Learning and forgetting in image classification and generation

Luis Herranz
Learning and machine perception (LAMP) group
Computer Vision Center (Barcelona)
About me

ETS Ing. Telecomunicación
Universidad Politécnica de Madrid (Spain)

(Ph.D) Escuela Politécnica Superior
Universidad Autónoma de Madrid (Spain)

Mitsubishi Electric R&D (UK)

Institute of Computing Technology,
Chinese Academy of Sciences (China)

Computer Vision Center,
Universitat Autònoma de Barcelona (Spain)

Research interests:
- Multimedia
- Computer vision
- Deep learning
- Multimodal representations
- Lifelong learning
Computer Vision Center (UAB campus)

Only Center in Europe fully devoted to Computer Vision

23 Years
+130 Staff
+20 Nationalities

M€2,3 Income / year
8 PhD thesis /year
+100 Intl publications / year
Learning and Machine Perception (LAMP) group

Senior PhDs

Joost van de Weijer
Leader

Postdocs

PhD students
Outline

• Introduction
• Transferring GANs (ECCV 2018)
• Rotated elastic weight consolidation (ICPR 2018)
• Memory Replay GANs (NIPS 2018)
• Mix and match networks (CVPR 2018)
Outline

• Introduction
• Transferring GANs (ECCV 2018)
• Rotated elastic weight consolidation (ICPR 2018)
• Memory Replay GANs (NIPS 2018)
• Mix and match networks (CVPR 2018)
Transfer learning and lifelong learning

Forgets source task, i.e. catastrophic forgetting (who cares?)

Forgets task 1 (big deal!!)

Transfer+adaptation

Lifelong learning

Discriminative models

Source task

Target task

Task 1

Task 2

Rotated elastic weight consolidation (ICPR 2018)

Liu et al., Rotate your Networks: Better Weight Consolidation and Less Catastrophic Forgetting, ICPR 2018
Transfer learning and lifelong learning (now with GANs for image generation)

Transfer+adaptation (generative)

Source domain

Target domain

Transferring GANs (ECCV 2018)

Lifelong learning (generative)

Task 1

Task 2

Memory Replay GANs (NIPS 2018)
Transfer learning and zero-shot learning

Source task

Target task

Zero-shot learning

Zero-shot classifier

Polar bear: white, water, eats fish

Zebra: black, white, stripes

Task 1

Task 2
Zero-pair image-to-image translations

Only these translations are trained (seen)

Train

$x^{(5)} = g_5(f_1(x^{(1)}))$

$x^{(1)} = g_1(f_5(x^{(5)}))$

Evaluate on these unseen translations (no training pairs)

Test

Mix and match networks (CVPR 2018)
Outline

• Introduction
• Transferring GANs (ECCV 2018)
• Rotated elastic weight consolidation (ICPR 2018)
• Memory Replay GANs (NIPS 2018)
• Mix and match networks (CVPR 2018)
Generative models: networks that imagine

Different approaches
- Density estimation
- Variational autoencoders
- Autoregressive models

Training data
(e.g. 64x64x3≈12K dims)

Sampling

Learning

- Generative adversarial networks (GANs)
Generative Adversarial Networks (GANs)

Goodfellow et al., “Generative Adversarial Networks”, NIPS 2014
Figure from https://deeplearning4j.org/generative-adversarial-network

Classify fake images vs real images

Generate fake samples to fool the discriminator

Goodfellow et al., “Generative Adversarial Networks”, NIPS 2014
Figure from https://deeplearning4j.org/generative-adversarial-network
Generative Adversarial Networks

Wasserstein GAN (WGAN-GP)

Progressive growing of GANs

and many more...
Transferring GAN representations

Source domain
Trainign data (ImageNet)

Target domain
Training data (LSUN bedrooms)

Wang et al., “Transfering GANs: generating images from limited data”, ECCV 2018
Transfer configuration

Training data

( LSUN bedrooms )

Target domain

Generator

z →

Discriminator

real/fake?

Training from scratch
- Discr: from scratch
- Gen: from scratch
Transfer configuration

Source domain
Training data
(ImageNet)

Discriminator
real/fake?

