

Sistema de Visión por Computadora para Reconocimiento de Plumas de Aves y Simulación de vision Tetrocromática inferencial a partir de imagenes RGB

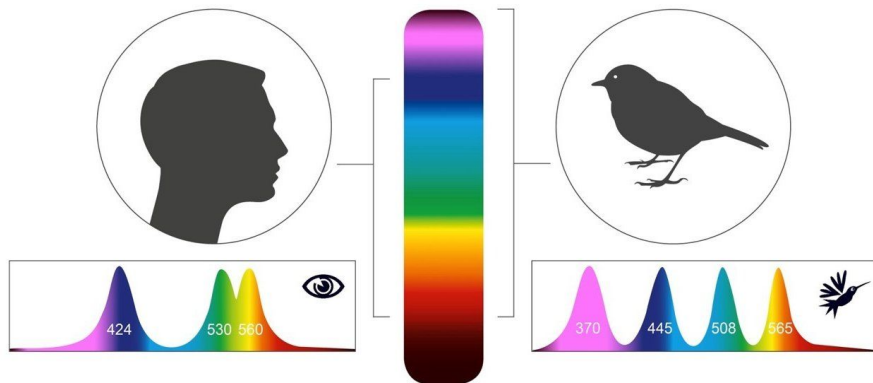
Juan Felipe Cardona

Motivación



Visión humana

Visión de aves



El objetivo principal es simular la percepción visual de las aves (cuatro canales: UV, R, G, B) y utilizar dicha información para:

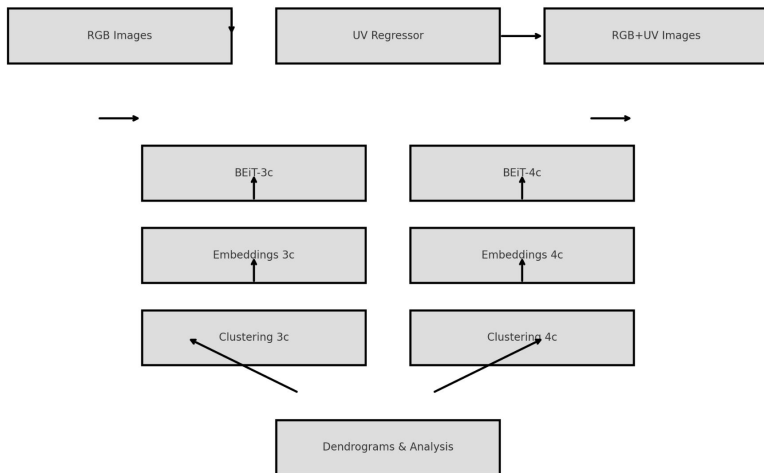
1. Reconocimiento y clasificación de especies.
2. Agrupación no supervisada basada en patrones de coloración y estrategias evolutivas.(BEIT ajustado)
3. El UVB añade información espectral que, en teoría, puede revelar patrones invisibles al ojo humano (por ejemplo, marcas de plumaje que solo se reflejan en el ultravioleta).

El proyecto se basa en dos fuentes de datos principales:

- LaSBiRD: Conjunto de imágenes de aves con bounding boxes y metadatos taxonómicos. rgb
- FeathersV1: Imágenes de plumas de aves en formato rgb
- BirdColorBase: base de datos de reflectancia de 300 a 700 nm de 25 especies de aves

PIPELINE DE TRABAJO

Pipeline de Trabajo: Comparación BEiT 3c vs BEiT 4c
y Análisis de Clusters/Patrones



A) Entrada: Imágenes RGB originales

B) Predicción de UVB: Paso por el modelo de regresión UV (dos fases previas: entrenamiento con datos espectrales, inferencia pixel-a-pixel)

