



Facultad de Ciencias  
**Astronómicas  
y Geofísicas**  
UNIVERSIDAD NACIONAL DE LA PLATA

Tesis de Grado para obtener la Licenciatura en Astronomía

**Detección y medición automática de galaxias ubicadas en la  
dirección del cúmulo de Fornax en imágenes de S-PLUS**

**Rodrigo Facundo Haack**

Directora: Analía Viviana Smith Castelli

Codirector: Favio Raúl Faifer

Firma y aclaración del Director o Codirector:

Analía V. Smith Castelli

Firma y aclaración de uno de los Jurados:

La Plata  
- 2023 -



*A mi familia por acompañarme y estar siempre en todo.*

*A Claudia, mi mamá, por la enorme paciencia, por escuchar y entenderme siempre. Por desearme lo mejor de lo mejor en cada cosa que haga, por no solo desearlo sino darlo todo para que intente lograrlo.*

*A Ernesto, mi papá, no tanto por la paciencia sino por la pasión. Por mostrarme que es posible preguntarme qué quiero para mi vida, por demostrarme cómo acompañar, confiar y cuidar incondicionalmente a alguien.*

*Estoy orgulloso de ustedes, de que sean mi mamá y papá, de que me hayan cuidado y educado como lo hicieron. A les dos, gracias por mostrarme que siempre se sigue aprendiendo, que viajar ayuda mucho y que no se puede vivir sin amor. Les amo muchísimo.*

*A Sofi, mi hermana, por la relación hermosa que construimos. Por mostrar que la palabra hermana no es solo por compartir a mamá y papá, sino por compartir nuestras propias vidas. Te amo y amo vivir la vida con vos.*

*A mis tíos, tías, primos y primas. A las abuelas y abuelos, que siguen estando desde otro lado. Por cuidarme y enseñarme a jugar, reír y soñar. Les amaré por siempre.*

*Con el mismo amor profundo, gracias a mis amigos. Son quienes elijo para compartir mi vida y gracias por permitirme ser parte de las suyas. A Mili, Terre, Peter y Azu, por ser hermanes de otra familia. A Juli, Tomi R, Mica, Flor, Anita, Tomi A y Juli T gracias por estar cerca estando tan lejos. A Gasti, Colo, Lu F, Jere, Seba Fede B, Pelu, Nati, Agus M, Nico R, Tati, Agus G, Mau, Tin y Palo por cada paso que dimos y damos juntas.*

*A Analía, por tenerme tanta paciencia y por escucharme constantemente. Por darme el espacio para decir y manifestar lo que siento. Espero que los próximos años sean disfrutables como los que ya compartimos. A Favio por sus valiosos aportes a este trabajo y también al jurado por sus correcciones y comentarios.*

*A quien lea esta Tesis.*

*A Inti por tantos años de lucha en las calles, por repensar constante y críticamente qué Facultad queremos construir, por cada instancia de reflexión y formación que tuvimos, por la horizontalidad y autogestión real como forma de organización. A Derecho al Cielo Nocturno, por educar entendiendo que observar el cielo es un derecho, por buscar fortalecer el derecho a la recreación y educación en las instituciones de encierro, por que la educación popular es necesariamente el camino, y por proponer colaborar en vez de competir. Al Encuentro Nacional de Estudiantes de Astronomía, por hacerme repensar el rol de la ciencia en la sociedad, por cada debate o taller y cada uno de los 7 encuentros que viví, por la gente de Córdoba y San Juan que conocí. Les agradezco porque sin haber formado parte de estos tres espacios no hubiera llegado hasta acá ni sería quien soy.*

*A la comunidad del Observatorio de La Plata de la cuál estoy orgulloso de pertenecer. Al personal docente y nodocente con quienes compartí aulas o charlas en el pasillo. A quienes forman y formaron parte de la Secretaría de Extensión y el Planetario. A Chacho por su dedicación incansable a la extensión universitaria crítica y cada intercambio que tuvimos. A Sixto por todos los viajes de campo, por el seminario y por la posibilidad de estudiar el cielo desde otro lado.*

*A la Universidad Nacional de La Plata y la educación pública, porque sin ella no estarían leyendo esto.*

*A Rayo, Saturnina y al amor interespecie. Al mate y la siesta.*

*A la Colaboración Internacional S-PLUS por los datos provistos y la ayuda brindada en distintas instancias del desarrolló de esta Tesis. Al Consejo Interuniversitario Nacional por otorgarme una Beca de Estímulo a la Vocación Científica (EVC-CIN) en el marco de la cual se llevó adelante este trabajo de Tesis.*

*Al cielo y su compañía constante.*

# Índice general

<b>1</b>	<b>Introducción</b>	<b>9</b>
1.1	Grandes relevamientos astronómicos	9
1.2	Southern Photometric Local Universe Survey (S-PLUS)	10
1.3	Cúmulo de Fornax	16
1.4	Motivación de esta Tesis	19
<b>2</b>	<b>Datos y herramientas</b>	<b>21</b>
2.1	Imágenes y catálogos de S-PLUS	21
2.2	SExtractor	23
2.2.1	Generalidades	23
2.2.2	Obtención de magnitudes totales	24
2.2.3	Separación estrella/galaxia	25
2.2.4	Ejecución de SExtractor	30
2.3	TOPCAT	32
<b>3</b>	<b>Metodología</b>	<b>33</b>
3.1	Esquema General	33
3.2	Descarga de campos a procesar	35
3.3	Construcción de las Imágenes Suma + Headers	35
3.4	Lectura y selección de parámetros	38
3.5	Archivos de configuración y ejecuciones de SExtractor	38
3.6	Modificación y fusión de catálogos	46
<b>4</b>	<b>Resultados</b>	<b>50</b>
<b>5</b>	<b>Síntesis y trabajo a futuro</b>	<b>54</b>
<b>6.</b>	<b>Bibliografía</b>	<b>57</b>

# Índice de Figuras

Figura 1.1	Diagrama de relevamientos ópticos	11
Figura 1.2	Comparación de relevamientos ópticos	12
Figura 1.3	T80S en Cerro Tololo	13
Figura 1.4	T80S y su cámara de campo amplio	14
Figura 1.5	Sistema de filtros Javalambre	15
Figura 1.6	Imagen en falso color RGB de NGC 1316	18
Figura 2.1	Distribución espacial de los campos utilizados	22
Figura 2.2	Diagrama de la red neuronal utilizada por SExtractor	29
Figura 2.3	Ejemplo de un archivo de entrada .param	31
Figura 3.1	Esquema básico de los cinco pasos del proceso de automatización.	34
Figura 3.2	Esquema completo de la estructura de subdirectorios	36
Figura 3.3	Comparación entre la imagen suma y la imagen del filtro "G"	37
Figura 3.4	Diagrama separación de un objeto compuesto	41
Figura 3.5	Comparación entre imágenes de abertura para las ejecuciones realizadas	45
Figura 3.6	Parámetros de salida de los catálogos con su respectiva descripción	48
Figura 3.7	Parámetros geométricos	48
Figura 4.1	Comparación de histogramas sobre CLASS_STAR_G entre Run 1 y Run 2.	51
Figura 4.2	Comparación de diagramas de completitud en G, R e I.	51
Figura 4.3	Diagrama color-magnitud comparativo entre Fornax y Virgo	53
Figura 5.1	Gráficos de predicción para muestras con y sin <i>data missing</i>	55

# Índice de Tablas

Tabla 1.1 Filtros de S-PLUS . . . . .	15
Tabla 3.1 Parámetros y sus correspondientes valores para la detección y extracción, diferenciados entre ejecuciones. . . . .	39
Tabla 3.2 Parámetros y sus correspondientes valores para la estimación del background, diferenciados entre ejecuciones. . . . .	42
Tabla 3.3 Parámetros del archivo de configuración compartidos por ambas ejecuciones . . . . .	43





# Resumen

En este trabajo de Tesis exploramos estrategias de detección automática de galaxias en imágenes obtenidas por el *Southern Photometric Local Universe Survey* (S-PLUS<sup>1</sup>). Para ello, se hizo uso del software **SExtractor** el cual se encuentra especialmente diseñado para realizar detección y fotometría rápida en imágenes de gran campo. Este software provee, entre otras cosas, catálogos que contienen información fotométrica y parámetros estructurales para todos los objetos detectados sobre las imágenes. El fin último de esta Tesis consiste en, a partir del uso de imágenes de la colaboración S-PLUS, adquirir el manejo del software **SExtractor** y habilidades de programación en **Python** que permitan obtener y manejar grandes catálogos con fotometría confiable de galaxias ubicadas en la dirección del Cúmulo de Fornax.

S-PLUS es un relevamiento fotométrico actualmente en desarrollo que, una vez finalizado, habrá mapeado  $\sim 9300 \text{ grad}^2$  del Hemisferio Sur Celeste. Este cartografiado se realizará a través de la obtención de imágenes de gran campo ( $1.4 \times 1.4 \text{ grad}^2$ ) en 12 bandas fotométricas, 5 de las cuales corresponden a los filtros de banda ancha del SDSS (u,g,r,i,z), y 7 adicionales que trazan distintas características espectroscópicas ([OII], Ca H+K, G-band, H $\delta$ , Mgb, H $\alpha$ , y CaT). El número de filtros involucrados sumado al gran campo de sus imágenes, constituye una particularidad por demás interesante que lo distingue de muchos de los relevamientos actuales y planeados a futuro. Por ejemplo, las 12 bandas fotométricas en las que se obtienen simultáneamente las imágenes, permitirán, entre otras cosas, obtener la distribución de energía espectral (SEDs por sus siglas en inglés) en el rango  $\sim 3200 - 10000 \text{ \AA}$ , así como estimar redshifts fotométricos, para millones de objetos.

En este contexto, el objetivo general de esta Tesis consiste en acercarse al entorno de trabajo relacionado con los grandes relevamientos astronómicos que se encuentran actualmente en desarrollo a nivel internacional, o que se planean iniciar en el futuro cercano. El paso inicial para conseguir esto consistió en la adquisición de manejo de algún software de detección y medición automática de galaxias en imágenes de gran campo. **SExtractor** se encuentra especialmente diseñado para tal fin. El segundo paso para concretar el objetivo planteado fue adquirir habilidades de programación en **Python** que permitieron automatizar la ejecución de **SExtractor** y manejar el volumen de información obtenido a partir de dicha ejecución.

---

<sup>1</sup><https://www.splus.iag.usp.br/>

El objetivo específico de esta Tesis consistió en obtener fotometría homogénea de un gran número de objetos ubicados en imágenes obtenidas con S-PLUS mediante la utilización de **SExtractor**. Dicho software trabaja mediante redes neuronales, y el usuario puede modificar distintos parámetros de entrada de modo de optimizar la detección y medición de objetos no resueltos o extendidos. En el caso particular de este trabajo, se realizaron diferentes pruebas utilizando diversos parámetros de entrada con el fin de optimizar la detección de galaxias reportadas en la literatura como miembros espectroscópicamente confirmados del Cúmulo de Fornax, o como candidatas a serlo por criterios morfológicos. Para ello, se trabajó sobre 43 campos observados como parte de la DR3 y de la iDR4 de S-PLUS, que cubren por completo la ubicación de dichas galaxias, abarcando una extensión total de  $\sim 2.5$  radios viriales ( $R_{\text{vir}}$ ) desde el centro del cúmulo. Una vez identificados dichos parámetros, se aplicaron para obtener catálogos con fotometría homogénea de 106 campos observados por S-PLUS que cubren no solo los 43 mencionados anteriormente, sino 63 campos adicionales que abarcan los alrededores de Fornax. Esperamos que durante la etapa de Doctorado, la fotometría homogénea obtenida para los 106 campos mencionados nos permita encarar distintos estudios astrofísicos como ser la identificación de nuevos miembros del cúmulo en regiones del cielo aún no exploradas en la literatura con ese fin.

# Capítulo 1

## Introducción

En esta Sección se introducirán nociones, tanto generales como específicas, sobre los grandes relevamientos astronómicos y su rol en la astronomía moderna, el Southern Photometric Local Universe Survey (S-PLUS) como el relevamiento astronómico utilizado en esta Tesis y el Cúmulo de Fornax como el objeto de estudio del presente trabajo.

### 1.1. Grandes relevamientos astronómicos

Los relevamientos astronómicos basados en la obtención de imágenes en tres o más bandas fotométricas se han generalizado en las últimas décadas, revolucionando, entre otros, los campos de la estructura a gran escala y evolución de galaxias. Con la llegada del Sloan Digital Sky Survey (SDSS; [York et al., 2000](#)), estudios que anteriormente se basaban en el análisis de tan solo cientos o miles de objetos, ahora involucran cientos de miles o incluso millones de ellos. Tales volúmenes de datos imponen la necesidad de determinar parámetros y realizar clasificaciones de objetos en forma automática. Por ejemplo, la clasificación morfológica de galaxias estuvo basada en la inspección visual de sus imágenes por parte de especialistas ([Nair y Abraham, 2010](#); [Ann et al., 2015](#)) o, incluso, de ciudadanos no formados profesionalmente en astronomía ([Lintott et al., 2011](#); [Willett et al., 2013](#); [Simmons et al., 2017](#)).

En las últimas décadas, se han utilizado algoritmos numéricos para determinar los principales parámetros morfológicos de las galaxias ([Spiekermann, 1992](#); [Storrie-Lombardi et al., 1992](#); [Walmsley et al., 2020](#)). Sin embargo, estos métodos presentan un problema común que consiste en que no resultarán eficientes para ser utilizados sobre la próxima generación de relevamientos astronómicos. Por ejemplo, los impresionantes volúmenes de datos esperados para el Legacy Survey of Space and Time (LSST, 17 TB por noche oscura; [Tyson et al., 2002](#); [Axelrod, 2006](#)) llevado adelante por el Observatorio Vera C. Rubin<sup>1</sup>, y los relevamientos que serán desarrollados por

---

<sup>1</sup><https://www.lsst.org/>

el telescopio espacial Nancy Grace Roman<sup>2</sup> (previamente llamado WFIRST, que mapeará 1000 grad<sup>2</sup> con una profundidad y calidad similar a la del Telescopio Espacial Hubble; Gehrels et al., 2015), transformarán en inviable la clasificación morfológica de galaxias realizada únicamente a partir de la inspección visual de sus imágenes por parte de humanos, entre otros métodos.

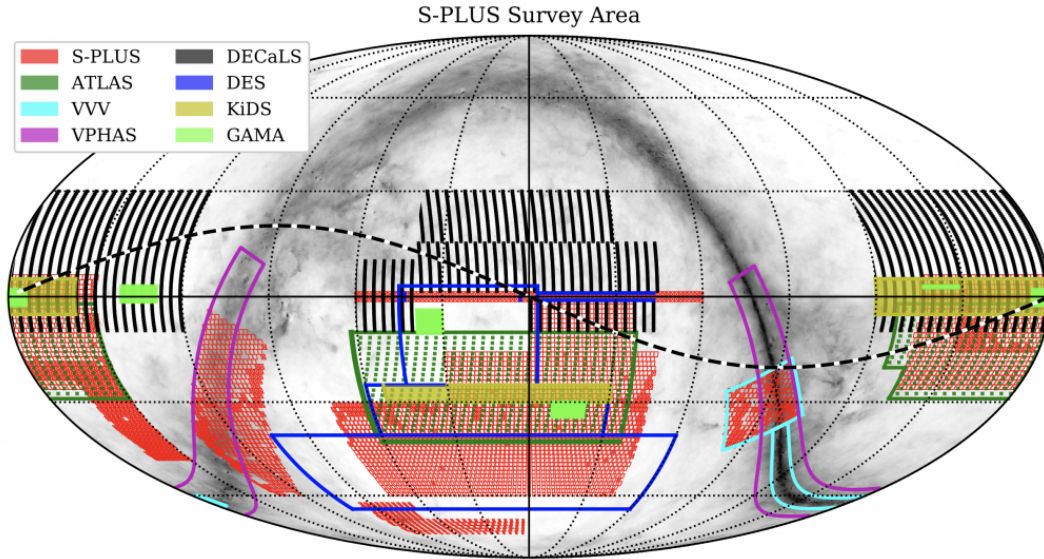
## 1.2. Southern Photometric Local Universe Survey (S-PLUS)

El Southern Photometric Local Universe Survey (S-PLUS) (Mendes de Oliveira et al., 2019) se encuentra tomando imágenes de  $\sim 9300$  grad<sup>2</sup> de la esfera celeste en doce bandas ópticas utilizando un telescopio robótico de 0,8 m, el T80-Sur, en el Observatorio Interamericano de Cerro Tololo, Chile. El relevamiento consta de cuatro zonas principales, que incluyen dos zonas no contiguas en latitudes galácticas altas ( $|b| > 30^\circ$ , 8000 grad<sup>2</sup>) y dos zonas del plano y el bulbo galácticos (para 1300 grad<sup>2</sup> adicionales). S-PLUS utiliza el sistema de magnitudes de 12 bandas de Javalambre que incluye los 5 filtros de banda ancha u, g, r, i, z y 7 filtros de banda estrecha centrados en líneas espectrales específicas. S-PLUS proporciona corrimientos al rojo fotométricos precisos ( $\delta z / (1 + z) = 0,02$  o mejor) para galaxias con  $r < 20$  mag y  $z < 0,5$ , produciendo así un mapa 3D del Universo local en un volumen de más de  $1 (\text{Gpc}/h)^3$ .

El catálogo final de S-PLUS también permitirá estudiar la formación de estrellas y las poblaciones estelares en la Vía Láctea y en las galaxias cercanas, así como también posibilitará la búsqueda de cuásares, fuentes variables y estrellas de baja metalicidad. Actualmente ya han sido liberados tres conjuntos de datos públicos y pronto se liberará el cuarto, el cual ya se encuentra disponible internamente para quienes participan de la colaboración. Al día de hoy ya se ha mapeado un 65 % de lo que se planificó inicialmente. En particular, la cuarta liberación de datos interna a la colaboración (iDR4) incluye 1629 campos que cubren 3000 grad<sup>2</sup> del cielo austral.

