Gebze Technical University

CSE 463

Introduction to Computer Vision

Assignment #2

Abdulsamet ASLAN

Project Overview

This project focuses on evaluating and comparing various feature extraction methods—**SIFT, ORB, and KAZE**—for object recognition within a dataset of images because they are not patented algorithms. The process involves loading and splitting the dataset into training and testing sets, applying each feature extraction method to detect keypoints and compute descriptors, and then using these descriptors for object recognition via the k-Nearest Neighbors (k-NN) matcher. The performance of each method is assessed and visualized through confusion matrices, providing insights into their accuracy and effectiveness in correctly identifying objects.

List of Objects

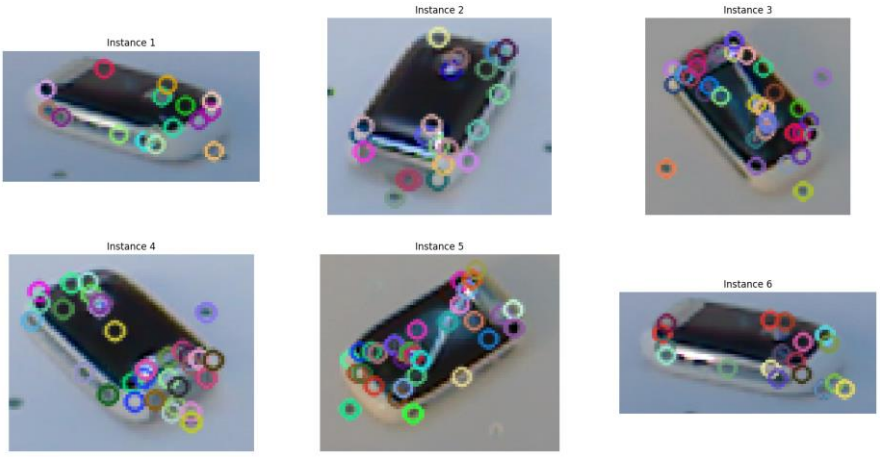
- The dataset, consisting of ten different everyday objects (**cell_phone_4, dry_battery_4, food_can_4, glue_stick_4, marker_4, notebook_4, sponge_4, stapler_4, toothbrush_4, and water_bottle_4**), is split into training (90%) and testing (10%) sets.

Extracting Keypoints and Descriptors

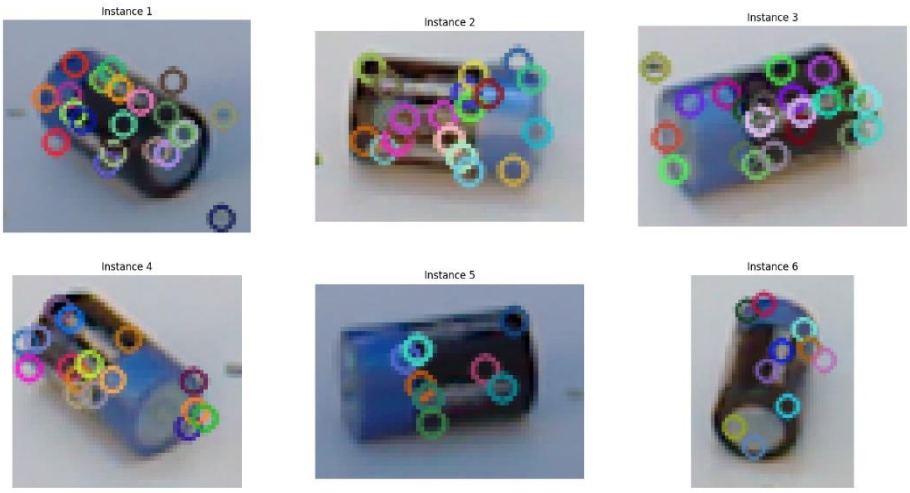
- Three feature extraction methods—SIFT (Scale-Invariant Feature Transform), ORB (Oriented FAST and Rotated BRIEF), and KAZE are applied separately to the training and testing datasets. Each method detects keypoints and computes descriptors for each image, returning a dictionary that includes these keypoints and descriptors.

- For visualization purposes, a subset of images with their keypoints is displayed using the Matplotlib library (`plt`) to demonstrate the feature detection capabilities of each method.

