

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
SAN JOAQUÍN - CHILE



“DESARROLLO DE UN PIPELINE BASADO EN DRAFTS
PARA DETECCIÓN Y CLASIFICACIÓN DE TRANSIENTES
DE RADIO CON ADAPTACIÓN Y EXTENSIÓN EN
REGÍMENES DE ALTA FRECUENCIA”

SEBASTIÁN SALGADO POLANCO

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN INFORMÁTICA

Profesor Guía: Daniela Opitz
Profesor Correferente: Marylin Cruces

Enero - 2025

DEDICATORIA

Considerando la importancia de este trabajo para los alumnos, este apartado es para que el autor entregue palabras personales para dedicar este documento. La extensión puede ser de máximo una hoja y se deben mantener este formato, tipo y tamaño de letra.

AGRADECIMIENTOS

Considerando la importancia de este trabajo para los alumnos, este apartado se podrá incluir en el caso de que el autor desee agradecer a las personas que facilitaron alguna ayuda relevante en su trabajo para la realización de este documento. La extensión puede ser de máximo una hoja y se deben mantener este formato, tipo y tamaño de letra.

RESUMEN

Resumen— El resumen y las palabras clave no deben superar la mitad de la página, donde debe precisarse brevemente: 1) lo que el autor ha hecho, 2) cómo lo hizo (sólo si es importante detallarlo), 3) los resultados principales, 4) la relevancia de los resultados. El resumen es una representación abreviada, pero comprensiva de la memoria y debe informar sobre el objetivo, la metodología y los resultados del trabajo realizado.

Palabras Clave— Cinco es el máximo de palabras clave para describir los temas tratados en la memoria, ponerlas separadas por punto y comas.

ABSTRACT

Abstract— Corresponde a la traducción al idioma inglés del Resumen anterior. Sujeto a la misma regla de extensión del Resumen.

Keywords— Corresponde a la traducción al idioma inglés de Palabras Clave anteriores.

GLOSARIO

Aquí se deben colocar las siglas mencionadas en el trabajo y su explicación, por orden alfabético. Por ejemplo:

DI: Departamento de Informática.

UTFSM: Universidad Técnica Federico Santa María.

ÍNDICE DE CONTENIDOS

RESUMEN	IV
ABSTRACT	IV
GLOSARIO	V
ÍNDICE DE FIGURAS	IX
ÍNDICE DE TABLAS	IX
CAPÍTULO 1: INTRODUCCIÓN	1
1.1 Contexto y motivación	1
1.2 Problema	2
1.3 Objetivos	2
1.4 Preguntas de investigación e hipótesis	2
1.5 Principales contribuciones técnicas	3
1.6 Alcance y limitaciones	3
1.7 Organización del documento	3
1.8 Posicionamiento frente a pipelines en tiempo real	4
CAPÍTULO 2: MARCO TEÓRICO	5
2.1 FRBs y pulsos individuales: DM, dispersión, ancho y S/N	5
2.2 Plano tiempo--DM y patrón ``bow-tie"	5
2.3 RFI: criterios de discriminación astrofísico vs terrestre	5
2.4 Dedispersión y waterfalls (sin y con dedispersión)	5
2.5 Particularidades a alta frecuencia (mm-wave)	6
2.6 Telescopios y datos relevantes	6
2.7 Convenciones de polarización (PSR/IEEE) y PSRFITS/PSRCHIVE	6
2.8 Trazabilidad temporal	6
CAPÍTULO 3: ESTADO DEL ARTE	7
3.1 Pipelines y herramientas de búsqueda de pulsos	7
3.2 DRAFTS: detector CenterNet y clasificador ResNet	7
3.3 Estudios a frecuencias altas	7
3.4 Síntesis del gap abordado	7
3.5 Arquitecturas operativas: CHIME/FRB, ASKAP/CRAFT, MeerTRAP	8
3.6 Mitigación de RFI	8
CAPÍTULO 4: REQUISITOS Y DISEÑO (ALTO NIVEL) DEL PIPELINE DRAFTS-MB	9
4.1 Requisitos funcionales del pipeline (qué debe cumplir)	9
4.2 Requisitos no funcionales (latencia, RAM, reproducibilidad)	11
4.3 Arquitectura: módulos, flujos y contratos (diseño de alto nivel)	11
4.4 Modelo de datos y trazabilidad temporal (timestamps y ventanas)	12

4.5	Estrategias de performance (chunking/slicing/overlap/decimado)	12
4.6	Criterios formales: paso de DM por smearing (diseño de Δ DM/overlap)	13
4.7	Control del trials factor y tasas de error (FAR/FPR)	14
4.8	DM--tiempo y dedispersión (fundamentos)	15
4.9	Coste computacional	15
4.10	Flujo de inferencia (t--DM \rightarrow detección \rightarrow clasificación)	16
4.11	Pseudocódigo del flujo (chunk \rightarrow slice \rightarrow detección \rightarrow clasificación)	16
4.12	Temporalidad precisa	17
CAPÍTULO 5: IMPLEMENTACIÓN		18
5.1	Lectura de PSRFITS/FIL y metadatos (implementación concreta)	18
5.2	Preprocesado	20
5.3	Generación de imágenes	20
5.4	Integración de modelos	20
5.5	Diagrama de bloques del pipeline	20
5.6	Pseudocódigo de la rama mm-wave (sin detector de objetos)	20
5.7	Aceleración y configuraciones	21
5.8	Polarización	21
5.9	Umbrales robustos	21
5.10	Cambio de estrategia de detección según banda	22
CAPÍTULO 6: METODOLOGÍA EXPERIMENTAL Y DATASETS		23
6.1	Datasets	23
6.2	Protocolo de experimentos	23
6.3	Métricas	23
6.4	Hardware y software	23
6.5	Inyecciones sintéticas	23
6.6	FAIR y manifests	24
6.7	Análisis del trials factor	24
CAPÍTULO 7: RESULTADOS EN BANDA L (VALIDACIÓN)		25
7.1	Detección y clasificación en test pulsar	25
7.2	FRB121102	25
7.3	Candidatos nuevos	25
7.4	Ablations	25
7.5	Comparación contra baseline	25
7.6	Cross-match con literatura	26
CAPÍTULO 8: EXTENSIÓN A FRECUENCIAS MILIMÉTRICAS (ALMA)		27
8.1	Desafío físico	27
8.2	Estrategia A: DM expandido	27
8.3	Estrategia B: Clasificación a DM cercana a cero	27
8.4	Experimentos y casos límite	27
8.5	Discusión	27
8.6	Rama alternativa basada en SNR	27
8.7	Viabilidad en ALMA phased	28
CAPÍTULO 9: DISCUSIÓN Y AMENAZAS A LA VALIDEZ		29
9.1	Interpretación de resultados	29

