Cosmology Cartography

Mapeo de redshift por píxel en imágenes multibanda Resultados de modelo de *deep learning* sobre HSC-SSP PDR3

Juan Camilo Sánchez

Mitacs Globalink · University of Victoria

Supervisor: Prof. Sébastien Fabbro

CoCo · Cosmology in Colombia · 2026





Contenido

1. Por qué un mapa de redshift por píxel
2. Datos: catálogo y etiquetas
3. Problema de regresión
4. Problema de clasificación (z a $P(z)$)
5. Balance, futuros pasos y conclusiones.

Parte 1

¿Por qué un mapa de redshift por píxel?

El redshift es la coordenada que le falta a la imagen

- Una imagen registra posición angular; el redshift aporta la dirección radial.
- Con él: tomografía de la red cósmica, $N(z)$ regional para lensing débil, correlaciones con lensing del CMB.
- Los métodos tradicionales dan una vista dispersa: sólo fuentes detectadas y catalogadas.
- Fallan donde aún hay información: fuentes débiles, *blending*, bajo brillo superficial.

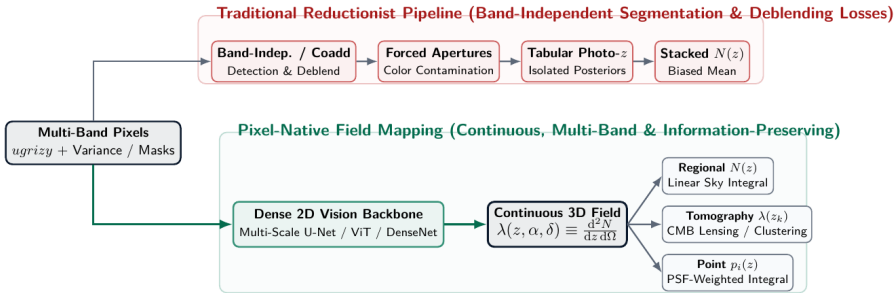
Objetivo de largo plazo

$$\lambda(z, \alpha, \delta) \equiv \frac{d^2N}{dz\, d\Omega} = \rho(x, y)\; p(z \mid x, y)$$

ρ: densidad proyectada de fuentes

$p(z \mid x, y)$: distribución de redshift en cada línea de visión

Método esperado



Arriba: detección, aperturas y photo-z por objeto. Abajo: una red densa predice el campo continuo directamente desde los píxeles; $N(z)$, cortes tomográficos y $p_i(z)$ salen como proyecciones lineales.

Figura: S. Fabbro, *Pixel-Native Redshift Density Mapping* (propuesta).

Objetivos de la pasantía (12 semanas)

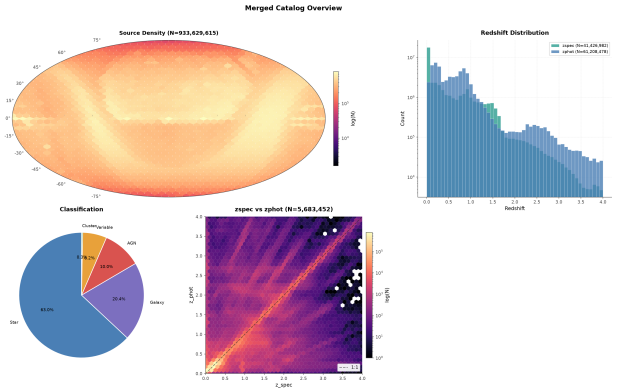
1. Preparar e inspeccionar un conjunto pequeño de imágenes multibanda.
2. Implementar o adaptar componentes de *deep learning* de línea base para redshift por píxel.
3. Contrastar las predicciones con redshifts de referencia, incluyendo su incertidumbre.
4. Evaluar si el método recupera estructura útil en fuentes débiles, *blended* o de bajo brillo superficial.

