

# Desarrollo de sistema generativo de imágenes de personajes con el estilo del videojuego Fortnite

Sergio Valdueza

Odreman Ferrer

Máster Deep Learning

8 Julio 2025

# Contexto



# FORTNITE

## Estilo visual único

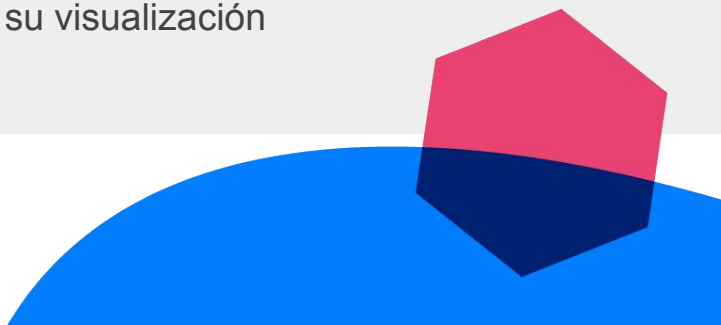
Parte del éxito del juego se basa en su estilo visual distintivo y la comercialización de skins de personajes.

## Fortniteando

Aplicación existente diseñada para la comunidad de Fortnite, donde se pueden visualizar los skins existentes en la Tienda Oficial.

# Objetivo

Crear un sistema de generación de imágenes de personajes con estilo visual de Fortnite usando modelos de difusión y lenguaje natural, e integrar su visualización en la app de Fortnitenado.



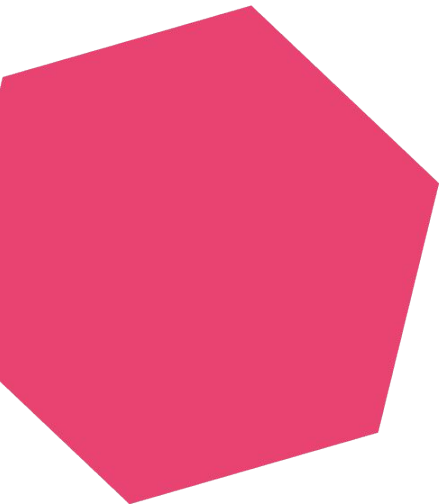
# Dataset



Se utilizó la API del juego para obtener los personajes e imágenes.

## Características:

- 1,525 personajes en resolución de 1024x1024 píxeles y 4 canales (RGBA)
- Diversidad extrema: características muy diferentes entre personajes
- Enfoque en estilo: Copiar el "estilo" visual, no personajes concretos





# Pre-procesamiento

El dataset original requirió un preprocesamiento intensivo para optimizar el entrenamiento del modelo.

## Selección de personajes

Eliminación de personajes que requerían una edición importante para usar.

1

## Edición y aislamiento

Separación del personaje de otros elementos y fondos

2

## Reducción de canales

Conversión de 4 canales (RGBA) a 3 canales (RGB)

3

## Data Augmentation x30

Multiplicación del dataset mediante transformaciones para aumentar variabilidad

4

## Etiquetado automático

Generación de descripciones textuales para entrenamiento texto-imagen.

5

# Fine-Tuning (Kohya's GUI)



Se realizó un fine-tuning intensivo para actualizar los pesos del modelo base y que aprendiera del nuevo dataset y sus etiquetas empleadas. El proceso demostró una convergencia rápida debido al tamaño relativamente pequeño del dataset para este tipo de modelos.

