td>
<td>63.87</td>
<td>70.71</td>
<td>64.66</td>
<td>85.55</td>
<td>71.20</td>
</tr>
<tr>
<td>DeepAll</td>
<td>63.40</td>
<td>66.10</td>
<td>56.60</td>
<td>88.50</td>
<td>68.70</td>
</tr>
<tr>
<td>Epi-FCR</td>
<td>64.70</td>
<td>72.30</td>
<td>65.00</td>
<td>86.10</td>
<td>72.00</td>
</tr>
<tr>
<td>DeepAll</td>
<td>64.91</td>
<td>64.28</td>
<td>53.08</td>
<td>86.67</td>
<td>67.24</td>
</tr>
<tr>
<td>MLDG</td>
<td>66.23</td>
<td>66.88</td>
<td>58.96</td>
<td>88.00</td>
<td>70.01</td>
</tr>
<tr>
<td>DeepAll</td>
<td>67.21</td>
<td>66.12</td>
<td>55.32</td>
<td>88.47</td>
<td>69.28</td>
</tr>
<tr>
<td>MetaReg</td>
<td>69.82</td>
<td>70.35</td>
<td>59.26</td>
<td>91.07</td>
<td>72.62</td>
</tr>
<tr>
<td>DeepAll</td>
<td>63.30</td>
<td>63.10</td>
<td>54.00</td>
<td>87.70</td>
<td>67.03</td>
</tr>
<tr>
<td>PAR</td>
<td>68.70</td>
<td>70.50</td>
<td>64.60</td>
<td>90.40</td>
<td>73.54</td>
</tr>
<tr>
<td>DeepAll</td>
<td>68.09</td>
<td>70.23</td>
<td>61.80</td>
<td>88.86</td>
<td>72.25</td>
</tr>
<tr>
<td>MMLD</td>
<td>66.99</td>
<td>70.64</td>
<td>67.78</td>
<td>89.35</td>
<td>73.69</td>
</tr>
<tr>
<td>DeepAll</td>
<td>66.50</td>
<td>69.65</td>
<td>61.42</td>
<td>89.68</td>
<td>71.81±0.26</td>
</tr>
</tbody>
</table>

**Jigsaw**

<table>
<thead>
<tr>
<th>PACS</th>
<th>Jigsaw</th>
<th>Rotation</th>
<th>Jigsaw+Rotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepAll</td>
<td>67.76</td>
<td>69.43</td>
<td>69.70</td>
</tr>
<tr>
<td>Jigsaw</td>
<td>70.79</td>
<td>69.40</td>
<td>71.00</td>
</tr>
<tr>
<td>Rotation</td>
<td>64.01</td>
<td>65.20</td>
<td>66.00</td>
</tr>
<tr>
<td>Jigsaw+Rotation</td>
<td>89.64</td>
<td>89.17</td>
<td>89.60</td>
</tr>
</tbody>
</table>

[Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019]
[Self-Supervised Learning Across Domains, ArXiv 2020]
[Deeper, Broader and Artier Domain Generalization, ICCV 2017]
Self-Supervision + Domain Generalization
Self-Supervision + Domain Generalization
Self-Supervision + Domain Generalization

\[
\arg \min_{\theta_f, \theta_c, \theta_p} \frac{1}{n_s} \sum_{i=1}^{n_s} \mathcal{L}_c(G_c(G_f(x_i^s)), y_i^s) + \alpha_s \frac{1}{K_s} \sum_{k=1}^{K_s} \mathcal{L}_p(G_p(G_f(z_k^s)); p_k^s)
\]

[Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019]
[Self-Supervised Learning Across Domains, ArXiv 2020]
Self-Supervision + Domain Adaptation

arg \min_{\theta_f, \theta_c, \theta_p} \frac{1}{n^s} \sum_{i=1}^{n^s} \mathcal{L}_c(G_c(G_f(x_i^s)), y_i^s) + \frac{\alpha_s}{K^s} \sum_{k=1}^{K^s} \mathcal{L}_p(G_p(G_f(z_k^s)), p_k^s) + \frac{\alpha_t}{K^t} \sum_{k=1}^{K^t} \mathcal{L}_p(G_p(G_f(z_k^t)), p_k^t).

[Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019]
[Self-Supervised Learning Across Domains, ArXiv 2020]
### Self-Supervision + Domain Adaptation

[Domain Generalization by Solving Jigsaw Puzzles, CVPR 2019]
[Self-Supervised Learning Across Domains, ArXiv 2020]
[Deeper, Broader and Artier Domain Generalization, ICCV 2017]

#### Table: PACS-DA

<table>
<thead>
<tr>
<th>Method</th>
<th>art_paint.</th>
<th>cartoon</th>
<th>sketches</th>
<th>photo</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resnet-18</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeepAll</td>
<td>74.70</td>
<td>72.40</td>
<td>60.10</td>
<td>92.90</td>
<td>75.03</td>
</tr>
<tr>
<td>Dial</td>
<td>87.30</td>
<td>85.50</td>
<td>66.80</td>
<td>97.00</td>
<td>84.15</td>
</tr>
<tr>
<td>DDiscovery</td>
<td>87.70</td>
<td>86.90</td>
<td>69.60</td>
<td>97.00</td>
<td>85.30</td>
</tr>
<tr>
<td><strong>Resnet-18</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeepAll</td>
<td>76.17</td>
<td>73.58</td>
<td>55.65</td>
<td>96.07</td>
<td>75.37±0.42</td>
</tr>
<tr>
<td>HAFN</td>
<td>84.95</td>
<td>79.64</td>
<td>64.24</td>
<td>97.70</td>
<td>81.63±0.50</td>
</tr>
<tr>
<td>SAFN</td>
<td>86.78</td>
<td>82.72</td>
<td>60.26</td>
<td>98.26</td>
<td>82.01±0.32</td>
</tr>
<tr>
<td>SAFN+ENT</td>
<td>89.22</td>
<td>87.39</td>
<td>60.02</td>
<td>98.14</td>
<td>83.69±0.17</td>
</tr>
<tr>
<td><strong>Resnet-18</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeepAll</td>
<td>77.83</td>
<td>74.26</td>
<td>65.81</td>
<td>95.71</td>
<td>78.40±0.28</td>
</tr>
<tr>
<td><strong>Jigsaw</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s = \alpha_t = 0.7$</td>
<td>84.49</td>
<td>82.07</td>
<td>79.86</td>
<td>97.98</td>
<td>86.10±0.26</td>
</tr>
<tr>
<td><strong>Rotation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s = \alpha_t = 0.8$</td>
<td>89.97</td>
<td>82.60</td>
<td>82.00</td>
<td>98.07</td>
<td>88.16±0.51</td>
</tr>
<tr>
<td><strong>Jigsaw+Rotation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>90.87</td>
<td>82.77</td>
<td>83.80</td>
<td>98.37</td>
<td>88.95±0.38</td>
</tr>
</tbody>
</table>
Self-Supervision + 3D Domain Adaptation

[Self-Supervision for 3D Real-World Challenges, TASK-CV Workshop ECCV 2020]
Self-Supervision + Open-Set DA

Stage I - known/unknown separation

Stage II - domain alignment

[On the Effectiveness of Image Rotation for Open Set Domain Adaptation, ECCV 2020]
Stage I - Known/Unknown Separation
Stage I - Known/Unknown Separation

\[ \mathcal{L}_{C_1} = - \sum_{j \in \mathcal{D}_s} y_j^s \cdot \log(\hat{y}_j^s) \]

\[ \mathcal{L}_{R_1} = \sum_{j \in \tilde{\mathcal{D}}_s} -\lambda_{1,1} z_j^s \cdot \log(\hat{z}_j^s) \]

\[ + \lambda_{1,2} ||v_j^s - \gamma(z_j^s)||_2^2 \]
Stage I - Known/Unknown Separation

Multi-Rotation Classifier

\((4 \times |C_s|)\) class labels
Stage I - Known/Unknown Separation

$D_s$ → $\triangledown$ → $\mathcal{E}$ → $C_1$ → $R_1$

Relative Multi-Rotation Classifier

$(4 \times |\mathcal{C}_s|)$ class labels

$\Delta$ Known  ○ Source
Normality Score

Stage I - known/unknown separation

$D_s$ $\rightarrow$ $\triangle$ $\rightarrow$ $\rightarrow$ $\rightarrow$ $E$ $\rightarrow$ $\rightarrow$ $C_1$ $\rightarrow$ $\rightarrow$ $R_1$ $\rightarrow$ $\rightarrow$ $D_t$ $\rightarrow$ $\rightarrow$ $D_{unk}$ $\rightarrow$ $\rightarrow$ $D_{tnw}$ $\rightarrow$ $\rightarrow$ $D_{tnw}$