9.2	Amenazas internas y externas	29
9.3	Generalización	29
9.4	Sesgos y políticas	29
CAPÍTULO 10: CONCLUSIONES Y TRABAJO FUTURO		30
10.0.1	Resumen de aportes y respuestas a las preguntas	30
10.0.2	Lecciones aprendidas de llevar DRAFTS a producción	30
10.0.3	Próximos pasos	30
10.0.4	Líneas de trabajo futuro	30
ANEXOS		31
REFERENCIAS BIBLIOGRÁFICAS		32

ÍNDICE DE FIGURAS

1.1	Ráfaga de Lorimer: observación de la primera ráfaga de radio rápida detectada, tal como la describió Lorimer en 2006.	1
4.2	Esquema conceptual del sistema de procesamiento por <i>chunks</i> y <i>slices</i> (con solapes).	13
5.3	Convención de metadatos asociados a cada bloque emitido por el lector. . . .	19

ÍNDICE DE TABLAS

CAPÍTULO 1

INTRODUCCIÓN

1.1 Contexto y motivación

En la última década, la radioastronomía ha revelado la existencia de fenómenos transitorios extremadamente breves y energéticos, entre los que destacan las *ráfagas rápidas de radio* (Fast Radio Bursts, FRBs). Las FRBs son pulsos de emisión de radio de duración del orden de milisegundos, generalmente originados a distancias extragalácticas. Su descubrimiento inicial en 2007 marcó un hito por la intensidad y lejanía de estas señales [Lorimer *et al.*, 2007]. El estudio de las FRBs es de gran relevancia científica: estas ráfagas pueden utilizarse como trazadores del medio intergaláctico, aportando información sobre la distribución de materia bariónica y sobre campos magnéticos a escalas cosmológicas, además de ofrecer nuevas oportunidades para la cosmología observacional [Petroff *et al.*, 2022].

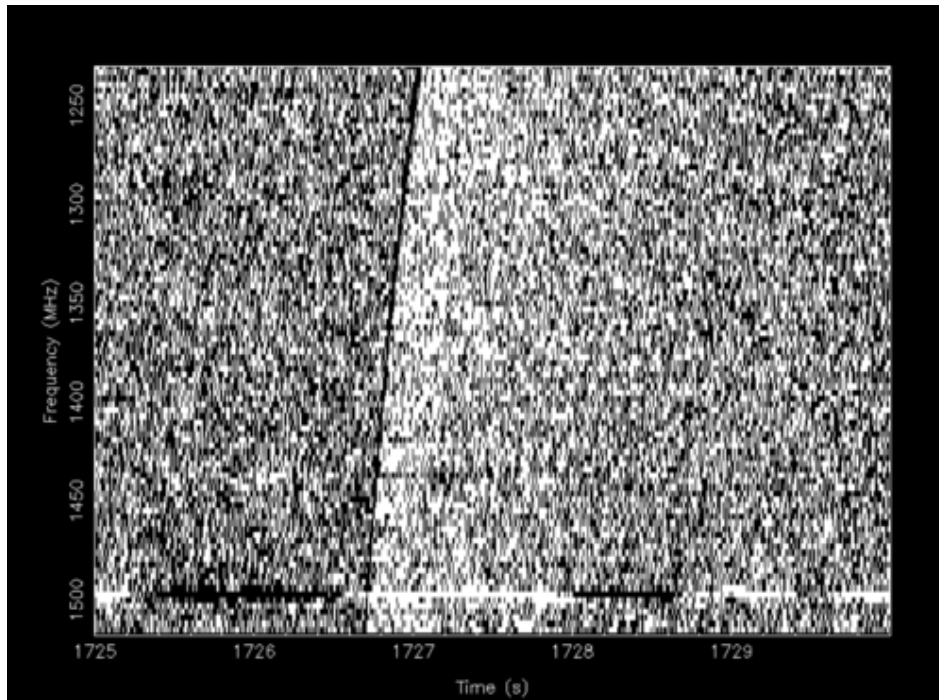


Figura 1.1: Ráfaga de Lorimer: observación de la primera ráfaga de radio rápida detectada, tal como la describió Lorimer en 2006.

1.2 Problema

Detectar FRBs en tiempo de procesamiento plantea desafíos considerables. Los radiotelescopios modernos generan volúmenes masivos de datos, lo que dificulta el procesamiento eficiente de observaciones en busca de eventos de milisegundos. A esto se suma la abundante interferencia de radiofrecuencia (RFI) de origen humano, que contamina las señales e imita pulsos astrofísicos, produciendo grandes listas de candidatos falsos. Los métodos tradicionales de búsqueda de pulsos individuales como los algoritmos implementados en suites clásicas tipo PRESTO y Heimdall se basan en dedispersión exhaustiva y umbrales fijos de detección, si bien han sido exitosos, son propensos a listas extensas de falsos positivos y requieren inspección manual intensiva, lo que limita su uso en operación en tiempo (casi) real [Cordes y McLaughlin, 2003, Ransom *et al.*, 2003, Barsdell *et al.*, 2012]. Por otra parte, **DRAFTS** aporta modelos de detección y clasificación, pero no un pipeline operativo extremo a extremo.

TODO (Problema específico): Completar con 3–4 frases que describan el vacío exacto en tu contexto (datos disponibles, limitaciones de herramientas actuales en tu entorno, necesidad de near-real-time, etc.).

1.3 Objetivos

Objetivo general: Diseñar e implementar un pipeline basado en DRAFTS para detección y clasificación de transientes de radio, y extenderlo a regímenes de alta frecuencia.

Objetivos específicos:

- Integrar y operacionalizar los modelos de DRAFTS en un pipeline reproducible.
- Diseñar módulos de ingesta, preprocesado, detección, clasificación y reporte.
- Adaptar la búsqueda a frecuencias milimétricas (mm-wave) considerando sus particularidades.
- Evaluar desempeño (recall, precision, F1, latencia) en banda L y en mm-wave.

TODO (Medición de objetivos): Indicar cómo verificarás cada objetivo (p.ej., latencia < N s/archivo de 32 GB; F1 > X en FRB121102).

1.4 Preguntas de investigación e hipótesis

¿Cómo se comporta la detección y clasificación de FRBs al comparar banda L con mm-wave?
Hipótesis: la atenuación del patrón bow-tie en mm-wave disminuye sensibilidad de detecto-

res basados en tiempo–DM, mitigable ampliando rango/step de DM y mediante validación por sub-bandas y clasificación a $DM \approx 0$.

TODO (Criterios de falsación): Expresar qué resultados refutarían/confirmarían la hipótesis (p.ej., recall relativo mm-wave/L-band, tasa de falsos $DM \approx 0$).

1.5 Principales contribuciones técnicas

- Pipeline DRAFTS-MB con módulos definidos, contratos I/O y configuración.
- Estrategias de chunking/slices/overlap con criterios formales de ΔDM por smearing.
- Integración CenterNet (detección) + ResNet (clasificación) con umbrales robustos (MAD/IQR).
- Extensión y validación en mm-wave y lineamientos para operación near-real-time.

