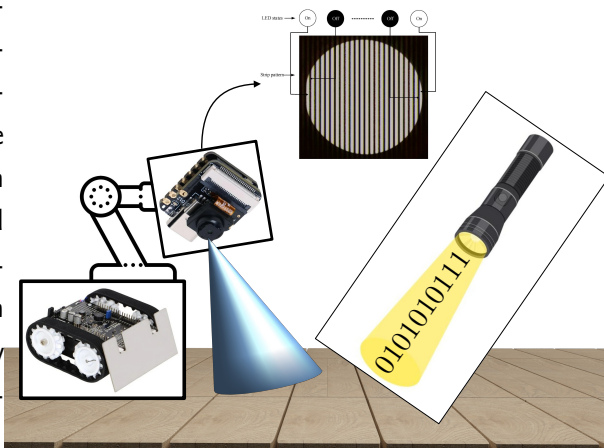


## Master Thesis Idea

### Title: Synthetic Data-Driven Learning for Optical Camera Communication Decoding

Optical Camera Communication (OCC) is an emerging wireless communication paradigm that leverages LEDs and camera sensors to transmit data. In rolling shutter-based OCC systems, information is encoded in high-frequency light modulation (e.g., On-Off Keying, OOK), which appears as stripe patterns across image rows. Decoding these patterns reliably is challenging due to noise, motion, exposure variations, and hardware limitations. Recently, Convolutional Neural Networks (CNNs) have shown strong potential for improving decoding robustness and reducing Bit Error Rate (BER), especially in constrained environments. However, training such models requires large labeled datasets where each image must be annotated with the corresponding transmitted bits. This labeling process is extremely time-consuming, error-prone, and difficult to scale, creating a major bottleneck for improving model performance and enabling deployment on embedded platforms.

This thesis addresses the challenge of reducing manual labeling effort in optical camera communication systems by exploring data-efficient learning techniques. In particular, the student will investigate the use of synthetic data generation, where realistic communication images are simulated with automatically generated ground-truth labels, and combine them with a limited set of labeled real images. The goal is to enable scalable dataset creation while maintaining or improving decoding accuracy and minimizing the bit error rate in practical systems.



### Goals:

- Design and implementation of a rolling shutter-based OCC simulation framework to generate realistic synthetic images with automatic ground-truth labels, including LED modulation and channel impairments.
- Development and training of a CNN-based decoder using both synthetic and real datasets, including pre-training and fine-tuning strategies with limited labeled data.
- Experimental evaluation of the trade-offs between labeling effort, dataset size, and decoding performance, with emphasis on Bit Error Rate (BER) and reduction of manual labeling.

### Contact:

- Juan Felipe Gutiérrez Gómez ([juan-felipe.gutierrez-gomez@tu-braunschweig.de](mailto:juan-felipe.gutierrez-gomez@tu-braunschweig.de))