Target domain
Training data
(LSUN bedrooms)

Generator
$z$

Transfer only discr.
- Discr: pretrained
- Gen: from scratch
Transfer configuration

Source domain
Training data (ImageNet)

Generator

Discriminator

Target domain

Training data (LSUN bedrooms)

Generator

Discriminator

Transfer only gen.
- Discr: from scratch
- Gen: pretrained
Transfer configuration

Source domain
Training data
(ImageNet)

Target domain
Training data
(LSUN bedrooms)

Generator

Discriminator
real/fake?

Transfer both
- Discr: pretrained
- Gen: pretrained
Transfer configuration

- **What should I transfer?**
  - Experiment: ImageNet to Bedrooms (100K images)

<table>
<thead>
<tr>
<th>Generator</th>
<th>Scratch</th>
<th>Pretrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discriminator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FID ($\mathcal{X}<em>{data}^{tgt}, \mathcal{X}</em>{gen}^{tgt}$)</td>
<td>32.87</td>
<td>30.57</td>
</tr>
<tr>
<td>IW ($\mathcal{X}<em>{val}^{tgt}, \mathcal{X}</em>{gen}^{tgt}$)</td>
<td>-4.27</td>
<td>-4.02</td>
</tr>
</tbody>
</table>

Lower better
Higher better

- Training is faster and images have better quality
  - Especially when data is limited
Learning with limited data

<table>
<thead>
<tr>
<th>Target dataset: Bedroom</th>
<th>From scratch</th>
<th>Pretrained (ImageNet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000 images</td>
<td><img src="image1.png" alt="Images from scratch" /></td>
<td><img src="image2.png" alt="Images pretrained" /></td>
</tr>
<tr>
<td>10000 images</td>
<td><img src="image3.png" alt="Images from scratch" /></td>
<td><img src="image4.png" alt="Images pretrained" /></td>
</tr>
<tr>
<td>100000 images</td>
<td><img src="image5.png" alt="Images from scratch" /></td>
<td><img src="image6.png" alt="Images pretrained" /></td>
</tr>
</tbody>
</table>

Wang et al., “Transferring GANs: generating images from limited data”, ECCV 2018
Pretrain model selection

Wang et al., “Transferring GANs: generating images from limited data”, ECCV 2018
Good/bad source datasets

- **Source**
  - Places (205 classes ~10K img/class)
  - Bedrooms (1 class 3M images)

- **Target**
  - Kitchen (50K)

Generative: very dense, diversity not so important (e.g. LSUN Bedrooms)

Discriminative: very diverse, medium density (e.g. ImageNet, Places)
Outline

• Introduction
• Transferring GANs (ECCV 2018)
• Rotated elastic weight consolidation (ICPR 2018)
• Memory Replay GANs (NIPS 2018)
• Mix and match networks (CVPR 2018)
When are neural networks good?
Sequential learning

Catastrophic forgetting

Task 1

Task 2

Task 3
Catastrophic interference and forgetting

\[ \mathcal{L}(\theta) = \mathcal{L}_B(\theta) \]

\[ \mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \frac{\lambda}{2} (\theta - \theta^*_A)^T F_A (\theta - \theta^*_A) \]

Elastic weight consolidation (EWC)

Task B

Kirkpatric et al., Overcoming catastrophic forgetting in neural networks, PNAS, 2017
Rotated elastic weight consolidation

Elastic weight consolidation in practice (with diagonal approx. of $F_A$)

Rotated elastic weight consolidation (R-EWC)

$$L(\theta) = L_B(\theta) + \frac{\lambda}{2} \sum_i (F_A)_i (\theta_i - (\theta^*_A)_i)^2$$

Size of the diagonal of $F_A$ is $\#\theta$

Liu, Masana et al., Rotate your Networks: Better Weight Consolidation and Less Catastrophic Forgetting, ICPR 2018
Rotating fully connected layers

