Follow-Me Project

Congratulations on reaching the final project of the Robotics Nanodegree!

Previously, you worked on the Semantic Segmentation lab where you built a deep learning network that locates a particular human target within an image. For this project, you will utilize what you implemented and learned from that lab and extend it to train a deep learning model that will allow a simulated quadcopter to follow around the person that it detects!

Most of the code below is similar to the lab with some minor modifications. You can start with your existing solution, and modify and improve upon it to train the best possible model for this task.

You can click on any of the following to quickly jump to that part of this notebook:

Data Collection

We have provided you with a starting dataset for this project. Download instructions can be found in the README for this project's repo. Alternatively, you can collect additional data of your own to improve your model. Check out the "Collecting Data" section in the Project Lesson in the Classroom for more details!

In [1]:

```
import os
import glob
import sys
import tensorflow as tf

from scipy import misc
import numpy as np

from tensorflow.contrib.keras.python import keras
from tensorflow.contrib.keras.python.keras import layers, models

from tensorflow import image

from utils import scoring_utils
from utils.separable_conv2d import SeparableConv2DKeras, BilinearUpSam
pling2D
from utils import data_iterator
from utils import plotting_tools
from utils import model_tools
```

FCN Layers

In the Classroom, we discussed the different layers that constitute a fully convolutional network (FCN). The following code will introduce you to the functions that you need to build your semantic segmentation model.

Separable Convolutions

The Encoder for your FCN will essentially require separable convolution layers, due to their advantages as explained in the classroom. The 1x1 convolution layer in the FCN, however, is a regular convolution. Implementations for both are provided below for your use. Each includes batch normalization with the ReLU activation function applied to the layers.

In [2]:
```
def separable_conv2d_batchnorm(input_layer, filters, strides=1):
    output_layer = SeparableConv2DKeras(filters=filters,kernel_size=3,
strides=strides,
                             padding='same', activation='relu')(input_
layer)

    output_layer = layers.BatchNormalization()(output_layer)
    return output_layer

def conv2d_batchnorm(input_layer, filters, kernel_size=3, strides=1):
    output_layer = layers.Conv2D(filters=filters, kernel_size=kernel_s
ize, strides=strides,
                      padding='same', activation='relu')(input_layer)

    output_layer = layers.BatchNormalization()(output_layer)
    return output_layer
```

Bilinear Upsampling

The following helper function implements the bilinear upsampling layer. Upsampling by a factor of 2 is generally recommended, but you can try out different factors as well. Upsampling is used in the decoder block of the FCN.

In [3]:
```
def bilinear_upsample(input_layer):
    output_layer = BilinearUpSampling2D((2,2))(input_layer)
    return output_layer
```

Build the Model

In the following cells, you will build an FCN to train a model to detect and locate the hero target within an image. The steps are:

- Create an `encoder_block`
- Create a `decoder_block`
- Build the FCN consisting of encoder block(s), a 1x1 convolution, and decoder block(s). This step requires experimentation with different numbers of layers and filter sizes to build your model.

Encoder Block

Create an encoder block that includes a separable convolution layer using the `separable_conv2d_batchnorm()` function. The `filters` parameter defines the size or depth of the output layer. For example, 32 or 64.

In [4]:
```
def encoder_block(input_layer, filters, strides):

    output_layer = separable_conv2d_batchnorm(input_layer, filters, st
rides=strides)

    return output_layer
```

Decoder Block

The decoder block is comprised of three parts:

- A bilinear upsampling layer using the upsample_bilinear() function. The current recommended factor for upsampling is set to 2.
- A layer concatenation step. This step is similar to skip connections. You will concatenate the upsampled small_ip_layer and the large_ip_layer.
- Some (one or two) additional separable convolution layers to extract some more spatial information from prior layers.

In [5]:
```
def decoder_block(small_ip_layer, large_ip_layer, filters):

    sampled = bilinear_upsample(small_ip_layer)

    # Concatenate the upsampled and large input layers using layers.co
ncatenate
    cat = layers.concatenate([sampled, large_ip_layer])

    # Add some number of separable convolution layers
    #conv1 = layers.Conv2D(filters=filters, kernel_size=1, strides=1,
padding='same', activation='relu')(cat)
    #conv2 = layers.Conv2D(filters=filters, kernel_size=1, strides=1,
padding='same', activation='relu')(conv1)
    #output_layer = layers.Conv2D(filters=filters, kernel_size=1, stri
des=1, padding='same', activation='relu')(conv2)

    conv1 = SeparableConv2DKeras(filters=filters,kernel_size=1, stride
s=1,padding='same', activation='relu')(cat)
    conv2 = SeparableConv2DKeras(filters=filters,kernel_size=1, stride
s=1,padding='same', activation='relu')(conv1)
    output_layer = SeparableConv2DKeras(filters=filters,kernel_size=1,
strides=1,padding='same', activation='relu')(conv2)

    return output_layer
```

Model

Now that you have the encoder and decoder blocks ready, go ahead and build your FCN architecture!

There are three steps:

- Add encoder blocks to build the encoder layers. This is similar to how you added regular convolutional layers in your CNN lab.
- Add a 1x1 Convolution layer using the conv2d_batchnorm() function. Remember that 1x1 Convolutions require a kernel and stride of 1.
- Add decoder blocks for the decoder layers.

In [6]:
```
def fcn_model(inputs, num_classes):

    layer1 = encoder_block(inputs, 32, strides=2)
    layer2 = encoder_block(layer1, 64, strides=2)
    layer3 = encoder_block(layer2, 64, strides=2)



    # Add 1x1 Convolution layer using conv2d_batchnorm().
    conv2d_batchnormed = conv2d_batchnorm(layer3, 64, kernel_size=1, s
trides=1)

    # Add the same number of Decoder Blocks as the number of Encoder B
locks

    layer4 = decoder_block(conv2d_batchnormed, layer2, 64)
    layer5 = decoder_block(layer4, layer1, 64)
    x = decoder_block(layer5, inputs, 32)

    # The function returns the output layer of your model. "x" is the
final layer obtained from the last decoder_block()
    output_layer = layers.Conv2D(num_classes, 3, activation='softmax',
padding='same')(x)

    return output_layer
```

Training

The following cells will use the FCN you created and define an ouput layer based on the size of the processed image and the number of classes recognized. You will define the hyperparameters to compile and train your model.

Please Note: For this project, the helper code in `data_iterator.py` will resize the copter images to 160x160x3 to speed up training.

In [7]:
```
"""
DON'T MODIFY ANYTHING IN THIS CELL THAT IS BELOW THIS LINE
"""

image_hw = 160
image_shape = (image_hw, image_hw, 3)
inputs = layers.Input(image_shape)
num_classes = 3

# Call fcn_model()
output_layer = fcn_model(inputs, num_classes)
```

Hyperparameters

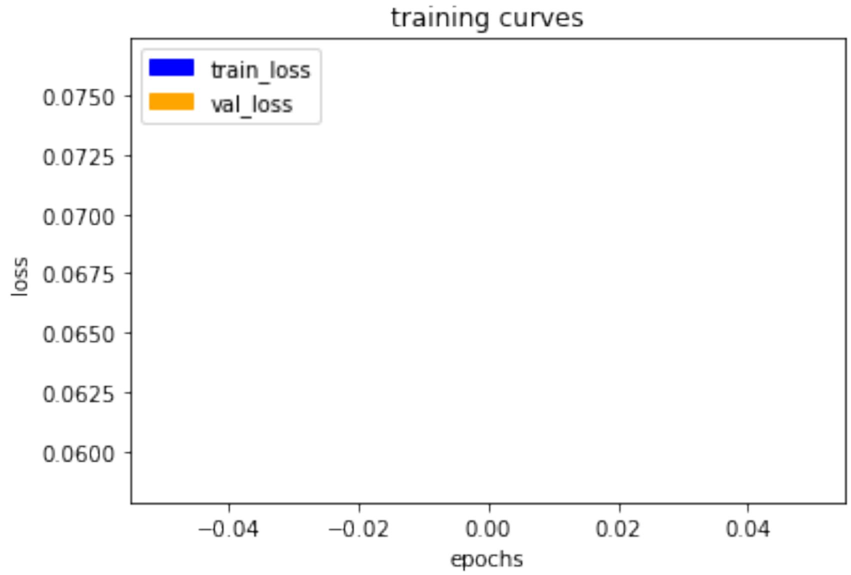
Define and tune your hyperparameters.