TODO (Evidencias de contribución): Enumerar outputs verificables (scripts, configs, artefactos, figuras clave, IDs de commit).

1.6 Alcance y limitaciones

El trabajo se enfoca en búsqueda de pulsos individuales (single-pulse) y no aborda timing ni localización interferométrica plena. Se trabaja con PSRFITS/FIL, con disponibilidad de GPU estándar. Limitaciones incluyen RFI a $DM \approx 0$, dependencia de metadatos fiables y variabilidad instrumental.

TODO (Límites explícitos): Añadir supuestos concretos (p.ej., resolución temporal mínima, anchos de banda soportados, formatos exactos, versiones de librerías).

1.7 Organización del documento

Este documento se organiza en diez capítulos más anexos. El Capítulo 2 presenta el marco teórico; el 3, el estado del arte; el 4, los requisitos y el diseño del pipeline; el 5, la implementación; el 6, la metodología experimental y los datasets; el 7, los resultados en banda L; el 8, la extensión a mm-wave (ALMA); el 9, la discusión y amenazas a la validez; y el 10, las conclusiones y trabajo futuro.

1.8 Posicionamiento frente a pipelines en tiempo real

Como motivación y punto de comparación, se discuten arquitecturas operativas de CHIME/FRB, ASKAP/CRAFT y MeerTRAP, destacando diferencias en adquisición, backend, mitigación de RFI y criterios de decisión, que orientan los requisitos y decisiones del presente pipeline.

CAPÍTULO 2

MARCO TEÓRICO

2.1 FRBs y pulsos individuales: DM, dispersión, ancho y S/N

Las FRBs son pulsos de duración típica de milisegundos, caracterizados por su *Medida de Dispersión* (DM), ancho intrínseco/observado y relación señal/ruido (S/N). La dispersión en el medio ionizado introduce un retardo en frecuencia que se aproxima por

$$\Delta t \text{ [ms]} \approx 4,15 \times 10^3 \text{ DM } (\nu_1^{-2} - \nu_2^{-2}),$$

donde ν se expresa en MHz. El ensanchamiento por dispersión y efectos instrumentales (*smearing*) limita la sensibilidad si excede el ancho del pulso.

2.2 Plano tiempo–DM y patrón “bow-tie”

La búsqueda de pulsos individuales se realiza en el plano tiempo–DM, donde una ráfaga astrofísica genera un patrón en “corbatín” (*bow-tie*). La apertura del patrón depende de la cobertura en frecuencia y de la DM del evento; el máximo S/N aparece cerca de la DM verdadera.

2.3 RFI: criterios de discriminación astrofísico vs terrestre

La RFI suele concentrarse en bandas estrechas, presenta DM cercana a cero y carece de la dispersión cromática propia de señales cósmicas. Criterios como consistencia en sub-bandas, simetría temporal, estabilidad de DM y persistencia espacial ayudan a discriminar eventos astrofísicos.

2.4 Dedispersión y waterfalls (sin y con dedispersión)

Los *waterfalls* crudos muestran la dispersión en frecuencia; tras dedispersar a una DM candidata, un pulso real se alinea temporalmente a través de canales. La inspección conjunta tiempo–DM y *waterfalls* dedispersados es clave para validar candidatos.

2.5 Particularidades a alta frecuencia (mm-wave)

A frecuencias milimétricas el retardo por dispersión es menor, atenuando el “bow-tie” y reduciendo la ganancia de S/N asociada a la dedispersión. Esto exige ampliar el rango y el paso de DM o estrategias de validación por sub-bandas y clasificación complementaria.

2.6 Telescopios y datos relevantes

Se consideran instrumentos y formatos representativos: FAST y Effelsberg (banda L), y ALMA en modo *phased* para frecuencias milimétricas. Formatos habituales: PSRFITS/SEARCH y FIL, con metadatos esenciales como TBIN, NCHAN, ancho de banda, frecuencias y tiempos de inicio en MJD.

2.7 Convenciones de polarización (PSR/IEEE) y PSRFITS/PSRCHIVE

Se adopta la convención PSR/IEEE para Stokes I , Q , U , V . La correcta interpretación y almacenamiento de polarización en PSRFITS y su tratamiento en PSRCHIVE son relevantes para análisis avanzados y mitigación de RFI.

2.8 Trazabilidad temporal

Se resguarda trazabilidad mediante sellos de tiempo absolutos (MJD topocéntrico) y consistencia con modelos temporales (p.ej., TEMPO2), asegurando coherencia entre bloques y reproducibilidad de ventanas temporales.

CAPÍTULO 3

ESTADO DEL ARTE

3.1 Pipelines y herramientas de búsqueda de pulsos

Se revisan suites como PRESTO y Heimdall, y otras herramientas relevantes para búsquedas de pulsos individuales, destacando fortalezas y limitaciones en escenarios operativos.

TODO: Añadir referencias y una tabla breve comparando PRESTO vs Heimdall vs otros (entradas, desempeño, salida, facilidad de integración).

3.2 DRAFTS: detector CenterNet y clasificador ResNet

Se describe el enfoque DRAFTS, su arquitectura de detección en tiempo-DM y la clasificación binaria, explicitando qué resuelve y qué queda fuera del alcance (pipeline operativo extremo a extremo).

TODO: Resumir hiperparámetros y tamaños de entrada de los modelos de DRAFTS empleados.

3.3 Estudios a frecuencias altas

Se resumen intentos y hallazgos en bandas altas, identificando brechas y desafíos abiertos.

TODO: Incorporar 1–2 trabajos recientes en mm-wave y su pertinencia para ALMA.

3.4 Síntesis del gap abordado

Se articula el vacío entre modelos de ML y su operacionalización, y la extensión a mm-wave propuesta en este trabajo.

3.5 Arquitecturas operativas: CHIME/FRB, ASKAP/CRAFT, MeerTRAP

Se reseñan arquitecturas en tiempo real/*near-real-time*, mecanismos de backend y publicación de candidatos, como referencia para requerimientos y comparación.

TODO: Esbozar 4–5 bullets por arquitectura con su pipeline y latencias.

3.6 Mitigación de RFI

Panorama de estrategias aceptadas (AOFlagger, enfoques morfológicos/estadísticos) y su relación con el preprocesado del pipeline.

CAPÍTULO 4

REQUISITOS Y DISEÑO (ALTO NIVEL) DEL PIPELINE DRAFTS-MB

4.1 Requisitos funcionales del pipeline (qué debe cumplir)

El sistema debe ejecutar, de forma reproducible y auditable, la cadena completa: **ingesta** → **preprocesado** → **representaciones** (tiempo–DM y waterfalls) → **detección** → **clasificación** → **visualización** → **reporte y trazabilidad**.