$\triangle$ Known $\bigcirc$ Source
$\bigcirc$ Unknown $\triangle$ Target
Known Classes = \{ \text{A, B, C} \}

For one sample \( T \) (The evaluation is done for each Target sample)

\[
\text{Score } T = [A^0, A^{90}, A^{180}, A^{270}, B^0, B^{90}, B^{180}, B^{270}, C^0, C^{90}, C^{180}, C^{270}]
\]
Known Classes = \{ A, B, C \}

Score T = \[ A^0 + A^90 + A^{180} + A^{270} + B^0 + B^90 + B^{180} + B^{270} + C^0 + C^90 + C^{180} + C^{270} \]

Entropy Score T = mean(Entropy(\alpha) + Entropy(\beta) + Entropy(\gamma) + Entropy(\delta))
Known Classes = \{A, B, C\}

For one sample T (The evaluation is done for each Target sample)

Score T = \[A^0 + A^{90} + A^{180} + A^{270} + B^0 + B^{90} + B^{180} + B^{270} + C^0 + C^{90} + C^{180} + C^{270}\]

Entropy Score T = mean(Entropy(\alpha) + Entropy(\beta) + Entropy(\gamma) + Entropy(\delta))

Normality Score T = max\{max_{\{A, B, C\}}(Score T), (1 - Entropy Score T)\}
The **Normality Score** gives the probability that each Target sample is from a **Known Class**.

\[
\begin{align*}
\begin{cases}
\mathbf{x}^t \in D^{\text{known}}_t & \text{if } N(\mathbf{x}^t) > \bar{N} \\
\mathbf{x}^t \in D^{\text{unk}}_t & \text{if } N(\mathbf{x}^t) < \bar{N}
\end{cases}
\end{align*}
\]

\[
\bar{N} = \frac{1}{N_t} \sum_{j=1}^{N_t} N_j
\]

The Threshold IS NOT an hyperparameter.
Stage II - Domain Alignment

Stage I - known/unknown separation

Stage II - domain alignment

[On the Effectiveness of Image Rotation for Open Set Domain Adaptation, ECCV 2020]
Stage II - Domain Alignment

Relative Rotation Classifier for the Domain Alignment (4 classes)

$$\mathcal{L}_{R_2} = -\lambda_{2,2} \sum_{j \in D_t^{knw}} q_j \cdot \log(\hat{q}_j)$$
Harmonic mean
of OS* and UNK

\[ OS = \frac{|C_s|}{|C_s| + 1} \times OS^* + \frac{1}{|C_s| + 1} \times UNK \]

New Open-Set DA Metrics

Number of Known Classes

Measure of the overall performance

Class accuracy over the Known Classes

Class accuracy over the Unknown Class

\[ HOS = 2 \frac{OS^* \times UNK}{OS^* + UNK} \]

It provides a high score only if the algorithm performs well both on known and on unknown samples, independently of $|C_s|$. 
Results on Office-Home Dataset

25 Known Classes
40 Unknown Classes

### Office-Home

<table>
<thead>
<tr>
<th>STA$_{sum}$ [24]</th>
<th>STA$_{max}$</th>
<th>OSBP [33]</th>
<th>UAN [45]</th>
<th>ROS</th>
<th>R$\rightarrow$ Ar</th>
<th>R$\rightarrow$ Pr</th>
<th>R$\rightarrow$ Cl</th>
<th>Cl$\rightarrow$ Rw</th>
<th>Cl$\rightarrow$ Ar</th>
<th>Cl$\rightarrow$ Pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr $\rightarrow$ Rw</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
<td>OS* UNK HOS</td>
</tr>
<tr>
<td>78.1</td>
<td>63.3</td>
<td>69.7</td>
<td>44.7</td>
<td>71.5</td>
<td>55.0</td>
<td>55.4</td>
<td>73.7</td>
<td>63.1</td>
<td>68.7</td>
<td>59.7</td>
</tr>
<tr>
<td>76.2</td>
<td>64.3</td>
<td>69.5</td>
<td>44.2</td>
<td>67.1</td>
<td>53.2</td>
<td>54.2</td>
<td>72.4</td>
<td>61.9</td>
<td>68.0</td>
<td>48.4</td>
</tr>
<tr>
<td>76.2</td>
<td>71.7</td>
<td>73.9</td>
<td>44.5</td>
<td>66.3</td>
<td>53.2</td>
<td>59.1</td>
<td>68.1</td>
<td>63.2</td>
<td>71.8</td>
<td>59.8</td>
</tr>
<tr>
<td>84.0</td>
<td>0.1</td>
<td>0.2</td>
<td>59.1</td>
<td>0.0</td>
<td>0.0</td>
<td>73.7</td>
<td>0.0</td>
<td>0.0</td>
<td>81.1</td>
<td>0.0</td>
</tr>
<tr>
<td>70.8</td>
<td>78.4</td>
<td>74.4</td>
<td>46.5</td>
<td>71.2</td>
<td>56.3</td>
<td>57.3</td>
<td>64.3</td>
<td>60.6</td>
<td>68.4</td>
<td>70.3</td>
</tr>
</tbody>
</table>