Parte 2

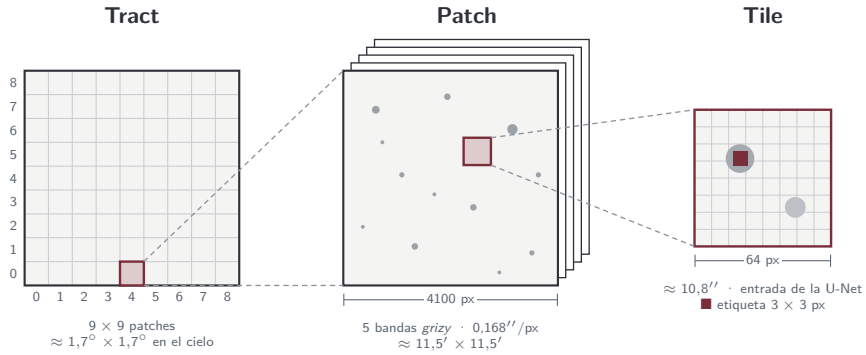
Datos: catálogo y etiquetas

zscrape: Catalogo unificado de redshifts

- Decenas de catálogos (DESI, SDSS, DEEP2/3, VIPERS, COSMOS2020, . . .) en un esquema común.
- Desduplicación espacial; formato HATS.
- **[+900M fuentes]**.
- Peso de calidad $W_{\text{spec}} \in [0, 1]$ por fuente, calibrado por consenso entre catálogos.



Organización de los datos: tract → patch → tile



`calexp-HSC-G-10055-4,0.fits`

banda · tract · patch (x, y) HDUs: imagen, máscara, varianza

60 patches, split por patch



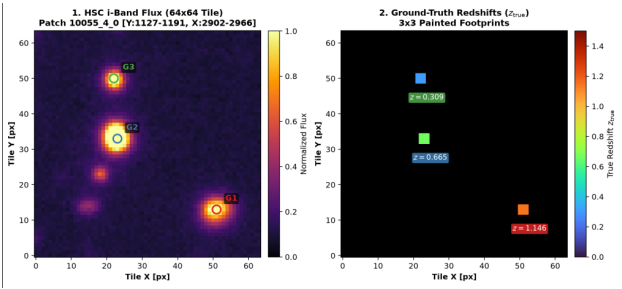
HSC-SSP PDR3 Deep+UltraDeep · un mapa de etiquetas de 6 planos por patch, en la misma malla WCS.

Imágenes y etiquetas

Imágenes: HSC-SSP PDR3 Deep+UltraDeep, patches de 4100×4100 px en *grizy*.

Etiquetas (misma malla WCS):

- 6 planos: z, z_{lo}, z_{hi}, `IS_SPEC`, máscara, W_{spec}
- Bloques de 3×3 px por fuente
- 1 530 mapas, $\sim$12,5 M fuentes



Parte 3

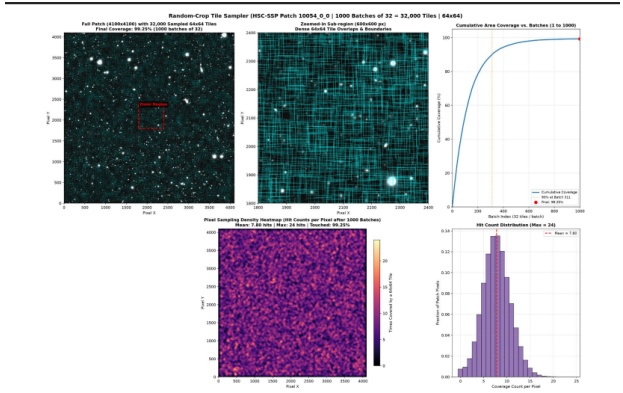
Problema de regresión

Pipeline de entrenamiento

- U-Net, 5 bandas $\rightarrow$ 1 mapa denso.
- Tiles de 64×64 px con esquinas aleatorias.
- Normalización interquantile conjunta:

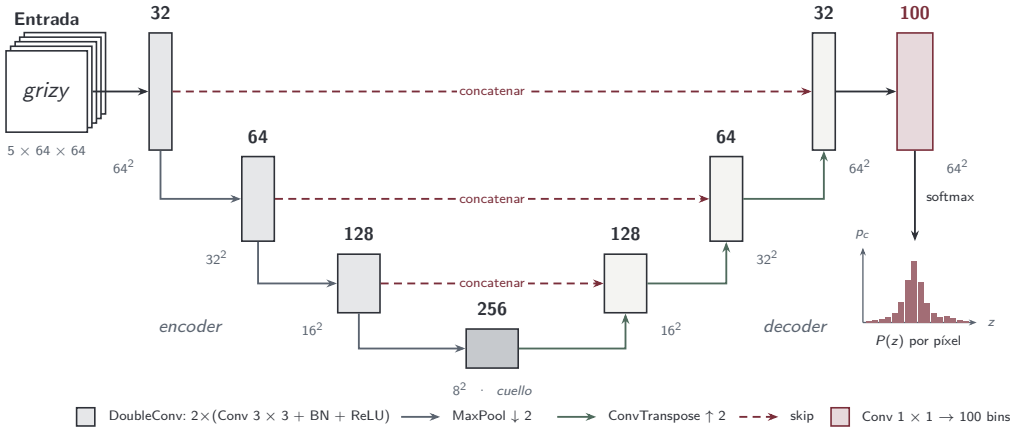
$$x' = \mathrm{clip}\left(\frac{x - Q_{1\,\%}}{Q_{99\,\%} - Q_{1\,\%}}, 0, 1 \right)$$

- Split por patch: 40 / 10 / 10.
- Entorno: CANFAR, GPU H100 (MIG, $\sim$11 GB).



Con 1 000 lotes de 32 tiles el muestreo cubre el 99,25 % del patch.

Arquitectura: U-Net de clasificación



Altura de cada bloque $\propto$ resolución espacial; ancho $\propto$ número de canales. Pérdida: entropía cruzada ponderada por W_{spec}.

Regresión: Resumen

- **Normalización**: 8 variantes; gana interquantile.
- **Pérdida**: MAE enmascarada sobre $\log(1+z)$, ponderada por calidad:

$$\mathcal{L} = \frac{\sum_i W_i\, M_i\, \left|\log(1+\hat{z}_i) - \log(1+z_i)\right|}{\sum_i W_i\, M_i}$$

- **Aumentos**: ruido gaussiano ($0{,}1$–$0{,}6\,\sigma_{\mathrm{sky}}$) y escalado aleatorio de una banda ($\pm 1\,\%$).
- Patch contaminado por un halo estelar detectado y reemplazado.

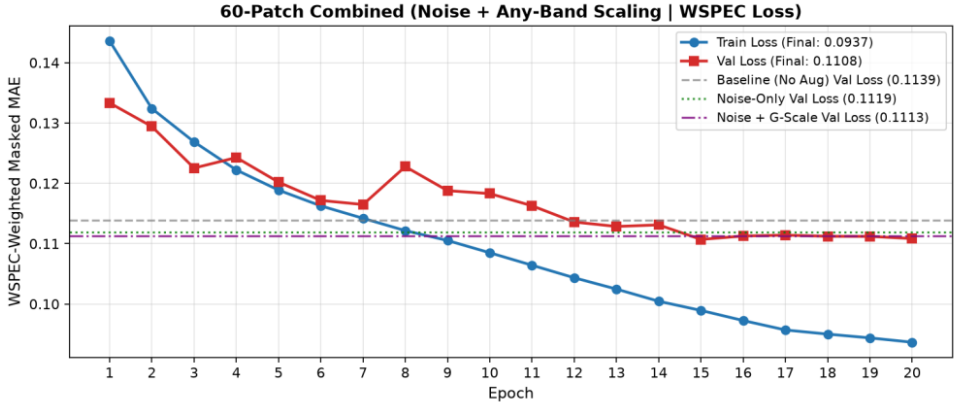
Mejor modelo de regresión (test)