- **batch_size**: number of training samples/images that get propagated through the network in a single pass.
- **num_epochs**: number of times the entire training dataset gets propagated through the network.
- **steps_per_epoch**: number of batches of training images that go through the network in 1 epoch. We have provided you with a default value. One recommended value to try would be based on the total number of images in training dataset divided by the batch_size.
- **validation_steps**: number of batches of validation images that go through the network in 1 epoch. This is similar to steps_per_epoch, except validation_steps is for the validation dataset. We have provided you with a default value for this as well.
- **workers**: maximum number of processes to spin up. This can affect your training speed and is dependent on your hardware. We have provided a recommended value to work with.

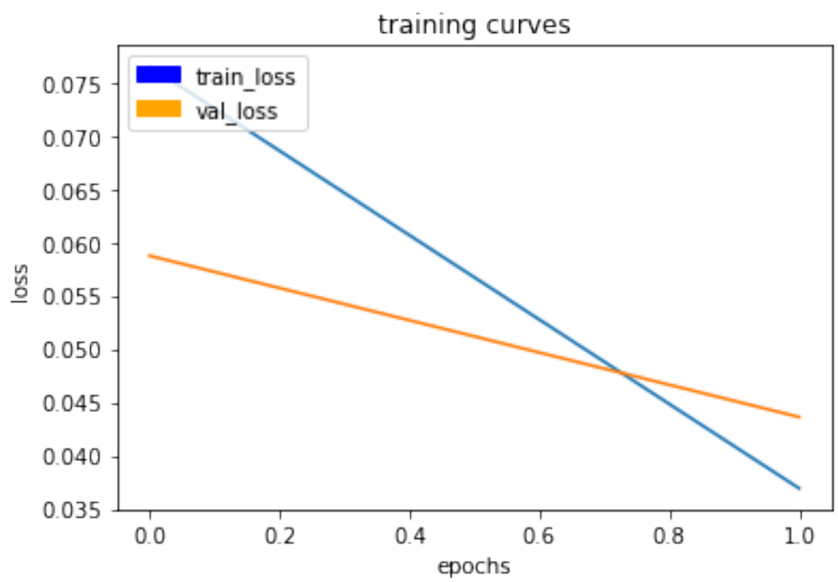
In [8]:
```
learning_rate = 0.01
batch_size = 64
num_epochs = 20
steps_per_epoch = 400
validation_steps = 50
workers = 2
```

In [9]:
```
"""
DON'T MODIFY ANYTHING IN THIS CELL THAT IS BELOW THIS LINE
"""
# Define the Keras model and compile it for training
model = models.Model(inputs=inputs, outputs=output_layer)

model.compile(optimizer=keras.optimizers.Adam(learning_rate), loss='ca
tegorical_crossentropy')

# Data iterators for loading the training and validation data
train_iter = data_iterator.BatchIteratorSimple(batch_size=batch_size,
                                               data_folder=os.path.joi
n('..', 'data', 'train'),
                                               image_shape=image_shape
,
                                               shift_aug=True)

val_iter = data_iterator.BatchIteratorSimple(batch_size=batch_size,
                                             data_folder=os.path.join(
'..', 'data', 'validation'),
                                             image_shape=image_shape)

logger_cb = plotting_tools.LoggerPlotter()
callbacks = [logger_cb]

model.fit_generator(train_iter,
                    steps_per_epoch = steps_per_epoch, # the number of
batches per epoch,
                    epochs = num_epochs, # the number of epochs to tra
in for,
                    validation_data = val_iter, # validation iterator
                    validation_steps = validation_steps, # the number
of batches to validate on
                    callbacks=callbacks,
                    workers = workers)
```

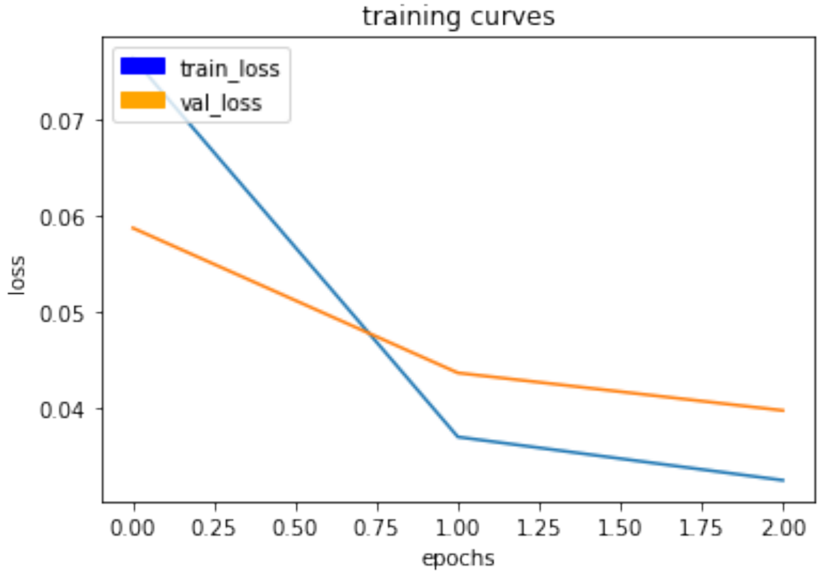
```
Epoch 1/20
399/400 [============================>.] - ETA: 0s - loss: 0.0764
```



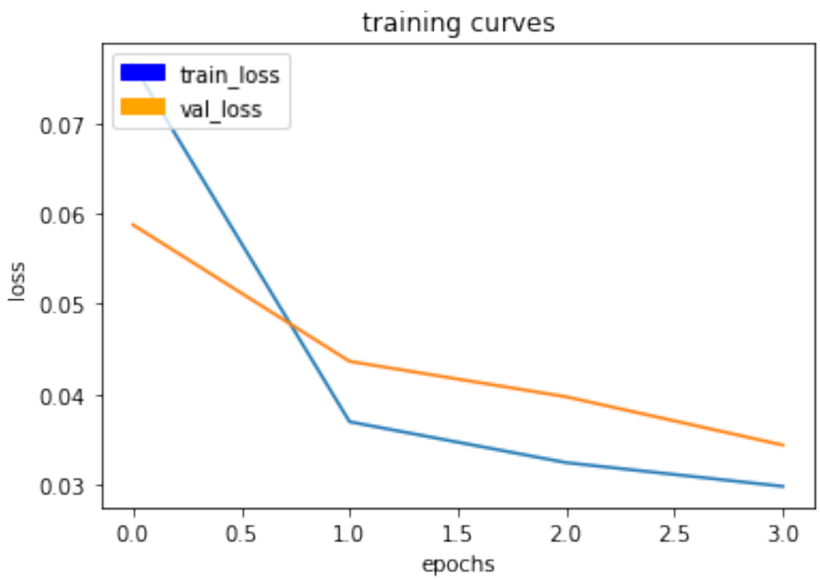

```
400/400 [==============================] - 370s - loss: 0.0763 - val
_loss: 0.0587
Epoch 2/20
399/400 [============================>.] - ETA: 0s - loss: 0.0370
```



