

# Hacia una Agricultura de Precisión en Chiloé.

Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares en Cultivos Hortícolas.



# “Hacia una Agricultura de Precisión en Chiloé: Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares en Cultivos Hortícolas”

Agricultura de Precisión

**Eduardo Opazo D. Guido Ríos C.**

Núcleo Data Science, Facultad de Ingeniería y Negocios Universidad de las Américas



# Introducción y Contexto

- **Contexto crítico en Chiloé:** alta humedad, lluvias frecuentes y variabilidad térmica favorecen enfermedades foliares en papa y tomate.
- **Impacto productivo:** pérdidas de hasta 40% y uso intensivo de agroquímicos.
- **Limitación actual:** detección tardía basada en inspección visual manual.
- **Propuesta:** sistema automatizado de **detección temprana** mediante **Deep Learning + IoT**.
- **Beneficios:** alertas predictivas, mejor toma de decisiones, **reducción de agroquímicos** y mayor **productividad y sostenibilidad**.



**Enfermedad foliar en papa - Isla de Quenac, Chiloé, Chile**

# Problema y Motivación

- **Problema crítico:** las enfermedades foliares reducen la productividad y afectan la sostenibilidad económica y ambiental de pequeños y medianos productores.
- **Limitación actual:** detección tardía basada en inspección visual y decisiones reactivas.
- **Solución tecnológica:** Visión Computacional + Deep Learning + IoT permiten detección temprana y monitoreo continuo, incluso antes de la percepción humana.

# Problema y Motivación

- **Impacto esperado:** mejor toma de decisiones agronómicas, mayor rendimiento, menor uso de insumos químicos.
- **Proyección:** impulso a la **Agricultura de Precisión 4.0** adaptada a las condiciones de Chiloé, fortaleciendo la sostenibilidad productiva y ambiental.



**Detección Temprana de Enfermedades Foliares mediante Drones en Chiloé**

Mapa de detección de enfermedad foliar

# Metodología Propuesta

- **Sistema de Agricultura de Precisión que integra Visión Computacional, Deep Learning e IoT para la detección temprana de enfermedades foliares.**
- **Monitoreo automatizado y continuo, mediante sensores microclimáticos y cámaras conectadas por LoRaWAN/4G, entrenado con datos locales de Chiloé.**
- **Alertas tempranas y visualización simple en plataformas web y móviles, apoyando decisiones fitosanitarias oportunas.**

# Metodología Propuesta

- **Impacto:** reducción de pérdidas productivas, optimización del uso de agroquímicos y beneficio directo a pequeños agricultores.
- **Proyección:** desarrollo de un **prototipo validado**, con potencial de **escalamiento y transferencia** al sector productivo regional.

# Esquema Conceptual de la Solución



# Esquema Conceptual de la Solución

- **Componentes tecnológicos protegibles**, que aseguran **diferenciación** y **ventaja competitiva** del sistema.
- **Uso de tecnologías open-source validadas**, garantizando **flexibilidad**, **seguridad jurídica** y viabilidad de **escalamiento comercial**.
- **Estrategia de PI y licenciamiento**: desarrollo propio, hardware abierto y **registro de software**, evitando restricciones de uso.
- **Impacto en transferencia**: facilita la **validación del prototipo** y su **adopción por el sector productivo**.



Trips occidentalis



Mosca blanca



Pulgones

Plagas de insectos afectando el cultivo de papa en Chiloé



# Uso de Imágenes Multiespectrales

- **Limitación actual:** la inspección visual detecta enfermedades en etapas tardías y con alta subjetividad.
- **Ventaja multiespectral:** el estrés vegetal altera la **reflectancia espectral** antes de ser visible al ojo humano, especialmente relevante en el clima de Chiloé.
- **Datos clave:** uso de **RGB, Red Edge y NIR**, junto a índices **NDVI, GNDVI** y **NDRE**, para detectar cambios tempranos en vigor y clorofila.

# Uso de Imágenes Multiespectrales

- **Integración inteligente:** imágenes + variables ambientales mediante IA e IoT, generando **mapas de salud del cultivo**.
- **Impacto:** detección temprana, decisiones agrícolas más precisas y **Agricultura de Precisión sostenible** en Chiloé.

# Hacia una Agricultura de Precisión en Chiloé

Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares en Cultivos Hortícolas



# Enfoque de Aprendizaje Supervisado

- **Problema de clasificación supervisada:** identificación automática del estado sanitario de hojas hortícolas en Chiloé.
- **Datos de entrada:** imágenes foliares y **variables ambientales IoT**, capturadas en condiciones reales de campo.
- **Modelos Deep Learning:** **redes neuronales convolucionales (CNN)** entrenadas para detectar enfermedades específicas.
- **Evaluación y extensión:** métricas estándar de clasificación y  **posible extensión a regresión de severidad.**
- **Impacto:** apoyo a **decisiones agronómicas tempranas y transferencia tecnológica.**

# Enfoque de Aprendizaje Supervisado

- Por tanto, el proyecto “**Hacia una Agricultura de Precisión en Chiloé: Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares en Cultivos Hortícolas**”
- Se va a modelar como un **problema de aprendizaje supervisado de clasificación**, con posibilidad de extensión a **regresión para estimar severidad**.

# Hacia una Agricultura de Precisión en Chiloé

Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares  
en Cultivos Hortícolas



## Tomadores de Decisión Beneficiados



• Productores Hortícolas



• Agrónomos / Fitopatólogos



• Cooperativas y Asociaciones



• Instituciones de Desarrollo  
y Gobierno

## Aporte del Modelo Propuesto



• Detección Temprana  
de Enfermedades



• Optimización  
de Rendimientos



• Reducción del Uso  
de Agroquímicos



• Toma de Decisiones  
Basada en Datos

Mayor Productividad y Sostenibilidad en la Agricultura de Chiloé

# Apoyo a la Toma de Decisiones

- **Publico Objetivo**

- Pequeños y medianos productores de la provincia de Chiloé.
- Alta dependencia de condiciones climáticas.
- Acceso limitado a asesoría técnica especializada en sanidad vegetal.

# Apoyo a la Toma de Decisiones

- **Aporte del Modelo Propuesto**
  - Detección temprana de enfermedades foliares mediante clasificación supervisada.
  - Generación de **alertas objetivas y oportunas** sobre el estado sanitario del cultivo.
  - Soporte directo a decisiones agronómicas, tales como:
    - ✓ Aplicación selectiva de tratamientos fitosanitarios.
    - ✓ Priorización de parcelas o sectores críticos.
    - ✓ Reducción de pérdidas productivas y uso innecesario de insumos.

# Apoyo a la Toma de Decisiones

- **Valor Agregado**
  - Mejora la eficiencia productiva.
  - Reduce costos operacionales.
  - Promueve prácticas agrícolas más sostenibles.

# Justificación del Desafío Computacional

Hacia una Agricultura de Precisión en Chiloé: Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares en Cultivos

## Justificación del Desafío Computacional



### Variabilidad Visual

- Diversidad de síntomas, colores y severidades de enfermedades foliares



### Contexto Ambiental

- Análisis y correlación de factores como humedad, temperatura, radiación y precipitación



### Condiciones Cambiantes

- Algoritmo robusto a variabilidad estacional y cambios climáticos



### Escasez de Datos Etiquetados

- Limitada disponibilidad de imágenes etiquetadas
- Promueve prácticas agrícolas más sostenibles

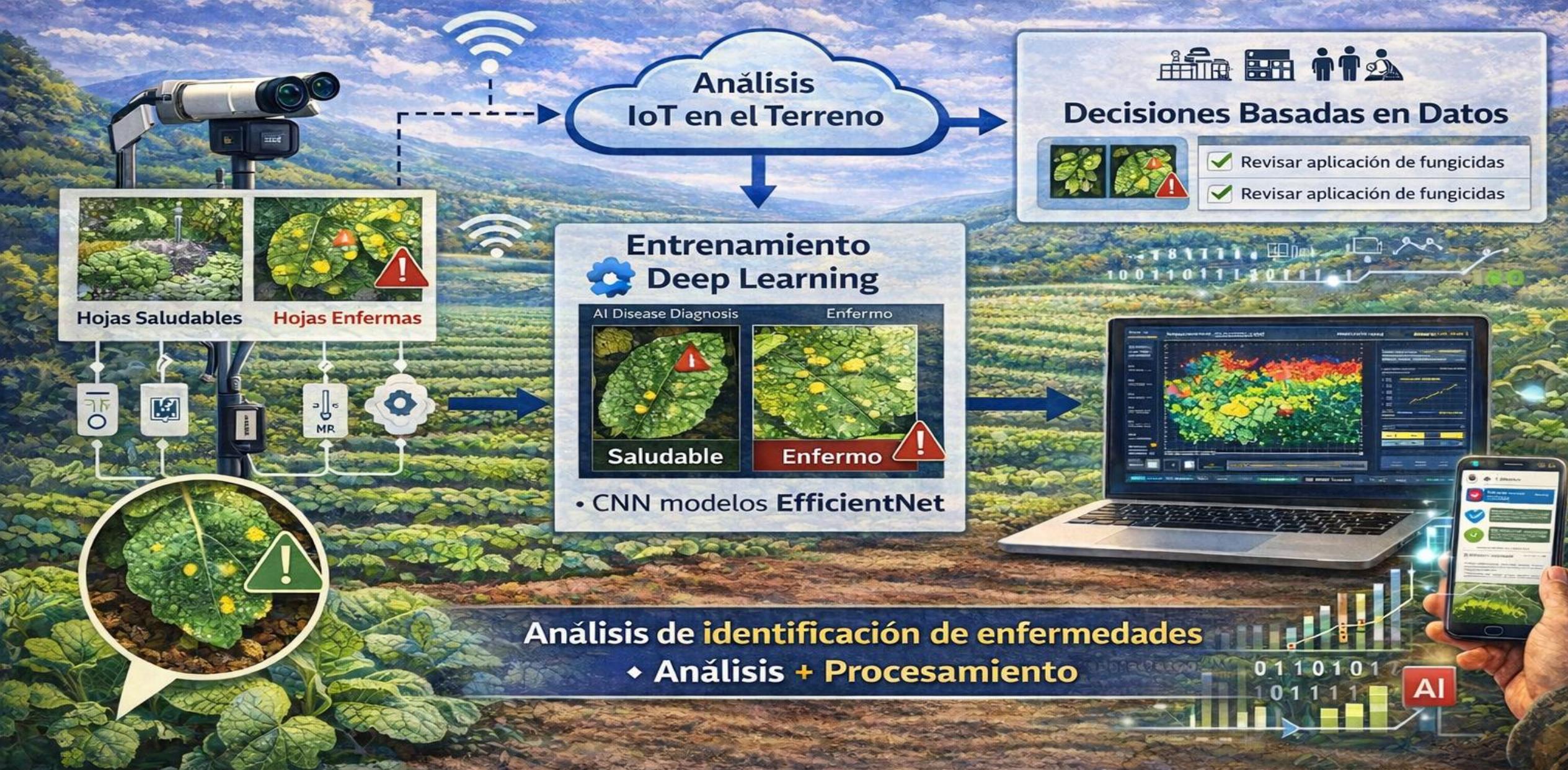
# Justificación del Problema

- **Complejidad del Entorno Agrícola**
  - Las enfermedades foliares no se ven siempre igual.
  - Las imágenes cambian según:
    - ✓ Luz.
    - ✓ Clima.
    - ✓ Etapa de crecimiento de la planta.
- **Limitaciones de Métodos Tradicionales**
  - No es posible usar reglas simples o umbrales fijos.
  - Las decisiones manuales son lentas y subjetivas.
  - Alta incertidumbre en condiciones reales de campo.

# Justificación del Problema

- **Rol del Modelo Propuesto**
  - Aprende de forma directa desde los datos reales.
  - Integra imágenes y variables climáticas.
  - Funciona de forma robusta ante ruido y variabilidad.
- **Resultado Esperado**
  - Detección temprana y confiable de enfermedades.
  - Apoyo efectivo a la toma de decisiones agrícolas.
  - Base tecnológica para la agricultura de precisión.

# Análisis y Procesamiento



# Pipeline de Preparación de Datos



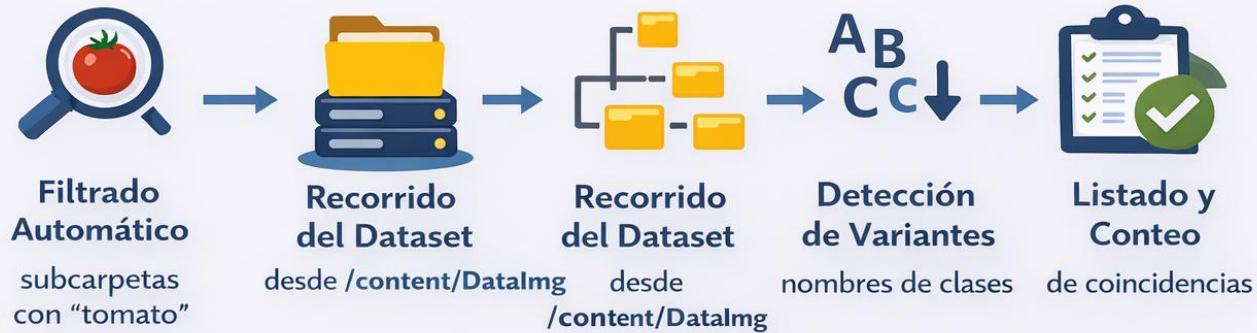
# Validación del Dataset



# Búsqueda de Clases “Tomato”

## Búsqueda de Clases “Tomato” en el Dataset

### Búsqueda de Clases “Tomato” en el Dataset

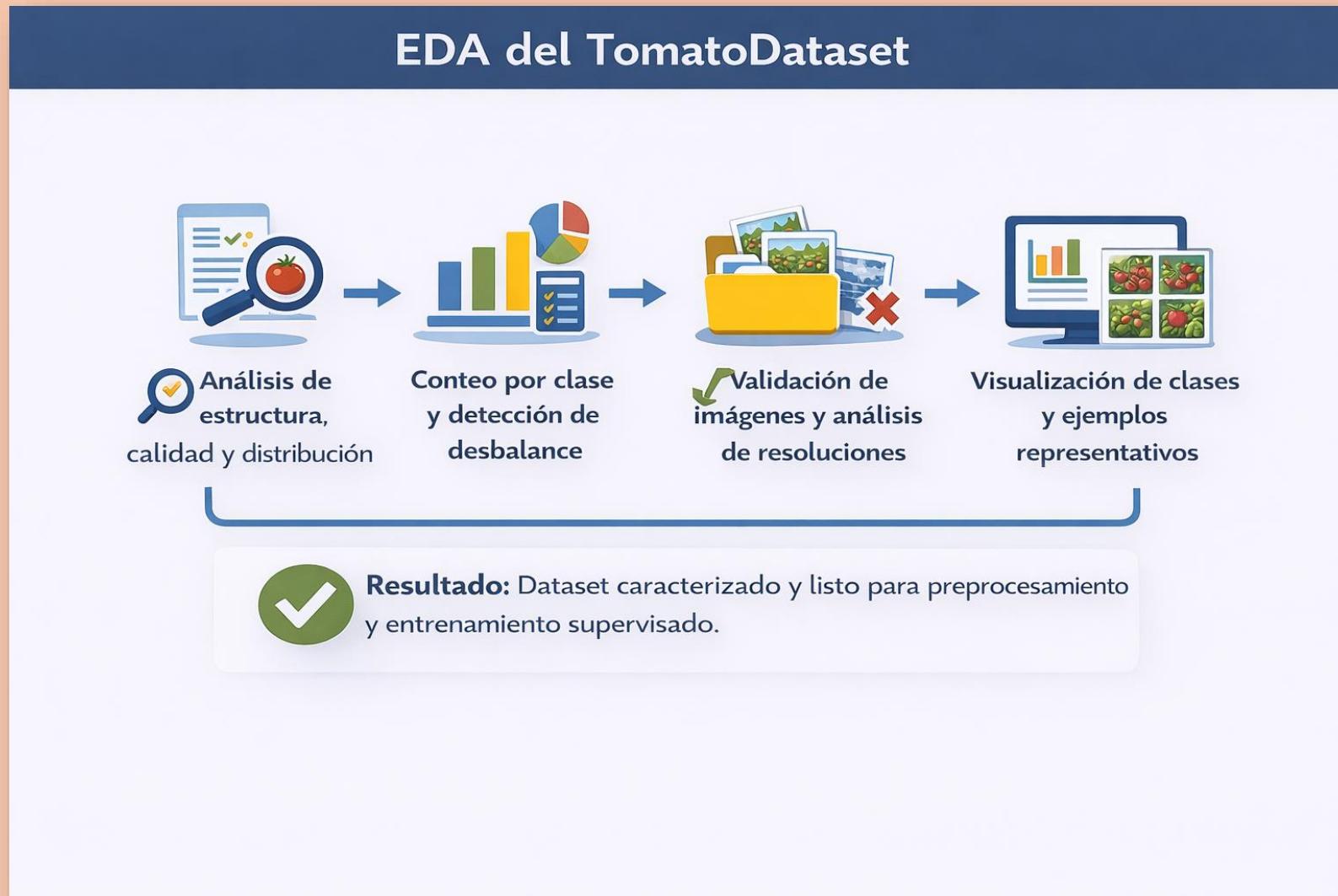


**Resultado:** Validación rápida de **clases de tomate** antes del entrenamiento supervisado.

# Construcción del Dataset de Tomate



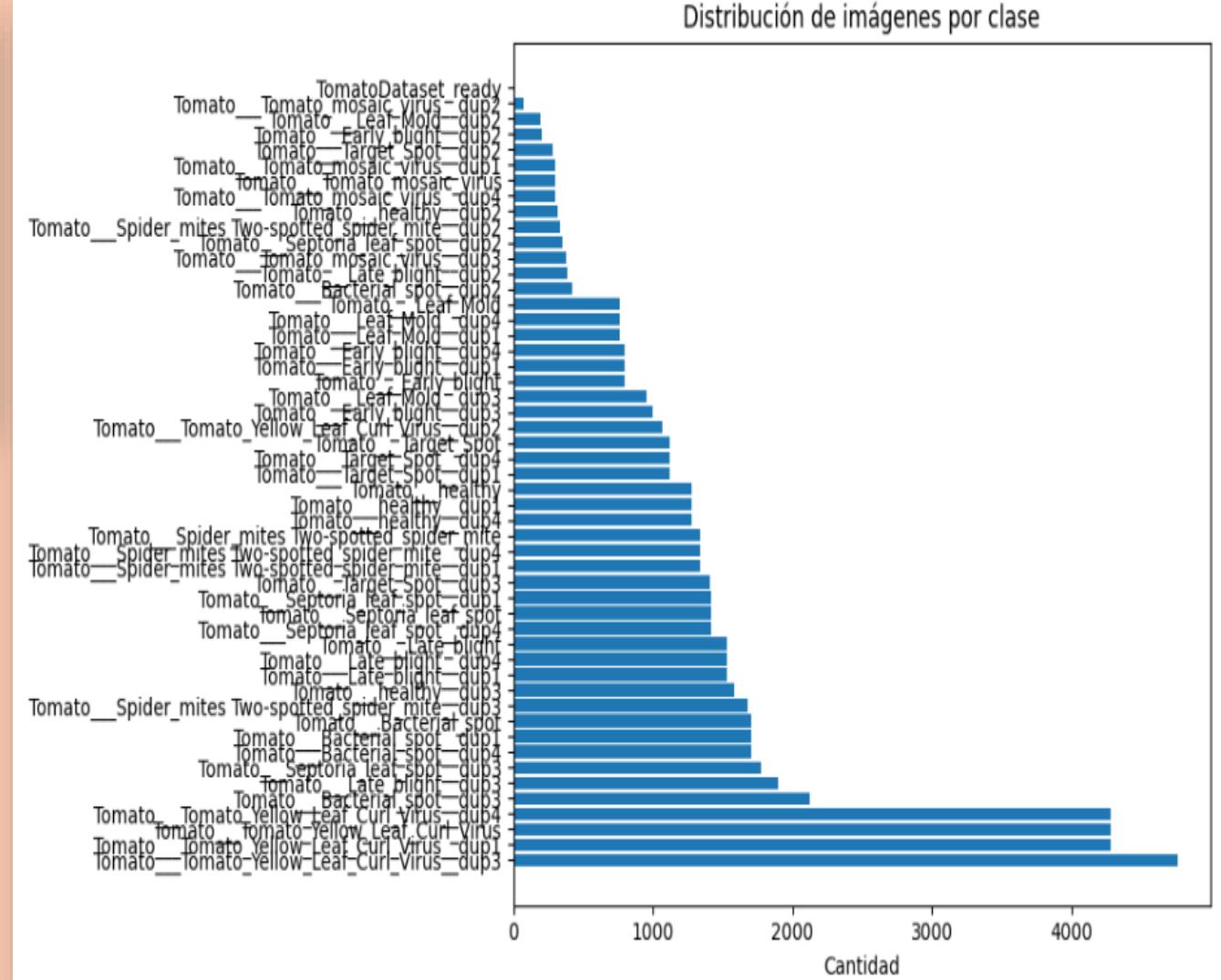
# EDA del TomatoDataset



# Visualizaciones

Resumen por clase:		
	Clase	Cantidad_imagenes
39	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup3	4761
37	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup1	4286
36	Tomato__Tomato_Yellow_Leaf_Curl_Virus	4286
40	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup4	4286
4	Tomato__Bacterial_spot_dup3	2126
14	Tomato__Late_blight_dup3	1897
24	Tomato__Septoria_leaf_spot_dup3	1771
5	Tomato__Bacterial_spot_dup4	1702
2	Tomato__Bacterial_spot_dup1	1702
1	Tomato__Bacterial_spot	1702
29	Tomato__Spider_mites Two-spotted_spider_mite_...	1676
49	Tomato__healthy_dup3	1585
12	Tomato__Late_blight_dup1	1527
15	Tomato__Late_blight_dup4	1527
11	Tomato__Late_blight	1527
25	Tomato__Septoria_leaf_spot_dup4	1417
21	Tomato__Septoria_leaf_spot	1417
22	Tomato__Septoria_leaf_spot_dup1	1417
34	Tomato__Target_Spot_dup3	1404
27	Tomato__Spider_mites Two-spotted_spider_mite_...	1341
30	Tomato__Spider_mites Two-spotted_spider_mite_...	1341
26	Tomato__Spider_mites Two-spotted_spider_mite	1341

Estadísticas de tamaño de imágenes:		
	Ancho	Alto
count	64761.0	64761.0
mean	256.0	256.0
std	0.0	0.0
min	256.0	256.0
25%	256.0	256.0
50%	256.0	256.0
75%	256.0	256.0
max	256.0	256.0



# Visualizaciones

Tomato\_\_Bacterial\_spot\_dup3



Tomato\_\_Late\_blight\_dup3



Tomato\_\_Septoria\_leaf\_spot\_d



Tomato\_\_Bacterial\_spot\_dup4



Tomato\_\_Bacterial\_spot\_dup1



Tomato\_\_Bacterial\_spot



Tomato\_\_Spider\_mites\_Two-spot



Tomato\_\_healthy\_dup3



Conclusiones automáticas:

- Total de clases: 51
- Total de imágenes: 64761
- Resolución promedio aproximada: 256 x 256
- Clases con menos imágenes:

	Clase	Cantidad_imagenes
18	Tomato__Leaf_Mold_dup2	191
43	Tomato__Tomato_mosaic_virus_dup2	74
0	TomatoDataset_ready	0

