

# Herramienta inteligente capaz de mejorar la resolución de una imagen digital de formato espacial con pérdida de información

Trabajo Terminal No. \_\_\_\_ - \_\_\_\_

*Alumnos: \*Alcaraz Fraga Ricardo, Cruz Vera José Rodrigo, Benítez Miranda Samuel Eduardo*

*Directores: Cruz Meza María Elena*

*\*e-mail: jcruzv1402@alumno.ipn.mx*

**Resumen** - Las imágenes digitales con formato espacial presentan de forma inherente una pérdida de información debido a los métodos de compresión de imágenes existentes. Así pues, los usuarios de dichas imágenes perciben una vista pobre de las mismas[1]. Existen métodos para mejorar la resolución de imágenes en formato espacial, pero éstas tratan únicamente con los valores numéricos de los píxeles y no con el objeto contenido en la imagen. En este trabajo se presenta una propuesta que pretende diseñar un algoritmo inteligente capaz de reconocer el objeto contenido en la imagen para poder reconstruirlo y obtener un resultado con mejor vista para el usuario.

**Palabras clave** - Aprendizaje automático, imágenes digitales, formato espacial, redes neuronales

## 1. Introducción

Con la evolución acelerada de la tecnología a lo largo del siglo XX, el uso de la cámara analógica se vio desplazado por el de la cámara digital, lo cual a su vez implicó un cambio en la manera en la que visualizamos nuestras fotografías. Este cambio dio paso a una forma más eficiente de almacenar las fotografías tomadas por el usuario, desafortunadamente, con el aumento en la resolución de la vista presentada al usuario, el almacenamiento de estas imágenes digitales se volvió costoso en cuanto a espacio, por lo que, como se menciona en [2] y [3], se buscó una forma de poder almacenarlas sin tener que guardar todas las componentes de la misma, encontrándose con dos tipos de almacenamiento de información para los tipos de imágenes con formatos espaciales: con compresión o sin compresión, donde los primeros son los que presentan el problema, ejemplo de estos son los formatos TIFF, GIF, JPG[5], entre otros.

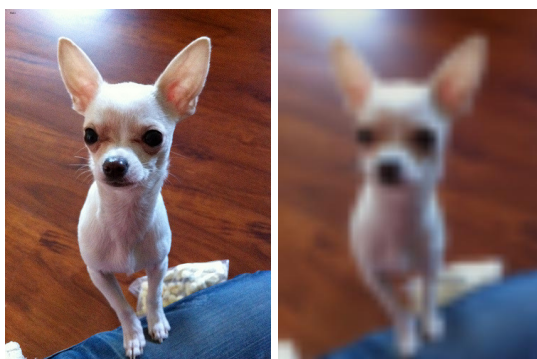


Figura 1. Imagen formato JPG con buena resolución (izquierda) e imagen con baja resolución (derecha)

Estos algoritmos, llamados de compresión, impactan en mayor o menor medida a la vista presentada al usuario, para corregir esta situación existen algoritmos para mejora de resolución de imágenes, estos tienen una gran desventaja, debido a que toman en cuenta únicamente los valores numéricos de los píxeles y no los objetos presentados en la escena, por lo que la mejora de resolución no alcanza todo el potencial posible, y considerando el avance actual de los sistemas de hardware con el uso de aprendizaje profundo [4], el cual se ha podido utilizar de manera práctica y eficiente, por lo que en este trabajo planteamos la posibilidad de utilizar una arquitectura de esta área de la inteligencia artificial para reconocer los objetos presentados en la escena y, de esta manera, poder mejorar la resolución de una manera mucho más consistente con el contenido presentado en las imágenes.

Algunos trabajos similares al planteado que se han desarrollado en esta área tanto comerciales, gratuitos y de investigación se muestran en la Tabla 1.

