

Instrucciones Práctica Inteligencia de Negocio

Base de Datos: Corales

January 15, 2021

1 Introduction

En este documento recogemos las instrucciones para los alumnos de la asignatura de Inteligencia de Negocio para la realización de la práctica alternativa al examen.

2 Conjunto de datos: Corales

La práctica consiste en un trabajo de clasificación sobre un conjunto de datos. En vuestro caso es sobre la base de datos Corales.

Esta base de datos presenta la siguiente estructura:

- Contiene 14 clases de imágenes a color de corales
- Presenta un total de 5 conjuntos de datos para realizar una 5 fold cross validation.

En primer lugar es importante que hagáis un estudio inicial del conjunto de datos, observando bien los datos que tiene y cada una de sus clases, por si fuera necesario alguna técnica de preprocesamiento previo a clasificar los datos del conjunto. La clasificación de Corales es un problema que se ha estudiado en la literatura y que se han creado bases de datos para clasificar los diferentes tipos de corales. Uno de los primeros conjuntos de datos lo podéis ver en [4] donde se realiza un primer estudio para la clasificación de Corales.

En [3] podéis ver un estudio en el que se compararon diferentes bases de datos sobre Corales usando distintas redes neuronales y diferentes tipos de preprocesamiento para que podáis tomar ideas para vuestra práctica.

En el grupo de investigación se ha creado una nueva base de datos de Corales que extiende a esta primera base de datos y que es la que tenéis que utilizar. Una compañera se encargó de realizar este trabajo el cual se recoge en [1]. En el se crea una nueva base de datos llamada StructureRMAS (SRSMAS) y además se construye un clasificador que es capaz de clasificar bien estos datos usando una validación cruzada de 5 fold. En dicho artículo se proporcionan varios métodos

para afrontar esta nueva base de datos que podéis explorar y usar para clasificar los diferentes tipos de corales.

Por último, os dejo un enlace a otro trabajo de esta compañera en el que también trabaja con corales y prueba diferentes tipos de redes neuronales en [2].

Mi recomendación es que estudiéis estos trabajos y veáis que redes se han utilizado y que técnicas de procesamiento/tratamiento de datos se han usado. Algunas ideas que podéis desarrollar son las siguientes:

- Estudio de diferentes técnicas de preprocesamiento para el conjunto de datos.
- Estudio de diferentes redes neuronales para ver cuál de ellas clasifica mejor este conjunto. Comparación de nuevas redes con redes ya usadas para este problema.
- Aplicar los dos conceptos anteriores de forma conjunta.
- Desarrollar métodos de ensemble para intentar superar los resultados actuales.

References

- [1] Anabel Gómez-Ríos et al. “Coral species identification with texture or structure images using a two-level classifier based on Convolutional Neural Networks”. In: *Knowledge-Based Systems* 184 (2019), p. 104891.
- [2] Anabel Gómez-Ríos et al. “Redes Neuronales Convolucionales para Una Clasificación Precisa de Imágenes de Corales”. In: Feb. 2019.
- [3] Anabel Gómez-Ríos et al. “Towards highly accurate coral texture images classification using deep convolutional neural networks and data augmentation”. In: *Expert Systems with Applications* 118 (2019), pp. 315–328.
- [4] *Primera BD Corales*. <https://data.mendeley.com/datasets/86y667257h/2>. Accessed: 2021-01-15.