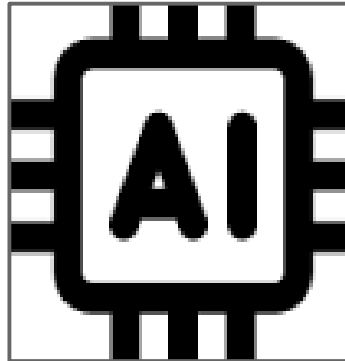


Introducción a la Inteligencia Artificial



Computer Vision

Objetivo

El objetivo de este módulo es introducir a los estudiantes en los fundamentos y aplicaciones prácticas de la Visión por Computadora (Computer Vision), una rama fundamental de la Inteligencia Artificial.

Se busca explorar desde los conceptos básicos del procesamiento de imágenes hasta las arquitecturas como CNNs y Transformers, con énfasis en aplicaciones del mundo real.

¿Qué es Computer Vision?

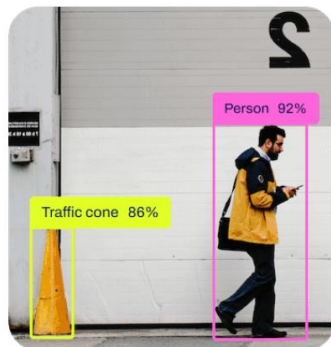
- *Rama de la IA que permite a las computadoras "ver" y entender contenido visual*
- *Busca replicar la capacidad humana de interpretar información visual*
- *Aplicaciones en múltiples industrias: medicina, automoción, seguridad, retail, agro*