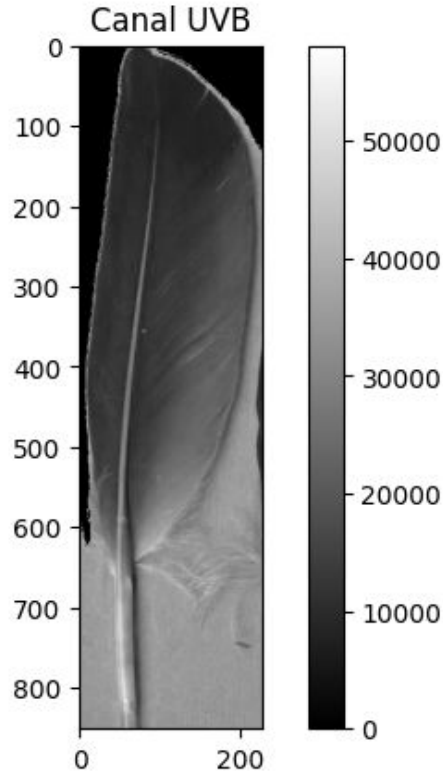
C) Generación de imágenes “RGB + UVB” (TIFF 4 canales)

D) Dos ramas de BEiT:

- **BEiT-3c** (input solo RGB) → Extracción de embeddings → Clustering 3-canal
- **BEiT-4c** (input RGB + UVB) → Extracción de embeddings → Clustering 4-canal

E) Análisis final: Dendrogramas y otras métricas para comparar los resultados de ambos enfoques y demostrar cómo el canal UVB enriquece los patrones y clusters.

Criterios de Exito



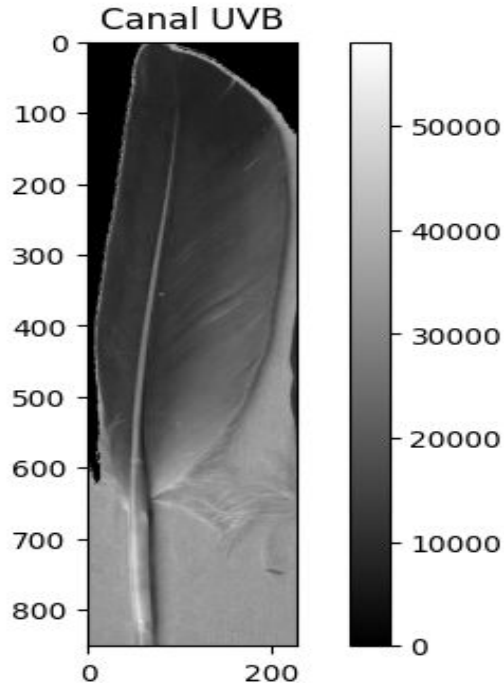
A) Modelo de regresión UVB robusto (RGB \rightarrow UVB) Objetivo mínimo: $R^2 \geq 0.75$ $R^2 \geq 0.75$ en el conjunto de test (datos independientes usados solo para evaluar).

B) Verificación de que las nuevas imágenes “sí contienen UVB” (enriquecimiento real)

C) Implementación de BEiT ajustado para 4 canales (RGB + UVB)

D) Identificación de nuevos patrones, clusters o “clusogramas” tetrocromáticos

PIPELINE DE TRABAJO: VERIFICAR UVB EN IMAGENES



A) Entrada: Imágenes RGB originales

B) Predicción de UVB

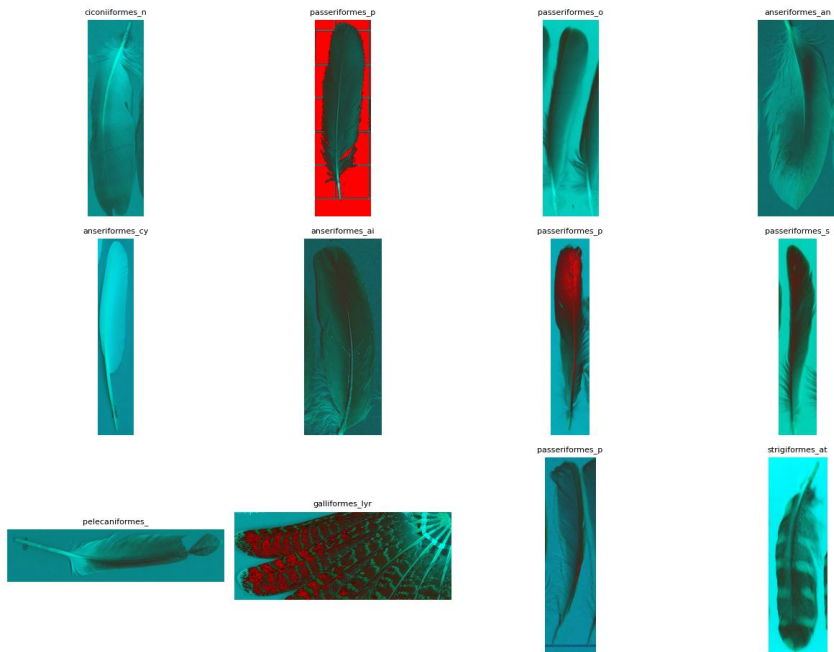
Se entrena un regresor (HistGradientBoosting con características polinómicas) usando datos espectrales (reflectancia 300–700 nm) BIRDCOLORBASE para aprender la relación entre (R, G, B) y la reflectancia ultravioleta (300–400 nm).

