

# Red Generativa Adversaria para la creación de rostros de anime

Bergna Aguilar, Diego  
Universidad Nacional de Ingeniería  
Lima, Perú  
diego.bergna.a@uni.pe

Mechan Osorio, Rodrigo  
Universidad Nacional de Ingeniería  
Lima, Perú  
rodrigo.mechan.o@uni.pe

Paima Mijahuanca, Kevin  
Universidad Nacional de Ingeniería  
Lima, Perú  
kpaimam@uni.pe

**Resumen**—Los animes han llegado a ser tantos que actualmente uno puede confundir el rostro de algún personaje de anime con alguno de otro anime y esto posiblemente debido a la falta de originalidad y creatividad o esfuerzo que le dedican los ilustradores. Todo esto por motivos como la tendencia de la forma de los rostros, o simplemente la gran cantidad de animes existentes, con lo que aumenta la probabilidad de que muchos de estos rostros se parezcan.

El presente trabajo implementa una red generativa adversaria (GAN - por sus siglas en inglés), la cual es capaz de generar nuevos rostros de personajes al implementar dos redes que compiten entre sí, la primera red es una generadora de rostros de anime mientras que la segunda es una red que discrimina la salida que da la red generadora, logrando así resolver el problema de no tener idea de qué rostro dar a un nuevo personaje para algún anime.

**Índice de Términos**— GAN, red generadora, red discriminadora, rostros de anime.

## I. INTRODUCCIÓN

El anime es una parte de la industria del entretenimiento japonesa el cual genera una gran cantidad de ingresos, logrando alcanzar una cantidad de 19,9 mil millones de dólares en el año 2019, sin embargo el ámbito comercial del anime se ha visto envuelta en problemas por derechos debido a la similitud que tienen los personajes de distintas empresas pudiendo generar un pérdidas en demandas entre las mismas.

Con ello como problemática, entre las posibles soluciones se encuentra la de idear un modelo de inteligencia artificial que pueda generar nuevos personajes de anime no existentes o en caso de que el personaje generado sea semejante, sea posible modificar parámetros que varíen las características del personaje generado.

Este proyecto se centra en la realización de un modelo de redes generativas adversarias con el fin de generar rostros de personajes de anime.

### I-A. Objetivos

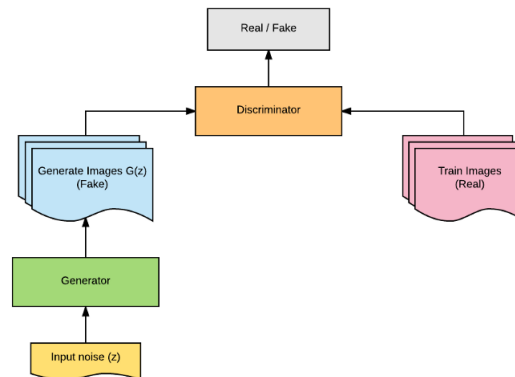
- Construir una red generativa adversaria para la generación de rostros

#### I-A1. Objetivos específicos:

- Implementación de la red generativa usando como referencia StyleGAN, StyleGAN2 y DCGAN.
- Implementación de la red discriminadora.

## II. GAN's

Las Redes Generativas Adversaria [4], mejor conocidas como GAN's por sus siglas en ingles, fueron introducidas para resolver la problemática de la complejidad de cálculo que tenían los modelos generativos profundos [2]. En el cual el modelo generativo(G) se enfrenta a un adversario, un modelo discriminatorio(D) el cual aprende a determinar si la muestra pertenece o no a los datos. Una analogía es que nuestro modelo generativo sea un falsificador de monedas mientras que nuestro modelo discriminatorio es la policía que intenta detectar esta moneda falsa.



Este proceso se repetirá hasta que nuestro modelo discriminatorio sea incapaz de detectar que la muestra creada por nuestro modelo generativo es falso.

## III. ESTADO DEL ARTE

Los GAN's están mejorando de tal manera que suplan ciertas limitaciones que se tenían en un principio. En esta sección veremos las distintas variantes que fueron creadas.