# Auditoria de Clases Duplicadas



# Visualización

Reporte Final de Duplicados:										
	Clase	Referencia	Comparado	Img_Ref	Img_Comp	En_comun	Solo_Ref	Solo_Comp	Hash_Iguales	Hash_Distintos
0	Tomato__Early_blight	Tomato__Early_blight_dup3	Tomato__Early_blight_dup4	998	800	798	200	2	200	0
1	Tomato__Early_blight	Tomato__Early_blight_dup3	Tomato__Early_blight	998	800	798	200	2	200	0
2	Tomato__Early_blight	Tomato__Early_blight_dup3	Tomato__Early_blight_dup2	998	200	200	798	0	200	0
3	Tomato__Early_blight	Tomato__Early_blight_dup3	Tomato__Early_blight_dup1	998	800	798	200	2	200	0
4	Tomato__Spider_mites Two-spotted_spider_mite	Tomato__Spider_mites Two-spotted_spider_mite_...	Tomato__Spider_mites Two-spotted_spider_mite	1676	1341	1341	335	0	200	0
5	Tomato__Spider_mites Two-spotted_spider_mite	Tomato__Spider_mites Two-spotted_spider_mite_...	Tomato__Spider_mites Two-spotted_spider_mite_...	1676	335	335	1341	0	200	0
6	Tomato__Spider_mites Two-spotted_spider_mite	Tomato__Spider_mites Two-spotted_spider_mite_...	Tomato__Spider_mites Two-spotted_spider_mite_...	1676	1341	1341	335	0	200	0
7	Tomato__Spider_mites Two-spotted_spider_mite	Tomato__Spider_mites Two-spotted_spider_mite_...	Tomato__Spider_mites Two-spotted_spider_mite_...	1676	1341	1341	335	0	200	0
8	Tomato__Tomato_Yellow_Leaf_Curl_Virus	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup3	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup4	4761	4286	3821	940	465	200	0
9	Tomato__Tomato_Yellow_Leaf_Curl_Virus	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup3	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup2	4761	1071	940	3821	131	200	0
10	Tomato__Tomato_Yellow_Leaf_Curl_Virus	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup3	Tomato__Tomato_Yellow_Leaf_Curl_Virus	4761	4286	3821	940	465	200	0
11	Tomato__Tomato_Yellow_Leaf_Curl_Virus	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup3	Tomato__Tomato_Yellow_Leaf_Curl_Virus_dup1	4761	4286	3821	940	465	200	0
12	Tomato__Tomato_mosaic_virus	Tomato__Tomato_mosaic_virus_dup2	Tomato__Tomato_mosaic_virus_dup1	74	299	0	74	299	0	0
13	Tomato__Tomato_mosaic_virus	Tomato__Tomato_mosaic_virus_dup2	Tomato__Tomato_mosaic_virus_dup4	74	299	0	74	299	0	0
14	Tomato__Tomato_mosaic_virus	Tomato__Tomato_mosaic_virus_dup2	Tomato__Tomato_mosaic_virus_dup3	74	373	74	0	299	74	0
15	Tomato__Tomato_mosaic_virus	Tomato__Tomato_mosaic_virus_dup2	Tomato__Tomato_mosaic_virus	74	299	0	74	299	0	0
16	Tomato__Bacterial_spot	Tomato__Bacterial_spot_dup4	Tomato__Bacterial_spot_dup1	1702	1702	1702	0	0	200	0
17	Tomato__Bacterial_spot	Tomato__Bacterial_spot_dup4	Tomato__Bacterial_spot	1702	1702	1702	0	0	200	0

# Fusión y Depuración del Dataset



# Visualización

```
Fuente: /content/TomatoDataset
Destino: /content/TomatoDataset_merged

Clase: TomatoDataset_ready
Orígenes: ['TomatoDataset_ready']
Copiados: 0
Omitidos por duplicado (hash): 0
Renombrados por colisión de nombre: 0

Clase: Tomato__Bacterial_spot
Orígenes: ['Tomato__Bacterial_spot', 'Tomato__Bacterial_spot_dup1', 'Tomato__Bacterial_spot_dup2', 'Tomato__Bacterial_spot_dup3', 'Tomato__Bacterial_spot_dup4']
Copiados: 2127
Omitidos por duplicado (hash): 5530
Renombrados por colisión de nombre: 0

Clase: Tomato__Early_blight
Orígenes: ['Tomato__Early_blight', 'Tomato__Early_blight_dup1', 'Tomato__Early_blight_dup2', 'Tomato__Early_blight_dup3', 'Tomato__Early_blight_dup4']
Copiados: 1000
Omitidos por duplicado (hash): 2598
Renombrados por colisión de nombre: 0

Clase: Tomato__Late_blight
Orígenes: ['Tomato__Late_blight', 'Tomato__Late_blight_dup1', 'Tomato__Late_blight_dup2', 'Tomato__Late_blight_dup3', 'Tomato__Late_blight_dup4']
Copiados: 1901
Omitidos por duplicado (hash): 4959
Renombrados por colisión de nombre: 0

Clase: Tomato__Leaf_Mold
Orígenes: ['Tomato__Leaf_Mold', 'Tomato__Leaf_Mold_dup1', 'Tomato__Leaf_Mold_dup2', 'Tomato__Leaf_Mold_dup3', 'Tomato__Leaf_Mold_dup4']
Copiados: 952
Omitidos por duplicado (hash): 2474
Renombrados por colisión de nombre: 0

Clase: Tomato__Septoria_leaf_spot
Orígenes: ['Tomato__Septoria_leaf_spot', 'Tomato__Septoria_leaf_spot_dup1', 'Tomato__Septoria_leaf_spot_dup2', 'Tomato__Septoria_leaf_spot_dup3', 'Tomato__Septoria_leaf_spot_dup4']
Copiados: 1771
Omitidos por duplicado (hash): 4605
Renombrados por colisión de nombre: 0

Clase: Tomato__Spider_mites Two-spotted_spider_mite
Orígenes: ['Tomato__Spider_mites Two-spotted_spider_mite', 'Tomato__Spider_mites Two-spotted_spider_mite_dup1', 'Tomato__Spider_mites Two-spotted_spider_mite_dup2', 'Tomato__Spider_mites Two-spotted_spider_mite_dup3', 'Tomato__Spider_mites Two-spotted_spider_mite_dup4']
Copiados: 1676
Omitidos por duplicado (hash): 4358
Renombrados por colisión de nombre: 0
```

# EDA Avanzado TomatoDataset



# Visualización

## RESUMEN GENERAL

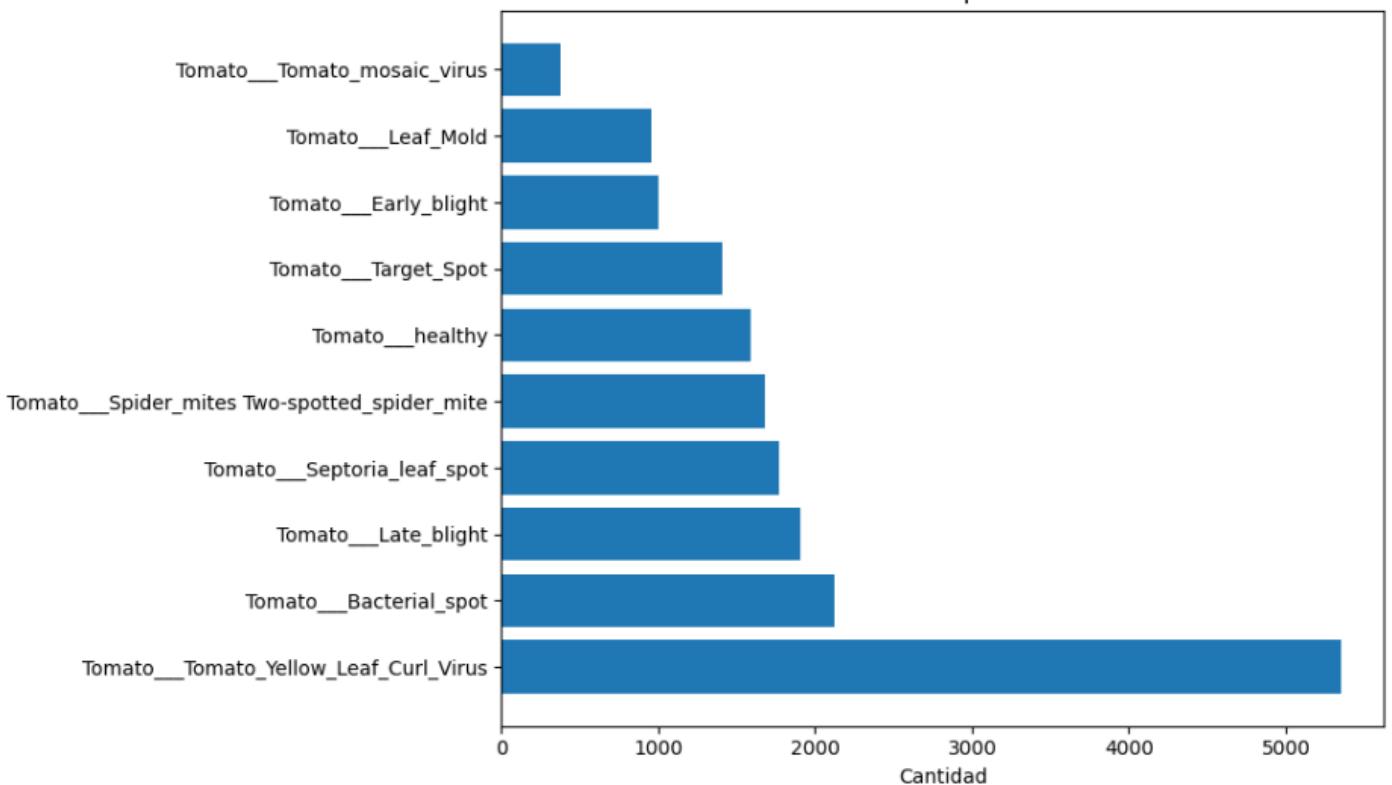
Total imágenes: 18146

Total clases: 10

Corruptas: 0

	Cantidad	Porcentaje
Clase		
Tomato__Tomato_Yellow_Leaf_Curl_Virus	5357	29.52
Tomato__Bacterial_spot	2127	11.72
Tomato__Late_blight	1901	10.48
Tomato__Septoria_leaf_spot	1771	9.76
Tomato__Spider_mites Two-spotted_spider_mite	1676	9.24
Tomato__healthy	1585	8.73
Tomato__Target_Spot	1404	7.74
Tomato__Early_blight	1000	5.51
Tomato__Leaf_Mold	952	5.25
Tomato__Tomato_mosaic_virus	373	2.06

Distribución por clase



# Visualización

ESTADÍSTICAS GLOBALES							
	Ancho	Alto	Aspect	Brillo	Contraste	Entropia	Blur
count	18146.0	18146.0	18146.0	18146.000000	18146.000000	18146.000000	18146.000000
mean	256.0	256.0	1.0	115.607121	38.163846	6.998742	4095.634562
std	0.0	0.0	0.0	16.330085	10.695764	0.517583	4434.449148
min	256.0	256.0	1.0	8.246216	10.629742	1.352320	5.019577
25%	256.0	256.0	1.0	106.267632	29.963157	6.764452	716.729044
50%	256.0	256.0	1.0	115.971916	38.307126	7.076588	2740.308527
75%	256.0	256.0	1.0	124.544544	46.177556	7.352979	6145.844635
max	256.0	256.0	1.0	194.555527	80.729056	7.791131	27109.052995

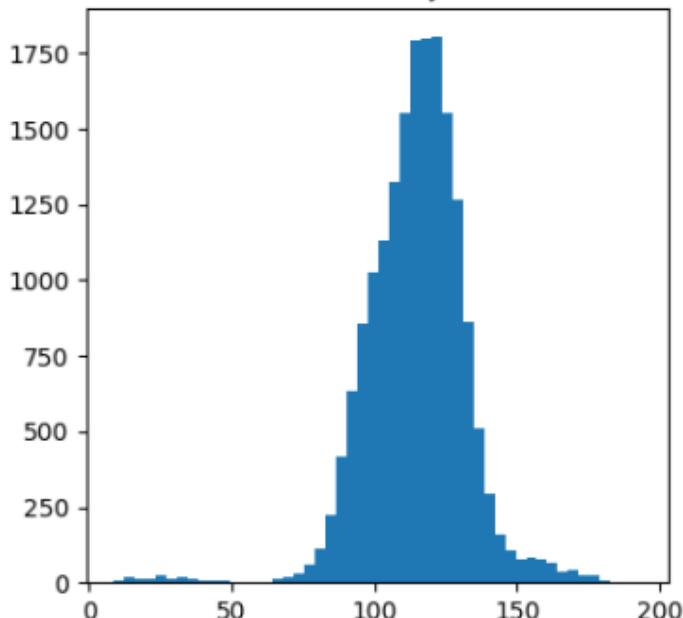
	Clase	Brillo	Contraste	Blur	Ancho	Alto
		Tomato__Bacterial_spot	104.92	33.41	1058.66	256.0 256.0
	Tomato__Early_blight	110.58	46.31	3025.38	256.0 256.0	
	Tomato__Late_blight	113.61	39.63	2544.53	256.0 256.0	
	Tomato__Leaf_Mold	115.41	36.82	3162.19	256.0 256.0	
	Tomato__Septoria_leaf_spot	130.90	37.84	4766.69	256.0 256.0	
	Tomato__Spider_mites Two-spotted_spider_mite	119.27	44.56	6781.24	256.0 256.0	
	Tomato__Target_Spot	118.75	45.10	7209.69	256.0 256.0	
	Tomato__Tomato_Yellow_Leaf_Curl_Virus	114.20	34.71	1797.99	256.0 256.0	
	Tomato__Tomato_mosaic_virus	117.35	34.79	6142.23	256.0 256.0	
	Tomato__healthy	116.24	38.34	12203.31	256.0 256.0	

# Visualización

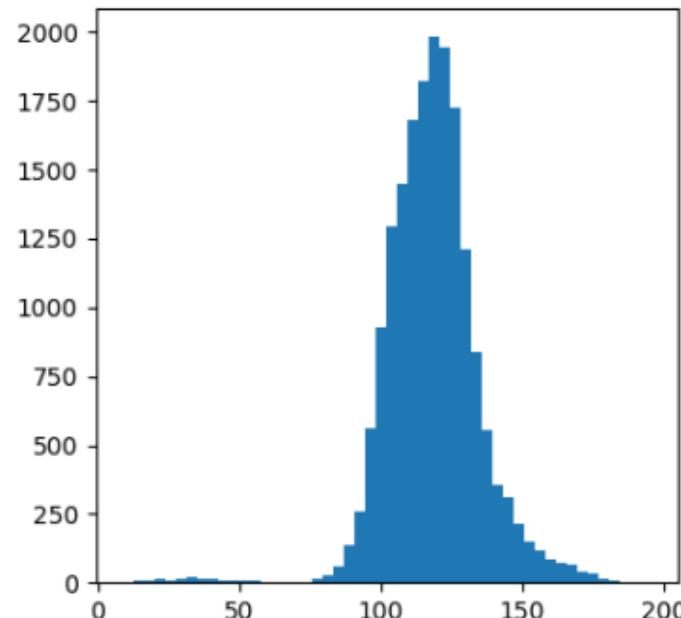
Muy borrosas:

	Clase		Archivo	Blur
14088	Tomato__Tomato_Yellow_Leaf_Curl_Virus	cd37cbce-e141-40a5-8af6-8fa6790c7ed3__UF.GRC....		5.019577
12304	Tomato__Tomato_Yellow_Leaf_Curl_Virus	a5b5d284-5081-4340-9cfa-be58ed74aea3__UF.GRC....		5.232528
14545	Tomato__Tomato_Yellow_Leaf_Curl_Virus	bdb46bf4-d226-4cd1-be9d-4ca36623fd5f__UF.GRC....		7.885695
11502	Tomato__Tomato_Yellow_Leaf_Curl_Virus	8a92e970-1eb5-43f1-852e-2778ddde1503__UF.GRC....		9.046097
12518	Tomato__Tomato_Yellow_Leaf_Curl_Virus	c6380761-954b-4a14-b686-05ce51868611__UF.GRC....		9.248927

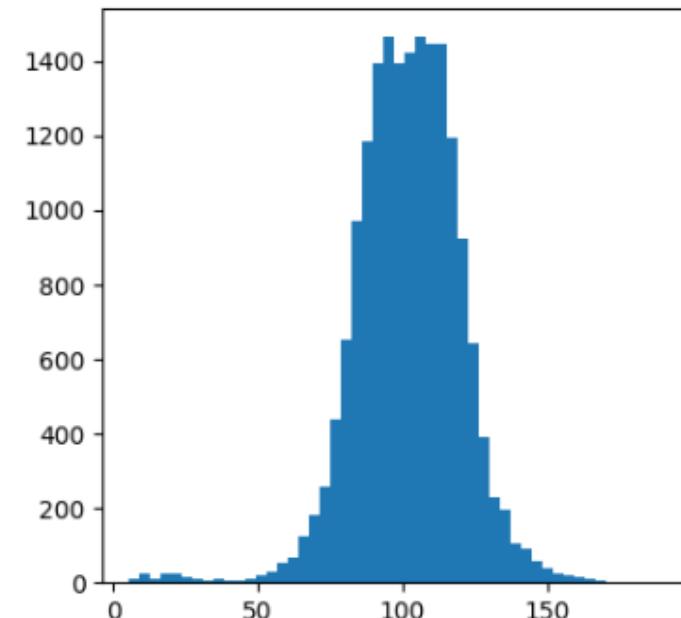
Canal Rojo



Canal Verde



Canal Azul



# Visualización

IMÁGENES MÁS PEQUEÑAS				
	Clase	Archivo	Ancho	Alto
18143	Tomato__healthy	724de01c-90cd-48e7-9360-749bb5b4713b__GH_HL L...	256	256
18142	Tomato__healthy	a424fe57-c3ff-400e-ac88-000ca1781300__RS_HL 9...	256	256
18141	Tomato__healthy	0cabff41-b00b-4692-98af-6b8b0aebe7aa__RS_HL 9...	256	256
18140	Tomato__healthy	ea0a8316-361e-4224-a700-5b86a8b93774__GH_HL L...	256	256
18139	Tomato__healthy	ec768816-02c1-4726-9374-b6d0ccce49d2__RS_HL 0...	256	256
IMÁGENES MÁS GRANDES				
	Clase	Archivo	Ancho	Alto
18145	Tomato__healthy	7c879fc2-09b8-47ab-af57-2508cc854861__GH_HL L...	256	256
0	Tomato__Bacterial_spot	c53b64d9-37ef-4269-b809-00e59c2a033f__GCREC_B...	256	256
18129	Tomato__healthy	0f9a114d-1df8-442c-90e7-d14d2f6efb23__RS_HL 9...	256	256
18128	Tomato__healthy	02d94768-da6d-4c1a-b8b3-fa2e5a2fcd4d__RS_HL 0...	256	256
18127	Tomato__healthy	19e31197-5f6d-4925-975c-884bd4a94bfe__RS_HL 0...	256	256

# Limpieza Automática del Dataset

## Limpieza Automática del Dataset

The diagram illustrates a four-step automatic dataset cleaning process:

- Filtrado de imágenes corruptas, pequeñas y borrosas**: Shows a folder with several images, one of which is being examined with a magnifying glass and has a red 'X' over it.
- Cuarentena inteligente por baja nitidez (blur)**: Shows a folder with a yellow warning sign and a dashed border around one image, indicating it's been flagged for low clarity.
- Reporte auditable + estadísticas y gráficos**: Shows a computer monitor displaying charts and a clipboard with a checklist.
- Recomendaciones automáticas para entrenamiento**: Shows a lightbulb with a brain inside and a clipboard with a checklist.

**Resultado:** TomatoDataset\_ready

Dataset limpio y confiable, listo para train/val/test y Deep Learning.

# Visualización

## LIMPIEZA COMPLETADA

- Total imágenes revisadas: 18146
- Mantener (READY): 17543 (96.68%)
- Cuarentena: 603 (3.32%)
- Descartadas: 0 (0.00%)

Dataset listo: /content/TomatoDataset\_ready

Cuarentena: /content/TomatoDataset\_quarantine

Reporte: /content/TomatoDataset\_cleaning\_report.csv

Vista rápida del reporte (primeras 10 filas):

	Clase	Archivo	Ruta_origen	Estado	Motivo	Ancho	Alto	Blur
0	Tomato__Bacterial_spot	00416648-be6e-4bd4-bc8d-82f43f8a7240__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	672.112518
1	Tomato__Bacterial_spot	0045ba29-ed1b-43b4-afde-719cc7adefdb__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	1545.125317
2	Tomato__Bacterial_spot	00639d29-2d1a-4fcf-9bd3-a2b3109c74c4__UF.GRC...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	785.756565
3	Tomato__Bacterial_spot	00728f4d-83a0-49f1-87f8-374646fcda05__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	1940.613708
4	Tomato__Bacterial_spot	00a7c269-3476-4d25-b744-44d6353cd921__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	1306.964077
5	Tomato__Bacterial_spot	00b7e89a-e129-4576-b51f-48923888bff9__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	1599.762820
6	Tomato__Bacterial_spot	01375198-62af-4c40-bddf-f3c11107200b__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	1254.932037
7	Tomato__Bacterial_spot	014b58ae-091b-408a-ab4a-5a780cd1c3f3__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	319.953421
8	Tomato__Bacterial_spot	01a3cf3f-94c1-44d5-8972-8c509d62558e__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	1210.490609
9	Tomato__Bacterial_spot	01a46cb5-d354-4f59-868e-e56186701541__GCREC_B...	/content/TomatoDataset_merged/Tomato__Bacteri...	keep	ok	256	256	970.336764

