

# Instituto Politécnico Nacional. Escuela Superior de Computación Ingeniería en Inteligencia Artificial



# **Phyp**

Proyecto "Operaciones básicas con imágenes"

Líder de proyecto: Soto Álvarez Alejandro
Aragón González Francisco Javier
Castillo Montoya Jessica Alejandra
Torres López Marco Antonio

### RESUMEN

Centrarse en un programa para concatenar las imágenes con dos operaciones sencillas, tomando en cuenta su importancia de aplicación y los antecedentes que nos llevan a dichas aplicaciones. Aunado al alcance del código, obteniendo una investigación completa y un programa funcional.

# INTRODUCCIÓN.

El enorme desarrollo que está viviendo la tecnología asociada a la Inteligencia Artificial (IA) está dando lugar en los últimos tiempos a nuevas herramientas y aplicaciones espectaculares.

Una de las áreas donde los avances han sido más notables es en el reconocimiento de imágenes, en parte gracias al desarrollo de nuevas técnicas de Deep Learning o aprendizaje profundo.

En la actualidad tenemos al alcance de nuestras manos sistemas más precisos que los propios humanos, en las tareas de clasificación y detección en imágenes.

Según un estudio reciente de O'Really sobre el mercado de la IA, el reconocimiento de imágenes es una de las áreas donde más empresas están invirtiendo en EEUU dentro de la inversión en IA.

Los casos de uso son muchos y en diversas industrias y sectores, algunos ejemplos interesantes serían los siguientes:

- Etiquetado de imágenes: extraer tags o keywords asociados a imágenes, para poder clasificar o buscar posteriormente. Múltiples aplicaciones en sector turismo o retail.
- Verificación de usuarios basada en rostro: seguridad, autenticación, perfilado/segmentación de clientes, identificación en tiendas físicas.
- Análisis de opinión: detección del sentimiento o la experiencia de compra en tiendas físicas.
- Análisis de clientes: conocer mejor al usuario a través de la detección de logos o texto en los productos que consume.
- Diagnóstico de enfermedades: diagnóstico por imagen en base a comparación con diagnósticos previos. Retinopatías, diabetes, imagen médica.
- Realidad aumentada: gaming, catálogo virtual, interacción avanzada con el medio. Detección de matrículas: seguridad, segmentación, identificación, etcétera.

### **OBJETIVO DEL PROYECTO**

El propósito de este proyecto se enfoca en el diseño y desarrollo de un programa capaz de analizar imágenes por medio de matrices, para el análisis de imágenes en blanco y negro a través de la grayscale para obtener dichas matrices con valores.

### **OBJETIVOS PARTICULARES**

- Desarrollar un programa funcional
- Realizar operaciones básicas con matrices
- Almacenar las operaciones de matrices en nuevo archivo para visualizar los resultados

# **JUSTIFICACIÓN**

Es importante tener herramientas de este estilo que nos permitan entender más del lenguaje máquina al lenguaje humano, además de la compresibilidad de los valores que están detrás de una imagen a color y de una blanca y negro, así como sus respectivos tamaños.

Por otra parte, la creación y desarrollo de este tipo de programas son importantes para el avance en el análisis de imágenes y su manera de ser entendidas por el lenguaje máquina para una mejor optimización.

### **ALCANCE**

El alcance de este proyecto es el poder hacer suma, resta y producto cartesiano de matrices, dadas a través de imágenes y la grayscale, así como sus concatenaciones una a una, con una interfaz sencilla y entendible para el usuario.

### IMPLICACIONES DE LA INTELIGENCIA ARTIFICIAL

Las aplicaciones de visión artificial y el análisis de imágenes son capaces de resolver varios problemas del sector industrial, científico o de seguridad en la actualidad. La clasificación de imágenes es muy útil en la automatización de procesos en una empresa. Para realizar una tarea de clasificación de imágenes se requiere hacer la extracción de características que identifiquen a cada tipo de imagen como, por ejemplo: color, forma, textura. Esta metodología de construcción e implementación del clasificador se puede usar en varias aplicaciones según las clases de imágenes de objetos a analizar en casos similares y automatizar procesos de clasificación y reconocimiento de objetos.

#### ESTADO DEL ARTE

El análisis del estado del arte que aquí se realiza se agrupa por trabajos y proyectos recientes que se asemejan al nuestro junto con sus alcances.

