

Multimedia Project SoSe 23

Assignment 5



Multimedia Computing Lab
Prof. Dr. Rainer Lienhart
Katja Ludwig
Julian Lorenz

Submission: May 31, 2023

Please take a look at assignment 0 on how to submit your work.

Exercise 5.1 Model

8 Points

In this assignment, you will implement a light version of the algorithm described in *SSD: Single Shot Multibox Detector*¹.

- (a) Create a class `class MmpNet` that inherits from `torch.nn.Module`. This class must contain all the layers defined in the following tasks. The constructor must have the following prototype: `def __init__(self, num_sizes, num_aspect_ratios)`
- (b) Use MobileNetV2 for the backbone. In your forward pass, only use MobileNet's features.
- (c) Implement the `def forward` method such that the backbone extracts the features which then flow through one or more additional layers. Make sure that your model is compatible with the label grid returned by the dataset. The last 4 dimensions of your model output must match your label grid. However, the output shape of your model and the label grid shape will not be exactly the same. Why is this the case?

Exercise 5.2 Training

5 Points

Write a function `def step(model, criterion, optimizer, img_batch, lbl_batch)` that processes a batch of data and executes one optimizer step. Return the calculated loss for the given batch as a `float`.

¹Liu, Wei, et al. SSD: Single Shot Multibox Detector. European conference on computer vision. Springer, Cham, 2016.

Use your data pipeline from assignment 4 to load batches of images with their associated label grid and pass the images through your model using the `step` function. Train your model for multiple epochs.

Exercise 5.3 Negative Mining

7 Points

Implement Negative Mining. Write a function

```
def get_random_sampling_mask(labels, neg_ratio)
```

that takes the labels as returned from your data loader and a desired ratio of negative examples and returns a PyTorch tensor. This tensor contains 0 or 1 and acts as a mask that selects only a few values from the loss. Hint: Think about what happens when you calculate `mask * loss`. Update the `step` function and include the new function. Explain the effect of random sampling.

Train your model again for multiple epochs with Negative Mining. Report and interpret your findings.