بسم الله الرحمن الرحیم

University of Tehran
College of Engineering

Digital Image Processing

Instructor: Dr. Hamid Soltanian-Zadeh

Homework Assignment 10:

Object Recognition

Due date: TO BE SET

Corresponding TA:
Abolfazl Danayi

adanayidet@gmail.com

Spring 2025

Contents

1 Instructions

Please answer the following questions based on Chapter 12 (*Object Recognition*) of the textbook by Gonzalez and Woods. Submit your solutions by the due date. Read the following carefully and follow these instructions when submitting your answers:

Requirement	Description	Consideration
Standard Due Date	TO BE SET- 23:59	
Elearn HW Upload	**Only**	Only use ELearn to submit your homeworks.
Email Address	adanayidet@gmail.com	Feel free to say hello!
Submit format	Read Note 4	Extremely important
Late Submit	Penalty	$5 - 10\%$ penalty per day

Note 1: *This request is not intended to impose any hardship on you, but to help with better management of the class. Your cooperation is greatly appreciated in making the process more efficient. Thank you!*

Note 2: *We will also have a Telegram group for Q&A. The link (clickable) is provided here:* [open the link]

Note 3: *Use camera scanners if you wish to submit your handwritten answers. While readability is always important, it is especially critical in the DIP course, which focuses on quality enhancement.*

Note 4: *You must send a zip or rar file named* `DIP-HWx-std.no` *where x is the number of the homework (e.g.* `DIP-HW10-810199034`*). Inside the zipped file will be a folder for each question containing its code and output images. There should also be a single pdf file in the root folder as your main report.*

Academic Integrity: *Plagiarism, cheating, or using unauthorized external resources (including AI tools, solution manuals, or copying from peers) is strictly prohibited. All assignments will be checked for originality, and any violations will result in academic penalties. Please submit only your own work.* It's easy to detect whether your answers are generated by AI or are truly your own work, so be sure to submit original solutions.

2 Object Recognition via Template Matching

Background

Template matching is a fundamental technique in object recognition. It involves sliding a template image over a larger image to find regions that match the template. This method is useful in scenarios where the object's appearance is consistent and can be directly compared to a predefined template.

Note

You are NOT allowed to use direct template matching functions and libraries in this exercise. Please develop your own code.

Note 2

The target image provided in `images/q2/image.png` is RGB. Convert it into 1 channel grayscale and solve the problem in grayscale format.

(a) Extract a template. Choose ONE of the windows in the main image. Discuss your choice.

(b) Apply template matching using normalized cross-correlation to locate instances (multiple) of the template in the image.

(c) Display the original image with rectangles drawn around detected matches.

(d) Discuss the effectiveness of template matching in this context and any limitations observed.

3 Bayesian Classification: Two-Class Problem

Background

Bayesian classifiers utilize probability distributions to assign patterns to classes, aiming to minimize classification errors. Assuming Gaussian distributions simplifies the computation of class probabilities. In this exercise, you'll implement a Bayesian classifier to distinguish between two classes based on extracted features.

Image Description

The MNIST dataset consists of 28x28 pixel grayscale images of handwritten digits. For this task, you'll focus on images labeled as '0' and '1'.

(a) Download the MNIST dataset. We aim to create a classifier that can differentiate handwritten digit 0 from 1. Take (at least) 100 samples of each class and store them as you will need them for prior calculations.

(b) What features do you suggest for this task? Explain and extract the feature map.

(c) Estimate the mean and covariance matrix for each shape class, assuming Gaussian distributions.

(d) Implement a Bayesian classifier using the estimated parameters to classify each shape.

(e) Evaluate the classifier's performance by computing the confusion matrix and accuracy.

4 Binary Classification of MNIST Digits Using Neural Networks

Background

Neural networks are powerful tools for pattern recognition tasks, including digit classification. In this exercise, you'll implement a simple feedforward neural network to distinguish between the digits '0' and '1' from the MNIST dataset. This task will help you understand the basics of neural network architecture, training, and evaluation.

(a) Load the MNIST dataset and filter it to include only images labeled '0' and '1'. Normalize the pixel values to the range [0, 1].

(b) Split the filtered dataset into training and testing sets (e.g., 80% training, 20% testing).

(c) Design a simple feedforward neural network with one hidden layer (e.g., 128 neurons with ReLU activation) and an output layer with a single neuron using the sigmoid activation function.

(d) Train the neural network on the training set using binary cross-entropy loss and an appropriate optimizer (e.g., Adam).

(e) Evaluate the trained model on the testing set, reporting metrics such as accuracy, precision, recall, and the confusion matrix.

(f) Discuss the model's performance and any challenges encountered during training.

(g) Compare with the result of previous exercise.

Textbook Questions

Please provide the solutions to the questions presented in the 12th chapter of the Image Processing book (Edition3): 1, 5, 7, 9, 12, 17