Systems Engineering
Project Year 2019/20

KÄRNTEN University of Applied Sciences

VISION CONTROL

Design, Train and Test a Deep Learning Powered Image Regression Sensor

Goal

The goal is to build a computer vision sensor, that is able to predict the angle of a motor shaft out of images using Machine Learning (ML).
This task can be subdivided into 2 major steps:

- Acquisition of a large dataset
- Training of a Convolutional Neural Network (CNN)

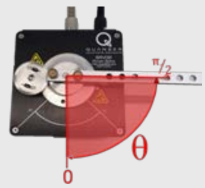


Figure 1: Servo Motor

Motivation

Machine Learning Algorithms learn abstract high-level patterns and relationships from datasets.
Benefits:

- Knowledge comes from Data, not the Engineer
- Very robust against noisy/changing inputs
- Able to generalize knowledge to data not seen during training

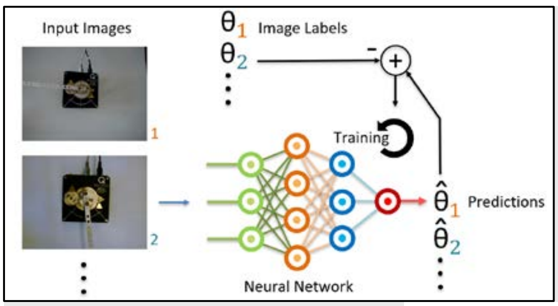


Figure 2: Training of CNN using Supervised Learning

Supervised Learning

Machine Learning provides computers with the ability to learn and improve from experience. In Supervised Learning, a sub-field of ML, this experience is provided by labelled training data, consisting of input-output pairs. Once a model has learned the desired relationship between in- and output (image -> angle) it should be able to generalize this knowledge to new examples, it had not seen during training.

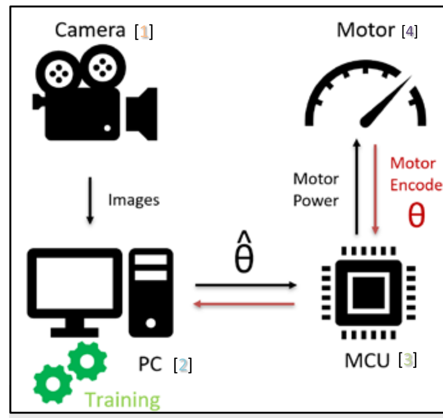


Figure 4: Schematic Setup – Data Acquisition and Sensor Predictions

Training

The training loop was implemented in TensorFlow 2.0 using its high-level API Keras. A CNN Architecture with 230.000 trainable parameters was used and trained on 80.000 images over 45 epochs, which corresponds to 3.600.000 training steps.

$$loss = \frac{1}{n} * (\arctan(\frac{\sin(y-\hat{y})}{\cos(y-\hat{y})})^2 + (y-\hat{y})^2)$$

[5]

Figure 3: Custom Loss Function

The model was optimized using Adaptive Momentum Optimization on a custom loss function, consisting of two loss-functions added together: Mean Squared Error and Mean Squared Angular Difference [5]

In order to improve the generalization abilities of the model, and to increase the effective amount of training data, the following augmentations were used:

height - / width - shifting, zooming, rotating, brightness shifting and a color transformation to grayscale.

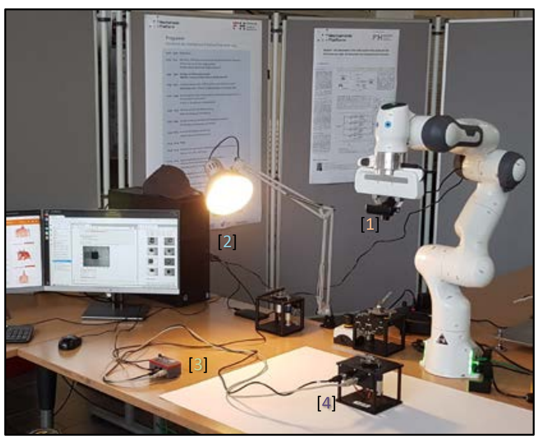
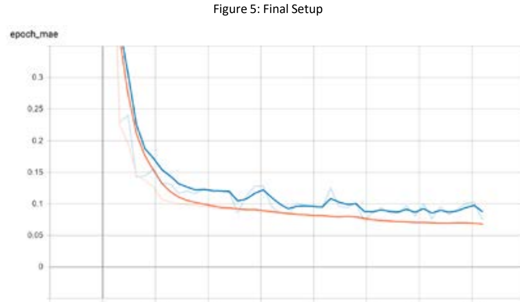


Figure 5: Final Setup

Automated Data Acquisition

- With the fully automated data acquisition system it was possible to save 100.000 labelled images in approximately 7h!
- Positional motor control using a PID control
- Communication of all parts allows synchronization and automation
- The robot arm moves the camera to get variation in perspective
- Multithreaded central script and distribution tasks allow for minimal acquisition times



metrics	train set	validation set	test set
custom loss [5]	0.0554	0.1055	0.0625
mean absolute error	0.0666	0.0747	0.0638

Figure 6: Training Results

Results

Having a MAE of only 0.067 radians, the final model converged to an accurate solution.
Tests with different motors than the one used for training and with different backgrounds and difficult lightening conditions, proofed the strong generalization capabilities of CNN models, as MAE did not increase further than 0.4 radians.

A big contribution to that success was our ability to automate the data acquisition, for which our application-focused engineering education can be credited for.

This strength of our university should be something to build upon in future projects to leverage more of Machine Learning strengths.

Team: Laurenz Hundgeburth
Gregor Fritz
Supervisors: FH-Prof. Dipl.-Ing. Dr. Wolfgang Werth
Heinz-Peter Liechtenecker, BSc MSc

Dataset available at:
https://github.com/LaurenzBeck/Vision-Control