Multimedia Project

SoSe 23

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

Assignment Submission 7

Jan Claar

Exercise 7.1

The images with the augmentations applied are shown in figures 1, 2, 3 and 4. The augmentations used are from the `torchvision.transforms.v2` API which supports bounding box data natively. The crop operation that is used is the `RandomIoUCrop`, which samples regions of the image with one of several

Exercise 7.2

The validation AP values for the different augmentation combinations are shown in figure 5. The combinations chosen are all combinations of augmentations from the previous exercise with at least one geometric augmentation enabled, since the geometric augmentations actually affect the bounding boxes in scale and position, which was thought to have the biggest effect on the data in comparison to the color augmentations.

As for the results, it seems that the crop augmentation has a detrimental effect on the final model performance, as the results of runs with crop enabled are markedly worse than the rest, while flip and solarize slightly increase the performance. This could be because the label grid is adapted to boxes in the full image size, which the scaling of the crop changes. This might prevent learning from the cropped patches.

Overall, the tested augmentation strategies don't improve the models performance much (blue baseline with no augmentations) which might indicate that the augmentation might not be strong enough or the model has already reached its maximum learning capacity in its current architecture.

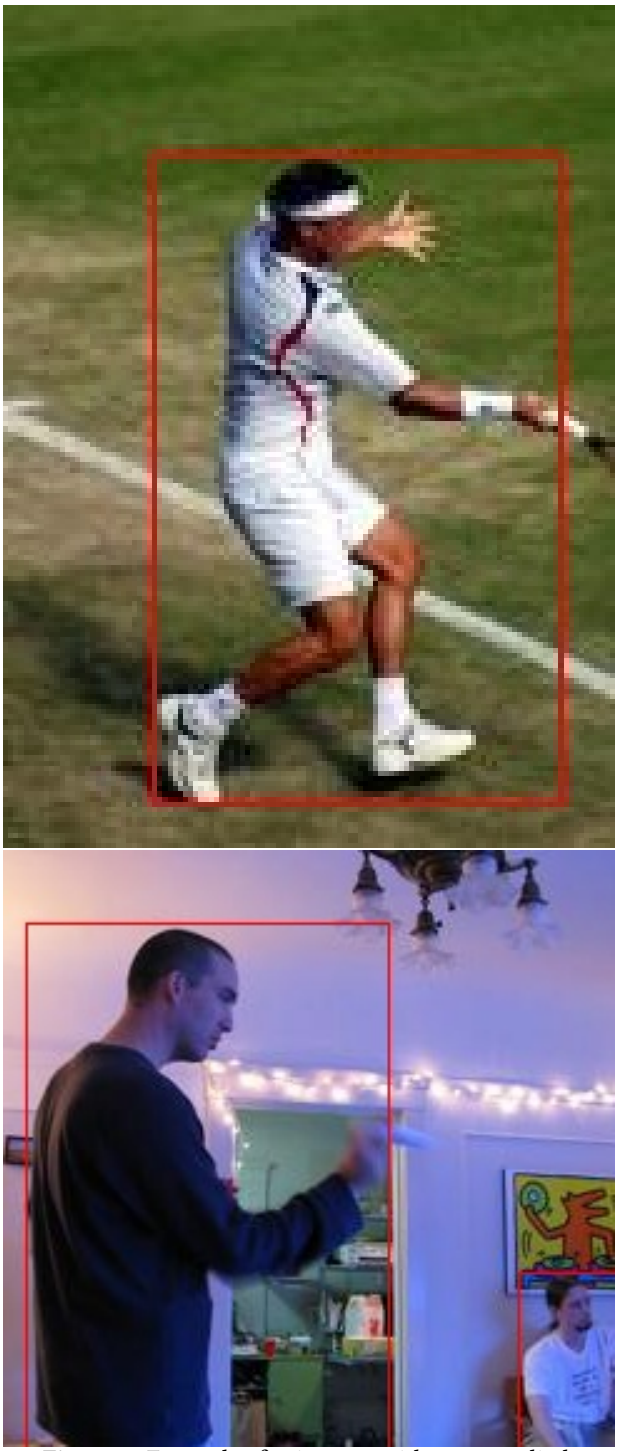

Figure 1: Examples for images with crop applied.