Fuentes	9 823
σ_{NMAD}	0,055
η_{15}	24,4 %
Bias	−0,003
MAE	0,211

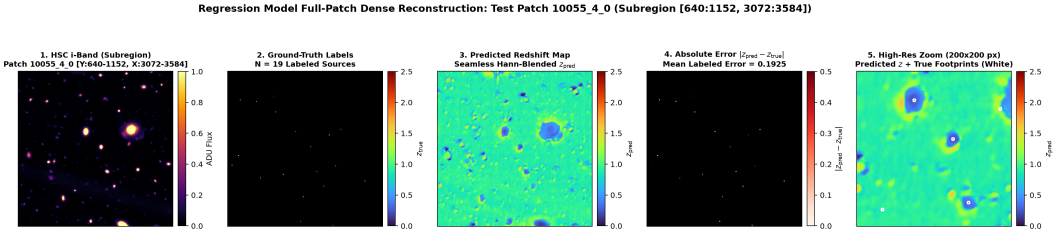
Por fuente, espectroscópicas de alta calidad (`IS_SPEC`$= 1$, $W_{\mathrm{spec}} > 0{,}7$); mapa reconstruido sin costuras, una predicción por píxel.

Limitación: casi no predice $z \gtrsim 2$, y da un solo número sin incertidumbre.