SIFT Keypoints for cell_phone_4



SIFT Keypoints for dry_battery_4



SIFT Keypoints for food_can_4

SIFT Keypoints for notebook_4

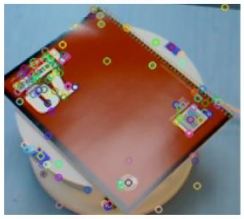
Instance 1

Instance 2

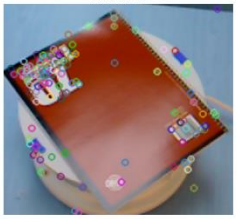
Instance 3

Instance 4

Instance 5



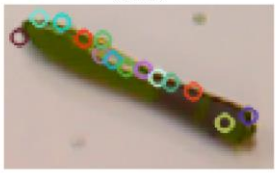
Instance 6



SIFT Keypoints for marker_4

Instance 1

Instance 2



Instance 3

Instance 4

Instance 5

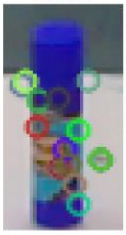
Instance 6

SIFT Keypoints for glue_stick_4

Instance 1

Instance 2

Instance 3

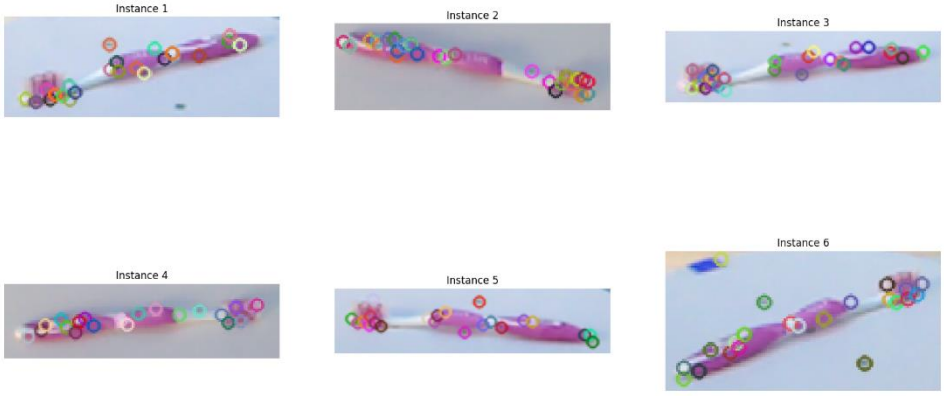


Instance 4

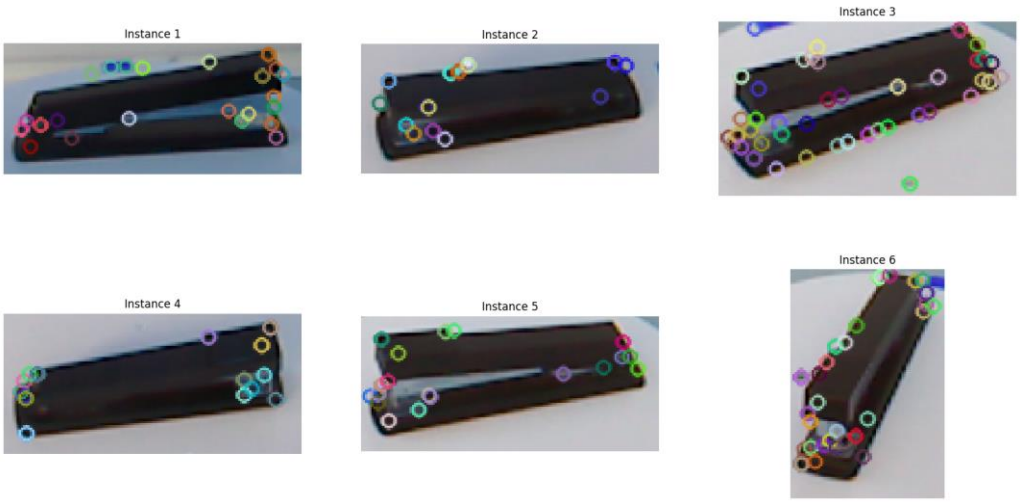
Instance 5

Instance 6

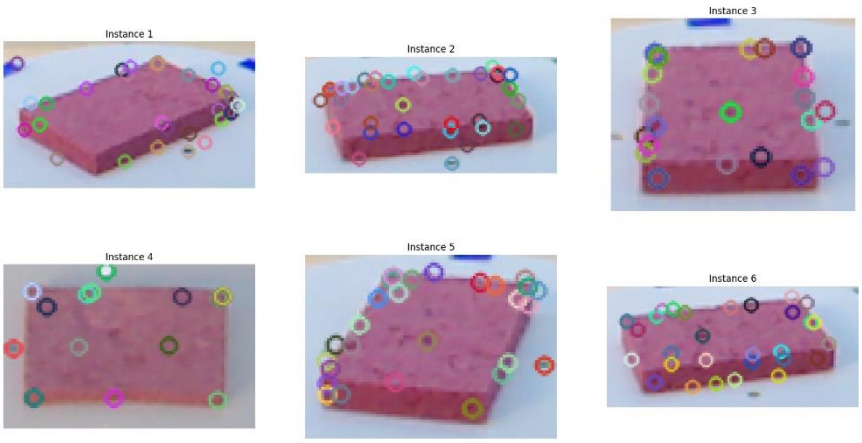
SIFT Keypoints for toothbrush_4



SIFT Keypoints for stapler_4



SIFT Keypoints for sponge_4



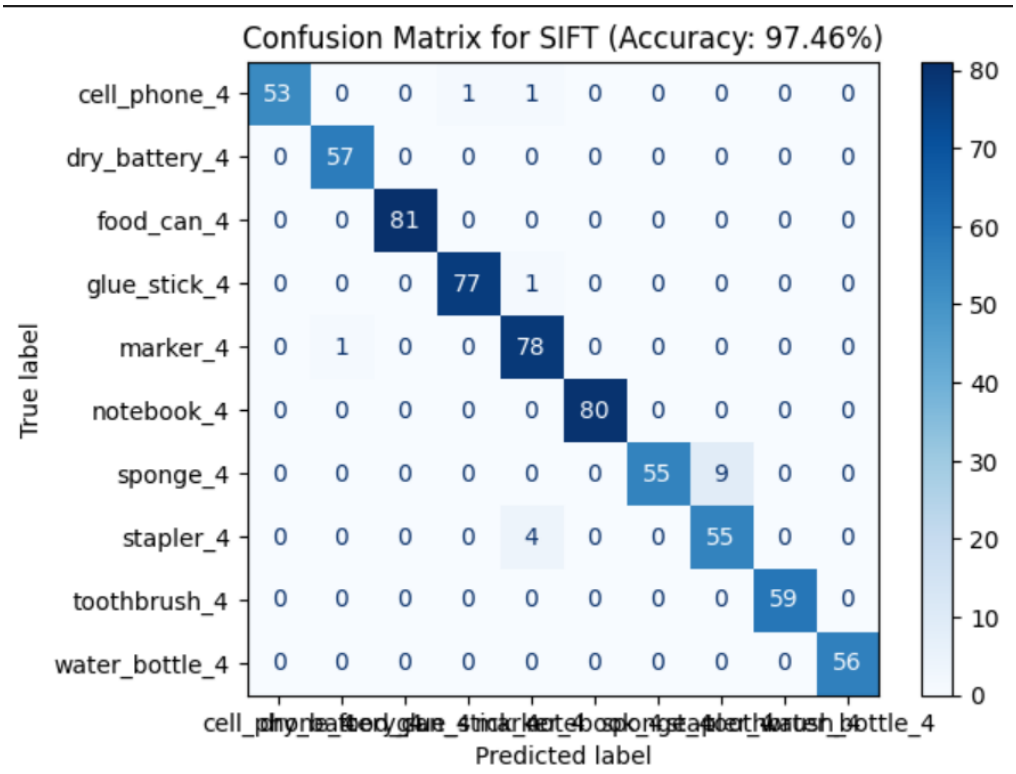
The details of object recognition

- The `recognize_objects` function identifies and classifies objects in the test dataset by comparing their descriptors to those in the training dataset using the k-Nearest Neighbors (k-NN) matching technique.
- It initializes a `BFMatcher` object, with the norm type based on the feature extraction method (SIFT, ORB, or KAZE). For each test image, the function finds the two nearest neighbors for each descriptor using `knnMatch` and applies a ratio test to retain good matches.
- It counts these matches for each training object category and determines the recognized object by identifying the category with the maximum number of good matches. The function collects actual and predicted labels for the test images to evaluate the performance of the recognition method.

