Supplementary Material - Appendix A

Unsupervised Adversarial Learning of Anomaly Detection in the Wild

September 5, 2019

1 Network Architectures

We used the implementation of pGAN by Karras et al. [1] available at `https://github.com/tkarras/progressive_growing_of_gans`. The encoder is identical to the discriminator except for the output layer. Detailed descriptions of network architectures for KTH-Cellvideos and CIFAR-10 experiments can be found below. Our code will be publicly available on GitHub if the paper is accepted.

1.1 KTH-Cellvideos

Generator	Params	OutputShape	WeightShape
latents in	-	(?, 512)	-
labels in	-	(?, 0)	-
lod	-	()	-
4x4/PixelNorm	-	(?, 512)	-
4x4/Dense	4194816	(?, 512, 4, 4)	(512, 8192)
4x4/Conv	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)
ToRGB lod4	513	(?, 1, 4, 4)	(1, 1, 512, 1)
8x8/Conv0 up	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
8x8/Conv1	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
ToRGB lod3	513	(?, 1, 8, 8)	(1, 1, 512, 1)
Upscale2D	-	(?, 1, 64, 64)	-
Grow lod3	-	(?, 1, 8, 8)	-
16x16/Conv0 up	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
16x16/Conv1	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
ToRGB lod2	513	(?, 1, 16, 16)	(1, 1, 512, 1)
Grow lod2	-	(?, 1, 16, 16)	-
32x32/Conv0 up	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv1	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
ToRGB lod1	513	(?, 1, 32, 32)	(1, 1, 512, 1)
Grow lod1	-	(?, 1, 32, 32)	-
64x64/Conv0 up	1179904	(?, 256, 64, 64)	(3, 3, 256, 512)
64x64/Conv1	590080	(?, 256, 64, 64)	(3, 3, 256, 256)
ToRGB lod0	257	(?, 1, 64, 64)	(1, 1, 256, 1)
Grow lod0	-	(?, 1, 64, 64)	-
images out	-	(?, 1, 64, 64)	-
Total	22485765		

Discriminator	Params	OutputShape	WeightShape
images in	-	(?, 1, 64, 64)	-
lod	-	()	-
FromRGB lod0	512	(?, 256, 64, 64)	(1, 1, 1, 256)
64x64/Conv0	590080	(?, 256, 64, 64)	(3, 3, 256, 256)
64x64/Conv1 down	1180160	(?, 512, 32, 32)	(3, 3, 256, 512)
Downscale2D	-	(?, 1, 4, 4)	-
FromRGB lod1	1024	(?, 512, 32, 32)	(1, 1, 1, 512)
Grow lod0	-	(?, 512, 32, 32)	-
32x32/Conv0	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv1 down	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
FromRGB lod2	1024	(?, 512, 16, 16)	(1, 1, 1, 512)
Grow lod1	-	(?, 512, 16, 16)	-
16x16/Conv0	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
16x16/Conv1 down	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
FromRGB lod3	1024	(?, 512, 8, 8)	(1, 1, 1, 512)
Grow lod2	-	(?, 512, 8, 8)	-
8x8/Conv0	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
8x8/Conv1 down	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)
FromRGB lod4	1024	(?, 512, 4, 4)	(1, 1, 1, 512)
Grow lod3	-	(?, 512, 4, 4)	-
4x4/MinibatchStddev	-	(?, 1, 4, 4)	-
4x4/Conv	2364416	(?, 512, 4, 4)	(3, 3, 513, 512)
4x4/Dense0	4194816	(?, 512)	(8192, 512)
4x4/Dense1	513	(?, 1)	(512, 1)
scores out	-	(?, 1)	-
labels out	-	(?, 0)	-
Total	22493441		

Encoder	Params	OutputShape	WeightShape
images in	-	(?, 1, 64, 64)	-
lod	-	()	-
FromRGB lod0	512	(?, 256, 64, 64)	(1, 1, 1, 256)
64x64/Conv0	590080	(?, 256, 64, 64)	(3, 3, 256, 256)
64x64/Conv1 down	1180160	(?, 512, 32, 32)	(3, 3, 256, 512)
Downscale2D	-	(?, 1, 4, 4)	-
FromRGB lod1	1024	(?, 512, 32, 32)	(1, 1, 1, 512)
Grow lod0	-	(?, 512, 32, 32)	-
32x32/Conv0	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv1 down	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
FromRGB lod2	1024	(?, 512, 16, 16)	(1, 1, 1, 512)
Grow lod1	-	(?, 512, 16, 16)	-
16x16/Conv0	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
16x16/Conv1 down	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
FromRGB lod3	1024	(?, 512, 8, 8)	(1, 1, 1, 512)
Grow lod2	-	(?, 512, 8, 8)	-
8x8/Conv0	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
8x8/Conv1 down	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)
FromRGB lod4	1024	(?, 512, 4, 4)	(1, 1, 1, 512)
Grow lod3	-	(?, 512, 4, 4)	-
4x4/MinibatchStddev	-	(?, 1, 4, 4)	-
4x4/Conv	2364416	(?, 512, 4, 4)	(3, 3, 513, 512)
4x4/Dense0	4194816	(?, 512)	(8192, 512)
z	-	(?, 512)	-
Total	22492928		