Liu, Masana et al., Rotate your Networks: Better Weight Consolidation and Less Catastrophic Forgetting, ICPR 2018
Computing the rotations

\[
F_W = \mathbb{E}_{p \sim \pi} \left[ (\frac{\partial L}{\partial y}) \right] xx^T \left( \frac{\partial L}{\partial y} \right)^T \right]
\]

\[
F_W \approx \mathbb{E}_{x \sim \pi} \left[ (\frac{\partial L}{\partial y}) \left( \frac{\partial L}{\partial y} \right)^T \right] \mathbb{E}_{x \sim \pi} \left[ xx^T \right]
\]

Assuming \( x \) and \( \delta L/\delta y \) independent

Using SVD

Liu, Masana et al., Rotate your Networks: Better Weight Consolidation and Less Catastrophic Forgetting, ICPR 2018
Rotating convolutional layers

Liu et al., Rotate your Networks: Better Weight Consolidation and Less Catastrophic Forgetting, ICPR 2018
Fisher Information matrix

No rotation (i.e. EWC)

Energy in the diagonal: 40%

After rotation (i.e. R-EWC)

Energy in the diagonal: 74%

Liu, Masana et al., Rotate your Networks: Better Weight Consolidation and Less Catastrophic Forgetting, ICPR 2018
Experimental results (2 tasks)

- **MNIST dataset. Two tasks: 0-4 and 5-9**

<table>
<thead>
<tr>
<th></th>
<th>$\lambda = 1$</th>
<th>$\lambda = 10$</th>
<th>$\lambda = 100$</th>
<th>$\lambda = 1000$</th>
<th>$\lambda = 10000$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Task 1</td>
<td>Task 2</td>
<td>Task 1</td>
<td>Task 2</td>
<td>Task 1</td>
</tr>
<tr>
<td>FT</td>
<td>6.1</td>
<td>97.6</td>
<td>6.1</td>
<td>97.6</td>
<td>6.1</td>
</tr>
<tr>
<td>EWC [5]</td>
<td>66.8</td>
<td>90.9</td>
<td>75.3</td>
<td>95.6</td>
<td><strong>85.8</strong></td>
</tr>
<tr>
<td>R-EWC - conv only</td>
<td>62.7</td>
<td>89.2</td>
<td>67.5</td>
<td>96.1</td>
<td>80.4</td>
</tr>
<tr>
<td>R-EWC - fc only</td>
<td>78.9</td>
<td>95.3</td>
<td>79.0</td>
<td>95.8</td>
<td>87.4</td>
</tr>
<tr>
<td>R-EWC - all</td>
<td>77.2</td>
<td>96.7</td>
<td><strong>91.7</strong></td>
<td><strong>91.2</strong></td>
<td>86.9</td>
</tr>
<tr>
<td>R-EWC - all no last</td>
<td>71.5</td>
<td>91.8</td>
<td>84.9</td>
<td>97.0</td>
<td><strong>91.6</strong></td>
</tr>
</tbody>
</table>

- **Several datasets**

<table>
<thead>
<tr>
<th></th>
<th>EWC [5] (T1 / T2)</th>
<th>R-EWC (T1 / T2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNIST</td>
<td>89.3 (85.8 / 92.8)</td>
<td><strong>93.1</strong> (91.6 / 94.5)</td>
</tr>
<tr>
<td>CIFAR-100</td>
<td>37.5 (23.5 / 51.5)</td>
<td><strong>42.5</strong> (30.2 / 54.7)</td>
</tr>
<tr>
<td>CUB-200 Birds</td>
<td>45.3 (42.3 / 48.6)</td>
<td><strong>48.4</strong> (53.3 / 45.2)</td>
</tr>
<tr>
<td>Stanford-40 Actions</td>
<td>50.4 (44.3 / 58.4)</td>
<td><strong>52.5</strong> (52.3 / 52.6)</td>
</tr>
</tbody>
</table>
Experimental results (4 tasks)

CIFAR 100

Stanford Actions

CUB 200 (birds)

Liu, Masana et al., Rotate your Networks: Better Weight Consolidation and Less Catastrophic Forgetting, ICPR 2018
Outline

• Introduction
• Transferring GANs (ECCV 2018)
• Rotated elastic weight consolidation (ICPR 2018)
• Memory Replay GANs (NIPS 2018)
• Mix and match networks (CVPR 2018)
Joint training (non-sequential)

Training data
- dog
- cat
- bird

Condition $c$

Latent vector $z$

Generator

Discriminator

Classifier $\hat{c}=\text{dog/cat}$?