# TRABAJOS DE INVESTIGACIÓN SOBRE OPERACIONES DE IMÁGENES

Dependiendo del lenguaje que se utilice es necesario utilizar diferente sintaxis, por ello en diferentes instituciones se explica cómo realiza las operaciones entre imágenes, mostrando los resultados de dicha interacción.

La BUAP nos explica cómo realizarlo en C y diversas páginas web nos proporcionan el software empleado en Python que nos da las bases para darle continuidad y ver su implementación en diversos campos como lo que se explica posteriormente.(PDIcursos)

# INTERPRETACIÓN

Los modelos de aprendizaje han demostrado ser una herramienta potente para solucionar problemas. Hoy en día esta herramienta se utiliza en una gran cantid ad de sectores, muchos de ellos de gran relevancia, desde la medicina hasta los coches autónomos.

Además, desde el punto de vista científico, hay una razón para utilizar las técnicas de interpretabilidad: la validación y depuración de modelos.(Sanchez, 2019)

### CASO: DETECCIÓN DE DAÑOS EN COCHES

Se realizaron pruebas sobre la detección de daños en coches a partir de imágenes. Los objetivos eran complejos: La gran variabilidad de modelos de coches y la mala calidad de imágenes al ser fotografiadas por los usuarios.

En modelo del proyecto para detectar si un coche estaba dañado o no, se usó LIME (Local Interpretable Model- Explanations) en la cual se basa en un aprendizaje automático. Esto permitió dar una explicación de cualquier predicción de un modelo en forma de modelo lineal. Para esto LIME crea superpixels de la imagen de entrada, va encendiendo y apagando los pixeles para comprobar en la predicción final.(A.Sánchez, 2019)

# RECONOCIMIENTO DE IMÁGENES E INTELIGENCIA EMOCIONAL

IBM anunció nuevas APIs cognitivas dirigidas a desarrolladores, las cuales mejoran sentidos visuales y emocionales de Watson, extendiendo aún más las capacidades del conjunto de herramientas y tecnologías cognitivas más grande y diverso de la industria.

Visual Recognition permite a los desarrolladores entrenar a Watson con clasificadores personalizados para las imágenes y construir apps que identifiquen visualmente conceptos e ideas únicas. Esto quiere decir que Visual Recognition ahora es personalizable con resultados hechos a la medida de las necesidades específicas de cada usuario.

Por ejemplo, una tienda puede crear una etiqueta específica para un estilo de pantalones de la nueva línea primaveral, de tal forma que pueda identificar cuando una imagen aparece en social media de alguien que use esos pantalones.(DiarioTI,2016)

# MEJORAR LOS FOTOGRAMAS DE UN VIDEO CON INTELIGENCIA ARTIFICIAL

Con el paso de los años se ha intentado mejorar los pixeles en las imágenes, para poder conseguir una mejor calidad de la imagen, mejores zoom, o simplemente para obtener un mejor resultado de la imagen final, todo esto utilizando una Inteligencia Artificial desde un ordenador, o si es el caso una I.A. en la cámara del celular.

Es decir, una inteligencia artificial que mezcla los diferentes pixeles obtenidos con la imagen real.

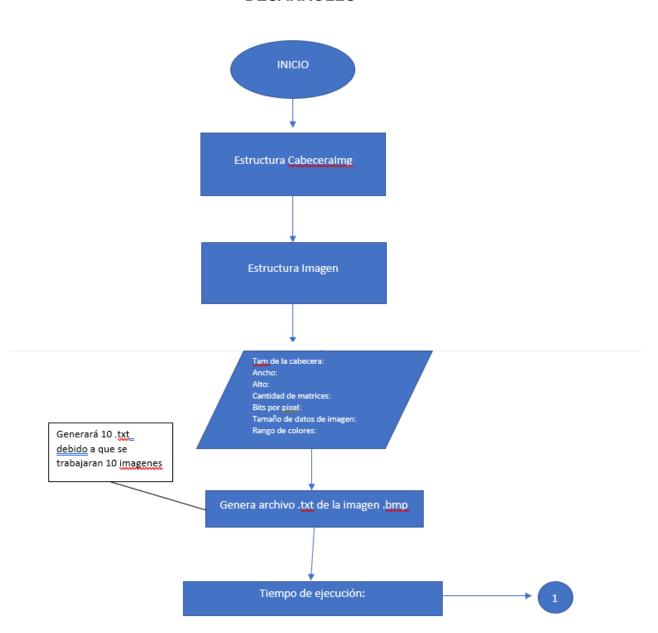
Por consecuencia, también se ha mejorado la forma de tomar videos o de mejorar los mismos videos existentes, utilizando la Inteligencia Artificial.