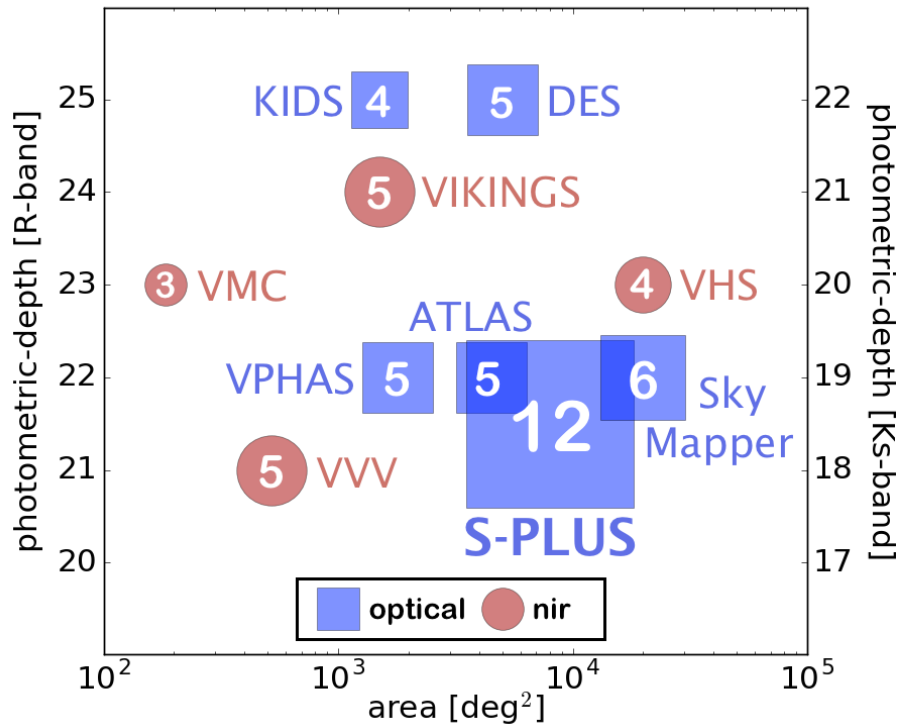
Aunque existen varios estudios actualmente en ejecución y planteados a futuro basados en la toma de imágenes de gran campo en el Hemisferio Sur Celeste, S-PLUS proporciona un muestreo único del espectro óptico gracias a sus siete filtros de banda estrecha. Las Figuras 1.1 y 1.2 muestran comparaciones de diferentes sondeos ópticos y en el infrarrojo cercano realizados con telescopios situados en el hemisferio sur, con respecto a su cobertura de área, profundidad fotométrica y número de filtros. S-PLUS también ofrecerá sinergias con la misión Gaia (Lindegren y Perryman, 1996) que en última instancia proporcionará espectrofotometría azul y roja de baja resolución para fuentes compactas obtenida a través de prismas, en un rango de longitudes de onda similares a las que mapea S-PLUS. Especialmente en el caso de

<sup>2</sup><https://roman.gsfc.nasa.gov/>



**Figura 1.1:** Diagrama en coordenadas ecuatoriales que muestra algunos de los principales relevamientos ópticos y en el infrarrojo cercano del Hemisferio Sur Celeste (se omiten los relevamientos que cubren totalmente el Hemisferio Sur Celeste o la totalidad de la esfera celeste). Para los relevamientos ópticos, ATLAS (Shanks et al., 2015) se muestra en verde sombreado, VPHAS+ (Drew et al., 2014) es el contorno rectangular rosa sobre el bulbo y el disco de la Galaxia, DECaLS está en negro sombreado y DES (Dark Energy Survey Collaboration et al., 2016) se muestra en contornos azules, KiDS (de Jong et al., 2015) en amarillo relleno, y GAMA (Driver et al., 2009) en áreas verdes rellenas. El único relevamiento en el infrarrojo cercano que se muestra en esta figura es VISTA-VVV (Minniti et al., 2010), en contornos azul claro, principalmente sobre el bulbo galáctico, solapándose con S-PLUS. El área cubierta por S-PLUS se muestra en rojo. La línea negra representa la eclíptica. La imagen de fondo es el mapa de extinción de Schlegel et al. (1998). Figura obtenida de (Mendes de Oliveira et al., 2019).

las galaxias resueltas, las imágenes de S-PLUS serán útiles para identificar qué áreas contribuyeron al espectro de Gaia, y qué información se está perdiendo. Además, la banda  $u$  de Javalambre, en combinación con los datos de Gaia, puede ser útil para mejorar la sensibilidad de Gaia en estas longitudes de onda. Cuando el Large Synoptic Survey Telescope (LSST) (Ivezić et al., 2019) entre en funcionamiento, proporcionará observaciones profundas del cielo observable desde CTIO con información temporal, pero todavía utilizando sólo cinco filtros de banda ancha. Por lo tanto, se prevé que los relevamientos multibanda de banda estrecha que utilizan incluso telescopios modestos como S-PLUS pueden seguir desempeñando un papel importante al proporcionar información espectral importante que se necesita para una amplia gama de aplicaciones astrofísicas. La tipificación estelar y los desplazamientos al rojo fotométricos de los relevamientos multibanda como S-PLUS proporcionarán un valioso recurso para la calibración de LSST y otros relevamientos. (Mendes de Oliveira et al., 2019) Es importante señalar que J-PLUS (Cenarro et al., 2019), realizado con el telescopio T80/JAST en España, ha estado generando datos durante los últimos años. T80-Sur y su cámara de gran formato, incluidos los filtros, son un duplicado de ese sistema instalado en el Cerro Javalambre. J-PLUS también es importante para calibrar J-PAS, el Javalambre Physics of the Accelerating Universe Survey que



**Figura 1.2:** Comparación de varios relevamientos ópticos y en el infrarrojo cercano del Hemisferio Sur Celeste. Las escalas de la izquierda y derecha de la figura muestran las profundidades aproximadas de los relevamientos ópticos (recuadros azules) e infrarrojos cercanos (círculos rojos), respectivamente, ambas en magnitudes AB. El número de cada recuadro indica el número de filtros del relevamiento; el tamaño del recuadro es proporcional al número de filtros. Figura obtenida de (Mendes de Oliveira et al., 2019).

utilizará 54 filtros de banda estrecha (145 Å de ancho) y 2 filtros de banda ancha que cubren todo el espectro óptico (Benitez et al., 2014). J-PAS se llevará a cabo con un telescopio dedicado de 2,5-m y una cámara de amplio campo de visión (FoV; por sus siglas en inglés) de  $\sim 4,7$  grad<sup>2</sup> en el Observatorio Astrofísico de Javalambre en España. Sin embargo, hasta ahora no se ha planificado un relevamiento de este tipo para el hemisferio sur.

El T80S está situado cerca de la cima de Cerro Tololo, en el centro de Chile, aproximadamente a 200 metros al noreste del telescopio Blanco de 4,0 m. La Figura 1.3 muestra una imagen del telescopio y sus alrededores. El telescopio se encuentra a una altitud de 2178 m sobre el nivel del mar, en la posición geodésica (Sistema Geodésico Mundial 84, latitud Sur y longitud Oeste son negativas) -30:10:04.31, -70:48:20.48. El CTIO tiene unas condiciones meteorológicas muy estables, con un 82,3% del tiempo utilizado para las observaciones del estudio de campo amplio durante el período 2013-2016. La mediana del seeing total es de 0,95".

El T80S tiene una montura ecuatorial alemana (modelo NTM1000), fabricada por la



**Figura 1.3:** T80S se encuentra en Cerro Tololo, junto a los telescopios PROMPT. En esta foto, tomada en octubre de 2017, T80S es la cúpula más grande de la izquierda. Figura obtenida de (Mendes de Oliveira et al., 2019).

empresa ASTELCO, bajo contrato con la empresa AMOS. Los diseños ópticos y del telescopio se realizaron en estrecha colaboración entre el CEFCA y AMOS/ASTELCO. El sistema óptico del T80S consiste en un RitcheyChretien de  $f/4,31$  con un plano focal Cassegrain axial y una apertura de 860 mm. Esto proporciona una escala de placas de  $55,56'' \text{ mm}^{-1}$ , un FoV total de 130 mm (que se traduce en un diámetro de 2 grados en el cielo), y un FoV óptimo de 110 mm (1,7 grados de diámetro en el cielo). La lente correctora de campo construida por AMOS garantiza una degradación de la aberración inferior al 1%. El detector es un arreglo de  $9232 \times 9216 \mu\text{m pixel}^{-1}$  y la escala del detector sobre el telescopio es de  $0.55'' \text{ pixel}^{-1}$ , traducándose en un FoV de  $1,4 \times 1,4 \text{ grad}^2$ . En la Figura 1.4 se muestra una imagen del telescopio y su cámara. El T80S está alojado en una cúpula Ash de 8 m. El telescopio puede girar entre dos posiciones opuestas del cielo en menos de 1,5 minutos, siendo el factor limitante el tiempo que tarda la cúpula en moverse entre las dos posiciones. Es operado robóticamente por el sistema de control remoto CHIMERA<sup>3</sup>. Desarrollado en Python, CHIMERA utiliza la biblioteca Pyro3 para convertir los subsistemas del observatorio en objetos Python accesibles a través de la red local de forma distribuida. Sobre este marco, un algoritmo supervisor se encarga de comprobar las condiciones meteorológicas y ejecuta las observaciones según las restricciones impuestas por las condiciones astronómicas.

El T80S está equipado con la cámara óptica T80Cam-S que consta de un siste-

<sup>3</sup><https://github.com/astroufsc/chimera>



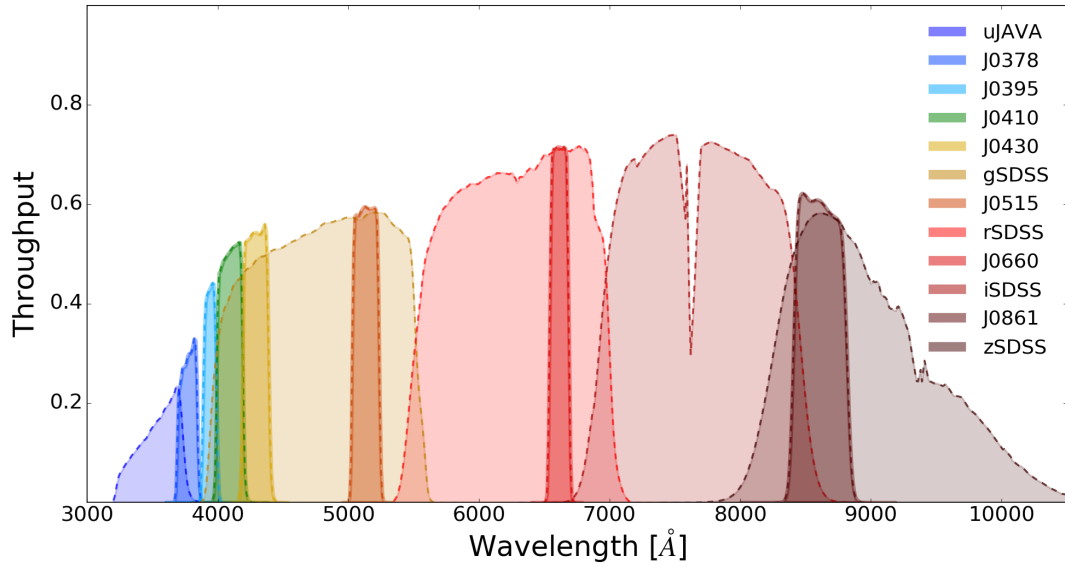
**Figura 1.4:** T80S y su cámara de campo amplio. Figura obtenida de (Mendes de Oliveira et al., 2019).

ma de 12 filtros distribuidos en dos ruedas de filtros, obturador, ventana de entrada, criostato, detector y el correspondiente sistema electrónico y de control. El CCD se lee con 16 amplificadores organizados en una matriz de  $8 \times 2$ . El detector puede funcionar a cuatro velocidades de lectura diferentes, y dos ganancias diferentes, con la opción de  $1 \times 1$  sin binneado o binneado  $2 \times 2$ .

Como se ilustra en la Figura 1.5, el sistema de filtros se compone de 7 filtros de banda estrecha (J0378, J0395, J0410, J0430, J0515, J0660, J0861) que coinciden, respectivamente, con las características [OII], Ca H+K,  $H\delta$ , banda G, triplete Mgb,  $H\alpha$  y triplete Ca. El sistema también incluye los filtros de banda ancha u, g, r, i y z que sirven para restringir el continuo espectral de las fuentes. Las bandas g, r, i y z son similares a las del SDSS, con algunas pequeñas diferencias de punto cero. El filtro de la banda u es el filtro u de Javalambre, que tiene una transmisión ligeramente más eficiente en comparación con la banda u de SDSS. Un resumen del sistema de filtros puede verse en la Tabla 1.1.

El proyecto S-PLUS cuenta con un survey principal, el Main Survey (MS), y otros 4 sub-surveys: el Ultra-Short Survey (USS), el Galactic Survey (GS), el Marble Field Survey (MFS) y el Variability Fields Survey (VFS). También cuenta con casi 90 proyectos internos específicos, direccionados por objeto/s astrofísico/s de interés, técnica/s a aplicar, objetivo/s científico/s a abordar, etc. Entre estos últimos se encuentra el proyecto interno #59 titulado “*The S-PLUS Fornax Project (S+FP)*”, coordinado por la Dra. Analía Smith Castelli (Smith Castelli et al., 2021), y dentro





**Figura 1.5:** El sistema Javalambre de 12 filtros. El eje y muestra la eficiencia total de los filtros S-PLUS, obtenida mediante la multiplicación de las curvas de transmisión media de los filtros, la transmisión atmosférica, la eficiencia del CCD y las curvas de reflectividad del espejo primario. Los distintos filtros están coloreados según las etiquetas que aparecen en la leyenda de la derecha. Figura obtenida de (Mendes de Oliveira et al., 2019).

**Tabla 1.1:** Resumen de los filtros S-PLUS. Tabla obtenida de (Mendes de Oliveira et al., 2019).

Nombre del filtro	$\lambda$ eff [Å]	$\Delta\lambda$ [Å]	Comentario
uJAVa	3574	330	Javalambre u
J0378	3771	151	[O II]
J0395	3941	103	Ca H+K
J0410	4094	201	H $\delta$
J0430	4292	200	G-band
gSDSS	4756	1536	SDSS-like g
J0515	5133	207	Mgb Triplet
rSDSS	6260	1462	SDSS-like r
J0660	6614	147	H $\alpha$
iSDSS	7692	1504	SDSS-like i
J0861	8611	408	Ca Triplet
zSDSS	8783	1072	SDSS-like z

del cual se enmarca el trabajo presentado en esta Tesis. El objetivo de dicho proyecto es realizar el estudio de la población de galaxias del cúmulo de Fornax utilizando los datos de S-PLUS.

### 1.3. Cúmulo de Fornax

La mayor concentración de galaxias en un radio menor a  $\sim 20$  Mpc de nuestra galaxia es, después de Virgo, el Cúmulo de Fornax. Este cúmulo, es la concentración de galaxias más importante del Hemisferio Sur Celeste y está centrado, aproximadamente, en  $\alpha \sim 3\text{h } 35\text{m}$  y  $\delta \sim -35,7^\circ$  (Ferguson, 1989). Se encuentra en la constelación de Fornax, que fue introducida en el siglo XVIII por Nicolas Louis de la Caille mientras cartografiaba los cielos australes desde el Cabo de Buena Esperanza, extremo sur de África, durante los años 1751-1753.

Por lo que sabemos, se lo identificó por primera vez como un auténtico cúmulo de galaxias en 1943 (Shapley, 1943), aunque ya se habían realizado catálogos extragalácticos en la dirección de Fornax (Baker (1933), Baker y Shapley (1937)). Sin embargo, aún faltaban muchos años para disponer de un catálogo más completo de los miembros del cúmulo (Jones y Jones, 1980). Incluso en 1956 NGC 1399, la luminosa galaxia cD/E0 situada en el centro dinámico del componente principal del cúmulo (Drinkwater et al., 2001), figuraba como galaxia no perteneciente al cúmulo en el catálogo de (Humason et al., 1956).

Los estudios sobre el Cúmulo de Fornax han seguido a menudo los pasos de estudios similares sobre el Cúmulo de Virgo. Por ejemplo, uno de los primeros estudios sobre Fornax reveló el descubrimiento de galaxias enanas (Hodge, 1959) después de que Reaves (1956) descubriera una población similar en Virgo. Los primeros catálogos de los miembros del cúmulo de Fornax (Hodge, 1960; Hodge et al., 1965) fueron actualizados por Ferguson (1989) a través de su Catálogo del Cúmulo de Fornax (FCC, por sus siglas en inglés), quien estudió 40 grados cuadrados utilizando, a gran escala, placas fotográficas del telescopio de 2,5 m del Observatorio de Las Campanas (Chile). Este mismo telescopio fue utilizado por Binggeli et al. (1985) unos años antes para construir el Catálogo del Cúmulo de Virgo (VCC, por sus siglas en inglés).

Los sistemas de cúmulos globulares (CGs) de las galaxias más brillantes de Fornax se detectaron por primera vez de forma provisional alrededor de NGC 1374, NGC 1379 y NGC 1399 llevando a cabo un recuento de estrellas en una única placa fotográfica profunda (Dawe y Dickens, 1976), dos décadas después de que se hablara por primera vez de una población similar en M 87, la elíptica gigante situada en el centro del cúmulo de Virgo (Baum, 1955). Más tarde, Hanes y Harris (1986) presentaron un estudio sistemático de los sistemas de cúmulos globulares en Fornax como parte de una serie de artículos sobre cúmulos globulares extragalácticos, que habían incluido el estudio de cúmulos globulares en galaxias de Virgo.

En la importantísima tarea de establecer la escala de distancias cosmológicas, las Cefeidas fueron utilizadas por el Proyecto Clave del Hubble Space Telescope (HST) sobre la Escala de Distancias Extragalácticas para derivar una distancia a Fornax (Madore et al., 1999) poco después de que el mismo programa hubiera identificado

Cefeidas en Virgo (Freedman et al., 1994; Ferrarese et al., 1996). Tras pruebas iniciales de viabilidad del indicador de distancia de las fluctuaciones del brillo superficial (SBF, por sus siglas en inglés; Tonry y Schneider, 1988) en M 32 y NGC 3379, el método se aplicó a un número significativo de galaxias del Cúmulo de Virgo (Tonry et al., 1989). Sin embargo, debido a las complicaciones derivadas de la considerable profundidad de la línea de visión de Virgo, la primera calibración precisa del método SBF tuvo que esperar a observaciones posteriores de galaxias en el cúmulo de Fornax (Tonry, 1991). Igualmente ya se habían realizado varias estimaciones de distancia a Fornax utilizando otras técnicas, por ejemplo CGs (de Vaucouleurs, 1977).

A pesar de su proximidad compartida con nuestra Galaxia, los cúmulos de Virgo y Fornax muestran algunas diferencias. El cúmulo de Fornax tiene una forma mucho más regular, y probablemente más evolucionado dinámicamente debido a su relativa baja dispersión de velocidades respecto a su homóloga del norte, lo cual ya había sido sugerido previamente (Reaves, 1964). A su vez, el cúmulo de Fornax muestra segregación en masa (Drinkwater et al., 2001), es decir, que las galaxias menos masivas/luminosas se distribuyen dentro del cúmulo diferente a las más masivas/brillantes; y sería más viejo dinámicamente que los Cúmulos de Coma y Perseo, nuevamente por estar relativamente más virializado (Struble, 1979). También es considerablemente más pequeño y denso que el cúmulo de Virgo, con un radio del núcleo  $\sim 40\%$  del de este último y una densidad central dos veces mayor. La masa total del cúmulo de Fornax es  $\approx 7 \times 10^{13} M_{\odot}$ , según se estima a partir de las velocidades radiales de las galaxias (Drinkwater et al., 2001), que es  $\sim 1/10$  la del cúmulo de Virgo. De esta forma el cúmulo de Fornax es más representativo de los grupos y cúmulos pobres en los que residen la mayoría de las galaxias del universo. Suponiendo que la función de luminosidad de las galaxias no es una función universal, Jones y Jones (1980) sugiere que en el cúmulo de Fornax existe una marcada diferencia entre la función de luminosidad de las galaxias de tipo temprano (E y SO) y la de las galaxias de tipo tardío. La primera tiene la misma forma que la función de luminosidad total del cúmulo, mientras que la segunda se asemeja más a una función gaussiana. Donde la función de luminosidad total del cúmulo es inconsistente con la función de Schechter (1976) y parece ser consistente con la sugerida por Holmberg (1969) y por Tammann y Kraan (1978).

Al ser un cúmulo mucho más compacto que Virgo, Fornax es un objetivo ideal para la calibración de indicadores de distancia. Sin embargo, como la mayoría de los cúmulos, el cúmulo de Fornax muestra subestructuras: el subcúmulo principal está centrado en NGC 1399, mientras que un subcúmulo que incluye a NGC 1316 (Fornax A) está centrado  $3^{\circ}$  al suroeste (Drinkwater et al., 2001). Fornax A fue identificada como una radio fuente (Shain, 1958) que inicialmente no fue asociada con NGC 1316, pero luego de una década se asoció con dicha galaxia (Searle, 1965). Una imagen en falso color de NGC 1316 puede verse en la Figura 1.6. Según Drinkwater et al. (2001), el subcúmulo de Fornax A tiene una velocidad de recesión de  $105 \pm 102 \text{ km s}^{-1}$  mayor que el subcúmulo principal. Aunque está aún en discusión, todo parece indicar que el subcúmulo de Fornax A está cayendo hacia la estructura principal.

El subcúmulo de Fornax A es un lugar de intensa formación estelar, que contiene el 15 % de todas las galaxias de la muestra total de Fornax, pero el 31 % de las galaxias con formación estelar. Este subcúmulo contiene dos de las cuatro galaxias del cúmulo con tasas de formación estelar excepcionalmente altas, NGC 1341 y FCC 33, y una gran cantidad de hidrógeno neutro (Putman et al., 1998), que probablemente no ha sobrevivido al paso por el núcleo del cúmulo. Además, las observaciones de Fornax A en el subcúmulo (Ekers et al., 1983) indican que tiene una velocidad proyectada de aproximadamente  $80 \text{ km s}^{-1}$  hacia el norte, consistente con una caída. Este grupo contiene a NGC 1404, que tiene una envoltura de rayos X distorsionada indicativa de la caída (Jones et al., 1997), y la galaxia irregular NGC 1427A, que muestra signos de haber sido perturbada en su primer paso por el cúmulo (Chanamé et al., 2000). Las morfologías de estas dos galaxias son indicativas de la presencia de material que se desprende de ellas a medida que se acercan al centro del cúmulo. La identificación de estas subestructuras en el cúmulo de Fornax han influido en la determinación de su distancia.



**Figura 1.6:** Imagen en falso color RGB de NGC 1316 obtenida a partir del paquete de Python `make_lupton_rgb`, en la cual se utilizó el filtro J660 en color rojo, gSDSS en verde y J378 combinado con uJAVA para el azul.

En lo que a relevamientos fotométricos en curso del cúmulo de Fornax se refiere, podemos mencionar el *Fornax Deep Survey (FDS)* (Iodice et al., 2016), el cual pre-

tende cubrir el cúmulo de Fornax hasta su radio virial ( $R_{vir} \sim 0.7$  Mpc; Drinkwater et al., 2001), con un área de unos  $26 \text{ grad}^2$  alrededor de la galaxia central NGC 1399, incluyendo la región donde se encuentra NGC 1316. El FDS da una visión sin precedentes de las estructuras de los miembros del cúmulo, que van desde galaxias gigantes de tipo temprano hasta pequeñas galaxias esferoidales. Las imágenes profundas multibanda permiten cartografiar la distribución de la luz y los colores de las galaxias del cúmulo hasta 8-10 radios efectivos, con el fin de estudiar los discos y barras interiores, el débil halo estelar, incluida la componente de luz difusa, y los restos de marea que evidencian recientes eventos de canibalismo. Además, FDS también proporciona una caracterización bastante completa de los sistemas estelares pequeños alrededor de las galaxias, como los CGs y galaxias enanas ultracompactas (UCD, por sus siglas en inglés). La profundidad esperada para el FDS es de  $28,0 \text{ mag}^{-2}$  en la banda u,  $28,6 \text{ mag}^{-2}$  en la banda g,  $28,1 \text{ mag}^{-2}$  en la banda r, y  $27,2 \text{ mag}^{-2}$  en la banda i.