Real/fake?

Auxiliary Classifier GAN (AC-GAN)

Odena et al., Conditional Image Synthesis With Auxiliary Classifier GANs, ICML 2017
Sequential learning for image generation

Task 1
- Task 2
- Task 3

Generator

Catastrophic forgetting

$z = 0.643$
$c = \text{dog}$

$z = 0.453$
$c = \text{cat}$

$z = 0.132$
$c = \text{bird}$
Sequential fine tuning and forgetting

LSUN 4 categories (4 tasks)

<table>
<thead>
<tr>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>c=bedroom</td>
<td>c=kitchen</td>
<td>c=church</td>
<td>c=tower</td>
</tr>
</tbody>
</table>

MNIST 10 categories (10 tasks)

After task 10
Memory Replay GANs (MeRGANs)

Generator

Task 1

Replay generator

Not trained:
- Remembers previous task
- Prevents forgetting

Generator

Task 2
MeRGAN-JTR: joint training w/ replay

Training data task 3

bird

dog

cat

Replayed data

c=dog,cat

previous gen. (after task 2)

Discrim.

real/fake?

Classifier

̂c=dog/cat/bird?

Current gen. (task 3)

c=dog,cat,bird

Step 1: replay previous tasks

Step 2: joint training
MeRGAN-RA: replay alignment

Step 1: replay previous tasks and align

We can do pixelwise alignment because for given $z$ and $c$ output is deterministic (thanks conditional GAN)
MeRGAN-RA: replay alignment

Training data task 3

bird

Discrim.

real/fake?

Current gen. (task 3)

c=bird

z

Step 2: learning new task
Digit generation (MNIST)

10 tasks (10 categories) 3 layers generator

<table>
<thead>
<tr>
<th>SFT</th>
<th>EWC</th>
<th>DGR</th>
<th>MeRGAN-JTR</th>
<th>MeRGAN-RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>9999</td>
<td>1111</td>
<td>0080</td>
<td>0000</td>
<td>0000</td>
</tr>
<tr>
<td>9999</td>
<td>1111</td>
<td>1111</td>
<td>1111</td>
<td>1111</td>
</tr>
<tr>
<td>9999</td>
<td>2222</td>
<td>2222</td>
<td>2222</td>
<td>2222</td>
</tr>
<tr>
<td>9999</td>
<td>3333</td>
<td>3333</td>
<td>3333</td>
<td>3333</td>
</tr>
<tr>
<td>9999</td>
<td>4444</td>
<td>4444</td>
<td>4444</td>
<td>4444</td>
</tr>
<tr>
<td>9999</td>
<td>5555</td>
<td>5555</td>
<td>5555</td>
<td>5555</td>
</tr>
<tr>
<td>9999</td>
<td>6666</td>
<td>6666</td>
<td>6666</td>
<td>6666</td>
</tr>
<tr>
<td>9999</td>
<td>7777</td>
<td>7777</td>
<td>7777</td>
<td>7777</td>
</tr>
<tr>
<td>9999</td>
<td>8888</td>
<td>8888</td>
<td>8888</td>
<td>8888</td>
</tr>
<tr>
<td>9999</td>
<td>9999</td>
<td>9999</td>
<td>9999</td>
<td>9999</td>
</tr>
</tbody>
</table>

Scene generation (LSUN)
4 tasks (4 categories) 18-layer ResNet generator

Sequential fine tuning
Task 1 Task 2 Task 3 Task 4

c=bedroom

c=kitchen

c=church

c=tower

Different bedrooms!
MeRGAN-JTR
Task 1 Task 2 Task 3 Task 4

Same bedroom!
MeRGAN-RA
Task 1 Task 2 Task 3 Task 4

Remembers the category

Remembers the instance
t-SNE visualizations (MNIST)

Generating digit 0 (i.e. first task) after learning 10 tasks

Samples from MeRGANs overlap with real data
Learning and forgetting in t-SNE (MNIST)

Generating digit 0 (i.e. task 1)

