

Programación Técnica y Científica. Ejercicio de organización de un dataset

En este ejercicio vamos a organizar un dataset para clasificación de flores en varias carpetas. Como datos de entrada se proporciona la carpeta “dataset” que contiene 80 imágenes de cada uno de 17 tipos de flores, en total 1360 imágenes numeradas de 1 a 1360 y un fichero “classes.txt” donde viene el nombre de cada clase. Cada 80 imágenes pertenecen a cada una de las 17 clases listadas en el fichero “classes.txt” según en el orden en el que aparecen en dicho fichero. Es decir, en “classes.txt” la primera clase se llama “daffodil” (narciso), por tanto las imágenes image_0001.jpg a image_0080.jpg son narcisos. Del image_0081.jpg a image_0160.jpg son “snowdrop”(campanillas) y así sucesivamente.

Hay que implementar un programa en Python que sea capaz de crear una carpeta por cada clase y copiar sus correspondientes imágenes a la vez que se renombran a "image_9999_clase_n.jpg". Es decir dentro de “daffodil” se deberían llamar image_0001_daffodil_1.jpg a image_0080_daffodil_80.jpg y así con cada clase.

Para realizar este ejercicio se recomienda usar los módulos *glob* y *os* de los cuales tenéis ejemplos en los script proporcionados de la práctica 2 Tkinter y Robótica. El dataset se encuentra en Prado. Antes de realizar la entrega se debe corregir el ejercicio con los ficheros de la carpeta Corregir ejercicio Flowers de Prado. El script a entregar se debe llamar “dividirFlowers.py”