Validación resultó en $R^2 \approx 0.9193$ y $MAE \approx 6.4 \times 10^{-5}$, confirmando alta precisión en UVB.

C) RESULTADO: Para cada imagen RGB original, el modelo predice un valor UVB por pixel, usando las intensidades R, G y B del pixel.

Enriquecer imágenes originales (JPEG 3-canales) **no** contienen mediciones UVB reales

False-color (UV→R, G→G, B→B)



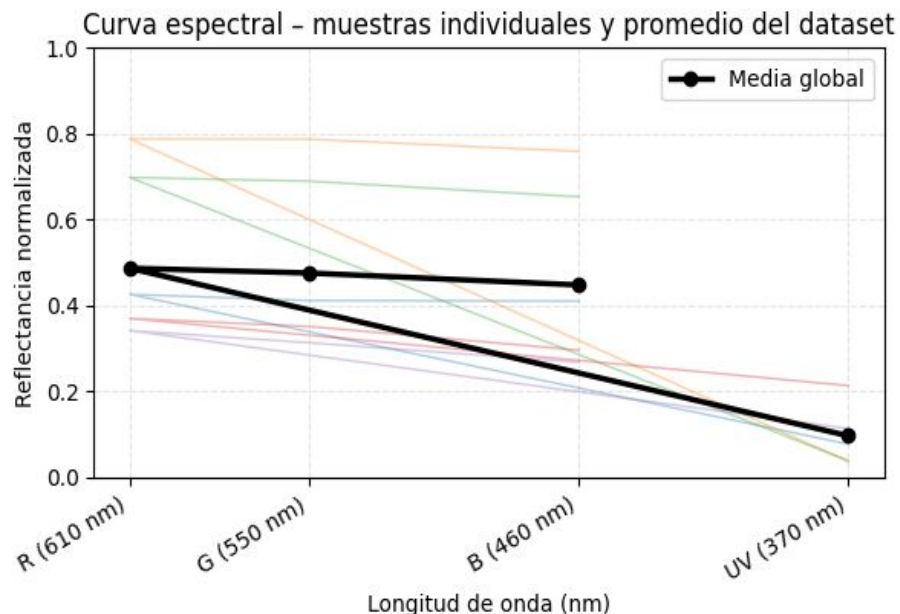
Originalmente las fotos no tenían UVB (eran sólo RGB).

El “canal UV” es una estimación generada por el modelo HGB entrenado con datos espectrales.

Gracias a esa estimación, las imágenes 4 canales sí quedan enriquecidas: ahora se puede ver, cuantificar y segmentar patrones UV que antes eran indetectables en la foto visible.

Este enriquecimiento es **útil para el nuevo EDA** porque permite explorar correlaciones “RGB vs. UV” y descubrir agrupamientos nuevos de plumajes que dependen del comportamiento en UV.

Curva espectral – muestras individuales y promedio del dataset



La curva promedio (línea negra): revela que el dataset, en su conjunto, refleja más en el rojo, luego decrece progresivamente hacia el azul y prácticamente no refleja en UV.