Posibilidades

Classify



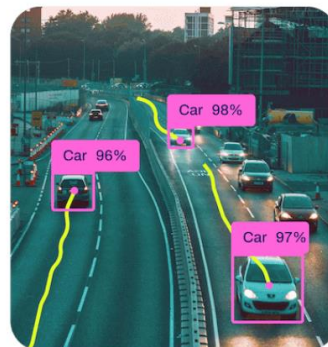
Detect



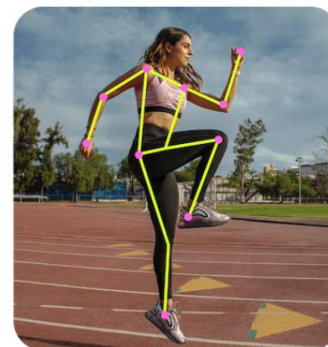
Segment



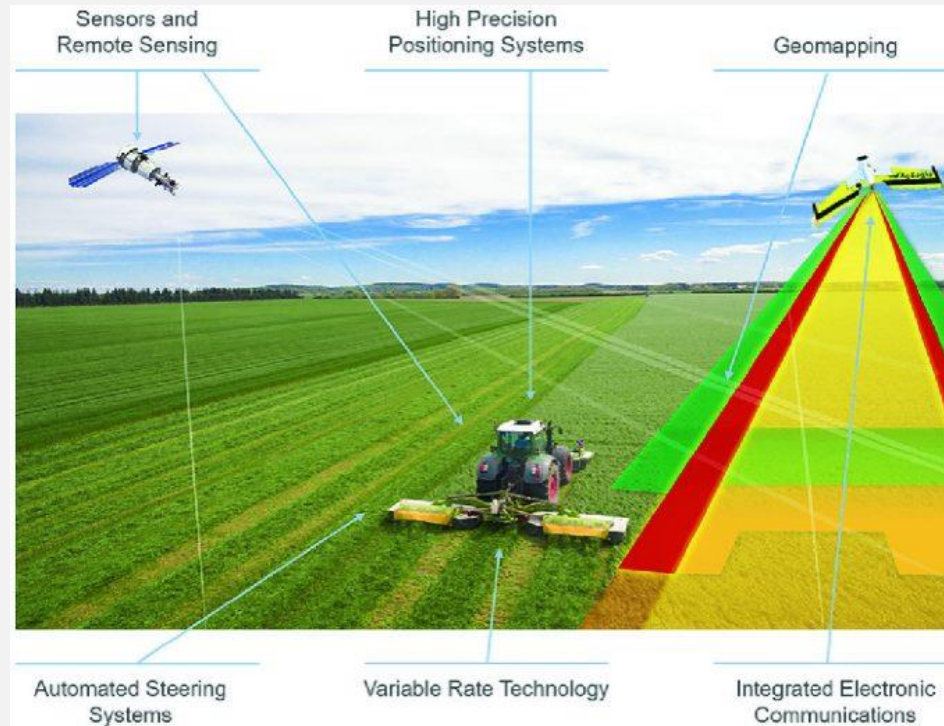
Track



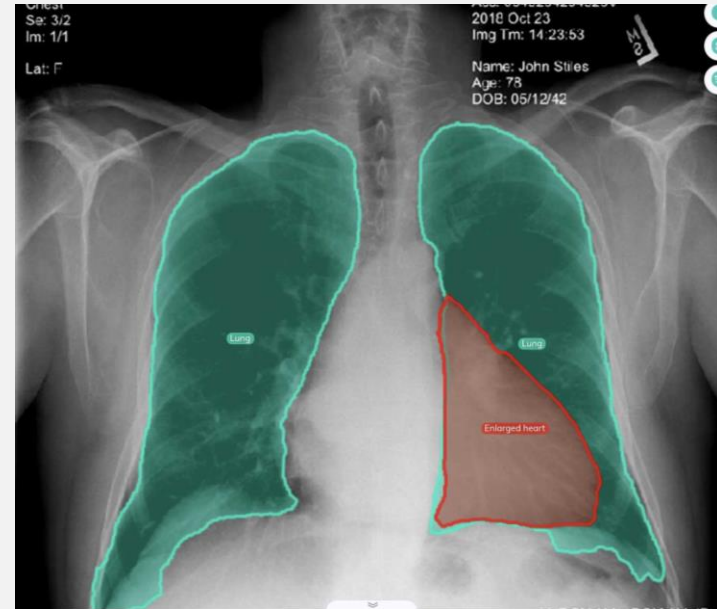
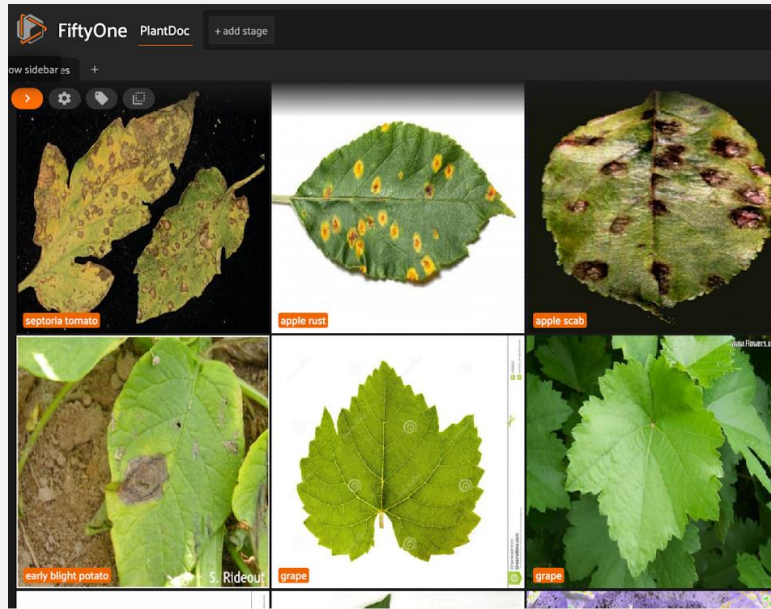
Pose



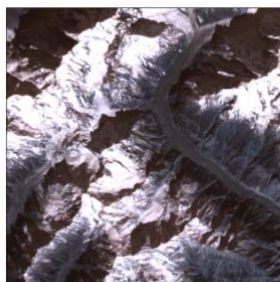
Ejemplo: predicción de cultivos y rindes