# Visualización

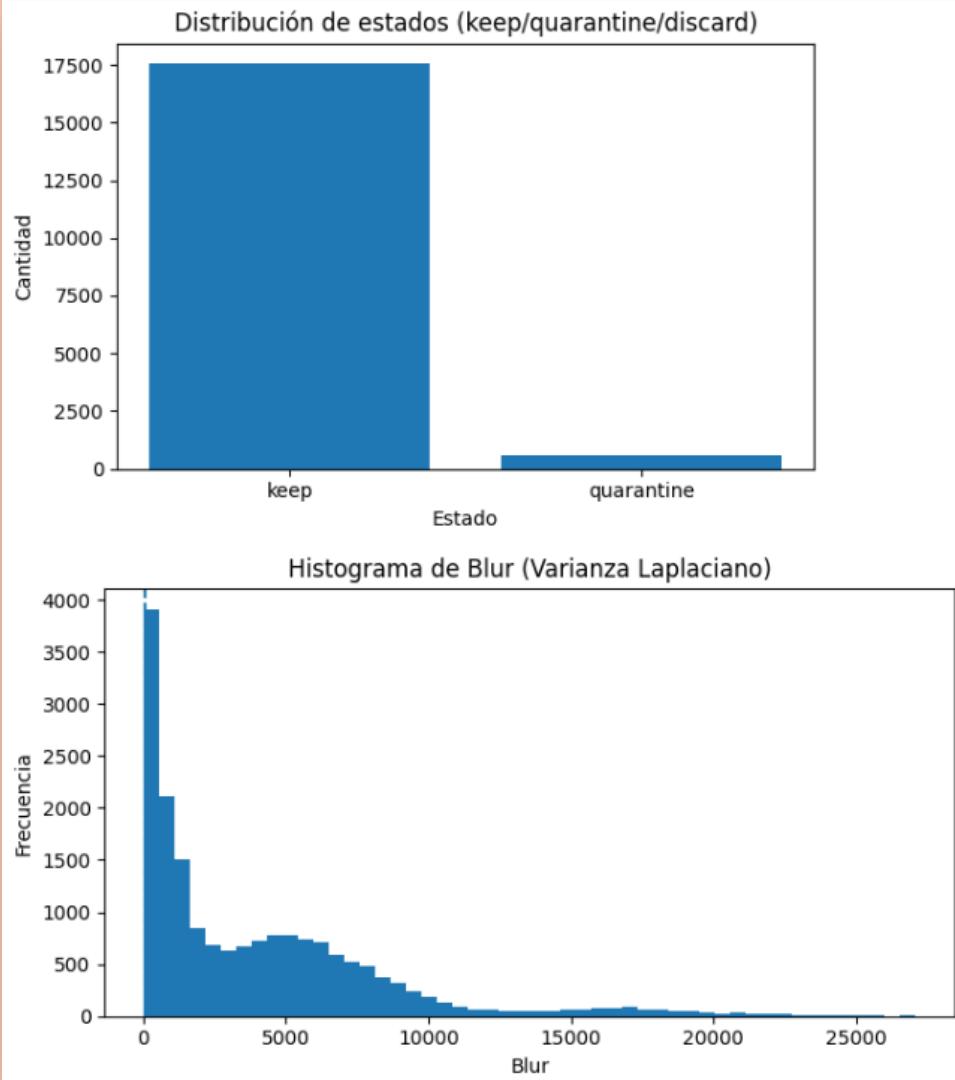
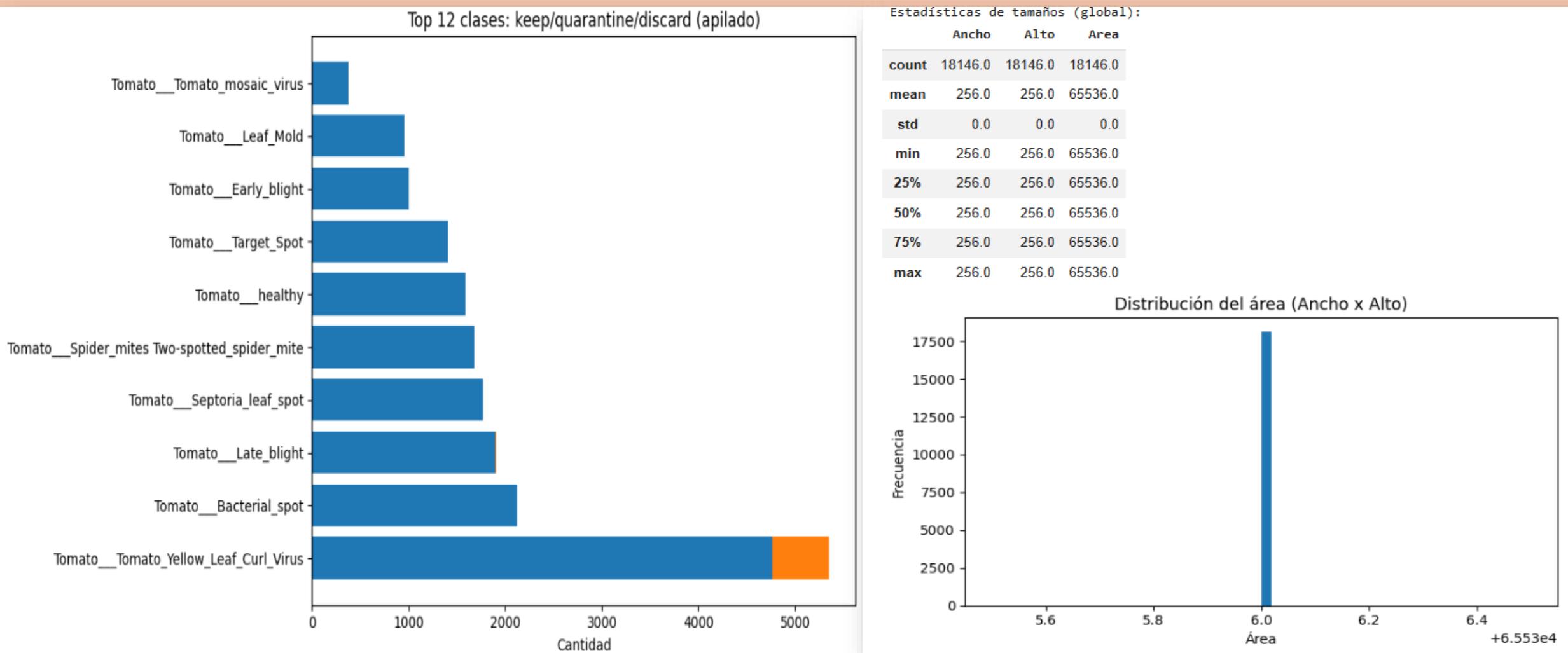


Tabla por clase (keep/quarantine/discard):

Clase	Estado	keep	quarantine	total
Tomato__Tomato_Yellow_Leaf_Curl_Virus	4761	596	5357	
Tomato__Bacterial_spot	2126	1	2127	
Tomato__Late_blight	1897	4	1901	
Tomato__Septoria_leaf_spot	1771	0	1771	
Tomato__Spider_mites Two-spotted_spider_mite	1676	0	1676	
Tomato__healthy	1585	0	1585	
Tomato__Target_Spot	1404	0	1404	
Tomato__Early_blight	998	2	1000	
Tomato__Leaf_Mold	952	0	952	
Tomato__Tomato_mosaic_virus	373	0	373	

# Visualización



# Split Estratificado Train / Val / Test



# Visualización

```
SPLIT COMPLETADO
Dataset split en: /content/TomatoDataset_split
Reporte guardado en: /content/TomatoDataset_split_report.csv

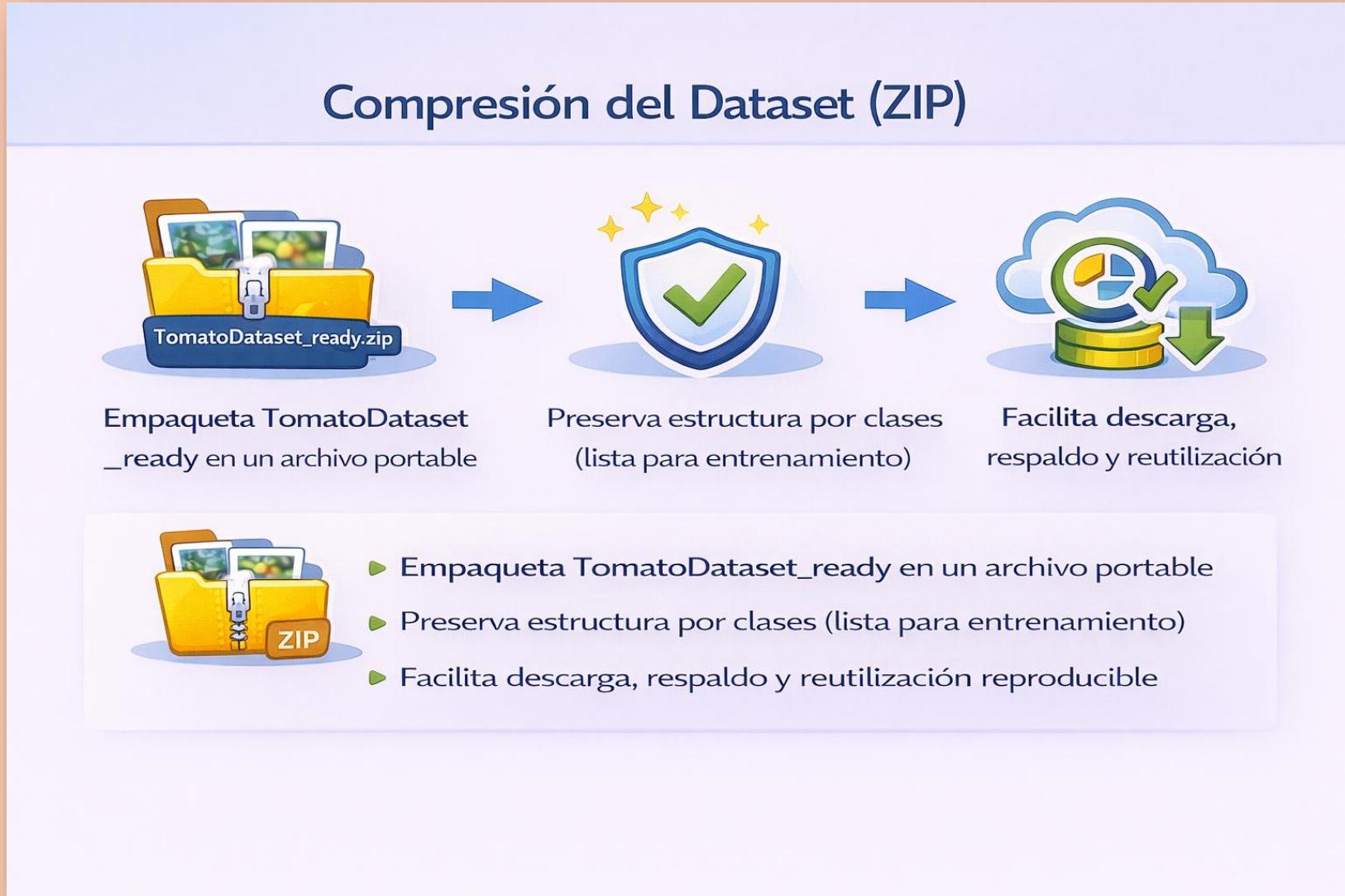
Resumen (primeras 15 clases):

          Clase  Total  Train  Val  Test  Train_%  Val_%  Test_%
0      TomatoDataset_ready     0     0     0     0    NaN    NaN    NaN
1  Tomato__Bacterial_spot  2126  1488   318   320   69.99  14.96  15.05
2  Tomato__Early_blight   998   698   149   151   69.94  14.93  15.13
3  Tomato__Late_blight  1897  1327   284   286   69.95  14.97  15.08
4  Tomato__Leaf_Mold     952   666   142   144   69.96  14.92  15.13
5  Tomato__Septoria_leaf_spot  1771  1239   265   267   69.96  14.96  15.08
6  Tomato__Spider_mites Two-spotted_spider_mite  1676  1173   251   252   69.99  14.98  15.04
7  Tomato__Target_Spot    1404   982   210   212   69.94  14.96  15.10
8  Tomato__Tomato_Yellow_Leaf_Curl_Virus  4761  3332   714   715   69.99  15.00  15.02
9  Tomato__Tomato_mosaic_virus    373   261    55    57   69.97  14.75  15.28
10  Tomato__healthy     1585  1109   237   239   69.97  14.95  15.08

Totales globales:
- Total imágenes: 17543
- Train: 12275
- Val: 2625
- Test: 2643

Estructura creada (ejemplo):
- /content/TomatoDataset_split/train/<clase>/
- /content/TomatoDataset_split/val/<clase>/
- /content/TomatoDataset_split/test/<clase>/
```

# Compresión del Dataset

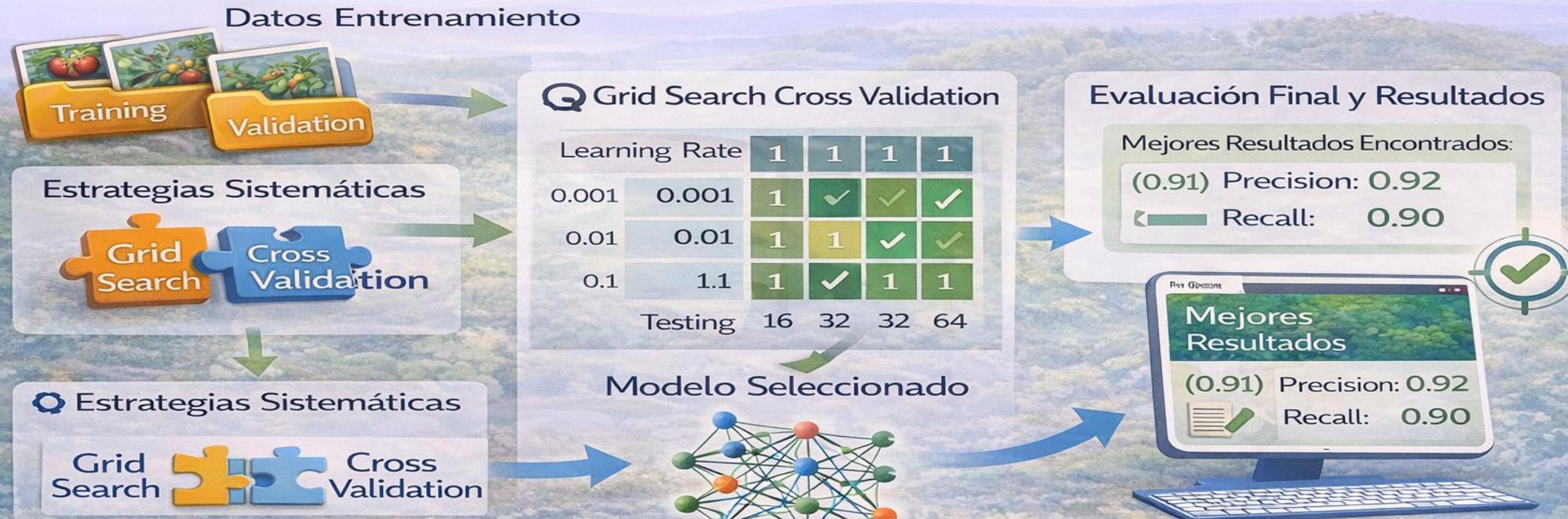


# Visualización

Árbol de directorios: /content/TomatoDataset\_ready

```
└── TomatoDataset_ready/
    ├── TomatoDataset_ready/
    ├── Tomato__Bacterial_spot/
    ├── Tomato__Early_blight/
    ├── Tomato__Late_blight/
    ├── Tomato__Leaf_Mold/
    ├── Tomato__Septoria_leaf_spot/
    ├── Tomato__Spider_mites Two-spotted_spider_mite/
    ├── Tomato__Target_Spot/
    ├── Tomato__Tomato_Yellow_Leaf_Curl_Virus/
    ├── Tomato__Tomato_mosaic_virus/
    └── Tomato__healthy/
```

# Hacia una Agricultura de Precisión en Chiloé: Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares en Cultivos



Selección sistemática de hiperparámetros utilizando Grid Search, Cross Validation y estrategias sistemáticas

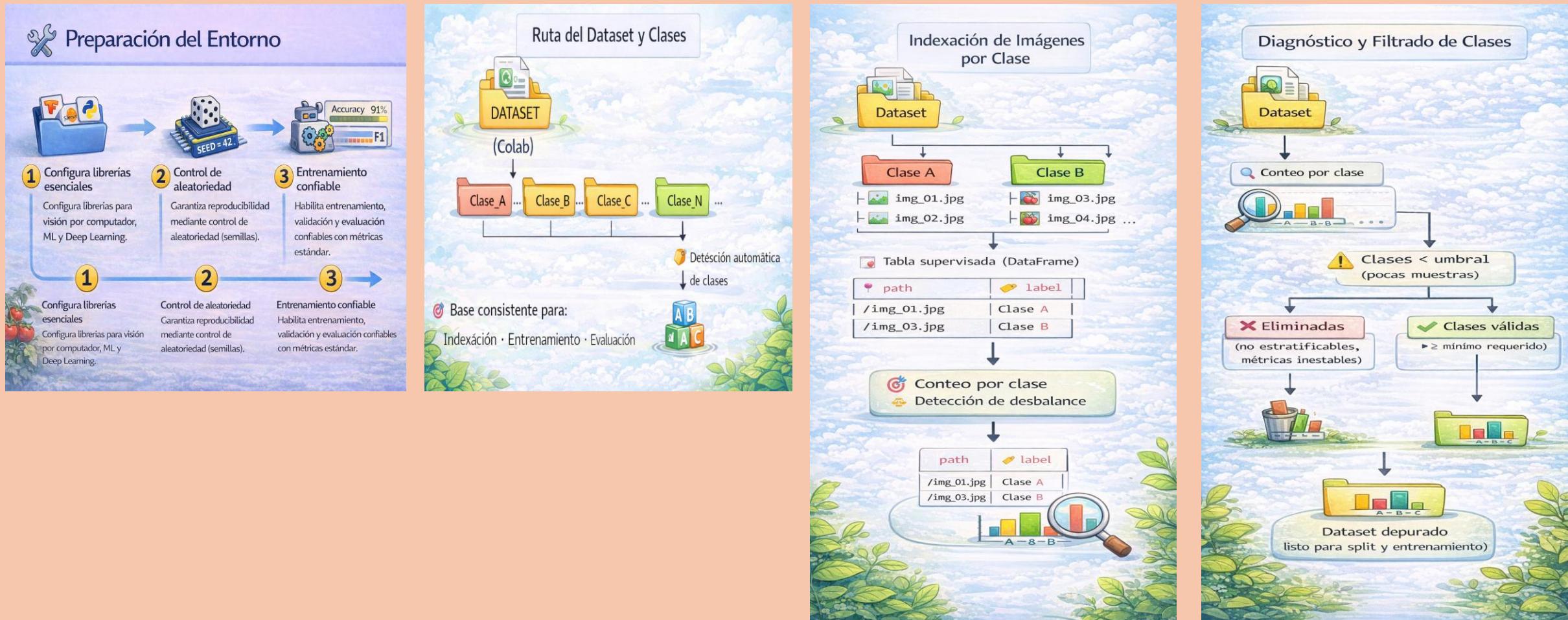
# Modelos de Implementación



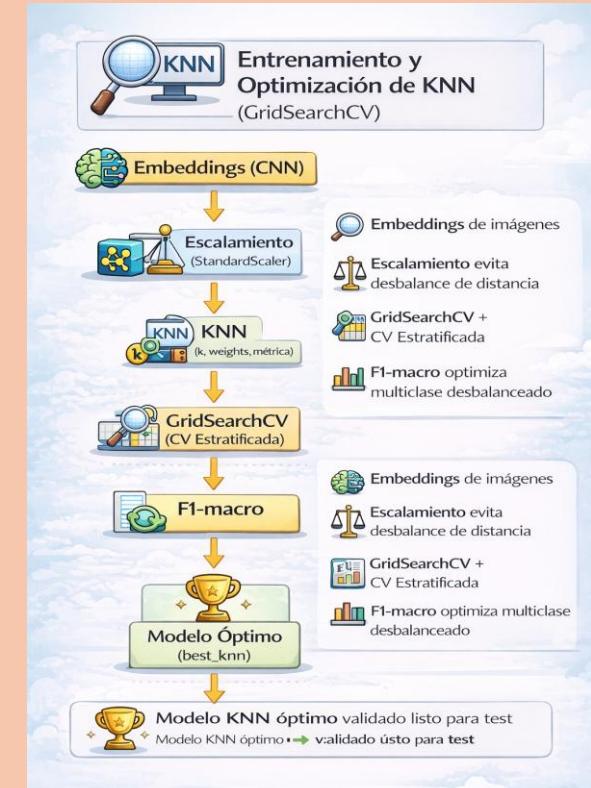
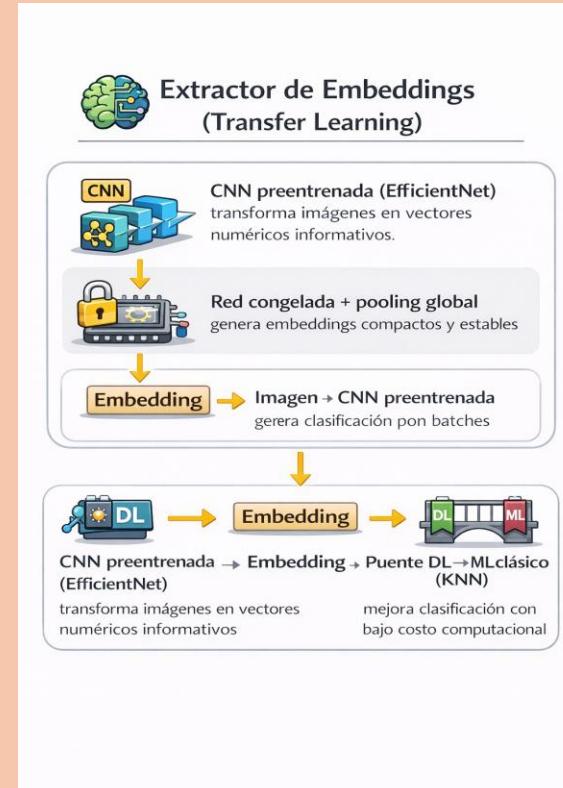
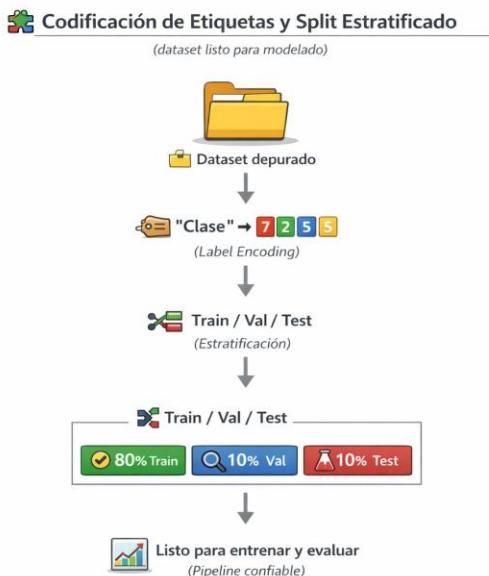
# Modelos de Implementación

- **KNN para Enfermedades del Tomate**
  - **Clasificación por similitud:** la hoja se etiqueta según sus **K hojas más cercanas** en el espacio de características.
  - **Datos de entrada:** imágenes foliares convertidas en **features o embeddings** (color, textura o CNN).
  - **Uso en el proyecto:** modelo base explicable para detectar **hojas sanas vs enfermas** y tipos de enfermedad.