### III-A. Conditional Generative Adversarial Network (cGAN)

CGAN propuesto en el artículo del mismo nombre [8], en donde se presento un modelo en el cual los GAN's puede extenderse a un modelo en donde tanto el modelo generador como el modelo discriminatorio son condicionados por alguna información  $y$ . Pudiendo

realizar un condicionamiento introduciendo  $y$  a ambos modelos como capa de entrada adicional.

### III-B. Deep Convolutional Generative Adversarial Network (DCGAN)

DCGAN es un modelo que fue publicado en el artículo **Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks** [10], en donde se utilizan las redes neuronales convolucionales para el aprendizaje no supervisado, planteando ciertas restricciones de tal manera que se hagan estable al momento de ser entrenados para la mayoría de los escenarios.

Pasos para una DCGAN estable:

- Reemplazar las capas de agrupación por convoluciones estriadas(modelo discriminatorio) y convoluciones parcialmente estriadas(modelo generativo).
- Usar normalización batch [5] en el modelo generativo y discriminatorio.
- Eliminar completamente todas las capas convolucionales conectadas.
- Usar la función de activación ReLU [9] en el modelo generador para todas las capas excepto para la salida, el cual usa la función hiperbólica.
- Usar la función de activación LeakyReLU [11] en el modelo discriminatorio para todas las capas.

### III-C. Information Maximizing Generative Adversarial Network (InfoGAN)

El modelo se presenta en el artículo **InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets** [3]

### III-D. Wasserstein Generative Adversarial Network (WGAN)

El modelo [1]

### III-E. Style-Based Generative Adversarial Network (StyleGAN)

Modelo [6] en el cual

### III-F. StyleGAN2

[7]

## IV. METODOLOGÍA

Procederemos a detallar el planteamiento del desarrollo del proyecto.

### IV-A. Proceso

### IV-B. Dataset

Previo a la creación de un modelo, se debe analizar el conjunto de datos con los que se trabajará.

Para propósito de nuestro proyecto optamos por un dataset de rostros de anime. El dataset elegido es [este](#) el cual consta de un conjunto de 63500 archivos aproximadamente con rostros de anime genéricos y conocidos,

muchos de ellos cuentan con artículos en el rostro e incluso algunos con rostros de colores fuera de lo común. Un defecto que encontramos en el dataset es que en su mayoría los rostros pertenecen a personajes femeninos, pero para fines prácticos no le daremos importancia por ahora.

### IV-C. Implementación de la red generativa y la red discriminadora

#### IV-C1. Red generativa:

IV-C2. Red discriminadora: En esta sección se explicará los pasos que se seguirán en la experimentación.

## V. EXPERIMENTACIÓN Y RESULTADOS

En esta sección se mostrarán resultados obtenidos para experimentaciones realizadas.

### V-A. Limpieza de datos

Lo primero en realizarse fue un análisis del dataset, de las 63500 imágenes muchas de ellas tuvieron dimensiones diferentes, por lo que fueron redimensionadas a imágenes en formato PNG de 128px x 128px.

Posterior a ello se realizó una búsqueda de duplicados y se encontraron alrededor de 21000 imágenes repetidas las cuales fueron eliminadas dejando solo un ejemplar de las mismas.

## VI. DISCUSIÓN DE RESULTADOS

En esta sección se hablará de los resultados y presentando argumentos que respalden los mismos.

## VII. CONCLUSIONES

## REFERENCIAS

- [1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein gan, 2017.
- [2] Ludovic Arnold and Yann Ollivier. Layer-wise learning of deep generative models, 2013.
- [3] Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets, 2016.
- [4] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.
- [5] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift, 2015.
- [6] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks, 2019.
- [7] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan, 2020.
- [8] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets, 2014.
- [9] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Icml*, 2010.
- [10] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2016.
- [11] Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li. Empirical evaluation of rectified activations in convolutional network, 2015.