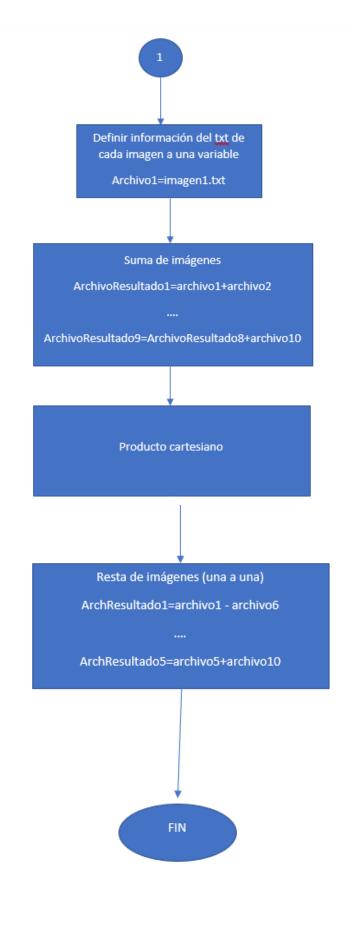
Mejorar los fotogramas individualmente de los videos, nos da como resultado una mejor fluidez en el video, y llegar incluso a los 60fps. El proceso para obtener este resultado, es trabajo de una Inteligencia Artificial, que se encargue de dividir el video en fotogramas, prediciendo en donde será el próximo fotograma y de esta manera lograr combinar los píxeles de la misma imagen, y al final combinar las imágenes para obtener el resultado de la fluidez en el video. (CSV, 2020)

Tabla comparativa de proyectos en el Mercado					
Trabajos del estado del arte	Agente inteligente clasificador de datos	CASO: DETECCIÓN DE DAÑOS EN COCHES	RECONOCIMIEN TO DE IMÁGENES E INTELIGENCIA EMOCIONAL	MEJORAR LOS FOTOGRAMAS DE UN VIDEO CON INTELIGENCIA ARTIFICIAL	
Autor	PHYP	Ting Neo	IBM	DAIN (Depth-Aware Video Frame Interpolation)	
¿Qué formato de imágenes puede procesar?	ВМР	JPG	JPG, JPEG, BMP, GIF, BPS, SVG, RAW	JPEG, BMP, PNG,GIF, BPS, SVG, RAW, MPEG-4, AVI, MOV, WMV, FLV, MKV	
¿El sistema tiene un clasificador interno?	No	Si	Si	Si	
¿Que sistema de Base de Datos utiliza?	Dinamica	Dinamica	Dinamica	Dinámica y Deductiva	
Es un sistema híbrido (necesita hardware y software específico)	No	Si	No	Si	
¿Cuantos desarrolladores participaron en el proyecto?	4	1	80 mil	6	
¿Requerimientos del sistema? (RAM,CPU, Etc.)	Intel I3 / 4 GB Ram / 256 GB de Disco duro	Intel® VT-x/48 TB de almacenamiento/768 GB de DRAM/GPU Tesla M60 de NVIDIA	Intel Dual core 2.6 Hrz/ Memoria RAM de 4 GB, HDD de 256 GB, Interl HD Family 4400	Memoria RAM de 16 GB, SSD de 256 GB, Nvidia GeForce GTX 1050, GCC PyTorch y Ubuntu	
¿Cuántas imágenes puede procesar de forma concurrente	10	Aproximadamente 300	De acuerdo a las necesidades del usuario	Por ser un software que mejora vídeos, las imágenes que procesa, dependerá de la duración del mismo, debido a la interpolacion de imagenes.	
Operaciones básicas que aplica a las imágenes (suma, resta, etc.)	Suma, resta, producto cartesiano	Sumas	Operaciones Básicas	Suma y Producto	
¿Máxima capacidad de procesamiento?	10	3,000,000 de imagenes	No tiene limite	No tiene límite,la única desventaja es el tiempo que tardará en mejorar el video.	
Tiempo de procesamiento de sus imágenes de mayor tamaño	512x512	50 segundos	Dependerá del tipo y cantidad de imagenes que use el usuario	Dependera del equipo, su memoria RAM y procesamiento	

