Gottfried Wilhelm Leibniz Universität Hannover
Summerterm 2023

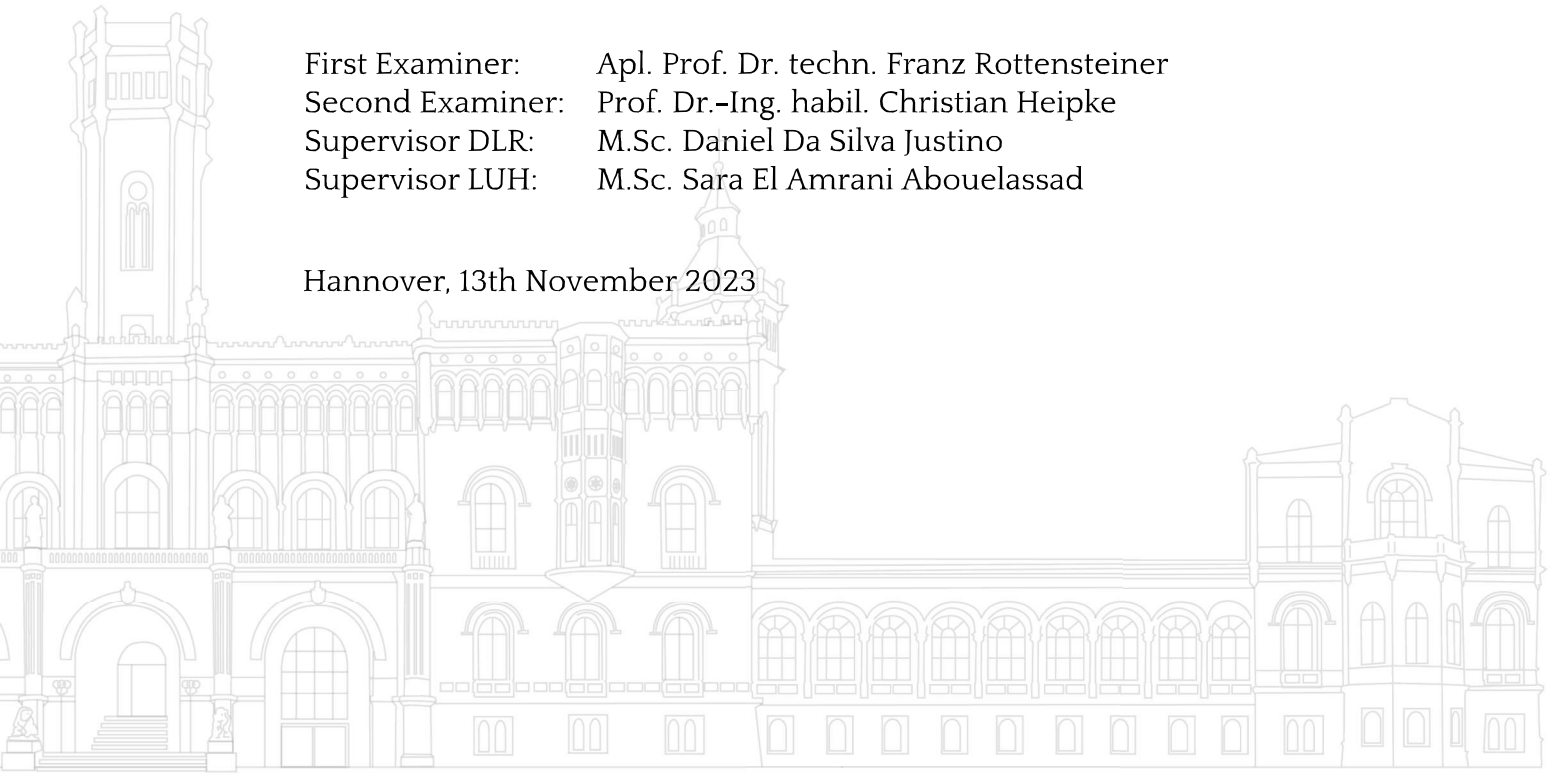
Estimation of the pose of a drone from RGB images acquired on board of an interceptor vehicle

Master Thesis
in the course Mechatronics and Robotics

Author: Stefan Kampmann

First Examiner:	Apl. Prof. Dr. techn. Franz Rottensteiner
Second Examiner:	Prof. Dr.-Ing. habil. Christian Heipke
Supervisor DLR:	M.Sc. Daniel Da Silva Justino
Supervisor LUH:	M.Sc. Sara El Amrani Abouelassad

Hannover, 13th November 2023



Abstract

Currently, consumer grade drones are widely utilized in modern conflicts. The double-use characteristic of this technology endangers the public space. As such, a safe solution for disposing a drone without imposing a risk to the immediate surroundings in mid-flight is required. Common counter unmanned aerial systems do not suffice for this task, as most lead to the nefarious drone being forcefully destroyed where it is, or landing right beneath where it is. Both options endanger the civilian space. At Deutsches Zentrum für Luft- und Raumfahrttechnik (DLR), one system is currently being researched which is supposed to catch target drones mid-flight and carry them off for safe disposal. For this maneuver, a $6D$ object pose prediction neural network is required enabling the grappling by providing accurate target pose estimation.
A network explicitly handling objects of this complexity with the pose-range and the size of the target drone does not exist within the specificity of the task. This thesis aims at finding a first approach to solving this problem by exploring the prediction head of GDRNPP, winner of the BOP2022 challenge, with a strictly depth-based simple feature map. Since there is no reference for this kind of feature map in combination with this type of prediction head, multiple feature map variants, network variants, and training strategies are tested.
The experimentation results show, that for horizontal and vertical translation prediction, as well as pitch angle prediction the results are reliably accurate. For the target depth, roll, and yaw prediction, however, the results are not accurate enough. The most promising results for translation were achieved with a combination of single-layered distance-based feature and weighted training, while the best rotational results were achieved by a three-layered depth-based feature map with the original PatchPnP architecture. However, the results were very close to each other in all experiments, a final answer cannot be given within the scope of this work.
This paper provides a first entry point into the topic within this specific research problem. Following works can explore the integration of a backbone, and further adapt the prediction head based on the provided experimentation results, ultimately resulting in a new end-to-end training routine for machine learning methods in drone pose estimation.