

GAX-Kepler: Clasificador de morfología galáctica mediante el uso de redes neuronales convolucionales

Brian Gonzalez, Dylan Jara

Universidad de Santiago de Chile

18 de diciembre de 2025



1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

1 Introducción

Introducción al problema

Justificación

Estado del arte

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

1 Introducción

Introducción al problema

Justificación

Estado del arte

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

Dataset Galaxy Zoo

- Una galaxia es un entorno en el cual las estrellas nacen y mueren, y las galaxias lejanas son faros luminosos que nos permiten explorar el universo lejano. (Silk y Bouwens, 2001)

Dataset Galaxy Zoo

- Una galaxia es un entorno en el cual las estrellas nacen y mueren, y las galaxias lejanas son faros luminosos que nos permiten explorar el universo lejano. (Silk y Bouwens, 2001)
- En este trabajo, usamos Galaxy Zoo (Lintott y cols., 2008; Willett y cols., 2013) para la clasificación y validación de nuestros modelos de clasificación.

Dataset Galaxy Zoo

- Una galaxia es un entorno en el cual las estrellas nacen y mueren, y las galaxias lejanas son faros luminosos que nos permiten explorar el universo lejano. (Silk y Bouwens, 2001)
- En este trabajo, usamos Galaxy Zoo (Lintott y cols., 2008; Willett y cols., 2013) para la clasificación y validación de nuestros modelos de clasificación.
- El proyecto original Galaxy Zoo fue lanzado en julio de 2007 y obtuvo más de 100,000 voluntarios en sus primeros diez días.

Dataset Galaxy Zoo

- Una galaxia es un entorno en el cual las estrellas nacen y mueren, y las galaxias lejanas son faros luminosos que nos permiten explorar el universo lejano. (Silk y Bouwens, 2001)
- En este trabajo, usamos Galaxy Zoo (Lintott y cols., 2008; Willett y cols., 2013) para la clasificación y validación de nuestros modelos de clasificación.
- El proyecto original Galaxy Zoo fue lanzado en julio de 2007 y obtuvo más de 100,000 voluntarios en sus primeros diez días.
- Los voluntarios clasificaron imágenes del Sloan Digital Sky Survey (SDSS) de galaxias en seis categorías: elíptica, espiral en sentido horario, espiral en sentido antihorario, edge-on , estrella/no se, o fusión de galaxias.

1 Introducción

Introducción al problema

Justificación

Estado del arte

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

¿De qué nos sirve?

- La clasificación morfológica de galaxias es fundamental para entender la formación y evolución de las galaxias.
- Basicamente facilita la exploración espacial. Y esto ayuda a entender mejor el universo y el origen de la vida.

1 Introducción

Introducción al problema

Justificación

Estado del arte

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

Estado del arte

- La clasificación morfológica de galaxias es una pieza clave de información para definir muestras de galaxias con el objetivo de estudiar la estructura a gran escala del universo. (Barchi y cols., 2020)

Estado del arte

- La clasificación morfológica de galaxias es una pieza clave de información para definir muestras de galaxias con el objetivo de estudiar la estructura a gran escala del universo. (Barchi y cols., 2020)
- Muchos autores (Abraham y cols., 1996; Conselice y cols., 2000) han estudiado y presentado resultados sobre medidas objetivas de morfología de galaxias con Concentración, Asimetría, Suavidad, Gini y M20 (sistema CASGM).

Estado del arte

- La clasificación morfológica de galaxias es una pieza clave de información para definir muestras de galaxias con el objetivo de estudiar la estructura a gran escala del universo. (Barchi y cols., 2020)
- Muchos autores (Abraham y cols., 1996; Conselice y cols., 2000) han estudiado y presentado resultados sobre medidas objetivas de morfología de galaxias con Concentración, Asimetría, Suavidad, Gini y M20 (sistema CASGM).
- Las Redes convolucionales profundas (CNN) son una metodología bien establecida para clasificar imágenes (Goodfellow y cols., 2016).

1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

Tipos de galaxias



Galaxia Espiral
Galaxia espiral
M61. Cúmulo de
Virgo.



Galaxia Elíptica
Galaxia elíptica
Messier 32.
Constelación de
Andrómeda.



Galaxia Irregular
Galaxia enana
UGC 4459.
Constelación
Osa Mayor.