Otro relevamiento fotométrico reciente del cúmulo de Fornax, con similar profundidad al FDS, es el *Next Generation Fornax Survey (NGFS)* (Muñoz et al., 2015; Eigenthaler et al., 2018). Este relevamiento pretende cubrir un área de  $30 \text{ grados}^2$  en las bandas ópticas u, g, i e infrarroja cercana  $K_s$  utilizando la *Dark Energy Camera (DECam)* acoplada al telescopio Blanco de 4 m en el Observatorio Interamericano de Cerro Tololo (CTIO, Chile) para las bandas ópticas u, g e i, y VISTA/VIRCAM (Sutherland et al., 2015) para la banda  $K_s$ . Hasta ahora, el NGFS ha publicado un catálogo de galaxias que cubre, igual que FDS, el área dentro del radio virial del cúmulo de Fornax (Eigenthaler et al., 2018; Ordenes-Briceño et al., 2018) contabilizando 643 galaxias enanas en total. Ambos relevamientos, aunque profundos, solo incluyen 4 filtros fotométricos de banda ancha en comparación a los 12 filtros de S-PLUS. Además, como se mencionó anteriormente, la cobertura areal tanto de FDS como de NGFS solo se restringe al área interna del radio virial, mientras que S-PLUS, por sus imágenes de gran campo, cuenta con la posibilidad de lograr una cobertura areal mucho más extensa, como se muestra en la Sección 2.

## 1.4. Motivación de esta Tesis

Como se mencionó anteriormente, el cúmulo de Fornax es el segundo cúmulo de galaxias más rico y más cercano ( $D \sim 20$  Mpc) después del cúmulo de Virgo, por lo que representa un entorno extraordinario en el hemisferio sur celeste. Aunque el cúmulo de Fornax históricamente ha sido muy estudiado, nunca ha sido analizado, simultáneamente, con imágenes de campo amplio ( $1.4 \times 1.4 \text{ grad}^2$ ) y 12 filtros ópticos (5 de banda ancha similares a los del SDSS y 7 de banda angosta que trazan características espectroscópicas como [OII],  $H\alpha$  y el triplete del Ca, entre otros), como los provistos por el relevamiento S-PLUS. El proyecto Fornax de S-PLUS (S+FP, por sus siglas en inglés) tiene como objetivo realizar un estudio exhaustivo de dicho cúmulo de galaxias utilizando imágenes y catálogos correspondientes a 43 campos observados como parte de la cuarta liberación de datos interna (iDR4) de S-PLUS,

---

que cubren un área de  $\sim 45 \text{ grad}^2$  hasta un radio de  $\sim 5$  grados desde el centro del cúmulo ( $\sim 2,5 R_{vir}$ ). A partir del cruce de los catálogos de la iDR4 de S-PLUS con catálogos de galaxias reportadas como miembros del cúmulo de Fornax en la literatura, notamos que algunas galaxias enanas y brillantes no estaban incluidas en los primeros. Esto se debe a que los parámetros del software **SExtractor** utilizados para obtener los catálogos de la iDR4 de S-PLUS tienen como objetivo proporcionar datos útiles tanto para proyectos galácticos como extragalácticos. En esta Tesis presentamos un esfuerzo por identificar el conjunto de parámetros de **SExtractor** que resulte más adecuado para detectar específicamente galaxias previamente catalogadas como miembros del cúmulo de Fornax en las imágenes de S-PLUS. Además, pretendemos automatizar el proceso de ejecución de **SExtractor** sobre las imágenes de S-PLUS mediante su implementación dentro de un código escrito en Python. Esto facilitará la ejecución de **SExtractor** sobre campos adicionales de S-PLUS correspondientes a otras zonas del hemisferio sur celeste.

# Capítulo 2

## Datos y herramientas

### 2.1. Imágenes y catálogos de S-PLUS

Hemos considerado dos conjuntos de datos para este trabajo. El primer conjunto incluye 43 campos S-PLUS observados como parte de la tercera liberación de datos de la colaboración (DR3):

- 23 campos que cubren galaxias reportadas en la literatura como miembros espectroscópicamente confirmados o miembros probables del cúmulo de Fornax.
- 20 campos que cubren los alrededores del cúmulo de Fornax.

El segundo conjunto incluye nuevos campos S-PLUS obtenidos a partir de la cuarta liberación de datos interna a la colaboración (iDR4):

- 20 campos que cubren galaxias de Fornax reportadas en la literatura no cubiertas por el conjunto anterior.
- 43 campos adicionales que cubren los alrededores del cúmulo de Fornax.

Por tanto, en el marco de esta Tesis, hemos trabajado sobre imágenes en 12 filtros obtenidas para un total de 106 campos de S-PLUS ( $12 \times 106 = 1272$  imágenes). Para todos ellos se cuenta con catálogos con fotometría homogénea obtenidos como parte de la iDR4 dado que la iDR4 incluye a la DR3, y pueden verse en la Figura 1. Entonces, el total de campos sobre los que se trabajó, incluyen:

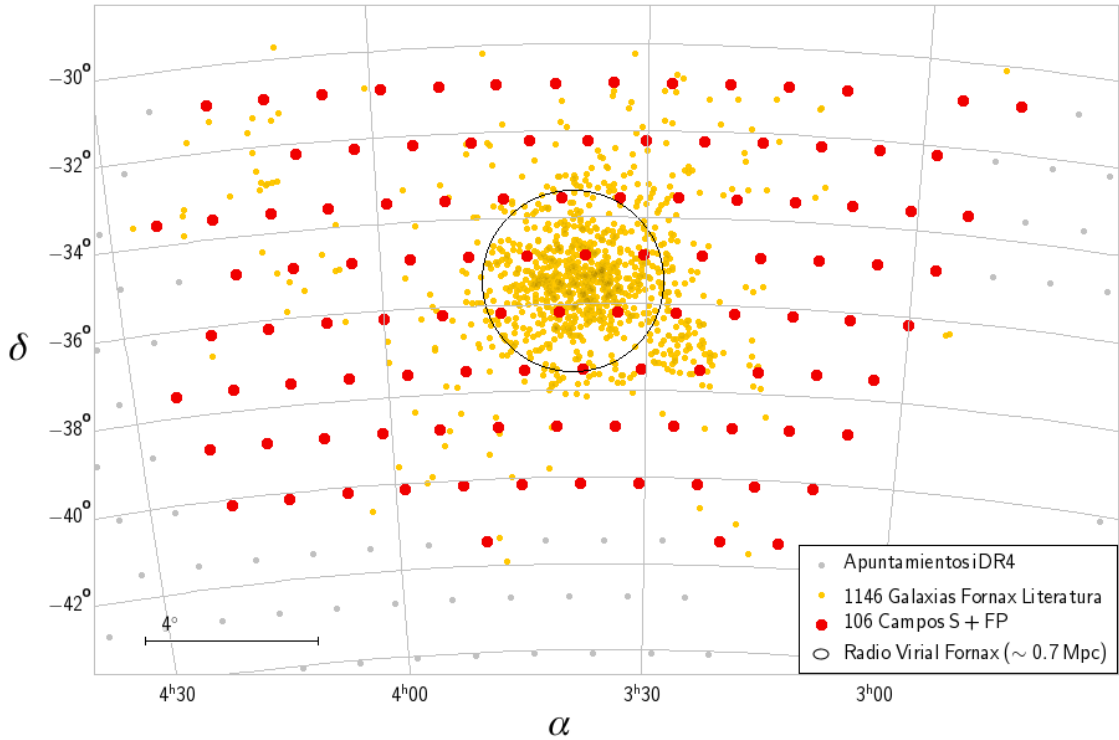
- 43 campos que cubren galaxias de Fornax reportadas en la literatura.
- 63 campos que cubren los alrededores del cúmulo de Fornax.

Cada una de las imágenes de S-PLUS con las que se trabajó tienen una resolución de  $0.55'' \text{ píxel}^{-1}$  y un tamaño total de  $11.000 \times 11.000$  píxeles, lo cual se traduce en un campo de  $\sim 1,4 \times 1,4 \text{ grad}^2$ . Vale la pena aclarar que durante la construcción del

código se utilizaron, en primera instancia, los 43 campos obtenidos como parte de la DR3, para los cuales se obtuvieron los catálogos finales y, luego de tener el código funcionando completamente, se agregaron los 63 campos que sumó la iDR4 y se volvieron a obtener nuevos catálogos. Este último paso fue muy sencillo por el nivel de automatización que posee el código. Simplemente hubo que proveer como entrada la lista de nuevos campos y ejecutar una vez más el código. A modo de referencia, cabe mencionar que el código fue ejecutado repetidamente en una computadora con un procesador Intel(R) Core(TM) i5-9400F CPU @ 2.90GHz y 32 GB de RAM, demorando 24 hs de ejecución en promedio.

Cabe mencionar que la identificación de galaxias de Fornax reportadas en la literatura sobre los campos de S-PLUS se hizo mediante la utilización de la recopilación de objetos incluidos en 16 catálogos de galaxias de Fornax realizada por la Dra. Smith Castelli en el marco del S+FP. Dicha lista incluye un total de 1006 galaxias, de las cuales  $\sim 850$  son identificables visualmente en las imágenes de S-PLUS debido a la profundidad de las mismas ( $\mu_R \sim 23 \text{ mag}^{-2}$ ).

Es necesario comentar que el proceso de calibración fotométrica de las imágenes puede encontrarse detallado en (Almeida-Fernandes et al., 2022). Es bueno destacar que: las magnitudes fotométricas se calibran al sistema estándar AB, no se realiza corrección por extinción interestelar y la precisión astrométrica utilizada es  $\Delta\alpha = 0.010 \text{ arcsec}$  y  $\Delta\delta = 0.027 \text{ arcsec}$ .



**Figura 2.1:** Distribución espacial proyectada de los campos obtenidos a partir de las DR3+iDR4 en Fornax+Alrededores y de las galaxias de Fornax reportadas en la literatura.



## 2.2. SExtractor

### 2.2.1. Generalidades

Source Extractor (SExtractor; Bertin y Arnouts, 1996) es un software que detecta, desdobra, mide y clasifica en estrella/galaxia las fuentes presentes en imágenes astronómicas. La separación estrella/galaxia se realiza utilizando una red neuronal entrenada con imágenes simuladas. Las características más destacadas de SExtractor incluyen su capacidad para trabajar con imágenes de gran tamaño (de hasta  $60.000 \times 60.000$  píxeles) y para tratar con una amplia variedad de objetos. Por lo tanto, es especialmente adecuado para el análisis de imágenes obtenidas en el marco de grandes relevamientos astronómicos. A continuación, se introducirán aspectos relevantes al presente trabajo, los cuales se encuentran explicados con mayor detalle en Bertin y Arnouts (1996).

Resumidamente, los pasos que realiza el software pueden enumerarse de la siguiente manera:

1. Medir el background y su ruido RMS.
2. Restar el background.
3. Filtrar (convolucionar con un perfil específico).
4. Encontrar objetos.
5. Separar detecciones (dividir la detección en diferentes objetos).
6. Medir formas y posiciones.
7. Limpiar (revisar detecciones, contabilizando las contribuciones de vecinos).
8. Realizar la fotometría.
9. Clasificar las detecciones (separación estrella/galaxia).
10. Generar catálogos con la información extraída de los objetos detectados y medidos en las imágenes.
11. Generar diferentes imágenes de control.

Los pasos que involucran medir y restar el background, filtrado, detección y separación o “deblending”, se detallan en la Sección 3.5.

### 2.2.2. Obtención de magnitudes totales

Además de las magnitudes isofotales y de abertura circular, **SExtractor** tiene la posibilidad de estimar la magnitud "total" de un objeto. La magnitud total aparente es el parámetro fotométrico más relevante desde el punto de vista físico que se puede extraer de las imágenes. Para los objetos no estelares, la fotometría isofotal y de abertura tienen tanto ventajas como desventajas cuando se utilizan para estimar las magnitudes totales. Se sabe que la fotometría de abertura es generalmente menos sesgada que la fotometría isofotal, pero sólo funciona en regiones no densamente pobladas. Estas limitaciones, y la necesidad de mantener el procesamiento lo más rápido posible, llevaron a adoptar el siguiente esquema. Para cada objeto, se calculan dos tipos de magnitud total: el primer tipo utiliza una abertura adaptativa y el segundo, una corrección isofotal.

- **Magnitudes de abertura adaptativa:** La rutina de fotometría de abertura adaptativa se inspira en el algoritmo de [Kron \(1980\)](#). Primero los momentos de segundo orden del perfil del objeto se utilizan para definir un perfil gaussiano equivalente con una desviación estándar media  $\sigma_{iso}$ . Luego, una abertura elíptica cuya elipticidad y ángulo de posición  $\theta$  están definidos por estos momentos, se escala a  $6 \sigma_{iso}$ , que corresponde aproximadamente a 2 "radios" isofotales. Finalmente, dentro de esta abertura calculamos el "primer momento" según:

$$r_1 = \frac{\sum rI(r)}{\sum I(r)} \quad (2.1)$$

[Kron \(1980\)](#) e [Infante \(1987\)](#) han demostrado que para las estrellas y perfiles de galaxias convolucionados con un seeing gaussiano se espera que una fracción casi constante del flujo se encuentre dentro de una abertura circular de radio  $kr_1$ , independientemente de su magnitud. Esto no cambia si consideramos una elipse con ejes principales  $\epsilon kr_1$  y  $kr_1/\epsilon$ . Se consigue, entonces, un equilibrio entre los errores sistemáticos y aleatorios para  $k \sim 2$ . Si se toma  $k = 2,5$ , la fracción media de flujo perdida es de aproximadamente del 6%. Cuando la relación S/N es muy baja, el algoritmo puede considerar aberturas erróneamente pequeñas. Es por ello que debe acotarse la menor abertura permitida a través de un valor  $R_{min}$  que, típicamente, se considera como  $3-4\sigma_{iso}$

- **Magnitudes isofotales corregidas:** Si hacemos la suposición de que los perfiles de intensidad de los objetos débiles registrados en las imágenes son aproximadamente gaussianos al atravesar la atmósfera, entonces la fracción  $\eta = \frac{I_{iso}}{I_{tot}}$  del flujo total encerrado en una isofota concreta esta dado por:

$$\left(1 - \frac{1}{\eta}\right) \ln(1 - \eta) = \frac{A.t}{I_{iso}} \quad (2.2)$$

donde  $A$  es el área y  $t$  el umbral relacionado con esta isofota. La ecuación [2.2](#) no es analíticamente invertible, pero es una muy buena aproximación a  $\eta$  y puede hacerse con el ajuste polinómico de segundo orden:

$$\eta \approx 1 - 0.1961 \frac{A \cdot t}{I_{\text{iso}}} - 0.7512 \left( \frac{A \cdot t}{I_{\text{iso}}} \right)^2 \quad (2.3)$$

Una estimación de la magnitud total  $m_{\text{tot}}$  es entonces:

$$m_{\text{tot}} = m_{\text{iso}} + 2.5 \log \eta \quad (2.4)$$

Está claro que esta sencilla corrección funciona mejor con las estrellas, pero da resultados bastante precisos con galaxias de disco. Aun así, las alas anchas de los perfiles de las galaxias esferoidales pueden conducir a grandes errores cuando se utiliza un umbral isofotal alto.

La estrategia que utiliza `SExtractor` para dar la mejor estimación de la magnitud total de un objeto es tomar el método de abertura adaptativa, excepto si se considera que un objeto vecino puede llegar a sesgar la magnitud en más de 0,1 mag. En estos casos, que normalmente son menos del 20% en campos de alta latitud galáctica y en exposiciones moderadamente profundas, se toma la magnitud isofotal corregida.

### 2.2.3. Separación estrella/galaxia

Una buena discriminación entre objetos resueltos y no resueltos o, para simplificar la terminología, entre galaxias y estrellas, resulta ser esencial para los estudios extragalácticos. La suposición principal es que las imágenes de las galaxias se ven más extendidas que las de las estrellas o cuasares (QSO, por sus siglas en inglés). Un gran número de diferentes clasificadores de estrellas/galaxias han sido descritos y aplicados hasta la actualidad ([Slater et al., 2020](#)). Entre los estimadores simples de 2 parámetros se encuentran el área isofotal, la magnitud isofotal y el brillo superficial central, los cuales proporcionan formas sencillas de separar las dos clases. Sin embargo, estos estimadores simples no utilizan toda la información disponible en el perfil de intensidad. Otros estimadores más sofisticados han sido desarrollados para hacerlo y han mejorado en este aspecto, pero comparten un defecto: la falta de robustez, especialmente cuando se enfrentan a objetos fusionados o con objetos que tienen vecinos cercanos.

Si uno considera a un objeto como un vector de parámetros, clasificar estrellas y galaxias de forma óptima no es más que encontrar la mejor hipersuperficie de frontera entre las dos clases en el espacio de los parámetros. En este punto, las redes neuronales (NN) proporcionan una herramienta extremadamente robusta para determinar la mejor frontera entre objetos resueltos y no resueltos. Pueden manejar fácilmente casos problemáticos como los objetos fusionados o los vecinos cercanos, siempre que hayan aprendido a reconocerlos. Sin embargo, estas redes neuronales necesitan ser entrenadas con un conjunto de prototipos antes de ser utilizadas. Una NN es básicamente un grupo de unidades conectadas llamadas neuronas cuyo comportamiento se inspira en las neuronas biológicas reales. Una de las características más interesantes de las NN es su capacidad de "aprender" y "generalizar" en un contexto determinado. Esto ha llevado a utilizarlas en un número cada vez mayor de problemas de reconocimiento de patrones complejos que van desde la identificación de palabras habladas o caracteres escritos a mano, hasta la discriminación de partículas en la física de altas energías.

La selección y etiquetado de prototipos a partir de imágenes reales es una tarea manual y bastante tediosa que sólo puede justificarse para una gran cantidad de datos homogéneos. Desgraciadamente, la mayoría de los estudios consisten en conjuntos de imágenes de calidad variable para los que cada uno requeriría en principio un entrenamiento diferente. Una solución a esto puede ser realizar el entrenamiento con un conjunto homogéneo y aplicar una transformación adecuada a los parámetros de entrada de otros conjuntos, para recuperar una distribución similar de objetos en el espacio de parámetros. Pero esto no es satisfactorio, ya que vuelve a plantear la necesidad de una parametrización, y sobre todo no hay garantía de que la transformación sea lo suficientemente precisa para recuperar todo el potencial de la clasificación neuronal, especialmente para los "valores atípicos". De hecho, las imágenes astronómicas producidas por conjuntos de detectores modernos comparten muchas similitudes: la escala de intensidad es lineal con una buena precisión en un amplio rango hasta que aparece una saturación brusca, el ruido de fondo es esencialmente ruido blanco, y el desenfoque atmosférico hace que el núcleo de la *point spread function* (PSF) sea considerada como una función gaussiana. Por lo tanto, las imágenes correctamente muestreadas, lo cual está dado por un "ancho-altura-media" o FWHM (por sus siglas en inglés) mayor a 2 píxeles, pueden describirse aproximadamente con tres números: la escala de píxeles, la profundidad o relación señal-ruido (S/N) y el FWHM. Asumiendo que las imágenes astronómicas actuales comparten muchas propiedades comunes, se puede pensar en utilizar imágenes artificiales similares, lo suficientemente precisas como para entrenar la red neuronal, en lugar de utilizar imágenes reales.

En el caso de la separación estrella/galaxia, SExtractor ha elegido una NN con arquitectura de red multicapa y un procedimiento de aprendizaje conocido como retropropagación. De todas las redes neuronales, es probablemente una de las mejor estudiadas, y se ha aplicado intensamente con éxito a muchas tareas de clasificación. Estas redes tienen una capa de entrada, una de salida y una o más capas "ocultas"

entre ellas. Las neuronas de una capa están conectadas a todas las neuronas de la capa anterior, pero no hay conexión entre las neuronas de una misma capa. En un clasificador, los nodos de entrada reciben los parámetros que definen el objeto, y los nodos de salida contienen "índices de afiliación" a cada clase. La red neuronal se utiliza en dos modos:

- Un modo de "entrenamiento" en el que la red "aprende" a asociar las entradas y las salidas deseadas que se le presentan repetidamente (aprendizaje supervisado).
- Un modo de "reproducción" en el que simplemente responde a nuevos patrones según el entrenamiento previo.

