Task 2: Basics of SSD object detector.

1 . The form of the target values are presented in .csv files in the form of “frame,xmin,xmax,ymin,ymax,class\_id”. The training dataset has more samples than the validation (1300 vs 541).

Training set is the dataset used to adjust the weights of neural networks. Its size is bigger than

validation set. While validation does not adjust the weights, it also has smaller size. It used to

monitor if you model is overfitting.

2. There are 7 convolutional blocks in the first convolutional part. Except the last layer comes without the MaxPooling2D, the first 6 blocks contain the following components:

* Conv2D
* BatchNormalization
* ELU
* MaxPooling2D

3. The model loss is a weighted sum between:

* Localization loss (e.g. Smooth L1) - The loss between the predicted box and the ground truth box.
* Confidence loss (e.g. Softmax) indicates how confident the object can be labeled for each class.

Task 3: YOLO for real time object detection