Módulos funcionales y contratos I/O

1) Ingesta de datos

- **Entradas:** archivos de búsqueda (p.ej., PSRFITS en modo SUBINT, filterbank), más una configuración de ejecución.
- **Salidas:** flujo de bloques temporales con ejes (tiempo, polarización, canal) y meta-datos mínimos: resolución temporal, número de canales, eje de frecuencias (orden ascendente), duración en muestras, y, cuando exista, tiempo absoluto de inicio (MJD) y correcciones.
- **Contratos:** forma de datos consistente, orden de frecuencias ascendente, preservación de tiempos absolutos/relativos de cada bloque, y detección explícita de errores de cabecera.

2) Preprocesado

- **Entradas:** bloque temporal + parámetros (p.ej., decimado, máscara RFI, correcciones de banda/baseline).
- **Salidas:** bloque normalizado y opcionalmente máscara/artefactos de QA.
- **Contratos:** no alterar la referencia temporal; la salida debe ser idempotente a igualdad de entrada y configuración; registrar parámetros efectivos aplicados.

3) Representaciones (tiempo–DM y waterfalls)

- **Entradas:** datos preprocesados, rango y paso de DM, ventana temporal.

- **Salidas:** imagen tiempo–DM (por *slice*) y waterfall dedispersado para DM candidata; tamaños compatibles con detector/visualización.
- **Contratos:** referencia temporal y eje de DM documentados; asociación inequívoca entre cada *slice* y sus figuras.

4) Detección

- **Entradas:** recortes de la imagen tiempo–DM.
- **Salidas:** candidatos con localización (caja/centro en tiempo y DM) y probabilidad/score; top-k tras supresión de no-máximos si aplica.
- **Contratos:** coordenadas en el mismo sistema de la representación; umbral(es) declarados; latencia por *slice* dentro del presupuesto.

5) Clasificación

- **Entradas:** recortes consistentes (tiempo–DM y/o waterfall dedispersado) centrados en el candidato.
- **Salidas:** etiqueta binaria (FRB/no FRB) o score, con sus umbrales operativos.
- **Contratos:** reproducibilidad (semillas, pesos), latencia por candidato y compatibilidad con auditoría visual.

6) Visualización y auditoría

- **Entradas:** candidato + recortes de sus representaciones.
- **Salidas:** figuras sincronizadas (waterfall dedispersado, tiempo–DM, línea temporal de *slices*).
- **Contratos:** leyendas con tiempos absolutos/relativos, DM estimada y parámetros de ventana; nomenclatura reproducible.

7) Reporte y trazabilidad

- **Entradas:** lista de candidatos con metadatos.
- **Salidas:** tabla de candidatos (p.ej., CSV/Parquet), manifiesto de ejecución (JSON/YAML) con parámetros, versiones y semillas; registros (logs) e índices de artefactos.
- **Contratos:** esquema de columnas estable (tiempos, DM, S/N, scores, rutas a figuras), hashes/DOIs cuando aplique.

Criterios de aceptación (DoD funcional)

- Completa la cadena de valor archivo→candidatos→figuras con artefactos e índices verificables.
- Mantiene trazabilidad temporal absoluta y relativa por bloque y candidato.
- Es reproducible (misma configuración/semillas \Rightarrow mismos candidatos y figuras).

Ejemplo mínimo end-to-end (plantilla)

1. **Entrada:** archivo de observación (p.ej., banda L), parámetros: rango/paso de DM, tamaño de *chunk/slice*, solape, umbrales de detección/clasificación, política de decimado.
2. **Ejecución:** ingesta y extracción de metadatos \Rightarrow preprocesado \Rightarrow generación de tiempo-DM por *slice* \Rightarrow detección \Rightarrow dedispersión óptima \Rightarrow clasificación \Rightarrow registro y visualización.
3. **Salidas:** (i) tabla de candidatos con campos mínimos [t_abs, t_rel, DM, S/N, scores, IDs de bloque y rutas a figuras]; (ii) figuras sincronizadas; (iii) manifiesto de parámetros/semillas/ versionado; (iv) logs con métricas por etapa.

TODO (qué completar aquí): especificar umbrales y objetivos de calidad (p.ej., FAR/h deseada, latencia por archivo/bloque), el esquema definitivo de columnas de la tabla de candidatos, y un ejemplo real instanciado con cifras de tu dataset (tamaños de ventanas, rango de DM, número de canales, tiempos esperados).

4.2 Requisitos no funcionales (latencia, RAM, reproducibilidad)

Operación en tiempo real/*near-real-time*, uso de RAM y reproducibilidad controlada (semillas, versiones, contenedores).

TODO (parámetros objetivo): Fijar límites de latencia por archivo/bloque, tope de RAM y CPU/GPU; incluir política de semillas/versionado (manifest con commit/pesos/entorno) y criterio de éxito.

4.3 Arquitectura: módulos, flujos y contratos (diseño de alto nivel)

Arquitectura en cinco fases: (1) ingesta de datos, (2) preprocesamiento, (3) procesamiento (detección y clasificación), (4) visualización, (5) resultados y registro. Definición de módulos,

interfaces I/O, configuración, registro (logging) y manejo de errores. Los modelos se usan como APIs consumibles para aislar dependencias.

TODO (diagrama): Bocetar las conexiones entre módulos (puertos/interfaces) y definir estados de error/retentos (qué hacer ante fallas parciales y reintentos).

4.4 Modelo de datos y trazabilidad temporal (timestamps y ventanas)

Timestamps absolutos (MJD), ventanas de visualización y metadatos mínimos para reproducibilidad.

TODO (checklist): Enumerar metadatos obligatorios por bloque (MJD, frecuencia de referencia, TBIN, NCHAN, ancho de banda, identificador de backend) y cómo se validan/coinciden entre ejecuciones.

Notación y variables

DM Medida de dispersión (pc cm^{-3}); $k_{\text{DM}} = 4,1488 \times 10^3 \text{ ms MHz}^2 \text{ pc}^{-1} \text{ cm}^3$.

$\nu_{\text{mín}}, \nu_{\text{máx}}, \nu_{\text{ref}}$ Frecuencias (MHz) mínima, máxima y de referencia; ν_c es la del canal c .

$\Delta\nu$ Ancho de canal en MHz; ν_{GHz} es la frecuencia en GHz.

Δt Resolución temporal (ms); Δt_{samp} resolución de muestreo.

$W, \tau_{\text{sc}}, \tau_{\text{chan}}$ Ancho intrínseco y ensanchamientos por scattering y canal.

D, C, S, N N° de DM, canales, muestras por slice y muestras totales en tiempo.

$H \times W$ Tamaño de entrada del detector; $E_{\text{det}}, E_{\text{cls}}$ costos de inferencia.

α, β Constantes de costo (dedispersión, extracción de rasgos/peaks).

O Solapamiento (muestras) entre ventanas; B, B' n° de candidatos por slice (ramas main/HF).