En la red multicapa que se ha considerado para **SExtractor**, una neurona  $i$  en la capa  $l$  tiene la siguiente salida:

$$\mathbf{o}_i^{(l)} = g\left(\mathbf{w}_i^{(l)} \cdot \mathbf{o}^{(l-1)} - b_i^{(l)}\right) \quad (2.5)$$

donde  $\mathbf{w}_i^{(l)}$  es el vector de pesos de las conexiones entre las neuronas de la capa anterior  $l - 1$  y la neurona  $i$ , y  $b_i^{(l)}$  es un término de sesgo que puede considerarse como el peso de una conexión con la salida de una neurona de salida. La función de activación  $g$  es la función sigmoidea habitual:  $g(u) = 1 / (1 + e^{-u})$ . El procedimiento de entrenamiento es organizado de la siguiente manera. En primer lugar, los pesos y los sesgos se inicializan con valores aleatorios. A continuación, para cada patrón, de un "conjunto de entrenamiento" cargado en la capa de entrada en la iteración  $t$ , se realiza una propagación hacia adelante a través de la red utilizando la ecuación [2.5](#). Los pesos y los sesgos se ajustan capa por capa hacia atrás, según la regla de aprendizaje de retropropagación de errores:

$$\Delta \mathbf{w}_j^{(l)}(t+1) = -\eta \nabla E\left(\mathbf{w}_j^{(l)}\right) + \alpha \Delta \mathbf{w}_j^{(l)}(t) \quad (2.6)$$

Aquí  $E$  es una función de coste cuadrática que mide el error de la red:

$$E = \frac{1}{2} |\mathbf{O} - \mathbf{o}|^2, \quad (2.7)$$

donde  $\mathbf{O}$  es el vector de salida deseado y  $\mathbf{o}$  la respuesta de la red al patrón de entrenamiento. La "tasa de aprendizaje" de la red viene determinada por  $\eta$ , mientras que el parámetro de impulso  $\alpha$  puede establecerse en un valor distinto de cero

(pero  $< 1$ ) para acelerar la convergencia en las regiones planas de la superficie de costes y evitar que la red se deje engañar por las fluctuaciones locales. En SExtractor, todos los aprendizajes se han realizado con  $\eta = 0,1$  y  $\alpha = 0.2$ .

A medida que avanza el aprendizaje,  $E$  disminuye lentamente. Con datos experimentales, a menudo se necesita un millón de pasos de aprendizaje para que  $E$  alcance un mínimo estable, lo que significa que es necesario un número grande de iteraciones en la etapa de entrenamiento.

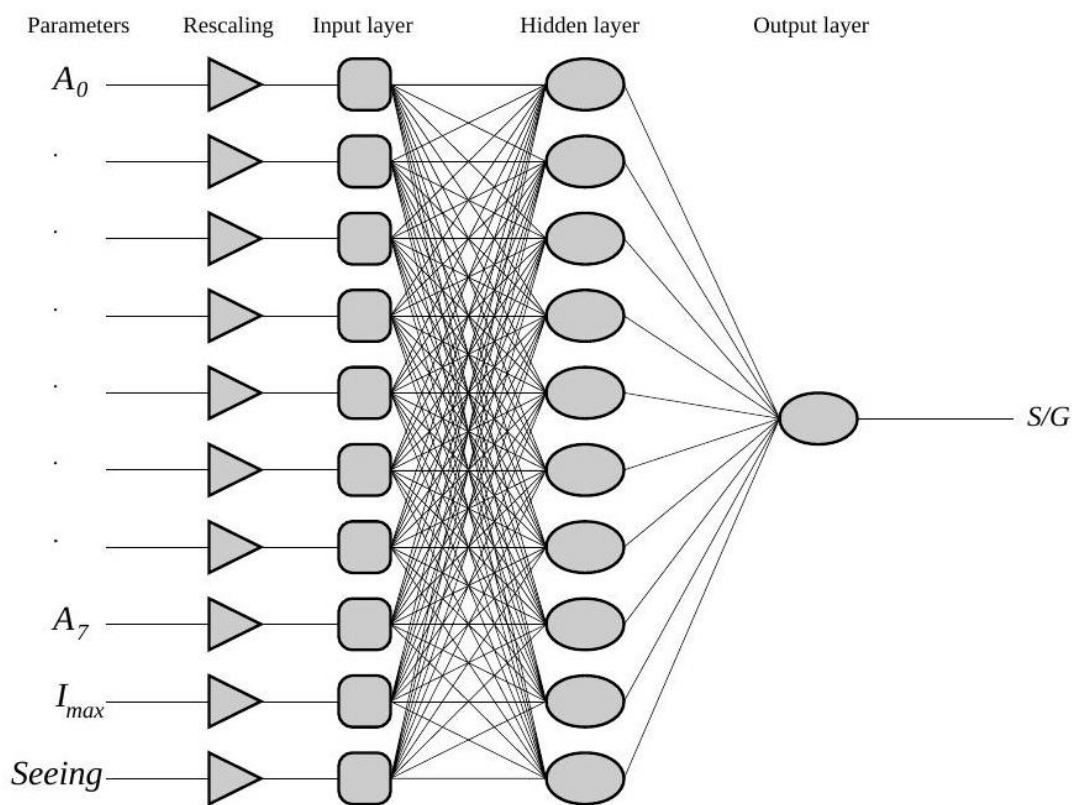
La elección de los parámetros de entrada correctos es un paso crucial en la separación estrella/galaxia. Para elegir los parámetros "óptimos" se ha tenido en cuenta que deban poder:

- i) Discriminar eficazmente las dos clases en toda la gama de rango de magnitudes accesible con cualquier detector lineal.
- ii) Ser invariantes a la traslación y a la rotación.
- iii) Tener cierta robustez con respecto al ruido, las distorsiones de la imagen y la influencia de los vecinos cercanos.
- iv) Ser lo más independiente posible de las características de la exposición.

Los parámetros simples que cumplen estos requisitos son las áreas isofotales. Una buena representación de los perfiles de la imagen puede hacerse combinando algunas áreas isofotales: A0 (correspondiente a la isofota más baja), A1, A2, etc. y el valor máximo del píxel sobre el cielo  $I_{\text{max}}$ . En el caso de objetos tenues, para que el clasificador funcione mejor, se toman varias áreas isofotales además de A0, ya que al ser objetos débiles  $I_{\text{max}}$  tendrá una incertidumbre relativamente alta, implicando poco contraste entre las dos clases. Esto justifica la necesidad de agregar parámetros para lograr una separación más precisa. En el caso de objetos brillantes, no es así dado que  $I_{\text{max}}$  llega a el nivel de saturación. No se incluye ninguna medida de elongación en el conjunto de parámetros para cumplir con iii); de hecho, la elipticidad y los parámetros geométricos más sofisticados no son de gran importancia en la separación estrella/galaxia, sobre todo para los objetos débiles.

Así, se termina con diez parámetros: ocho áreas isofotales, la intensidad de pico, y un "parámetro de control" que es el *seeing*. Las áreas isofotales se dan a la red en unidades de "seeing FWHM" al cuadrado, lo que elimina la necesidad de información de la escala de píxeles y se asume que las imágenes están correctamente muestreadas. No debería haber ninguna diferencia en cuanto a los perfiles extraídos, excepto el ruido, entre una exposición profunda analizada con un umbral alto, y otra menos profunda analizada con un umbral más bajo, por lo que se puede dar la información de profundidad expresando la intensidad del pico en unidades del umbral de extracción. Debido al amplio rango de luminosidad de los objetos que pueden aparecer en la imagen, que son resultado del amplio rango dinámico de las

imágenes astronómicas actuales, la red neuronal se alimenta con los logaritmos de los parámetros. En la Figura 2.2 se muestra un diagrama esquemático de la red neuronal que se utiliza para realizar la clasificación estrella/galaxia.



**Figura 2.2:** Diagrama de la red neuronal utilizada por SExtractor para realizar la clasificación estrella/galaxia. Figura obtenida de (Bertin y Arnouts, 1996).

La capa de salida contiene una única neurona dado que "estrella" y "galaxia" son dos clases mutuamente excluyentes. El parámetro de salida que indica esta clasificación es un "índice de estelaridad" denominado `CLASS_STAR`, el cuál es una variable punto flotante que puede tomar valores entre 0 y 1, y representa la probabilidad de que el objeto clasificado sea una fuente no resuelta o fuente puntual. Un valor para `CLASS_STAR` cercano a 0 significa que es muy probable que el objeto clasificado sea una galaxia mientras que un valor cercano a 1 indica que el objeto es muy probablemente una estrella. Si bien en la realidad los datos reales difieren al menos levemente de la muestra de entrenamiento y el valor de `CLASS_STAR` resulta ser una aproximación pobre de la probabilidad esperada a posteriori, un valor de este parámetro cercano a 0 o 1 usualmente puede ser considerado como una clasificación confiable.

### 2.2.4. Ejecución de SExtractor

SExtractor puede ser ejecutado de dos maneras diferentes: en modo simple o modo dual.

El modo simple consiste en ejecutar SExtractor solo sobre una imagen. El software necesita una serie de parámetros para ejecutar el programa, y éstos se pueden dar en una línea de comando o mediante un archivo de configuración. Para ejecutar SExtractor en modo simple el usuario puede escribir:

```
1 sex image -c config_file.txt -PARAMETER1 val1 -PARAMETER2 val2
```

donde:

- *'sex'* es la sentencia de ejecución de SExtractor.
- *'image'* es el nombre de la imagen de detección y medición.
- *'config\_file.txt'* es el directorio del archivo de configuración a utilizar.
- *'-PARAMETER1/-PARAMETER2/etc.'* son los nombres de los parámetros que se desea modificar del archivo de configuración para realizar la ejecución.
- *'val1/val2/etc.'* son los valores que se le asignarán a los parámetros mencionados en el ítem anterior, respectivamente.

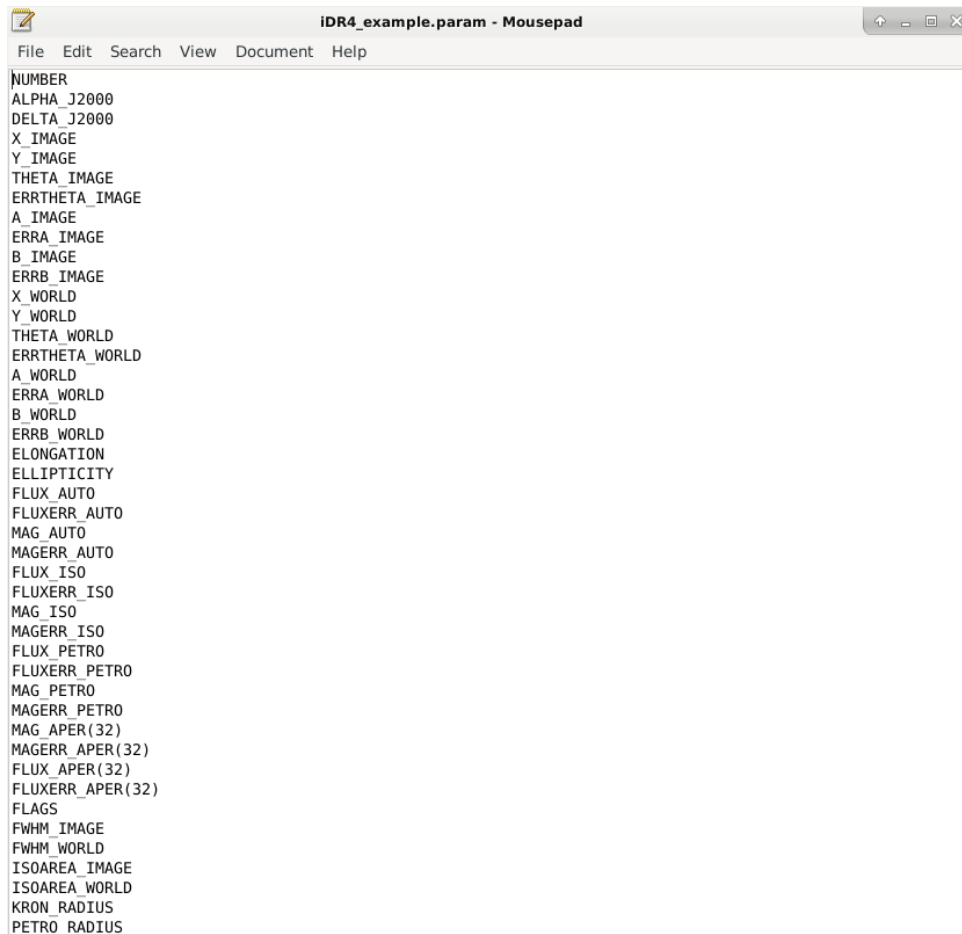
Nótese que, manteniendo la misma estructura sintáctica, es posible modificar la cantidad de parámetros que se desee. Si no existe un archivo de configuración dado, SExtractor tratará de leer el archivo *'default.sex'* del directorio local.

El modo dual se da al ejecutar SExtractor en imágenes separadas para detección y fotometría. El software permite encontrar las posiciones y aberturas de todos los objetos en una primer imagen y luego utilizar esa información sobre una segunda imagen para obtener la fotometría. Para realizar este proceso las imágenes deben estar correctamente registradas (coherentemente recortadas, alineadas y con su respectivo header asignado), y el ID que se asigna a un objeto es único en todos los catálogos que resultan de cada ejecución en modo dual, salvo que se cambie algún parámetro entre las ejecuciones. De este modo, la identidad de cada detección es única e independiente del filtro en que se esté midiendo, lo que facilita el manejo de grandes volúmenes de datos. Si se desea realizar la ejecución en modo dual, al igual que en el modo simple, se debe explicitar el archivo de configuración, opcionalmente asignar valores a parámetros específicos, pero en este caso deben darse como entrada ambas imágenes: la de detección y la de medición. Para ello, el usuario puede escribir:

```
1 sex image1,image2 -c config_file.txt
```



Las columnas que contendrán los catálogos resultantes de la ejecución de **SExtractor** se encuentran definidas en un archivo de entrada (formato ASCII plano) de extensión ".param" que puede ser editado por el usuario de acuerdo a sus necesidades. Un ejemplo puede verse en la Figura 2.3. **SExtractor** provee un archivo inicial denominado 'default.param' que puede ser tomado como referencia de posibles parámetros de salida. Una descripción de los parámetros más relevantes puede encontrarse en la Figura 3.6 en el siguiente Capítulo.



```
NUMBER
ALPHA_J2000
DELTA_J2000
X_IMAGE
Y_IMAGE
THETA_IMAGE
ERRTHETA_IMAGE
A_IMAGE
ERRA_IMAGE
B_IMAGE
ERRB_IMAGE
X_WORLD
Y_WORLD
THETA_WORLD
ERRTHETA_WORLD
A_WORLD
ERRA_WORLD
B_WORLD
ERRB_WORLD
ELONGATION
ELLIPTICITY
FLUX_AUTO
FLUXERR_AUTO
MAG_AUTO
MAGERR_AUTO
FLUX_ISO
FLUXERR_ISO
MAG_ISO
MAGERR_ISO
FLUX_PETRO
FLUXERR_PETRO
MAG_PETRO
MAGERR_PETRO
MAG_APER(32)
MAGERR_APER(32)
FLUX_APER(32)
FLUXERR_APER(32)
FLAGS
FWHM_IMAGE
FWHM_WORLD
ISOAREA_IMAGE
ISOAREA_WORLD
KRON_RADIUS
PETRO_RADIUS
```

Figura 2.3: Ejemplo de un archivo de entrada .param

## 2.3. TOPCAT

*Tool for Operations on Catalogues And Tables* (TOPCAT, Taylor (2017)) es un visor y editor gráfico interactivo para datos tabulados. Su objetivo es proporcionar la mayoría de las facilidades que los astrónomos necesitan para el análisis y la manipulación de catálogos de fuentes y otras tablas, aunque también puede utilizarse para datos no astronómicos. Entiende varios formatos astronómicos importantes (incluyendo FITS, CSV, VOTable y CDF) y se pueden añadir más formatos. Es especialmente bueno en la exploración interactiva de grandes tablas (varios millones de filas y un gran número de columnas).

Ofrece una gran variedad de formas de visualizar y analizar las tablas, incluyendo un navegador para los datos de las celdas en sí, visores para la información sobre los metadatos de las tablas y las columnas, y facilidades para la visualización interactiva sofisticada de 1, 2, 3 y más dimensiones, el cálculo de estadísticas y la unión de tablas mediante algoritmos de correspondencia flexibles. Mediante un potente y extensible lenguaje de expresión basado en Java, se pueden definir nuevas columnas y seleccionar subconjuntos de filas para su análisis por separado. Los datos y metadatos de la tabla pueden editarse y la tabla modificada resultante puede escribirse en una amplia gama de formatos de salida. Además, permite generar histogramas, gráficos 2D, 3D, etc. de forma simple y rápida.

TOPCAT es una aplicación autónoma que funciona perfectamente sin conexión a la red. Sin embargo, como utiliza los estándares del Observatorio Virtual (OV), puede cooperar sin problemas con otras herramientas, servicios y conjuntos de datos del universo del OV y fuera de él.

El programa está escrito en Java puro y está disponible bajo la licencia pública general GNU, aunque parte del código de la biblioteca es LGPL. Se ha desarrollado principalmente en el Reino Unido en el marco de varios proyectos británicos y de Euro-VO (Starlink, AstroGrid, VOTech, AIDA, GAVO, GENIUS, DPAC) y con subvenciones del PPARC y del STFC. Los paquetes relacionados STIL y STILTS proporcionan las facilidades para el procesamiento de las tablas subyacentes.

En el marco de este trabajo de Tesis, hemos utilizado TOPCAT cada vez que necesitamos desplegar los catálogos que se fueron generando, los cuales tienen volúmenes del orden de los 5 Gbytes. Además, en las últimas instancias del proceso de estandarización de los catálogos finales, se utilizó TOPCAT para realizar operaciones entre columnas. Las operaciones realizadas se encuentran explicadas en la Sección [3.6](#).

# Capítulo 3

## Metodología

### 3.1. Esquema General

En este capítulo se desarrollan los aspectos metodológicos del trabajo presentado en esta Tesis. Se explica en detalle cada uno de los pasos que realiza el código. Además se explicitan las decisiones tomadas y sus justificaciones, así como algunos de los obstáculos y problemas encontrados, y la forma de resolverlos.

Una de las complejidades que presenta el trabajo de esta Tesis es que por cada campo observado, existen 12 imágenes correspondientes, cada una de ellas, a uno de los filtros de S-PLUS. De esta manera, por cada campo, se deben editar 12 archivos de configuración de entrada para poder ejecutar **SExtractor** sobre las imágenes individuales. Esto implica que, previo al trabajo específico con **SExtractor**, se requerirá de una primera etapa de programación en **Python** para automatizar la edición de la totalidad de los archivos de configuración ( $106 \times 12 = 1272$  archivos de configuración). Asimismo, la ejecución de **SExtractor** en modo dual involucra, en nuestro caso, un conjunto adicional de imágenes denominadas "de detección". Estas imágenes suelen generarse a partir de la suma de varias de las imágenes individuales o considerando un promedio de las mismas. De este modo, existe una segunda etapa previa a la ejecución propiamente dicha de **SExtractor** que consistirá en la construcción de dichas imágenes de detección para los 106 campos sobre los que se planea trabajar.

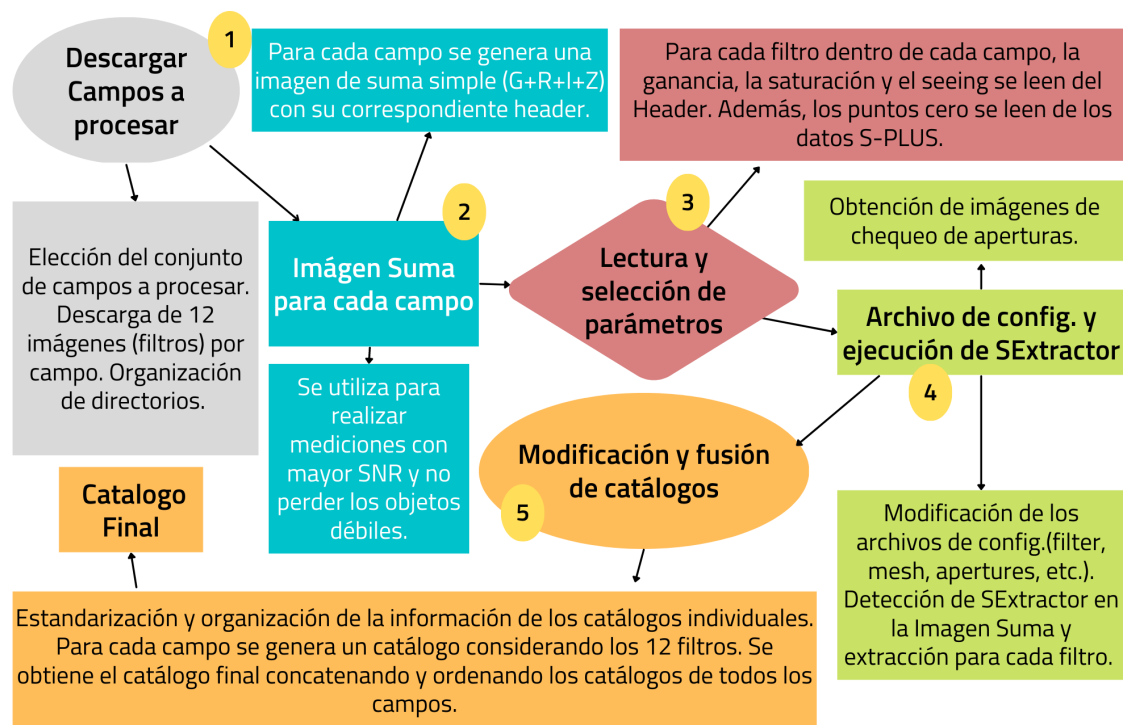
Una vez obtenidas las imágenes de detección y editados todos los archivos de configuración, se procederá a la ejecución de **SExtractor**. Como resultado de dicha ejecución se obtendrá, por un lado, un nuevo conjunto de imágenes de control (imágenes de apertura en nuestro caso particular) que mostrarán la calidad de detección de los distintos objetos. Por otro lado, se generarán catálogos conteniendo la información fotométrica correspondiente a la medición de dichos objetos en cada una de las imágenes consideradas en este trabajo de Tesis. Por cada uno de los 106 campos considerados, se obtendrá una imagen de apertura y 12 catálogos (uno por cada filtro

de S-PLUS), debiendo estos últimos ser concatenados para obtener como resultado final un único catálogo por campo.

