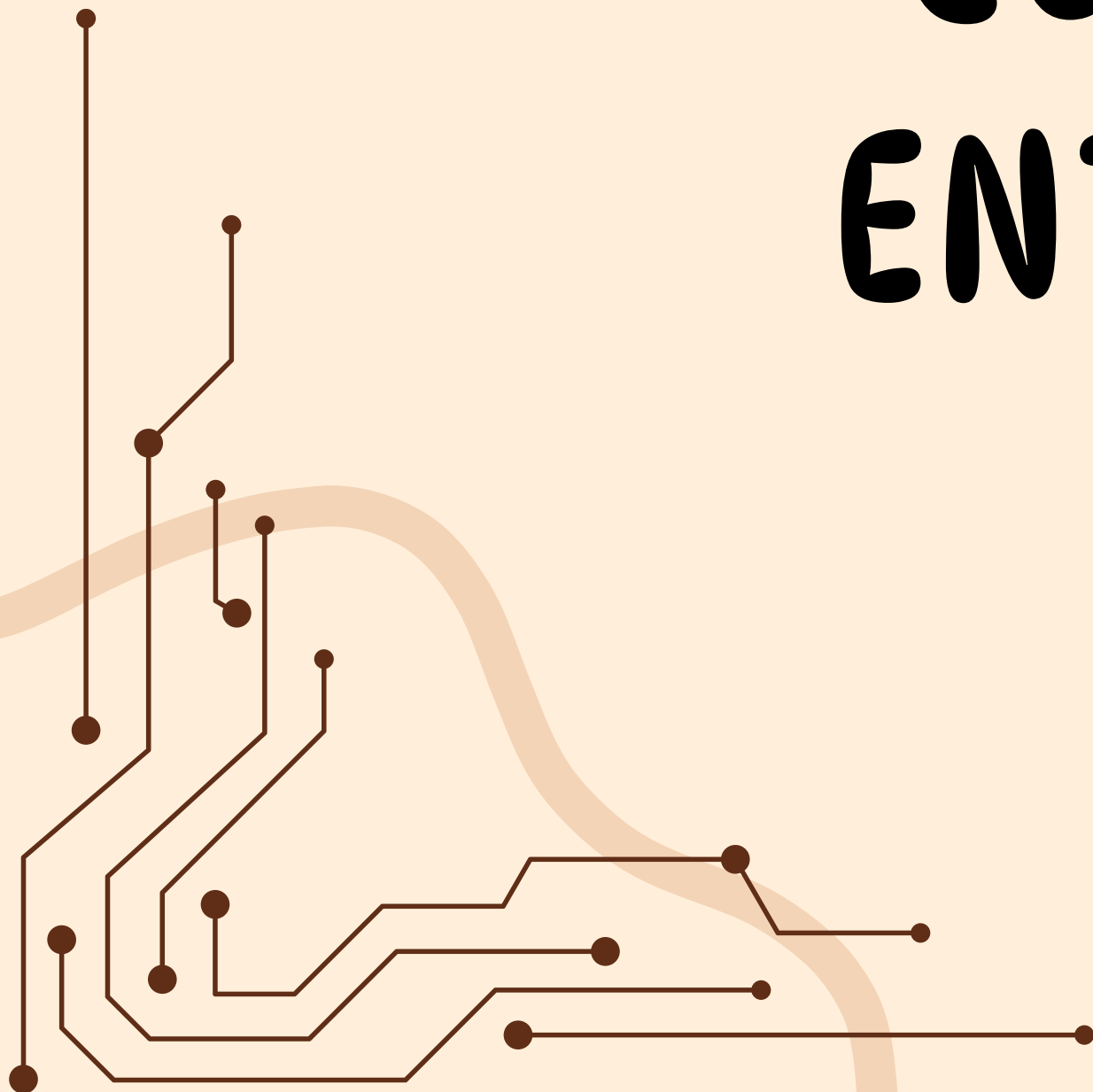
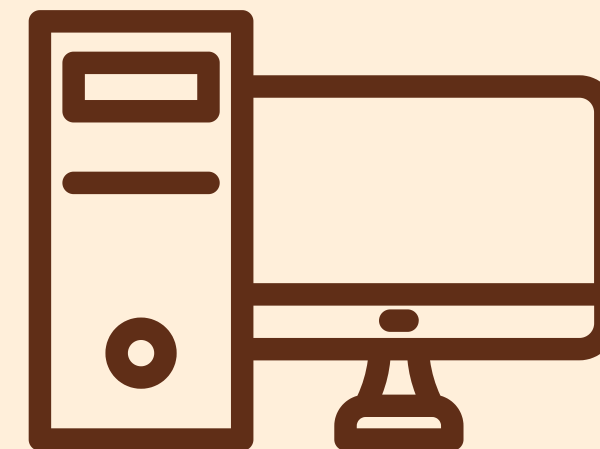
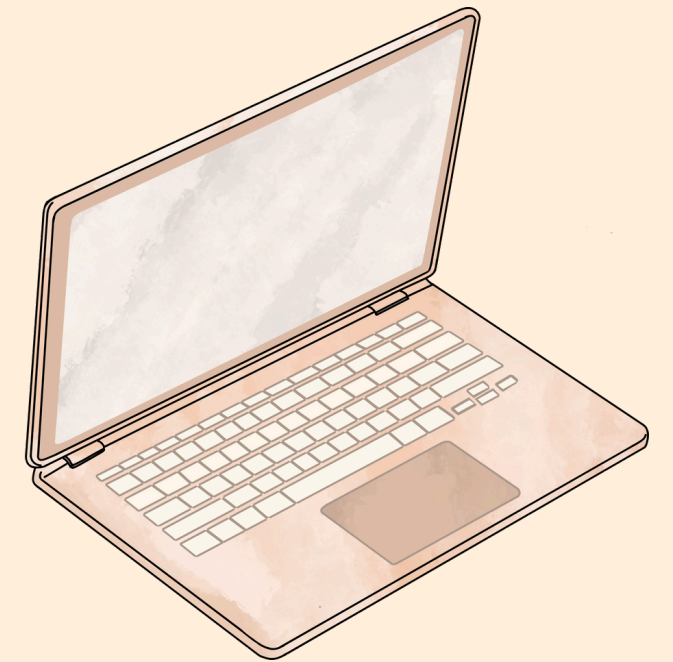
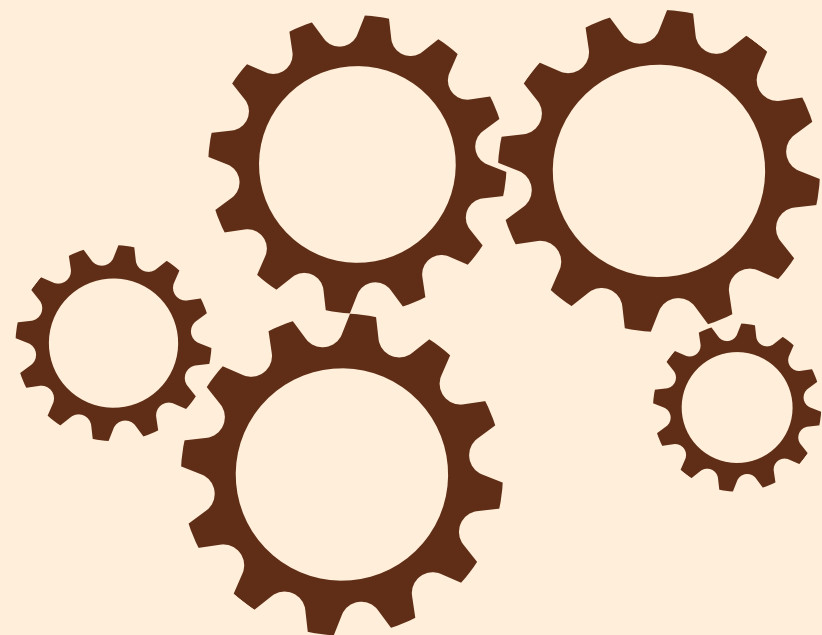