# Pasos para Implementar KNN



# Pasos para Implementar KNN



# Pasos para Implementar KNN



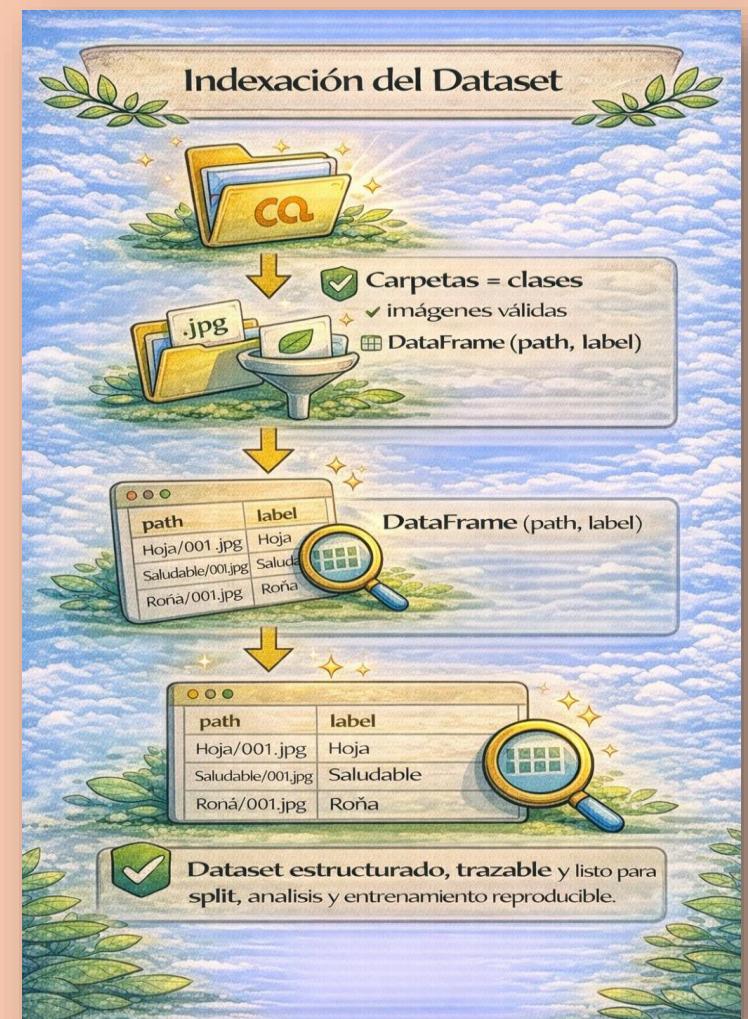
# Modelos de Implementación



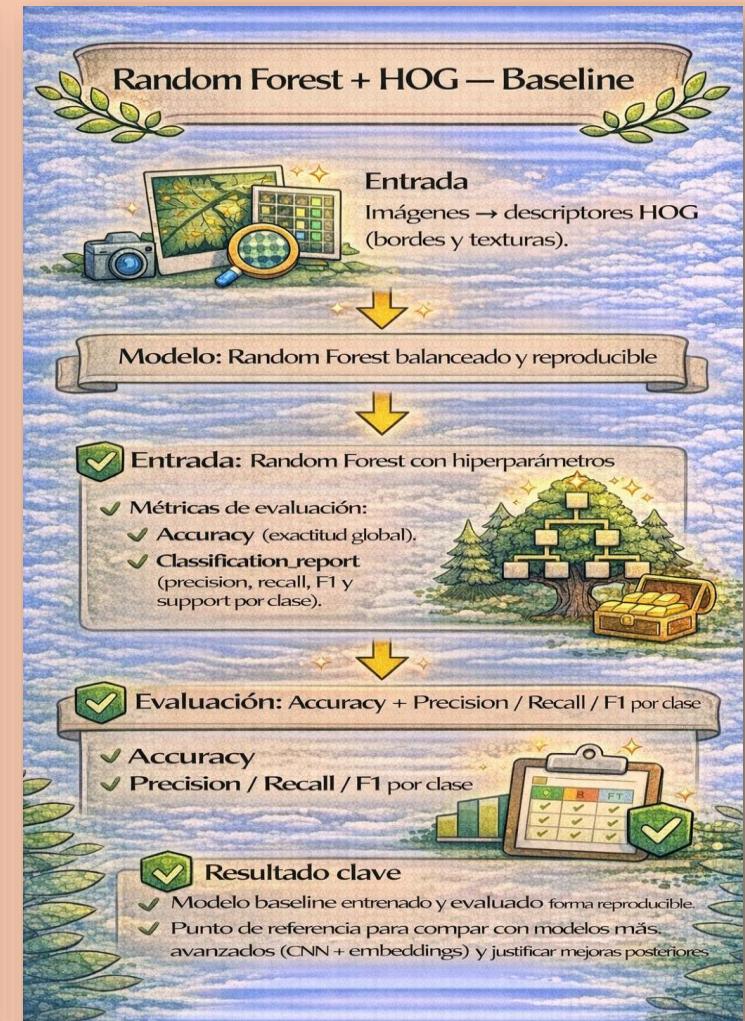
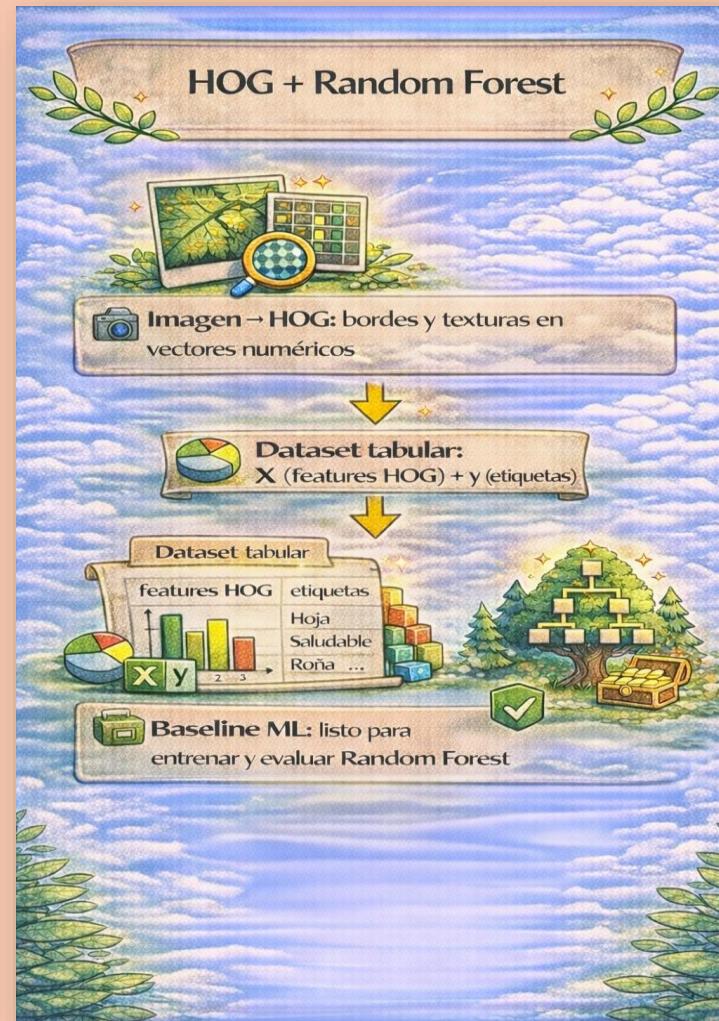
# Modelos de Implementación

- **Random Forest para Enfermedades del Tomate**
  - **Ensamble de árboles** → decisión por votación.
  - **Entrada:** *features / embeddings* de hojas.
  - **Robusto y estable** → menos sobreajuste, buen rendimiento multiclas.

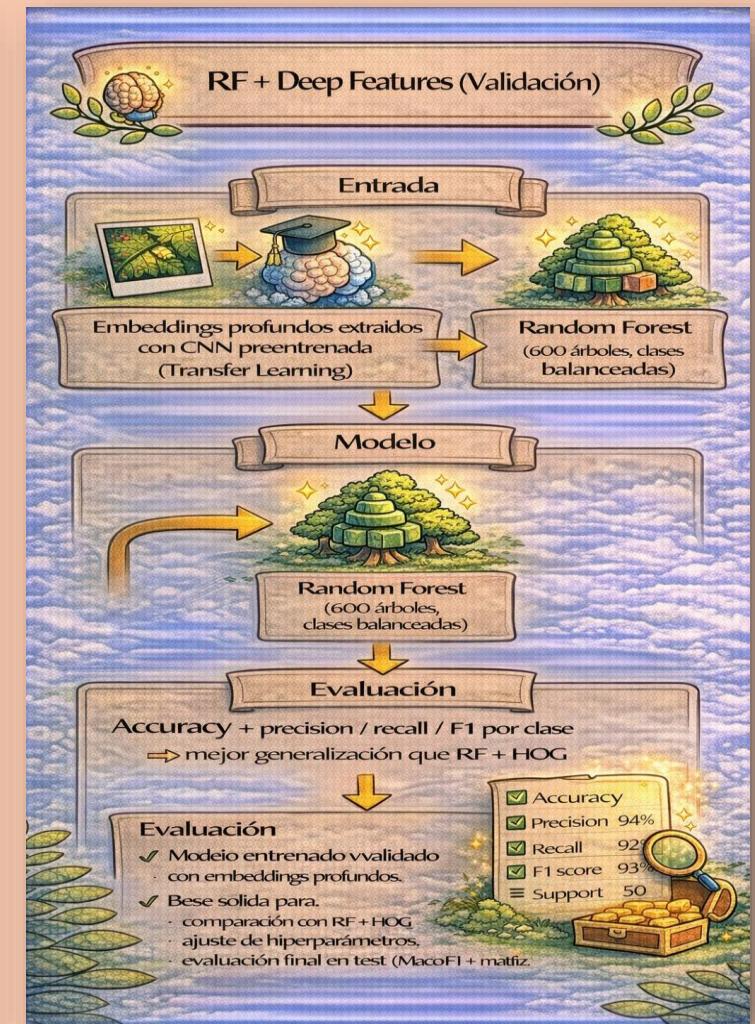
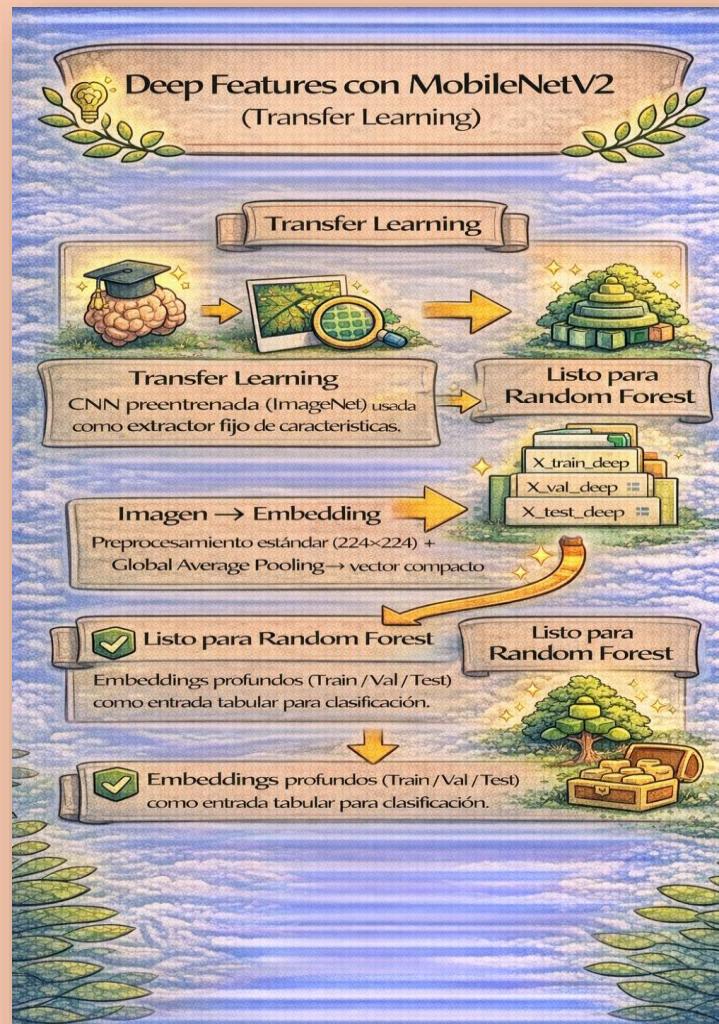
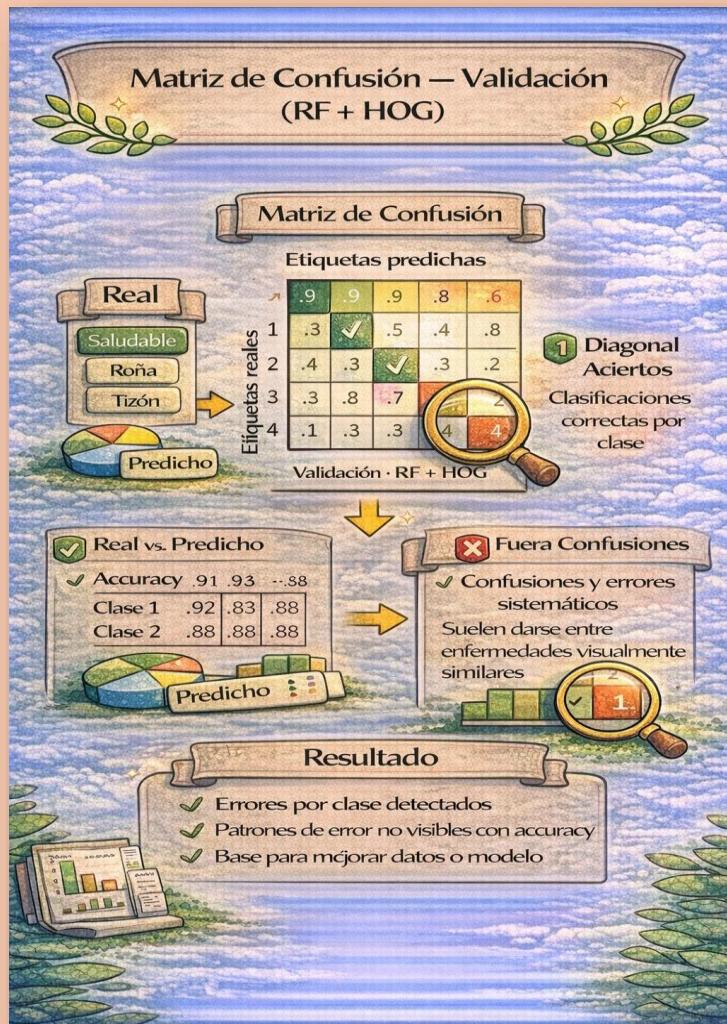
# Pasos para Implementar RF



# Pasos para Implementar RF

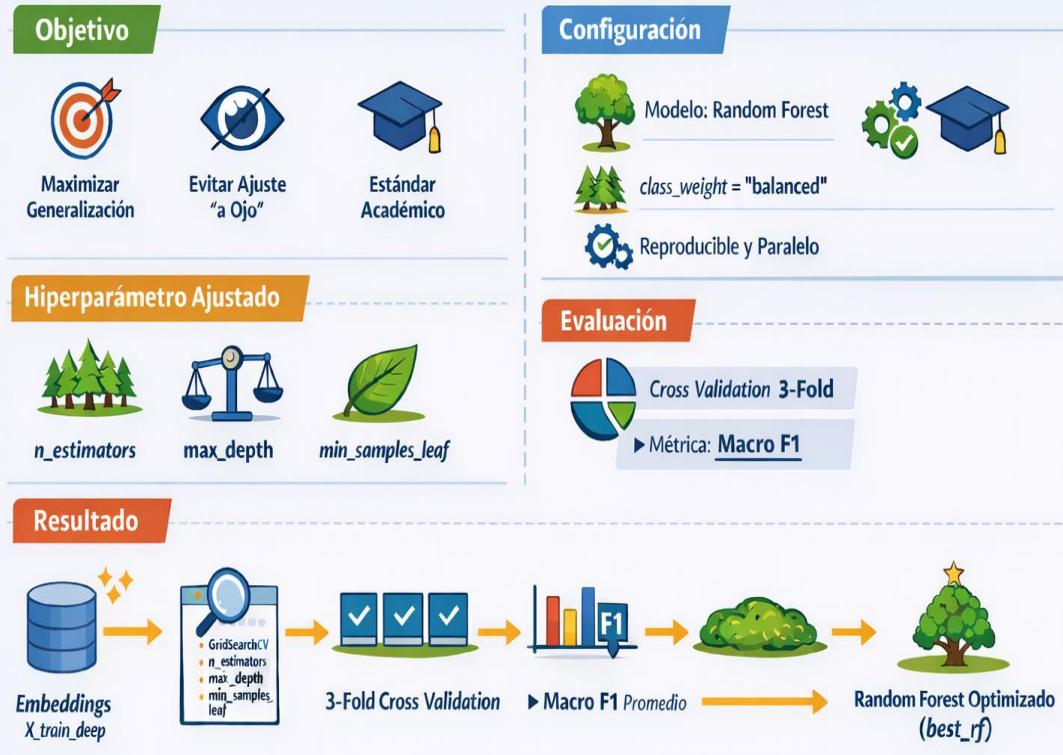


# Pasos para Implementar RF

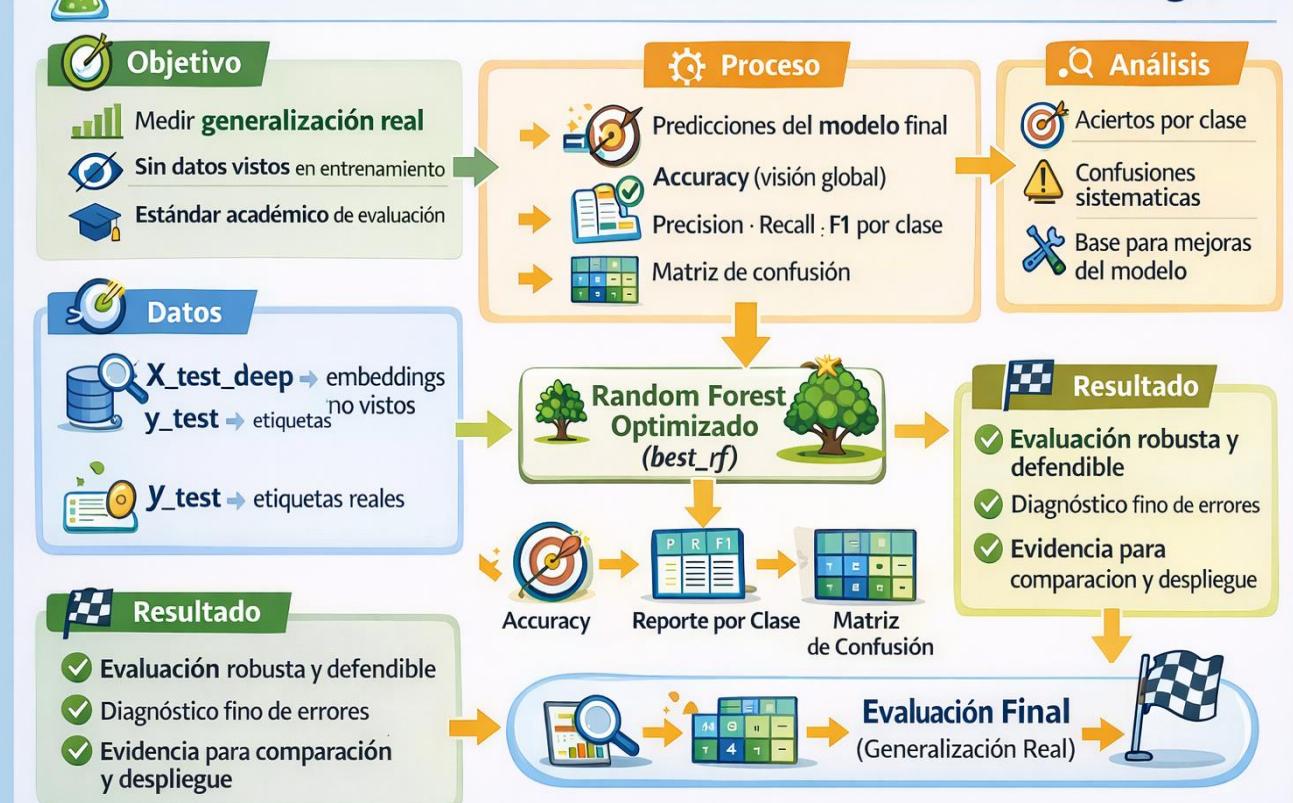


# Pasos para Implementar RF

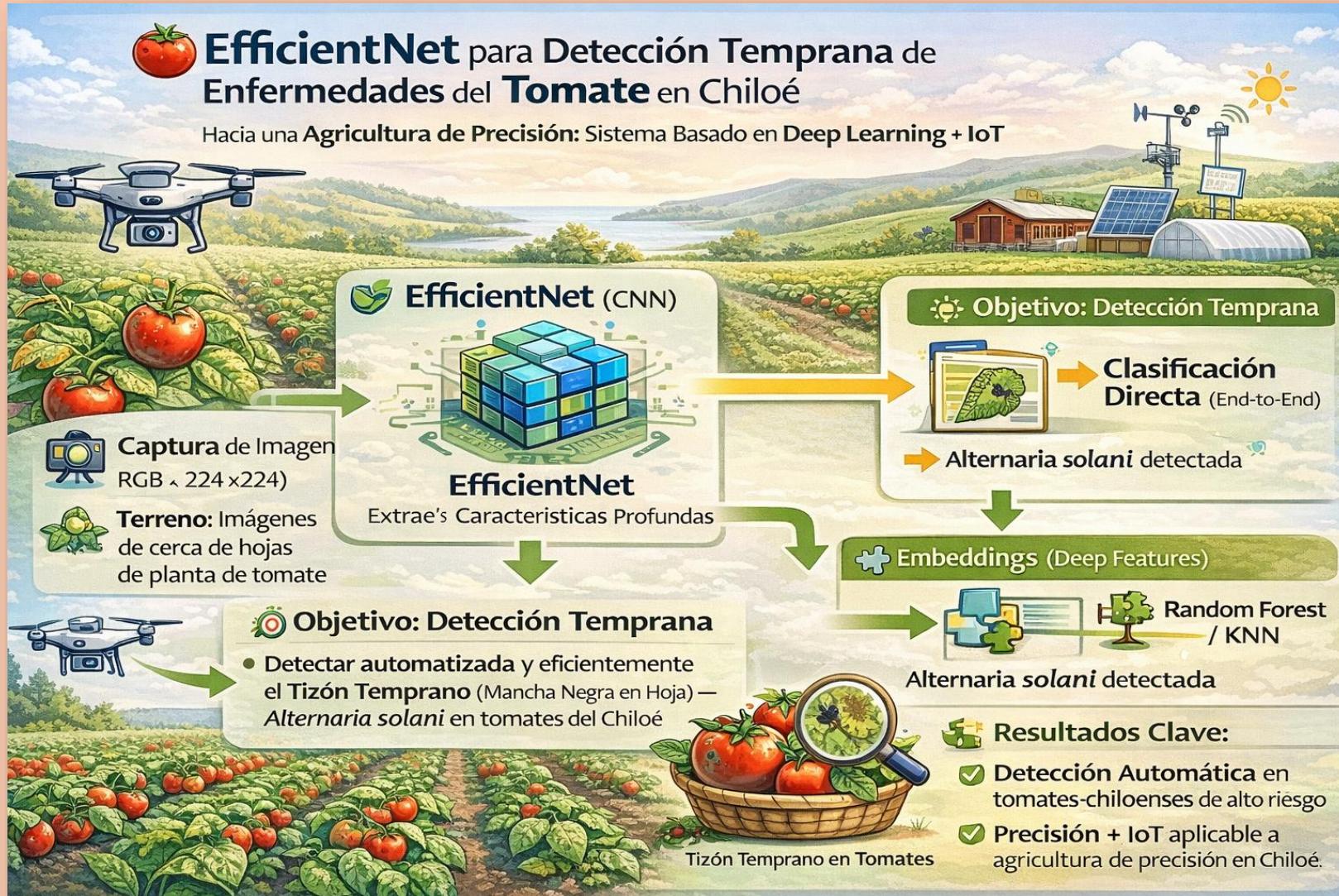
## Ajuste de Hiperparámetro con GridSearchCV para RF



## Evaluación Final en TEST – Random Forest (Embeddings)



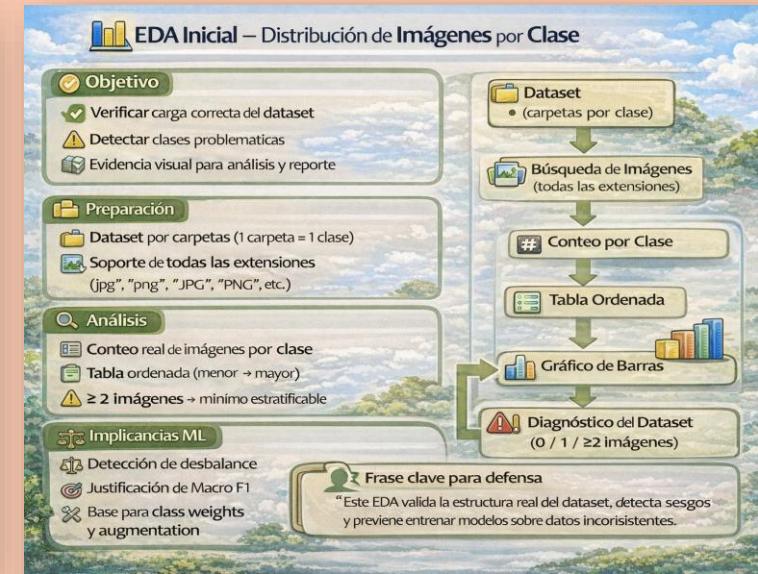
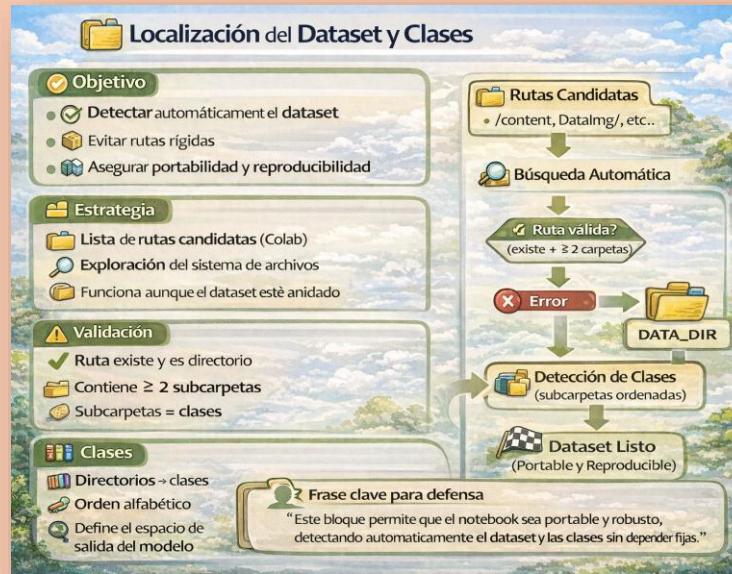
# Modelos de Implementación



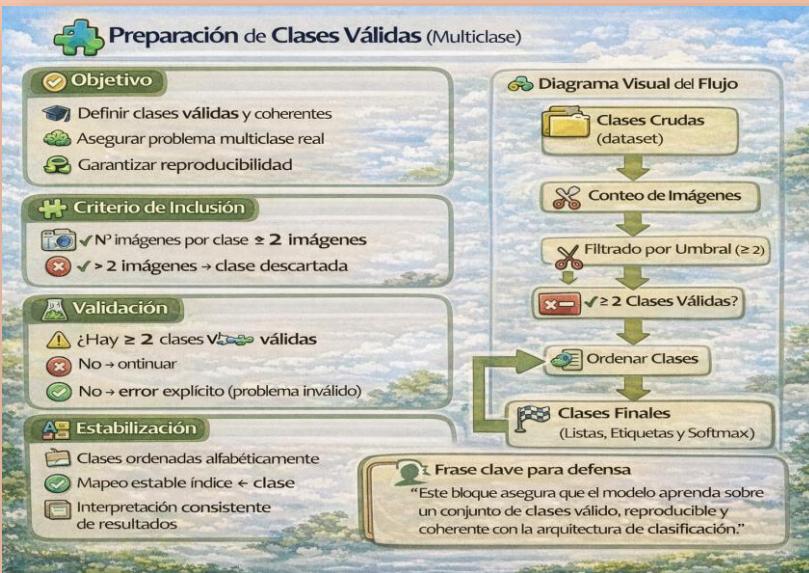
# Modelos de Implementación

- **EfficientNet para Enfermedades del Tomate**
  - **Captura en campo (dron/terreno)**: imágenes RGB de hojas de tomate → entrada al sistema.
  - **EfficientNet (CNN)**: extrae **Deep features** y permite **clasificación directa o embeddings + RF/KNN**.
  - **Resultado: detección temprana** de enfermedades del tomate (p. ej., tizón temprano) para agricultura de precisión en Chiloé.

