Computer Vision

Alien vs Predator Classification

Roman Garayev

Dataset. Artificially was expanded by Image Generator

Represents Images of Aliens and Predators.

- Format: JPG images, various thumbnail sizes (around 250 x 250 px).

Architecture and Method

1. First convolutional layer kernel is larger that others. Usually 3x3 kernels perform better that 5x5, but if 5x5 kernel goes first it can reduce spatial dimension of an image. Moreover, it will not lose to much valuable information.

2.Convolutional layer follows by Max Pooling.This structure is typical

3 Number of filters grows as we go deeper. It was made to catch low-level features

Also, because of pooling layer divided spatial dimension by 2, there was not exploding of parametres

Accuracy 76%

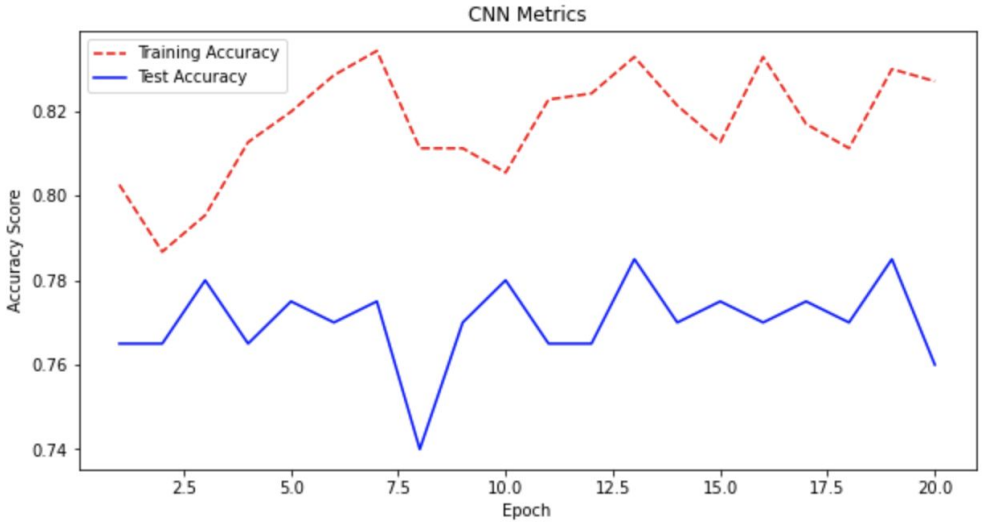
Optimizer RMSprop

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 256, 256, 64)	4864
max_pooling2d (MaxPooling2D)	(None, 128, 128, 64)	0
conv2d_1 (Conv2D)	(None, 128, 128, 64)	36928
max_pooling2d_1 (MaxPooling2	(None, 64, 64, 64)	0
conv2d_2 (Conv2D)	(None, 64, 64, 64)	36928
max_pooling2d_2 (MaxPooling2	(None, 32, 32, 64)	0
conv2d_3 (Conv2D)	(None, 32, 32, 128)	73856
max_pooling2d_3 (MaxPooling2	(None, 16, 16, 128)	0
conv2d_4 (Conv2D)	(None, 16, 16, 128)	147584
max_pooling2d_4 (MaxPooling2	(None, 8, 8, 128)	0
flatten (Flatten)	(None, 8192)	0
dense (Dense)	(None, 512)	4194816
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 512)	262656
dense_2 (Dense)	(None, 1)	513

Learning Rate Scheduling

1-cycle LR scheduling method was introduced in 2018 year paper.It Starts by increasing the initial LR linearly to some value during the first half of training. Next, it decrease it linearly as well during the second half of training. Finally, in last few epochs it dropping rate down by several orders of magnitude linerally too

In this particular case it did not help to improve the performance but in general in larger datasets it helps.



Transfer Learning. 91% Accuracy

For transfer learning as pretrained model I decided to use Xception model and added average pooling layer. Perfomance increased dramatically, as it was expected :)