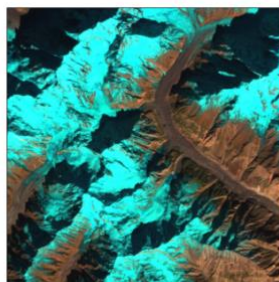
Ejemplo: detección de enfermedades



Ejemplo: monitoreo de glaciares



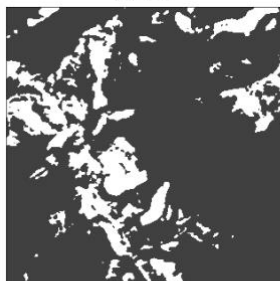
(a) RGB



(b) B5 B4 B2 Color Composite



(c) Glacier Labels



(d) Mask of Binary Class
IoU = 0.37

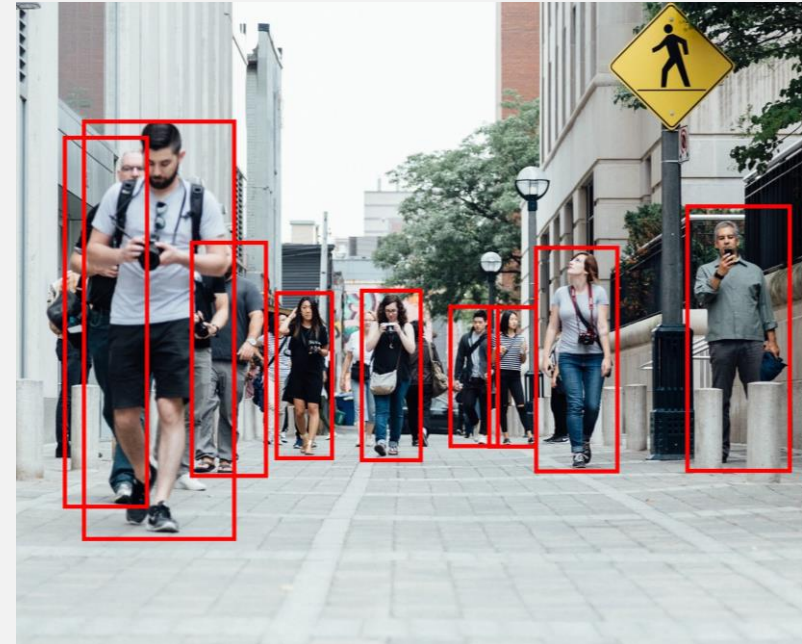
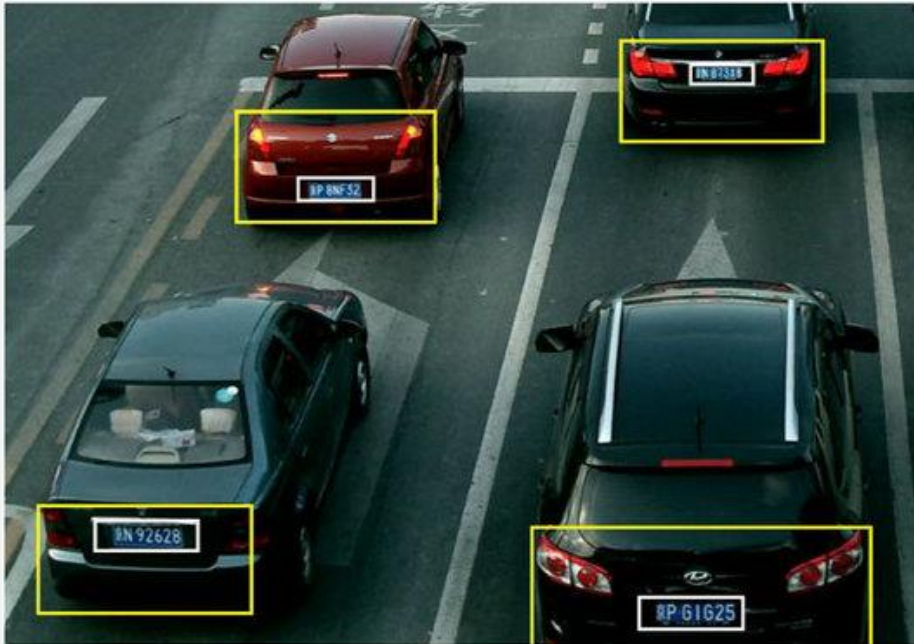