1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

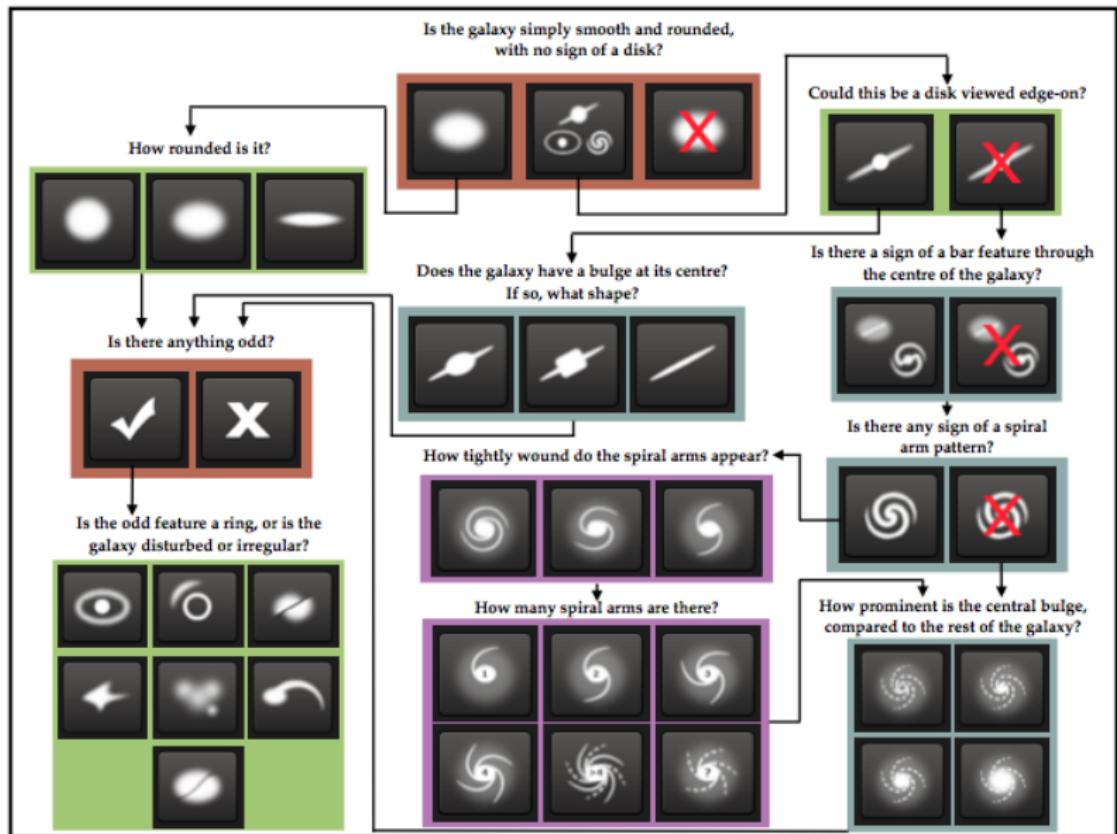
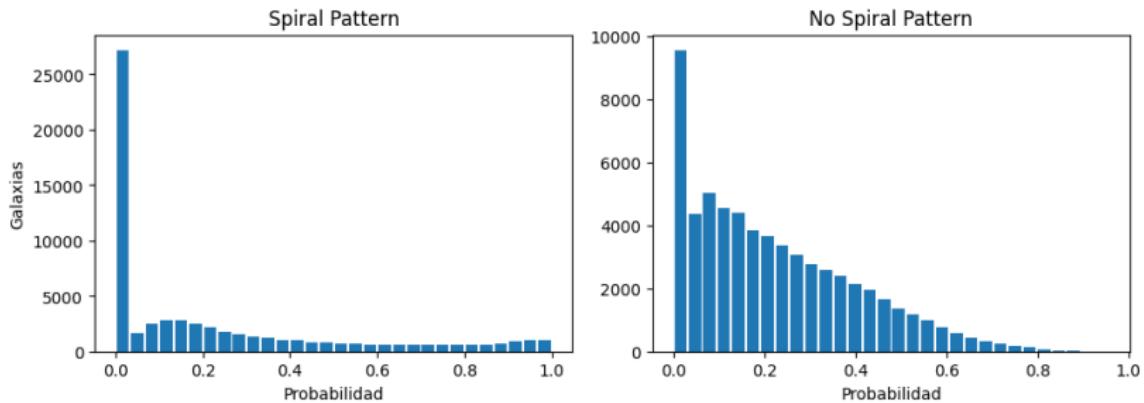


Figure 1. Flowchart of the classification tasks for GZ2, beginning at the top centre. Tasks are colour-coded by their relative depths in the decision tree. Tasks outlined in brown are asked of every galaxy. Tasks outlined in green, blue, and purple are (respectively) one, two or three steps below branching points in the decision tree. Table 2 describes the responses that correspond to the icons in this diagram.