<b>SOFTWARE</b>	<b>CARACTERÍSTICAS</b>	<b>PRECIO EN EL MERCADO</b>
Topaz GigaPixel AI <a href="https://topazlabs.com/gigapixel-ai/">https://topazlabs.com/gigapixel-ai/</a>	Uso de Inteligencia Artificial para el mejoramiento de imágenes. Reescalamiento de hasta 16x	\$2,113.32 MXN
Let's Enhance <a href="https://letsenhance.io/">https://letsenhance.io/</a>	Reescalamiento de hasta 16x. Reconstrucción facial por medio de Inteligencia Artificial. Emplea redes neuronales convolucionales profundas.	\$0.00 MXN
Upscale Pics <a href="https://upscalepics.com/">https://upscalepics.com/</a>	Uso de Inteligencia Artificial para el mejoramiento de imágenes. Aplicación web.	\$0.00 MXN
Paper: HiFaceGan[6]	Uso de una arquitectura anidada para una renovación de cara multietapas. Uso de supresión colaborativa y reposición	No aplica

**Tabla 1.** Resumen de trabajos similares

## 2. Objetivo

Diseñar una herramienta inteligente capaz de mejorar la resolución de una imagen digital en formato espacial, con pérdida de información, mediante técnicas de análisis de imágenes e inteligencia artificial, que le permita una mejor vista al usuario.

### 2.1 Objetivos Específicos

- Comprender las redes neuronales generativas adversarias y el proceso que siguen para lograr generar salidas deseadas.
- Analizar y seleccionar el conjunto de datos y el modelo de machine learning a utilizar de acuerdo con las necesidades del problema planteado a resolver.
- Aplicar técnicas de preprocesamiento de imágenes al conjunto de datos seleccionado para simular las condiciones en las cuales los usuarios ingresarán las imágenes, que formarán el conjunto de testeo.
- Desarrollar el algoritmo inteligente que permita mejorar la calidad de una imagen de forma automática.
- Realizar el análisis de la funcionalidad del algoritmo diseñado mediante pruebas con el conjunto de testeo diseñado.
- Diseñar e implementar una interfaz de usuario que permita al usuario final visualizar e interactuar con la herramienta de forma sencilla.

### 3. Propuesta de Solución

Como se puede observar en [7], es posible diseñar modelos de aprendizaje profundo que, dada una entrada, generen una salida condicionada con esta entrada en forma de texto, imagen, señales, etc. Estos modelos se conocen como redes neuronales generativas adversarias. Estos modelos, como se explica en [4], son un trabajo en conjunto de dos modelos: un discriminador y un generador.

La solución que se propone, utilizará el esquema que la mayoría de los modelos existentes tienen, por lo que se requerirá de un generador que estará encargado de producir salidas condicionadas por un ruido inicial (no necesariamente aleatorio), mientras que el discriminador estará encargado de discernir entre imágenes reales (provenientes del conjunto de datos de entrenamiento) e imágenes generadas artificialmente (provenientes del modelo generador) (Fig. 2).

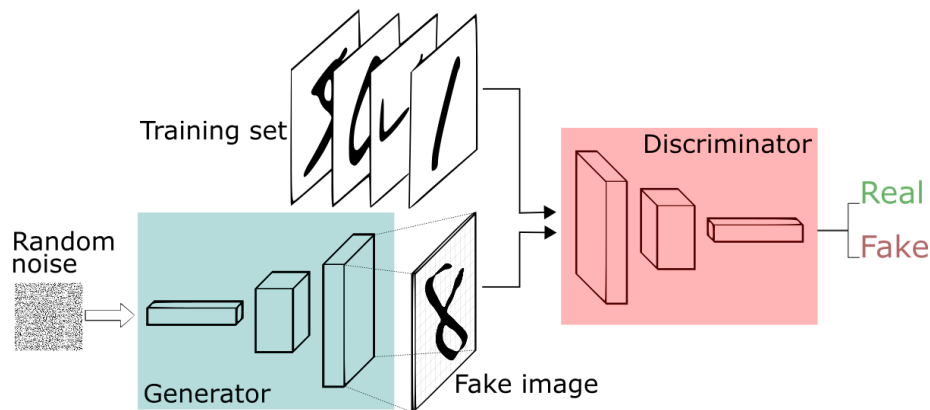


Figura 2. Funcionamiento de una GAN

Como también se menciona en [7], si el ruido inicial proviene de una capa convolucional alimentada con la imagen ingresada, la salida estará condicionada por esta imagen, por lo que, si como entrada al modelo generador pasamos la imagen con una compresión “muy agresiva” y a la salida indicamos que queremos la misma imagen pero con una mejor vista, el modelo generador aprenderá las relaciones necesarias para producir la salida que deseamos.