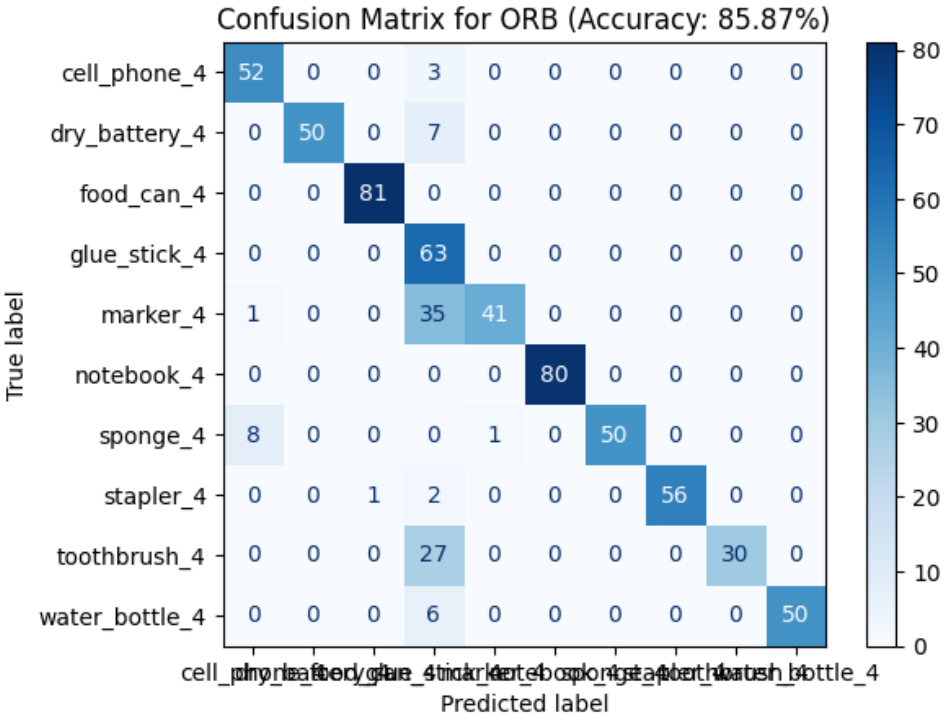
Results

- The performance of each method is evaluated using confusion matrices, which provide insights into the accuracy and effectiveness of the methods in recognizing and classifying the objects in the test dataset.

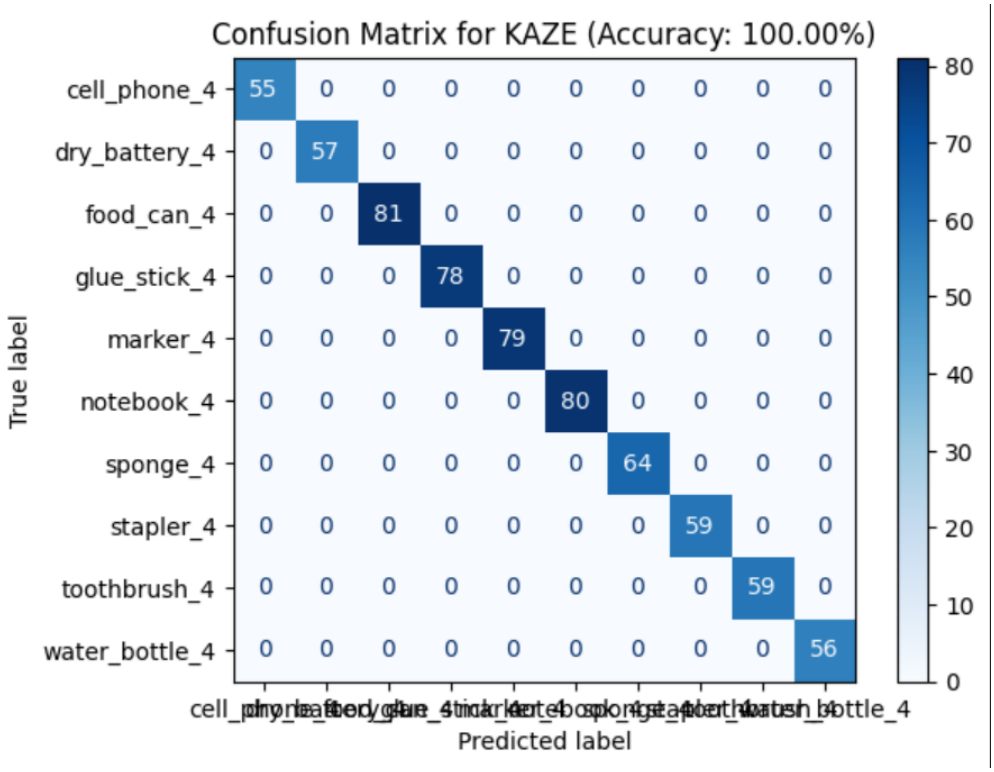
SIFT Result (244s)



ORB Result (238s)



KAZE Result (700s)



,

Analysis

1. KAZE algorithms take more time compared to SIFT and ORB primarily due to the computationally intensive process of building nonlinear scale spaces and performing nonlinear diffusion filtering. As it is seen in the confusion matrices, Kaze gives more accurate results. This results in more accurate and robust keypoints and descriptors, which can enhance the performance of object recognition tasks, as observed in this project. However, this improved performance comes at the cost of increased processing time.

2. The performance of SIFT and ORB are not similar. Both methods are capable of effectively detecting and describing features within the specific dataset and conditions used. However, ORB parameters are very hard to adjust for small size images that's why it gives less accuracy for small size images.

3. For applications where processing time is critical, methods like ORB or SIFT might be preferred despite potentially lower accuracy, while KAZE is better suited for scenarios where accuracy is more important than computational efficiency.

4. Wrong predictions, especially for similar objects like sponges, can be due to factors such as image size, image quality, or the similarity in object shape.