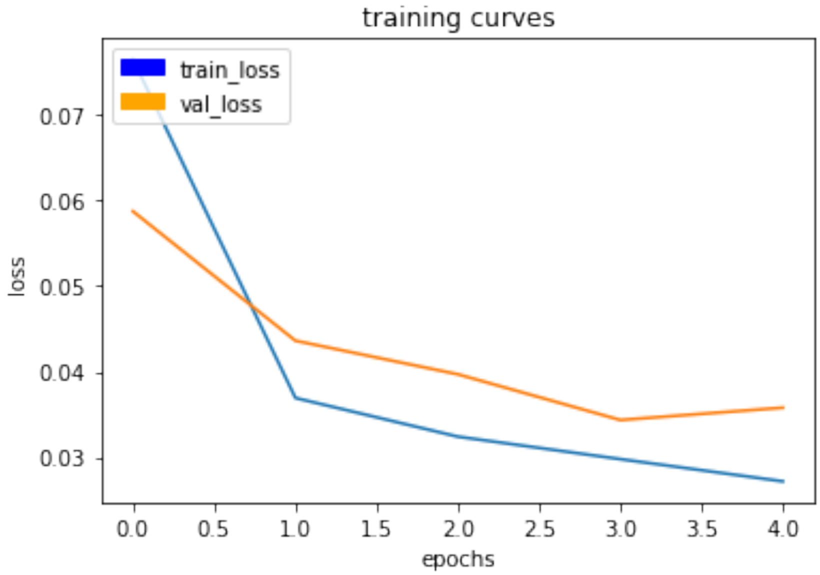

```
400/400 [==============================] - 362s - loss: 0.0370 - val
_loss: 0.0436
Epoch 3/20
399/400 [============================>.] - ETA: 0s - loss: 0.0325
```



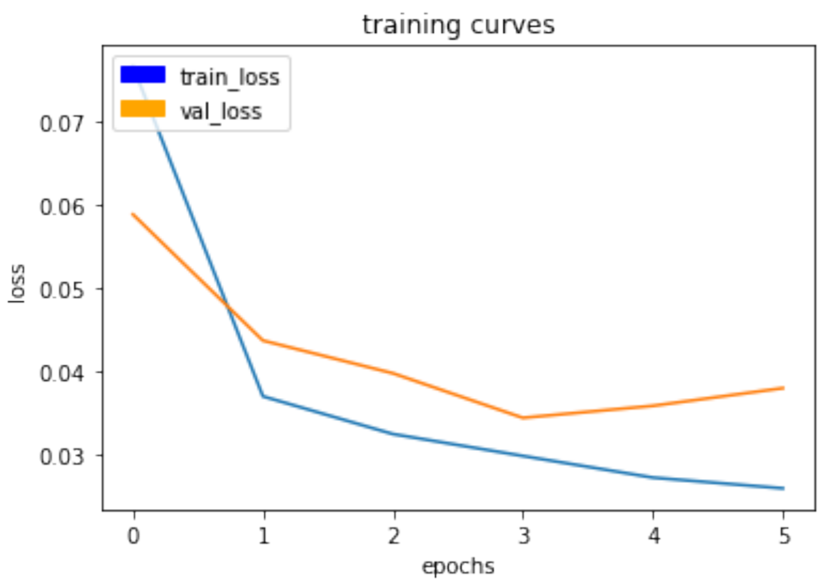

```
400/400 [==============================] - 363s - loss: 0.0324 - val
_loss: 0.0397
Epoch 4/20
399/400 [============================>.] - ETA: 0s - loss: 0.0298
```



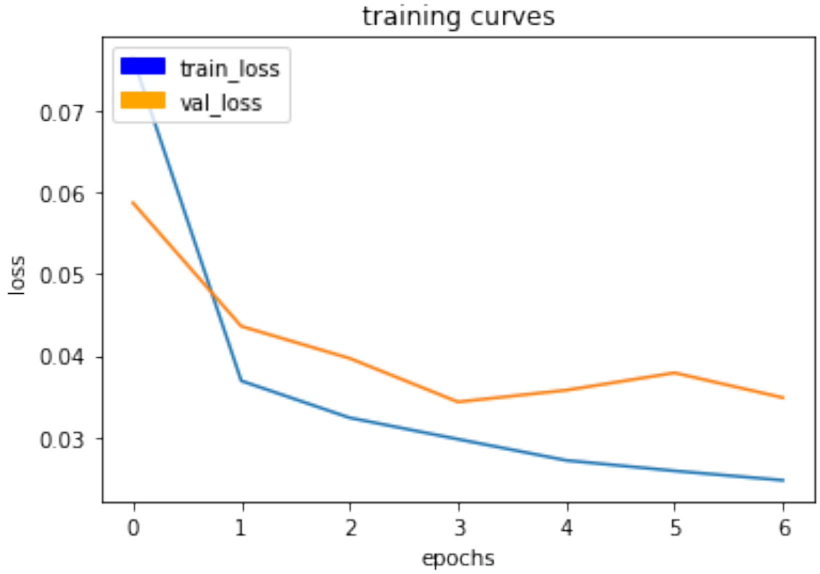

```
400/400 [==============================] - 363s - loss: 0.0298 - val
_loss: 0.0343
Epoch 5/20
399/400 [============================>.] - ETA: 0s - loss: 0.0272
```



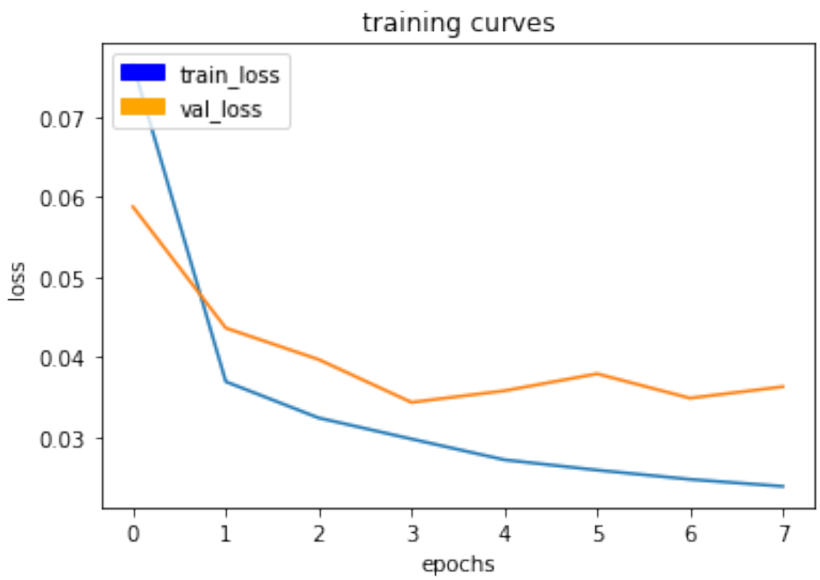

```
400/400 [==============================] - 363s - loss: 0.0272 - val
_loss: 0.0358
Epoch 6/20
399/400 [============================>.] - ETA: 0s - loss: 0.0259
```



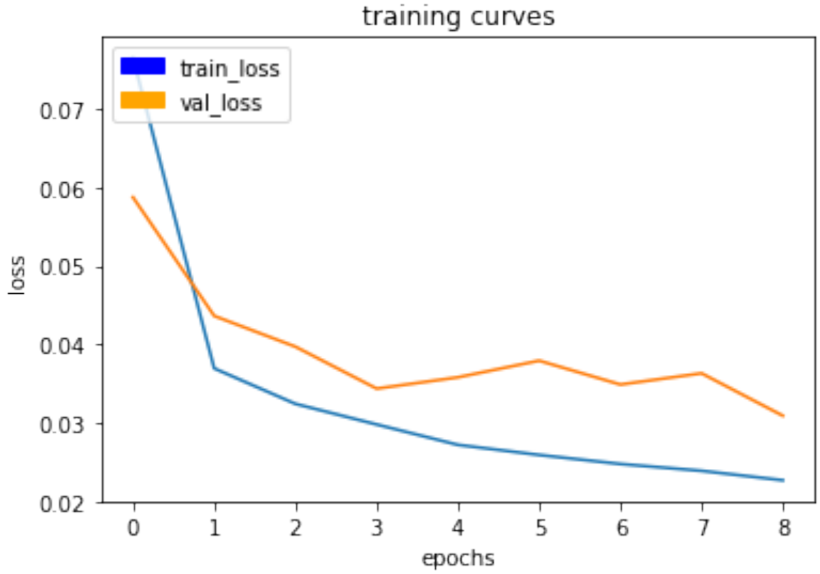

```
400/400 [==============================] - 363s - loss: 0.0259 - val
_loss: 0.0379
Epoch 7/20
399/400 [============================>.] - ETA: 0s - loss: 0.0247
```



