

Assignment 03

Simon Franken

2025-11-05

Course: Multimedia Project WiSe 2526

Exercise 3.1 Dataset Parsing

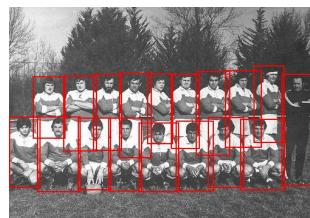


Figure 1: 02254418.jpg with 18 annotations

c)

Exercise 3.3 Training

Batch size	32
Training epoches	10
Loss	0.3719
Accuracy	78.90 %

The accuracy metric isn't ideal because it's quite intuitive to classify whether an object appears once or more in an image. In this case, the model might not have guessed the correct number of objects (e.g., 3 instead of 2), but the result will still be interpreted as accurate. This also makes it more challenging to backpropagate. A better approach would be to output the actual number of objects (in this case, humans).