Experiment No: 9

Name:Mohit Raje

Roll number: 33

Aim: To Creating and Training an Object Detector

Objective: Bag of Words (BOW) in computer vision - Detecting cars in a scene

Theory:

Creating and Training an Object Detector

Object detection is a crucial task in computer vision, with applications ranging from autonomous driving to surveillance systems. In this experiment, we aim tocreate and train an object detector specifically for detecting cars in a given scene. One of the techniques we'll use for this purpose is the Bag of Words (BOW) model.

Bag of Words (BOW) in Computer Vision:

The Bag of Words model, borrowed from natural language processing, has beenadapted for use in computer vision. It is a popular image representation technique used for object detection and image classification.

The BOW model works as follows:

1. Feature Extraction: Extract local features, such as SIFT (Scale-Invariant Feature Transform) or ORB (Oriented FAST and Rotated BRIEF), from a set oftraining images containing the object of interest (in this case, cars).

2. Create a Vocabulary: Cluster the extracted features into a vocabulary ofvisual words using clustering techniques like K-means. These visual words represent common patterns or features found in the training images.

3. Histogram Representation: For each image, create a histogram that countsthe occurrences of visual words in the image. This histogram is known as the Bag of Words representation.

4. Train a Classifier: Train a machine learning classifier (e.g., SVM or Random Forest) on the Bag of Words representations of the training images, labelingthem as positive (contains the object) or negative (does not contain the object).

5. Object Detection: Apply the trained classifier to new, unseen images to detectthe object of interest.

Detecting Cars:

In our experiment, we will focus on detecting cars in a scene using the BOW model. We'll use a dataset of images containing cars and background scenes fortraining.

Example:

Let's take a look at a simplified example of how this process might work in Python:

Code:

```
import cv2
import os
import numpy as np


car_images = [cv2.imread(car) for car in os.listdir("car_dataset")]
car_features = [extract_features(image) for image in car_images]
vocabulary = create_vocabulary(car_features)
training_data = create_bow_representation(car_images, vocabulary)
classifier = train_classifier(training_data, labels)
test_image = cv2.imread('test_scene.jpg')
test_features = extract_features(test_image)
test_representation = create_bow_representation([test_image], vocabulary)
result = classifier.predict(test_representation)
if result == 'car':
    print('Car detected!')
else:
    print('No car detected.')
```

Input Image:

Output:

Car detected!

Conclusion:

The Bag of Words (BoW) model has been a significant tool in computer vision for detecting cars in scenes. BoW involves extracting features from images, creating a visual vocabulary by clustering these features, and representing images as histograms of visual word occurrences. In the context of car detection, this approach is particularly useful due to its ability to handle variations in car appearance, scale, and orientation. When combined with a classifier like SVM, BoW facilitates the identification of cars in complex scenes. However, the success of BoW largely depends on the quality of feature extraction and the size and diversity of the training dataset. In practice, more advanced techniques such as deep learning have gained prominence, but BoW remains a fundamental concept in computer vision.