# Pasos para Implementar EfficientNet



# Pasos para Implementar EfficientNet



# Pasos para Implementar EfficientNet

## Class Weights

Corrección de desbalance: clases raras pesan más, clases frecuentes menos.

Efecto: ➡ mejora Recall y Macro F1 en multiclase.

Más peso      Menos peso

Uso práctico: pesos calculados en `train` y aplicados en Keras/TensorFlow.

Resultado del Bloque:

- ✓ Pesos por clase listos para entrenamiento
- ✓ Evidencia cuantitativa del desbalance
- ✓ Modelo menos sesgado y más justo en multiclase

## tf.data

X crudo → tensores: carga, resize, one-hot.

Pipeline eficiente: map paralelo · batch · prefetch (shuffle solo en train)

Salida: `train_ds, val_ds, test_ds` listos para GPU.

Salida: `train_ds, val_ds, test_ds` listos para GPU.

## Data Augmentation

Regularización y Generalización

Objetivo

- Aumentar la **variabilidad efectiva** del conjunto de entrenamiento mediante transformaciones aleatorias, reduciendo el **sobreajuste** y mejorando la generalización del modelo.

Pipeline de Augmentación (Keras)

- Módulo secuencial integrado al modelo.
- Ejecutado dentro del grafo de TensorFlow.
- Reutilizable, eficiente y reproducible.

Transformaciones Aplicadas

- Flip horizontal: invariancia a orientación
- Rotación leve: variación real de ángulos.
- Zoom controlado: distintas escalas de lesión.

Buenas Prácticas

- Augmentation solo en **entrenamiento**.
- Validación y test sin alteraciones.
- Actúa como regularización implícita.

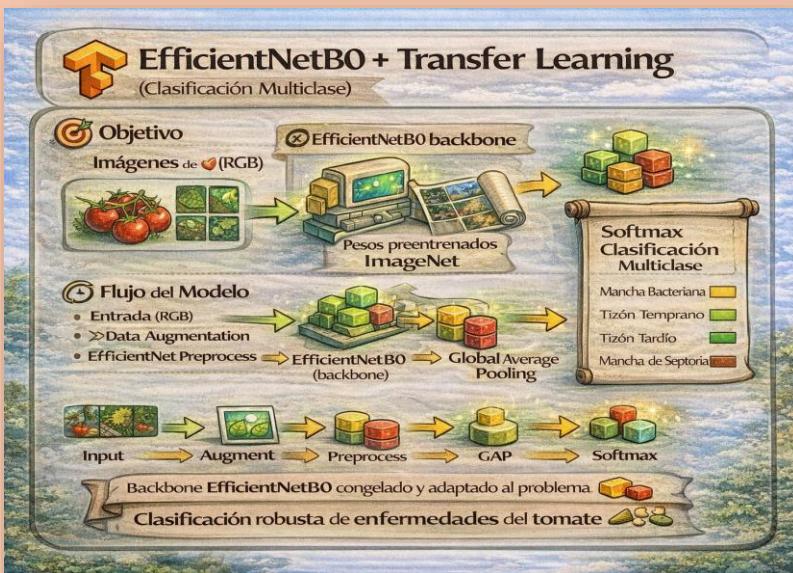
Impacto Esperado

- Mejor desempeño en datos no vistos
- Máior robustez ante variabilidad real
- Mejora de métricas balanceadas (ej. Macro F1)

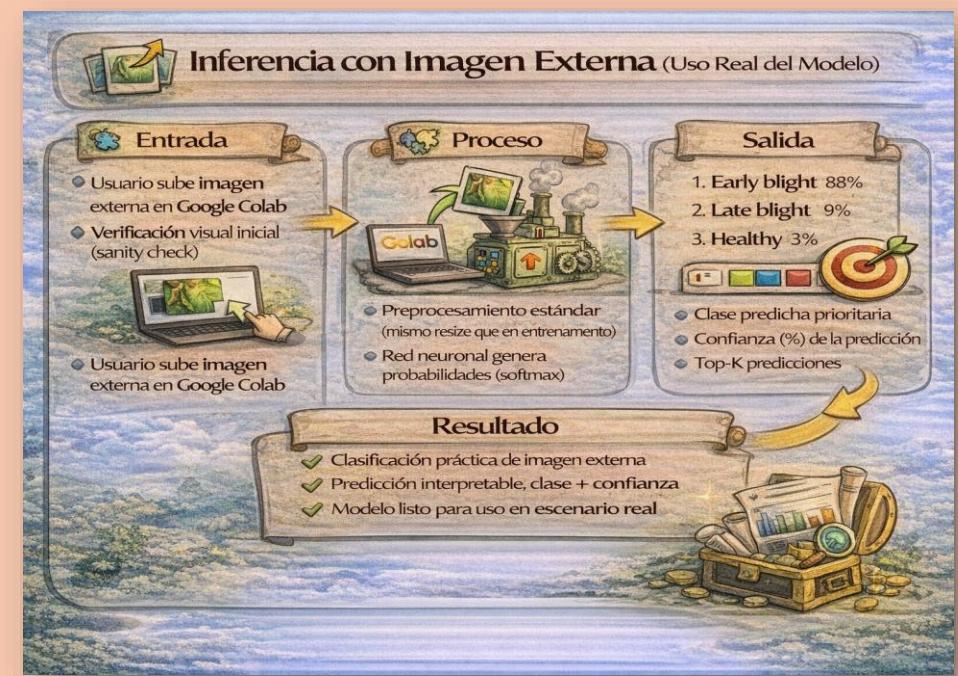
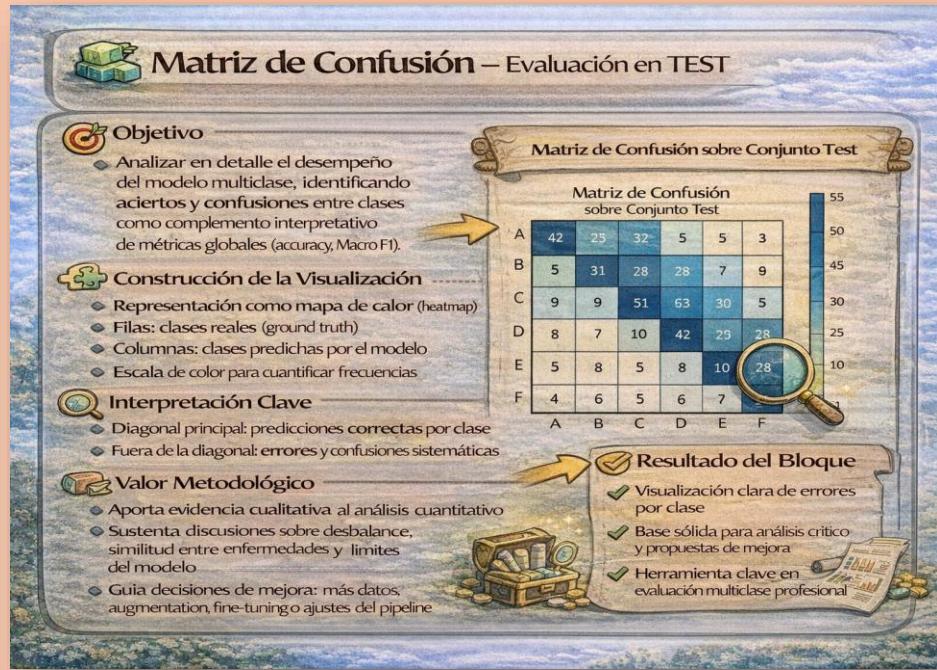
Resultado del Bloque:

- Módulo de **Data Augmentation** listo para integrarse en **EfficientNet/CNN**. (ortaleciendo el aprendizaje en clasificación de enfermedades foliares en tomate).

# Pasos para Implementar EfficientNet



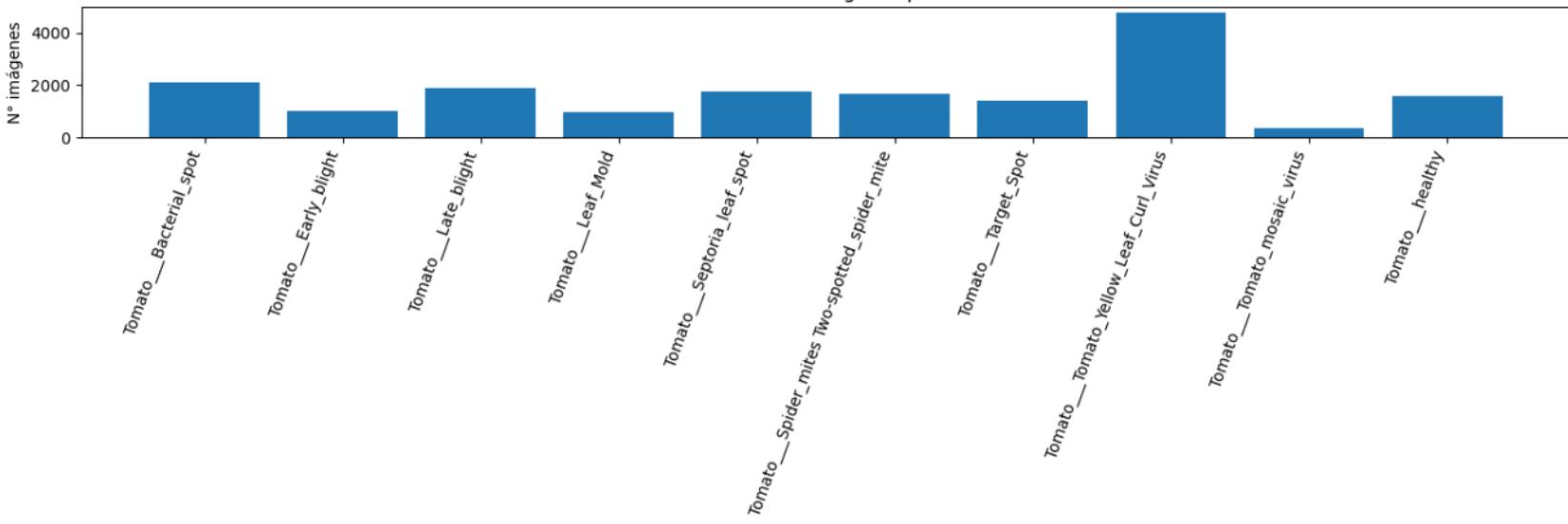
# Pasos para Implementar EfficientNet



# Visualizaciones

```
Conteo por clase:  
Tomato__Tomato_mosaic_virus: 373  
Tomato__Leaf_Mold: 952  
Tomato__Early_blight: 998  
Tomato__Target_Spot: 1404  
Tomato__healthy: 1585  
Tomato__Spider_mites Two-spotted_spider_mite: 1676  
Tomato__Septoria_leaf_spot: 1771  
Tomato__Late_blight: 1897  
Tomato__Bacterial_spot: 2126  
Tomato__Tomato_Yellow_Leaf_Curl_Virus: 4761
```

Distribución de imágenes por clase



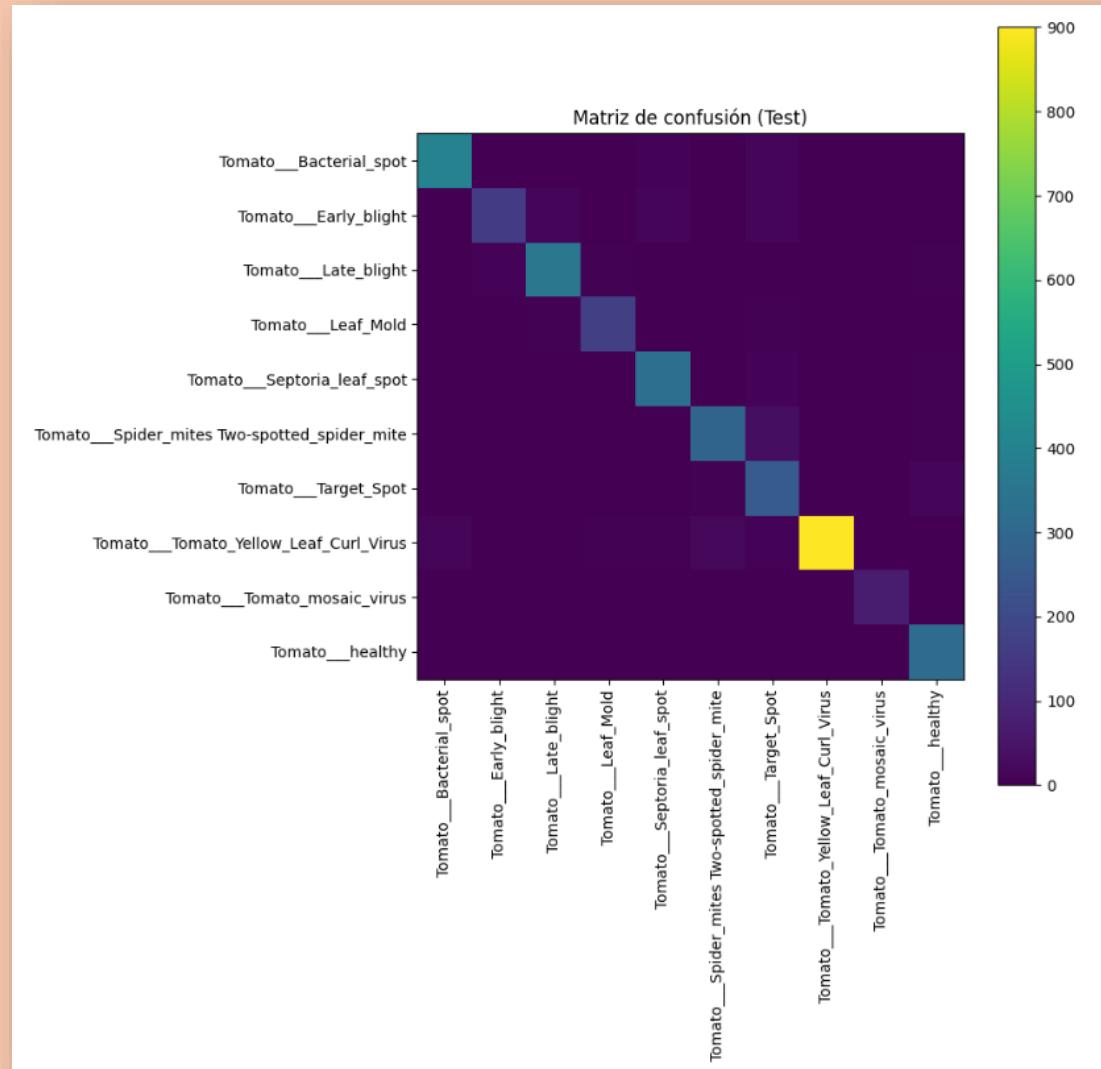
# Visualizaciones

```
Test loss: 0.2398664355278015  
Test accuracy: 0.9244799017906189  
Macro F1: 0.9103535555182836
```

Reporte de clasificación:

		precision	recall	f1-score	support
	Tomato__Bacterial_spot	0.96	0.94	0.95	425
	Tomato__Early_blight	0.92	0.78	0.84	200
	Tomato__Late_blight	0.94	0.94	0.94	380
	Tomato__Leaf_Mold	0.91	0.90	0.90	190
	Tomato__Septoria_leaf_spot	0.90	0.93	0.92	354
Tomato__Spider_mites	Two-spotted_spider_mite	0.90	0.87	0.88	335
	Tomato__Target_Spot	0.75	0.92	0.83	281
Tomato__Tomato_Yellow_Leaf_Curl_Virus		1.00	0.95	0.97	952
	Tomato__Tomato_mosaic_virus	0.94	0.91	0.93	75
	Tomato__healthy	0.90	0.99	0.94	317
	accuracy			0.92	3509
	macro avg	0.91	0.91	0.91	3509
	weighted avg	0.93	0.92	0.93	3509

# Visualizaciones



# Visualizaciones

Elegir archivos Ningún archivo seleccionado Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.  
Saving 0a9b3ff4-5343-4814-ac2c-fdb3613d4e4d\_\_Crn1\_L.Mold 6559.JPG to 0a9b3ff4-5343-4814-ac2c-fdb3613d4e4d\_\_Crn1\_L.Mold 6559.JPG  
Imagen cargada: 0a9b3ff4-5343-4814-ac2c-fdb3613d4e4d\_\_Crn1\_L.Mold 6559.JPG

Imagen subida



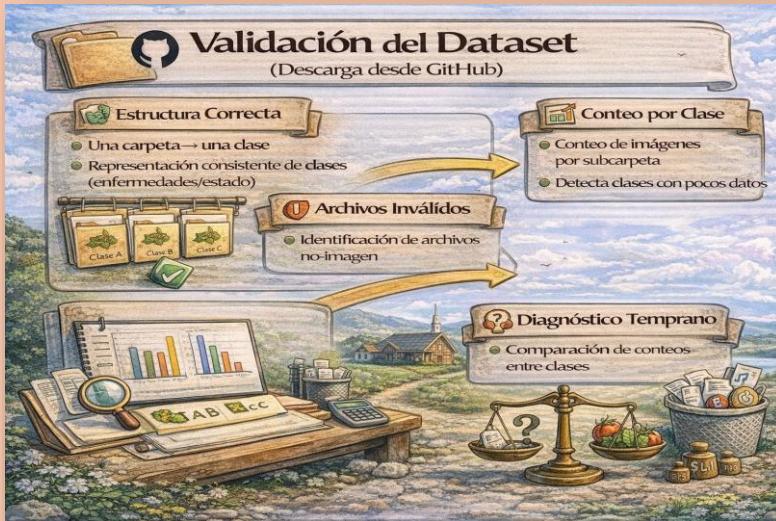
Predicción: Tomato\_\_Leaf\_Mold

Confianza (prob): 0.9899

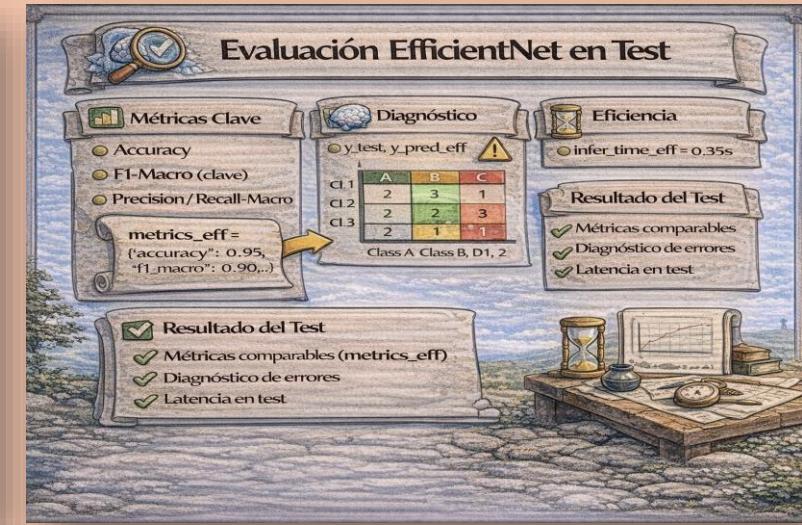
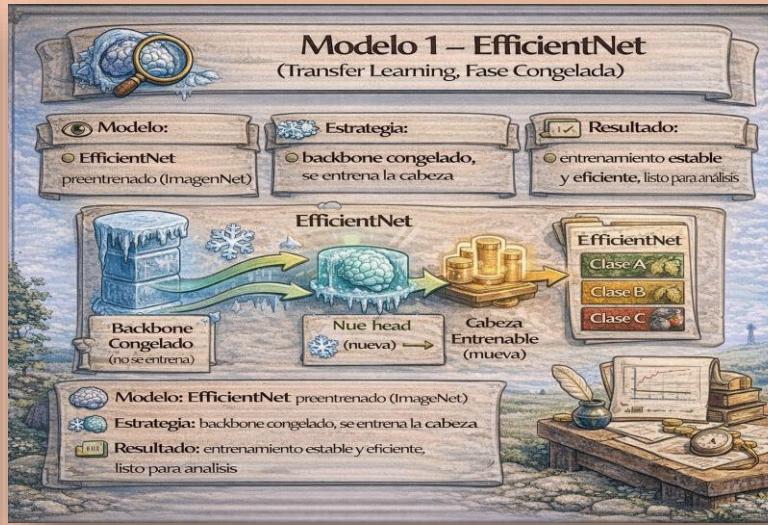
Top 5 predicciones:

- Tomato\_\_Leaf\_Mold prob=0.9899
- Tomato\_\_Late\_blight prob=0.0039
- Tomato\_\_Tomato\_Yellow\_Leaf\_Curl\_Virus prob=0.0021
- Tomato\_\_Spider\_mites Two-spotted\_spider\_mite prob=0.0014
- Tomato\_\_Early\_blight prob=0.0012