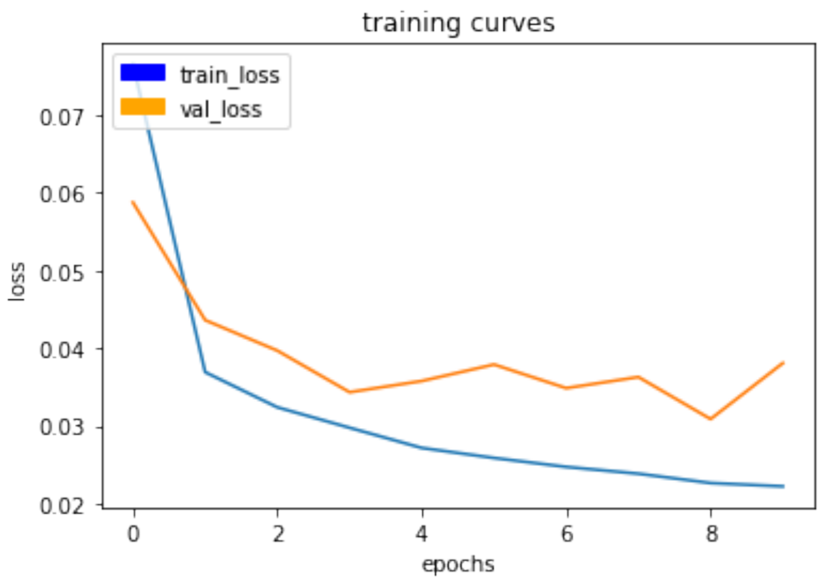

```
400/400 [==============================] - 362s - loss: 0.0248 - val
_loss: 0.0349
Epoch 8/20
399/400 [============================>.] - ETA: 0s - loss: 0.0239
```



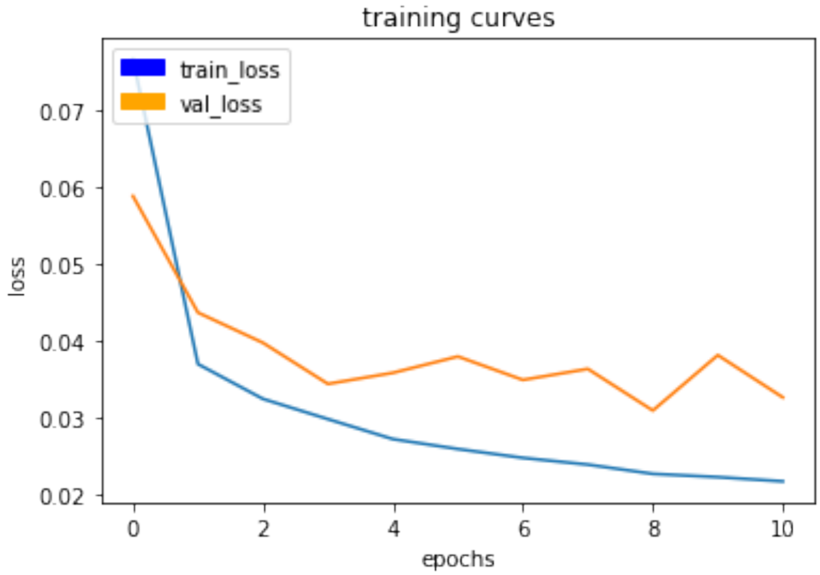

```
400/400 [==============================] - 362s - loss: 0.0238 - val
_loss: 0.0363
Epoch 9/20
399/400 [============================>.] - ETA: 0s - loss: 0.0226
```



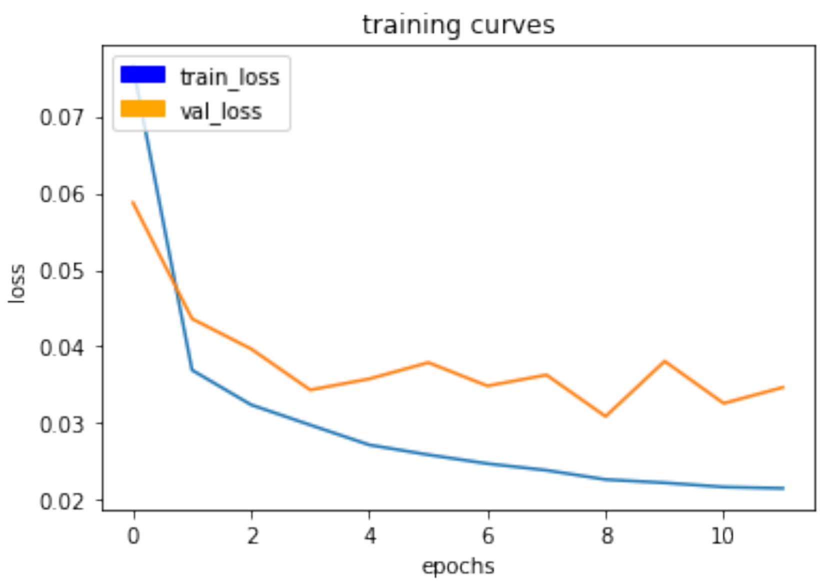

```
400/400 [==============================] - 362s - loss: 0.0227 - val
_loss: 0.0309
Epoch 10/20
399/400 [============================>.] - ETA: 0s - loss: 0.0223
```



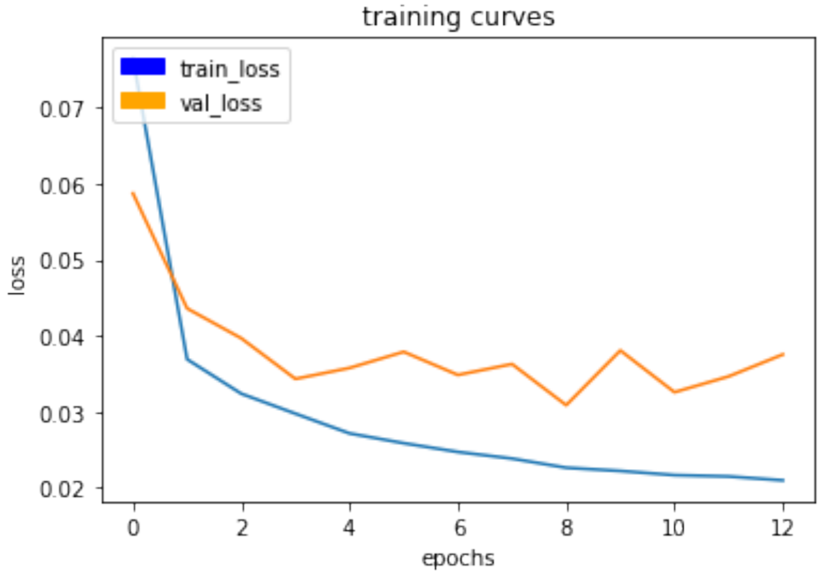

```
400/400 [==============================] - 363s - loss: 0.0223 - val
_loss: 0.0381
Epoch 11/20
399/400 [============================>.] - ETA: 0s - loss: 0.0217
```