Con base en lo anterior, la propuesta consistirá en diseñar una red neuronal convolucional, previamente entrenada junto al modelo discriminador, que maximice la resolución de la imagen, aprovechando la estrategia de este tipo de algoritmos, ya que una convolución, según [4], “Las convoluciones consisten en tomar grupos de píxeles cercanos de la imagen de entrada e ir operando matemáticamente (producto escalar) contra una pequeña matriz que se llama kernel. Ese kernel, supongamos de tamaño  $3 \times 3$  pixels, recorre todas las neuronas de entrada (de izquierda-derecha, de arriba-abajo) y genera una nueva matriz de salida, que en definitiva será nuestra nueva capa de neuronas ocultas.”

Entendido esto, la propuesta de solución al problema planteado consiste en un sistema capaz de recibir una imagen con mala resolución que será procesada por el algoritmo inteligente, el cual hará uso del generador obtenido al término de la etapa de entrenamiento y que conectado a una interfaz gráfica permitirá recibir y desplegar dicha imagen con una mejor vista al usuario de una manera sencilla. La herramienta consistirá en tres módulos principales: la interfaz gráfica que permitirá recibir y desplegar las imágenes ingresadas a la herramienta, una API que se encargará de ser la conexión entre el modelo y la interfaz, y finalmente, el modelo de aprendizaje profundo que se encargará del procesamiento de las imágenes retornando así al usuario imágenes mucho más útiles con una mejor resolución y vista general. (figura 3)

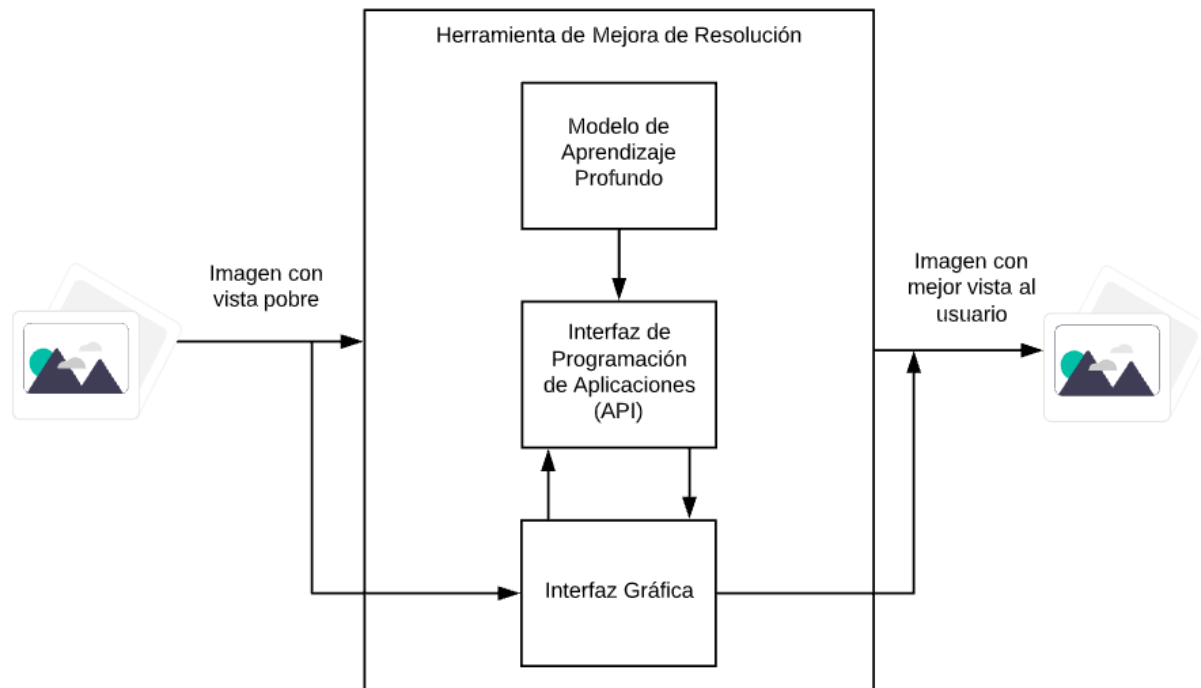


Figura 3. Propuesta de Arquitectura del Sistema

#### 4. Justificación

El área de mejora de resolución de imágenes con formato espacial es un área con mucho potencial, como se menciona en [2], con diversos problemas tales como presencia de ruido, desenfoque, entre otros; el de interés para esta propuesta es la mejora de la resolución de una imagen digital con el formato espacial, con pérdida de información. Este trabajo presenta una propuesta con otro enfoque, parecido al presentado en [7], de solución al problema de aumentar la resolución de las imágenes con formato espacial, con el cual se aproveche de mejor manera la información presentada por la imagen, debido a que, como se ha descrito en párrafos previos, las actuales soluciones toman en cuenta exclusivamente los valores numéricos de los píxeles y no los objetos presentados en la escena, por lo que la mejora de resolución no alcanza todo el potencial posible.