# COMPARACIÓN ENTRE GPU, CPU Y TPU



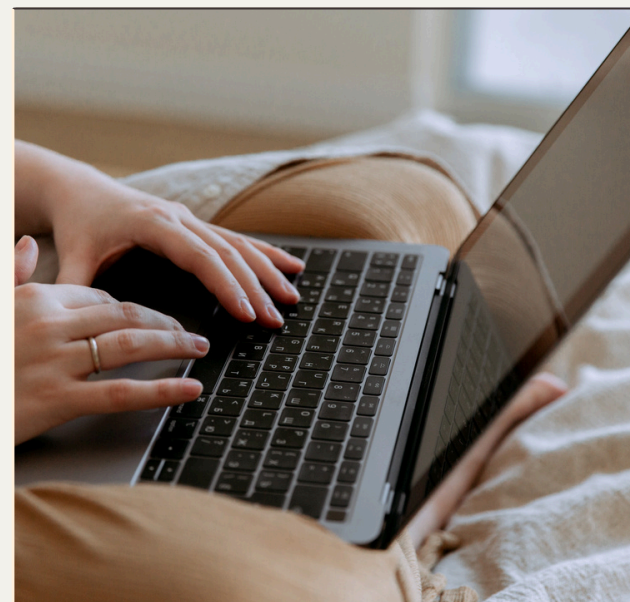


# MIEMBROS



**Julián Ramírez**

[julramirezdi@unal.edu.co](mailto:julramirezdi@unal.edu.co)



