Data Generation Methods for Multi-Label Images of LEGO Bricks

Author: Berend Kam , Supervisor: Attila Lengyel, Responsible Professor: Jan van Gemert / 25-06-2020 / CSE3000 Research Project

1. CONTRIBUTION

(a) Real data (b) Rendered data (c) Cut&paste data

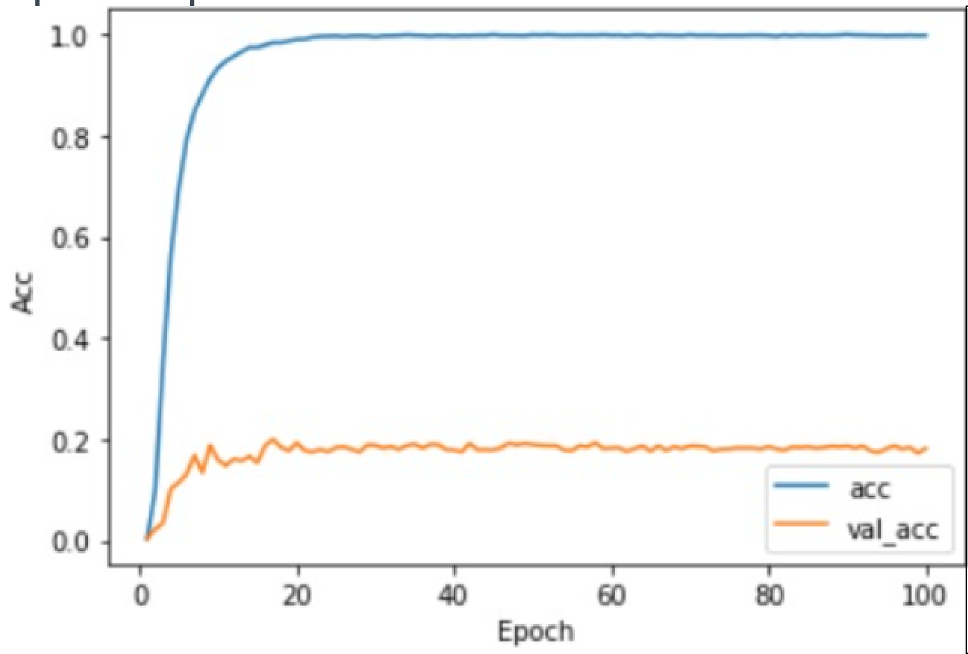
We introduce three and compare differently generated datasets of LEGO bricks to be used for computer vision purposes: one traditionally generated dataset and two synthetically generated datasets.

4. SPEEDUPS

To speed up 'real' data collection we:

1. Photograph a selection of LEGO bricks multiple times, shuffling the bricks in between
2. Create a new selection of bricks by adding or removing a single brick . This can be done up to 10 times

This adds imbalances to the dataset which decrease relative with an increase in data. We investigate the affect of these speedups on multi-label classification accuracy during training.



Training accuracy

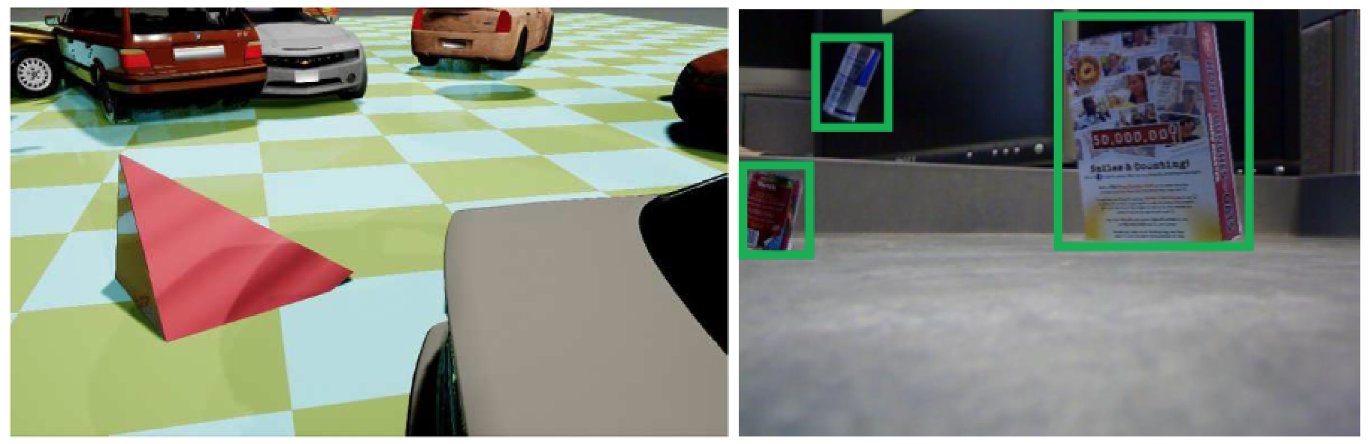
The validation set was selected on selections of LEGO bricks the classifier had never seen before. Low F1 validation accuracy indicates the network overfit.