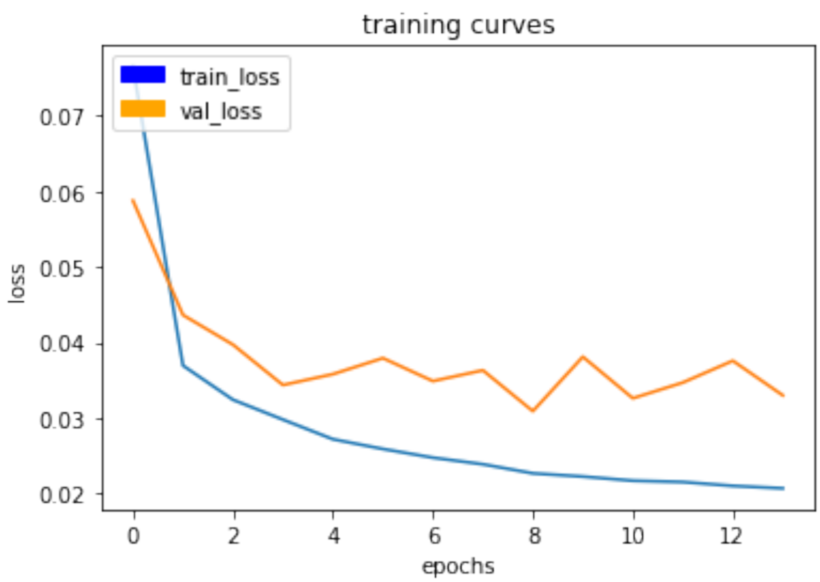

```
400/400 [==============================] - 363s - loss: 0.0217 - val
_loss: 0.0326
Epoch 12/20
399/400 [============================>.] - ETA: 0s - loss: 0.0215
```



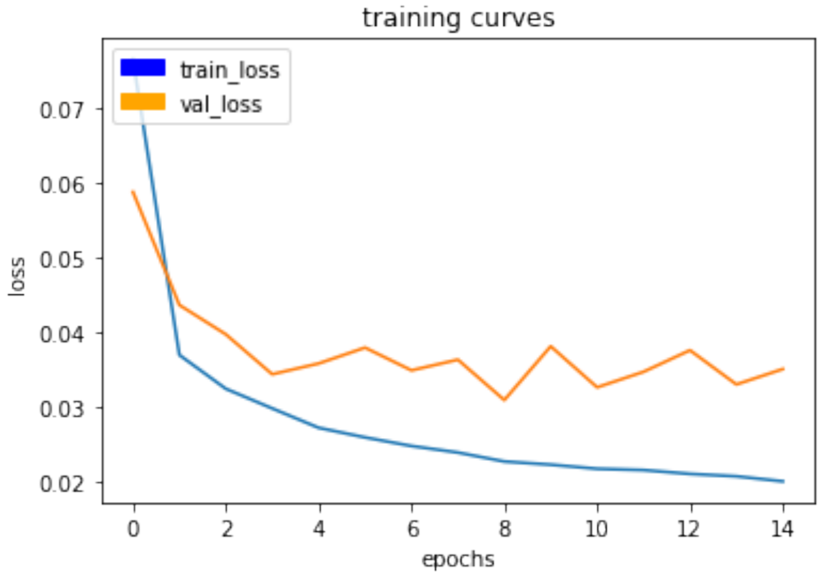

```
400/400 [==============================] - 362s - loss: 0.0215 - val
_loss: 0.0347
Epoch 13/20
399/400 [============================>.] - ETA: 0s - loss: 0.0210
```



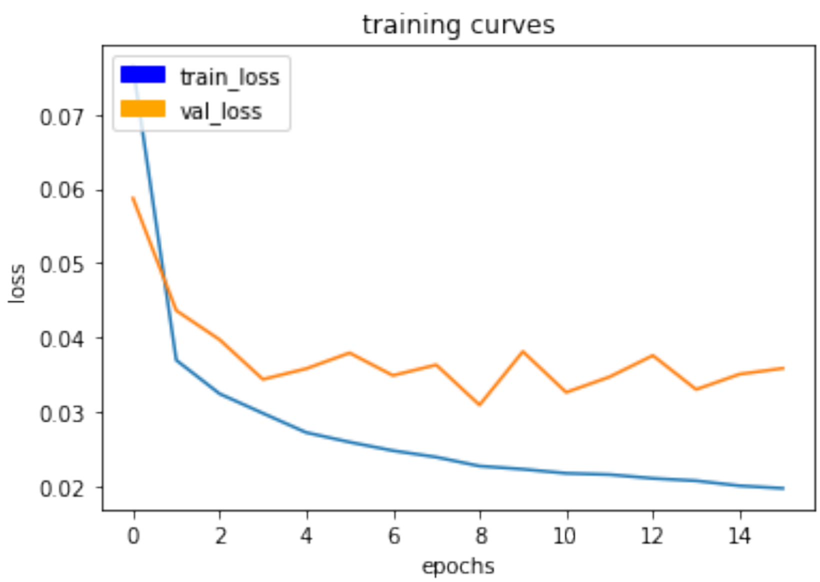

```
400/400 [==============================] - 362s - loss: 0.0210 - val
_loss: 0.0376
Epoch 14/20
399/400 [============================>.] - ETA: 0s - loss: 0.0207
```



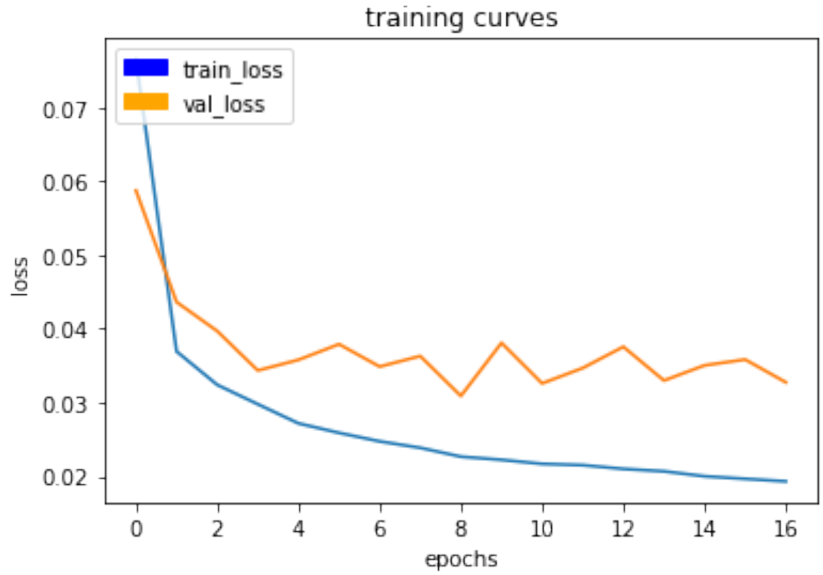

```
400/400 [==============================] - 362s - loss: 0.0207 - val
_loss: 0.0330
Epoch 15/20
399/400 [============================>.] - ETA: 0s - loss: 0.0200
```



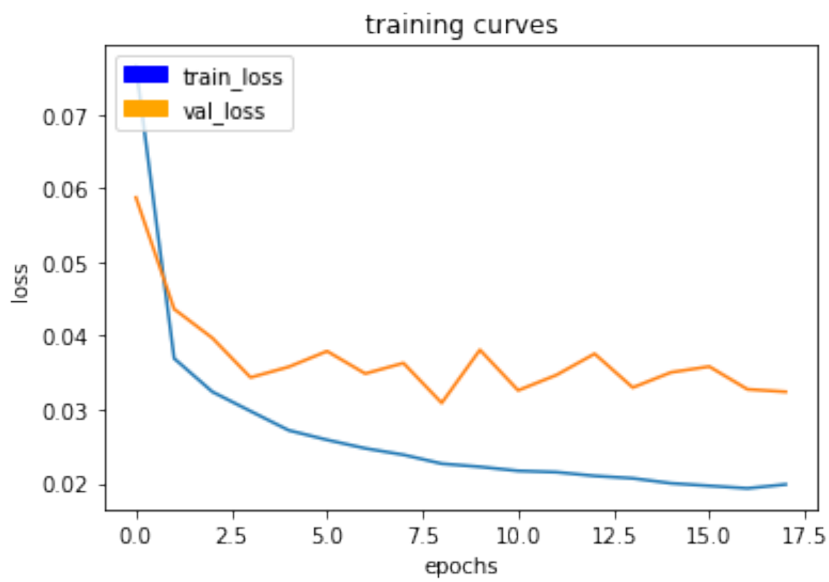

```
400/400 [==============================] - 362s - loss: 0.0200 - val
_loss: 0.0350
Epoch 16/20
399/400 [============================>.] - ETA: 0s - loss: 0.0197
```