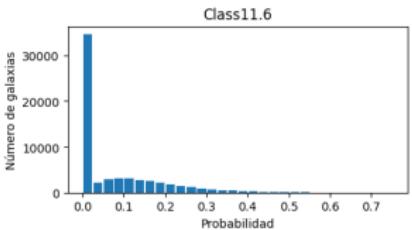
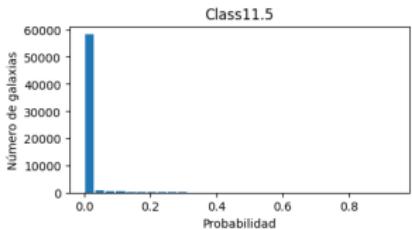
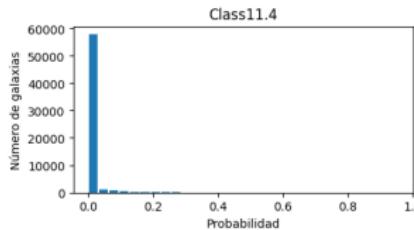
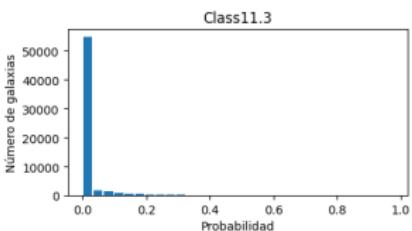
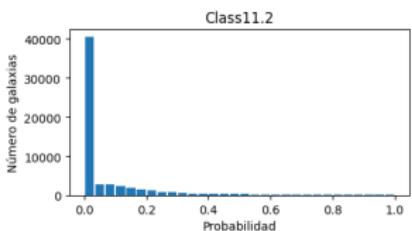
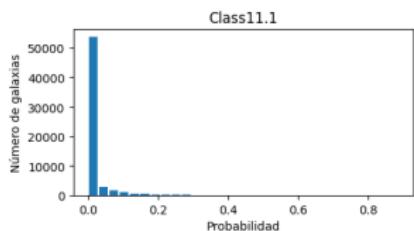
Probabilidades de Q4

Distribución de patrón espiral (Q4)



Probabilidades de Q11

Distribución de probabilidades - Número de brazos espirales (Q11)



1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

¿Qué queremos lograr?

5 Metodología

6 Conclusiones e Ideas Futuras

1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

¿Qué queremos lograr?

5 Metodología

6 Conclusiones e Ideas Futuras

¿Qué queremos lograr?

- Una red neuronal convolucional que pueda aprender representaciones jerárquicas de imágenes de galaxias y predecir de manera efectiva las probabilidades de sus características morfológicas con un bajo error cuadrático medio

