

Appendix for *Adversarial oversampling for multi-class imbalanced data classification with convolutional neural networks*

Adam Wojciechowski      Mateusz Lango

## 1 Model architectures

Below we present model architectures for every dataset experimented on.

Layer	Dimensions	Kernel size	Stride
Input	(28x28x1)	-	-
Convolution	(28x28x32)	(3x3)	(1x1)
Max pooling	(14x14x32)	(2x2)	(2x2)
Convolution	(14x14x64)	(3x3)	(1x1)
Max pooling	(7x7x64)	(2x2)	(2x2)
Convolution	(7x7x64)	(3x3)	(1x1)
Fully Connected	(1x1x10)	-	-

Table 1: MNIST architecture

Layer	Dimensions	Kernel size	Stride
Input	(32x32x3)	-	-
Convolution	(32x32x32)	(3x3)	(1x1)
Batch Normalization	(32x32x32)	-	-
Convolution	(30x30x32)	(3x3)	(1x1)
Batch Normalization	(30x30x32)	-	-
Max pooling	(15x15x32)	(2x2)	(2x2)
Convolution	(15x15x64)	(3x3)	(1x1)
Batch Normalization	(15x15x64)	-	-
Convolution	(13x13x64)	(3x3)	(1x1)
Batch Normalization	(13x13x64)	-	-
Max pooling	(6x6x64)	(2x2)	(2x2)
Fully Connected	(1x1x10)	-	-

Table 2: CIFAR-10 architecture

Layer	Dimensions	Kernel size	Stride
Input	(150x150x3)	-	-
Convolution	(150x150x32)	(3x3)	(1x1)
Batch Normalization	(150x150x32)	-	-
Convolution	(148x148x32)	(3x3)	(1x1)
Batch Normalization	(148x148x32)	-	-
Max pooling	(74x74x32)	(2x2)	(2x2)
Convolution	(74x74x64)	(3x3)	(1x1)
Batch Normalization	(74x74x64)	-	-
Convolution	(72x72x64)	(3x3)	(1x1)
Batch Normalization	(72x72x64)	-	-
Max pooling	(36x36x64)	(2x2)	(2x2)
Convolution	(36x36x64)	(3x3)	(1x1)
Batch Normalization	(36x36x64)	-	-
Convolution	(34x34x64)	(3x3)	(1x1)
Batch Normalization	(34x34x64)	-	-
Max pooling	(17x17x64)	(2x2)	(2x2)
Fully Connected	(1x1x6)	-	-

Table 3: Intel image classification architecture

Layer	Dimensions	Kernel size	Stride
Input	(240x200x64)	-	-
Convolution	(240x200x64)	(3x3)	(1x1)
Batch Normalization	(240x200x64)	-	-
Convolution	(238x198x64)	(3x3)	(1x1)
Batch Normalization	(238x198x64)	-	-
Max pooling	(119x99x64)	(2x2)	(2x2)
Convolution	(119x99x128)	(3x3)	(1x1)
Batch Normalization	(119x99x128)	-	-
Convolution	(117x97x128)	(3x3)	(1x1)
Batch Normalization	(117x97x128)	-	-
Max pooling	(58x48x128)	(2x2)	(2x2)
Convolution	(58x48x256)	(3x3)	(1x1)
Batch Normalization	(58x48x128)	-	-
Convolution	(56x46x128)	(3x3)	(1x1)
Batch Normalization	(56x46x128)	-	-
Max pooling	(28x23x128)	(2x2)	(2x2)
Fully Connected	(1x1x13)	-	-

Table 4: MIT indoor scene recognition