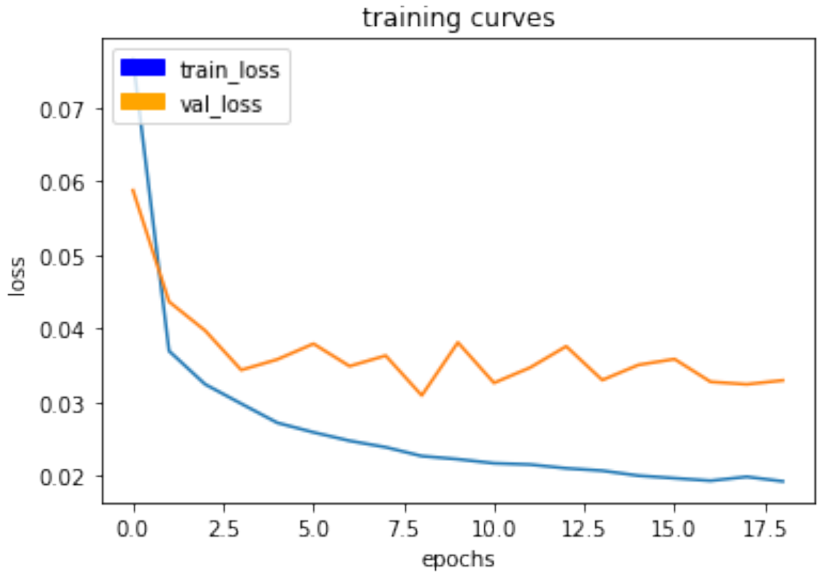

```
400/400 [==============================] - 362s - loss: 0.0196 - val
_loss: 0.0358
Epoch 17/20
399/400 [============================>.] - ETA: 0s - loss: 0.0193
```



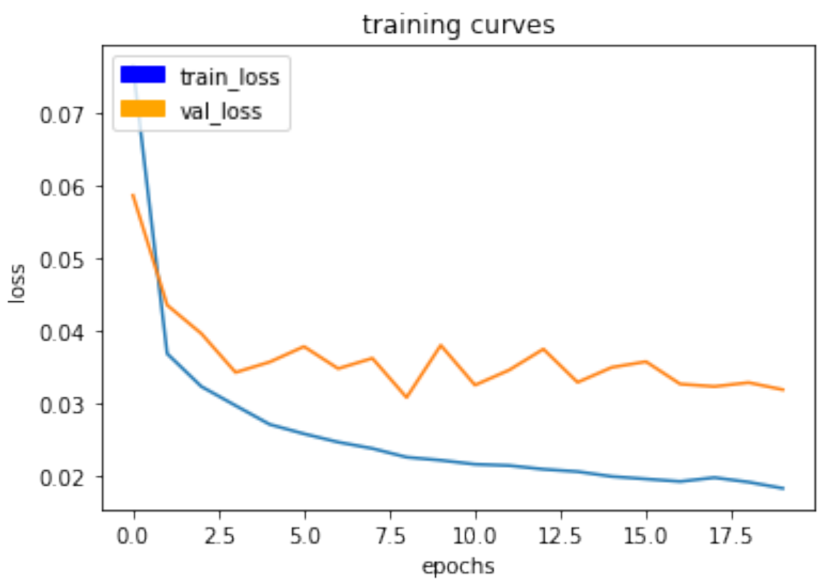

```
400/400 [==============================] - 362s - loss: 0.0193 - val
_loss: 0.0327
Epoch 18/20
399/400 [============================>.] - ETA: 0s - loss: 0.0198
```




```
400/400 [==============================] - 363s - loss: 0.0199 - val
_loss: 0.0324
Epoch 19/20
399/400 [============================>.] - ETA: 0s - loss: 0.0193
```




```
400/400 [==============================] - 361s - loss: 0.0193 - val
_loss: 0.0329
Epoch 20/20
399/400 [============================>.] - ETA: 0s - loss: 0.0184
```




```
400/400 [==============================] - 361s - loss: 0.0184 - val
_loss: 0.0320
```

Out[9]: <tensorflow.contrib.keras.python.keras.callbacks.History at 0x7f5887
015fd0>

In [10]:

```
# Save your trained model weights
weight_file_name = 'my_aws_model_weights'
model_tools.save_network(model, weight_file_name)
```

Prediction

Now that you have your model trained and saved, you can make predictions on your validation dataset. These predictions can be compared to the mask images, which are the ground truth labels, to evaluate how well your model is doing under different conditions.

There are three different predictions available from the helper code provided:

- **patrol_with_targ**: Test how well the network can detect the hero from a distance.
- **patrol_non_targ**: Test how often the network makes a mistake and identifies the wrong person as the target.
- **following_images**: Test how well the network can identify the target while following them.

In [11]:
```
# If you need to load a model which you previously trained you can unc
omment the codeline that calls the function below.

weight_file_name = 'model_weights'
#model = model_tools.load_network(weight_file_name)
```

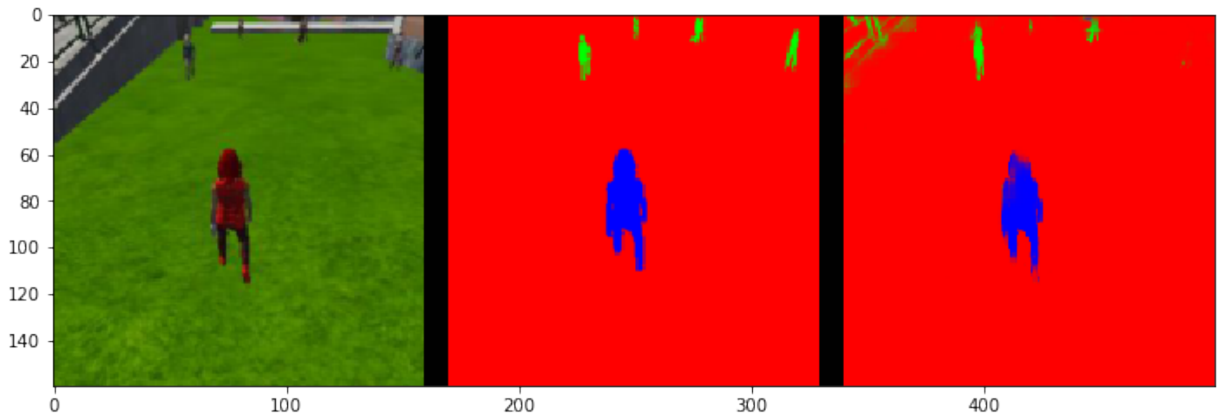
The following cell will write predictions to files and return paths to the appropriate directories. The `run_num` parameter is used to define or group all the data for a particular model run. You can change it for different runs. For example, 'run_1', 'run_2' etc.