De este modo, posteriormente a la ejecución de **SExtractor**, también se requerirá de un trabajo de programación adicional con el fin de concatenar los catálogos individuales y así obtener un catálogo maestro conteniendo la fotometría completa en los 12 filtros de S-PLUS de todos los objetos ubicados en un dado campo. Cabe mencionar que la realización de una ejecución de **SExtractor** puede conducir a resultados no satisfactorios, lo cual implicará modificar los parámetros de los archivos de configuración con el fin de obtener un mejor resultado. De este modo, el procedimiento de edición de los archivos de configuración y de ejecución de **SExtractor** puede llegar a repetirse varias veces. En ese sentido, como producto final en cuanto a programación, hemos generado un código en **Python** que integra todas las etapas anteriormente mencionadas y cuyos datos de entrada correspondan al archivo de configuración general a utilizar y al conjunto de imágenes sobre las cuales trabajar.

En la Figura 3.1 se muestra un esquema resumido de los cinco pasos del proceso de automatización. Cada uno de los pasos será detallado y comentado en las siguientes subsecciones. Si se desea hacer un seguimiento exhaustivo, el código completo se encuentra accesible en el siguiente link:

<https://colab.research.google.com/drive/1AGaGAM7Ujxqh0keazDP1s6vKShDq59X3?>



**Figura 3.1:** Esquema básico de los cinco pasos del proceso de automatización.

## 3.2. Descarga de campos a procesar

En este primer paso, el código realiza la descarga de las imágenes de los campos sobre los que se trabajará. Para ello, como primera instancia, el código lee una lista que contiene las designaciones de dichos campos. S-PLUS designa los diferentes campos observados con la nomenclatura general *SPLUS-sXXsXX*<sup>1</sup>. El código organiza esta descarga de la siguiente manera:

- Se genera una carpeta “*Campos*”.
- Para cada campo incluido en la lista de entrada, y dentro del directorio “*Campos*”, el código genera, a su vez, una subcarpeta con el nombre de cada campo.
- Dentro de la subcarpeta “*Campos/SPLUS-sXXsXX*”, crea a su vez una nueva carpeta denominada “*Imágenes Originales*”.
- El código se conecta a la base de datos de S-PLUS (<https://splus.cloud/>), y dentro de la carpeta “*Campos/SPLUS-sXXsXX/Imágenes Originales*” procede a descargar y descomprimir, una a una, las 12 imágenes correspondientes al campo *SPLUS-sXXsXX*.

La conexión a <https://splus.cloud/> se realiza mediante la instrucción:

```
1 splusdata.connect('user', 'password')
```

donde `splusdata` es un script escrito en Python especialmente desarrollado por la colaboración S-PLUS. Una vez que el usuario se loguea a <https://splus.cloud/>, se utiliza la función `get_field(campo, filtro)` para descargar cada una de las 12 imágenes de un campo en particular. Cabe mencionar que, dentro del código, se fuerza a que las imágenes sean descargadas con el *header* largo o completo, ya que a partir de dichos *headers* el código extraerá, posteriormente, valores de parámetros para ser utilizados con los distintos archivos de configuración de `SExtractor`.

Para lograr seguir la estructura de directorios y archivos, en la Figura 3.2 se muestra un esquema completo de los subdirectorios generados hasta el momento y de los que se generarán a la largo de todo el proceso de automatización.

## 3.3. Construcción de las Imágenes Suma + Headers

En este paso, para cada campo, el código crea una carpeta “*Imágenes*” al mismo nivel que la previamente generada “*Imágenes Originales*”, donde se copian las

<sup>1</sup>La primera letra ‘s’ designa el hemisferio ecuatorial y la segunda letra ‘s’ designa el hemisferio galáctico. El primer número de dos cifras designa simbólicamente a la posición en declinación en la grilla de apuntamientos. De esta manera, cuanto más cercano a 0 es este número, más cercano al ecuador se encuentra el campo. De manera similar, el segundo número simboliza la posición en ascensión recta dentro de la grilla de apuntamientos.

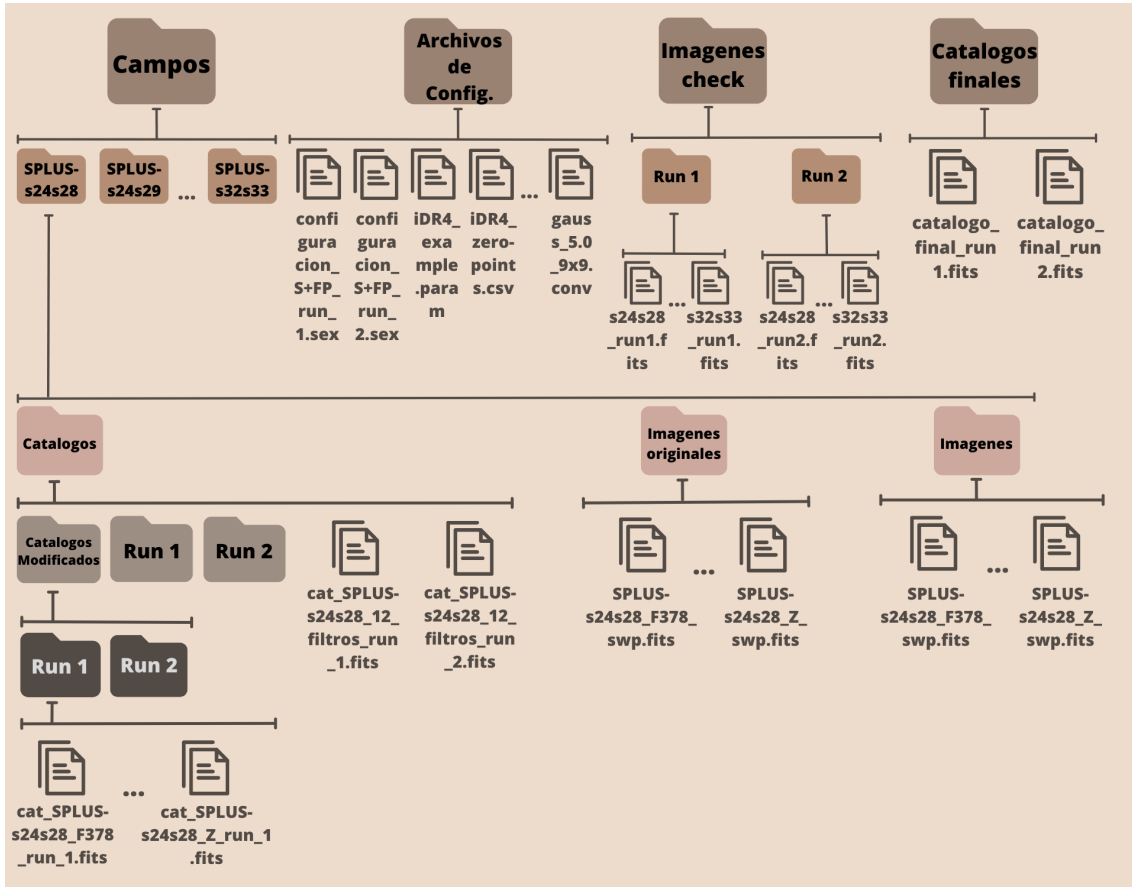


Figura 3.2: Esquema completo de la estructura de subdirectorios y archivos generados.

12 imágenes anteriormente descargadas y descomprimidas, que será el directorio de trabajo con SExtractor. De este modo, el directorio “*Imágenes Originales*” se transforma en el directorio de respaldo de la descarga y posterior descompresión de las imágenes. Para realizar esto, el código utiliza los paquetes “os” y “shutil”.

Luego, para cada campo, el código lee, considerándolas arreglos de píxeles, las imágenes correspondientes a los filtros de banda ancha *gSDSS*, *rSDSS*, *iSDSS* y *zSDSS* (ver Tabla 1.1). Por simplicidad, en el código los denominamos “*G*”, “*R*”, “*I*” y “*Z*”, respectivamente. Como tenemos cuatro arreglos, podemos simplemente sumarlos y así generar una imagen suma, no ponderada, donde los cuatro filtros tienen el mismo peso.

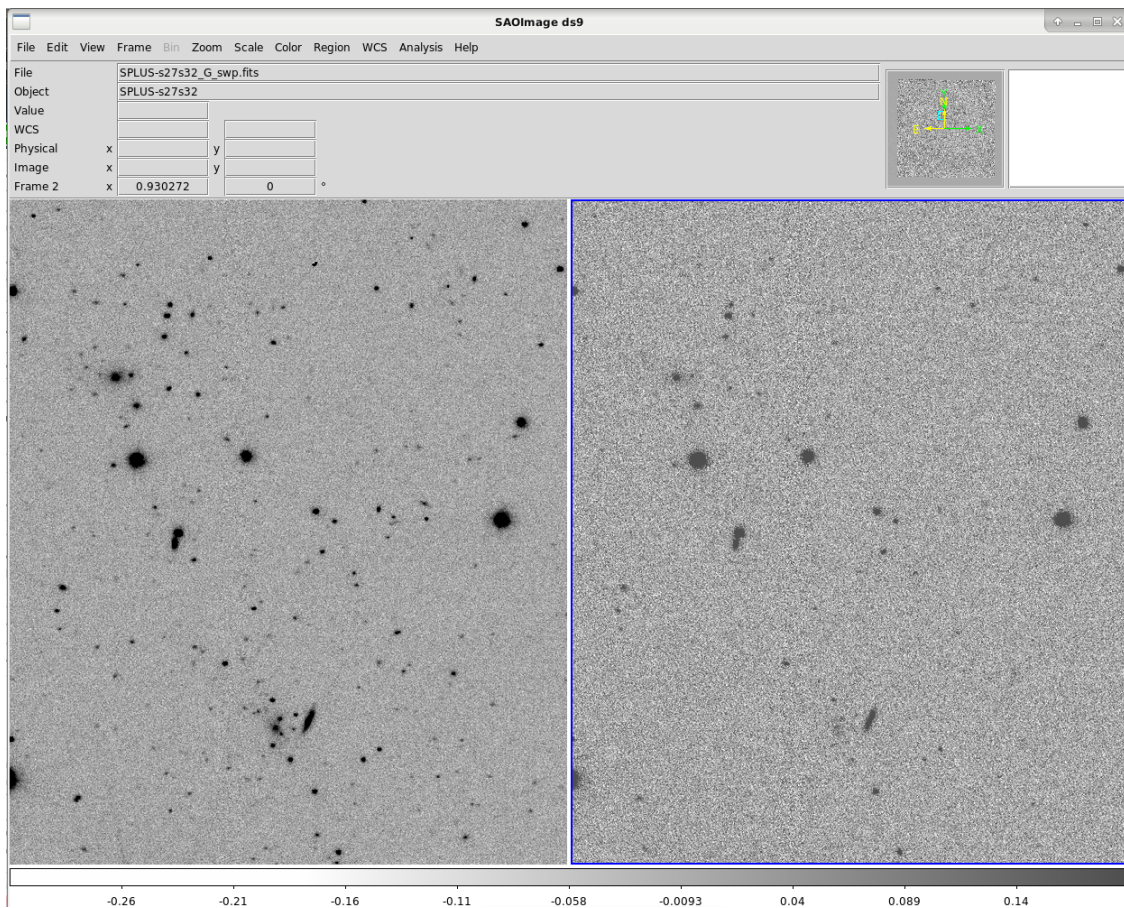
Además de leer la imagen correspondiente al filtro “*I*”, el código obtiene el *header* de esta imagen, lo duplica, y modifica los siguientes parámetros:

```
1 FILTER = "G+R+I+Z"
2 FILENAME = "Campos/{campo}/Imágenes/{campo}\_suma.fits"
```

Este *header* modificado es asignado como *header* de la imagen suma, la cual es luego

guardada en la carpeta “*Imágenes*”, junto a las imágenes de los 12 filtros individuales. Se realiza el mismo proceso para todos los campos.

La imagen suma generada de esta manera será utilizada para ejecutar **SExtractor** cumpliendo la función de imagen de detección; esto es, será la imagen donde **SExtractor** identificará las fuentes a extraer y medir. La principal motivación de construir una imagen suma es aumentar la relación S/N y, de esa forma, poder detectar una mayor cantidad de objetos débiles y medirlos con mayor seguridad. Cabe mencionar que la detección de los objetos en la iDR4 de S-PLUS se realizó también sobre una imagen suma de los mismos cuatro filtros considerados aquí, aunque ponderada.



**Figura 3.3:** Comparación entre la imagen suma (a la izquierda) y la imagen del filtro "G" (a la derecha) del campo *SPLUS-s27s32*.

En la Figura [3.3](#) se muestra cómo la imagen suma (panel izquierdo) permite detectar mayor cantidad de objetos, mientras que en la imagen correspondiente al filtro "G" (panel derecho) muchos de los objetos visibles en la imagen suma no se observan a simple vista. En ambos paneles se muestra la misma zona del cielo.

## 3.4. Lectura y selección de parámetros

En este paso, el código comienza recorriendo los directorios de imágenes de los distintos campos uno por uno y, dentro de cada uno de ellos, construye una lista con los nombres de las 12 imágenes de cada filtro e identifica y aparta la imagen suma. A su vez, crea una carpeta “*Catalogos/Gen*” al mismo nivel que “*Imágenes*” e “*Imágenes Originales*”. Aquí la designación “*Gen*” hace referencia a una ejecución de SExtractor “general” ya que, a futuro, se planea realizar las ejecuciones de SExtractor presentadas en esta Tesis pero considerando 32 aberturas circulares concéntricas con el fin de obtener perfiles de brillo en forma automática (ver Sección 5). En el presente trabajo solo se ha obtenido la fotometría correspondiente a 5 aberturas de distinto tipo (ver Sección 3.5).

A continuación, el código lee por completo una tabla previamente descargada de <https://splus.cloud/>, que contiene los puntos de cero (ZP, los cuales sirven para calibrar las magnitudes al sistema estándar fotométrico), para todos los campos incluidos en la iDR4. Esta tabla incluye, en cada fila, el nombre del campo y los valores de los ZPs correspondientes a los 12 filtros de S-PLUS para ese campo (13 columnas en total).

Una vez leída la tabla, el código la recorre hasta encontrar el nombre del campo sobre el que va a trabajar, y toma los 12 ZPs del campo en cuestión. Por otro lado, se extraen del *header* de cada una de las 12 imágenes, los siguientes parámetros:

```
1 SATURATE = hdr['SATURATE']
2 GAIN = hdr['GAIN']
3 SEEING\_FWHM = hdr['HIERARCH OAJ PRO FWHMSEXT']
```

Esto se realiza con la intención de que SExtractor realice las mediciones de los objetos en las imágenes de los 12 filtros de S-PLUS, utilizando en el archivo de configuración correspondiente los valores de dichos parámetros, que son específicos de cada imagen. Esto resulta necesario ya que las imágenes proporcionadas por S-PLUS son construidas a partir de diferentes números de exposiciones que son tomadas con distintos tiempos de exposición, dependiendo del filtro. Por otro lado, las condiciones de observación pueden variar a lo largo de una noche de observación.

## 3.5. Archivos de configuración y ejecuciones de SExtractor

En este paso se definen los archivos de configuración con los que se realizarán las ejecuciones de SExtractor. En nuestro caso particular nos hemos concentrado en dos configuraciones cuyos parámetros generales provienen de experiencias previas con el uso de SExtractor por parte de la directora de esta Tesis (ver, por ejemplo,



Smith Castelli et al., 2008), y de mejoras para la detección y medición de cierto tipo de objetos que fueron testeadas por el Dr. Favio R. Faifer (Co-Director de este trabajo). Por lo tanto, lo que se presenta de aquí en adelante estará restringido a dos archivos de configuración base, los cuales designaremos como **Run 1** y **Run 2**.

Los archivos de configuración incluyen parámetros que guiarán la detección y extracción de los objetos presentes en una imagen, y determinarán la obtención de la fotometría para cada uno de ellos. Asimismo, incluyen la opción de obtener distintos tipos de imágenes de control del proceso realizado.

Para la detección y extracción de los objetos, se consideraron los parámetros que se muestran en la Tabla 3.1. Los parámetros `DETECT_MINAREA` y `DETECT_THRESH` representan parámetros de umbral para la identificación de objetos por parte de `SExtractor`. Esto es, para que `SExtractor` considere que un dado conjunto de píxeles conforma un objeto, se deben cumplir los siguientes requisitos:

- El conjunto debe estar formado por un número mínimo de píxeles adyacentes dado por `DETECT_MINAREA`
- El brillo de todos los píxeles del conjunto debe estar por encima del valor de fondo o *background*, dado por `DETECT_THRESH`.

El parámetro `ANALYSIS_THRESH`, que en nuestro caso se encuentra dado en unidades de  $\sigma$  pero podría estar expresado en ADUs, determina el umbral de brillo que se considerará para obtener la fotometría del objeto e influye en la determinación de los parámetros de salida `CLASS_STAR` y el FWHM. Cabe mencionar que `DETECT_THRESH = 1.5` establece el umbral de detección a  $1.5 \sigma$  por sobre el background local, donde  $\sigma$  es la desviación estándar de ese background. Por otro lado, es importante aclarar que la desviación estándar nombrada corresponde a la imagen sin filtrar. Por su parte, y como se mencionó anteriormente, el parámetro `DETECT_MINAREA` establece la mínima cantidad de píxeles conectados que `SExtractor` considerará para realizar la detección. Estratégicamente, este parámetro puede usarse a su vez como `DETECT_THRESH` para detectar solamente objetos

**Tabla 3.1:** Parámetros y sus correspondientes valores para la detección y extracción, diferenciados entre ejecuciones.