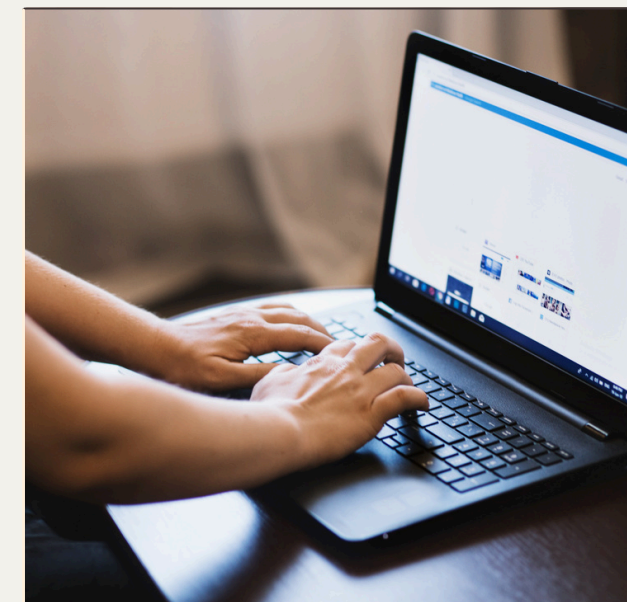
**María F. Cala**

[mcalar@unal.edu.co](mailto:mcalar@unal.edu.co)



**Xamir Rojas**

[xerojasga@unal.edu.co](mailto:xerojasga@unal.edu.co)



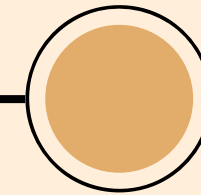
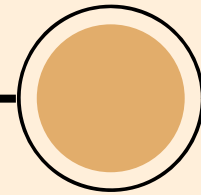
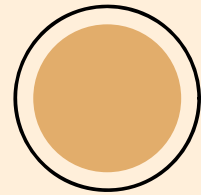
**Julián Rincón**

[jurinconor@unal.edu.co](mailto:jurinconor@unal.edu.co)

# ÍNDICE

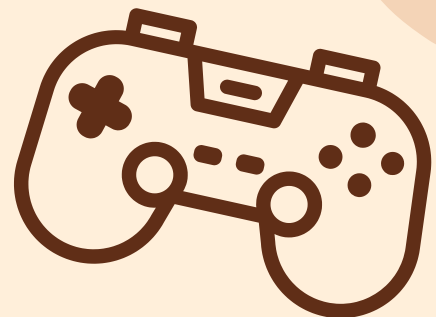
- 1. Introducción a unidades de procesamiento**
- 2. Arquitectura y propósito de cada unidad**
- 3. Medidas de rendimiento**
- 4. Tabla comparativa de modelos**
- 5. Aplicaciones específicas**
- 6. Cuándo usar cada unidad**
- 7. Conclusiones**

# 1. UNIDADES DE PROCESAMIENTO



## CPU

- Procesador general de la computadora
- Diseñada para ejecutar múltiples tareas secuenciales

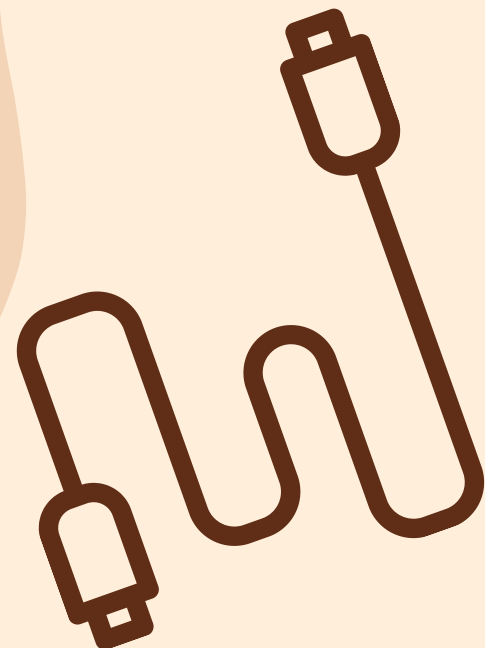


## GPU

- Especializada en procesamiento paralelo
- Optimizada para renderización de gráficos