Las curvas individuales: muestran que existen plumas con reflectancias muy altas en R (0.7–0.8) y plumas muy oscuras en UV (cercanas a 0), pero algunos pocos ejemplares tienen un aumento “relativamente mayor” en UV (pero aun así bajo en términos absolutos).

Impacto del UV: aunque la reflectancia promedio en UV sea pequeña, las ligeras variaciones que aparecen en las líneas finas de ese segmento pueden ser diagnósticas para distinguir especies. Esa poca variabilidad puede ser crucial para un modelo de clasificación o un estudio biológico.

En definitiva, la gráfica te muestra que **el canal UV no es dominante** en términos de reflectancia (casi todas las plumas reflejan muy poca luz a 370 nm), pero **su ligera variación** ofrece información complementaria importante.

BEIT 4 CHANNELS VS BEIT BASE 3 CANALES

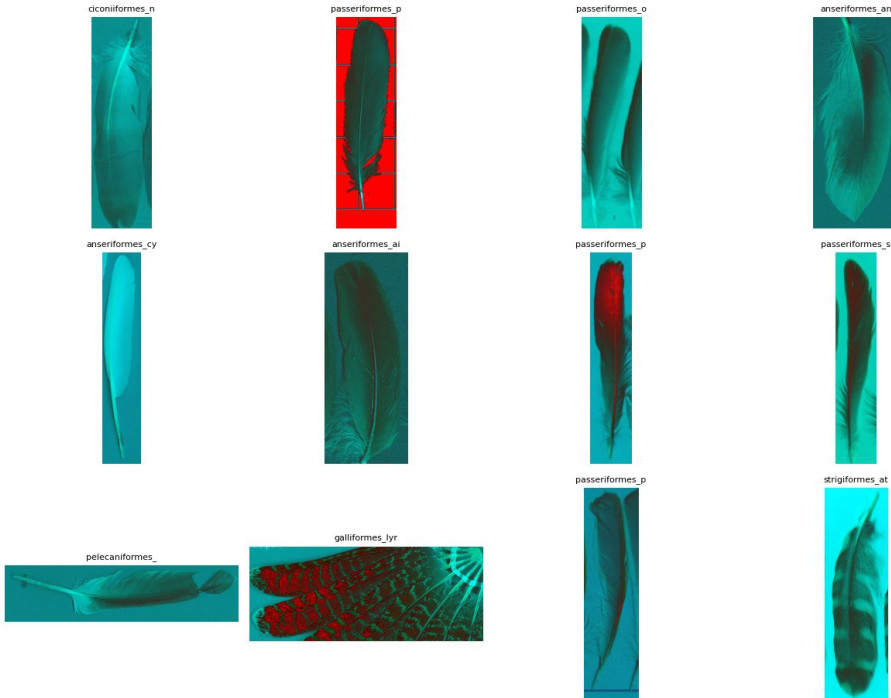
```
10
11 # 1. Cargar BEiT preentrenado en RGB
12 beit_model = timm.create_model('beit_base_patch16_224', pretrained=True)
13 beit_model.head = nn.Identity() # Usar como extractor de embeddings
14
15 # 2. Clonar para adaptar a 4 canales (RGB + UVB)
16 beit_4ch = copy.deepcopy(beit_model)
17
18 # 3. Modificar la capa de entrada Conv2D: pasar de 3 a 4 canales
19 old_proj = beit_4ch.patch_embed.proj
20 new_proj = nn.Conv2d(
21     in_channels=4,
22     out_channels=old_proj.out_channels,
23     kernel_size=old_proj.kernel_size,
24     stride=old_proj.stride,
25     padding=old_proj.padding,
26     bias=old_proj.bias is not None
27 )
28
29 # 4. Inicializar pesos:
30 with torch.no_grad():
31     new_proj.weight[:, :3] = old_proj.weight # copiar pesos preentrenados RGB
32     new_proj.weight[:, 3:] = old_proj.weight[:, :1] * 0.01 # inicializar UVB con ruido pequeño (mejor que ceros)
33     if old_proj.bias is not None:
34         new_proj.bias[:] = old_proj.bias
35
```

- Se obtiene el modelo base `beit_base_patch16_224` mediante la librería `timm` con sus pesos ya entrenados sobre imágenes RGB y se hace una copia
- **Crear una nueva capa de convolución** de parches con `in_channels=4`.
- **Copiar los pesos RGB preentrenados** a las tres primeras posiciones de los nuevos filtros, e **inicializar el canal UVB** con un valor pequeño heredado del filtro rojo.
- **Congelar** la mayoría de parámetros y **descongelar** solo las primeras capas (patch embeddings y dos bloques iniciales) para que se ajusten a la nueva dimensión del tensor de entrada.