La propuesta de este trabajo terminal involucra comprender todas las fases del proceso de desarrollo de una herramienta de software, además de los conocimientos, habilidades, capacidades y destrezas adquiridos a lo largo de la carrera; en conjunto con nuevos tópicos que serán necesarios para la implementación de la propuesta, como lo son el aprendizaje de máquina, el cual es un tema de alta demanda e interés actual.

De igual forma, desarrollar la propuesta planteada en este documento enriquecerá nuestra formación integral, para proponer, analizar, diseñar, desarrollar, implementar, gestionar y administrar sistemas computacionales usando tecnologías de cómputo de vanguardia para mejorar la visualización de las imágenes digitales con formato espacial. Al concluir con este trabajo esperamos haber adquirido los conocimientos, habilidades, capacidades y destrezas necesarios para resolver problemáticas que la sociedad demande.

Por otra parte, creemos que esta herramienta podría ser utilizada de forma cotidiana en el embellecimiento de imágenes tomadas por usuarios comunes, como se puede observar en [8], a su vez que representar un ahorro en el almacenamiento (y por lo tanto en el costo) del almacenamiento de los conjuntos de datos utilizados para entrenar modelos de reconocimiento de imágenes, los cuales pueden ocupar mucho espacio, como el Images of LEGO Bricks[11] que utiliza un total de 1.02GB. Este ahorro se observaría al comprimir las imágenes para su almacenamiento.

## **5. Productos o Resultados Esperados**

Con el fin de cumplir y cubrir los requerimientos establecidos por la escuela, los productos que se entregarán al finalizar de cursar las unidades de aprendizaje Trabajo Terminal I y II, se espera obtener la siguiente lista de productos:

1. El código fuente de la herramienta (algoritmo propuesto y aplicación)
2. La documentación técnica de la herramienta (algoritmo propuesto y aplicación)
3. El manual de usuario

## **6. Metodología**

Según [10] el uso de una metodología adecuada ha probado ser un pilar para el desarrollo de un proyecto de construcción de Software, por lo que elegir el más adecuado es una decisión muy importante que debe tomarse con base en distintos criterios, como lo pueden ser el tamaño del proyecto, el tiempo con el que se dispone para realizarlo, así como la relación con el equipo de trabajo, etc.

De acuerdo a las características del proyecto y al tiempo con el que contamos para diseñarlo e implementarlo hemos decidido que la mejor metodología para trabajar es SCRUM, siendo esta una metodología de tipo ágil. Esta metodología sugiere la necesidad de diseñar y desarrollar sistemas informáticos en periodos cortos de tiempo.

Como bien se menciona en [10], las principales características que benefician al desarrollo del proyecto utilizando SCRUM son:

- El tiempo: Se permiten iteraciones de dos a cuatro semanas lo cual permitirá un avance constante de trabajo.
- El orden: Permite al programador decidir qué características cumplir durante cada iteración.
- La retroalimentación: Existe una buena retroalimentación al final de cada iteración para resolver incidencias y posibles problemas que se presenten.

