

PEER REVIEW – ‘EFFECTS OF DIFFERENT ARRANGEMENTS IN VISUAL INPUT DATA ON OBJECT DETECTION ACCURACY’

1. Content

The paper ‘Effects of different arrangements in visual input data on object detection accuracy’ written by Ahmad Hashir and Rifa Khan describes different types of data arrangements for RGB images. Hashir and Khan describe classical and advanced methods, the AutoAugment algorithm and fully rendered data sets. [1]

They classify geometric methods which change positions of pixels like flipping, rotating and cropping and photometric methods which change the intensity of some pixels like color jittering, edge enhancement and fancy PCA as classical methods. [1-II]

In addition to classical methods they mentioned some advanced methods: Random Erasing which selects rectangle-regions in an image and changes the pixel’s intensity values in this region to random numbers. Hide-and Seek which randomly hides parts of each image and Patch Gaussian that adds random noise to specific part of the images. [1-III]

Moreover, they describe AutoAugment which is a method to find automatically the right augmentation method and introduce fully virtual rendered data like for example the VKITTI dataset. Finally, they conclude that the methods they describe are able to improve data-sets for learning networks. [1-IV, V, Conclusion]

Special to this paper is the general and widely description of different methods for data augmentation, which includes a wide range of different methods starting with classical methods and ending with fully rendered data sets. Moreover, it is also exceptional that the paper refers to the application of autonomous driving many times.

2. Evaluation and possible improvements

2.1 References

In general, Hashir and Khan gave most of the references correct and complete. Only for a few network architectures like ‘AlexNet’ and ‘GoogleLeNet’ in section III-B and the tables I, III, VI and VII the references should be added. In addition to that, some of the arguments should be empowered by adding references. For example, the first paragraph in Section 1 claims that the problem of autonomous driving could be converged to one specific problem of computer vision. Such a strong claim should be supported with references.

2.2 Conclusions

For the final conclusion that the methods shown in the paper are advisable and convenient to improve on datasets available for learning networks, Hashir and Khan provide a lot of evidence in Sections II to IV.

In general, the paper focuses on describing different data augmentation techniques and could benefit from more conclusions and analysis. For example, in section II-C, III-A, IV- A, IV-B only tables with results are added but no analysis of the results is given.

In the beginning of section III Hashir and Khan conclude that classical methods do not lead significant improvements. That could be supported with a longer reasoning. In Table I for example the Top-5 Accuracy improves by about 15 % when cropping is applied. It is not clear why this is not significant.

In the section ‘Conclusion’ a recommendation of one of the compared methods or a guideline to choose the right method for data augmentation could be added.

2.3 Figures

The figures used in the paper are all easy to read and positioned close to the text passages which belong to them. Most of the figures are described considerably and are thereby self-explaining. Just for figure 3 and 4 a longer explanation would be helpful.

2.4 Title and Structure

Hashir and Khan structured their paper in a very clear and consistent way. They classified the augmentation methods into classical, advanced method and mentioned AutoAugment and augmented reality as special sections. Starting with classical methods and going on with advanced methods, AutoAugment and augmented reality seems to be very senseful and makes the paper understandable.

Since every subsection ends with the influence of the method on the accuracy ‘Effects of different arrangements in visual input data on object detection accuracy’ is a feasible title.

3. Conclusion

Overall Hashir and Khan come up with a detailed description of different data augmentation techniques and composed their paper predominantly in a good habit of scientific work. Some more analysis and conclusion could be beneficial for the paper.

REFERENCES

- [1] Effects on different arrangements in visual input data on object detection accuracy.
Ahmad Hashir, Rifa Khan. 2019.