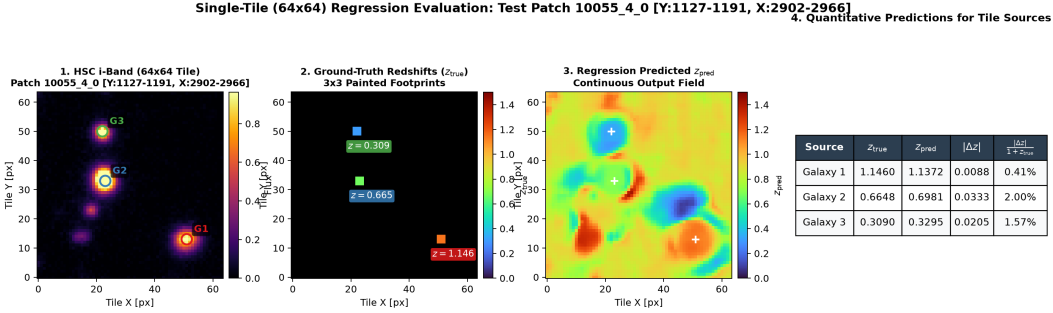
Loss function



Resultados Full Patch

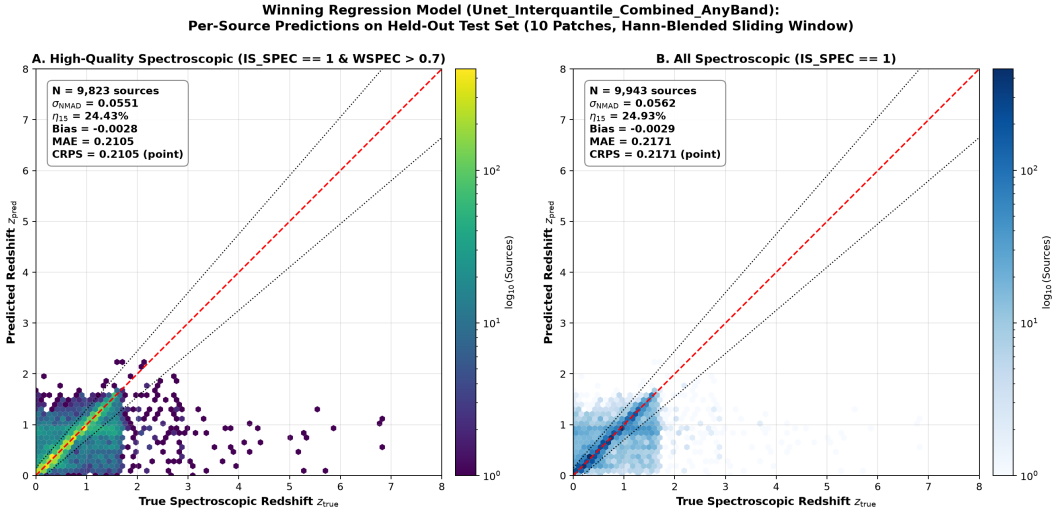


Resultados para 1 Tile



Source	z_{true}	z_{pred}	$\lvert\Delta z\rvert$	$\frac{\lvert\Delta z\rvert}{1+z_{true}}$
Galaxy 1	1.1460	1.1372	0.0088	0.41%
Galaxy 2	0.6648	0.6981	0.0333	2.00%
Galaxy 3	0.3090	0.3295	0.0205	1.57%

Resultados


Winning Regression Model (Unet_Interquantile_Combined_AnyBand):
Per-Source Predictions on Held-Out Test Set (10 Patches, Hann-Blended Sliding Window)
A. High-Quality Spectroscopic (IS_SPEC == 1 & WSPEC > 0.7)
N = 9,823 sources
σ_{NMAD} = 0.0551
η_{15} = 24.43%
Bias = -0.0028
MAE = 0.2105
CRPS = 0.2105 (point)
Predicted Redshift z_{pred}
True Spectroscopic Redshift z_{true}
$\log_{10}$(Sources)
B. All Spectroscopic (IS_SPEC == 1)
N = 9,943 sources
σ_{NMAD} = 0.0562
η_{15} = 24.93%
Bias = -0.0029
MAE = 0.2171
CRPS = 0.2171 (point)
Predicted Redshift z_{pred}
True Spectroscopic Redshift z_{true}
$\log_{10}$(Sources)

Parte 4

Problema de clasificación (z a $P(z)$)

Clasificación: PDF por píxel

- El redshift se discretiza en 100 bins; la salida es un softmax por píxel.
- Etiqueta: gaussiana integrada en cada bin,

$$q_c = \Phi\left(\frac{z_{c+1}-z}{\sigma}\right) - \Phi\left(\frac{z_c-z}{\sigma}\right)$$

  con $\sigma = \sigma_{\text{spec}}$ o $(z_{\text{hi}} - z_{\text{lo}})/2$.
- Pérdida: entropía cruzada ponderada,

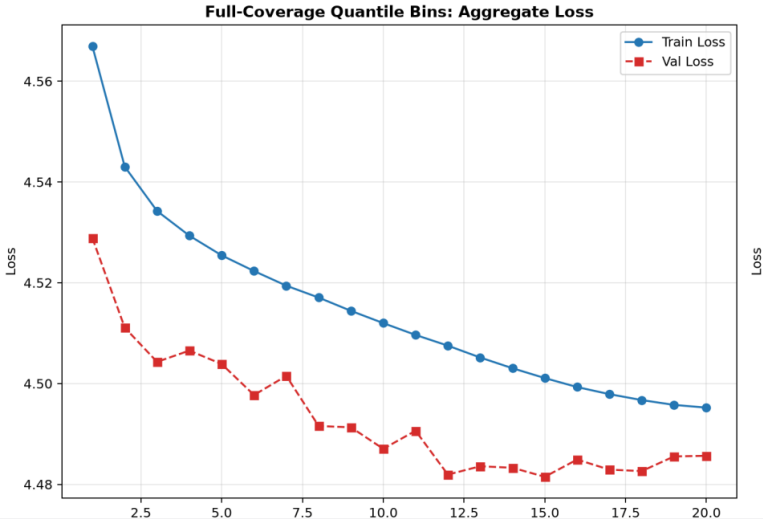
$$\mathcal{L} = -\frac{\sum_i W_i \sum_c q_{ic} \log p_{ic}}{\sum_i W_i}$$