Sequential fine tuning

MeRGAN-JTR

GAN with EWC

MeRGAN-RA
Outline

• Introduction
• Transferring GANs (ECCV 2018)
• Rotated elastic weight consolidation (ICPR 2018)
• Memory Replay GANs (NIPS 2018)
• Mix and match networks (CVPR 2018)
Unseen image-to-image translations

Only these translations are trained (seen)

\[
x^{(5)} = g_5 \left( f_1 \left( x^{(1)} \right) \right)
\]

\[
x^{(1)} = g_1 \left( f_5 \left( x^{(5)} \right) \right)
\]

Evaluate on these unseen translations

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Cascading image-to-image translators

Train

\[
x^{(5)} = g_5 \left( f_1 \left( x^{(1)} \right) \right) \\
x^{(1)} = g_1 \left( f_5 \left( x^{(5)} \right) \right)
\]

Test

Two cascaded translators (via an auxiliary seen domain)

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Mix and match networks

Key problem: encoder-decoder alignment

Mix&match encoder-decoders (they haven't seen each other during training)

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Application: many-to-many translations

We could train all possible translators

Problems:
- No sharing
- Poor scalability: number of networks $O(N^2)$

20 encoders, 20 decoders (each of these networks is different from the others)

$\begin{align*}
  x^{(5)} &= g_5 \left( f_1 \left( x^{(1)} \right) \right) \\
  x^{(1)} &= g_1 \left( f_5 \left( x^{(5)} \right) \right)
\end{align*}$

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Mix and match networks

Unseen encoder–decoder alignment
- Latent representation should be domain-independent
- Achieved using shared encoder/decoders and autoencoders

5 encoders, 5 decoders
- Scalable: number of networks $O(N)$

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Example: scalable recolorization

Unpaired translation
11 colors (i.e. 11 domains)

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Example: scalable recolorization

Unpaired translation
11 colors (i.e. 11 domains)

Requires training 11 encoders and 11 decoders

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Example: scalable recolorization

Unpaired translation
11 colors (i.e. 11 domains)

Requires training 11 encoders and 11 decoders

CycleGANs for all combinations would require 55 encoders and 55 decoders

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Example: scalable style transfer

Unpaired translation
Five domains
(photo, Monet, van Gogh, Ukiyo-e, Cezanne)

(5 encoders and 5 decoders)

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Application: zero-pair translation

Cross-modal translation setting (RGB, segmentation and depth)
Paired data available for (RGB, segm.) and (RGB, depth)
Evaluate on the unseen zero-pair translations (depth, segm.)

Disjoint sets

Unseen (no pairs for training)

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with two cascaded pix2pix (paired translations)

In theory

In practice

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with CycleGAN (unpaired translation)

In theory:

Depth-to-segmentation (CycleGAN)

In practice:

Depth-to-segmentation is too complex for CycleGAN

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with mix and match networks

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with mix and match networks

Training for encoder-decoder alignment: Autoencoders

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with mix and match networks

Training for encoder-decoder alignment:

Autoencoders

Shared encoder/decoders

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with mix and match networks

Training for encoder-decoder alignment:

- $L_{SEM}$
- $L_{RGB}$
- $L_{DEPTH}$

Autoencoders

Shared encoder/decoders

Latent losses

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with mix and match networks

Training for encoder-decoder alignment:

- $L_{SEM}$
- $L_{RGB}$
- $L_{DEPTH}$

Autoencoders
- Shared encoder/decoders

Latent losses
- Robust side information (pooling indices)

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Zero-pair translation with mix and match networks

Test on zero-pair translation depth-to-segmentation

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Comparison: depth-to-segmentation

Figure 1: Zero-pair depth→segmentation, trained on (depth,RGB) and (RGB,segmentation).

