IS IT A BIRD? IS IT A PLANE? IT'S...

AN OVERVIEW OVER OBJECT RECOGNITION PROBLEMS AND ADVANCES.

Lotte Steenbrink
Hamburg University of Applied Sciences, Dept. Computer Science,
Berliner Tor 7
20099 Hamburg, Germany
lotte.steenbrink@haw-hamburg.de

ABSTRACT

Recognizing an categorizing objects in an image is one of the problems which are harder to solve for computers than for humans. However, image recognition is imperative to make human-computer interaction more natural and improve the way information contained in images is stored and handled. In recent years, the field has seen a lot of progress. My homework would aim to provide an overview over the challenges faced by image recognition software, and available solutions as well as their limitations.

1. SOURCES

There are plenty of sources discussing various aspects of object or image recognition. Below, three key papers will be discussed, each covering a different aspect. The first source provides an overview over the entire topic. The second one discusses recent, state-of-the-art advances in object recognition. the third offers an outlook into the possibilities of interdisciplinary applications of object recognition.

1.1. Overview

[1] provides an overview over the field of computer vision, including the basics of object recognition and categorization as well as current strategies for object recognition.

1.2. Recent advances in object categorization

[2] discusses the use of convolutional neural networks for object recognition. Its results are superior to most other research and can be considered state-of-the-art.

1.3. Outlook

[3] connects recent advances in computer vision and natural language processing in order to produce natural-language sentences describing the content of an image.

2. REFERENCES

[1] D. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Always learning. Pearson, 2012.

[2] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, Eds., pp. 1097–1105. Curran Associates, Inc., 2012.

[3] Oriol (Google) Vinyals, Alexander (Google) Toshev, Samy (Google) Bengio, and Dumitru (Google) Erhan, "Show and Tell: A Neural Image Caption Generator," 2014.