Heatmap-based Object Detection and Tracking with a Fully Convolutional Neural Network

Matura Paper by Fabian Amherd and Elias Rodriguez

BETREUUNG: M.BUCHER UND C.LATKOCZY

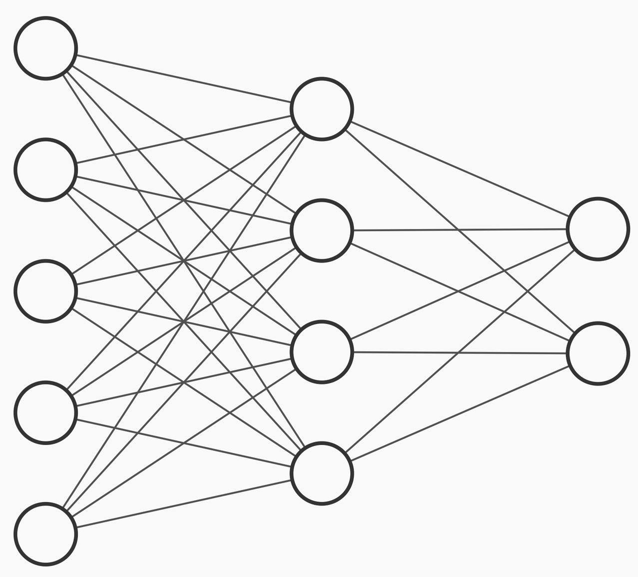
Kantonsschule Stadelhofen

Abstract

The main topic of this project was to provide an overview of the field of Artificial Intelligence. The core of the paper was a practical implementation of an algorithm for object detection and tracking. The ability to detect and track fast-moving objects is crucial for various applications of Artificial Intelligence like autonomous driving, ball tracking in sports, robotics or object counting. As part of this paper the Fully Convolutional Neural Network "CueNet" was developed. It detects and tracks the cueball on a labyrinth game robustly and reliably. While CueNet V1 has a single input image, the approach with CueNet V2 was to take three consecutive 240 x 180-pixel images as an input and transform them into a probability heatmap for the cueball's location. The network was tested with a separate video that contained all sorts of distractions to test its robustness. When confronted with our testing data, CueNet V1 predicted the correct cueball location in 99.6% of all frames, while CueNet V2 had 99.8% accuracy (See "Results" section for more details).

Method

To tackle this problem, we started to create a large data set. We took many images of the maze while the ball was rolling. We labeled these images or rather started to manually collect the position of the ball. The individual images and the corresponding label were added together as a pair to the dataset. With the help of data augmentation, we were able to artificially increase the size of our dataset many times over. With this data we fed our designed network, which takes three (one real time, the other two from the past) images as input with which it should learn something about the rolling behavior of a ball. Thus, in case of a bad image, it can try to estimate the position of the ball by the past two images. Using gradient descent and backpropagation, our network gradually approximated our data. Finally, we found a Local Minimum that gave us satisfactory results.



Motivation

AI could unravel to be the most important invention of humankind during our time or even of all time. Because of this and the rapid development during this decade it was an easy decision for us to choose AI as the subject of our Matura Paper. We believe that the future lies in AI and therefore it is worthwhile to have engaged with it. As potential computer science students we could imagine our career path unfold in the field of AI. So far only the tip of the iceberg has been scratched, the field of AI will take a big step in the direction of progress during our lifetimes and we strive to contribute to this development.

Experiment

The task for this matura paper was to develop a neural network to reliably detect and track the ball on the "Brio labyrinth game" in figure 1. The challenge of the labyrinth is to roll the ball to a destination without it falling into one of the holes on the board. The ball is moved by tilting the labyrinth with two knobs (one on each axis). In the context of this work we developed a fully convolutional neural network, with which the first step for a fully automated labyrinth game was developed. The data was collected and labelled by ourselves with a camera and suitable labelling software.

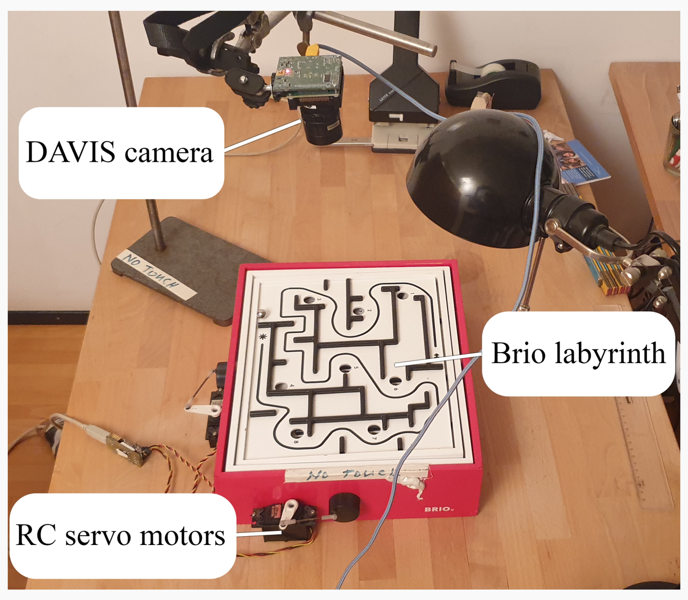
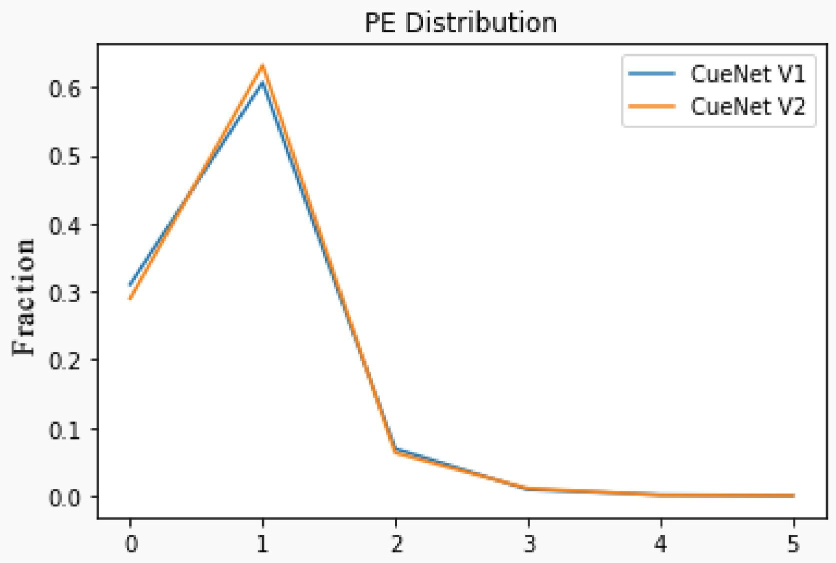


Figure 1: Experiment setup

Results



In this project we successfully developed a Fully Convolutional Neural Network named CueNet that can effectively detect and track the cueball on the used labyrinth. We also found that CueNet V2 outperforms CueNet V1, which indicates that using consecutive images as an input to the CNN increases its accuracy overall and especially in difficult situations.

Our Network Architecture

CueNet V2