(e) Mask of Multiple Classes
IoU = 0.44



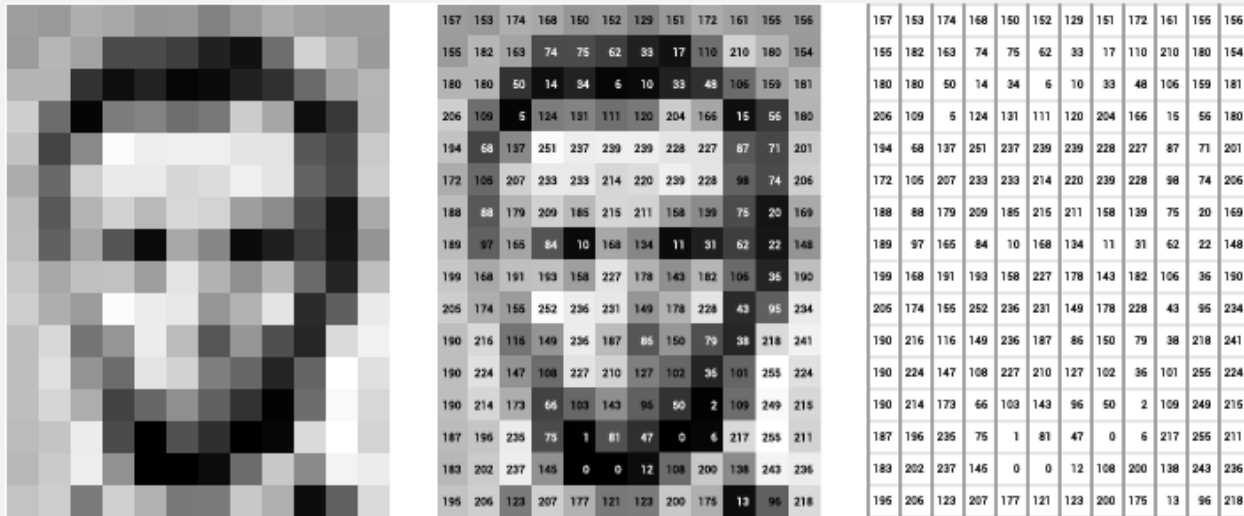
(f) Mask of Combined Two Models
IoU = 0.41

Ejemplo: tracking de objetos/personas



Imágenes digitales

Las imágenes son matrices de píxeles



Fuente: Towards Data Science

Contacto: ia@frh.utn.edu.ar

Imágenes digitales

Que luego se convierten en vectores

```
{157, 153, 174, 168, 150, 152, 129, 151, 172, 161, 155, 156,  
155, 182, 163, 74, 75, 62, 33, 17, 110, 210, 180, 154,  
180, 180, 50, 14, 34, 6, 10, 33, 48, 106, 159, 181,  
206, 109, 5, 124, 131, 111, 120, 204, 166, 15, 56, 180,  
194, 68, 137, 251, 237, 239, 239, 228, 227, 87, 71, 201,  
172, 105, 207, 233, 233, 214, 220, 239, 228, 98, 74, 206,  
188, 88, 179, 209, 185, 215, 211, 158, 139, 75, 20, 169,  
189, 97, 165, 84, 10, 168, 134, 11, 31, 62, 22, 148,  
199, 168, 191, 193, 158, 227, 178, 143, 182, 106, 36, 190,  
205, 174, 155, 252, 236, 231, 149, 178, 228, 43, 95, 234,  
190, 216, 116, 149, 236, 187, 86, 150, 79, 38, 218, 241,  
190, 224, 147, 108, 227, 210, 127, 102, 36, 101, 255, 224,  
190, 214, 173, 66, 103, 143, 96, 50, 2, 109, 249, 215,  
187, 196, 235, 75, 1, 81, 47, 0, 6, 217, 255, 211,  
183, 202, 237, 145, 0, 0, 12, 108, 200, 138, 243, 236,  
195, 206, 123, 207, 177, 121, 123, 200, 175, 13, 96, 218};
```