2. BACKGROUND

Could an image classifier detect all bricks within a pile of LEGO bricks? A first step in to uncovering this answer is the creation of a multi-label image dataset of LEGO bricks to gain a sense of feasibility. However, data is one of computer vision's greatest constrictions; it strongly relies on many man hours to collect images and many more to annotate them. Image data can be generated synthetically as an alternative to manual data collection. While less realistic, synthetic data is much cheaper to collect and annotate than real data. Past research has been able to accurately train networks using synthetic data to identify real data [1,2]. These synthetic methods are:

(a) Rendering scenes of objects

(b) Creating photoshopped 'cut&paste' scenes of objects



(a) rendered traffic scene[1] (b) cut&paste scene[2]

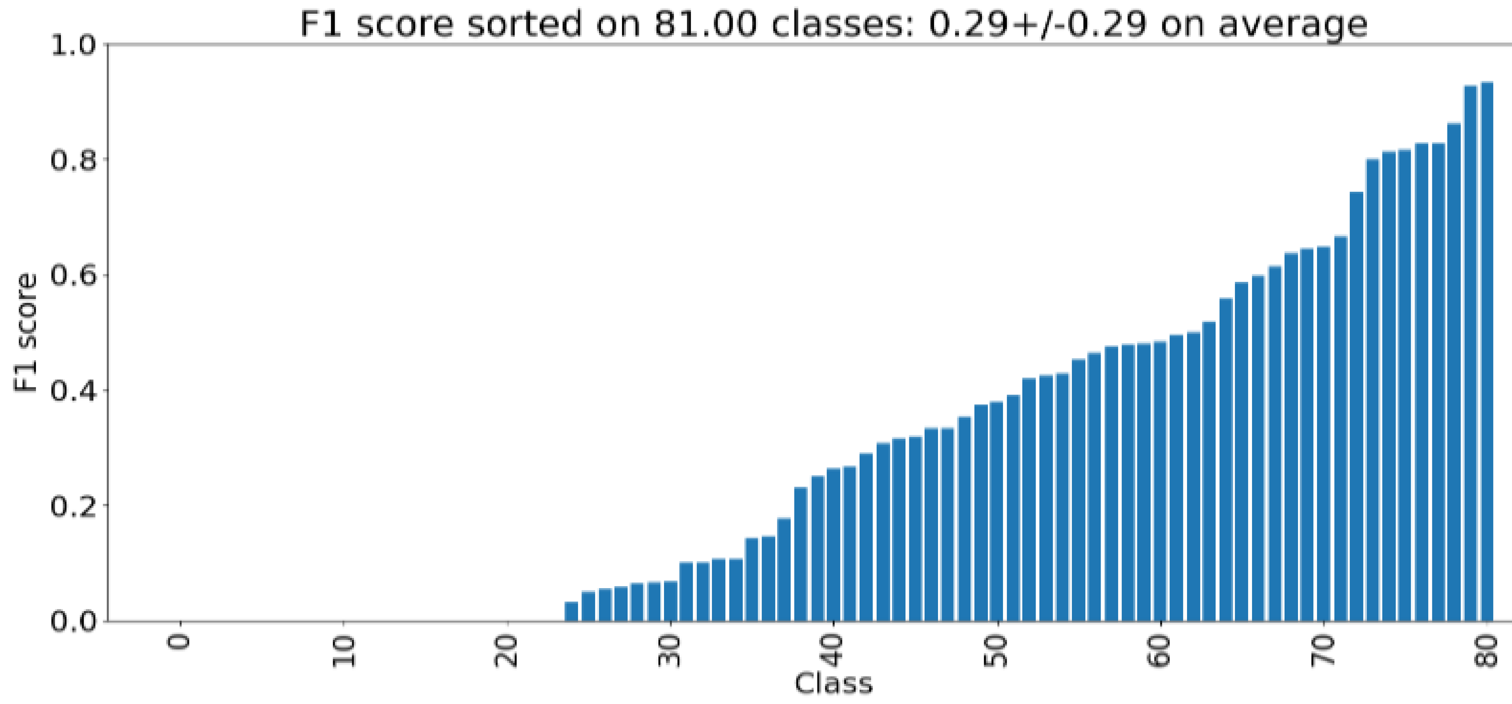
5. CONCLUSION

We find that all datasets should be improved further for practical applications. We suspect:

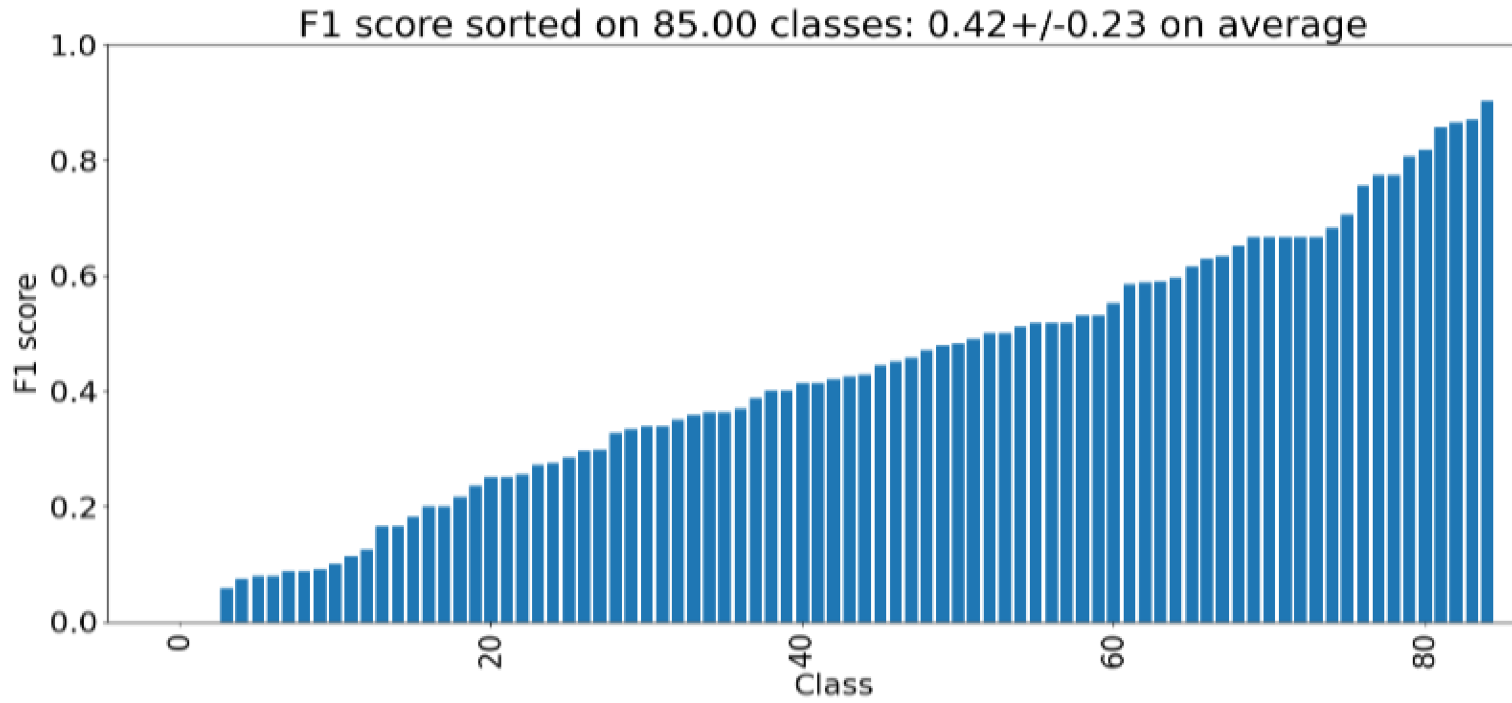
1. The real data should be expanded or diversified
2. The rendered data would benefit most from more realistic 3D models of bricks as well as more diverse lighting conditions
3. The cut&paste data would benefit most from more diverse angle coverage and improved 'blending' of objects into scenes.

Finally, future works may consider adding more bricks per image.

3. DOMAIN SHIFT



(a) Rendered5000



(b) CP5000

F1 scores of a multi-label classifying network trained on two random selections of 5000 synthetic images tested on 300 real images. F1 scores combine precision and recall.

We investigate the domain shifts between different datasets. While some classes of brick are not affected by the domain shift, others are. Results indicate that cut&paste data is generally affected less.

6. WORKS CITED

1. J. Tremblay, A. Prakash, D. Acuna, M. Brophy, V. Jam-pani, C. Anil, T. To, E. Cameracci, S. Boochoon, andS. Birchfield, “Training deep networks with syntheticdata: Bridging the reality gap by domain randomiza-tion,”CoRR, vol. abs/1804.06516, 2018.
2. D. Dwibedi, I. Misra, and M. Hebert, “Cut, paste andlearn: Surprisingly easy synthesis for instance detec-tion,”CoRR, vol. abs/1708.01642, 2017.