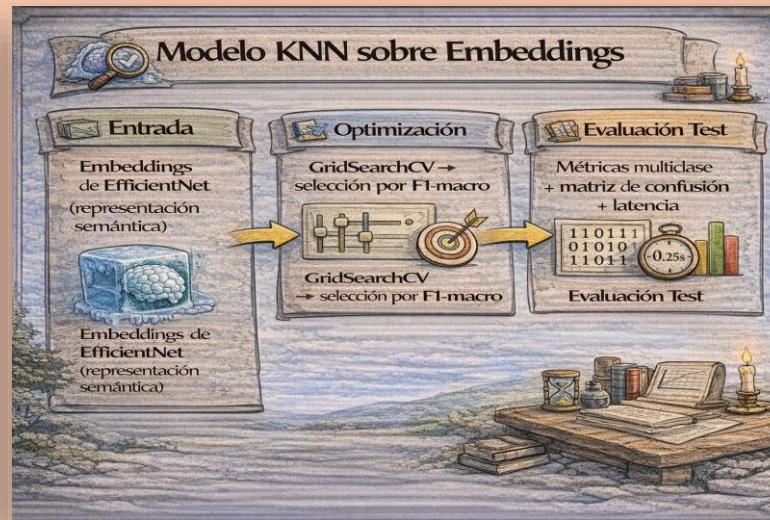
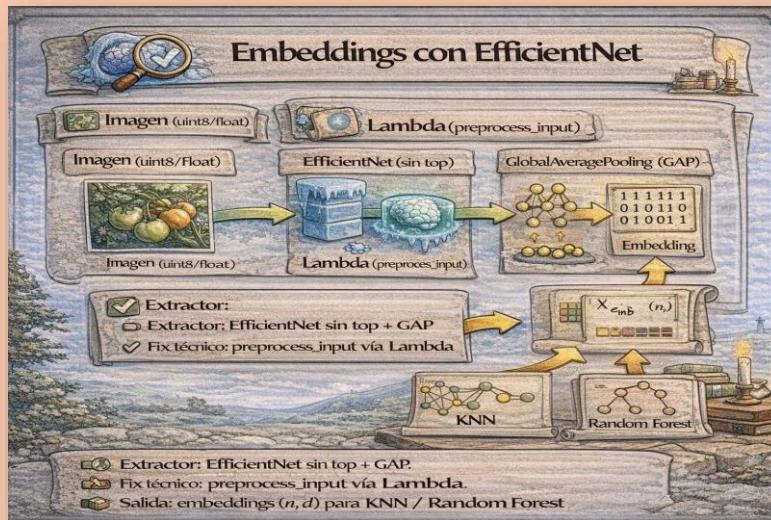
# Proceso de Comparación



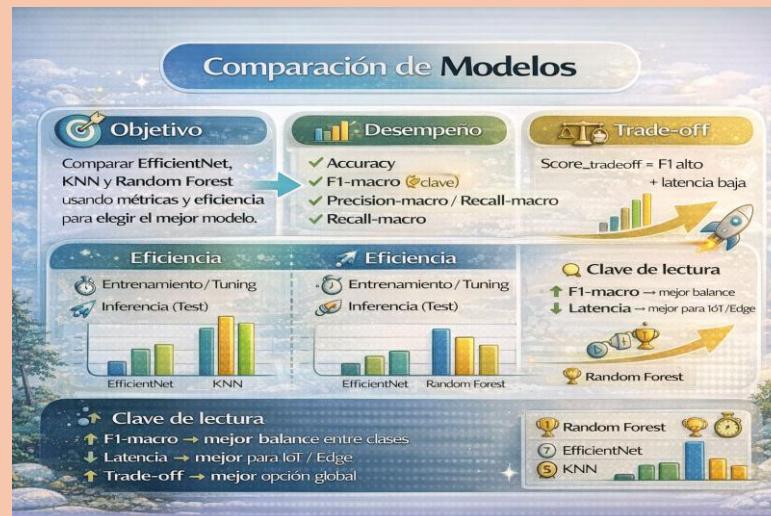
# Proceso de Comparación



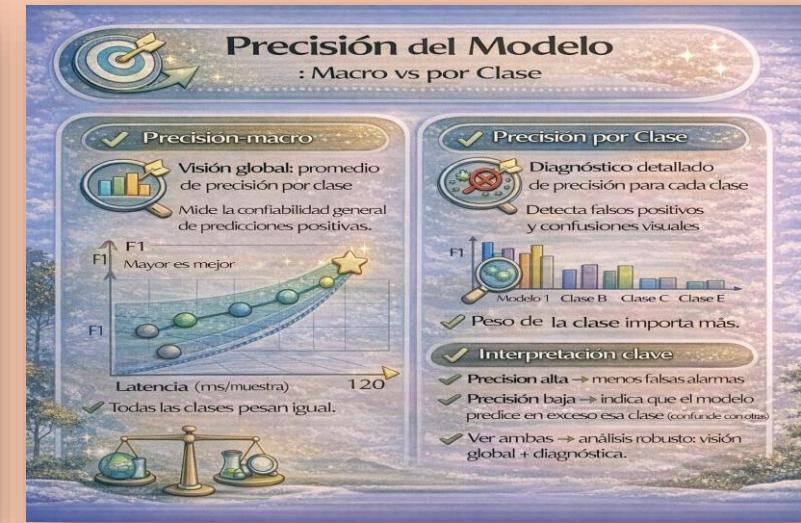
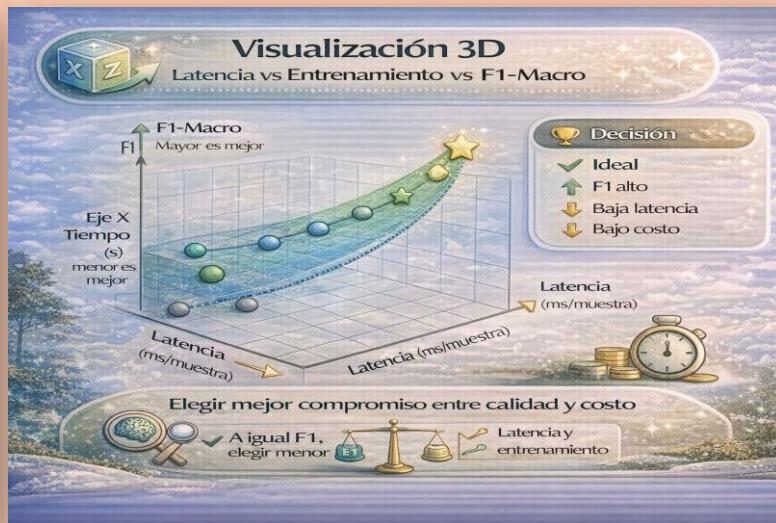
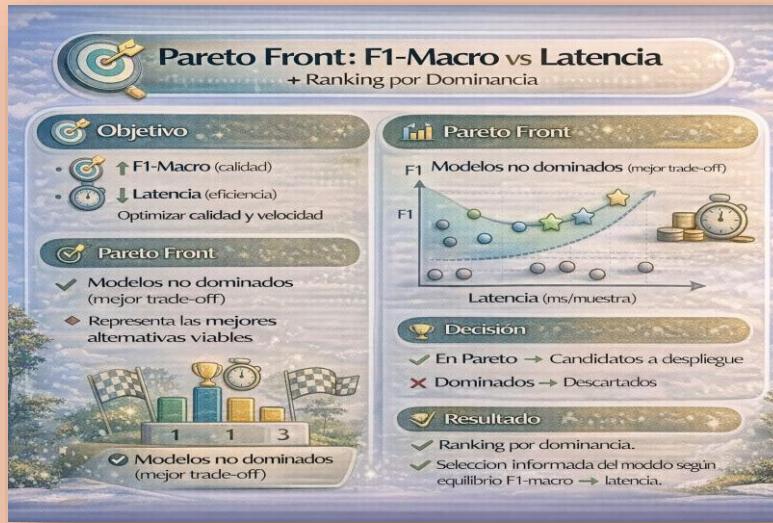
# Proceso de Comparación



# Proceso de Comparación



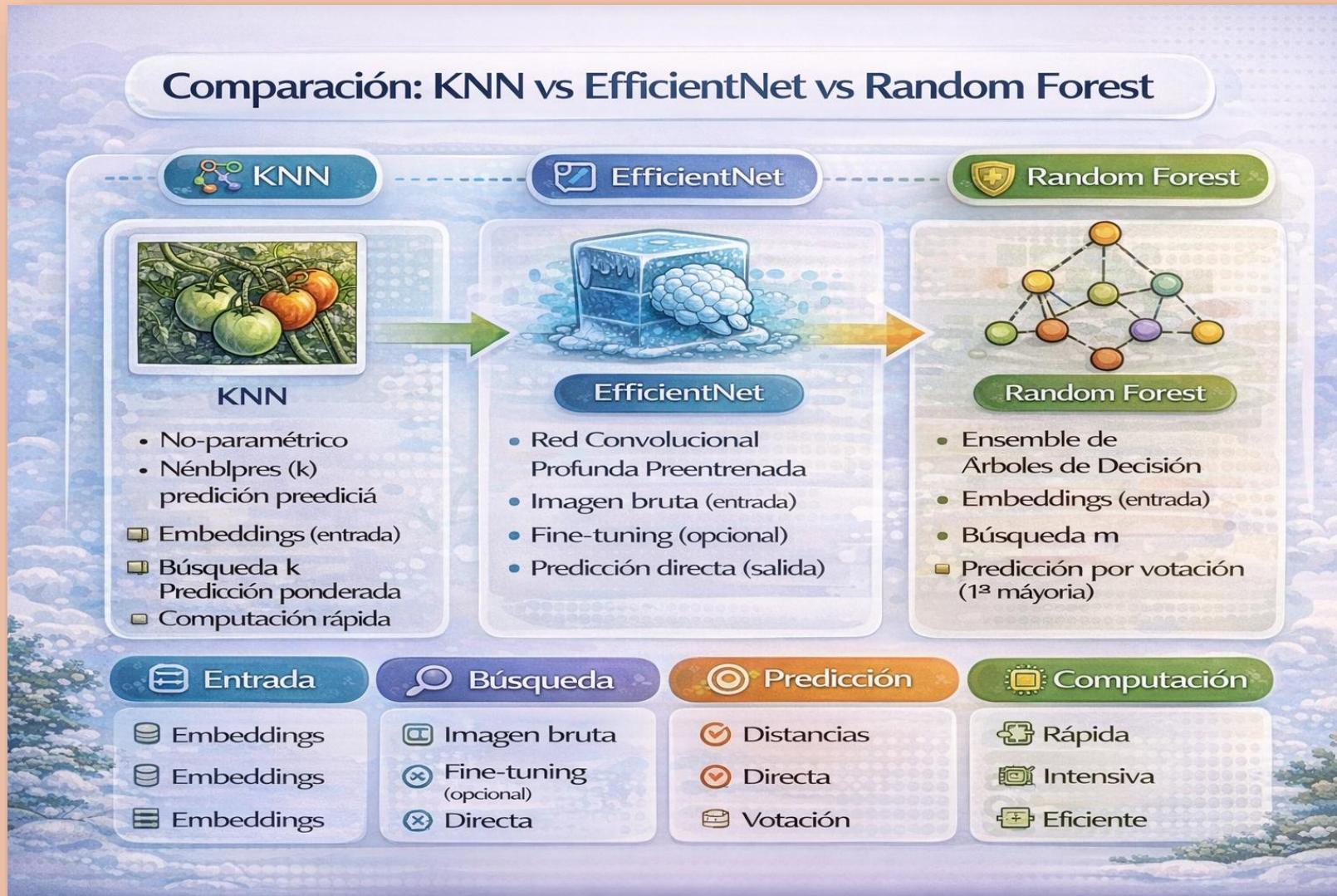
# Proceso de Comparación



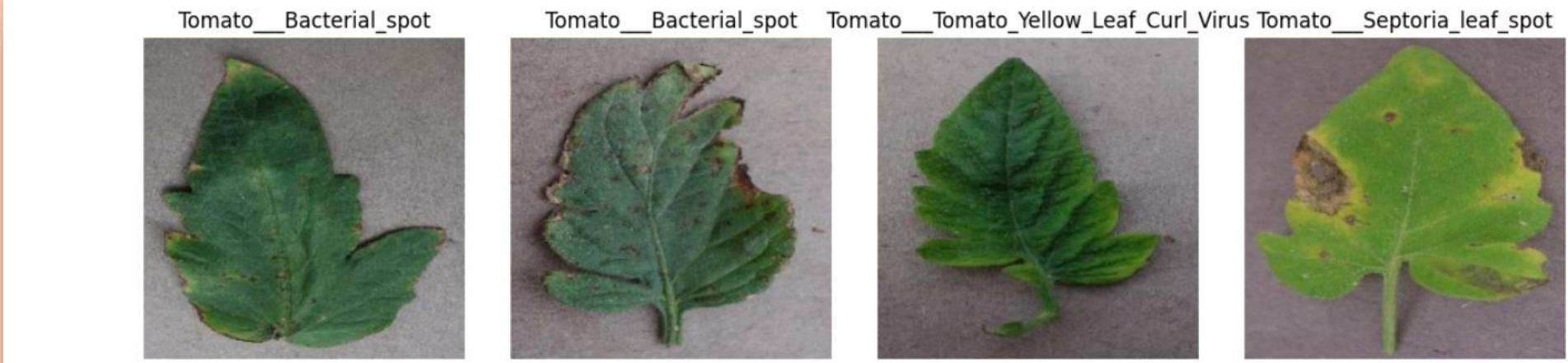
# Proceso de Comparación



# Proceso de Comparación 3 Modelos



# Visualizaciones



Tomato\_Bacterial\_spot

Tomato\_Bacterial\_spot

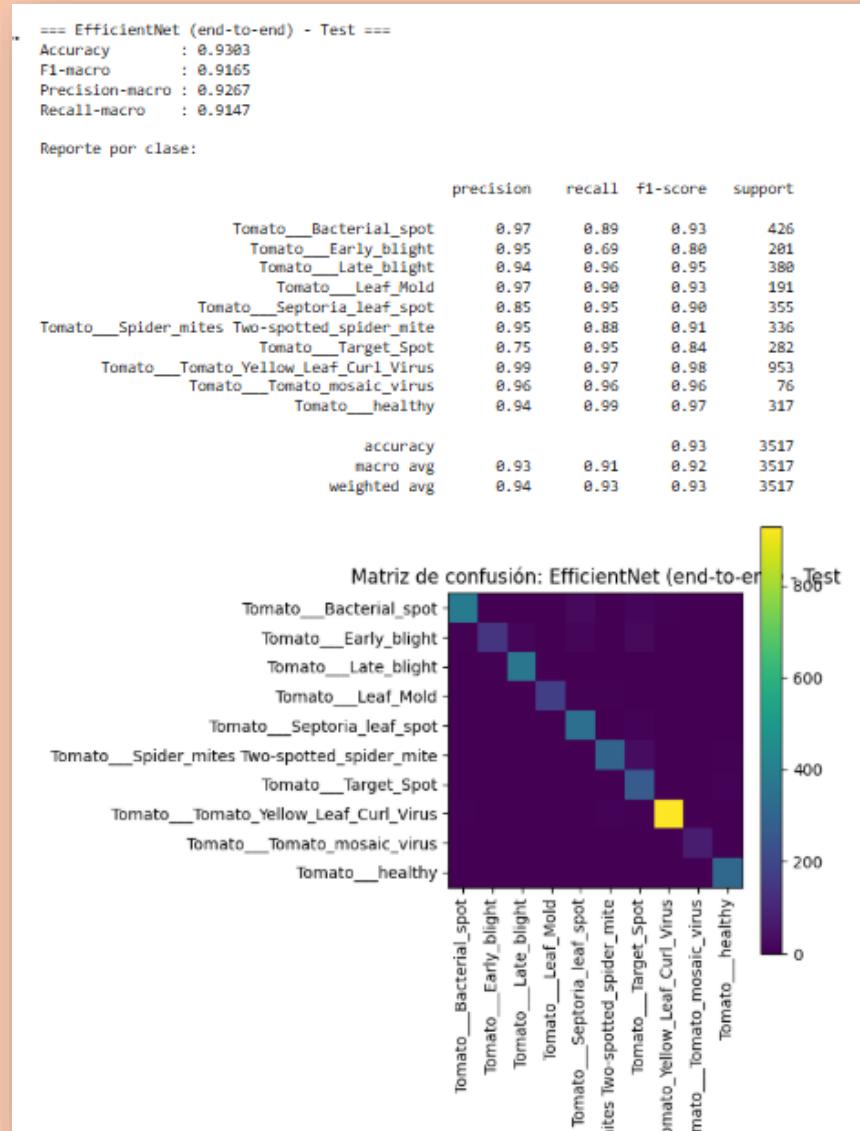
Tomato\_Tomato\_Yellow\_Leaf\_Curl\_Virus

Tomato\_Septoria\_leaf\_spot

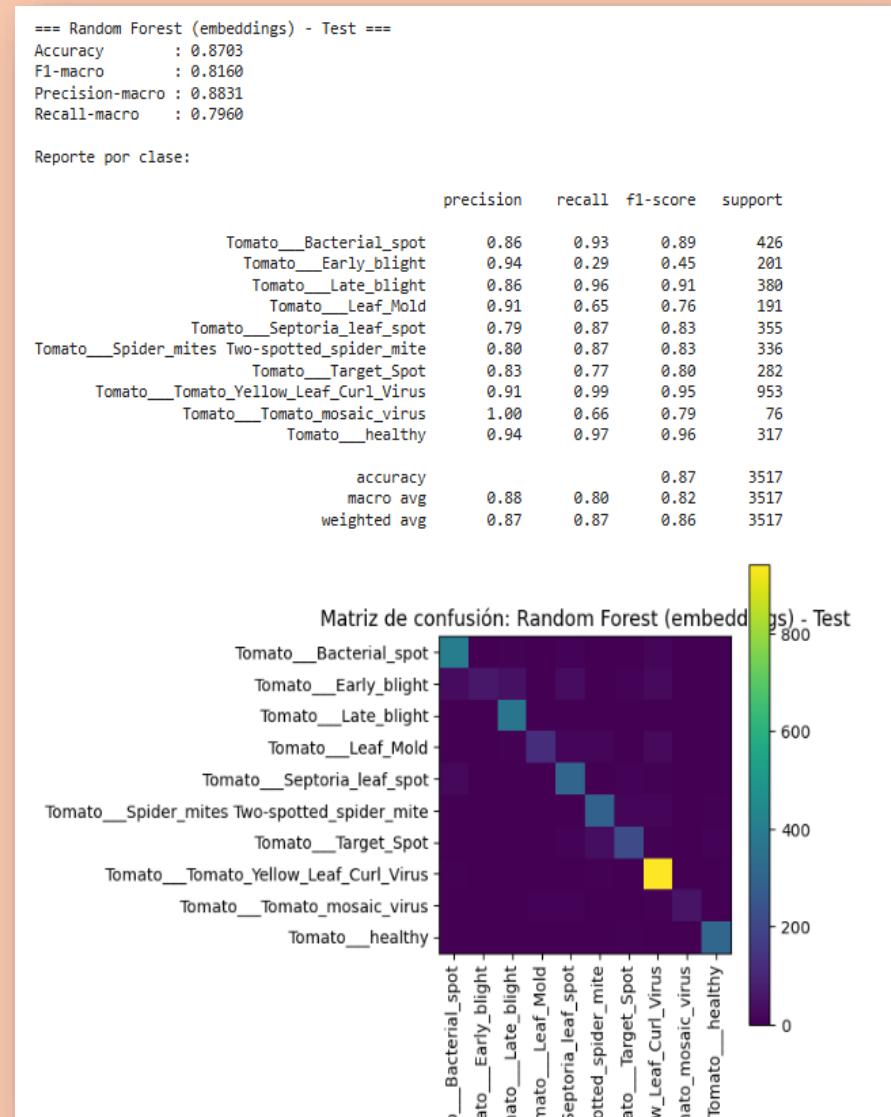
Mapping índice -> clase:

- 0: Tomato\_Bacterial\_spot
- 1: Tomato\_Early\_blight
- 2: Tomato\_Late\_blight
- 3: Tomato\_Leaf\_Mold
- 4: Tomato\_Septoria\_leaf\_spot
- 5: Tomato\_Spider\_mites Two-spotted\_spider\_mite
- 6: Tomato\_Target\_Spot
- 7: Tomato\_Tomato\_Yellow\_Leaf\_Curl\_Virus
- 8: Tomato\_Tomato\_mosaic\_virus
- 9: Tomato\_healthy

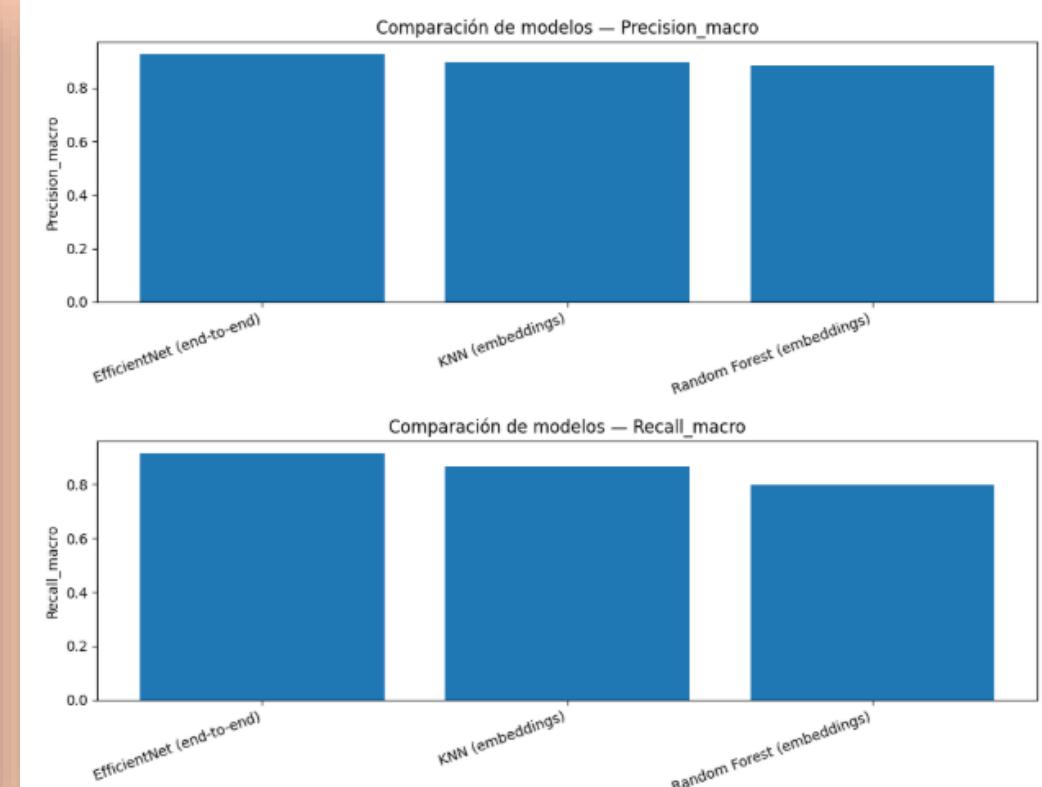
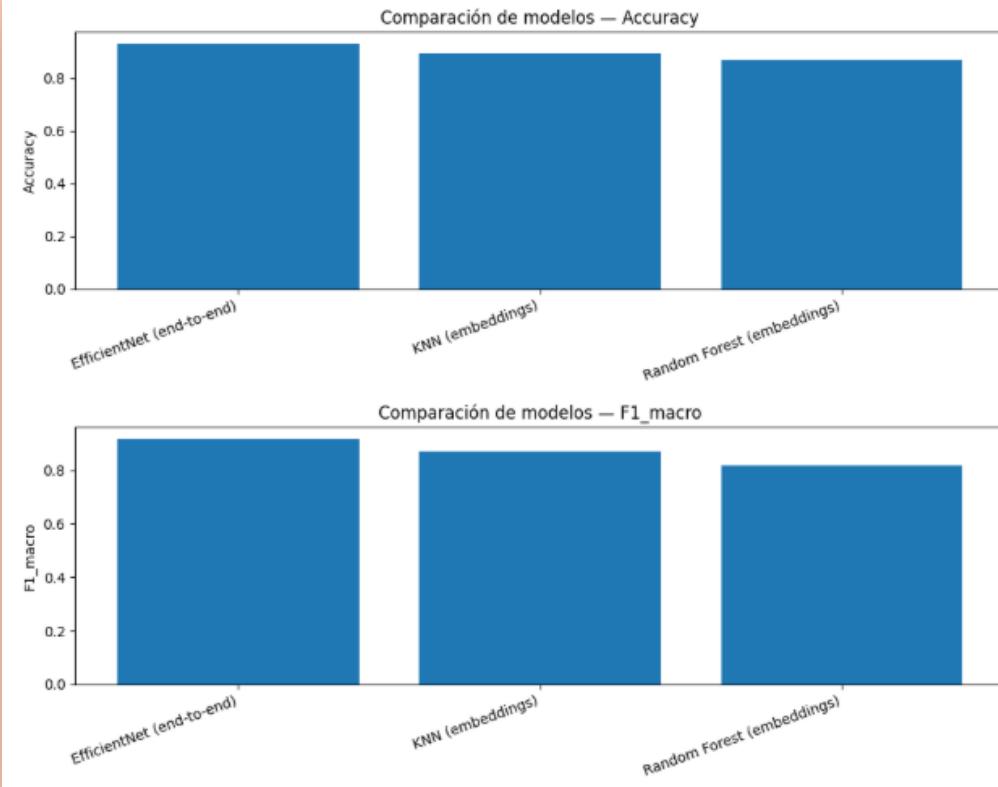
# Visualizaciones