Con este ajuste, el modelo conserva gran parte de su conocimiento previo en RGB y simultáneamente aprende a integrar la señal UVB

RETOS TECNICOS

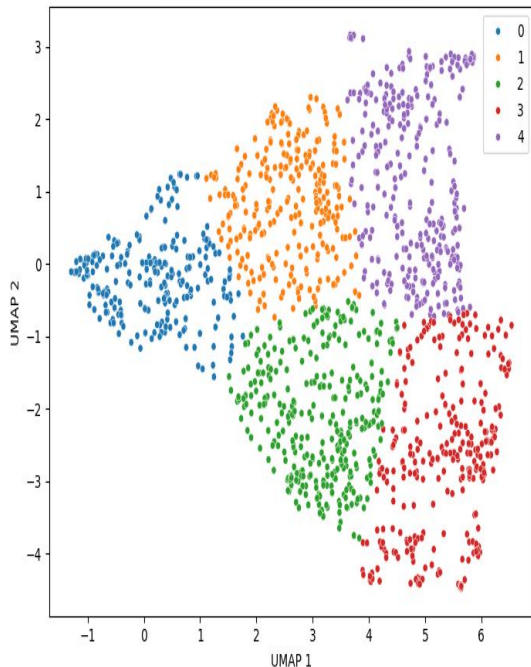
False-color (UV→R, G→G, B→B)



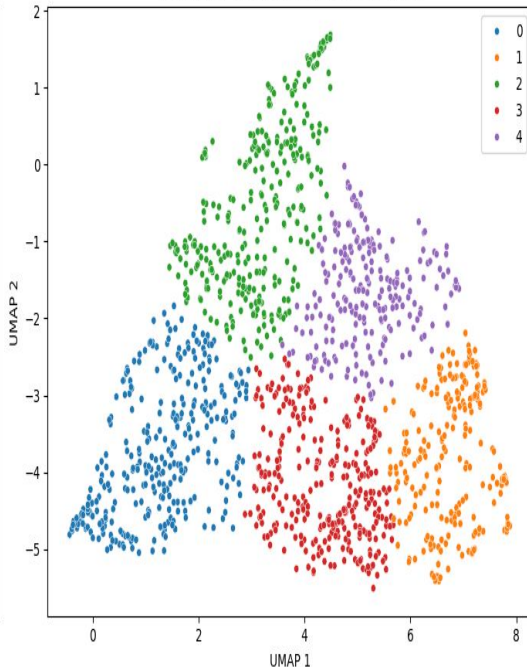
- Simular la percepción tetracromática de las aves añadiendo un canal ultravioleta (UV) proxy a imágenes originalmente capturadas en RGB.
- Segmentar automáticamente cada foto de plumas del fondo usando rembg un modelo de segmentación (normalmente U-2-Net o variantes) para detectar el sujeto en primer plano y generar un canal alfa que lo aísla del resto
- Normalizar y preprocesar las imágenes recortadas:
 - Estandarización (z-score) por canal (UV, R, G, B).
 - Se dejan las imágenes en 224x224 píxeles manteniendo los 4 canales
 - Se usa resize
- Extraer descriptores de color, textura y forma:
 - Se usa un BEiT Preentrenado para analizar las imágenes de 3 canales y también un BEiT adaptado para 4 canales
 - Se extraen los embeddings para rgb y rgb+uvb
 - Se compara y mide la diferencia entre embeddings
- Agrupar especies usando técnicas de clustering y reducción de dimensiones (PCA, UMAP + KMeans/HDBSCAN).
- Clasificar especies mediante métodos supervisados:

Métricas internas de calidad de clusters

UMAP con RGB (n_clusters=5)
Silhouette=0.39, DB=0.90, CH=1468.10



UMAP con RGB+UVB (n_clusters=5)
Silhouette=0.42, DB=0.81, CH=1500.28



	Silhouette ↑	Davies–Bouldin ↓	Calinski–Harabasz ↑
RGB	0.39	0.90	1468.10
RGB + UVB	0.42	0.81	1500.28

Los tres indicadores coinciden en que **el clustering con el canal UVB (es decir, RGB + UVB) es ligeramente mejor** que usando solo RGB.

