

# Clasificación de especies de Frailejón del páramo del Sumapaz mediante la morfología foliar.

Juan David Leal Campuzano

Andrés Julián Bermúdez García

Universidad Nacional de Colombia, Machine Learning 2021-1, Colombia.

julealc@unal.edu.co—anbermudezg@unal.edu.co



## Resumen

Determining the species of Frailejón to which a leaf of this species belongs is a fairly complex work that requires the advice of many times personnel specialized in botany and genetics. In this context, this work attempts to construct an algorithm capable of performing such a classification based on the morphology of the leaves of the different species of Frailejón (*E. Grandiflora*, *E. Argentea*, *E. Sumapacis*) present in a set of tagged photographic data based on previous classification work morphology of different types of plant species. Using the recommended features in the mentioned works we create a set of data from which we extract the most relevant characteristics and which will be used in different classification algorithms in order to find the one with the greatest accuracy.

## Introducción

La necesidad de distinguir el ecosistema en donde se trabaja genera la necesidad de identificar las diferentes especies que componen el mismo, para el estudio de caso, dado el predominio de la Espeletia en los páramos, es de gran importancia identificar esas diferentes especies y sus características morfológicas.

## Construcción del Conjunto de Datos y extracción de características



Figura 1: Las fotografías en su mayoría presentaba discrepancias en la toma de muestras tales como: Distancia del objetivo a la cámara, tamaño de la imagen en píxeles, iluminación, saturación, contaminación de los objetivos, fondos no adecuados para la separación de objetos, entre otros. Este conjunto de datos fué proporcionado por Iván Jiménez en su trabajo [11]

## Esquema de trabajo

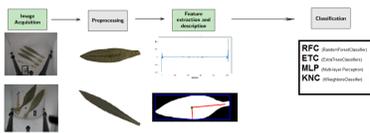


Figura 2: Metodología adoptada para el desarrollo de este trabajo

## Características geométricas.

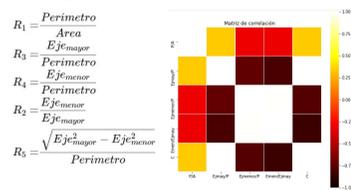


Figura 3: Primer espacio de características y sus correlaciones Pearson.

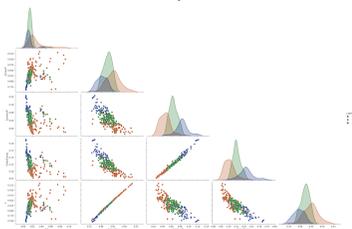
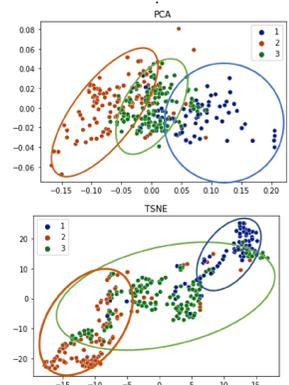


Figura 4: Datos interpretados a partir de pareja de características



## Características de contorno

### Contornos de Fourier



Figura 5: DFT aplicada a los puntos que conforman el contorno de la hoja.

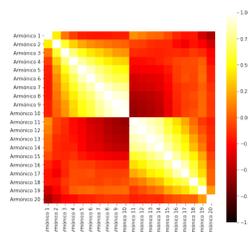


Figura 6: Donde efectivamente se observa que muchos de ellos poseen un grado alto de correlación (visto desde el valor absoluto de estos coeficientes).

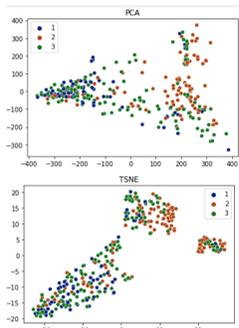


Figura 7: PCA y TSNE separados por grupos de especie usando el espacio de los armónicos

## Análisis de los datos obtenidos

### Modelos

#### Creación de los espacios de análisis

- Espacio 1 = Componentes Geométricos.
- Espacio 2 = Componentes de contorno.
- Espacio 3 = Unión de estos espacios.

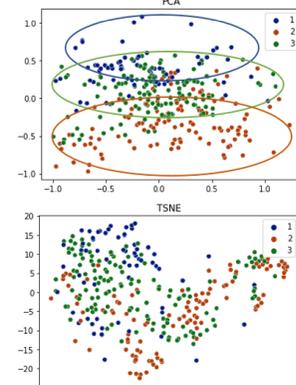


Figura 8: PCA y TSNE, Reducción de dimensión aplicado a la unión de los dos espacios E1 y E2

## Aplicación de los modelos de clasificación

### Resultados.

#### Espacio 1:

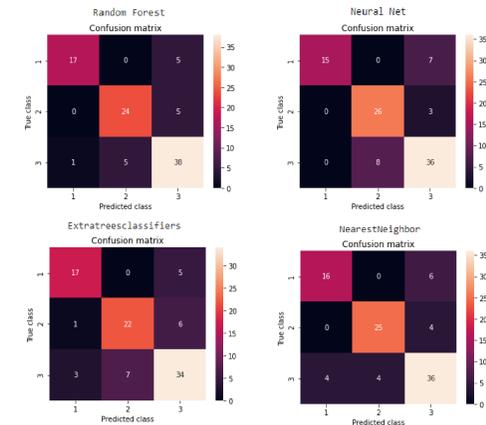


Figura 9: Matriz de confusión para cada modelo entrenado con el espacio 1 completo.

