The first step of this research focuses on acquiring essential environmental data, including aerial imagery and the corresponding global coordinates of the camera at the time of image capture. This data is critical for converting local pixel positions in the images into global coordinates, a process necessary for accurate spatial analysis. Several camera systems, such as the Sequoia, RedEdge, Micasense, and DJI Zenmuse X7, were considered based on their technical specifications and the resources available for this study. After careful evaluation, the GoPro Hero 4 was chosen due to its ability to capture high-resolution RGB images at a high frame rate. Although this camera lacks near-infrared (NIR) and thermal imaging capabilities, its RGB imagery is well-suited to the image processing algorithms employed in this research. In addition to aerial images, global positioning data is acquired using a Ublox-Neo-6m GPS module, which transmits the camera’s location to the processing unit. To further streamline the system, the camera’s altitude is predefined, rather than measured and transmitted in real-time, allowing for a more straightforward setup without compromising the precision needed for the study.

In response to the need for live waypoint generation in agricultural fields, a hardware setup is required to provide the necessary environmental information. This system has the role of capturing and transmitting data to the software unit, where the processing will take place. In the context of this project, a lightweight camera is crucial to take high-resolution aerial images of the target field which will be the main data source for the image processing algorithm. Several cameras including Sequoia b9-b4-b7-b6, Micasense RedEdge b9-b14-, and DJI Zenmuse X7 b5 were utilized in similar projects. Considering the specifications of these cameras and crop row detection algorithms developed based on their aerial images, it was concluded that having high-resolution RGB images meets the needs of this project. Although the NIR band contains valuable information about the field and can be used for vegetation segmentation, they are too sensitive to environmental conditions such as temperature and might lead to the poor performance of the software. b-5. In conclusion, the Gopro HERO4 camera was acquired for this project. The specifications of this camera including its lightweight, high-quality RGB images, and durability make it a suitable choice for aerial imagery.

Besides the camera, a radio transmission system consisting of power supplies, a transmitter, and a receiver was developed to provide live data transfer between the mounted camera on the UAV and the computer with an image processing program.

Also, the last step of this project, which is the conversion of pixel coordination on the images to global coordination, requires the global coordination of the camera at the time of photo capturing as a reference. A Ublox-Neo-6m GPS module must be mounted on the camera UAV to provide this information.

This section focuses on the crucial task of crop detection within the captured images, which is essential for identifying crop rows and determining the path for subsequent operations. The process begins with a preprocessing step that converts the raw images into a format compatible with the models used for detection. Two crop detection methods, K-means and Unet, are then implemented, with a detailed explanation of each approach. The performance of these models is compared to select the most effective one for this application. Additionally, insights from previous studies using these algorithms inform the analysis.

The main contribution of this paper lies within this section, to develop an image-processing pipeline to determine the position of pixels which are located in the path between the crops. The Figure shows the procedure proposed in this paper. Initially, the aerial images captured with the camera are split into sub-images, and then during a semantic segmentation process, a binary mask is generated for each sub-image showing the positions of vegetation and background. In the next step, using Hough Transform, crop rows are detected. After defining the path line, a desired number of equally spaced points are selected on this line, and their position is recorded for later steps. Finally, the sub-images and their local waypoint coordinates are reconstructed and form the original input image.

High-resolution aerial images covering large areas are typically too large to use in machine learning algorithms directly. To address this, the images are initially split into equal-sized sections, making them manageable for algorithmic processing. The size of these sections is determined by the specific requirements of the algorithms used, ensuring compatibility and optimized performance. Maintaining uniform image sizes across all sections enhances processing speed and efficiency.