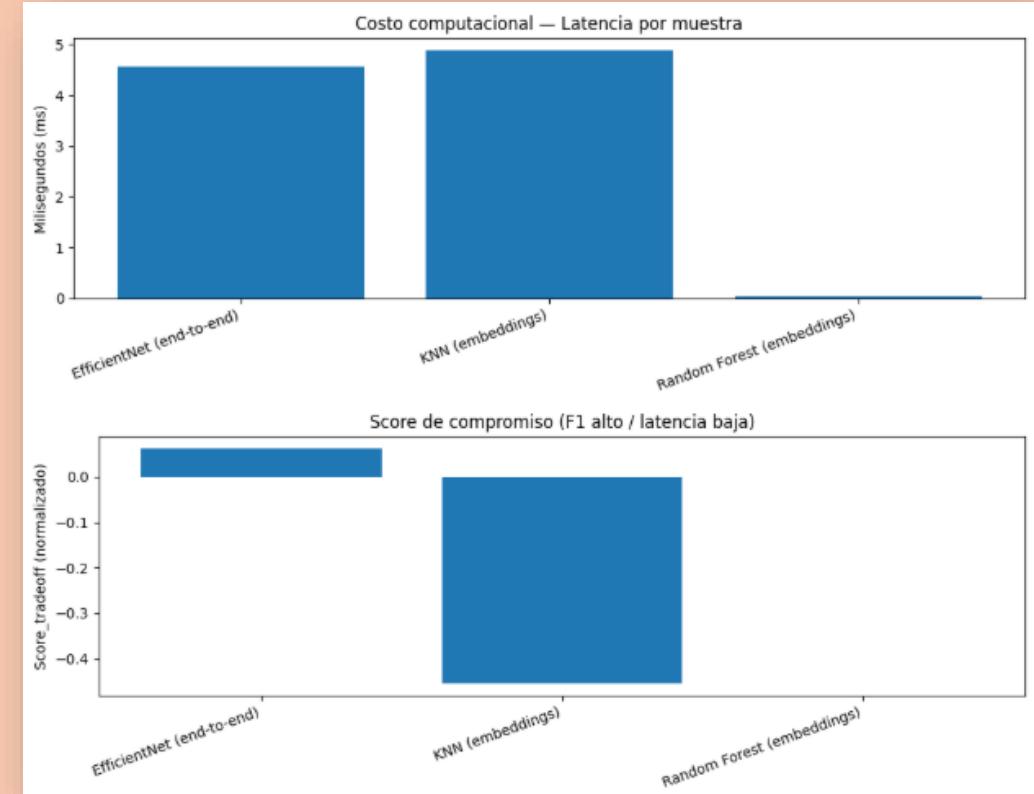
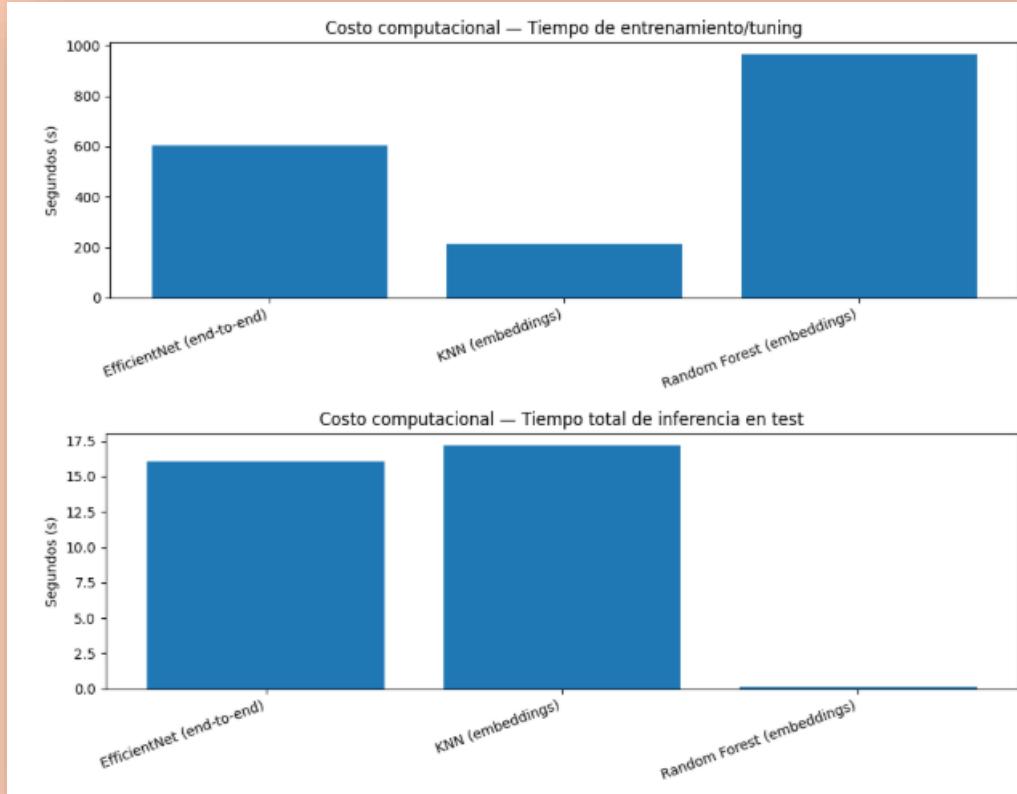
# Visualizaciones



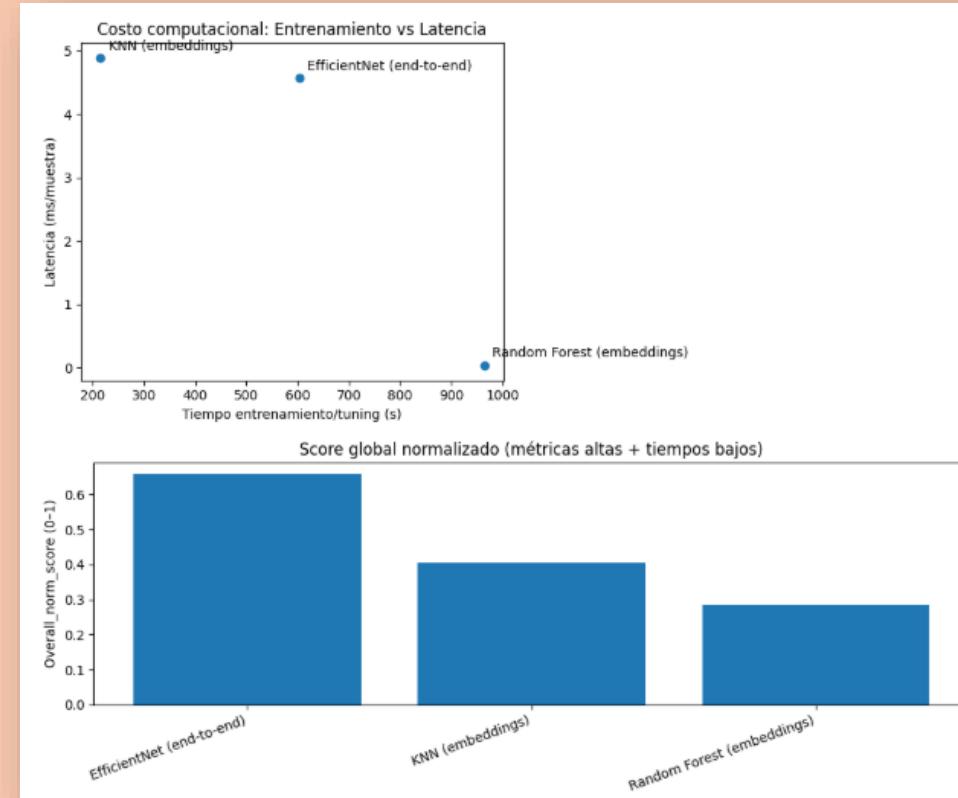
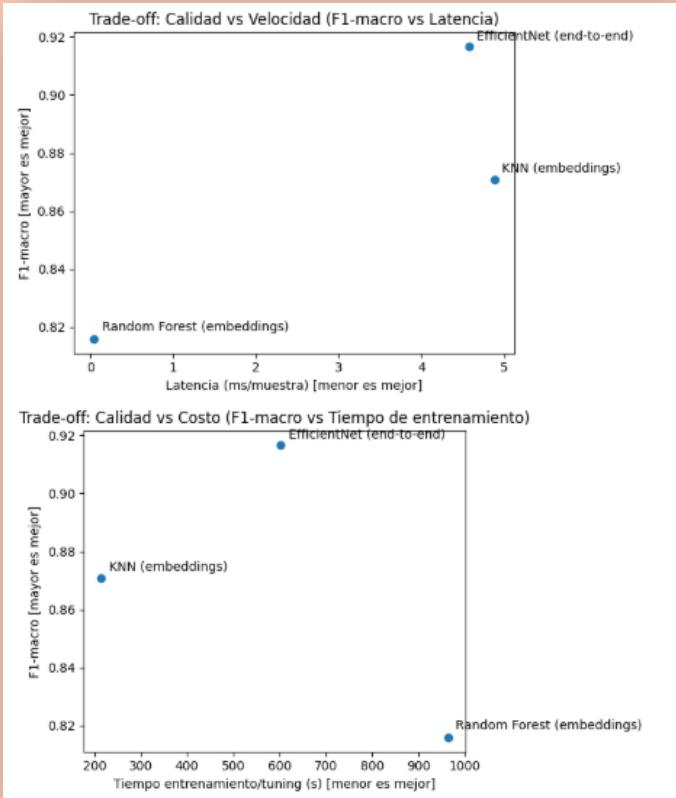
# Visualizaciones



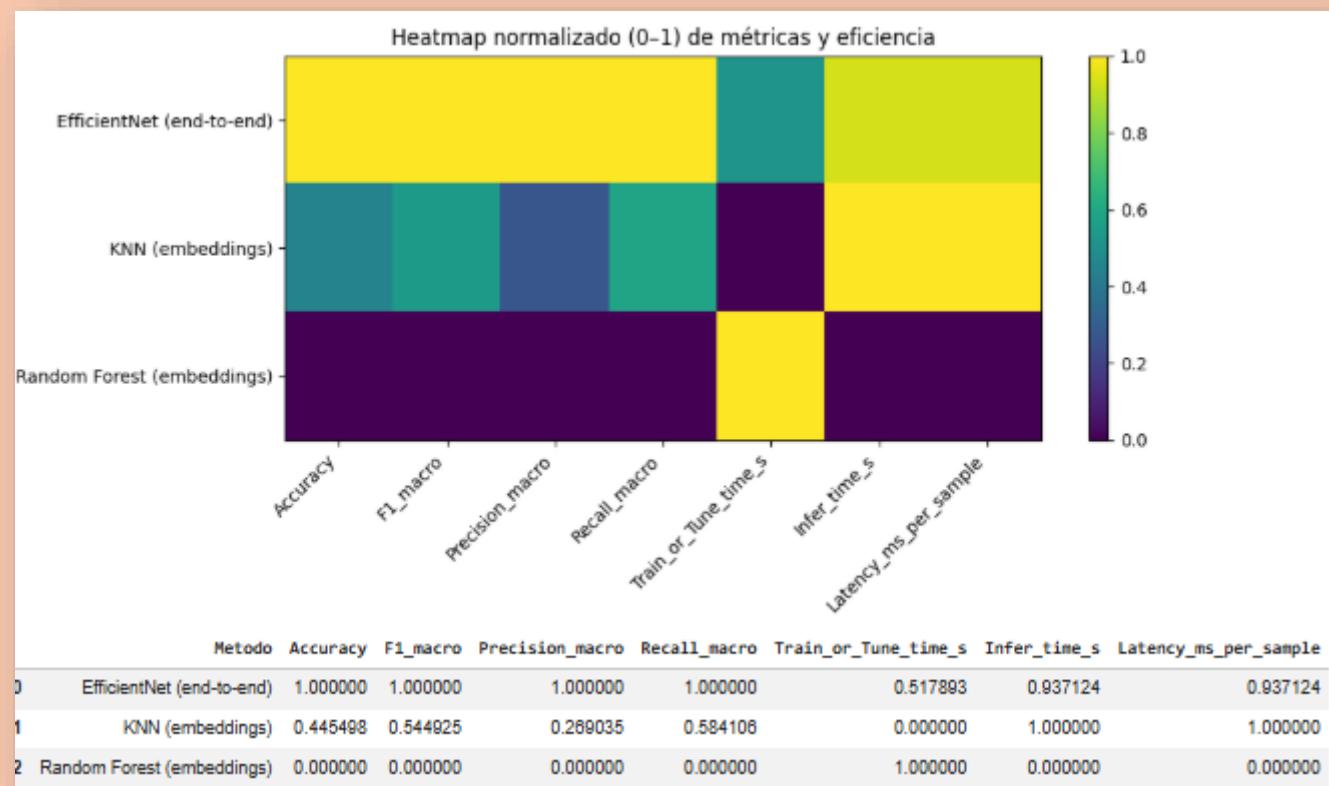
# Visualizaciones



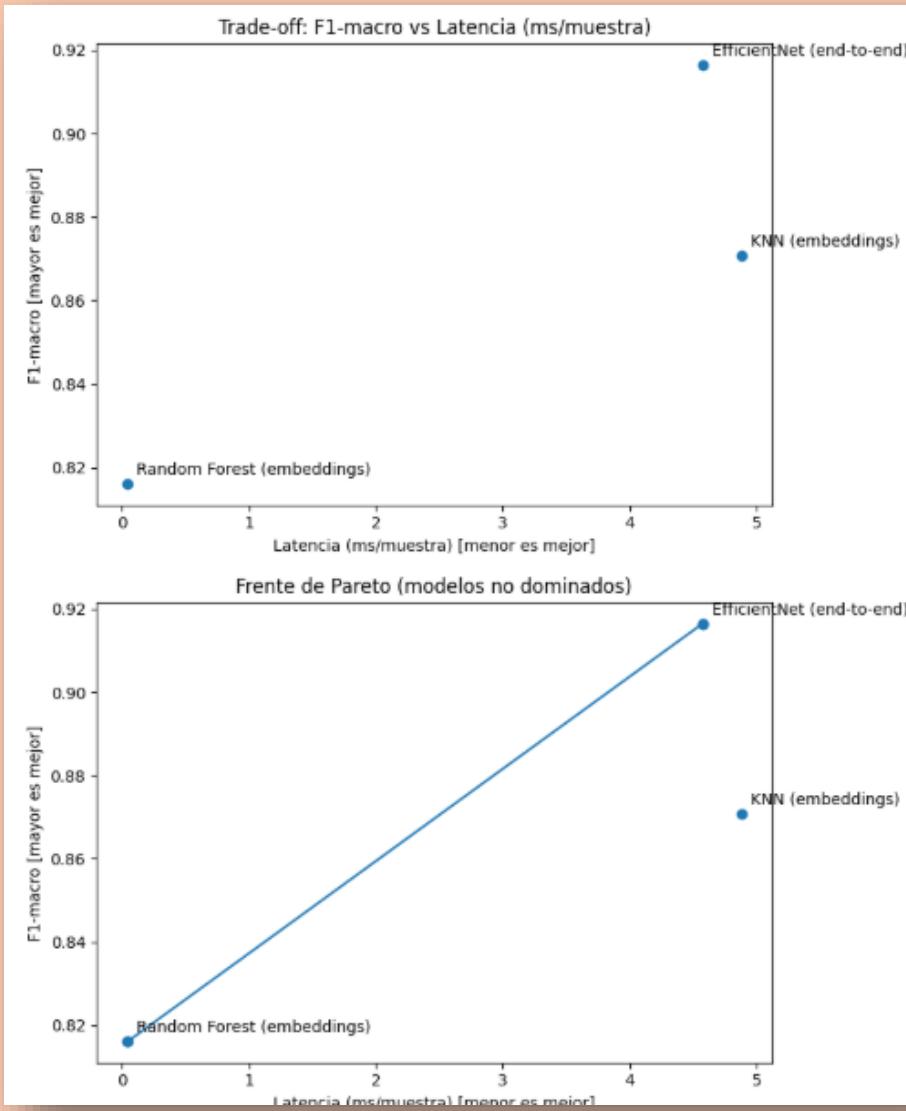
# Visualizaciones



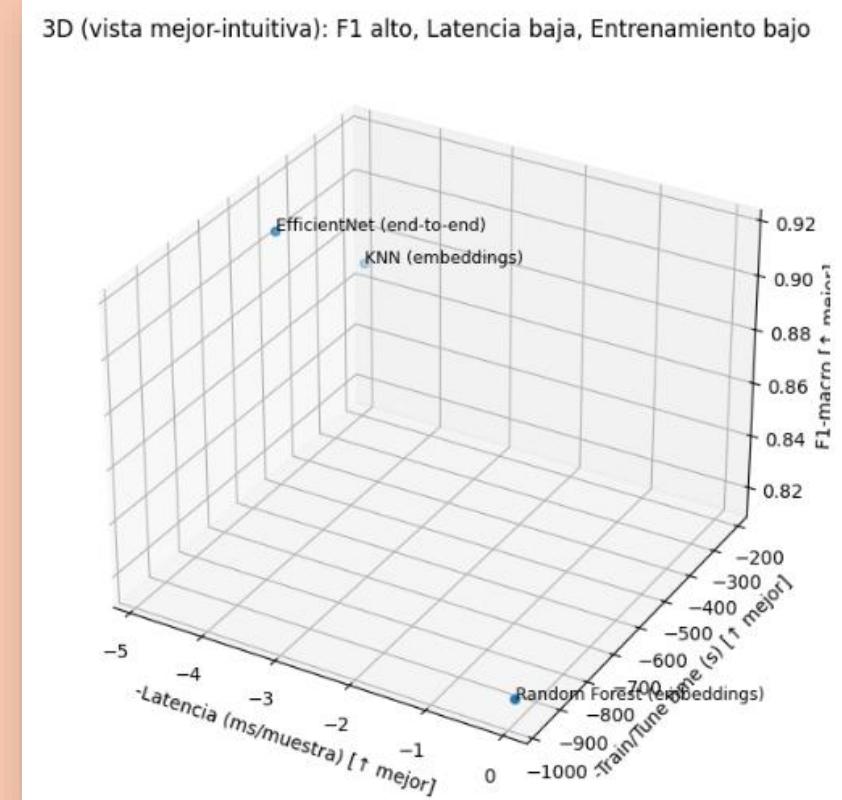
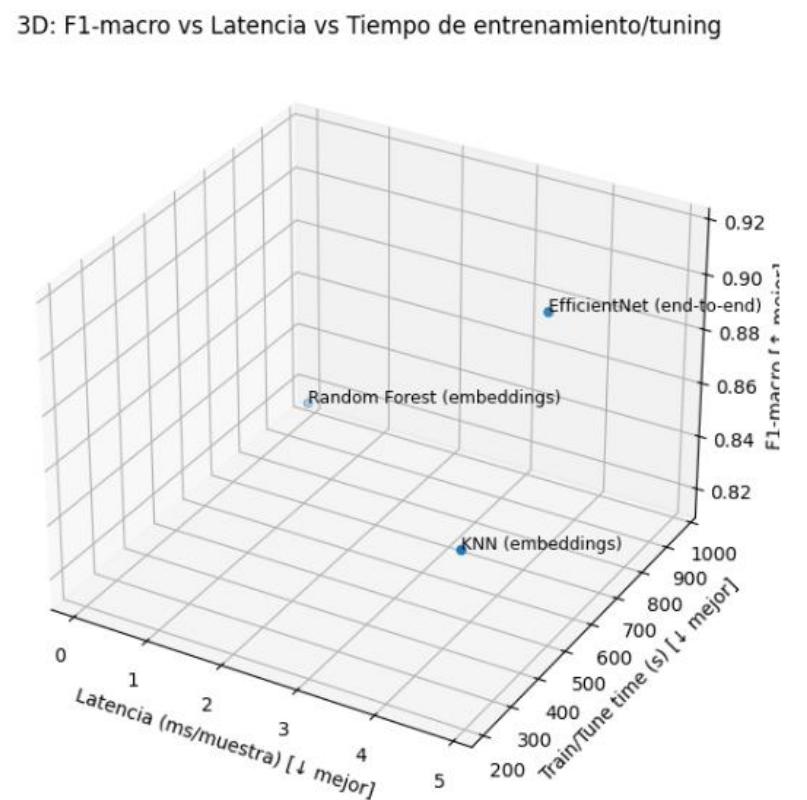
# Visualizaciones