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Side information in mix and match networks

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Quantitative evaluation

<table>
<thead>
<tr>
<th>Method</th>
<th>Conn.</th>
<th>$L_{SEM}$</th>
<th>Bed</th>
<th>Book</th>
<th>Ceiling</th>
<th>Chair</th>
<th>Floor</th>
<th>Furniture</th>
<th>Object</th>
<th>Picture</th>
<th>Sofa</th>
<th>Table</th>
<th>TV</th>
<th>Wall</th>
<th>Window</th>
<th>mIoU</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baselines</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CycleGAN [34]</td>
<td>SC</td>
<td>CE</td>
<td>2.79</td>
<td>0.00</td>
<td>16.9</td>
<td>6.81</td>
<td>4.48</td>
<td>0.92</td>
<td>7.43</td>
<td>0.57</td>
<td>9.48</td>
<td>0.92</td>
<td>0.31</td>
<td>17.4</td>
<td>15.1</td>
<td>6.34</td>
<td>14.2</td>
</tr>
<tr>
<td>2×pix2pix [10]</td>
<td>SC</td>
<td>CE</td>
<td>34.6</td>
<td>1.88</td>
<td>70.9</td>
<td>20.9</td>
<td>63.6</td>
<td>17.6</td>
<td>14.1</td>
<td>0.03</td>
<td>38.4</td>
<td>10.0</td>
<td>4.33</td>
<td>67.7</td>
<td>20.5</td>
<td>25.4</td>
<td>57.6</td>
</tr>
<tr>
<td>M&amp;MNet $D \rightarrow R \rightarrow S$</td>
<td>PI</td>
<td>CE</td>
<td>0.02</td>
<td>0.00</td>
<td>8.76</td>
<td>0.10</td>
<td>2.91</td>
<td>2.06</td>
<td>1.65</td>
<td>0.19</td>
<td>0.02</td>
<td>0.28</td>
<td>0.02</td>
<td>58.2</td>
<td>3.3</td>
<td>5.96</td>
<td>32.3</td>
</tr>
<tr>
<td>M&amp;MNet $D \rightarrow R \rightarrow S$</td>
<td>SC</td>
<td>CE</td>
<td>25.4</td>
<td>0.26</td>
<td>82.7</td>
<td>0.44</td>
<td>56.6</td>
<td>6.30</td>
<td>23.6</td>
<td>5.42</td>
<td>0.54</td>
<td>21.9</td>
<td>10.0</td>
<td>68.6</td>
<td>19.6</td>
<td>24.7</td>
<td>59.7</td>
</tr>
<tr>
<td><strong>Zero-pair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M&amp;MNet $D \rightarrow S$</td>
<td>PI</td>
<td>CE</td>
<td>50.8</td>
<td>18.9</td>
<td>89.8</td>
<td>31.6</td>
<td>88.7</td>
<td>48.3</td>
<td>44.9</td>
<td>62.1</td>
<td>17.8</td>
<td>49.9</td>
<td>51.9</td>
<td>86.2</td>
<td>79.2</td>
<td>55.4</td>
<td>80.4</td>
</tr>
<tr>
<td><strong>Multi-modal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M&amp;MNet $(R, D) \rightarrow S$</td>
<td>PI</td>
<td>CE</td>
<td>49.9</td>
<td>25.5</td>
<td>88.2</td>
<td>31.8</td>
<td>86.8</td>
<td>56.0</td>
<td>45.4</td>
<td>70.5</td>
<td>17.4</td>
<td>46.2</td>
<td>57.3</td>
<td>87.9</td>
<td>79.8</td>
<td>57.1</td>
<td>81.2</td>
</tr>
</tbody>
</table>

Table 3: Zero-pair depth-to-semantic segmentation. SC: skip connections, PI: pooling indexes, CE: cross-entropy

Wang, van de Weijer, Herranz, “Encoder-decoder alignment for zero-pair image-to-image translation”, CVPR 2018
Outline

• Introduction
• Transferring GANs (ECCV 2018)
• Rotated elastic weight consolidation (ICPR 2018)
• Memory Replay GANs (NIPS 2018)
• Mix and match networks (CVPR 2018)
• Other works
Other works

Domain-adaptive network compression (ICCV 2017)

(Pre)training

Target domain/task

Source domain/task

Domain-adaptive low rank (DALR) matrix decomposition

Contextual food recognition and analysis (TMM17, TMM18)

RGB-D deep representations (AAAI17, IJCAI17, TIP18)
THANK YOU!

lherranz@cvc.uab.es
www.lherranz.org