

HW7 DL Applications

Slack (Recommend): Lin Qiuxia
TA Email: qiuxia.lin@u.nus.edu

Content

- Three python notebooks
 - Homework_07_Q1.ipynb, Homework_07_Q2.ipynb, Homework_07_Q3.ipynb
- Point assignment (10 points)
 - Q1 (4 points), Q2 (3 points), Q1 (3 points)

Cautions

- DO NOT copy the code from the internet, e.g. GitHub.
- DO NOT use external libraries like Tensorflow, keras in your implementation.
- Submission
 - Please zip your answers under a folder named with your STUDENT ID: eg. `A0123456G.zip` and submit the zipped folder to Canvas. If we unzip the file, the structure should be like this:
A0123456G/
 Homework_07_Q1.ipynb
 Homework_07_Q2.ipynb
 Homework_07_Q3.ipynb

follow them strictly, otherwise it will cause penalty☹

Q1 Object detection

- Fast R-CNN

- Truck&bus detection
- What you need to do
 - Implement the structure of FRCNN
 - RoIPool, classifier, boundary regressor
 - Implement the forward function
 - Implement the loss calculation part
 - Classifier: `nn.CrossEntropyLoss()`
 - Boundary regressor: `nn.L1Loss()`

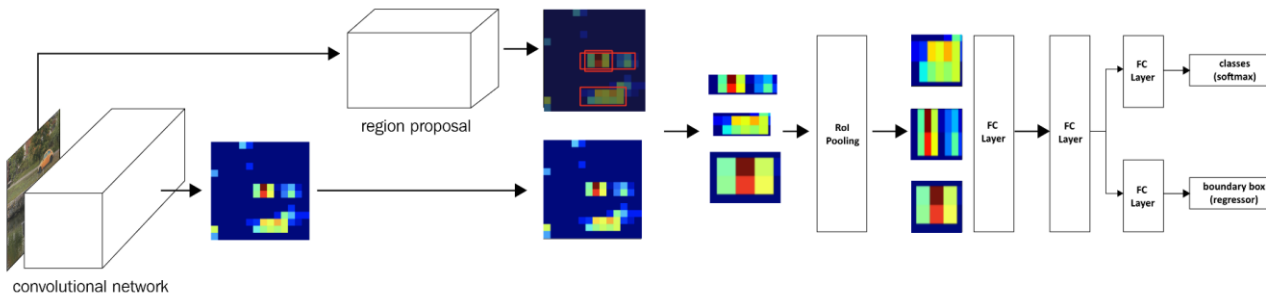


Fig. Architecture guideline for Fast RCNN taken from cs5242 lecture10

ROIPOOL

```
CLASS torchvision.ops.RoIPool(output_size: None, spatial_scale:  
float) [SOURCE]
```

See `roi_pool()`.

```
forward(input: torch.Tensor, rois: torch.Tensor) → torch.Tensor [SOURCE]
```

- **input** (`Tensor[N, C, H, W]`) – The input tensor, i.e. a batch with `N` elements. Each element contains `C` feature maps of dimensions `H × W`.
- **boxes** (`Tensor[K, 5]` or `List[Tensor[L, 4]]`) – the box coordinates in `(x1, y1, x2, y2)` format where the regions will be taken from. The coordinate must satisfy $0 \leq x1 < x2$ and $0 \leq y1 < y2$. If a single Tensor is passed, then the first column should contain the index of the corresponding element in the batch, i.e. a number in `[0, N - 1]`. If a list of Tensors is passed, then each Tensor will correspond to the boxes for an element `i` in the batch.
- **output_size** (`int` or `Tuple[int, int]`) – the size of the output after the cropping is performed, as (height, width)
- **spatial_scale** (`float`) – a scaling factor that maps the box coordinates to the input coordinates. For example, if your boxes are defined on the scale of a `224x224` image and your input is a `112x112` feature map (resulting from a `0.5x` scaling of the original image), you'll want to set this to `0.5`. Default: `1.0`

Fig. The official intro about function ROIPOOL [\[link\]](#)

Q2 Semantic Segmentation

- U-Net
 - What you need to do
 - Implement network architecture
 - Skip connection, up-conv, conv to aggregate
 - Implement forward function

