# Naslov:

Razpoznavanje predmetov na površini jedilne mize z uporabo RGB-D kamere

# Slika:

A picture containing colorful

Description automatically generated

# Abstract:

In this paper, we present a real-time method for 3D bounding box object detection and recognition in tabletop kitchen scenarios using RGBD cameras, with potential applications in robotics. Our approach combines color and depth information to accurately detect and classify a variety of kitchen objects. While the accuracy of our method is not the highest among all approaches, it is able to achieve good performance in a real-time setting, making it suitable for use in robotics applications. We evaluate the performance of our method on the SUN-RGBD dataset, a widely used benchmark for 3D object recognition, and demonstrate its real-time capabilities using Intel RealSense camera. The results of our experiments show that our method is able to achieve competitive accuracy while maintaining a high frame rate, making it a promising solution for object detection and recognition in robotics.

# Intruduction:

RGBD (Red, Green, Blue, Depth) sensors use a combination of traditional RGB (Red, Green, Blue) cameras, which capture colour information about an object, and depth sensors, which measure the distance from the sensor to the object. In the context of a dining table, RGBD sensors can be used to allow the robot to recognise and classify a variety of objects, including dishes, glasses, utensils and food items. By combining the colour information from the RGB cameras with the depth information from the depth sensors, the robot is able to create a detailed 3D model of the objects on the table, allowing it to understand their shape, size and orientation.

RGBD object detection and classification algorithms have a number of advantages over traditional object detection algorithms that rely solely on colour information. One advantage of using RGBD information is that the algorithm can better handle occlusion and clutter in the scene. Occlusion occurs when an object blocks the view of another object, and clutter refers to the presence of multiple objects in the scene that may be difficult to distinguish. By using depth information, the algorithm can better understand the spatial relationships between objects and more accurately detect and classify objects even when they are partially occluded or surrounded by clutter. Example of occlusion on a dining table is when a plate is partially hidden behind a glass. Another benefit of using RGBD information is that it enables the algorithm to better distinguish between objects that are similar in colour but different in shape or size. This can be particularly useful when there are multiple objects of the same colour in the scene, as the algorithm can use the depth information to distinguish them based on their shape and size. It can also help us distinguish between colourful tablecloths as a background and objects sitting on them.

However, using RGBD object detection is still less common than using RGB object detection. One reason is that RGB sensors are more common and less expensive than RGBD sensors. As a result, there are more and better quality RGB than RGBD data sets. This has partially changed with the appearance of commercial sensors such as Microsoft Kinect and Intel Real-Sense. In addition, algorithms using RGBD data often require more computing power and resources to analyse the data, which can complicate implementation in real-time applications or on resource-constrained devices. In addition, RGBD data can be noisy and inaccurate, especially in low light or at long distances, which can affect the performance of the object detection algorithm. This can make it difficult to achieve the same level of accuracy and robustness as with the RGB object detection algorithms. This can be partially corrected with indoor artificial light, but is difficult under natural conditions.

We developed an algorithm that uses an RGBD sensor to detect bounding boxes around objects sitting on a dining table in real time. The algorithm also classifies the objects within these bounding boxes and accurately identifies the type of each object. This system is designed to work effectively in a dynamic environment, such as a dining room, where items may be frequently added to or removed from the table. The ultimate goal of this project is to develop a system that can assist in tasks such as identifying and organising objects on the dining table, or providing a user with information about the table's contents.

**Kaj je goal naše raziskave?**

* **Zaznati objekte na površini mize**
* **Klasificirati objekte na površini mize**
* **Delovanje v realnem času**
* **Applikativnost v robotskih aplikacijah**

# Related Work:

Uporaba RGBD kamere:

* Uveljavljena je zaznava objektov z 2D kamero in CNN
* RGBD se je razvil z pojavitvijo cenejših prosto dostopnih kamer in posledično novih RGBD datasetov
* Amodal object detection (parts of objects occluded)

Metode zaznave:

* Tradicionalne metode
* Metode globokega učenja