Esto sugiere que incorporar la información espectral UVB aporta señales útiles para separar naturalmente las plumas en distintos grupos en el espacio reducido.

Tanto el Silhouette más alto, el Davies–Bouldin más bajo y el Calinski–Harabasz más alto en el gráfico derecho indican mayor cohesión interna y mejor separación de grupos

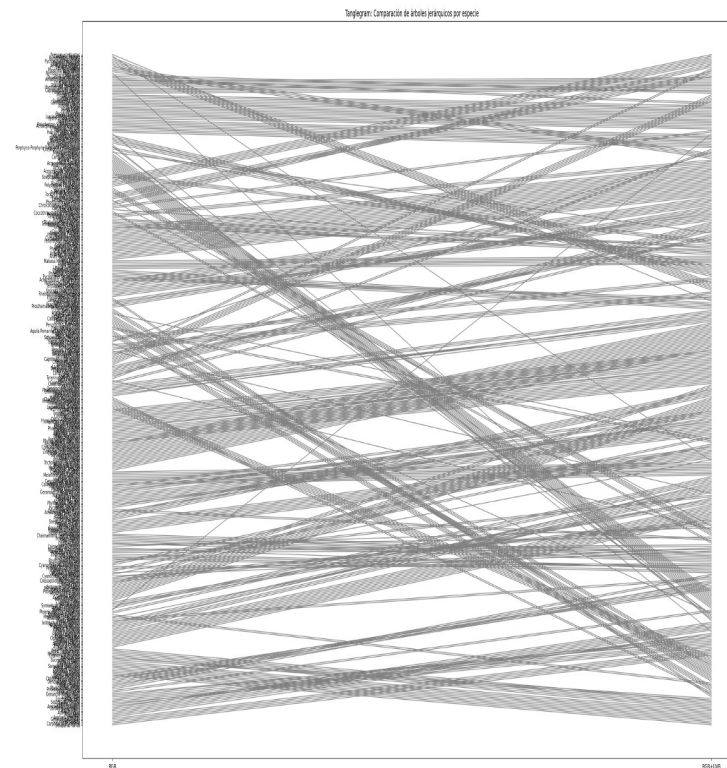
Métricas para comparar dendrogramas (estructuras jerárquicas)

Es una métrica que se utiliza para evaluar la calidad de un dendrograma obtenido mediante clustering jerárquico.