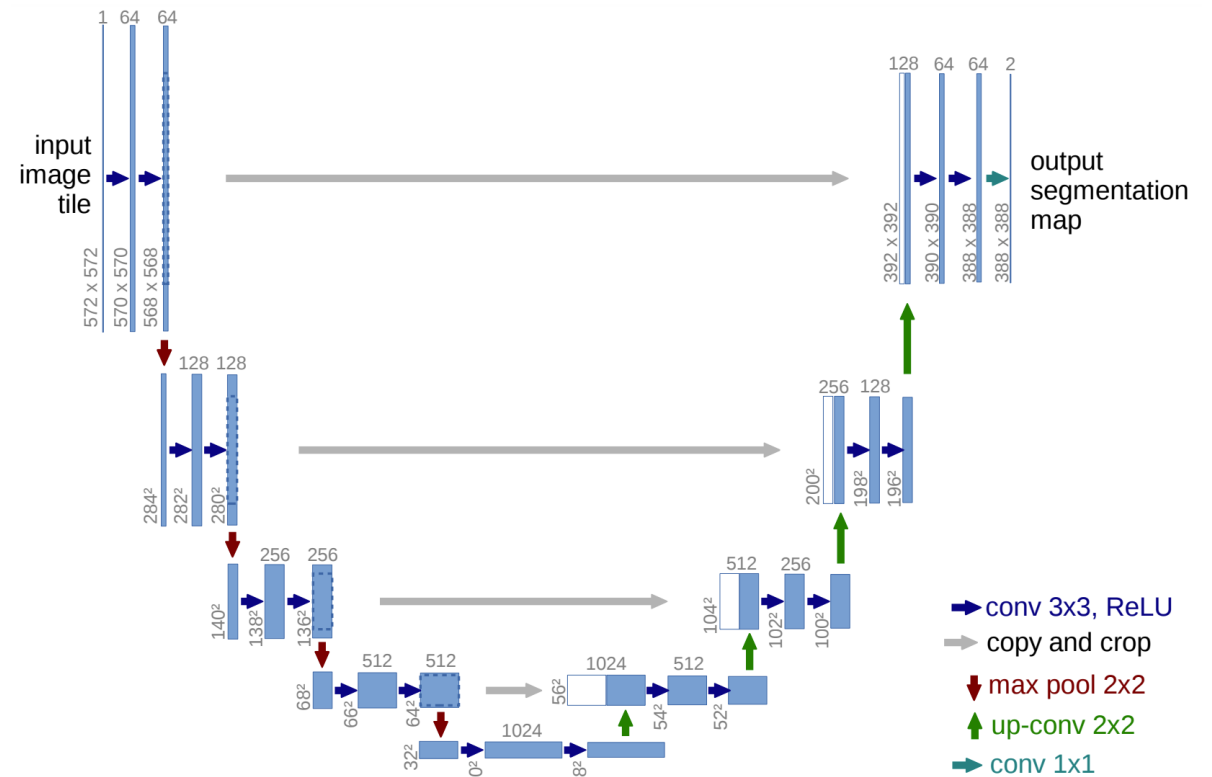


Fig. The overall architecture of U-Net taken from paper [\[link\]](#)

Q3 GANs

- Deep Convolutional GAN
 - Human face generation
 - What you need to do
 - Implement the structure of Discriminator
 - Implement the structure of Generator

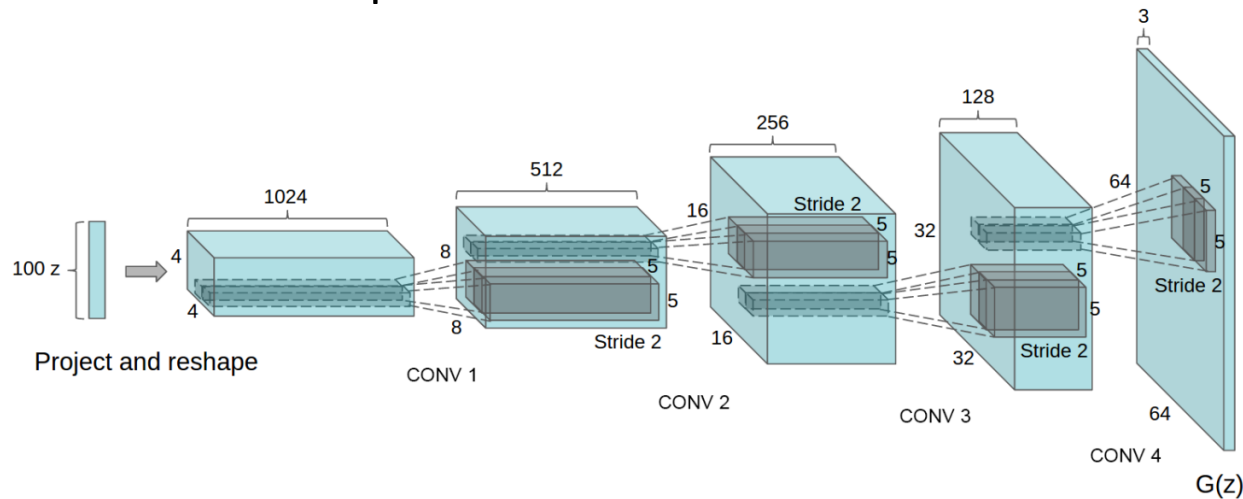


Fig. The overall architecture of DCGAN generator taken from paper[\[link\]](#)

Types of GAN

- DCGAN (Deep Convolutional GAN)
 - Use batch norm for most layers of D and G
 - except last layer of G and first layer of D
 - For D, use strided convolutions instead of pooling layers
 - For G, use transpose convolutions to upsample the latent vector to the generated image
 - No fully-connected layers with the exception of the last layer of D
 - LeakyRelu activations for all the layers (except the output layer) of D and G
 - Output of G: Tanh
 - Output of D: Logistic
 - Use Adam



Fig. Architecture guideline for DCGAN taken from cs5242 lecture11

Deadline

- 6 Nov 2022, 23:59

S	M	T	W	T	F	S
Oct						
23	24	25	26	27	28	29
		Nov				
30	31	1	2	3	4	5

6 After due date, 15% off per day late 😞