[Deep Hashing Network for Unsupervised Domain Adaptation, CVPR 2017]
...with less known classes
Self-Supervision +
One-Shot Unsupervised Cross-Domain Detection
Cross-Domain Object Detection

Wait to collect feeds

Source - VOC

Diversify

Match

 Wait to collect feeds

[A domain adaptive representation learning paradigm for object detection, CVPR 2019]
One-Shot Unsupervised Cross-Domain Detection

Pretrained model

Source - VOC

Per-sample unsupervised adaptation
One-Shot Unsupervised Cross-Domain Detection

-One-Shot Unsupervised Cross-Domain Detection, ECCV2020-

OSHOT

1. start from FasterCNN

[Towards real-time object detection with region proposal networks, NIPS 2015]
One-Shot Unsupervised Cross-Domain Detection

[One-Shot Unsupervised Cross-Domain Detection, ECCV2020]

OSHOT

1. start from FasterCNN
2. add a self-supervised task: rotation recognition

argmin_{\theta_f, \theta_d, \theta_r} \sum_{i=1}^{N} L_d(G_d(G_f(x_i^s)), y_i^s) + \lambda \sum_{j=1}^{M} L_r(G_r(G_f(R(x_i^s)_j)), q_j^s)

[Towards real-time object detection with region proposal networks, NIPS 2015]
One-Shot Unsupervised Cross-Domain Detection

[One-Shot Unsupervised Cross-Domain Detection, ECCV2020]

**OSHOT**

1. start from FasterCNN
2. add a self-supervised task: rotation recognition
3. no need to further access the source

[Towards real-time object detection with region proposal networks, NIPS 2015]
One-Shot Unsupervised Cross-Domain Detection

OSHOT

1. start from FasterCNN
2. add a self-supervised task: rotation recognition
3. no need to further access the source
4. sample-guided self-supervised fine-tuning

[Towards real-time object detection with region proposal networks, NIPS 2015]
One-Shot Unsupervised Cross-Domain Detection

Towards real-time object detection with region proposal networks, NIPS 2015

\[ y^t = G_d(G_f(x^t)) \]
### Adapting to a stream of Social Images

[One-Shot Unsupervised Cross-Domain Detection, ECCV2020]

<table>
<thead>
<tr>
<th>Method</th>
<th>Person</th>
<th>Bicycle</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRCNN</td>
<td>67.7</td>
<td>56.6</td>
<td>62.1</td>
</tr>
<tr>
<td><strong>OSHOT</strong> $\left(\gamma = 0\right)$</td>
<td>72.1</td>
<td>52.8</td>
<td>62.4</td>
</tr>
<tr>
<td><strong>OSHOT</strong> $\left(\gamma = 30\right)$</td>
<td>69.4</td>
<td>59.4</td>
<td>64.4</td>
</tr>
<tr>
<td>DivMatch [28]</td>
<td>63.7</td>
<td>51.7</td>
<td>57.7</td>
</tr>
<tr>
<td>SW [42]</td>
<td>63.2</td>
<td>44.3</td>
<td>53.7</td>
</tr>
</tbody>
</table>

[28] [Diversify and match: A domain adaptive representation learning paradigm for object detection, CVPR 2019]
[42] [Saito et al: Strong-weak distribution alignment for adaptive object detection, CVPR 2019]
From Real to Artistic Images

[One-Shot Unsupervised Cross-Domain Detection, ECCV2020]

