HaarCNN: Detection and Recognition of Comic Strip Characters with Deep Learning and Cascade Classifiers

Bartlomiej Kotlicki

b.m.kotlicki@student.tudelft.nl

Faculty of Electrical Engineering, Mathematics and Computer Science, TU Delft

Research Project CSE3000

Supervisors: Lydia Chen, Zilong Zhao

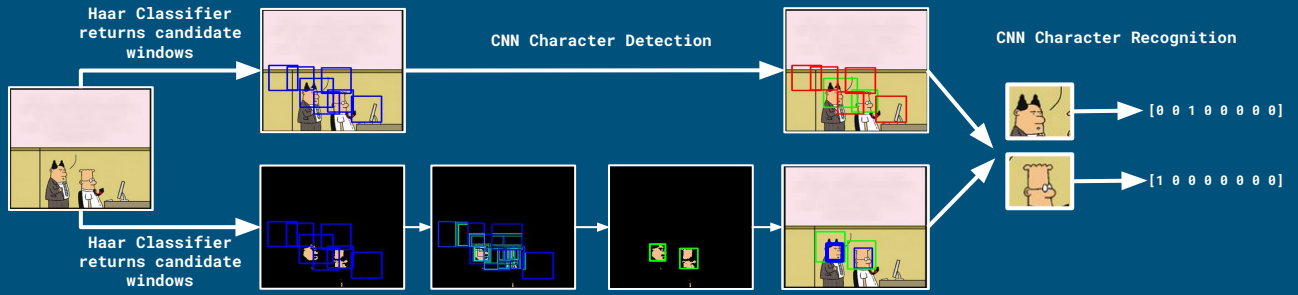
TU Delft

PROBLEM

- annotation of comic strips with characters present in them
- rapid character detection
- character recognition
- achieve low computational overhead

SOLUTIONS

- Haar Feature-Based Cascade Classifier[1] for rapidly providing candidate windows
- Convolutional Neural Networks[2] (CNNs) for conclusive character detection & recognition
- Selective Search[3] for providing more precise regions of interest



Selective Search of human characters performed on candidate windows + Non-Maximum Suppression[4] (NMS) + CNN Character Detection

RESULTS

Approach	HaarCNN	HaarCNN + Selective Search (H)	HaarCNN + Selective Search (HD)	Selective Search on the full image (H)	Sliding window + NMS + pyramid of scaled images
Precision	0.88	0.92	0.88	0.89	0.60
Recall	0.78	0.74	0.79	0.74	0.80
Inference time (seconds)	199	200	221	755	10457

Table 1: results of running different types of a pipeline on the first set 750 images shown in first three columns. The other two show baselines we compared the pipeline to. (H) -human-like character, (D) - white dog character.

Set of 750 images	Portion of eight main characters recognized
#1	0.88
#2	0.87

Table 2: results of recognition of eight main characters annotated in the extracted images. Experiment ran on two sets of 750 images.

CONCLUSION

- both HaarCNN alternatives led to fast inference times compared to other baselines
- achieved precision of ~90%, recall of ~80% in character detection and ~88% of eight main characters correctly recognized
- given more characters to be recognized with specialized convolutional neural networks in the HaarCNN with Selective Search step approach, the recall of characters detected increased

REFERENCES

[1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001,vol. 1, pp. I–I, IEEE, 2001.

[2] I. Goodfellow, Y. Bengio, and A. Courville,Deep Learning.MIT Press, 2016. http://www.deeplearningbook.org

[3] J. R. Uijlings, K. E. Van De Sande, T. Gevers, and A. W.Smeulders, "Selective search for object recognition,"International journal of computer vision, vol. 104, no. 2,pp. 154–171, 2013.

[4]S.K.,"Non-maximumsuppression(nms)."https://towardsdatascience.com/non-maximum-suppression-nms-93ce178e177c.