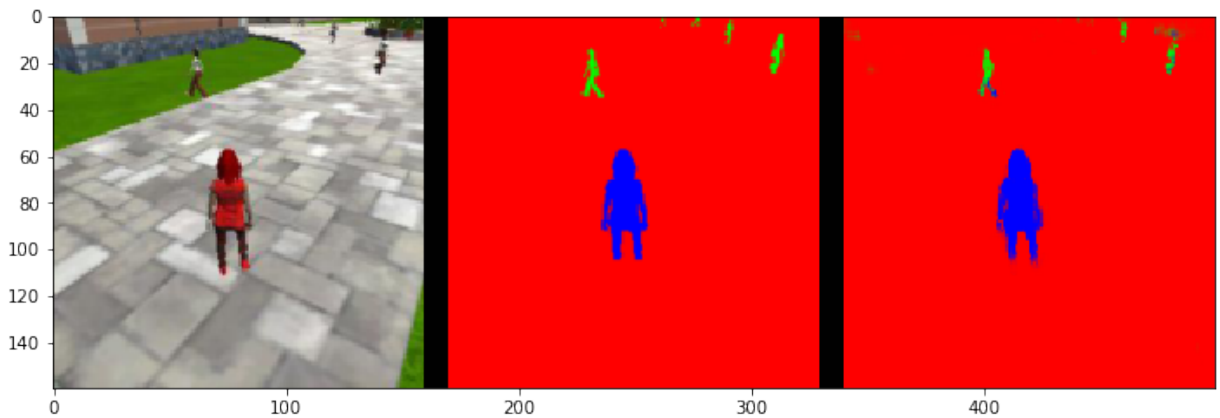
In [12]:
```
run_num = 'run_2'

val_with_targ, pred_with_targ = model_tools.write_predictions_grade_se
t(model,
                                        run_num,'patrol_with_targ', 's
ample_evaluation_data')

val_no_targ, pred_no_targ = model_tools.write_predictions_grade_set(mo
del,
                                        run_num,'patrol_non_targ', 'sa
mple_evaluation_data')

val_following, pred_following = model_tools.write_predictions_grade_se
t(model,
                                        run_num,'following_images', 's
ample_evaluation_data')
```

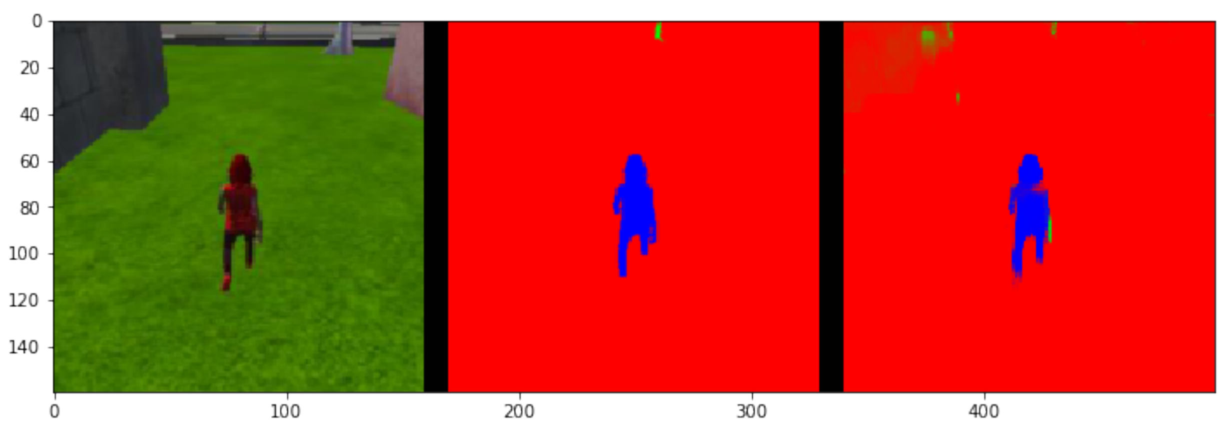
Now lets look at your predictions, and compare them to the ground truth labels and original images. Run each of the following cells to visualize some sample images from the predictions in the validation set.

In [13]:

```
# images while following the target
im_files = plotting_tools.get_im_file_sample('sample_evaluation_data',
'following_images', run_num)
for i in range(3):
    im_tuple = plotting_tools.load_images(im_files[i])
    plotting_tools.show_images(im_tuple)
```