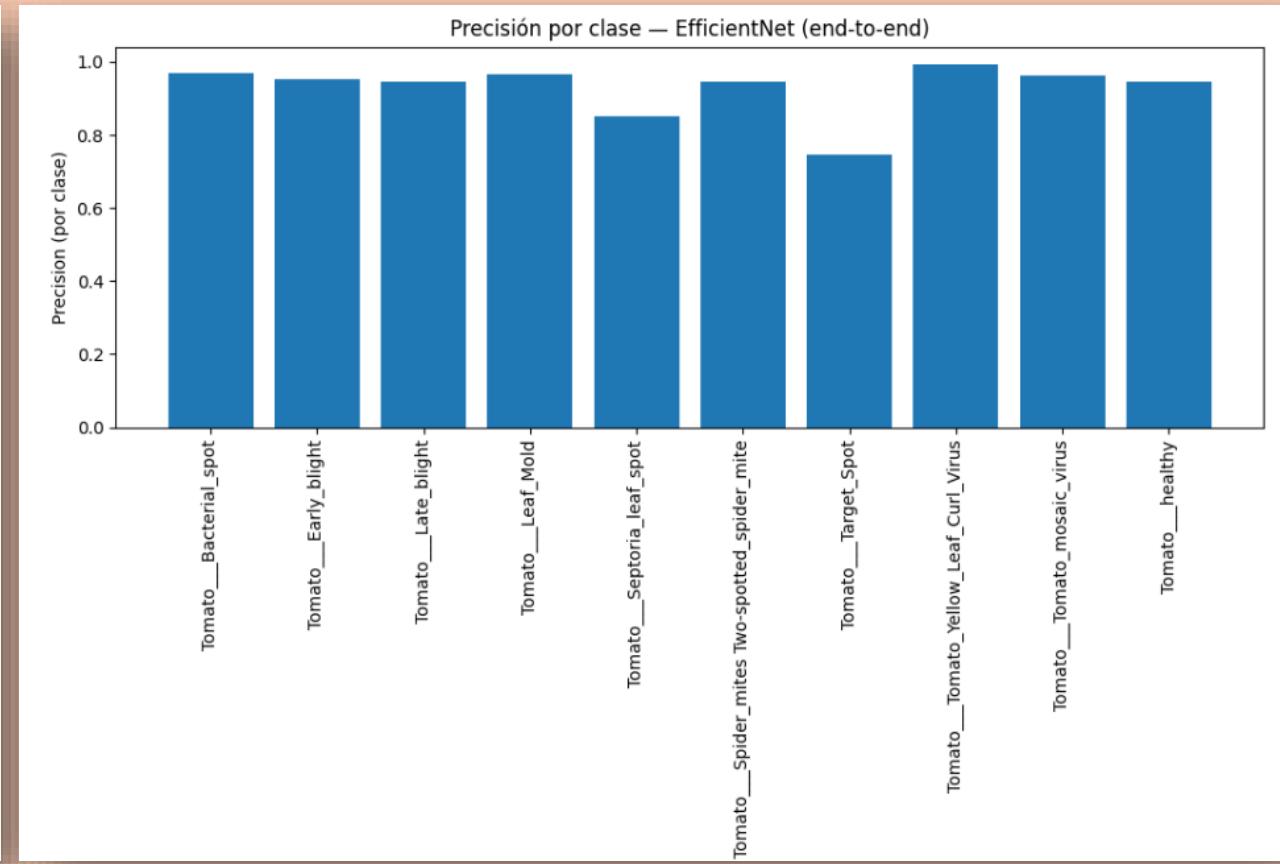
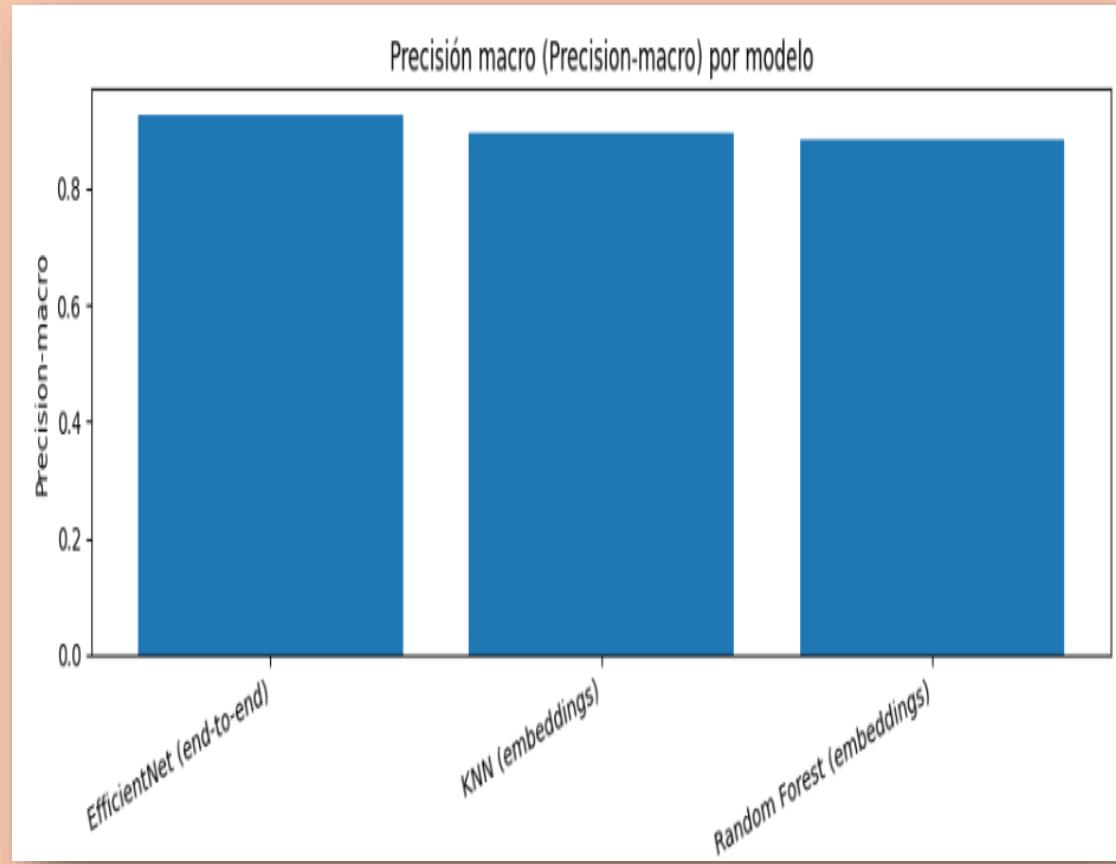
# Visualizaciones



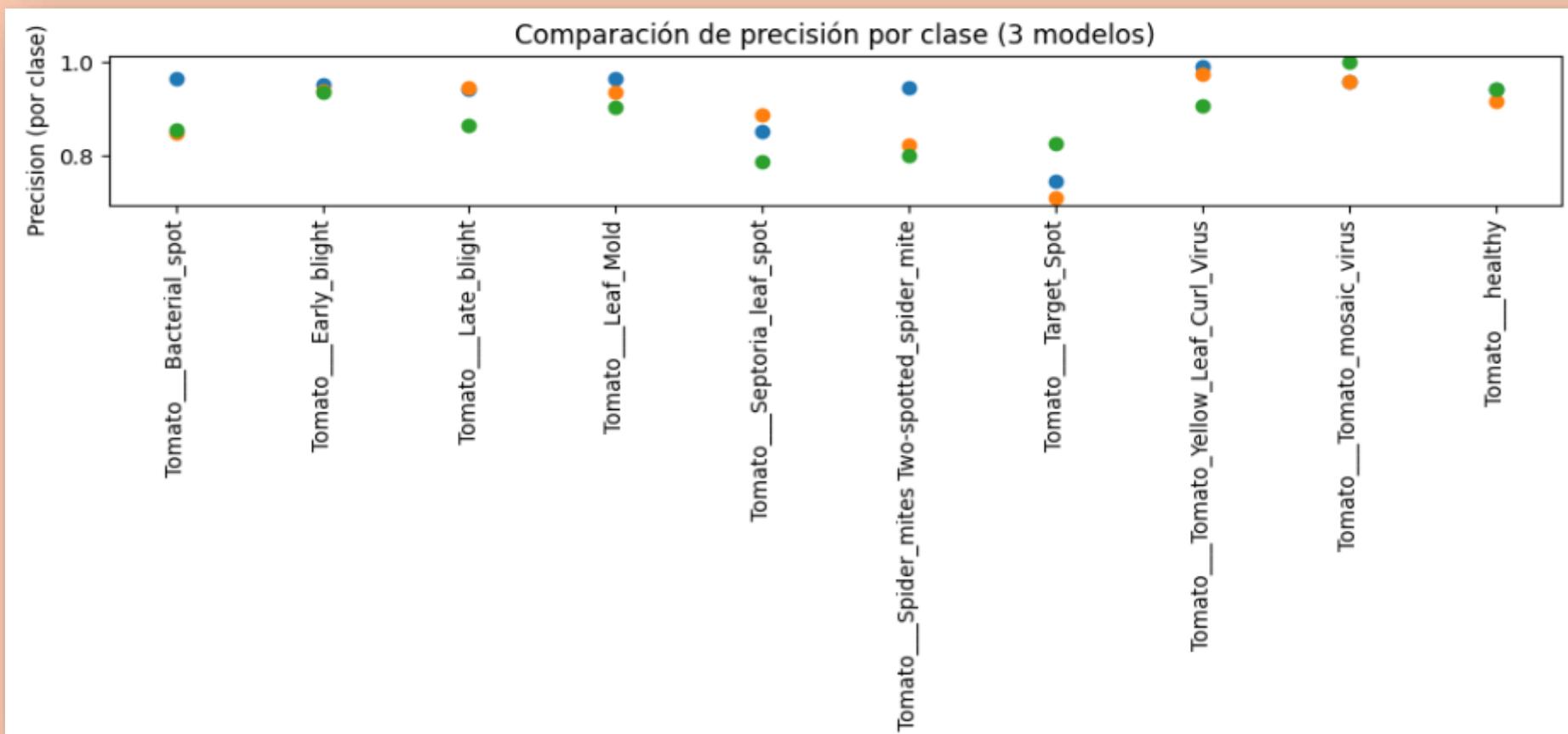
# Visualizaciones



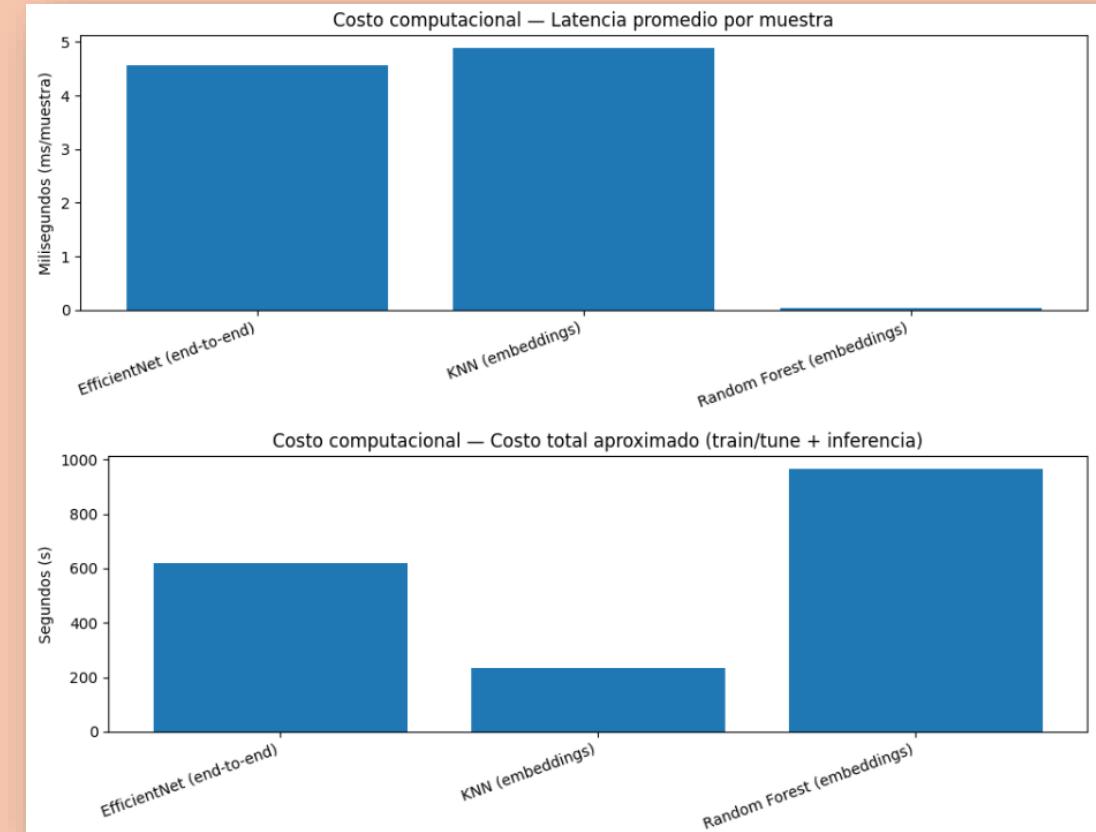
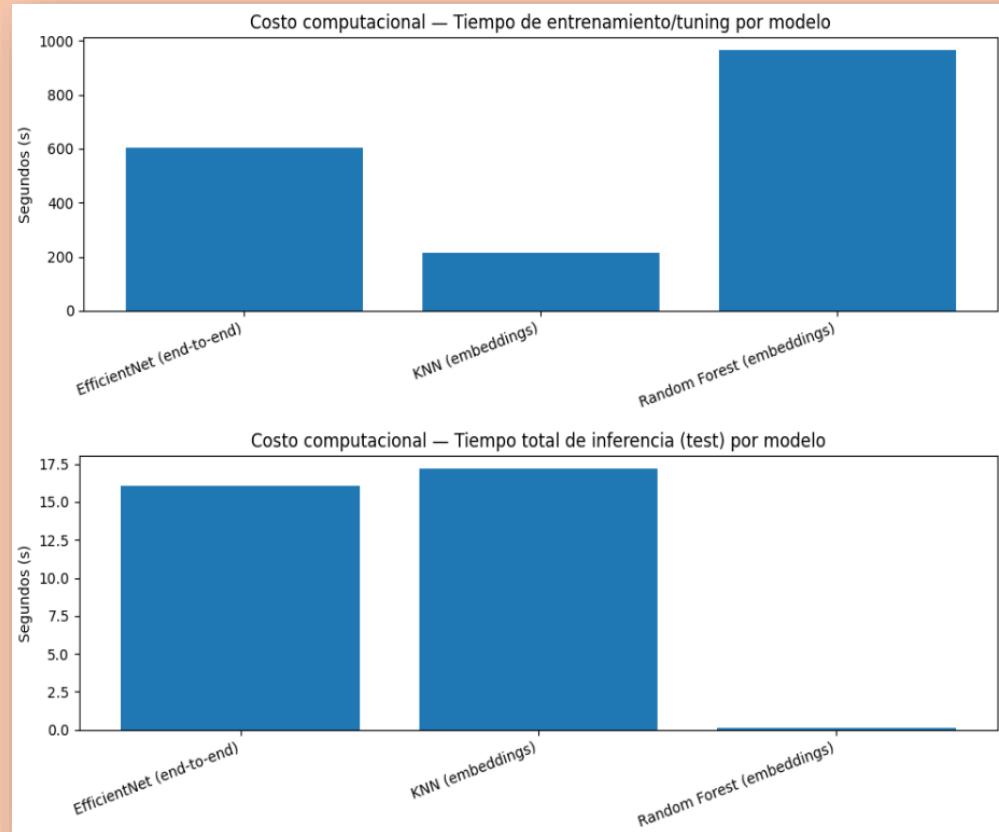
# Visualizaciones



# Visualizaciones



# Visualizaciones



# Visualizaciones

Imagen cargada: septoria\_leaf.JPG



```
== EfficientNet (end-to-end) ==
Predicción: Tomato__Septoria_leaf_spot
Top-5:
- Tomato__Septoria_leaf_spot: 0.9951
- Tomato__Tomato_mosaic_virus: 0.0023
- Tomato__Target_Spot: 0.0022
- Tomato__Bacterial_spot: 0.0002
- Tomato__Leaf_Mold: 0.0001

== KNN (embeddings) ==
Predicción: Tomato__Septoria_leaf_spot
Top-5:
- Tomato__Septoria_leaf_spot: 1.0000
- Tomato__healthy: 0.0000
- Tomato__Tomato_mosaic_virus: 0.0000
- Tomato__Tomato_Yellow_Leaf_Curl_Virus: 0.0000
- Tomato__Target_Spot: 0.0000

== Random Forest (embeddings) ==
Predicción: Tomato__Septoria_leaf_spot
Top-5:
- Tomato__Septoria_leaf_spot: 0.5685
- Tomato__Target_Spot: 0.1051
- Tomato__healthy: 0.0676
- Tomato__Bacterial_spot: 0.0605
- Tomato__Late_blight: 0.0452
```

# Comparación de Modelos

- **Enfermedades Foliares del Tomate**
  - **EfficientNet (Deep Learning):**
    - ✓ Máxima precisión y mejor detección de enfermedades complejas, a costa de mayor tiempo de entrenamiento e inferencia.
  - **Random Forest (embeddings):**
    - ✓ Mejor balance entre desempeño y eficiencia; buen F1-macro con baja latencia, ideal para despliegue práctico.
  - **KNN (embeddings):**
    - ✓ Modelo simple y rápido de entrenar; útil como referencia, pero menos robusto frente a clases similares.

# Comparación de Modelos

## Comparación de Modelos – Enfermedades Foliares del Tomate

**EfficientNet (Deep Learning):**

- Máxima precisión y mejor detección de enfermedades complejas, a costa de **mayor tiempo de entrenamiento** e inferencia.

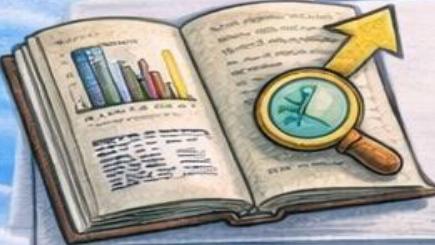
**Random Forest (embeddings):**

- Mejor balance entre desempeño y eficiencia; buen F1-macro con baja latencia, ideal para despliegue práctico.

**KNN (embeddings):**

- Modelo simple y rápido de entrenar; útil como referencia, pero menos robusto frente a clases similares.

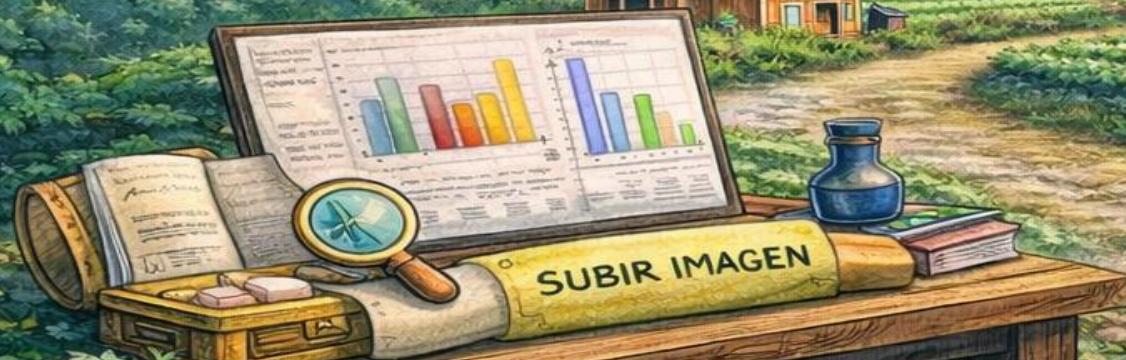
# Análisis Crítico de Resultados y Despliegue del Modelo



**Hacia una Agricultura de Precisión en Chiloé:**  
Sistema Basado en Deep Learning e IoT para la Detección Temprana  
de Enfermedades Foliares en Cultivos Hortícolas



**Enfoque Integral  
y Aplicación Real**



# Análisis Crítico del Modelo

- *En este proyecto no se priorizó únicamente el rendimiento métrico, sino el valor operativo del modelo en un entorno agrícola real.*
- *En agricultura de precisión, un modelo es útil en la medida en que permite detección temprana, mantiene capacidad de generalización en condiciones de campo y entrega soporte confiable a la toma de decisiones agronómicas.*

# Análisis Crítico del Modelo

- En la comparación de modelos, **EfficientNet** exhibió la mayor capacidad predictiva al modelar patrones visuales complejos, aunque con riesgos asociados de sobreajuste y limitada interpretabilidad.
- **Random Forest** sobre embeddings alcanzó un balance favorable entre desempeño y estabilidad, posicionándose como una alternativa más controlable para despliegue.
- **KNN** se utilizó principalmente como baseline y herramienta diagnóstica, mostrando alta sensibilidad al ruido y al desbalance de clases.

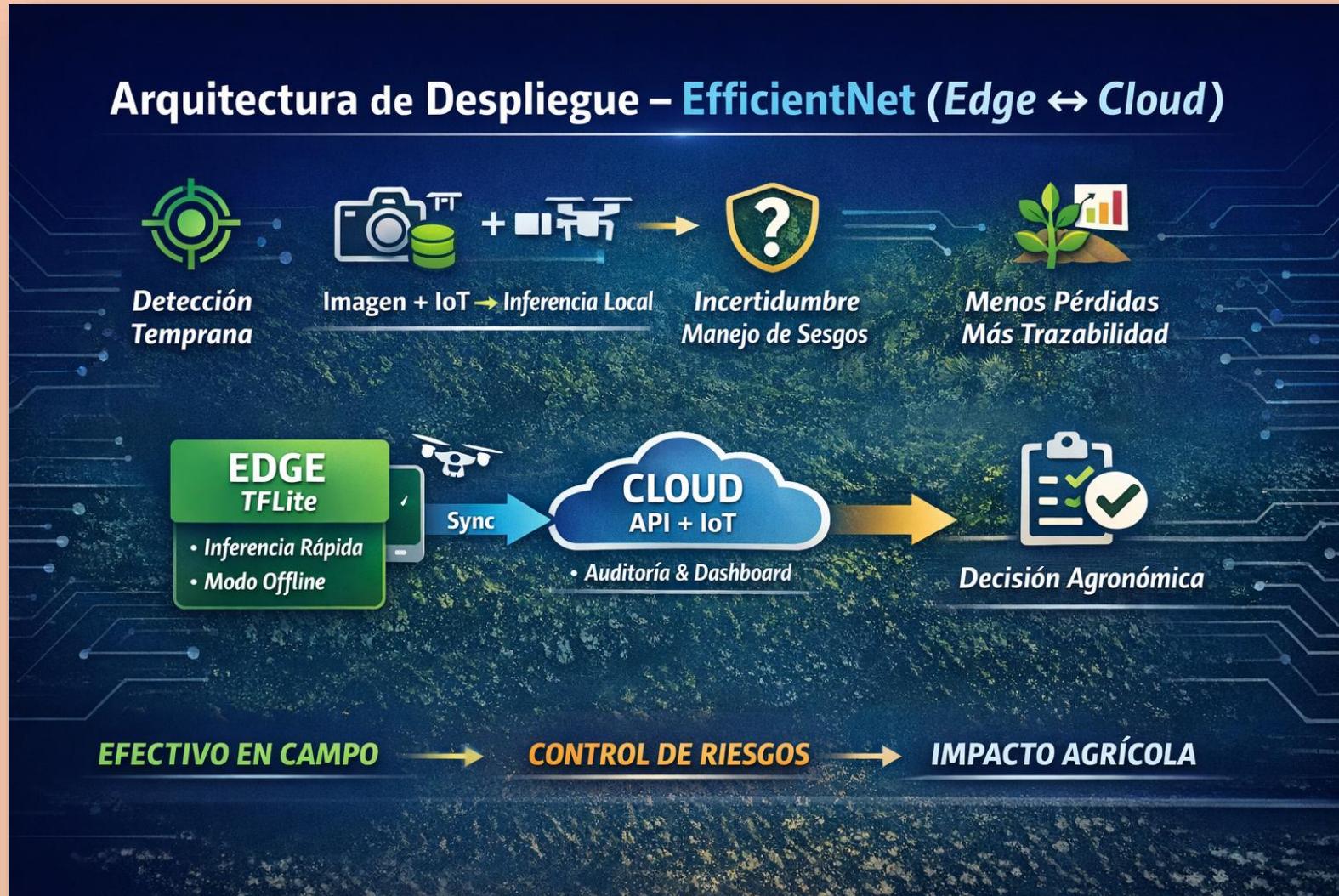
# Análisis Crítico del Modelo

- Se *identificaron limitaciones críticas* asociadas a sesgos de *iluminación y fondo, desbalance de clases, ruido de etiquetado y riesgos de pérdida de generalización* frente a *nuevos predios, dispositivos o condiciones estacionales*.
- *Desde una perspectiva operativa, los falsos negativos favorecen la propagación de enfermedades, mientras que los falsos positivos incrementan costos y erosionan la confianza en el sistema.*

# Análisis Crítico del Modelo

- Se concluye que el *impacto del sistema* está determinado menos por el rendimiento en test y más por su robustez en condiciones reales, el manejo explícito de la incertidumbre y la incorporación del contexto IoT, consolidando a **EfficientNet** como el modelo de mayor capacidad predictiva, a **Random Forest** como la alternativa más balanceada y al sistema como una herramienta efectiva de apoyo a la decisión agrícola.

# Despliegue Operacional del Modelo



# Despliegue Operacional del Modelo

- **Detección Temprana**
  - Identificación precoz de enfermedades para intervenir antes de pérdidas productivas.
- **Imagen + IoT → Inferencia Local**
  - Procesamiento en terreno de imágenes y contexto ambiental con baja latencia.
- **Incertidumbre / Sesgos**
  - Umbrales de confianza y controles para evitar decisiones no confiables.
- **Menos Pérdidas / Trazabilidad**
  - Reducción de daños y generación de evidencia histórica por lote.

# Despliegue Operacional del Modelo

- **EDGE – TFLite**
  - Inferencia rápida y autónoma, incluso sin conectividad.
- **Sync**
  - Sincronización diferida de datos hacia la nube.
- **CLOUD – API + IoT**
  - Auditoría, visualización y mejora continua del sistema.
- **Decisión Agronómica**
  - Soporte directo a la acción técnica en campo.

# Despliegue Operacional del Modelo

- **Efectivo en Campo**
  - Diseñado para operar bajo condiciones reales y variables.
- **Control de Riesgos**
  - Gestión explícita del costo de error y la incertidumbre.
- **Impacto Agrícola**
  - Optimización del manejo sanitario y sostenibilidad productiva.

# Hacia una Agricultura de Precisión en Chiloé: Sistema Basado en Deep Learning e IoT para la Detección Temprana de Enfermedades Foliares en Cultivos Hortícolas