1.2 CIFAR-10

Generator	Params	OutputShape	WeightShape
latents in	-	(?, 512)	-
labels in	-	(?, 0)	-
lod	-	()	-
4x4/PixelNorm	-	(?, 512)	-
4x4/Dense	4194816	(?, 512, 4, 4)	(512, 8192)
4x4/Conv	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)
ToRGB lod4	513	(?, 1, 4, 4)	(1, 1, 512, 1)
8x8/Conv0 up	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
8x8/Conv1	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
ToRGB lod3	513	(?, 1, 8, 8)	(1, 1, 512, 1)
Upscale2D	-	(?, 1, 64, 64)	-
Grow lod3	-	(?, 1, 8, 8)	-
16x16/Conv0 up	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
16x16/Conv1	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
ToRGB lod2	513	(?, 1, 16, 16)	(1, 1, 512, 1)
Grow lod2	-	(?, 1, 16, 16)	-
32x32/Conv0 up	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv1	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
ToRGB lod0	1539	(?, 3, 32, 32)	(1, 1, 512, 3)
Grow lod0	-	(?, 3, 32, 32)	-
images out	-	(?, 3, 32, 32)	-
Total	20719628		

Discriminator	Params	OutputShape	WeightShape
images in	-	(?, 3, 32, 32)	-
lod	-	()	-
FromRGB lod0	2048	(?, 512, 32, 32)	(1, 1, 3, 512)
32x32/Conv0	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv1 down	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
Downscale2D	-	(?, 3, 4, 4)	-
FromRGB lod1	2048	(?, 512, 16, 16)	(1, 1, 3, 512)
Grow lod0	-	(?, 512, 16, 16)	-
16x16/Conv0	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
16x16/Conv1 down	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
FromRGB lod2	2048	(?, 512, 8, 8)	(1, 1, 3, 512)
Grow lod1	-	(?, 512, 8, 8)	-
8x8/Conv0	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
8x8/Conv1 down	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)
FromRGB lod3	2048	(?, 512, 4, 4)	(1, 1, 3, 512)
Grow lod2	-	(?, 512, 4, 4)	-
4x4/MinibatchStddev	-	(?, 1, 4, 4)	-
4x4/Conv	2364416	(?, 512, 4, 4)	(3, 3, 513, 512)
4x4/Dense0	4194816	(?, 512)	(8192, 512)
4x4/Dense1	513	(?, 1)	(512, 1)
scores out	-	(?, 1)	-
labels out	-	(?, 0)	-
Total	20726785		

Encoder	Params	OutputShape	WeightShape
images in	-	(?, 3, 32, 32)	-
lod	-	()	-
FromRGB lod0	2048	(?, 512, 32, 32)	(1, 1, 3, 512)
32x32/Conv0	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv1 down	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
Downscale2D	-	(?, 3, 4, 4)	-
FromRGB lod1	2048	(?, 512, 16, 16)	(1, 1, 3, 512)
Grow lod0	-	(?, 512, 16, 16)	-
16x16/Conv0	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
16x16/Conv1 down	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
FromRGB lod2	2048	(?, 512, 8, 8)	(1, 1, 3, 512)
Grow lod1	-	(?, 512, 8, 8)	-
8x8/Conv0	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
8x8/Conv1 down	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)
FromRGB lod3	2048	(?, 512, 4, 4)	(1, 1, 3, 512)
Grow lod2	-	(?, 512, 4, 4)	-
4x4/MinibatchStddev	-	(?, 1, 4, 4)	-
4x4/Conv	2364416	(?, 512, 4, 4)	(3, 3, 513, 512)
4x4/Dense0	4194816	(?, 512)	(8192, 512)
z	-	(?, 512)	-
Total	20726272		

2   Training Configuration

Hyperparameters for the networks were equal for both KTH-Cellvideos and CIFAR-10.

β_1	0.0
β_2	0.99
ϵ	1e-08
num_ gpus	1
random_ seed	1000
minibatch_ base	4
minibatch_ dict	{4: 128, 8: 128, 16: 128, 32: 64, 64: 32, 128: 16, 256: 8, 512: 4}
G_ lrate_ dict	{1024: 0.0015}
D_ lrate_ dict	{1024: 0.0015}
max_ minibatch_ per_ gpu	{256: 16, 512: 8, 1024: 4}

References

[1] T. Karras, T. Aila, S. Laine, and J. Lehtinen. Progressive Growing of GANs for Improved Quality, Stability, and Variation. In *ICLR 2018*, oct 2017.