## Principales hallazgos:

- **Convergencia rápida:** Dataset de 16k imágenes permite convergencia acelerada del modelo
- **Learning rates variables:** LR bajo (15 epochs), medio (10-12 epochs), alto (5-7 epochs)
- **Relación personaje-etiqueta:** Aprendizaje efectivo entre nombres y características visuales
- **Captura de estilo:** Muy efectivo adoptando la apariencia y estilo visual del juego
- **Overfitting temprano:** Tendencia a memorizar ejemplos específicos en pocas épocas
- **Pérdida de diversidad:** Reducción gradual de variabilidad en generaciones tardías



# Adaptaciones con LoRAs (Kohya's GUI)

## Comida

**27 Imágenes** - 20 repeticiones  
4 imágenes/batch  
30 epochs - 4050 steps  
Learning rate =  $8e-5$   
network\_dim = 16 / network\_alpha = 8

## Animales

**42 Imágenes** - 10 repeticiones  
4 imágenes/batch  
40 epochs - 4000 steps  
Learning rate =  $8e-5$   
network\_dim = 16 / network\_alpha = 8

## Fuzzy Bear

**8 Imágenes** - 50 repeticiones  
2 imágenes/batch  
20 epochs - 4000 steps  
Learning rate =  $1e-4$   
network\_dim = 32 / network\_alpha = 16

## Star Wars

**19 Imágenes** - 40 repeticiones  
4 imágenes/batch  
20 epochs - 3800 steps  
Learning rate =  $1e-4$   
network\_dim = 16 / network\_alpha = 16

## Robots

**15 Imágenes** - 40 repeticiones  
2 imágenes/batch  
15 epochs - 4500 steps  
Learning rate =  $1e-4$   
network\_dim = 8 / network\_alpha = 4



# Recursos Computacionales (Vast.ai)



## FINE TUNING Modelo SDXL

**1x A100 SXM4**

500Gb Disco Duro

80Gb RAM

**Imagen Kohya GUI**

\$0,8 - \$1,2 por hora

8 - 12 horas por Fine Tune



## LoRAs Modelo Fine Tuned

**1x A100 SXM4**

300Gb Disco Duro

40Gb / 80Gb RAM

**Imagen Kohya GUI**

\$0,8 - \$1,2 por hora

1,5 - 2 horas por LoRA



## Inferencias

**1x RTX 4090**

200Gb Disco Duro

24Gb RAM

**Imagen ComfyUI**

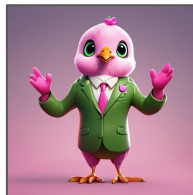
\$0,3 - \$0,6 por hora



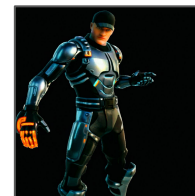
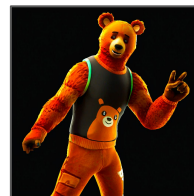
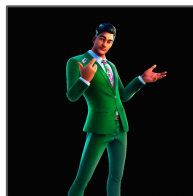
# Resultados (ComfyUI)



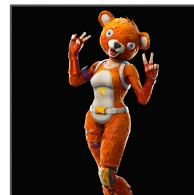
Modelo SD XL



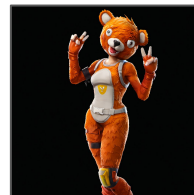
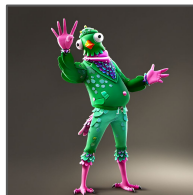
Fine Tuned



Con LoRAs



Ensamblado y  
Refinado



# Pipeline de Puesta en Producción

Servicio ComfyUI + API



- 1 Solicitud de Imagen en la Aplicación (Prompt)
- 2 Traducción al inglés del Prompt (LLM)
- 3 Clasificación del Prompt (LLM)
- 4 Llamada a ComfyUI para Generación Imagen
- 5 Refinamiento de la Imagen (LLM multimodal)
- 6 Visualización de Personaje en la Aplicación

# ¡Pruébame!

Se construyó una UI simple que consume la API desarrollada, permitiendo a los usuarios interactuar directamente con el modelo de generación.



Generador de Personajes Fortnite AI

Crea personajes únicos con inteligencia artificial

Descripción del personaje

Describe tu personaje ideal: jonesy con lentes oscuros y pantalón verde...

Calidad de imagen

Alta calidad (1024x1024)

Nivel de detalle

20

Precisión del estilo

11

Generar personaje

Error generando imagen



¿Preguntas?