Preprocesamiento es crucial

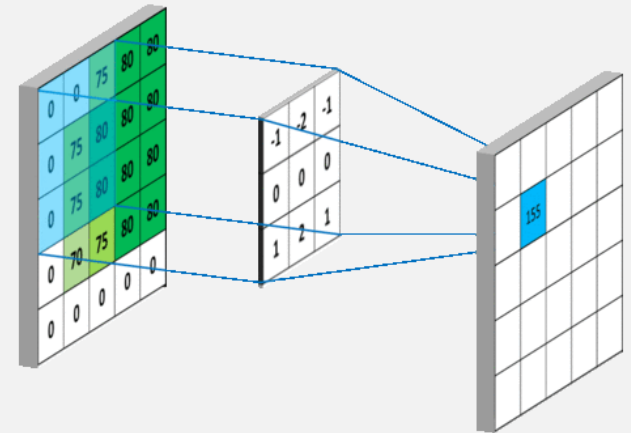
- *Normalización*
- *Redimensionamiento*
- *Aumentación de datos*

¿Cómo procesamos imágenes?

- *Redes Neuronales Convolucionales (CNN)*
- *Transformers*
- *Transfer Learning*
- *Bonus: Hugging Face*

CNN

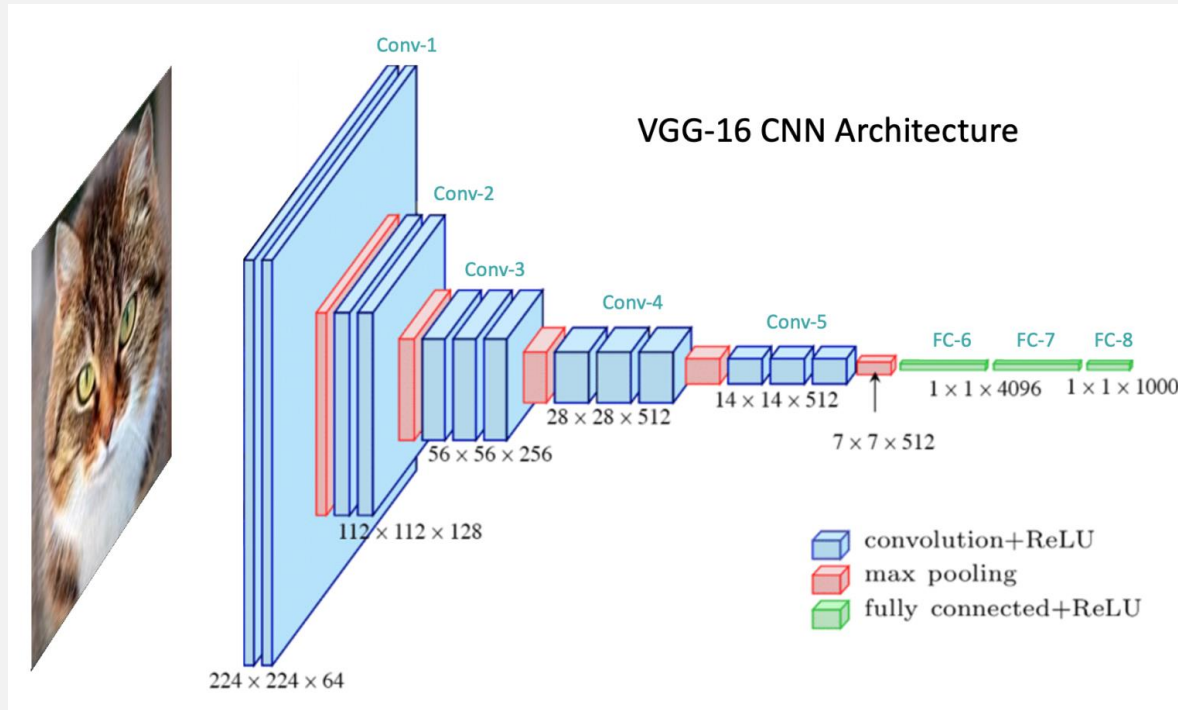
- *Arquitectura especializada para procesar datos en grid*
- *Componentes principales:*
 - *Capas convolucionales*
 - *Pooling*
 - *Capas completamente conectadas*



¿Cómo funcionan las CNN?