Run 1		Run 2	
<code>DETECT_MINAREA</code>	5	<code>DETECT_MINAREA</code>	10
<code>DETECT_THRESH</code>	1.5	<code>DETECT_THRESH</code>	1.5
<code>ANALYSIS_THRESH</code>	1.5	<code>ANALYSIS_THRESH</code>	1.5
<code>FILTER</code>	Y	<code>FILTER</code>	Y
<code>FILTER_NAME</code>	default.conv	<code>FILTER_NAME</code>	gauss_5.0_9x9.conv
<code>DEBLEND_NTHRESH</code>	32	<code>DEBLEND_NTHRESH</code>	64
<code>DEBLEND_MINCONT</code>	0.005	<code>DEBLEND_MINCONT</code>	0.001

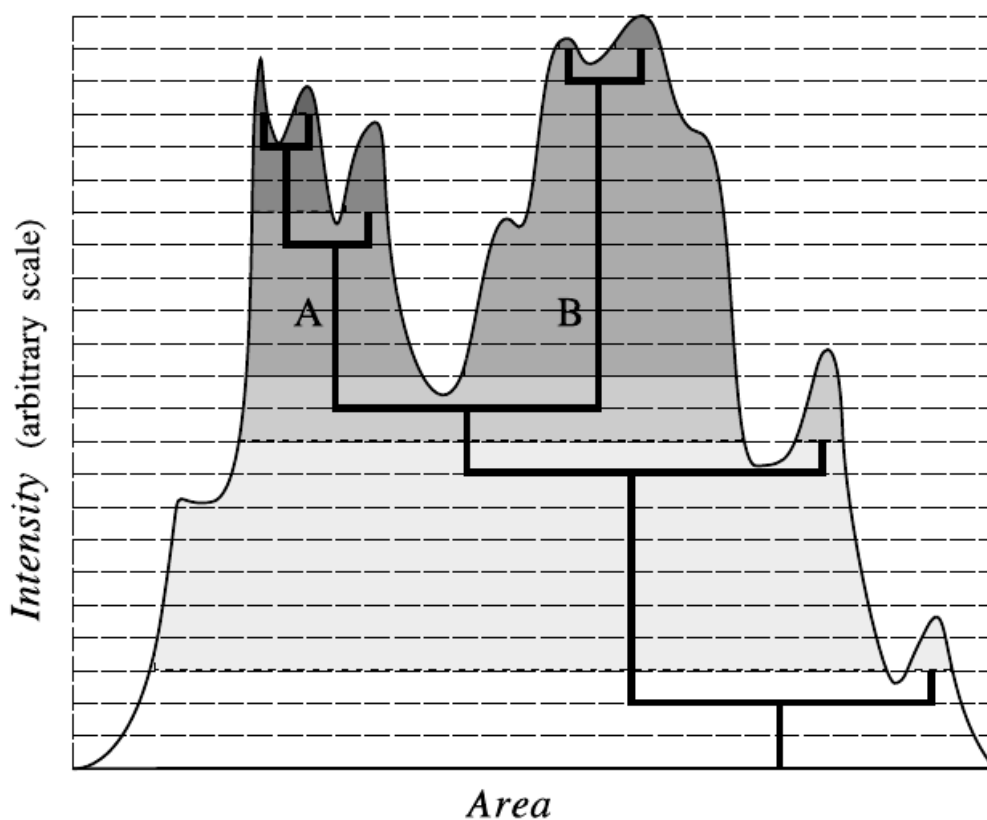
brillantes y extendidos, o para incrementar la fiabilidad de la detección. En nuestro caso, la única diferencia en estos parámetros entre **Run 1** y **Run 2** es el valor de `DETECT_MINAREA`. De este modo, **Run 2** no detectará objetos que cubran menos de 10 píxeles los cuales sí serán detectados por **Run 1** siempre y cuando superen el límite de 5 píxeles.

Para la detección de objetos sobre un dado umbral, existe la opción de aplicar un filtro, básicamente convolucionar la imagen con cierta superficie. Esto puede ayudar en la detección de objetos débiles y extendidos. Aquí es donde entran en juego los parámetros `FILTER` y `FILTER_NAME`. En nuestro caso, tanto **Run 1** como **Run 2** utilizan filtrado. En el caso de **Run 1**, se utiliza un filtro por defecto “*default.conv*”, a través de una función piramidal pequeña. Por su parte **Run 2** utiliza un filtro gaussiano “*gauss\_5.0\_9x9.conv*”, imponiendo una máscara de convolución  $9 \times 9$  de una PSF gaussiana con  $\text{FWHM}=5,0$  píxeles. Esto favorecerá que **Run 2** detecte objetos brillantes y extendidos.

Los parámetros `DEBLEND_NTHRESH` y `DEBLEND_MINCONT` están vinculados al proceso de “deblending” o “separación” de objetos, es decir, a través de ellos **SExtractor** toma la decisión de si un grupo de píxeles adyacentes por encima de `DETECT_THRESH` corresponden a un único objeto o constituyen la superposición de varios. Por ejemplo, si consideramos una pequeña isla de píxeles por encima del umbral, para poder identificar si constituye un objeto independiente o es parte de otro objeto más extendido, **SExtractor** define un número de niveles entre el umbral y el número máximo de cuentas en el objeto. El encargado de fijar esto es el parámetro `DEBLEND_NTHRESH`. Los niveles están separados de manera exponencial. Entonces **SExtractor** construye un “árbol” de las fuentes, ramificándose cada vez que haya píxeles por encima del umbral separados por píxeles debajo del mismo (ver Figura 3.4). Una rama es considerada un objeto diferente siempre que:

1. El número de cuentas en la rama (A en la figura) este por encima de una cierta fracción del número total de cuentas en toda la “isla”.
2. Hay al menos otra rama (B en la figura) por encima del mismo nivel que también supera esta fracción.

La fracción usada es definida por `DEBLEND_MINCONT`. El algoritmo va de arriba hacia abajo, desde las puntas de las ramas, y decide si en cada intersección se extraen dos o más objetos, o sigue su camino hacia abajo. Para objetos débiles, la eficiencia del “deblending” se ve limitada por el seeing y el muestreo. Además, el número de umbrales de “deblending” queda configurado por `DEBLEND_NTHRESH`, y el contraste por `DEBLEND_MINCONT`. Elegir `DEBLEND_MINCONT` como 0, representa que incluso el pico local más débil en el perfil será considerado como objeto separado. Si se establece como 1 significa que el “deblending” no se autoriza. En ese sentido, los parámetros tomados en la Tabla 3.1 permiten a **Run 1** detectar objetos cercanos y débiles, mientras que **Run 2** busca que los objetos brillantes que presenten estructura resuelta no sean fragmentados como objetos separados.



**Figura 3.4:** Diagrama esquemático del método utilizado por SExtractor para separar un objeto compuesto. El perfil de área del objeto (curva suave) puede ser descrito de una manera estructurada en árbol (líneas gruesas). La decisión de considerar o no una rama como un objeto distinto se determina de acuerdo a su intensidad relativa integrada (zona de color). En este caso, el objeto original se divide en dos componentes: A y B. Figura tomada de Bertin y Arnouts (1996).

Por otro lado, para la estimación del background se tomaron los parámetros de la Tabla 3.2. El valor que se mide en cada píxel es la suma de la señal del background y de la señal proveniente de los objetos. Para lograr detectar objetos débiles y poder medir con precisión sus flujos, es necesario contar con una estimación precisa del background en cualquier lugar de la imagen. Se deberá construir un mapa de background, y para ello SExtractor realiza un primer pasaje a través de los píxeles, calculando un estimador para el background local en cada subregión de una grilla que cubre toda la imagen.

Una vez que la grilla queda configurada, un filtrado puede usarse para quitar posibles sobre estimaciones locales debido a estrellas brillantes. En simultáneo a la construcción del mapa de background, se produce un “mapa background RMS”, el cual es

**Tabla 3.2:** Parámetros y sus correspondientes valores para la estimación del background, diferenciados entre ejecuciones.

Run 1		Run 2	
BACK_SIZE	64	BACK_SIZE	524
BACK_FILTERSIZE	3	BACK_FILTERSIZE	3
BACKPHOTO_TYPE	LOCAL	BACKPHOTO_TYPE	GLOBAL
		BACK_TYPE	MANUAL
		BACK_VALUE	0.0

un mapa del ruido de background. Este se utilizará si el parámetro WEIGHT\_TYPE está configurado diferente de NONE.

El tamaño de cada subregión de la grilla donde se calcula el background está dado por el parámetro BACK\_SIZE. Si el valor de este parámetro es muy pequeño, la estimación del background es afectada por la presencia de objetos y el ruido aleatorio. En contraposición, si el tamaño de la subregión de la grilla es muy grande, SExtractor no logra reproducir las variaciones a pequeña escala espacial del background. Típicamente, para imágenes razonablemente muestreadas, se usa un ancho de 32 a 256 píxeles. En nuestro caso **Run 2** adopta el valor 524 por el tamaño en píxeles de algunos objetos brillantes.

Otro parámetro interesante es el BACK\_FILTERSIZE, el cual permite dar el tamaño del filtro y dará un suavizado de la imagen de background. Este parámetro también sirve para eliminar las variaciones resultantes de objetos brillantes o extendidos. El parámetro BACK\_TYPE puede ser AUTO o MANUAL. Si se encuentra configurado en la segunda de esta opciones, es posible imponer el valor del background a través del parámetro BACK\_VALUE.

Las diferencias sustanciales de la estimación del background en nuestro caso tienen que ver con que **Run 1** efectúa una estimación local con un tamaño de la subregión de la grilla acorde a los valores frecuentemente utilizados. Por su parte **Run 2** aplica una estimación global, usando un tamaño de subregión de la grilla muy grande (más del doble del máximo de los valores frecuentemente utilizados), además de establecer manualmente el valor mínimo en 0. Esto último evitará obtener valores negativos al medir sobre la imagen. Estas diferencias están vinculadas con la idea de que **Run 1** espera detectar objetos débiles, mientras que **Run 2** apunta a la detección de objetos brillantes. La estimación del background está estrechamente ligada a este tipo de intencionalidad.

Con respecto a los parámetros que rigen la fotometría y la separación estrella/-galaxia en los archivos de configuración, no hay diferencias entre **Run 1** y **Run 2** ya que ambas utilizan los parámetros mostrados en la Tabla [3.3](#). Las aberturas

que se han considerado en ambas ejecuciones para estimación de magnitudes totales son AUTO, PETRO e ISO. En particular, las dos cantidades asignadas al parámetro PHOT\_AUTOPARAMS corresponden a la constante de proporcionalidad  $k$  y al  $R_{min}$  presentados anteriormente en la Sección 2.2.2. Además de las tres magnitudes mencionadas anteriormente, se han considerado dos provenientes de la medición dentro de aberturas circulares de diámetros dados por el parámetro PHOTO\_APERTURES, de 5.45 píxeles y 10.9 píxeles correspondientes, en unidades angulares, a 3" y 6" respectivamente, valores usualmente utilizados en fotometría de apertura automática. La escala de las imágenes se da a través del parámetro PIXEL\_SCALE y corresponde a  $0.55'' \text{ pix}^{-1}$  para las imágenes de S-PLUS. Se utilizó un archivo por defecto para los pesos de la NN en la separación estrella/galaxia. Se calculan, además, los radios conteniendo el 20 %, 50 %, 70 % y 90 % de la luz de los objetos mediante los 4 valores asignados al parámetro PHOT\_FLUXFRAC.

**Tabla 3.3:** Parámetros del archivo de configuración compartidos por ambas ejecuciones

Run 1 y Run 2	
PHOT_APERTURES	5.45, 10.90
PHOT_AUTOPARAMS	3.0, 1.82
PHOT_PETROPARAMS	2.0, 2.73
PHOT_FLUXFRAC	0.2,0.5,0.7,0.9
PIXEL_SCALE	0.55
STARNNW_NAME	default.nnw
PARAMETERS_NAME	iDR4_example.param

Con respecto al archivo que define la información sobre los parámetros que se desea que aparezcan en el catálogo de salida, se descargó de <https://splus.cloud/> el archivo “*iDR4\_example.param*” que fuera utilizado por la colaboración S-PLUS como parte de la iDR4, el cual es el archivo ejemplo mostrado en la Figura 2.3. Utilizamos dicho archivo para definir nuestros parámetros de salida ya que los catálogos que deseamos obtener serán luego estandarizados al formato de catálogos provistos por S-PLUS.

Luego de dar estos detalles sobre los archivos de configuración, volvemos a retomar lo que se realiza en el presente paso del código. Una vez elegido el archivo de configuración, se genera una dirección dual (directorio que incluye tanto a la imagen de detección como a la de medición). Como ya se mencionó anteriormente, la imagen de detección será la imagen suma y será la misma para todos los filtros de un dado campo, mientras que la imagen de medición irá variando filtro a filtro. Para ejecutar SExtractor desde Python se utilizó el paquete “**subprocess**”, el cual interactúa con una terminal bash para poder ejecutar SExtractor. La ejecución en modo dual en el código Python tiene la siguiente forma:

```
1 subprocess.run(["sextractor", dir_dual, "-c", f"{dir_entrada_configu
```

```

2 racion}", f"-CATALOG_NAME {dir_salida_gen}", f"-GAIN {GAIN}",
3 f"-SEEING_FWHM {SEEING_FWHM}", f"-SATUR_LEVEL {SATURATE}", f"-MAG_ZE
4 ROPOINT {zeropoint}"])
```

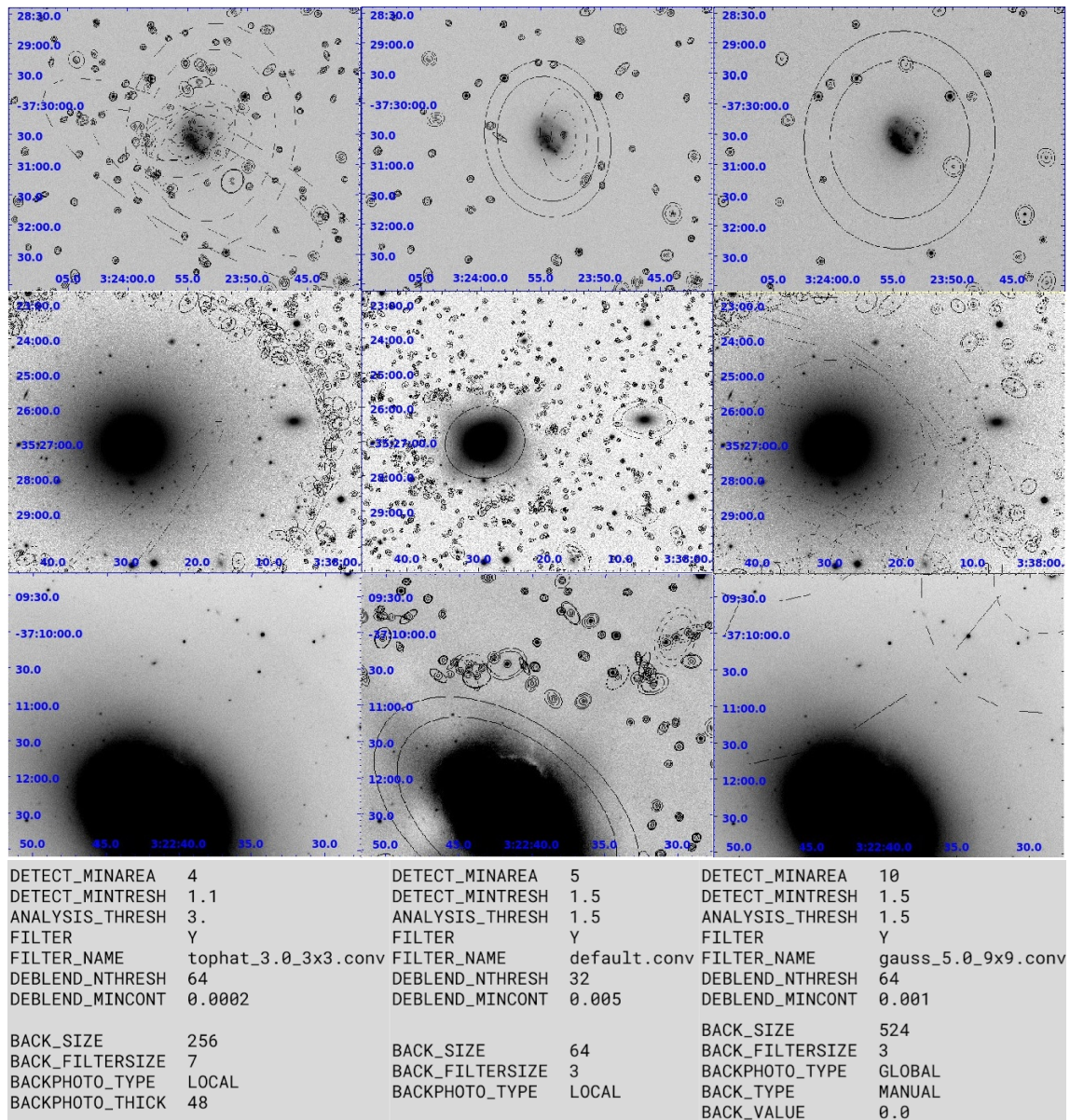
La salida de esta ejecución genera un catálogo por cada filtro que es guardado en “/Campos/campo/Catalogos/Gen/cat\_gen\_filtro”. Es decir, en esta instancia, por cada campo, habrá 12 catálogos. Además, en este paso del código, se obtienen también las imágenes de chequeo, CHECK\_IMAGES. Para ello, simplemente, se realiza una única ejecución de SExtractor en modo simple sobre la imagen de detección correspondiente al campo sobre el que se está trabajando, esto es, sobre la imagen suma obtenida anteriormente. Para obtener la imagen de chequeo, se ejecutan las siguientes líneas:

```

1 subprocess.run(["sextractor", dir_entrada_suma, "-c", "../Archivos_d
2 e_Configuracion/configuracion_S+FP_gen.sex", "-CATALOG_NAME", "../c
3 at_2.fits", "-CHECKIMAGE_TYPE", "APERTURES", "-CHECKIMAGE_NAME",
4 f"../{dir_check}"])
5
6 os.makedirs(f"../Imagenes_check/Run_1/", exist_ok=True)
7
8 shutil.move(f"../{dir_check}", f"../Imagenes_check/Run_1/")
```

donde, luego de la ejecución de SExtractor, se crea un directorio “Imagenes\_check” al mismo nivel que la carpeta inicial “Campos”, que a su vez contendrá dos subdirectorios, “Imagenes\_check/Run\_1” o “Imagenes\_check/Run\_2”, dentro de los cuales se guardarán las imágenes de chequeo según corresponda. Por lo tanto, las imágenes de chequeo de todos los campos quedan guardadas diferenciadas por **Run 1** o **Run 2**.

En el panel superior de la Figura 3.5 se muestran 9 imágenes de chequeo, donde las 3 imágenes de la columna de la izquierda corresponden a las imágenes de chequeo de la iDR4 de S-PLUS, la columna central a 3 imágenes de chequeo de **Run 1**, mientras que en la columna de la derecha las 3 imágenes de chequeo corresponden a **Run 2**. Cabe aclarar que en cada fila se muestra la misma zona del cielo. Lo que se observa aquí es lo anticipado en la Sección 1.4 como motivación de este trabajo: con los parámetros utilizados para obtener la fotometría incluida en los catálogos de la iDR4, se tiende a subdividir excesivamente las galaxias que presentan regiones de formación estelar y a pasar por alto las galaxias débiles y los objetos compactos cercanos a galaxias brillantes. En el caso de **Run 1**, se detectan correctamente esas galaxias débiles y objetos compactos pero se subestiman los tamaños de las aberturas de las galaxias brillantes, subestimando de esta manera sus magnitudes totales. Por su parte, **Run 2** logra detectar las galaxias brillantes con aberturas más apropiadas a sus extensiones reales en las imágenes de S-PLUS, pero no consigue detectar las galaxias débiles y objetos compactos cercanos a galaxias brillantes, al igual que sucedía con la iDR4. En el panel inferior de la Figura 3.5 se presentan



**Figura 3.5:** *Paneles Superiores:* Imágenes de apertura obtenidas a partir de las ejecuciones de SExtractor considerando los parámetros utilizados para obtener los catálogos fotométricos de S-PLUS iDR4 (izquierda), **Run 1** (centro) y **Run 2** (derecha). *Panel Inferior:* Cuadro comparativo de los parámetros de SExtractor para obtener la fotometría de la iDR4 (izquierda), de **Run 1** (centro) y **Run 2** (derecha).

las diferencias en los parámetros de los archivos de configuración de la iDR4 con respecto a los utilizados en **Run 1** y **Run 2**.