4.5 Estrategias de performance (chunking/slicing/overlap/decimado)

Procesamiento por *chunks* y *slices* con *overlap*, gestión de memoria y paralelismo. Diseño consciente de RAM: coordinación chunk–slice, solapamiento entre chunks para evitar pér-

didadas en bordes por rangos de DM, y controles de resolución/decimado ajustables por el usuario.

TODO (valores guía): Proponer valores por defecto y rangos recomendados para chunk, slice, overlap y decimados; justificar según ecuaciones de smearing y presupuesto de memoria.

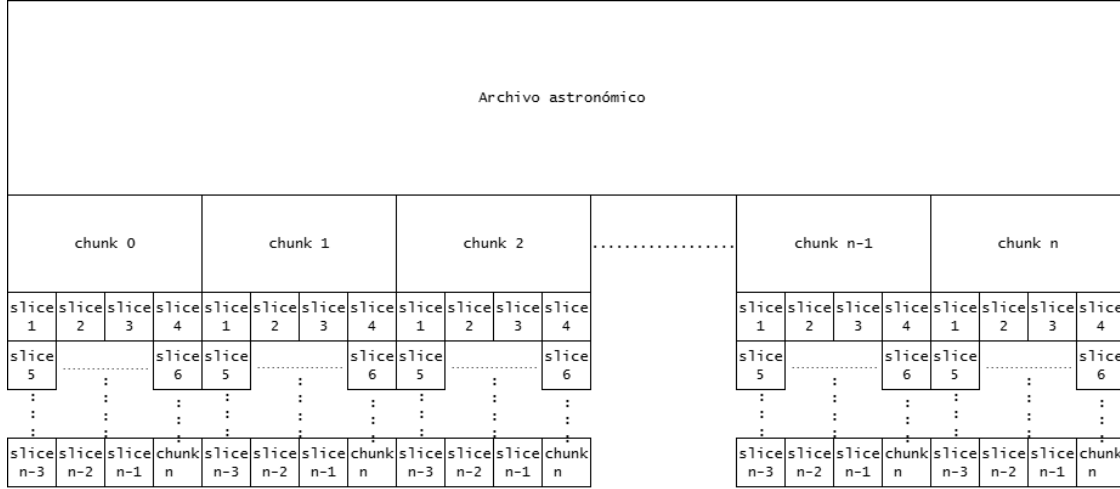


Figura 4.2: Esquema conceptual del sistema de procesamiento por *chunks* y *slices* (con solapes).

Fuente: Elaboración propia

Notas sobre S/N y decimado

Relación S/N (aprox.) al sumar muestras/canales tras dedispersión:

$$\text{SNR} \propto \frac{A}{\sigma} \sqrt{N_{\text{sum}}}. \quad (4.1)$$

Elección de decimado k y filtro tipo boxcar:

$$W_{\text{eff}} \approx k \Delta t_{\text{samp}}, \quad \text{maximizar S/N con ancho de filtro} \approx W_{\text{eff}}. \quad (4.2)$$

4.6 Criterios formales: paso de DM por smearing (diseño de $\Delta\text{DM}/\text{overlap}$)

Se definen pasos de DM y *overlap* mínimo con base en preservar el S/N de pulsos de ancho W . Una relación práctica es mantener el ensanchamiento por dispersión inferior al ancho objetivo.

TODO (ejemplo numérico): Aplicar las fórmulas a tu banda (ν_{\min} , ν_{\max}), TBIN y ancho objetivo W para derivar ΔDM y overlap mínimos; incluir un caso de referencia.

Ecuaciones relevantes

Retardo de dispersión entre dos frecuencias (ν en MHz):

$$\Delta t_{\text{DM}} [\text{ms}] \approx 4,1488 \times 10^3 \text{ DM } (\nu_1^{-2} - \nu_2^{-2}). \quad (4.3)$$

Smearing intra-canal (ν en GHz, $\Delta\nu$ en MHz):

$$\tau_{\text{smear}} [\mu\text{s}] \approx 8,3 \text{ DM } \frac{\Delta\nu_{\text{MHz}}}{\nu_{\text{GHz}}^3}. \quad (4.4)$$

Error por DM desajustada en la banda $[\nu_{\min}, \nu_{\max}]$:

$$\tau_{\text{err}}(\Delta\text{DM}) [\text{ms}] \approx 4,1488 \Delta\text{DM } (\nu_{\min}^{-2} - \nu_{\max}^{-2}). \quad (4.5)$$

Condición de diseño para el paso de DM (con $f \in [0,2, 0,5]$):

$$\tau_{\text{err}}(\Delta\text{DM}) \leq f W_{\text{eff}} \Rightarrow \Delta\text{DM} \leq \frac{f W_{\text{eff}}}{4,1488 (\nu_{\min}^{-2} - \nu_{\max}^{-2})}. \quad (4.6)$$

Solape mínimo entre *chunks*:

$$\text{overlap}_{\min} \geq \Delta t_{\text{DM}}(\text{DM}_{\max \text{ ventana}}) + \text{margen}_{\text{filtro}}. \quad (4.7)$$

4.7 Control del trials factor y tasas de error (FAR/FPR)

Estimación del *trials factor* en el plano DM×tiempo y definición preliminar de FAR/FPR para la toma de decisiones.

TODO (umbrales): Especificar FAR/h objetivo y derivar el umbral z asociado (con $N_{\text{trials/h}}$); documentar impacto en recall/precision.

Ecuaciones relevantes

Número de pruebas aproximado:

$$N_{\text{trials}} \approx N_{\text{DM}} N_{\text{tiempo}} N_{\text{anchos}}. \quad (4.8)$$

Para ruido gaussiano, probabilidad de falso por prueba con umbral z :

$$p_0 \approx Q(z), \quad Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-u^2/2} du. \quad (4.9)$$

Tasa de falsas alarmas por hora (FAR) y elección de z :

$$\text{FAR} \approx N_{\text{trials/h}} p_0 \Rightarrow z = Q^{-1}\left(\frac{\text{FAR}}{N_{\text{trials/h}}}\right). \quad (4.10)$$

4.8 DM–tiempo y dedispersión (fundamentos)

Retardo por canal respecto de una referencia ν_{ref} (frecuencias en MHz):

$$\Delta t_{\text{DM}}(c; \text{DM}) [\text{ms}] = 4,1488 \times 10^3 \text{DM} (\nu_c^{-2} - \nu_{\text{ref}}^{-2}). \quad (4.11)$$

Transformación temporal:

$$t' = t + \Delta t_{\text{DM}}(c; \text{DM}). \quad (4.12)$$

Cubo tiempo–DM (suma en frecuencia):

$$X(d, t) = \sum_{c=1}^C x\left(c, t - \frac{\Delta t_{\text{DM}}(c; \text{DM}_d)}{\Delta t}\right). \quad (4.13)$$