(a) VOC $\rightarrow$ Clipart

<table>
<thead>
<tr>
<th>Method</th>
<th>aero</th>
<th>bike</th>
<th>bird</th>
<th>boat</th>
<th>bottle</th>
<th>bus</th>
<th>car</th>
<th>cat</th>
<th>chair</th>
<th>cow</th>
<th>table</th>
<th>dog</th>
<th>horse</th>
<th>mbike</th>
<th>person</th>
<th>plant</th>
<th>sheep</th>
<th>sofa</th>
<th>train</th>
<th>tv</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRCNN</td>
<td>18.5</td>
<td>43.3</td>
<td>20.4</td>
<td>13.3</td>
<td>21.0</td>
<td>47.8</td>
<td>29.0</td>
<td>16.9</td>
<td>28.8</td>
<td>12.5</td>
<td>10.5</td>
<td>17.1</td>
<td>23.8</td>
<td>40.6</td>
<td>34.9</td>
<td>34.7</td>
<td>9.1</td>
<td>18.3</td>
<td>40.2</td>
<td>38.0</td>
<td>26.4</td>
</tr>
<tr>
<td>OSHT ($\gamma = 0$)</td>
<td>23.1</td>
<td>55.3</td>
<td>22.7</td>
<td>21.4</td>
<td>26.8</td>
<td>53.3</td>
<td>28.9</td>
<td>4.6</td>
<td>31.4</td>
<td>9.2</td>
<td>27.8</td>
<td>9.6</td>
<td>30.9</td>
<td>47.0</td>
<td>38.2</td>
<td>35.2</td>
<td>11.1</td>
<td>20.4</td>
<td>36.0</td>
<td>33.6</td>
<td>28.3</td>
</tr>
<tr>
<td>OSHT ($\gamma = 10$)</td>
<td>25.4</td>
<td>61.6</td>
<td>23.8</td>
<td>21.1</td>
<td>31.3</td>
<td>55.1</td>
<td>31.6</td>
<td>5.3</td>
<td>34.0</td>
<td>10.1</td>
<td>28.8</td>
<td>7.3</td>
<td>33.1</td>
<td>59.9</td>
<td>44.2</td>
<td>38.8</td>
<td>15.9</td>
<td>19.1</td>
<td>39.5</td>
<td>33.9</td>
<td>31.0</td>
</tr>
<tr>
<td>OSHT ($\gamma = 30$)</td>
<td>25.4</td>
<td>56.0</td>
<td>24.7</td>
<td>25.3</td>
<td>36.7</td>
<td>58.0</td>
<td>34.4</td>
<td>5.9</td>
<td>34.9</td>
<td>10.3</td>
<td>29.2</td>
<td>11.8</td>
<td>46.9</td>
<td>70.9</td>
<td>52.9</td>
<td>41.5</td>
<td>21.1</td>
<td>21.0</td>
<td>38.5</td>
<td>31.8</td>
<td>33.9</td>
</tr>
<tr>
<td>DivMatch [28]</td>
<td>19.5</td>
<td>57.2</td>
<td>17.0</td>
<td>23.8</td>
<td>14.4</td>
<td>25.4</td>
<td>29.4</td>
<td>2.7</td>
<td>35.0</td>
<td>8.4</td>
<td>22.9</td>
<td>14.2</td>
<td>30.0</td>
<td>55.6</td>
<td>50.8</td>
<td>30.2</td>
<td>1.9</td>
<td>12.3</td>
<td>37.8</td>
<td>37.2</td>
<td>26.3</td>
</tr>
<tr>
<td>SW [42]</td>
<td>21.5</td>
<td>39.9</td>
<td>21.7</td>
<td>20.5</td>
<td>32.7</td>
<td>34.1</td>
<td>25.1</td>
<td>8.5</td>
<td>33.2</td>
<td>10.9</td>
<td>15.2</td>
<td>3.4</td>
<td>32.2</td>
<td>56.9</td>
<td>46.5</td>
<td>35.4</td>
<td>14.7</td>
<td>15.2</td>
<td>29.2</td>
<td>32.0</td>
<td>26.4</td>
</tr>
</tbody>
</table>

