Transfer Learning and Domain Adaptation
Biggest Criticism of Computer Vision

Works on constructed datasets, but not in the real world...

... and that's also true for deep learning
### E.g., Multi-Dataset Efforts

<table>
<thead>
<tr>
<th></th>
<th>Stereo</th>
<th>MVS</th>
<th>Flow</th>
<th>Depth</th>
<th>Semantic</th>
<th>Instance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middlebury</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KITTI</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MPI Sintel</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETH3D</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HD1K</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ScanNet</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cityscapes</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>WildDash</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Robust Vision Challenge: CVPR’18 [Geiger/Niessner/Pollefeys/Rother et al.]
Transfer Learning & Domain Adaptation

- Task
  - Image Classification
  - Image Segmentation
  - Object Instance Segmentation
  - ...

- Domain
  - Real data
    - Real != real: webcam model 1 vs webcam model 2; day vs night
  - Synthetic data
    - E.g., rasterization vs
  - ...

Prof. Leal-Taixé and Prof. Niessner
Transfer Learning & Domain Adaptation

- Same Source and Target Marginal Distributions on X
- Yes
  - Same Tasks on Source and Target Domains
    - Yes
      - "Usual" Learning Setting
    - No
      - Inductive Transfer Learning
- No
  - Same Tasks on Source and Target Domains
    - Yes
      - Transductive Transfer Learning
    - No
      - Unsupervised Transfer Learning

Domain Adaptation

Source: wikipedia
Transfer Learning

Same domain, different task

• Pre-trained Image Net (visual domain of real images)
  – Train on image classification

• Fine-tune on new task
  – E.g., semantic image segmentation
  – > keep 'backbone the same, fine-tune 'head' layers
  – > assumption: visual features generalize within domain
Transfer Learning

Same task, different domain

• Pre-trained Image Net (visual domain of real images)
  – Train on image classification

• Fine-tune on new task
  – Now need to train *entire* network, cuz input features will be different
  – Training only a few layers at the end is less likely to fundamentally solve it
Fine Tuning

• How much labeled data in the target domain?
  – Zero-shot learning
  – One-shot learning
  – Few-shot learning

• Just ‘throwing in as much data as we can’ seems somewhat unsatisfactory…
Domain Adaption
Applications to different types of domain shift

- From dataset to dataset
- From RGB to depth
- From simulated to real control
- From CAD models to real images

Slide Credit: Kate Saenko
Adversarial domain adaptation

Source Data + Labels

Unlabeled Target Data

Classifier

classification loss

Slide Credit: Kate Saenko
Adversarial domain adaptation

Slide Credit: Kate Saenko
Adversarial domain adaptation
Adversarial domain adaptation

**Source Data + Labels**
- backpack
- chair
- bike

**Unlabeled Target Data**
- ?

**Encoder**

**Discriminator**

**Classifier**

- classification loss

**Adversarial loss**

The encoder can be shared between the source and target domains.

Slide Credit: Kate Saenko
Results on Cityscapes to SF adaptation

Before domain confusion

After domain confusion

Cycle-Consistent Adversarial Domain Adaptation

Reconstructed Source Image

Cycle loss

Source Image

$G_{T\rightarrow S}$

$G_{S\rightarrow T}$

Source Image Stylized as Target

Target Image

$D_T$

$D_{feat}$

$f_T$

GAN loss

GAN loss

CyCADA [Hoffman et al. 2018]
CyCADA [Hoffman et al. 2018]
Exam

• Slides provide additional references (use them)

• Look up the important papers that we discussed

• Understanding of
  – high-level concepts
  – underlying math
  – architecture design
Administrative

• Deadline for final projects
  – Wed Feb 6th, 11:59pm
  – Submission via moodle
  – Submission must contain
    • Code (results must be replicable)
    • 2-3 pages of final report (at most 1 page of text, rest results; i.e., images and tables)
Administrative

• Poster presentation
  – Friday Feb 8\textsuperscript{th}, 1pm-3pm
  – Location:
    • Magistrale (preliminary – will update if it changes)
    • In the area next to the back entrance (parking lot direction)
  – Poster stands will be provided
  – You need to print posters yourself (poster@in.tum.de)
  – Hang posters 15 mins before presentation session starts
Guest Speakers

• Oriol Vinyals:
  – https://ai.google/research/people/OriolVinyals
  – Time: January 31st, 6pm – 8pm
  – Location: HS-1 (CS building – the big one)
Next Lectures

This was the last lecture 😊