

Synthesising Abstract Deep Learning Training Data

Søren Winkel Holm, s183911

December 16, 2022

Contents

1	Introduction	2
2	Method	2
2.1	Synthetic Scene Rendering	2
2.2	Deep Image Segmentation	2
3	Results	3
4	Discussion	3

1 Introduction

In the field of computer vision, deep learning is often used to solve problems such as object detection and image segmentation. Two main problems are (1) the high cost of annotating the image by producing the label maps (such as this) (2) the high cost of training the deep learning models. Automatic generation of scenes with corresponding label maps might be usable even if the scenes are abstract. Such data could be used to pretrain the deep learning model and act as an inductive bias such that less training and less data is required when actually training on the real dataset. For this to work, the generated scenes should have objects with object having some assigned class which determines some visual characteristics of the object. The scenes do not have to have anything to do with the actual task but need to make the model learn some basics of object detection, classification and segmentation. An easy first solution would be having the class determine the colour of the object but it is much more realistic if the objects are textured with noise functions and the parameters of the procedurally generated noise were determined by the class.

2 Method

2.1 Synthetic Scene Rendering

Noise Functions

The Simple Scene

Random Scenes

The Full Scene With Label Maps

2.2 Deep Image Segmentation

Baseline Task

Addition of Synthetic Pretraining Data

3 Results

4 Discussion

1. Powerful benchmark for different methods generalization towards different difficulties: Explainability