1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

Arquitectura de la red

6 Conclusiones e Ideas Futuras

1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

Arquitectura de la red

6 Conclusiones e Ideas Futuras

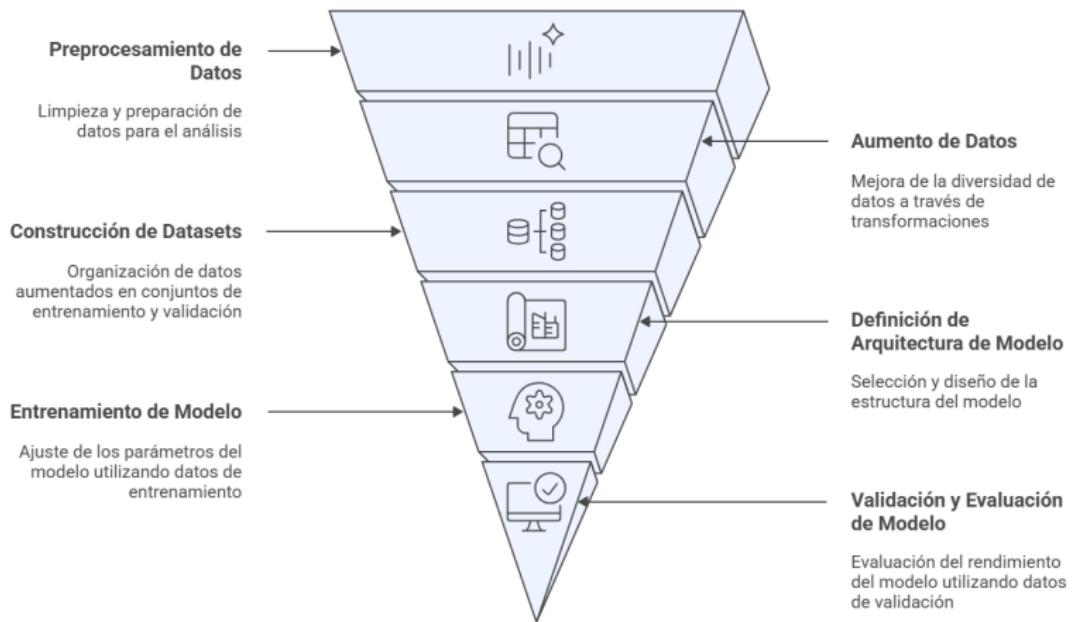
Arquitectura del Modelo CNN

Capa (tipo)	Forma de salida	Parámetros
Input (Sequential)	(64, 64, 3)	0
Conv2D (64 filtros)	(62, 62, 64)	1,792
Batch Normalization	(62, 62, 64)	256
MaxPooling2D	(31, 31, 64)	0
Conv2D (128 filtros)	(29, 29, 128)	73,856
Batch Normalization	(29, 29, 128)	512
MaxPooling2D	(14, 14, 128)	0
Conv2D (256 filtros)	(12, 12, 256)	295,168
Batch Normalization	(12, 12, 256)	1,024
MaxPooling2D	(6, 6, 256)	0

Arquitectura del Modelo CNN II

Capa (tipo)	Forma de salida	Parámetros
GlobalAveragePooling2D	(None, 256)	0
Dense	(None, 256)	65,792
Dropout	(None, 256)	0
Dense (Output)	(None, 37)	9,509
Total params		447,909
Trainable params		447,013
Non-trainable params		896

Esquema Metodológico



1 Introducción

2 Tipos de galaxias

3 Descripción de los datos de estudio

4 Hipótesis

5 Metodología

6 Conclusiones e Ideas Futuras

Ideas Futuras

- Usar GPU para entrenar más rápido.
- Método Divide y vencerás para distribuir.
- Usar la CNN para extracción de datos y clasificar con Decision Tree, SVM o Random Forest.

Conclusiones

- ① Las CNN son una herramienta poderosa para la clasificación de imágenes
- ② Con ayuda de un buen entendimiento en la teoría del problema, se pueden lograr muy buenos resultados.
- ③ La clasificación automática de galaxias puede acelerar significativamente el análisis de grandes conjuntos de datos astronómicos.

Referencias I

- Abraham, R. G., Tanvir, N. R., Santiago, B. X., Ellis, R. S., Glazebrook, K., y Bergh, S. v. d. (1996, 04). Galaxy morphology to $i=25$ mag in the hubble deep field. *Monthly Notices of the Royal Astronomical Society*, 279(3), L47-L52. Descargado de
<https://doi.org/10.1093/mnras/279.3.L47> doi: 10.1093/mnras/279.3.L47
- Barchi, P. H., de Carvalho, R., Rosa, R. R., Sautter, R., Soares-Santos, M., Marques, B. A., ... Moura, T. (2020). Machine and deep learning applied to galaxy morphology-a comparative study. *Astronomy and Computing*, 30, 100334.

Referencias II

- Conselice, C. J., Bershady, M. A., y Jangren, A. (2000). The asymmetry of galaxies: physical morphology for nearby and high-redshift galaxies. *The Astrophysical Journal*, 529(2), 886.
- Goodfellow, I., Bengio, Y., y Courville, A. (2016). Deep Learning. MIT Press. Descargado de
<http://www.deeplearningbook.org>
- Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., ... others (2008). Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 389(3), 1179–1189.

Referencias III

- Silk, J., y Bouwens, R. (2001). The formation of galaxies. *New Astronomy Reviews*, 45(4), 337-350. Descargado de <https://www.sciencedirect.com/science/article/pii/S1387647300001640> doi: [https://doi.org/10.1016/S1387-6473\(00\)00164-0](https://doi.org/10.1016/S1387-6473(00)00164-0)
- Willett, K. W., Lintott, C. J., Bamford, S. P., Masters, K. L., Simmons, B. D., Casteels, K. R., ... others (2013). Galaxy Zoo 2: detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 435(4), 2835–2860.