(b) VOC $\rightarrow$ Comic

<table>
<thead>
<tr>
<th>Method</th>
<th>bike</th>
<th>bird</th>
<th>car</th>
<th>cat</th>
<th>dog</th>
<th>person</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRCNN</td>
<td>25.2</td>
<td>10.0</td>
<td>21.1</td>
<td>14.1</td>
<td>11.0</td>
<td>27.1</td>
<td>18.1</td>
</tr>
<tr>
<td>OSHT ($\gamma = 0$)</td>
<td>26.9</td>
<td>11.6</td>
<td>22.7</td>
<td>9.1</td>
<td>14.2</td>
<td>28.3</td>
<td>18.8</td>
</tr>
<tr>
<td>OSHT ($\gamma = 10$)</td>
<td>35.5</td>
<td>11.7</td>
<td>25.1</td>
<td>9.1</td>
<td>15.8</td>
<td>34.5</td>
<td>22.0</td>
</tr>
<tr>
<td>OSHT ($\gamma = 30$)</td>
<td>35.2</td>
<td>14.4</td>
<td>30.0</td>
<td>14.8</td>
<td>20.0</td>
<td>46.7</td>
<td>26.9</td>
</tr>
<tr>
<td>DivMatch [28]</td>
<td>27.1</td>
<td>12.3</td>
<td>26.2</td>
<td>11.5</td>
<td>13.8</td>
<td>34.0</td>
<td>20.8</td>
</tr>
<tr>
<td>SW [42]</td>
<td>21.2</td>
<td>14.8</td>
<td>18.7</td>
<td>12.4</td>
<td>14.9</td>
<td>43.9</td>
<td>21.0</td>
</tr>
</tbody>
</table>

(c) VOC $\rightarrow$ Watercolor

<table>
<thead>
<tr>
<th>Method</th>
<th>bike</th>
<th>bird</th>
<th>car</th>
<th>cat</th>
<th>dog</th>
<th>person</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRCNN</td>
<td>62.5</td>
<td>39.7</td>
<td>43.4</td>
<td>31.9</td>
<td>26.7</td>
<td>52.4</td>
<td>42.8</td>
</tr>
<tr>
<td>OSHT ($\gamma = 0$)</td>
<td>70.2</td>
<td>46.7</td>
<td>45.5</td>
<td>31.2</td>
<td>27.2</td>
<td>55.7</td>
<td>46.1</td>
</tr>
<tr>
<td>OSHT ($\gamma = 10$)</td>
<td>70.2</td>
<td>46.7</td>
<td>48.1</td>
<td>30.9</td>
<td>32.3</td>
<td>59.9</td>
<td>48.0</td>
</tr>
<tr>
<td>OSHT ($\gamma = 30$)</td>
<td>77.1</td>
<td>44.7</td>
<td>52.4</td>
<td>37.3</td>
<td>37.0</td>
<td>63.3</td>
<td>52.0</td>
</tr>
<tr>
<td>DivMatch [28]</td>
<td>64.6</td>
<td>44.1</td>
<td>44.6</td>
<td>34.1</td>
<td>24.9</td>
<td>60.0</td>
<td>45.4</td>
</tr>
<tr>
<td>SW [42]</td>
<td>66.3</td>
<td>41.1</td>
<td>41.1</td>
<td>30.5</td>
<td>20.5</td>
<td>52.3</td>
<td>42.0</td>
</tr>
</tbody>
</table>

[28] [Diversify and match: A domain adaptive representation learning paradigm for object detection, CVPR 2019]
[42] [Saito et al: Strong-weak distribution alignment for adaptive object detection, CVPR 2019]
To Summarize...

zero-shot DG
multi-source Domain Generalization
predictive DA

evaluate DA

online DA

Transductive Unsupervised Domain Adaptation

zero-shot DA

semi-supervised DA

source-free DA
federated DA

one-shot unsupervised

Annotated Source data

Annotated Target data

Transfer Learning
To Summarize...

- **Self-Supervised Learning**
  - Single-source DG
  - Multi-source Domain Generalization
  - Predictive DA
  - Online DA
  - ZS-DG
  - ZS-DA

- **Transductive, Unsupervised Domain Adaptation**
  - Partial DA
  - Universal DA

- **Open-Set DA**

- **Unsupervised Domain Adaptation**

- **S classes = T classes**

- **# S classes > # T classes**
Thanks for your attention

Domain Adaptation for Visual Applications
Part 3: Beyond Classical Domain Adaptation

Tatiana Tommasi
tatiana.tommasi@polito.it