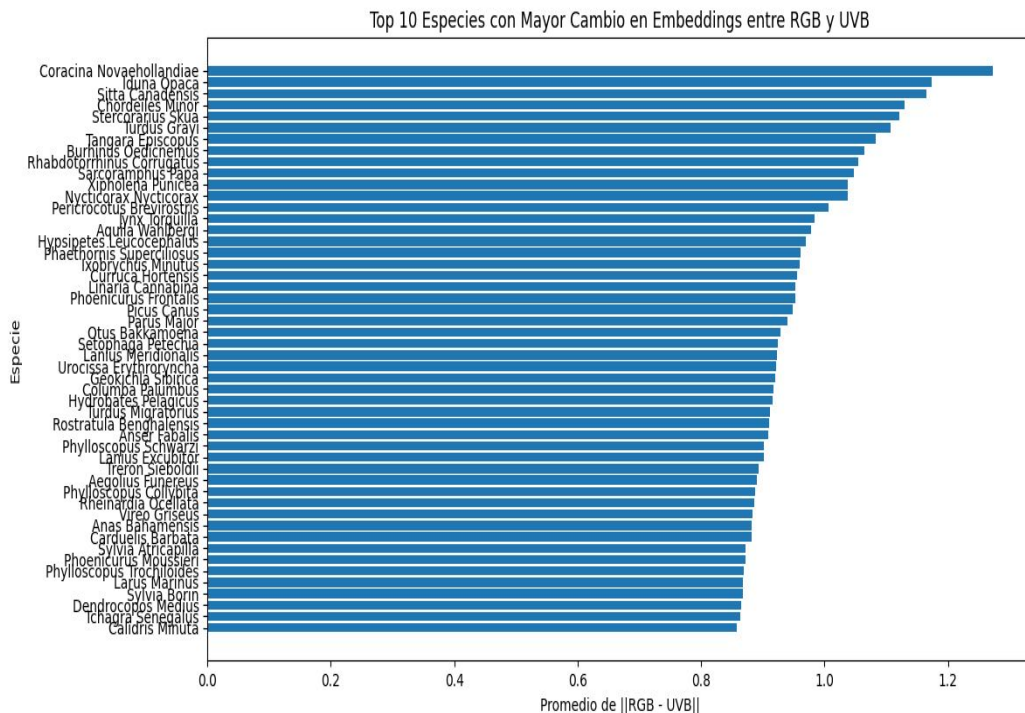
Cophenetic correlation RGB: 0.3501

Cophenetic correlation RGB+UVB:
0.2591

Al agregar el canal UVB, estás moviendo cada punto en el espacio de características (de 3 dimensiones RGB a 4 dimensiones RGB+UVB). Eso modifica las distancias entre pares de imágenes y, por ende, **el árbol jerárquico (el dendrograma) resultante será distinto.**