- *Las convoluciones detectan patrones locales*
- *Jerarquía de características:*
 - *Primeras capas: bordes, texturas*
 - *Capas intermedias: formas, patrones*
 - *Últimas capas: objetos completos*

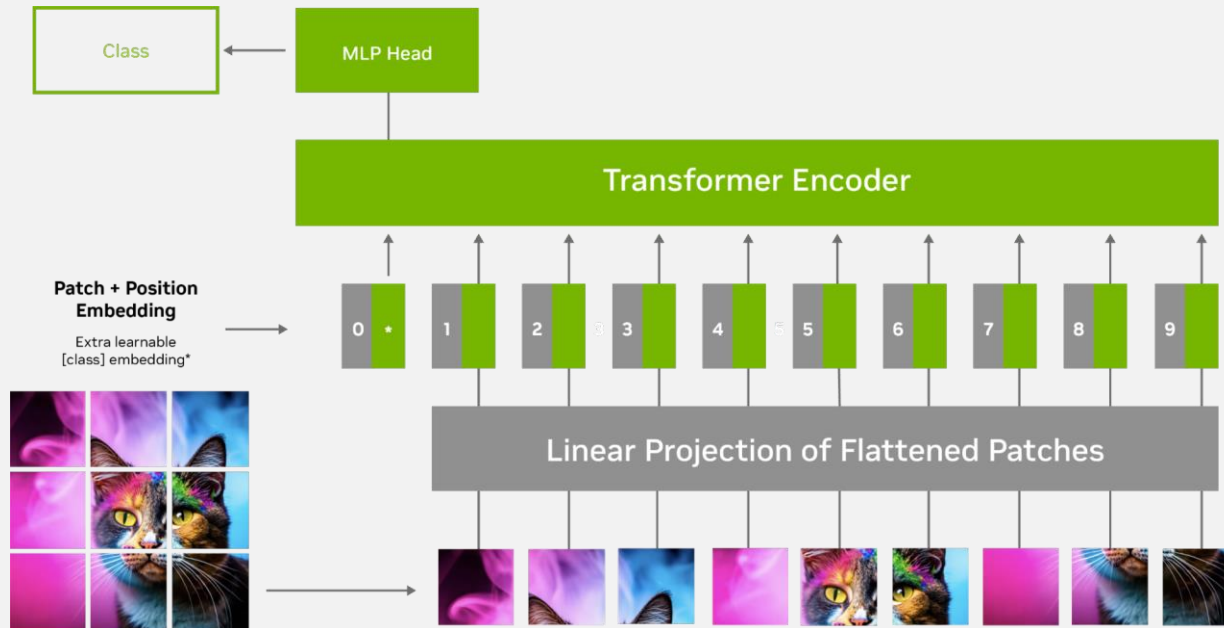
¿Cómo funcionan las CNN?



Transformers en Computer Vision

- *Originalmente diseñados para NLP*
- *Vision Transformer (ViT):*
 - *División de imagen en parches*
 - *Paralelización de tareas*
 - *Contexto global*
 - *Estado del arte en muchas tareas*