Ancho efectivo del pulso:

$$W_{\text{eff}}^2 = W^2 + \tau_{\text{sc}}^2 + \tau_{\text{chan}}^2, \quad \tau_{\text{chan}} \approx 4,1488 \text{ ms DM} (\nu_{\text{mín}}^{-2} - \nu_{\text{máx}}^{-2}). \quad (4.14)$$

Espaciado de DM (pérdida \lesssim una muestra):

$$\Delta \text{DM} \lesssim \frac{\Delta t}{4,1488 \text{ ms} (\nu_{\text{mín}}^{-2} - \nu_{\text{máx}}^{-2})}. \quad (4.15)$$

4.9 Coste computacional

Tiempo por *slice* (rama principal):

$$T_{\text{slice}}^{(\text{main})} \approx \underbrace{\alpha D C S}_{\text{dedispersión}} + \underbrace{E_{\text{det}}(H, W)}_{\text{infer. CenterNet}} + \underbrace{B E_{\text{cls}}}_{\text{infer. ResNet}}. \quad (4.16)$$

Total por archivo:

$$T_{\text{tot}}^{(\text{main})} \approx \frac{N}{S} \left(\alpha D C S + E_{\text{det}} + B E_{\text{cls}} \right) = \alpha N D C + \frac{N}{S} \left(E_{\text{det}} + B E_{\text{cls}} \right). \quad (4.17)$$

Alta frecuencia (rama sin detector de objetos):

$$T_{\text{slice}}^{(\text{HF})} \approx \alpha D C S + \beta D S + B' E_{\text{cls}}, \quad (4.18)$$

$$T_{\text{tot}}^{(\text{HF})} \approx \alpha N D C + \beta N D + \frac{N}{S} B' E_{\text{cls}}. \quad (4.19)$$

4.10 Flujo de inferencia (t-DM → detección → clasificación)

Generar mapa tiempo-DM para detección (CenterNet) → extraer el candidato y dedispersarlo óptimamente → clasificar (ResNet) → registrar y visualizar (*waterfall* y tiempo-DM consistentes).

TODO (trazabilidad de bloques): Indicar cómo se registra cada etapa (ID de bloque, tiempos relativos/absolutos, parámetros de DM y ventanas) para permitir auditoría posterior.

4.11 Pseudocódigo del flujo (chunk → slice → detección → clasificación)

Entrada: archivo PSRFITS/FIL, config {DM_min, DM_max, DM_step, chunk_ms, slice_ms, overlap_ms}
Salida: lista de candidatos {t_abs, DM, SNR, bbox, clase, figuras}

```
for chunk in ventanas_de_tiempo(archivo, tamaño=chunk_ms, solape=overlap_ms):
    # 1) Preprocesado por chunk
    datos = cargar(chunk)
    datos = corregir_bandpass_y_baseline(datos)
    mascara = generar_mascara_RFI(datos) # AOFlagger/robusta
    datos = aplicar_mascara(datos, mascara)

    # 2) Dedispersión y mapa tiempo-DM (slicing temporal)
    mapa_tDM = crear_mapa_tiempo_DM(datos, DM_min, DM_max, DM_step)
    for slice in ventanas_de_tiempo(mapa_tDM, tamaño=slice_ms, solape=overlap_ms):
        # 3) Detección (CenterNet) sobre el slice del mapa t-DM
        bboxes = det_model.predict(slice)
        for bbox in bboxes_filtradas(bboxes, umbral_det):
            # 4) Extracción y dedispersión óptima del candidato
            t0, DM0 = bbox.centro_t, bbox.centro_DM
            wf = extraer_waterfall(datos, ventana_alrededor=t0)
            wf_dedisp = dedispersar(wf, DM0)

            # 5) Clasificación (ResNet) con recortes consistentes (t-DM y waterfall)
            recorte_tDM = recortar(mapa_tDM, bbox)
            puntaje = cls_model.predict({recorte_tDM, wf_dedisp})
            if puntaje.clase == "FRB" y puntaje.conf > umbral_cls:
                registrar_candidato(t_abs=t0_abs, DM=DM0, SNR=medir_SNR(wf_dedisp),
                                    bbox=bbox, clase=puntaje, figuras={wf_dedisp, recorte_tDM})
```

4.12 Temporalidad precisa

Continuidad de tiempo absoluto entre chunks, corrección temporal tipo PRESTO ante cambios de frecuencia de referencia y metadatos exactos por bloque.

CAPÍTULO 5

IMPLEMENTACIÓN

5.1 Lectura de PSRFITS/FIL y metadatos (implementación concreta)

El pipeline soporta dos familias de formatos de búsqueda de pulsos: PSRFITS (SUBINT) y filterbank SIGPROC. En ambos casos se persigue el mismo objetivo: (i) identificar correctamente la resolución temporal, el número de canales y la duración total; (ii) construir el eje de frecuencias con el orden físico adecuado; (iii) seleccionar o derivar la polarización de trabajo; y (iv) garantizar trazabilidad temporal absoluta (MJD) y coherencia entre bloques.

PSRFITS (SUBINT): extracción y coherencia (detalles prácticos)

- **Resolución temporal y tamaño:** se usa el tiempo por muestra declarado (TBIN) y el producto muestras-por-subintegración (NSBLK) por número de subintegraciones (NAXIS2) para estimar la longitud total del archivo en muestras.
- **Canales y polarización:** se obtiene el número de canales (NCHAN) y el modo de polarización (NPOL/POL_TYPE); por defecto se trabaja con intensidad (Stokes I) o con la componente indicada, manteniendo un arreglo de datos con ejes (tiempo, polarización, canal).
- **Eje de frecuencias:** si el archivo incluye la columna de frecuencias por canal, se adopta directamente; de lo contrario, se reconstruye a partir de los parámetros WCS (valor de referencia, incremento y pixel de referencia). Se invierte el eje si las frecuencias vienen en orden descendente para trabajar internamente en orden ascendente.
- **Tiempo absoluto (MJD):** se calcula el inicio de la observación combinando día/milisegundos y posibles offsets, y se corrige por subintegraciones inicialmente omitidas (NSUBOFFS) o por desplazamientos de la primera fila (OFFS_SUB), lo que preserva la continuidad temporal entre bloques y permite reporte con cronología absoluta.
- **Emisión por bloques:** la lectura se realiza en ventanas temporales (chunks) con solapamientos configurables para evitar pérdidas en bordes; junto a cada bloque de datos se emiten metadatos con los índices válidos, el tamaño efectivo de la ventana, el tiempo relativo (en segundos) y, cuando existe, el tiempo absoluto asociado (MJD corregido).

Filterbank (.fil): cabeceras y construcción del eje de frecuencias (detalles prácticos)

- **Resolución temporal y tamaño:** se toma la resolución temporal por muestra (tsamp) y el número de muestras totales; si el número de muestras no está en cabecera, se deduce a partir del tamaño del archivo y la configuración de canales, bits por muestra y polarizaciones.
- **Canales y frecuencias:** con el número de canales y la pareja (frecuencia inicial, paso por canal) se construye el eje de frecuencias; si el paso por canal es negativo, se invierte el orden para trabajar de forma ascendente.
- **Emisión por bloques:** de manera análoga a PSRFITS, se leen ventanas con solape y se acompaña cada bloque con metadatos que describen su posición temporal relativa y su geometría en frecuencia.