## 3.6. Modificación y fusión de catálogos

En esta instancia se realiza el proceso de estandarización, unión y fusión de los catálogos. El código inicia esta etapa recorriendo nuevamente el árbol de directorios campo por campo y leyendo los catálogos anteriormente obtenidos, filtro por filtro. Luego crea un directorio dentro de la carpeta “/Campos/campo/Catalogos/” llamado “Catalogos\_modificados” donde se copian los 12 catálogos correspondientes a los 12 filtros de S-PLUS que fueron generados en la Sección 3.5, y que se encuentran en el directorio “/Campos/campo/Catalogos/Gen/”. A estas copias de los catálogos se le realizan varias modificaciones. Por un lado, se despliegan todas las columnas que se hayan generado en formato de vectores, como es el caso de MAG\_APER y sus errores (vectores de dimensión 2), FLUX\_APER y sus errores (vector de dimensión 2) y FLUX\_RADIUS (vector de dimensión 4). Para ello se utilizó la siguiente línea:

```
1 subprocess.run(["stilts", "tpipe", f'in={dir_1}', "cmd = explodeall",
2 f'out={dir_1}'])
```

donde se utilizó nuevamente el paquete “**subprocess**” y dentro de el, la función “**stilts**” para obtener un nuevo catálogo modificado por cada filtro. Luego, a esa primera modificación, se le agrega una modificación en el nombre de ciertas columnas. Esto se realiza buscando, por un lado, estandarizar los nombres para que coincidan con los dados en los catálogos de S-PLUS y, por otro lado, para agregarles el nombre del filtro explícitamente a las columnas de, por ejemplo, las magnitudes, que en los catálogos por filtro simplemente hacen referencia al tipo de magnitud. Para ello se ejecuta lo siguiente:

```
1 dat = Table.read(f"../Campos/{campo}/Catalogos/Catalogos_modificad
2 os/Gen_explode_RUN1/{catalogo}.fits")
3
4 df = dat.to_pandas()
5
6 df.rename(columns={'ALPHA_J2000':'RA', 'DELTA_J2000':'DEC', 'FLUX_AU
7 TO':f'FLUX_AUTO_{filtro}', 'FLUXERR_AUTO':f'FLUXERR_AUTO_{filtro}',
8 'MAG_AUTO':f'{filtro}_AUTO', 'MAGERR_AUTO':f'e_{filtro}_AUTO',
9 'FLUX_ISO':f'FLUX_ISO_{filtro}', 'FLUXERR_ISO':f'FLUXERR_ISO_{filtro
10 }', 'MAG_ISO':f'{filtro}_ISO', 'MAGERR_ISO':f'e_{filtro}_ISO',
11 'FLUX_PETRO':f'FLUX_PETRO_{filtro}', 'FLUXERR_PETRO':f'FLUXERR_PETRO_
12 {filtro}', 'MAG_PETRO':f'{filtro}_PETRO', 'MAGERR_PETRO':f'e_{filtro}
13 _PETRO', 'MAG_APER_1':f'{filtro}_APER_3', 'MAG_APER_2':f'{filtro}_APER
14 _6', 'MAGERR_APER_1':f'e_{filtro}_APER_3', 'MAGERR_APER_2':f'e_{filtro
15 }_APER_6', 'FLUX_APER_1':f'FLUX_APER_3_{filtro}', 'FLUX_APER_2':f'FLUX
16 _APER_6_{filtro}', 'FLUXERR_APER_1':f'FLUXERR_APER_3_{filtro}', 'FLUX
17 ERR_APER_2':f'FLUXERR_APER_6_{filtro}', 'FLAGS':f'FLAGS_{filtro}',
18 'FWHM_IMAGE':f'FWHM_IMAGE_{filtro}', 'FWHM_WORLD':f'FWHM_WORLD_{fil
19 tro}', 'ISOAREA_IMAGE':f'ISOAREA_IMAGE_{filtro}', 'ISOAREA_WORLD':
20 f'ISOAREA_WORLD_{filtro}', 'KRON_RADIUS':f'KRON_RADIUS_{filtro}',
21 'PETRO_RADIUS':f'PETRO_RADIUS_{filtro}', 'FLUX_RADIUS_1':f'FLUX_RADI
22 US_20_{filtro}', 'FLUX_RADIUS_2':f'FLUX_RADIUS_50_{filtro}', 'FLUX_RA
```



```

23 DIUS_3':f'FLUX_RADIUS_70_{filtro}','FLUX_RADIUS_4':f'FLUX_RADIUS_90_
24 {filtro}','FLUX_MAX':f'FLUX_MAX_{filtro}', 'SNR_WIN':f'SNR_WIN_{filt
25 ro}','MU_THRESHOLD':f'MU_THRESHOLD_{filtro}','THRESHOLD':f'THRESHOLD
26 _{filtro}','MU_MAX':f'MU_MAX_{filtro}','CLASS_STAR':f'CLASS_STAR_{fi
27 ltro}','BACKGROUND':f'BACKGROUND_{filtro}'}}, inplace=True)
28
29 t = Table.from_pandas(df)
30
31 t.write(f"..Campos/{campo}/Catalogos/Catalogos_modificados/Gen_expl
32 ode_RUN1/{catalogo}.fits", overwrite=True)

```

donde se utiliza, por primera vez, el paquete “**pandas**” para leer y modificar el catálogo como un DataFrame. Nuevamente, este catálogo modificado sobrescribe al anterior.

Posteriormente, se procede a fusionar todos los catálogos por filtro modificados correspondientes a un dado campo, en un único catálogo maestro por campo. Para ello, utilizando nuevamente “**pandas**”, se crea una lista de los 12 DataFrames correspondientes a los 12 catálogos modificados a fusionar y, previo a la fusión propiamente dicha, se extrae la información de las primeras columnas, que son comunes a los 12 catálogos por filtro y que corresponde a la astrometría y geometría de los objetos detectados, los cuales se encuentran listados y descriptos brevemente en la Figura 3.6, algunos de estos parámetros se encuentran esquematizados en la Figura 3.7

Luego sobre la lista de DataFrames se realiza la fusión, bajo una búsqueda de coincidencia exacta doble sobre RA y DEC, mediante:

```

1 df_merged = reduce(lambda left,right: pd.merge(left,right, on=['RA'
2 , 'DEC'], how='inner', copy=False), data_frames)

```

La coincidencia es exacta debido a que la información astrométrica es compartida con exactitud entre filtros del mismo campo, ya que comparten la imagen de detección (imagen suma).

Se agregan como primeras columnas “ID” y “Field”, se ordena según RA y se le da un formato específico a los nombres de los objetos en la columna ID. Ejecutando:

```

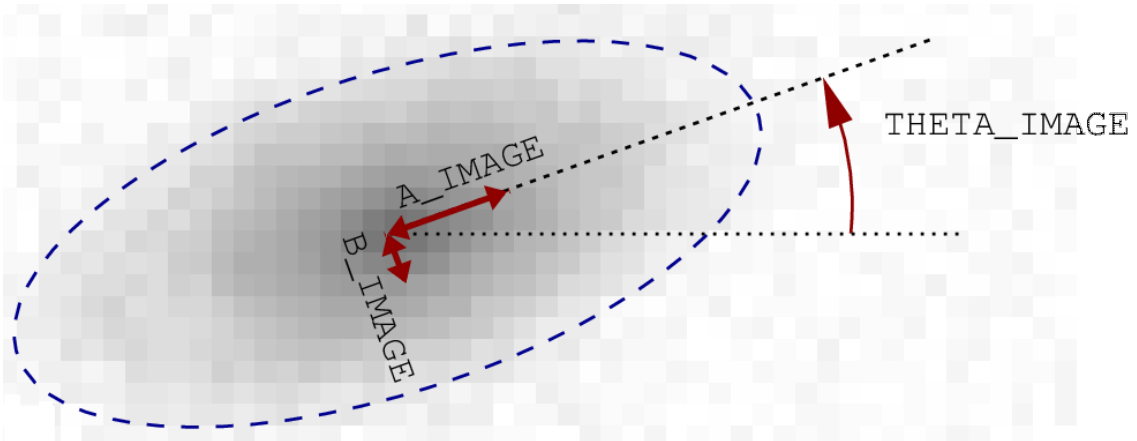
1 df_merged.insert(0, "ID", lista_number, True)
2 df_merged.insert(1, "Field", f"{campo}", allow_duplicates=False)
3 df_merged = df_merged.sort_values('RA', ascending=True)
4 for ID in range(len(df_merged.ID)+1):
5     df_merged.ID = df_merged.ID.replace({ID:f"iDR4_FORNAX_RUN1_{cam
6     po}.{str(ID).rjust(5, '0')}"})

```

El último paso antes de obtener los catálogos finales, será leer los catálogos maestros de cada campo, transformarlos en DataFrame, e ir agregándolos uno por uno a una lista de DataFrames. Una vez, construida esta lista se procede a concatenar todos los catálogos de todos los campos, para lograr tener un único catalogo final

Identification and Position		Flags and classification Columns	
<b>Field</b>	Name of the S-PLUS field of the observation	<b>calibration_flag</b>	Indication of the reference catalogue for calibration
<b>ID</b>	Observation ID in DR2	<b>PhotoFlag_{filter}</b>	SExtractor 'FLAGS' in the Detection image
<b>RA</b>	Right Ascension (J2000) [deg]	<b>PhotoFlagDet</b>	SExtractor 'FLAGS' in the Detection image
<b>DEC</b>	Declination (J2000) [deg]	<b>CLASS_STAR_{filter}</b>	Star classification in the Detection image
<b>X</b>	CCD X-axis position (reduced image) [pixel]	<b>CLASS_STAR</b>	Star classification in the Detection image
<b>Y</b>	CCD Y-axis position (reduced image) [pixel]		
Morphology		Photometry Columns	
<b>ISOarea</b>	Isophotal area above 1.1 sigma threshold	<b>{filter}_{aperture}</b>	AB magnitude
<b>MU_MAX</b>	Peak surface brightness above background [mag/arcsec <sup>2</sup> ]	<b>e_{filter}_{aperture}</b>	Magnitude error
<b>A</b>	Isophotal image major axis [pixel]	<b>s2n_{filter}_{aperture}</b>	Source's S/N
<b>B</b>	Isophotal image minor axis [pixel]	<b>nDet_{aperture}</b>	Number of detections
<b>THETA</b>	Isophotal image position angle [deg]	<b>s2n_Det_{aperture}</b>	Source's S/N in the detection image
<b>ELONGATION</b>	A/B		
<b>ELLIPTICITY</b>	1 - B/A		
<b>FLUX_RADIUS</b>	Radius containing (0.2,0.5,0.7,0.9) fraction of the light [pixel]		
<b>KRON_RADIUS</b>	Kron apertures in units of A or B		
<b>FWHM</b>	FWHM assuming a Gaussian core [pixel]		
<b>FWHM_n</b>	Normalized FWHM		
<b>FWHM_{filter}</b>	FWHM at each filter [pixel]		

**Figura 3.6:** Parámetros de salida de los catálogos. Figura tomada de (Almeida-Fernandes et al. 2022).



**Figura 3.7:** Ejemplificación de A\_IMAGE, B\_IMAGE y THETA\_IMAGE para un objeto extendido. Figura tomada de (Bertin y Arnouts, 1996).

estandarizado por cada una de las ejecuciones de SExtractor. A dicho catalogo se lo ordena por ID para que, conforme al formato dado para los ID, se ordene a los objetos por bloques de campos. Es decir, si se despliega el catalogo final, el mismo comenzara con todos los objetos correspondientes al primer campo según el orden secuencial en su denominación, ordenados de acuerdo al ID asignado originalmente por SExtractor Para ello, se ejecutan las líneas que se muestran a continuación:

```
1 os.makedirs(f'Catalogos_finales', exist_ok=True)
2 data_frames = []
3 lista_catalogos_a_concat = []
4 for campo in lista_campos:
5     print("Procesando:", campo)
6     archivos_campo_aux = os.listdir(f'Campos/{campo}/Catalogos')
7     lista_catalogo_campo = list(filter(lambda filtro: filtro.count(
8     f'_{run}.fits'), archivos_campo_aux))
9     catalogo_campo_aux = lista_catalogo_campo[0]
10    catalogo_campo = f'Campos/{campo}/Catalogos/'+catalogo_campo_aux
11    print("Agregando:", catalogo_campo)
12    lista_catalogos_a_concat.append(catalogo_campo)
13    lector = Table.read(catalogo_campo)
14    df = lector.to_pandas()
15    data_frames.append(df)
16
17 print("Realizando la concatenación total")
18 df_concat = pd.concat(data_frames)
19 print("Paso la concatenación")
20 df_concat = df_concat.sort_values('ID', ascending=True)
21 t = Table.from_pandas(df_concat)
22 t.write(f'Catalogos_finales/catalogo_final_{run}.fits',
23 overwrite=True)
```

Cabe mencionar que, por una cuestión de limitación en el tamaño de la memoria RAM disponible, este procedimiento se ha debido ejecutar varias veces para poder obtener los catálogos finales en etapas. Dado el volumen de información generado, con frecuencia nos hemos encontrado en la situación de que la memoria disponible en la computadora de trabajo no era suficiente para concatenar la información de todos los campos a la vez.

Una vez realizado este proceso han quedado construidos dos catálogos finales, uno para **Run 1** y otro para **Run 2**, conteniendo 2.900.926 y 1.390.237 fuentes entre objetos resuelto y no resueltos, respectivamente.

# Capítulo 4

## Resultados

Como primer resultado se obtuvieron dos catálogos conteniendo fotometría homogénea para un total de **2.900.926** fuentes considerando objetos resueltos y no resueltos en el caso de **Run 1**, y un total de **1.390.237** fuentes considerando objetos resueltos y no resueltos en el caso de **Run 2**, ubicadas todas ellas en un área total del cielo austral de  $\sim 208 \text{ grad}^2$  en dirección al cúmulo de galaxias de Fornax. Se espera que dichos catálogos sean publicados como parte de la DR4 de S-PLUS, la cual se hará disponible a la comunidad internacional durante 2023, en carácter de *Value Added Catalogs (VACs)* o catálogos de valor agregado. Hasta que se hagan públicos, ambos catálogos pueden encontrarse en:

[https://drive.google.com/drive/folders/1\\_E0SaQAuAdMkTtv8ZZo2N5OT110HNW0u](https://drive.google.com/drive/folders/1_E0SaQAuAdMkTtv8ZZo2N5OT110HNW0u)

Para validar la fotometría de ambos catálogos, se restringió cada uno de ellos bajo las siguientes condiciones:

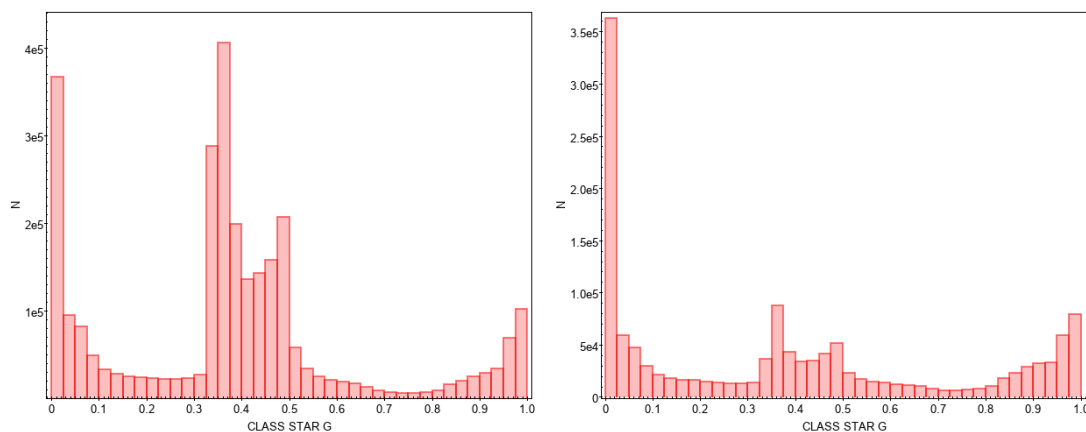
1.  $\text{CLASS\_STAR\_G} < 0.35$   $\text{CLASS\_STAR\_R} < 0.35$   $\text{CLASS\_STAR\_I} < 0.35$ .

Este criterio permite seleccionar con un buen grado de confianza objetos resueltos, eliminando una buena proporción de objetos no resueltos con alta probabilidad de ser estrellas. En la Figura 4.1 se muestra una comparación de histogramas sobre  $\text{CLASS\_STAR\_G}$  entre **Run 1** (izquierda) y **Run 2** (derecha). El histograma de **Run 1** tiene una distribución con varios máximos, de los cuales el máximo absoluto se encuentra en torno a 0.35. Por su parte, para **Run 2**, la distribución presenta un máximo en valores cercanos a 0 (alta probabilidad de ser galaxia). Además, los histogramas reflejan que, en total, el catálogo obtenido con **Run 1** contiene más objetos que **Run 2** y que **Run 1** tiende a detectar objetos más compactos que **Run 2**.

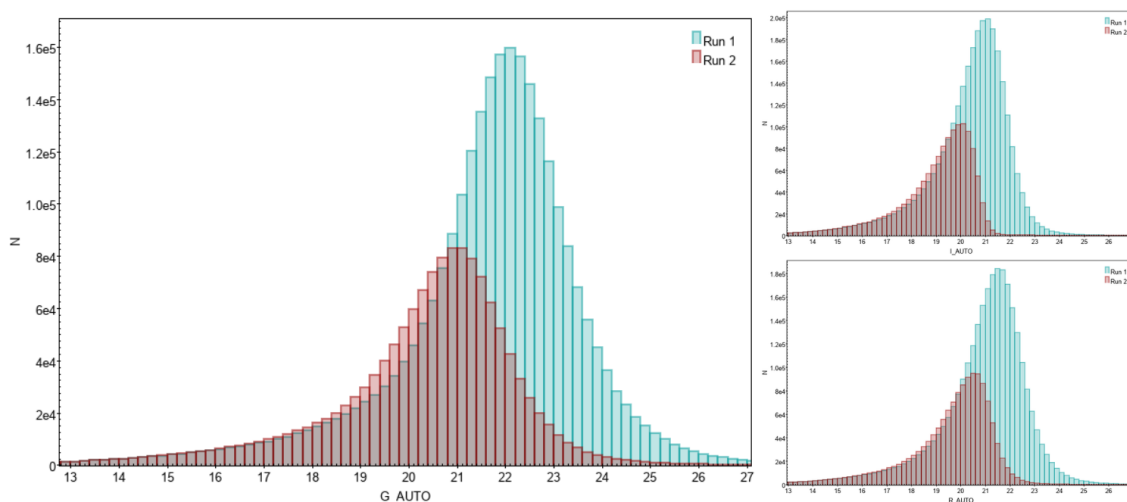
2.  $\text{G\_AUTO} < 22$   $\text{R\_AUTO} < 22$   $\text{I\_AUTO} < 22$ .

Este criterio permite seleccionar objetos con brillos esperables para galaxias teniendo en cuenta la profundidad de los datos de S-PLUS, y descarta objetos que no han sido bien medidos por **SExtractor**, a los cuales se les ha asignado valores de 99 en estas cantidades. En la Figura 4.2 se muestran tres diagramas de completitud, los cuales justifican el valor adoptado para la restricción. Nuevamente en los tres diagramas se aprecia que el catálogo obtenido con **Run 1** es más numeroso que el obtenido con **Run 2**. Asimismo se observa que en cualquiera de los tres filtros **Run 1** logra identificar objetos más débiles (magnitudes

mayores) que **Run 2**. Para cualquiera de los filtros y tanto para **Run 1** como para **Run 2**, las distribuciones llegan a un máximo en magnitudes, el cual difiere filtro a filtro y entre **Run 1** y **Run 2**. Para magnitudes mayores a dicho máximo la cantidad de fuentes detectadas comienza a decaer, lo cual nos indica que la muestra se encuentra completa hasta el máximo de la distribución. El máximo de estas distribuciones también se encuentran asociados a la profundidad del relevamiento S-PLUS. Si S-PLUS fuera más profundo, el máximo de las distribuciones se daría para magnitudes mayores (brillos más débiles).



**Figura 4.1:** Comparación de histogramas sobre CLASS\_STAR\_G entre **Run 1** (izquierda) y **Run 2** (derecha).



**Figura 4.2:** Comparación de diagramas de completitud en G (izquierda), R (derecha/arriba) e I (derecha/abajo). Objetos **Run 1** (verde) y **Run 2** (rojo).

De este modo obtenemos un catálogo **Run 1** restringido que contiene 350.078 objetos (12% del total incluido en el catálogo original) con alta probabilidad de ser fuentes resueltas con brillos acordes a la profundidad de S-PLUS para galaxias, y un catálogo **Run 2** restringido que contiene 357.574 objetos (26% del total incluido en el catálogo original) con las mismas características que los objetos incluidos en el

catálogo **Run 1** restringido. De aquí en más, nos referiremos a estos catálogos como **Run 1<sub>restr</sub>** y **Run 2<sub>restr</sub>**, respectivamente.

Luego de este proceso se tomó como referencia el catálogo **Run 2<sub>restr</sub>** dado que, para objetos extendidos incluidos tanto en el catálogo **Run 1<sub>restr</sub>** como en el **Run 2<sub>restr</sub>**, este último contendrá las magnitudes más confiables gracias a su buena determinación de las aberturas dentro de las cuales dichas cantidades fueron medidas. A continuación, se seleccionaron los objetos incluidos en el catálogo **Run 1<sub>restr</sub>** que no se encontraran en **Run 2<sub>restr</sub>** para sumar aquellas galaxias débiles y objetos compactos que **Run 2<sub>restr</sub>** no incluirá por su construcción original. Adicionando a **Run 2<sub>restr</sub>** los objetos de **Run 1<sub>restr</sub>** que no están incluidos en **Run 2<sub>restr</sub>**, obtenemos un catálogo general (**Run 1 + Run 2**)<sub>restr</sub>.