Transformers en Computer Vision



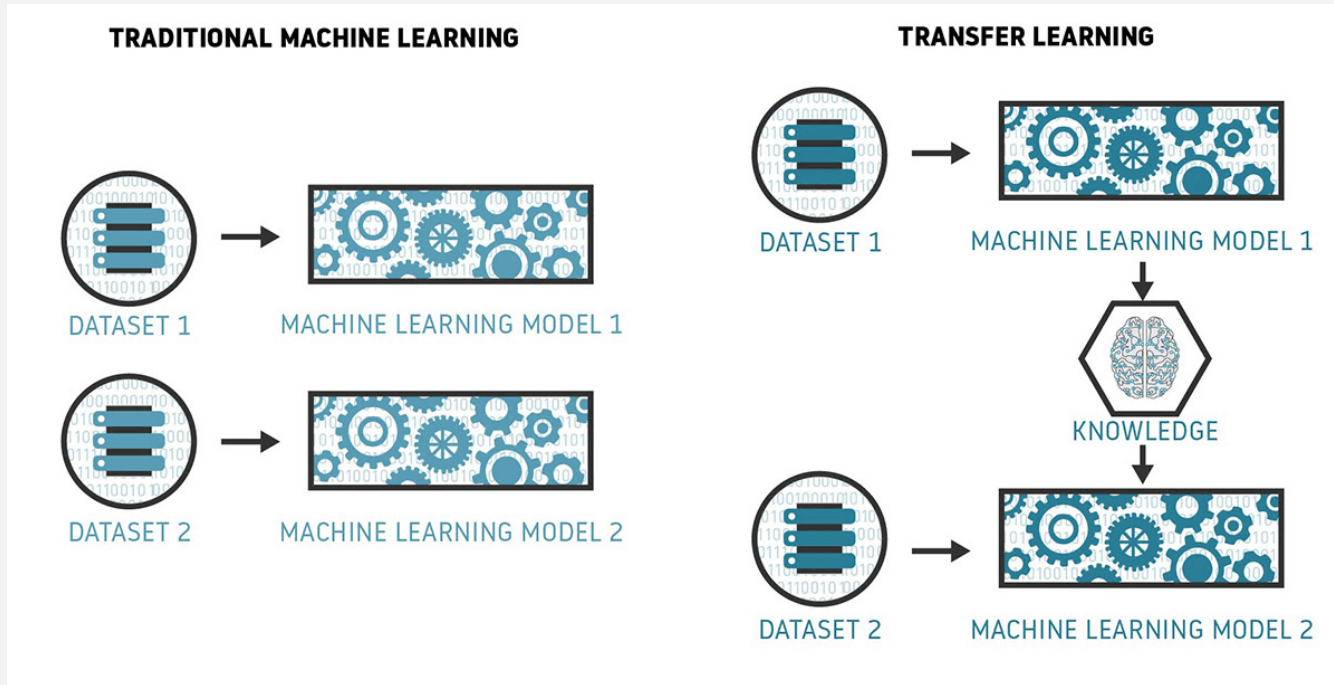
Fuente: NVIDIA

Contacto: ia@frh.utn.edu.ar

Transfer Learning

- *Reutilización de modelos pre-entrenados*
- *Ventajas:*
 - *Menor tiempo de entrenamiento*
 - *Menos datos necesarios*
 - *Mejor rendimiento inicial*

Transfer Learning



Fuente: datascience.aero

Contacto: ia@frh.utn.edu.ar

Técnicas de Transfer Learning

- *Feature Extraction*
 - *Congelar capas pre-entrenadas*
 - *Entrenar solo las últimas capas*
- *Fine-tuning*
 - *Ajustar todo el modelo*
 - *Learning rate bajo*

Hugging Face: Hub de IA

- *Plataforma líder para modelos de ML*
- *Beneficios:*
 - *Modelos pre-entrenados listos para usar*
 - *Documentación extensa*
 - *Comunidad activa*

Desafíos

- *Robustez ante variaciones*
- *Eficiencia computacional*
- *Interpretabilidad*
- *Privacidad y ética*
- *Integración con otros campos de IA (multimodalidad)*

Como siempre, los datos reales son imperfectos



Fuente: NVIDIA

Contacto: ia@frh.utn.edu.ar

Referencias

- *PyImageSearch: [Blog y Tutoriales](#)*
- *[The Serious Computer Vision Blog](#)*
- *[Introduction to Transfer Learning](#)*
- *Hugging Face: [Vision Tutorials](#)*
- *Kaggle: [Computer Vision Challenges](#)*