Metadatos mínimos por bloque (para auditoría y reproducibilidad)

De cara a la reproducibilidad y auditoría, cada bloque emitido debe llevar al menos: (i) índices de inicio y fin válidos dentro del archivo; (ii) número de muestras efectivas en la ventana; (iii) resolución temporal de la muestra y tiempos relativos de inicio/fin; (iv) número de canales y forma del bloque; (v) indicadores de solape aplicado a izquierda/derecha; y, cuando corresponde, (vi) tiempo absoluto de inicio de la observación (MJD) y su corrección por subintegraciones omitidas. Esta convención permite trazar cada candidato a su ventana de origen y cruzar resultados entre ejecuciones.

(Diagrama conceptual de bloque y metadatos)

Dato	Tensor (tiempo, polarización, canal)
Índices	start_sample, end_sample, actual_chunk_size
Tiempo relativo	tbin_sec, t_rel_start_sec, t_rel_end_sec
Tiempo absoluto	tstart_mjd, tstart_mjd_corr (si aplica)
Geometría	nchans, shape, dtype
Solape	overlap_left, overlap_right

Figura 5.3: Convención de metadatos asociados a cada bloque emitido por el lector.

Fuente: Elaboración propia

5.2 Preprocesado

Reducción de muestreo (down-sampling), corrección de banda/baseline y máscaras de RFI (AOFlagger y estrategias robustas).

TODO: Enlazar 'src/preprocessing/' (bandpass, baseline, RFI) y parámetros por defecto.

5.3 Generación de imágenes

Producción de mapas tiempo-DM y *waterfalls* dedispersados para inspección y entrada a modelos.

TODO: Enlazar funciones para mapa t-DM y waterfall con formatos de salida.

5.4 Integración de modelos

Los modelos de DRAFTS (CenterNet para detección y ResNet para clasificación) se integran como *APIs* consumibles. El flujo estándar (bajas frecuencias) detecta en tiempo-DM (CenterNet) y clasifica con ResNet, aplicando umbrales y NMS.

TODO: Apuntar a 'src/models/' (carga de pesos, versiones/hashes) y 'src/detection/', 'src/analysis/'.

5.5 Diagrama de bloques del pipeline

Ingesta → Preprocesado (bandpass/baseline, RFI) → Mapa tiempo-DM → Detección (CenterNet) → Dedispersión óptima → Clasificación (ResNet) → Registro y visualización (waterfall + t-DM)

5.6 Pseudocódigo de la rama mm-wave (sin detector de objetos)

```
for chunk in ventanas(archivo, chunk_ms, overlap_ms):
    datos = preprocesar(chunk)
    wf = waterfall(datos)
    picos = detectar_picos_SNR(wf, umbral_robusto(MAD/IQR))
    for pico in agrupar_picos(picos):
        bbox_sint = construir_bbox_alrededor(pico)
```

```
recorte_tDM = sintetizar_recorte_tDM(datos, pico) # opcional
puntaje = cls_model.predict({recorte_tDM, wf_alrededor(pico)})
if puntaje.clase == "FRB" y puntaje.conf > umbral:
    registrar_candidato(...)
```

Pérdidas y decisión (para referencia)

Pérdida de detección (CenterNet):

$$\mathcal{L}_{\text{det}} = \sum_{i,j} \text{FL}(\hat{H}_{ij}, H_{ij}) + \lambda_{\text{off}} \|\hat{\Delta}_{ij} - \Delta_{ij}\|_1 + \lambda_{\text{size}} \|\hat{S}_{ij} - S_{ij}\|_1. \quad (5.20)$$

Pérdida de clasificación (ResNet):

$$\mathcal{L}_{\text{cls}} = -\frac{1}{B} \sum_{b=1}^B \left[y_b \log \hat{p}_b + (1 - y_b) \log(1 - \hat{p}_b) \right]. \quad (5.21)$$

Umbrals de decisión:

$$\hat{p}^{(\text{det})} \geq \tau_{\text{det}}, \quad \hat{p}^{(\text{cls})} \geq \tau_{\text{cls}}, \quad \text{Score}(b) = (\hat{p}_b^{(\text{det})})^\lambda (\hat{p}_b^{(\text{cls})})^{1-\lambda}. \quad (5.22)$$

5.7 Aceleración y configuraciones

Uso de CPU/GPU, organización del código y parámetros clave de ejecución.

TODO: Enlazar 'src/config/' y 'src/core/' (planificador, colas, lotes) y flags de ejecución.

5.8 Polarización

Manejo coherente de Stokes I, Q, U, V conforme a PSR/IEEE cuando corresponda.

TODO: Indicar si tu dataset usa I únicamente o incluye Q/U/V y cómo se procesa.

5.9 Umbrals robustos

Estimadores robustos ($\text{MAD} \rightarrow \sigma = 1.4826, \text{IQR}/1.349$) para umbrales en ramas basadas en S/N.

TODO: Fijar umbrales por defecto y cómo se calibran (experimentos del Cap. 6).

5.10 Cambio de estrategia de detección según banda

Pipeline estándar (bajas frecuencias).

- Usa detector de objetos (`det_model`) sobre imagen tiempo–DM.
- Genera cajas de detección automáticamente y envía recortes al clasificador.

Pipeline de alta frecuencia (mm-wave).

- Elimina el detector de objetos (no usa `det_model`).
- Emplea detección por picos de SNR directamente del *waterfall*.
- Genera cajas sintéticas centradas en picos SNR para la clasificación.

CAPÍTULO 6

METODOLOGÍA EXPERIMENTAL Y DATASETS

6.1 Datasets

Conjunto de datos: pulsar de prueba (p.ej., B0355+54), FRB 121102; rangos de DM, TBIN, NCHAN y preparación.

TODO: Completar tabla con archivos, tamaños (GB), MJD, TBIN, NCHAN, BW.

6.2 Protocolo de experimentos

Semillas, ablations y barridos de parámetros; número de repeticiones y control de aleatoriedad.

TODO: Especificar seeds fijas, número de corridas por configuración y parámetros barridos.

6.3 Métricas

Recall, precision, F1, S/N y latencia por archivo/bloque, además de FAR estimada.

TODO: Definir cómo se calcula FAR/h y qué umbrales aplican por etapa.

6.4 Hardware y software

Descripción de GPU/CPU, versiones de software y mecanismos de reproducibilidad (conda/contenedores).

TODO: Añadir 'environment.yml'/Docker y versiones exactas (CUDA/cuDNN, PyTorch).