#### Subespacios:

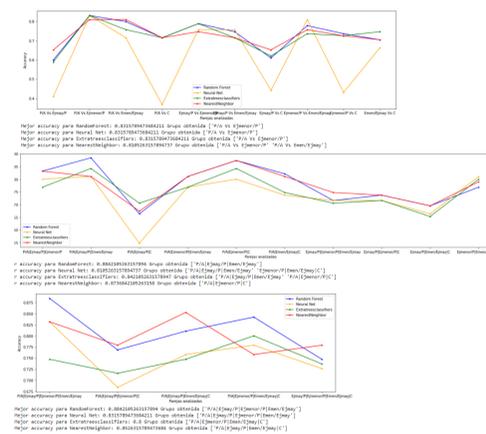


Figura 10: Análisis por subgrupos, arriba por parejas, en el medio por tripletas y abajo por cuartetos

#### Espacio 2:

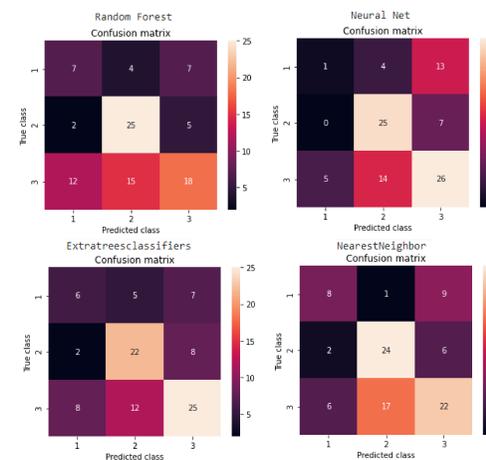


Figura 11: Matriz de confusión para cada modelo entrenado con el espacio 1 completo.

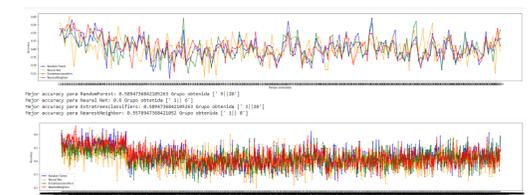


Figura 12: Análisis por subgrupos, arriba por parejas, abajo por tripletas

#### Subespacios: Espacio 3:

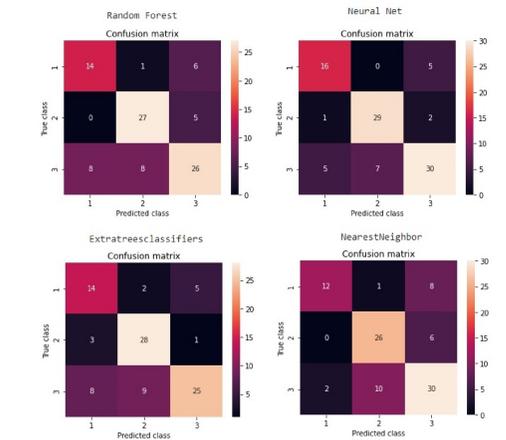


Figura 13: Matriz de confusión para cada modelo entrenado con el espacio 3 completo.

## Conclusiones

- Se puede concluir que a pesar de no tener un precisión de más del 90 por ciento, se obtuvo una buena precisión, esto dado que los datos o las imágenes no tenían la calidad que esperaba y además la identificación de las especies estaba hecha por un libro bastante antiguo y estás imágenes fueron tomadas hace pocos años, como menciona [11] el paso del tiempo y el cambio del entorno para estas plantas ha hecho que se creen syngameones, término que se usa cuando un conjunto de especies del mismo género o familia ha combinado sus características fenotípicas a tal punto de no ser fácilmente distinguibles como es el caso de la Espeletia Grandiflora y la Espeletia Sumapacis, incluso pudiendo haber otras especies que aún no se hayan reconocido en la misma muestra.
- Estos tipos de estudios vienen de gran importancia en las áreas como la Botánica y la Ecología porque la identificación de estos tipos de especies que habitan estas zona permiten caracterizar el ecosistema, teniendo imágenes de buena calidad y quizás un poco más de imágenes se podría llegar a un buen resultado lo cual acortaría tiempo y gastos a las personas que se encargan de este tipo de investigaciones.

[Link del video](#)

## Referencias

- [1] Leen B. Devellder C. Dhaene T. Deschrijver D. Automated classification of appliances using elliptical fourier descriptors.
- [2] Bobadilla J. *Machine Learning y Deep Learning: Usando Python, Scikit y Keras*. Ra-ma, 2020.
- [3] Cuatrecasas J. *A systematic study of the subtribe Espeletiinae*. The New York Botanical Garden Press, 2013.
- [4] Olivares L, Victorino J., and Gómez F. Automatic leaf shape category discovery. *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016.
- [5] Olivares L, Victorino J., and Gómez F. Contour analysis for interpretable leaf shape category discover. 2016.
- [6] Zhang S, Cheng D, Zong M, Gao L. Self-representation nearest neighbor search for classification.
- [7] Cardenas M. *Ecophysiology of paramos in Colombia: vulnerability to climate change and land use*. PhD thesis, Universidad Nacional de Colombia Sede Medellín Facultad de Minas Escuela de Geociencias y Medio Ambiente.
- [8] Van den Poel D, Prinzie A. Random forests for multiclass classification: Random multinomial logit. 2008.
- [9] Zhang S., Cheng D., Zong M., and Gao L. Self-representation nearest neighbor search for classification. 2016.
- [10] Burger W. and Burge MJ. *Principios del procesamiento de imágenes digitales. Temas de Pregrado en Informática*.
- [11] Pineda Y., Cortés A., Madrián S., and Jiménez I. The nature of espeletia species. 2020.