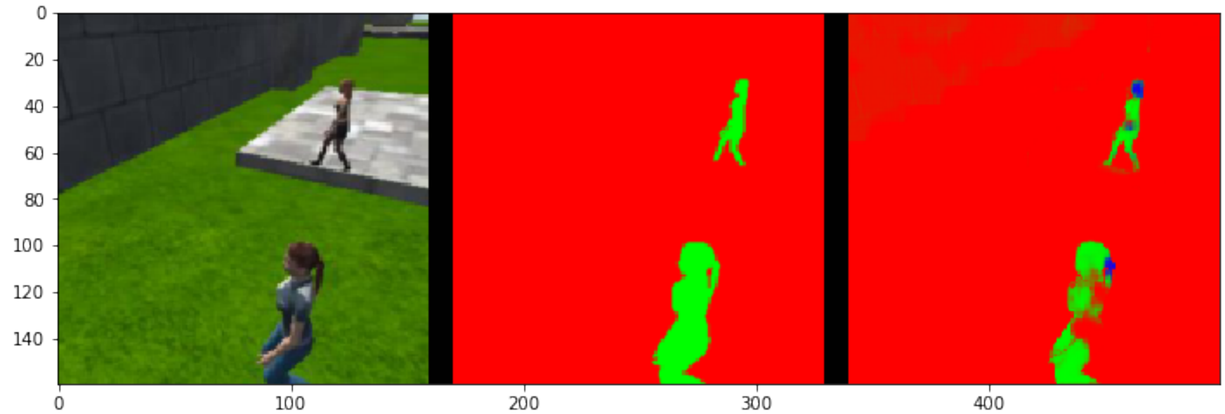


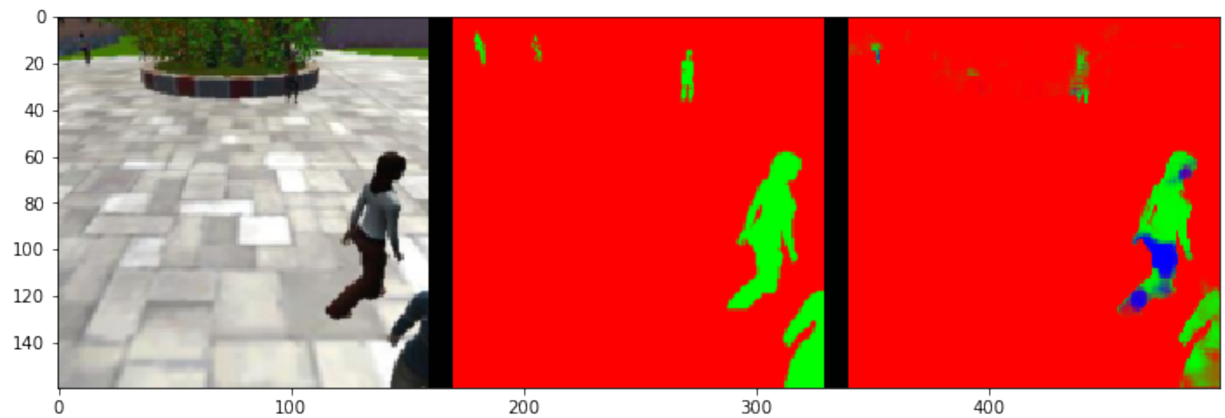


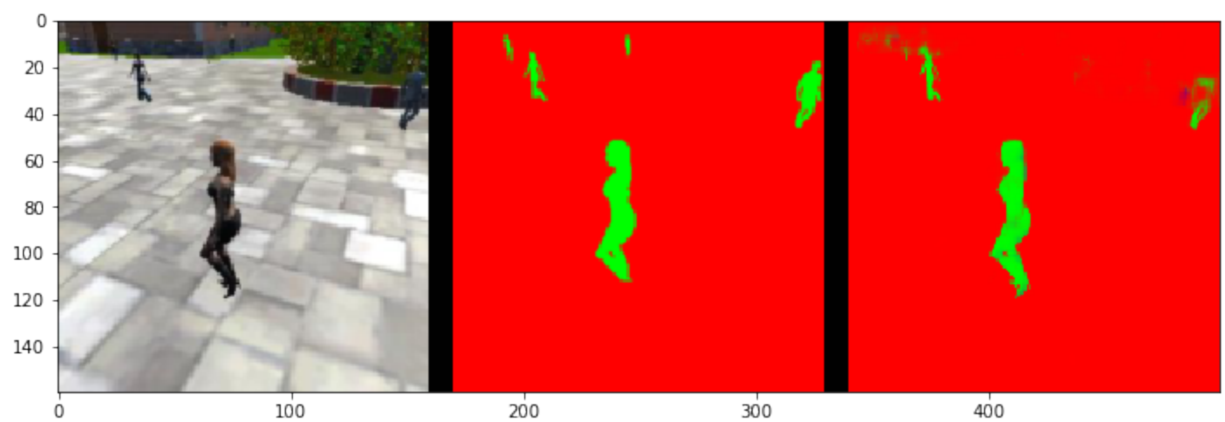


In [14]:

```
# images while at patrol without target
im_files = plotting_tools.get_im_file_sample('sample_evaluation_data',
'patrol_non_targ', run_num)
for i in range(3):
    im_tuple = plotting_tools.load_images(im_files[i])
    plotting_tools.show_images(im_tuple)
```

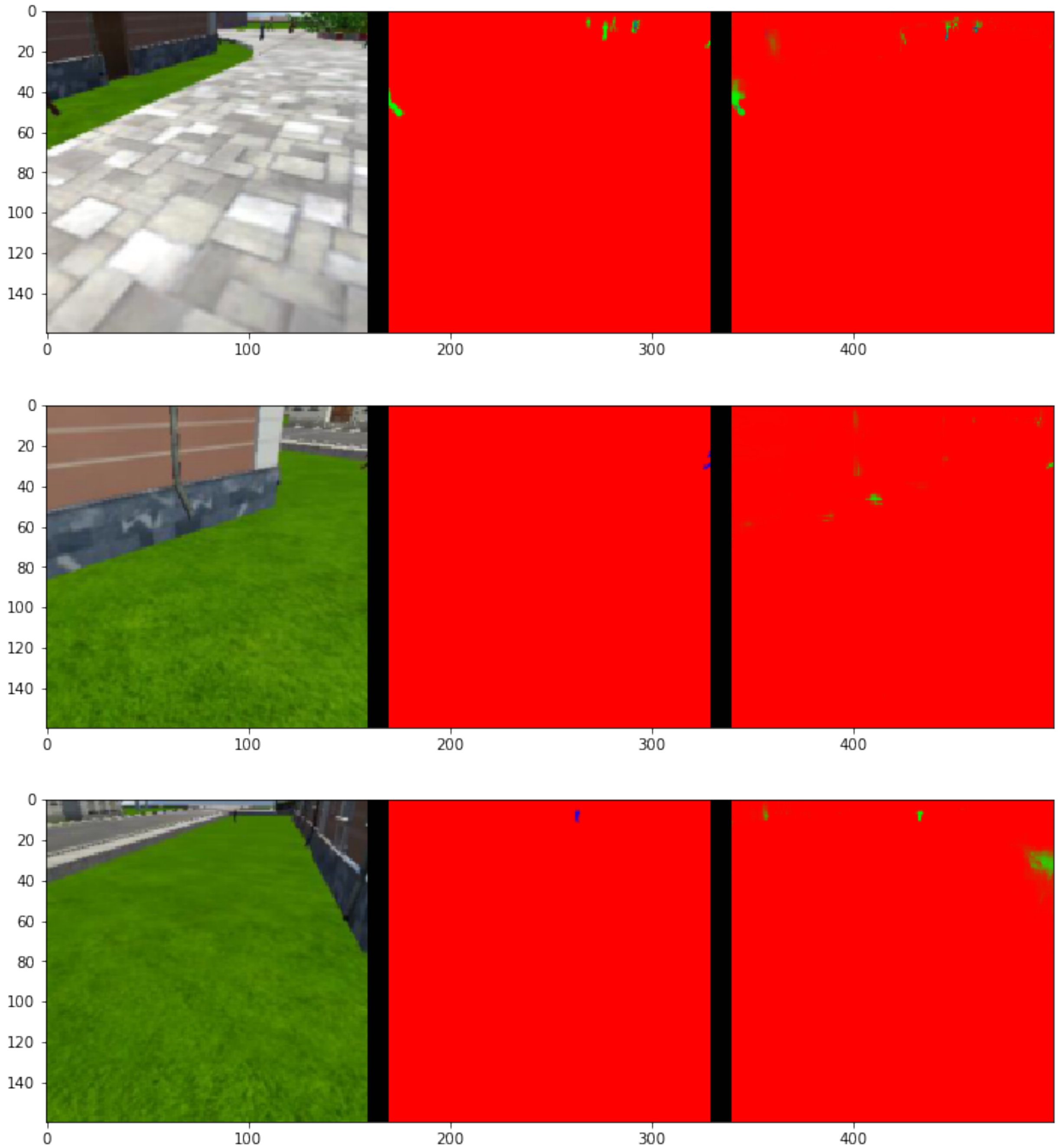






In [15]:

```
# images while at patrol with target
im_files = plotting_tools.get_im_file_sample('sample_evaluation_data',
'patrol_with_targ', run_num)
for i in range(3):
    im_tuple = plotting_tools.load_images(im_files[i])
    plotting_tools.show_images(im_tuple)
```



Evaluation

Evaluate your model! The following cells include several different scores to help you evaluate your model under the different conditions discussed during the Prediction step.

In [16]:
```
# Scores for while the quad is following behind the target.
true_pos1, false_pos1, false_neg1, iou1 = scoring_utils.score_run_iou(
val_following, pred_following)
```

```
number of validation samples intersection over the union evaulated o
n 542
average intersection over union for background is 0.9945721842931567
average intersection over union for other people is 0.29802721156424
944
average intersection over union for the hero is 0.8897191676744332
number true positives: 539, number false positives: 0, number false
negatives: 0
```

In [17]:
```
# Scores for images while the quad is on patrol and the target is not
visable
true_pos2, false_pos2, false_neg2, iou2 = scoring_utils.score_run_iou(
val_no_targ, pred_no_targ)
```

```
number of validation samples intersection over the union evaulated o
n 270
average intersection over union for background is 0.9841569551047916
average intersection over union for other people is 0.67160431027854
66
average intersection over union for the hero is 0.0
number true positives: 0, number false positives: 108, number false
negatives: 0
```

In [18]:
```
# This score measures how well the neural network can detect the targe
t from far away
true_pos3, false_pos3, false_neg3, iou3 = scoring_utils.score_run_iou(
val_with_targ, pred_with_targ)
```

```
number of validation samples intersection over the union evaulated o
n 322
average intersection over union for background is 0.9954061608049553
average intersection over union for other people is 0.38558296997191
444
average intersection over union for the hero is 0.2174408739725172
number true positives: 160, number false positives: 2, number false
negatives: 141
```

In [19]:
```
# Sum all the true positives, etc from the three datasets to get a wei
ght for the score
true_pos = true_pos1 + true_pos2 + true_pos3
false_pos = false_pos1 + false_pos2 + false_pos3
false_neg = false_neg1 + false_neg2 + false_neg3

weight = true_pos/(true_pos+false_neg+false_pos)
print(weight)
```

```
0.7357894736842105
```

In [20]:
```
# The IoU for the dataset that never includes the hero is excluded fro
m grading
final_IoU = (iou1 + iou3)/2
print(final_IoU)
```

```
0.553580020823
```

In [21]:
```
# And the final grade score is
final_score = final_IoU * weight
print(final_score)
```

```
0.407318352164
```

In [ ]:
```

```