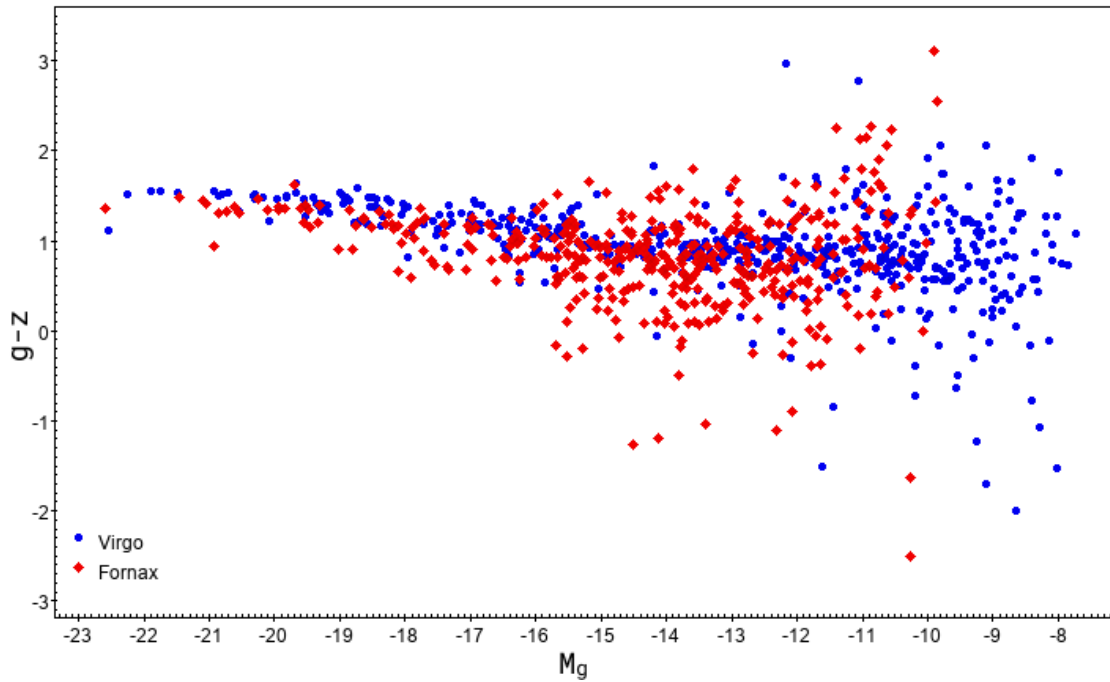
Con el fin de evaluar la calidad de la fotometría obtenida a partir de nuestras ejecuciones de **SExtractor**, se realizó un cruce del catálogo (**Run 1 + Run 2**)<sub>restr</sub> con la lista de las galaxias de Fornax de la literatura para recuperar los miembros del cúmulo que fueron detectados por **SExtractor** con cierto grado de confiabilidad en las imágenes de S-PLUS.

A partir de este cruce se se comparó el diagrama color-magnitud obtenido para Fornax a partir del catálogo (**Run 1 + Run 2**)<sub>restr</sub>, con el resultante de la fotometría SDSS para el Cúmulo de Virgo obtenida por [Chen et al. \(2010\)](#) dada la similitud entre los filtros g,r,i,z de SDSS y S-PLUS. Esta comparación tiene sentido ya que la relación color-magnitud de los cúmulos de galaxias en el Universo cercano resulta ser una relación fotométrica universal [Smith Castelli et al. 2012](#) y referencias allí incluidas. En la Figura [4.3](#) puede apreciarse el buen acuerdo entre ambas fotometrías. Cabe mencionar que, dado este resultado acerca de la confiabilidad de la fotometría obtenida en esta Tesis, el catálogo (**Run 1 + Run 2**)<sub>restr</sub> fue recientemente utilizado para seleccionar objetos a observar en la región del cúmulo de Fornax en el marco del proyecto CHANCES<sup>1</sup> de la colaboración 4MOST<sup>2</sup>.

Otro resultado importante obtenido en el marco de esta Tesis, es que se encontraron dos conjuntos de parámetros de **SExtractor** que nos permiten detectar y medir dos conjuntos diferentes de objetos astronómicos. Los parámetros de **Run 1** detectan correctamente galaxias débiles y objetos compactos ubicados en las cercanías de las galaxias brillantes. Sin embargo, subestiman los tamaños de las aberturas de las galaxias brillantes. Por su parte, los parámetros de **Run 2** detectan las galaxias brillantes con aberturas más apropiadas pero tienden a perder los objetos débiles cercanos a galaxias brillantes. Por lo tanto, los objetos compactos como las galaxias ultra-compactas (UCDs por sus siglas en inglés) o los cúmulos globulares en los alrededores de las galaxias de Fornax, se detectarían mejor con **Run 1** mientras que los objetos más brillantes y extendidos de Fornax se caracterizarían mejor utilizando

<sup>1</sup><https://www.4most.eu/cms/science/extragalactic-community-surveys/>

<sup>2</sup><https://www.4most.eu/cms/>



**Figura 4.3:** Diagrama color-magnitud comparativo mostrando las galaxias de Fornax con fotometría S-PLUS obtenida en esta Tesis, y las galaxias de Virgo con fotometría SDSS obtenida por [Chen et al. \(2010\)](#).

los parámetros de **Run 2**. Cabe mencionar que se planea que ambos conjuntos de parámetros sean utilizados para obtener la fotometría de la DR5 de S-PLUS.

Finalmente, se logró escribir un código Python que nos permitió automatizar casi la totalidad del trabajo, el cual fue pensado para los campos S-PLUS de Fornax y sus alrededores pero es útil, sin modificaciones, para cualquier conjunto de campos de S-PLUS. Dicho código se encuentra disponible en varios repositorios online<sup>3,4</sup>, con su respectiva descripción y licencia para ser compartido públicamente.

<sup>3</sup><https://github.com/rodrihaack/Tesis>

<sup>4</sup><https://colab.research.google.com/drive/1AGaGAM7Ujxqh0keazDP1s6vKShDq59X3?>

# Capítulo 5

## Síntesis y trabajo a futuro

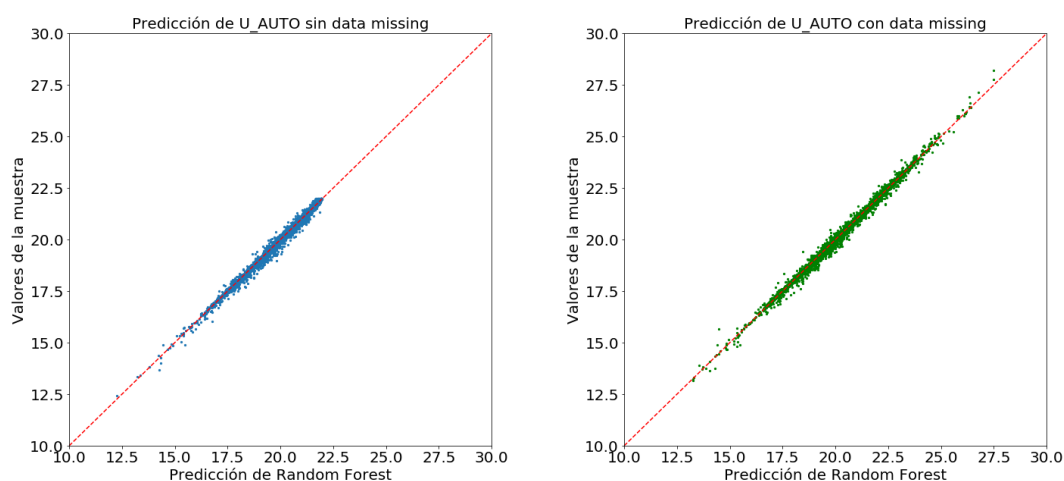
A modo de síntesis, luego del trabajo realizado en el contexto de esta Tesis, se lograron cumplir tanto los objetivos generales como específicos planteados inicialmente. Nos adentramos en la dinámica de trabajo que involucran los grandes volúmenes de datos generados por los grandes relevamientos astronómicos que se encuentran actualmente en desarrollo. Incorporamos herramientas para la obtención de tales datos, como ser el software de detección y medición automática **SExtractor**, estudiando su funcionamiento y posibilidades de ejecución. Adquirimos habilidades de programación en **Python** que nos permitieron obtener y distribuir públicamente un código propio que automatiza la ejecución de **SExtractor** permitiendo su aplicación sobre miles de imágenes de gran campo. Obtuvimos catálogos con fotometría homogénea en 5 filtros de banda ancha y 7 filtros de banda angosta, para  $\sim 3 \times 10^6$  objetos resueltos y no resueltos ubicados en un área del cielo austral de  $\sim 208 \text{ grad}^2$  en la dirección del Cúmulo de Fornax. Para ello hallamos dos conjuntos de parámetros apropiados para detectar y medir en forma automática diferentes tipos de objetos astronómicos: por un lado, objetos débiles y/o compactos cercanos a galaxias brillantes y, por otro, objetos brillantes y muy extendidos. Finalmente, vale la pena resaltar que durante el periodo de tiempo en el que se desarrolló esta Tesis se ha interactuado estrechamente con miembros de la colaboración internacional S-PLUS, lo cual resulta ser un aspecto clave para la próxima etapa de doctorado.

En cuanto a trabajo a futuro se puede mencionar que los catálogos resultantes de esta Tesis cuentan con la información fotométrica en los 12 filtros de S-PLUS para cada uno de los objetos del catálogo, siempre y cuando **SExtractor** haya sido capaz de obtener esa información a partir de las imágenes de los mencionados 12 filtros. Para muchos de los objetos incluidos en los catálogos obtenidos, **SExtractor** no ha logrado medir los flujos en alguno/s de los filtro/s y, en esos casos, a esas magnitudes se les asigna convencionalmente el valor de "99". Este tipo de información faltante es conocido como *data missing*, el cual, en los relevamientos en general y en S-PLUS en particular, es más significativo en los filtros azules del espectro.

En función de ello, y en el marco del curso *Introduction to Machine Learning in*



*Astronomy*<sup>1</sup>, se implementó un script de Machine Learning entrenando una muestra bien comportada, es decir, nos quedamos solo con fuentes que tienen magnitudes no nulas en los 12 filtros del catálogo (**Run 1 + Run 2**)<sub>restr</sub> con el objetivo de lograr una buena estimación en la magnitud de un determinado filtro aprendiendo de los flujos, las magnitudes y los errores obtenidos para el resto de los filtros. Esto se logró utilizando la técnica de *Random Forest Regressor* que luego también se aplicó a una submuestra del mismo catálogo que presentaba *data missing*. *Random Forest* (Breiman, 2001) se puede traducir en bosques aleatorios. La idea principal es que un bosque aleatorio es un estimador que ajusta una serie de árboles de decisión clasificatorios en varias submuestras del conjunto de datos y utiliza el promedio para mejorar la precisión predictiva y controlar el sobre-ajuste. Se utilizó la biblioteca *sklearn* (Pedregosa et al., 2011) y se establecieron como parámetros de entrada: magnitudes, flujos, y sus respectivos errores para los 12 filtros. Los resultados para ambas estimaciones realizadas se muestran en la Figura 5.1.



**Figura 5.1:** Ambos gráficos (izquierda y derecha) muestran la predicción de la magnitud en el filtro U hecha bajo la aplicación de la técnica *Random Forest Regressor*. Para la muestra sin *data missing* (izquierda) se obtuvo un coeficiente de determinación  $R^2=0.9944$  y un error cuadrático medio  $RMSE=0.07$ , mientras que para la muestra con *data missing* (derecha), se obtuvo  $R^2=0.9999$  y  $RMSE=0.08$ . Ambos parámetros se utilizan para calcular el error general del modelo. Se busca que el valor de  $RMSE$  sea lo más bajo posible, mientras que para  $R^2$  se desean valores cercanos a 1.

Ambas muestras fueron filtradas con los mismos criterios que se restringieron los catálogos finales para validar la fotometría. En ese sentido, el proceso de entrenamiento del algoritmo se realizó sobre muestras de objetos extendidos con un 65 % de probabilidad de ser galaxias, de acuerdo al parámetro `CLASS_STAR` que provee *SExtractor* para establecer la separación estrella/galaxia.

<sup>1</sup>Curso dictado por el Dr. Laerte Sodré (IAG-USP) en diciembre de 2022 en la Facultad de Ciencias Astronómicas y Geofísicas de la Universidad Nacional de La Plata, como parte del programa de Profesores Visitantes, invitado por la Dra. Smith Castelli

Aunque ambas predicciones resultan ser satisfactorias, llama la atención que el coeficiente de determinación de la muestra con *data missing* sea mayor al obtenido a partir de la muestra sin *data missing*. El coeficiente de determinación  $R^2$  es un estadístico que toma valores entre 0 y 1. Cuanto más cerca de 1 se sitúe su valor, mayor será el ajuste del modelo a la variable que estamos intentando explicar. De forma inversa, cuanto más cerca de cero, menos ajustado estará el modelo y, por tanto, menos fiable será. En el futuro se planea seguir indagando en este aspecto a través de la interacción con miembros de la colaboración S-PLUS, especialistas en el tema. Contar con catálogos que no presenten este efecto permitiría llevar adelante análisis de poblaciones estelares aplicando técnicas de ajuste de distribución espectral de energía (SED), u obtener redshifts fotométricos, para un mayor número de objetos que el inicialmente esperado a partir de los catálogos originales. Este será el punto inicial de la Tesis de Doctorado continuación de la presente Tesis de Licenciatura, la cual se desarrollará en el Instituto de Astrofísica de La Plata con una Beca Doctoral de CONICET recientemente otorgada bajo la dirección de la Dra. Analía Smith Castelli.

Por otro lado, se planea extender la ejecución de **SExtractor** con los parámetros correspondientes a **Run 1** y **Run 2** a otras zonas del cielo incluidas en la iDR4 de S-PLUS. Como paso inicial de ese objetivo, recientemente se han obtenido los catálogos **Run 1** y **Run 2** para 9 campos S-PLUS que cubren la región de Hydra+Antlia, ha pedido de los responsables de dos subproyectos de S-PLUS que involucran esas áreas del cielo. Asimismo, nos encontramos trabajando en las ejecuciones de **SExtractor** con los parámetros correspondientes a **Run 1** y **Run 2**, pero considerando un total de 32 aberturas circulares concéntricas con el fin de obtener, en forma automática, perfiles de brillo para un gran número de galaxias ubicadas en la dirección del cúmulo de Fornax.

# Bibliografía

- Almeida-Fernandes, F., SamPedro, L., Herpich, F. R., et al.: 2022, MNRAS **511(3)**, 4590
- Ann, H. B., Seo, M., y Ha, D. K.: 2015, ApJS **217(2)**, 27
- Axelrod, T. S.: 2006, en C. Gabriel, C. Arviset, D. Ponz, and S. Enrique (eds.), Astronomical Data Analysis Software and Systems XV, Vol. 351 en Astronomical Society of the Pacific Conference Series, p. 103
- Baker, R. H.: 1933, Annals of Harvard College Observatory **88(3)**, 77
- Baker, R. H. y Shapley, H.: 1937, Annals of Harvard College Observatory **88(6)**, 163
- Baum, W. A.: 1955, PASP **67(398)**, 328
- Benitez, N., Dupke, R., Moles, M., et al.: 2014, arXiv e-prints p. arXiv:1403.5237
- Bertin, E. y Arnouts, S.: 1996, A&AS **117**, 393
- Binggeli, B., Sandage, A., y Tammann, G. A.: 1985, AJ **90**, 1681
- Breiman, L.: 2001, Machine Learning **45**, 5
- Cenarro, A. J., Moles, M., Cristóbal-Hornillos, D., et al.: 2019, A&A **622**, A176
- Chanamé, J., Infante, L., y Reisenegger, A.: 2000, ApJ **530(1)**, 96
- Chen, C.-W., Côté, P., West, A. A., Peng, E. W., y Ferrarese, L.: 2010, ApJS **191(1)**, 1
- Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al.: 2016, MNRAS **460(2)**, 1270
- Dawe, J. A. y Dickens, R. J.: 1976, Nature **263**, 395
- de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al.: 2015, A&A **582**, A62
- de Vaucouleurs, G.: 1977, Nature **266(5598)**, 126

- Drew, J. E., Gonzalez-Solares, E., Greimel, R., et al.: 2014, MNRAS **440(3)**, 2036
- Drinkwater, M. J., Gregg, M. D., y Colless, M.: 2001, ApJL **548(2)**, L139
- Driver, S. P., GAMA Team, Baldry, I. K., et al.: 2009, en J. Andersen, Nordströara, B. m, and J. Bland-Hawthorn (eds.), The Galaxy Disk in Cosmological Context, Vol. 254, pp 469–474
- Eigenthaler, P., Puzia, T. H., Taylor, M. A., et al.: 2018, ApJ **855(2)**, 142
- Ekers, R. D., Goss, W. M., Wellington, K. J., et al.: 1983, A&A **127**, 361
- Ferguson, H. C.: 1989, AJ **98**, 367
- Ferrarese, L., Freedman, W. L., Hill, R. J., et al.: 1996, ApJ **464**, 568
- Freedman, W. L., Madore, B. F., Mould, J. R., et al.: 1994, Nature **371(6500)**, 757
- Gehrels, N., Spergel, D., y WFIRST SDT Project: 2015, en Journal of Physics Conference Series, Vol. 610 en Journal of Physics Conference Series, p. 012007
- Hanes, D. A. y Harris, W. E.: 1986, ApJ **309**, 564
- Hodge, P. W.: 1959, PASP **71(418)**, 28
- Hodge, P. W.: 1960, PASP **72(426)**, 188
- Hodge, P. W., Pyper, D. M., y Webb, C. J.: 1965, AJ **70**, 559
- Holmberg, E.: 1969, Arkiv for Astronomi **5**, 305
- Humason, M. L., Mayall, N. U., y Sandage, A. R.: 1956, AJ **61**, 97
- Infante, L.: 1987, A&A **183**, 177
- Iodice, E., Capaccioli, M., Grado, A., et al.: 2016, ApJ **820(1)**, 42
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al.: 2019, ApJ **873(2)**, 111
- Jones, C., Stern, C., Forman, W., et al.: 1997, ApJ **482(1)**, 143
- Jones, J. E. y Jones, B. J. T.: 1980, MNRAS **191**, 685
- Kron, R. G.: 1980, ApJS **43**, 305
- Lindgren, L. y Perryman, M. A. C.: 1996, A&AS **116**, 579
- Lintott, C., Schawinski, K., Bamford, S., et al.: 2011, MNRAS **410(1)**, 166
- Madore, B. F., Freedman, W. L., Silbermann, N., et al.: 1999, ApJ **515(1)**, 29
- Mendes de Oliveira, C., Ribeiro, T., Schoenell, W., et al.: 2019, MNRAS **489(1)**, 241

- Minniti, D., Lucas, P. W., Emerson, J. P., et al.: 2010, New A. **15(5)**, 433
- Muñoz, R. P., Eigenthaler, P., Puzia, T. H., et al.: 2015, ApJL **813(1)**, L15
- Nair, P. B. y Abraham, R. G.: 2010, ApJS **186(2)**, 427
- Ordenes-Briceño, Y., Eigenthaler, P., Taylor, M. A., et al.: 2018, ApJ **859(1)**, 52
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al.: 2011, Journal of Machine Learning Research **12**, 2825
- Putman, M. E., Bureau, M., Mould, J. R., Staveley-Smith, L., y Freeman, K. C.: 1998, AJ **115(6)**, 2345
- Reaves, G.: 1956, AJ **61**, 69
- Reaves, G.: 1964, AJ **69**, 556
- Schechter, P.: 1976, ApJ **203**, 297
- Schlegel, D. J., Finkbeiner, D. P., y Davis, M.: 1998, ApJ **500(2)**, 525
- Searle, L.: 1965, Nature **207(5003)**, 1282
- Shain, C. A.: 1958, Australian Journal of Physics **11**, 517
- Shanks, T., Metcalfe, N., Chehade, B., et al.: 2015, MNRAS **451(4)**, 4238
- Shapley, H.: 1943, en Galaxies, p. 50
- Simmons, B. D., Lintott, C., Willett, K. W., et al.: 2017, MNRAS **464(4)**, 4420
- Slater, C. T., Ivezić, Ž., y Lupton, R. H.: 2020, AJ **159(2)**, 65
- Smith Castelli, A. V., Bassino, L. P., Richtler, T., et al.: 2008, MNRAS **386(4)**, 2311
- Smith Castelli, A. V., Cellone, S. A., Faifer, F. R., et al.: 2012, MNRAS **419(3)**, 2472
- Smith Castelli, A. V., Mendes de Oliveira, C., Herpich, F., et al.: 2021, Boletín de la Asociación Argentina de Astronomía La Plata Argentina **62**, 180
- Spiekermann, G.: 1992, AJ **103**, 2102
- Storrie-Lombardi, M. C., Lahav, O., Sodre, L., J., y Storrie-Lombardi, L. J.: 1992, MNRAS **259**, 8P
- Struble, M. F.: 1979, AJ **84**, 27
- Sutherland, W., Emerson, J., Dalton, G., et al.: 2015, A&A **575**, A25

- 
- Tammann, G. A. y Kraan, R.: 1978, en M. S. Longair and J. Einasto (eds.), Large Scale Structures in the Universe, Vol. 79, p. 71
- Taylor, M.: 2017
- Tonry, J. y Schneider, D. P.: 1988, AJ **96**, 807
- Tonry, J. L.: 1991, ApJL **373**, L1
- Tonry, J. L., Ajhar, E. A., y Luppino, G. A.: 1989, ApJL **346**, L57
- Tyson, T., Wittman, D., Hennawi, J., y Spergel, D.: 2002, en APS April Meeting Abstracts, APS Meeting Abstracts, p. Y6.004
- Walmsley, M., Smith, L., Lintott, C., et al.: 2020, MNRAS **491(2)**, 1554
- Willett, K. W., Lintott, C. J., Bamford, S. P., et al.: 2013, MNRAS **435(4)**, 2835
- York, D. G., Adelman, J., Anderson, John E., J., et al.: 2000, AJ **120(3)**, 1579