Cairo University
Faculty of engineering
Computer Engineering Department

Cairo University

Pattern Recognition Project

Introduction

We will implement a classifier that takes a bitmap file with black background and a handwritten drawn shape with a white brush as shown in figure 1, The classifier should take a decision and classify the shape correctly into diamond, ellipse or line.

Figure1
The exact size of the bmp file will be similar to those of the training set provided

Design

Step 0

You will be provided with a set that consists of three folders

- Diamond folder which consists of the diamonds' training patterns.
- Ellipse folder which consists of the ellipses' training patterns.
- Line folder which consists of the lines' training patterns.

See the set given to you to know the format of the folders and the files to be able to deal with them in your code.

Step 1

Raw data is not useful, we need to choose powerful features to represent each bitmap file.

Use the internet to find useful features, and try to invent new ones if necessary.

Step 2

- You are supposed to divide the set given to you into a Training set and a Validation set.
- Use the training set to train your classifier using different algorithms and different parameters you have taken in the lectures. Each time calculate the efficiency of your classifier on the training set.
- For each classifier, find the efficiency on your validation set.
- Choose the best performing classifier on the validation set to deliver it.

Delivery

- Delivery date and time will be announced later.
- You will be required to deliver a report containing the features with histograms illustrating the usefulness of each feature, stages of your design, algorithms and parameters used and the final classifier chosen
- On the delivery day you will be given an unseen test set of three folders similar to those of the original set in format (any change will be announced before delivery by a sufficient time) and your program should be able to calculate the efficiency on each folder separately, show which of the files were misclassified as what.