RETOS TECNICOS



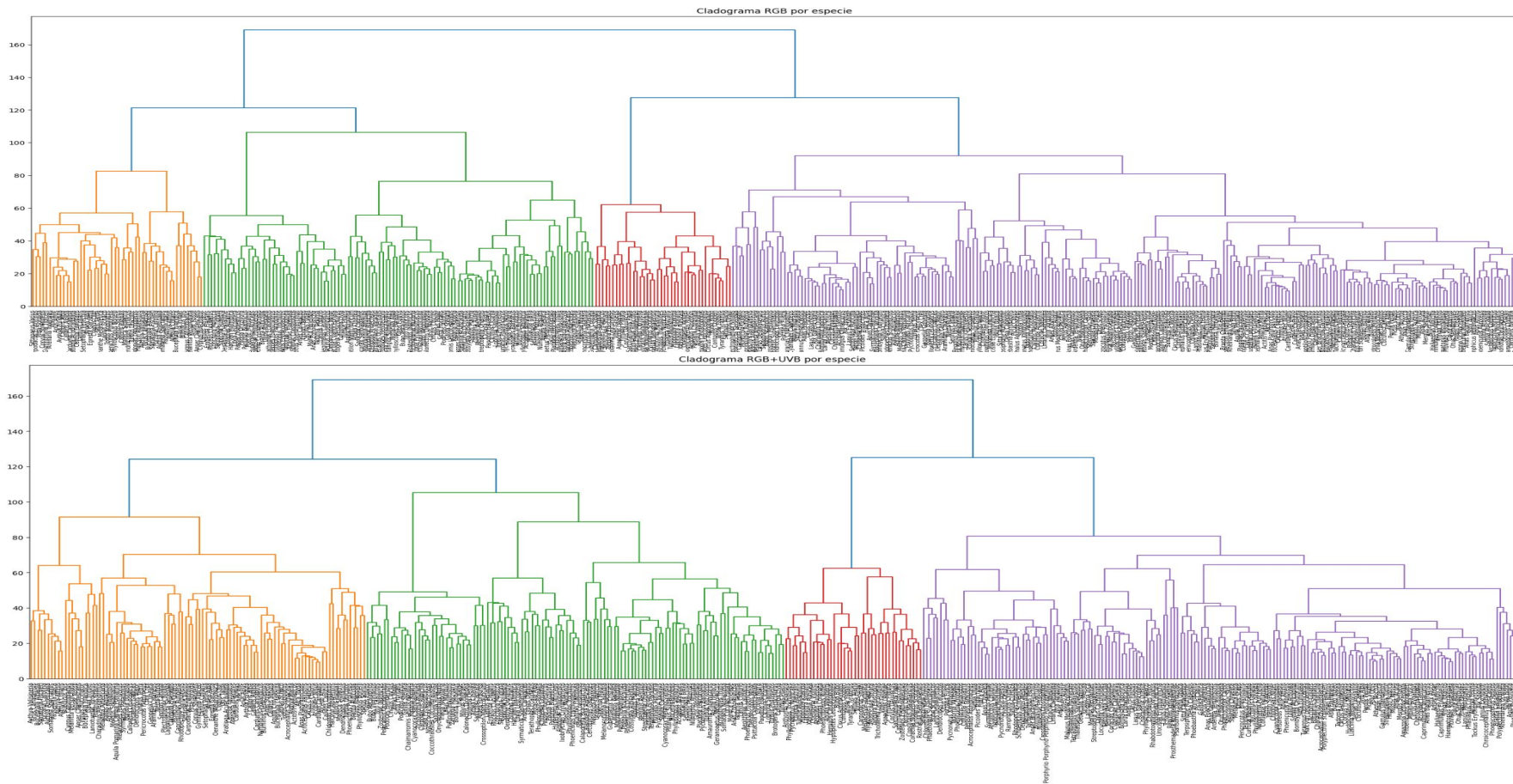
Contextualizar el objetivo final

- **Explorar posibles patrones espectrales adicionales** que no se ven en RGB (por ejemplo, diferencias de reflectancia UV que algunas especies o subespecies muestran), entonces puede tener sentido que el dendrograma cambie: quizá salga un agrupamiento más acorde a la biología.

“¿El UVB sugiere un dendrograma diferente?”

- Sí, absolutamente. Agregar UVB altera las distancias entre muestras y, por tanto, el orden y las alturas de fusión en el dendrograma cambian. Esto se ve reflejado en que la cophenetic correlation baja de 0.3501 (solo RGB) a 0.2591 (RGB+UVB).

CLADOGRAMA RGB VS RGB +UVB



RESULTADOS FINALES

Métrica / Espacio	Solo RGB	RGB + UVB	Comentario
Silhouette Score	0.39	0.42	Mejora de cohesión y separación
Davies–Bouldin Index	0.90	0.81	Disminuye → clusters más compactos
Calinski–Harabasz Index	1468	1500	Aumenta → mejor relación intra/inter
Adjusted Rand Index	—	0.32	Baja coincidencia de particiones
NMI	—	0.45	Baja mutua información entre grupos
Cophenetic Corr.	0.3501	0.2591	Menor fidelidad jerárquica en UVB

Conclusiones

El canal UVB genera un dendrograma distinto, La correlación cophenética desciende de **0.3501** (solo RGB) a **0.2591** (RGB + UVB), lo cual confirma que la estructura jerárquica resultante en el espacio enriquecido con UVB es diferente y menos fiel a las distancias originales de RGB.

UVB aporta señal espectral adicional, Incluir UVB permite capturar características espectrales que permanecen “invisibles” para el ojo humano, enriqueciendo el espacio de características con información biológicamente relevante.

Clusters más nítidos y cohesionados La métrica de Silhouette aumenta de **0.39** a **0.42** al pasar de RGB a RGB + UVB,

Validación a través de métricas internas y visualización, **La proyección UMAP muestra que los cinco grupos en RGB + UVB están más claramente diferenciados; existen menos zonas de solapamiento y los núcleos de cada cluster** se observan más definidos en comparación con el espacio solo RGB.

Confirmación de patrones ultravioleta como información discriminativa, Las mejoras simultáneas en Silhouette, Davies–Bouldin, Calinski–Harabasz y la separación visual en UMAP confirman que el canal UVB aporta datos relevantes que refuerzan la cohesión interna de cada cluster.

En conjunto, añadir el canal UVB enriquece el espacio de análisis, genera clusters más robustos y sugiere la existencia de patrones espectrales ultravioleta que complementan la información visible. Estos hallazgos corroboran que la inclusión de UVB es una vía eficaz para descubrir características “ocultas” en el plumaje de aves.