Breaking the cycle—Colleagues are all you need

Supplementary #4—Architecture

Ori Nizan
Technion, Israel
snizori@campus.technion.ac.il

Ayellet Tal
Technion, Israel
ayellet@ee.technion.ac.il

Generator–Encoder							
Layer	# filters	kernel_size	stride	padding	norm	activation	pad type
Conv2d	64	7	1	3	*in*	ReLU	*zero*
Conv2d	128	4	2	1	*in*	ReLU	*zero*
Conv2d	256	4	2	1	*in*	ReLU	*zero*
ResBlock 1							
Conv2d	256	3	1	1	*in*	ReLU	*zero*
Conv2d	256	3	1	1	*in*	None	*zero*
ResBlock 2							
Conv2d	256	3	1	1	*in*	ReLU	*zero*
Conv2d	256	3	1	1	*in*	None	*zero*
ResBlock 3							
Conv2d	256	3	1	1	*in*	ReLU	*zero*
Conv2d	256	3	1	1	*in*	None	*zero*
ResBlock 4							
Conv2d	256	3	1	1	*in*	ReLU	*zero*
Conv2d	256	3	1	1	*in*	None	*zero*
ResBlock 5							
Conv2d	256	3	1	1	*in*	ReLU	*zero*
Conv2d	256	3	1	1	*in*	None	*zero*

Figure 1: **Encoder architecture.**

Generator–Decoder							
Layer	# filters	kernel_size	stride	padding	norm	activation	pad type
ResBlock 1							
Conv2d	256	3	1	1	*adain*	ReLU	*zero*
Conv2d	256	3	1	1	*adain*	None	*zero*
ResBlock 2							
Conv2d	256	3	1	1	*adain*	ReLU	*zero*
Conv2d	256	3	1	1	*adain*	None	*zero*
ResBlock 3							
Conv2d	256	3	1	1	*adain*	ReLU	*zero*
Conv2d	256	3	1	1	*adain*	None	*zero*
ResBlock 4							
Conv2d	256	3	1	1	*adain*	ReLU	*zero*
Conv2d	256	3	1	1	*adain*	None	*zero*
ResBlock 5							
Conv2d	256	3	1	1	*adain*	ReLU	*zero*
Conv2d	256	3	1	1	*adain*	None	*zero*
Upsample	scale_factor=2				mode=nearest		
Conv2d	128	3	1	1	*adain*	ReLU	*zero*
Conv2d	128	3	1	1	*adain*	ReLU	*zero*
Upsample	scale_factor=2				mode=nearest		
Conv2d	128	3	1	1	*adain*	ReLU	*zero*
Conv2d	128	3	1	1	*adain*	ReLU	*zero*
Conv2d	64	1	1	0	*adain*	ReLU	*zero*
Conv2d	64	1	1	0	*adain*	ReLU	*zero*
Conv2d	# masks $*4$	1	1	0	None	*tanh*	*zero*
Mask	see caption						
MLP	see Table 2						

Table 1: **Decoder architecture.** This table describes the details of the generator's decoder. Here, Mask = $\tanh(10 * netOutput[:-1 * \#masks]) \longrightarrow$ last dimensions of the network output.

Layer	Input dimension	Output dimension	Normalization	Activation
LinearBlock	64	256	None	ReLU
LinearBlock	256	256	None	ReLU
LinearBlock	256	256	None	ReLU

Table 2: **MLP architecture.** This table shows the details of the last line in Table 2

Discriminator							
Layer	# filters	kernel_size	stride	padding	norm	activation	pad type
Conv2d	64	4	2	1	None	LReLU	$zero$
Conv2d	128	4	2	1	None	LReLU	$zero$
Conv2d	256	4	2	1	None	LReLU	$zero$
Conv2d	512	4	2	1	None	LReLU	$zero$
Conv2d	1	1	1	0			

Table 3: **Discriminator architecture.** Our discriminator is multi-scale, similarly to [21], with two scales.

Council Discriminator							
Layer	# filters	kernel_size	stride	padding	norm	activation	pad type
Conv2d	64	3	1	1	none	LReLU	zero
Conv2d	64	4	2	1	none	LReLU	zero
Conv2d	128	4	2	1	none	LReLU	zero
Conv2d	256	4	2	1	none	LReLU	zero
Conv2d	512	4	2	1	none	LReLU	zero
Conv2d	512	1	1	0			

Table 4: **Council discriminator architecture.** Recall that the input is a concatenation of the output and the input.