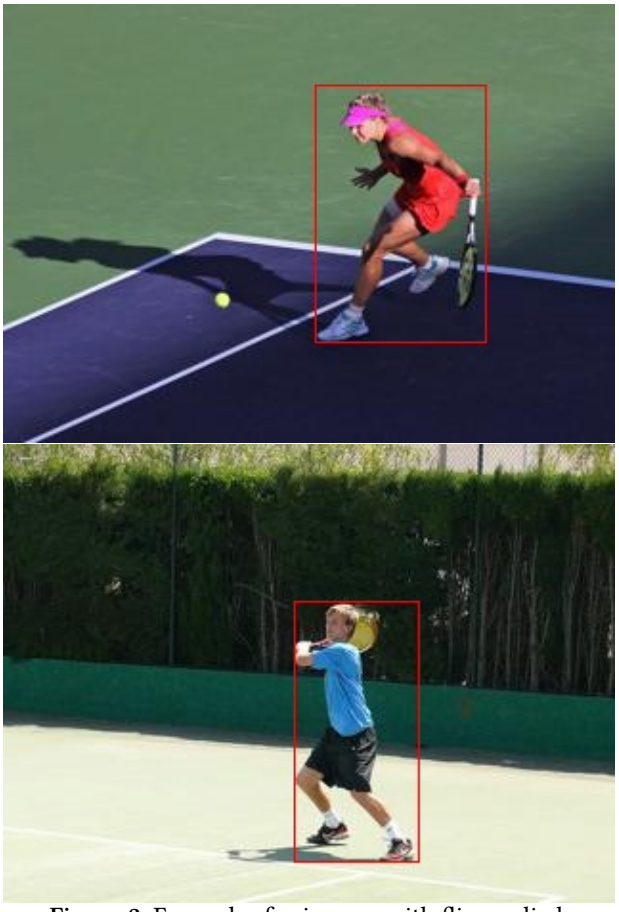

Figure 2: Examples for images with flip applied.

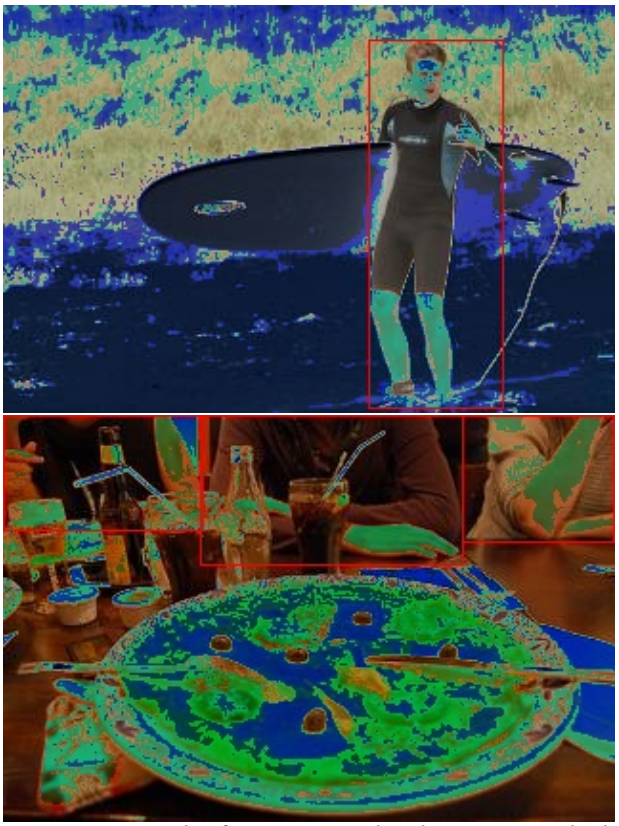

Figure 3: Examples for images with solarization applied.

Figure 4: Examples for images with gaussian blur applied.

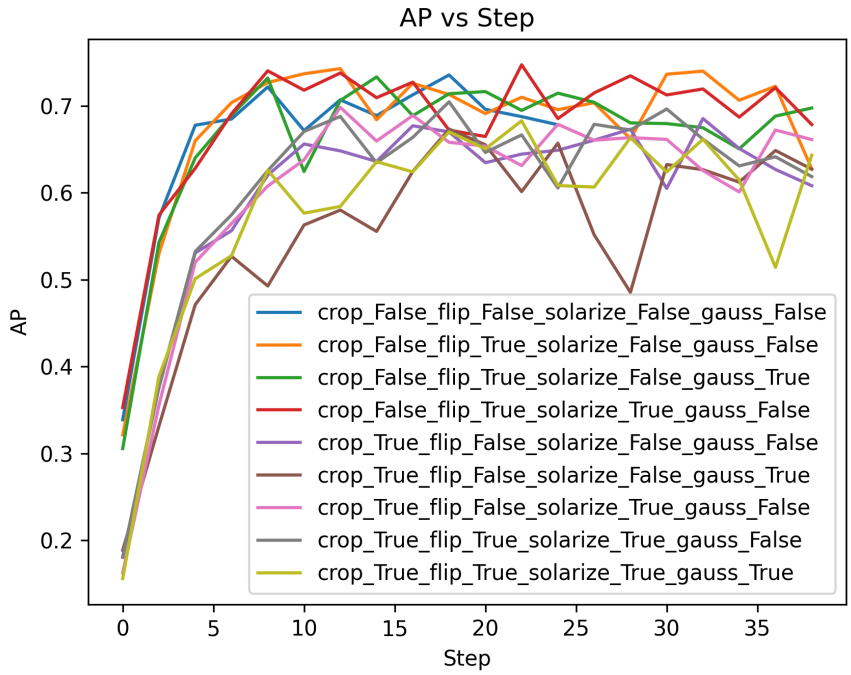

Figure 5: AP vs Step for the different augmentations combinations.