Dos esquemas de bins

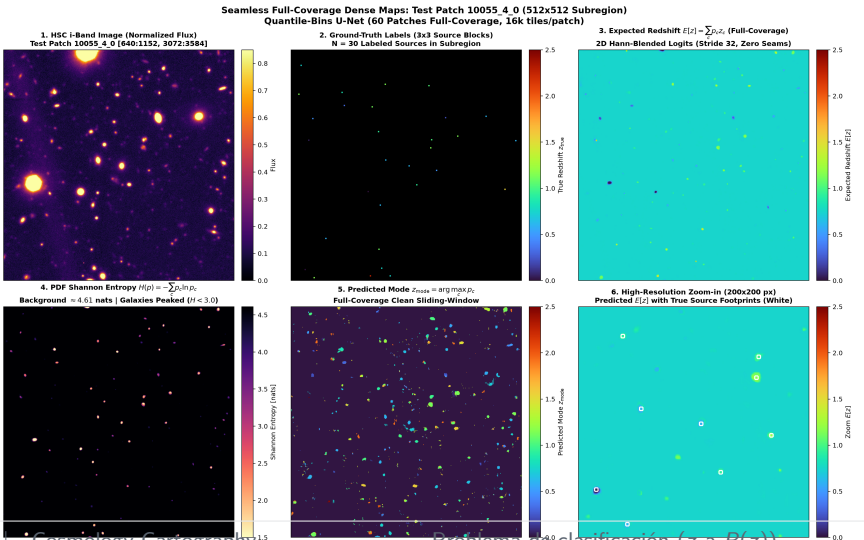
- **Fijos**, $\Delta z = 0{,}1$ en $[0, 10]$: las espectroscópicas caen en un solo bin y 75 bins quedan casi vacíos.
- **Cuantiles** de las z_{spec} de entrenamiento: bins finos donde hay datos.

Los cuantiles son la versión discreta de los nudos M-spline de la propuesta.