## 7. Cronograma

CRONOGRAMA Nombre del alumno: Alcaraz Fraga Ricardo

TT No.: \_\_\_\_\_

Título de TT: Herramienta inteligente capaz de mejorar la resolución de una imagen digital de formato espacial con pérdida de información

[illegible]

CRONOGRAMA Nombre del alumno: Cruz Vera José Rodrigo TT No.:

Título de TT: Herramienta inteligente capaz de mejorar la resolución de una imagen digital de formato espacial con pérdida de información

[illegible]

CRONOGRAMA Nombre del alumno: Benitez Miranda Samuel Eduardo TT No.:

TT No.:

Título de TT: Herramienta inteligente capaz de mejorar la resolución de una imagen digital de formato espacial con pérdida de información

[illegible]



## 8. Referencias

- [1] S. Singh, Reduction of Blocking Artifacts in JPEG Compressed Image, Computer Science and Engineering Department, Guru Nanak Dev Engineering College, Ludhiana, Punjab, India, 2012.
- [2] R. González, R. Woods, Digital Image Processing, Second Edition, Prentice Hall, pg 409.
- [3] C. Parmar, K. Pancholi, A Review of Image Compression Techniques, Department of Electronics and Communication, L.J. Institute of Engineering and Technology, Gujarat, India, 2012.
- [4] S. Skansi, Introduction to Deep Learning From Logical Calculus to Artificial Intelligence, UTiCS, Springer, 2018.
- [5] About JPEG, Joint Photographics Experts Group, Disponible en: <https://jpeg.org/about.html>
- [6] L. Yang, C. Liu, P. Wang, S. Wang, P. Ren, S. Ma, W. Gao, HiFaceGan: Face Renovation via Collaborative Suppression and Replenishment, Peking, China, 2020.
- [7] P. Isola, J. Zhu, T. Zhou, A. Efros, Image-to-Image Translation with Conditional Adversarial Networks, Berkeley, California, 2018.
- [8] B. Hariharan, R. Girshick, Low-shot Visual Recognition by Shrinking and Hallucinating Features, Facebook AI Research (FAIR), 2017.
- [9] I. Sommerville, Ingeniería del Software, Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Alicante, 2005.
- [10] C. Rodríguez, R. Dorado, ¿Por qué implementar SCRUM?, Universidad Nacional de Colombia, Colombia, 2015.
- [11] J. Hazalzet. Images of LEGO Bricks, Disponible en: <https://www.kaggle.com/joosthazelzet/lego-brick-images>

## 9. Alumnos y Directores

*Alcaraz Fraga Ricardo.*- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta:2018630891, Tel. 5586867569, email [ralcarazf1700@alumno.ipn.mx](mailto:ralcarazf1700@alumno.ipn.mx)

Firma: \_\_\_\_\_

*Cruz Vera José Rodrigo.*- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta:2015030295, Tel. 5587065717, email [jcruzv1402@alumno.ipn.mx](mailto:jcruzv1402@alumno.ipn.mx)

Firma: \_\_\_\_\_

*Benitez Miranda Samuel Eduardo.*- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta:2018631141, Tel. 7773475949, email [sbenitezml701@alumno.ipn.mx](mailto:sbenitezml701@alumno.ipn.mx)

Firma: \_\_\_\_\_

*Cruz Meza María Elena.*- Cruz Meza María Elena.  
- M. en C. de la Computación, por el Centro de Investigación y Computación (CIC) del IPN en 2006. Ingeniero en Sistemas Computacionales por el Instituto Tecnológico de Tuxtla Gutiérrez, Chiapas en 1998. Áreas de interés: Inteligencia Artificial, Reconocimiento de Patrones, Aprendizaje de Máquina, Ciencia de Datos, Algoritmos Genéticos. Tel. 5521716717, email: [mcruzmm@ipn.mx](mailto:mcruzmm@ipn.mx)

Firma: \_\_\_\_\_

CARÁCTER: Confidencial  
FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública.  
PARTES CONFIDENCIALES: Número de boleta y teléfono.

## Informe de Protocolo de Trabajo Terminal

1



Maria Elena Cruz Meza

Lun 09/11/2020 01:48 PM

Para: Jose Rodrigo Cruz Vera



**Acuse de recibido.**

Enterada y de acuerdo.

**Ma. en C. María Elena Cruz Meza**

**Academia de ISC**

**Depto. de ISC.**

**Escom-IPN**

...

Gracias.

Gracias por confirmar.

Agradezco su respuesta.

¿Las sugerencias anteriores son útiles? ☐ Sí ☐ No

Responder

Reenviar



Ricardo Alcaraz Fraga

Dom 08/11/2020 05:59 PM

Para: Jose Rodrigo Cruz Vera



Estoy de acuerdo.

...



Jose Rodrigo Cruz Vera

Dom 08/11/2020 05:58 PM

Para: Jose Rodrigo Cruz Vera



Acuso de recibido

...