## TPU

- Diseñada específicamente para inteligencia artificial
- Optimizada para operaciones matriciales de redes neuronales





## 2. ARQUITECTURA Y PROPÓSITO

### CPU

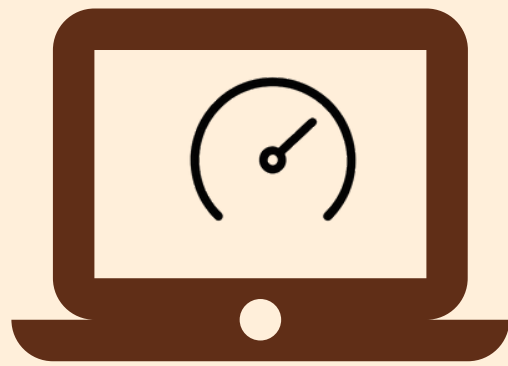
- **Arquitectura:** Pocos núcleos potentes (4-64)
- **Propósito:** Procesamiento secuencial versátil
- **Características:** Alta frecuencia, gran caché, ejecución fuera de orden
- **Diseño:** Optimizada para latencia y tareas generales

### GPU

- **Arquitectura:** Miles de núcleos simples (1000-5000+)
- **Propósito:** Procesamiento masivamente paralelo
- **Características:** Optimizada para cálculos vectoriales/matriciales
- **Diseño:** Maximiza el rendimiento en operaciones simultáneas idénticas

### TPU

- **Arquitectura:** Matriz sistólica de multiplicadores matriciales
- **Propósito:** Acelerar específicamente operaciones de IA/ML
- **Características:** Optimizada para redes neuronales
- **Diseño:** Circuitos específicos para operaciones tensores/matrices



### 3. MEDIDAS DE RENDIMIENTO

#### CPU

- **GHz (2-5-5.0 GHz)** : Frecuencia de reloj
- **IPC (Instrucciones por ciclo)**: Eficiencia de procesamiento
- **Número de núcleos**: Capacidad de multitarea

#### GPU

- **TFLOPS (5-60)** : Billones de operaciones de punto flotante por segundo
- **Memoria VRAM (8-48 GB)** : Capacidad para datos visuales
- **Ancho de banda memoria (500-1500 GB/s)**: Velocidad acceso a datos

#### TPU

- **TOPS (45-275)**: Tera operaciones por segundo
- **Precisión reducida (INT8/INT4)**: Optimizada para inferencia
- **Memoria HBM (8-32 GB)**: Alto ancho de banda

# 4. TABLA COMPARATIVA DE MODELOS

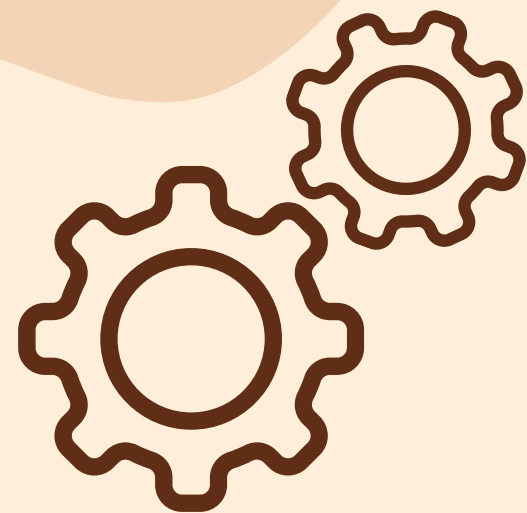
Modelo	Núcleos	Rendimiento	Memoria	Consumo energético
CPU Intel Core i9-12900K	16 (8P+8E)	5.2 GHz	30MB caché	125-241W
CPU AMD Ryzen 9 5950X	16	4.9 GHz	64MB caché	105W
GPU NVIDIA RTX 4090	16,384 CUDA	82.6 TFLOPS	24GB GDDR6X	450W
GPU AMD Radeon RX 7900 XTX	12,288	61.6 TFLOPS	24GB GDDR6	355W
TPU v4	-	275 TOPS	32GB HBM	175W
TPU v3	-	90 TOPS	16GB HBM	200W