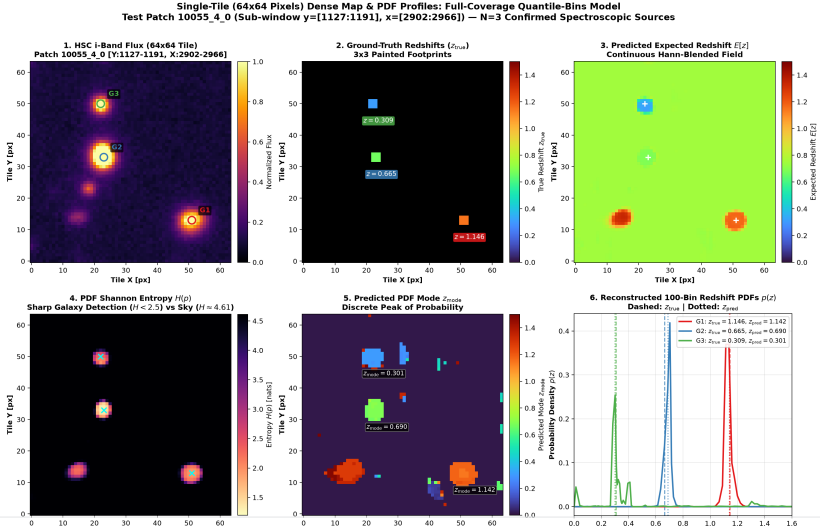
Loss function



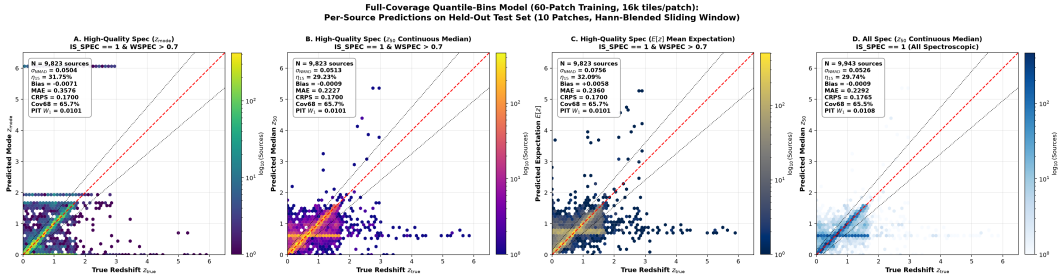
Resultados Full Patch



Resultados 1 Tile



Resultados


Full-Coverage Quantile-Bins Model (60-Patch Training, 16k tiles/patch):
Per-Source Predictions on Held-Out Test Set (10 Patches, Hann-Blended Sliding Window)
A. High-Quality Spec (z_{mode})
IS_SPEC == 1 & WSPEC > 0.7
N = 9,823 sources
σ_{NMAD} = 0.0504
σ_{25} = 31.75%
Bias = -0.0071
MAE = 0.3576
CRPS = 0.1700
CovE8 = 85.7%
PIT W_1 = 0.0101
B. High-Quality Spec (z_q, Continuous Median)
IS_SPEC == 1 & WSPEC > 0.7
N = 9,823 sources
σ_{NMAD} = 0.0513
η_{15} = 29.23%
Bias = -0.0009
MAE = 0.2227
CRPS = 0.1700
CovE8 = 85.7%
PIT W_1 = 0.0101
C. High-Quality Spec ($E[z]$ Mean Expectation)
IS_SPEC == 1 & WSPEC > 0.7
N = 9,823 sources
σ_{NMAD} = 0.0756
σ_{25} = 32.09%
Bias = +0.0058
MAE = 0.2360
CRPS = 0.1700
CovE8 = 85.7%
PIT W_1 = 0.0101
D. All Spec (z_q, Continuous Median)
IS_SPEC == 1 (All Spectroscopic)
N = 9,943 sources
σ_{NMAD} = 0.0526
η_{15} = 29.74%
Bias = -0.0009
MAE = 0.2292
CRPS = 0.1765
CovE8 = 85.5%
PIT W_1 = 0.0108
Predicted Mode z_{mode}
Predicted Median z_{50}
Predicted Expectation $E[z]$
True Redshift z_{true}
$\log_{10}$(Sources)

Calibración de las PDFs (60 patches, test)

	Cobertura 32 %		Cobertura 97 %
	Bins fijos	Cuantiles	Cuantiles
σ_{NMAD} (moda)	0,066	0,066	0,050
η_{15} (moda)	27,8 %	31,9 %	31,8 %
CRPS	0,723	0,181	0,170
Cov_{68} (ideal 68 %)	60,0 %	70,1 %	65,7 %
PIT W_1	0,139	0,019	0,010

9 837 fuentes espectroscópicas de alta calidad, por fuente. Con la mediana z_{50}, el modelo de cobertura completa da $\sigma_{\mathrm{NMAD}} = 0{,}051$, $\eta_{15} = 29{,}2\,\%$, MAE $= 0{,}223$.

- Los cuantiles calibran mucho mejor que los bins fijos: CRPS 4 veces menor.
- Con cobertura completa, la dispersión baja un 23 % y las PDFs se hacen más angostas sin perder calibración.
- Cov_{68} pasa de 70 % (algo subconfiado) a 66 % (apenas sobreconfiado).

La red puede detectar fuentes

Detección implícita

78 %
de las fuentes etiquetadas detectadas

98,6 %
de las detecciones coinciden con HSC

Magnitud i	Fuentes	Recall
< 24	3 565	89 %
24–25,5	3 822	91 %
$> 25{,}5$	964	56 %

La entropía marca la falta de información

En el 31 % de las fuentes de test la PDF es plana ($H > 4{,}3$, cielo: $\ln 100 = 4{,}61$).