6.5 Inyecciones sintéticas

Inyección de pulsos sintéticos para curvas de eficiencia (recall vs S/N, ancho, DM).

TODO: Especificar distribución de S/N, anchos, DM y cantidad de inyecciones por archivo.

6.6 FAIR y manifests

Buenas prácticas FAIR: DOIs, hashes de pesos, contenedores/environment y manifests de experimentos.

6.7 Análisis del trials factor

Estimación del *trials factor* ($DM \times \text{tiempo}$) y su impacto en FAR por hora.

CAPÍTULO 7

RESULTADOS EN BANDA L (VALIDACIÓN)

7.1 Detección y clasificación en test pulsar

Resultados de detección y clasificación sobre pulsar de prueba (p.ej., B0355+54), con métricas S/N y precisión de clasificador.

TODO: Añadir figura (waterfall dedispersado) y tabla con S/N y tiempos detectados.

7.2 FRB121102

Validación en banda L sobre FRB121102 (archivos de 32 GB): paridad con eventos reportados en la literatura y evidencia de detecciones adicionales. Reporte por archivo: MJD, rango de DM, S/N y figuras (tiempo–DM y *waterfalls* dedispersados).

TODO: Completar lista de archivos, eventos replicados y nuevos candidatos (IDs, ventanas, figuras).

7.3 Candidatos nuevos

Evidencia mínima replicable (ID de archivo, ventanas temporales, figuras) y discusión de plausibilidad.

TODO: Incluir criterios de aceptación/rechazo y justificación por sub-bandas.

7.4 Ablations

Efecto de *chunking/slices/overlap* sobre la sensibilidad y latencia.

7.5 Comparación contra baseline

Comparativa con un pipeline tradicional (si aplica), incluyendo tablas de recall/precision/F1 y FAR estimada.

7.6 Cross-match con literatura

Cruce con listas publicadas para FRB 121102 y discusión de consistencias/diferencias con trabajos de referencia.

CAPÍTULO 8

EXTENSIÓN A FRECUENCIAS MILIMÉTRICAS (ALMA)

8.1 Desafío físico

Atenuación del *bow-tie* en mm-wave y reducción del retardo por dispersión.

8.2 Estrategia A: DM expandido

Ampliar rango y paso de DM para “abrir” el *bow-tie* y recuperar sensibilidad.

TODO: Especificar rango y step de DM propuestos para mm-wave y justificación.

8.3 Estrategia B: Clasificación a DM cercana a cero

Ejecutar clasificador a DM cercana a cero y validar con $DM > 0$ y consistencia por sub-banda/*chunk*.

TODO: Definir criterios de consistencia en sub-bandas y umbrales asociados.

8.4 Experimentos y casos límite

Resultados y análisis de falsos positivos a DM cercana a cero, incluyendo RFI.

8.5 Discusión

Cuándo falla el detector, cuándo acierta el clasificador y límites de la estrategia.

8.6 Rama alternativa basada en SNR

Detección por picos en Stokes I con umbrales robustos (MAD 1.4826, IQR/1.349), agrupamiento de picos y cajas sintéticas para el clasificador.

TODO: Documentar cálculo de umbral robusto (MAD/IQR) y ventana de agrupamiento.

8.7 Viabilidad en ALMA phased

Fork del pipeline para validar DRAFTS en ALMA en tres modos: intensidad, polarización circular y polarización lineal. A partir de estos resultados, iterar en nuevos *plots* y futuros entrenamientos. Resultados/viabilidad con magnetar PSR J1745–2900 y motivación para buscar transientes en Band 3.

CAPÍTULO 9

DISCUSIÓN Y AMENAZAS A LA VALIDEZ

9.1 Interpretación de resultados

Discusión de desempeño en banda L versus mm-wave y su significado.

9.2 Amenazas internas y externas

Datos, hiperparámetros, selección de ventanas y RFI como posibles fuentes de sesgo.

9.3 Generalización

Transferencia a otros telescopios y bandas; dependencias instrumentales.

9.4 Sesgos y políticas

Sesgos por *trials factor*, falsos positivos a $DM \approx 0$, dependencia de hardware (latencia/throughput) y políticas de datos (periodos propietarios en ALMA) con buenas prácticas FAIR.

CAPÍTULO 10

CONCLUSIONES Y TRABAJO FUTURO

10.0.1. Resumen de aportes y respuestas a las preguntas

Se sintetizan las contribuciones técnicas del pipeline DRAFTS-MB, su desempeño en banda L y su extensión a frecuencias milimétricas, respondiendo las preguntas de investigación sobre sensibilidad y precisión comparativas.

10.0.2. Lecciones aprendidas de llevar DRAFTS a producción

Se discuten hallazgos sobre integración de modelos, control del *trials factor*, robustez a RFI y gestión de memoria/latencia en operación *near-real-time*.

10.0.3. Próximos pasos

Se delinean reentrenos específicos para mm-wave, adaptación de dominio, procesamiento en flujo continuo (streaming) e integración con pipelines de observatorios.

TODO: Priorizar 3 acciones con plazos tentativos (p.ej., reentrenos mm-wave, pruebas en Band 3, integración con backend).

10.0.4. Líneas de trabajo futuro

Se proponen líneas para localización y consistencia por sub-bandas, validaciones cruzadas con pipelines externos (CHIME/FRB, MeerTRAP, ASKAP/CRAFT), y estrategias FAIR para datos/modelos. Si la detección falla en mm-wave por falta de *bow-tie* (línea casi recta en tiempo–DM), reentrenar/adaptar el modelo o desarrollar uno nuevo específico para regímenes milimétricos con *plots* característicos.

ANEXOS

En los Anexos se incluye todo aquel material complementario que no es parte del contenido de los capítulos de la memoria, pero que permiten a un lector contar con un contenido adjunto relacionado con el tema.

REFERENCIAS BIBLIOGRÁFICAS

- [Barsdell *et al.*, 2012] Barsdell, B. R., Bailes, M., Barnes, D. G., y Fluke, C. J. (2012). Accelerating incoherent dedispersion. *Monthly Notices of the Royal Astronomical Society*, 422(1):379–392.
- [Cordes y McLaughlin, 2003] Cordes, J. M. y McLaughlin, M. A. (2003). Searches for fast radio transients. *The Astrophysical Journal*, 596(2):1142–1154.
- [Lorimer *et al.*, 2007] Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., y Crawford, F. (2007). A bright millisecond radio burst of extragalactic origin. *Science*, 318(5851):777–780.
- [Petroff *et al.*, 2022] Petroff, E., Hessels, J. W. T., y Lorimer, D. R. (2022). Fast radio bursts at the dawn of the 2020s. *The Astronomy and Astrophysics Review*, 30(1).
- [Ransom *et al.*, 2003] Ransom, S. M., Cordes, J. M., y Eikenberry, S. S. (2003). A new search technique for short orbital period binary pulsars. *The Astrophysical Journal*, 589(2):911–920.