# CPU

## 5. APLICACIONES ESPECÍFICAS

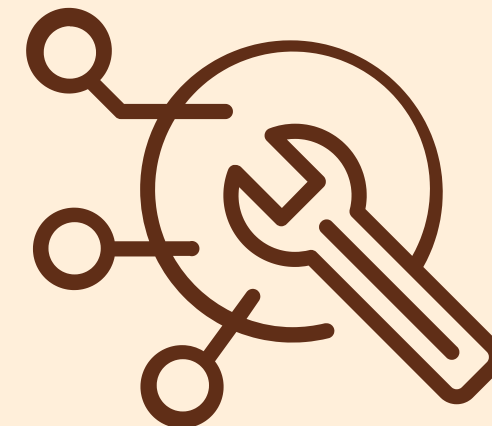
# TPU

- **Sistemas operativos y software general**
- **Tareas de un solo hilo (single-thread)**
- **Procesamiento secuencial**
- **Operaciones que requieren alta precisión**



# GPU

- **Renderizado 3D y videojuegos**
- **Procesamiento de imágenes y video**
- **Entrenamiento de modelos de IA**
- **Simulaciones visuales complejas**

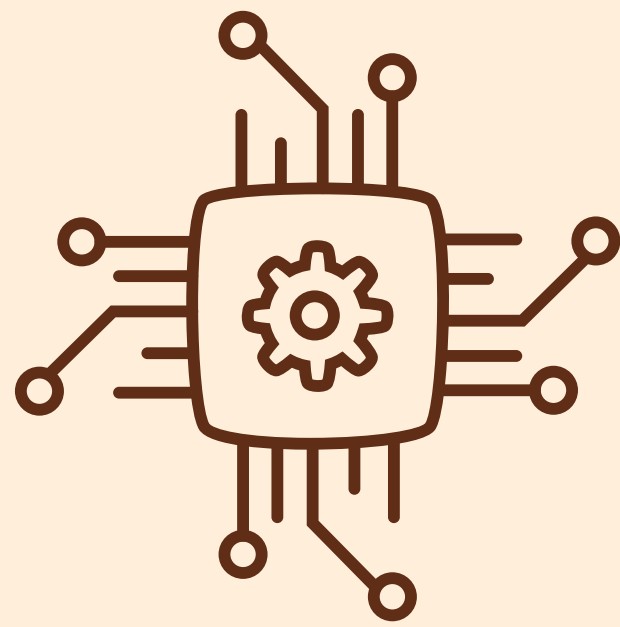


- **Inferencia de redes neuronales**
- **Procesamiento de lenguaje natural**
- **Visión por computadora**
- **Entrenamiento de modelos de aprendizaje profundo**



# 6. ¿CUÁNDO USAR CADA UNIDAD EN COMPUTACIÓN VISUAL?

tarea	Mejor opción	Razón
Edición básica de fotos	CPU	Suficiente para tareas simples
Renderizado 3D profesional	GPU	Cálculos paralelos masivos
Reconocimiento de objetos	TPU	Optimizado para redes neuronales convolucionales
Videojuegos	GPU	Renderizado en tiempo real
Edición de vídeo 4K	GPU	Procesamiento paralelo de frames
Procesamiento de imágenes médicas con IA	TPU	Eficiencia en inferencia de modelos complejos



## CPU

- **Sistemas operativos y software general**
- **Tareas de un solo hilo (single-thread)**
- **Procesamiento secuencial**
- **Operaciones que requieren alta precisión**

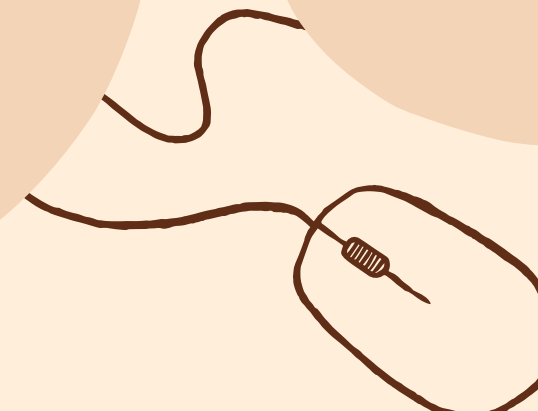
# CONCLUSIONES

## GPU

- **Renderizado 3D y videojuegos**
- **Procesamiento de imágenes y video**
- **Entrenamiento de modelos de IA**
- **Simulaciones visuales complejas**

## TPU

- **Inferencia de redes neuronales**
- **Procesamiento de lenguaje natural**
- **Visión por computadora**
- **Entrenamiento de modelos de aprendizaje profundo**



# GRACIAS