	PDF plana (31 %)	Resto (69 %)
Clasificación (z_{50})		
η_{15}	70 %	10,5 %
σ_{NMAD}	—	0,025
Regresión		
η_{15}	56 %	10,2 %
σ_{NMAD}	—	0,031

Ahí caen 3 de cada 4 outliers en ambos modelos; sólo la clasificación lo señala.

Parte 5

Balance,Futuros pasos y conclusiones

Balance frente a los objetivos

	Objetivo	Estado
✓	Preparar e inspeccionar el conjunto de datos	Catálogo unificado, 1 530 mapas de etiquetas, control de calidad
✓	Componentes de línea base	U-Net de regresión y de clasificación (PDF de 100 bins por píxel)
✓	Contraste con referencias, con incertidumbre	$\sigma_{\mathrm{NMAD}} = 0{,}050$ y PDFs calibradas (PIT $W_1 = 0{,}010$), evaluación por fuente
●	Fuentes débiles, *blended* y de bajo brillo	Detección implícita con recall $\sim$90 % hasta $i \approx 24$; *blending* y bajo brillo, pendientes

Siguientes pasos

1. **Escalar** el entrenamiento a los $\sim$1 500 patches disponibles.
2. **Separar detección y redshift**: una salida $\hat{\rho}(x, y)$ para la presencia de fuentes y otra $p(z \mid x, y)$ para su redshift.
3. **Agregar aumentos de datos** (rotaciones, *flips*, *blur*).
4. **Incorporar la banda** u de CLAUDS como sexta entrada.
5. **Lograr** estimar valores de redshift mayores a 2 de forma adecuada.

Con estos pasos se cierra la primera etapa del proyecto:

Conclusiones

1. Se construyo un pipeline que toma datos del HSC-SSP y de redshift y devuelve un mapa denso con una PDF para cada pixel.

2. La U-Net de clasificación alcanza $\sigma_{\mathrm{NMAD}} = 0{,}050$ con PDFs calibradas (PIT $W_1 = 0{,}010$, $\mathrm{Cov}_{68} = 66\,\%$). Que son métricas aceptables.

3. La red detecta fuentes sin que se le pida y, con la entropía de su PDF, marca dónde no tiene información: algo que un estimador puntual no puede hacer.

Siguiente paso: separar detección y redshift, hacia el campo $\lambda(z, \alpha, \delta) = \rho(x, y)\, p(z \mid x, y)$.

Gracias

Preguntas

Juan Camilo Sánchez · Mitacs Globalink · University of Victoria · OAN

Respaldo: métricas

Con $\Delta = (\hat{z} - z)/(1+z)$:

$$\text{bias} = \text{med}(\Delta), \qquad \sigma_{\text{NMAD}} = 1{,}4826\ \text{med}\big|\Delta - \text{med}(\Delta)\big|, \qquad \eta_{15} = \Pr\big(|\Delta| > 0{,}15\big)$$

$$\text{CRPS} = \int \big(\hat{F}(z') - \mathbb{1}[z' \geq z]\big)^2\, dz', \qquad \text{PIT} = \hat{F}(z_{\text{true}})$$

- PIT uniforme $\Rightarrow$ PDFs calibradas; forma de U $\Rightarrow$ sobreconfianza; joroba central $\Rightarrow$ subconfianza.
- Cov_{68}: fracción de z_{true} dentro del intervalo central del 68 %.
- Local W_1: distancia del PIT a la uniforme en vecindades de $k = 100$ fuentes en el espacio de flujos *grizy*.

Respaldo: cobertura del muestreo

Muestreo uniforme de N tiles de 64×64 en un patch de 4100×4100:

$$\lambda = \frac{N \cdot 64^2}{4100^2}, \qquad \text{cobertura} \approx 1 - e^{-\lambda}$$

	N	λ	cobertura medida
Clasificación	1 600	0,39	32,4 %
Regresión	16 000	3,90	96,9 %

En 1 patch con split por tiles, $P(\text{píxel de val visto en train}) \approx 1 - e^{-0,8\lambda}$: casi todos. Por eso esos resultados no miden generalización.