Lenguaje de programación utilizado	С	Python	APIS	Python
Utiliza alguna red neuronal?¿De que tipo?	No	No	No	No

# **DESARROLLO**





#### RESULTADOS

[@][@]-77 [@][1]-89 [@][2]-113 [@][3]-97 [@][4]-121 [@][5]-150 [@][6]-178 [@][7]-178 [@][9]-174 [@][9]-170 [@][10]-154 [@][11]-169 [@][12]-170 [@][13]-138 [@][13]-147 [@][23]-121 [@][23]-139 [@][23]-139 [@][23]-147 [@][23]-139 [@][23]-149 [@][23]

[@[@]-150 [@][1]-150 [@][2]-153 [@][3]-150 [@][4]-150 [@][5]-150 [@][6]-152 [@][7]-143 [@][8]-121 [@][9]-99 [@][10]-83 [@][11]-82 [@][12]-85 [@][13]-77 [8]-150 [@][8

### 훩 Python 3.8.3 Shell

```
File Edit Shell Debug Options Window Help
     [1][0]=120 [1][1]=120 [1][2]=100 [1][3]=78 ...
                [2][1]=118 [2][2]=102 [2][3]=82 ...
     [2][0]=118
     [3][0]=105 [3][1]=105 [3][2]=89 [3][3]=77 ...
     [4][0]=79 [4][1]=79 [4][2]=68 [4][3]=61 [4...
459
     [459][0]=41
                  [459][1]=41
                               [459][2]=44
     [460][0]=43
                  [460][1]=43
                               [460][2]=48
461
     [461][0]=49
                  [461][1]=49
                               [461][2]=53
                                            [461][3...
462
     [462][0]=53
                  [462][1]=53
                               [462][2]=53
                                            [4621[3...
    [463][0]=52 [463][1]=52 [463][2]=53 [463][3...
[464 rows x 1 columns]
                 [0][1]=120 [0][2]=100 [0][3]=78 ...
     [01 [01 = 120 ]
     [1][0]=118
                 [1][1]=118 [1][2]=102 [1][3]=82 ...
     [2][0]=105 [2][1]=105 [2][2]=89 [2][3]=77 ...
     [3][0]=79 [3][1]=79 [3][2]=68 [3][3]=61 [3...
[4][0]=55 [4][1]=55 [4][2]=51 [4][3]=48 [4...
                  [458][1]=41 [458][2]=44 [458][3...
458
     [458][0]=41
     [459][0]=43
                  [459][1]=43
                               [459][2]=48
459
                                            [459][3...
     [460][0]=49
                  [460][1]=49 [460][2]=53
                                            [460][3...
461
     [461][0]=53
                  [461][1]=53
                               [461][2]=53
                                            [461][3...
                  [462][1]=52 [462][2]=53
462
    [462][0]=52
                                            [462][3...
[463 rows x 1 columns]
     [0][0]=118 [0][1]=118 [0][2]=102 [0][3]=82 ...
     [1][0]=105 [1][1]=105 [1][2]=89 [1][3]=77 ...
     [2][0]=79 [2][1]=79 [2][2]=68 [2][3]=61 [2...
     [3][0]=55
               [3][1]=55 [3][2]=51 [3][3]=48 [3...
     [4][0]=48 [4][1]=48 [4][2]=48 [4][3]=48 [4...
                  [457][1]=41 [457][2]=44
                                            [457][3...
     [457][0]=41
458
     [458][0]=43
                  [458][1]=43
                               [458][2]=48
                                            [458][3...
     [459][0]=49
                  [459][1]=49
                               [459][2]=53
459
                                            [459][3...
460
     [460][0]=53
                  [460][1]=53 [460][2]=53
                                            [460][3...
     [461][0]=52 [461][1]=52 [461][2]=53 [461][3...
[462 rows x 1 columns]
```

```
from PIL import Image
import numpy as np
import pandas as pd
archivo1_0 = pd.read_csv('Imagen1.txt', header = None)
print (archivol 0)
archivo2_0 = pd.read_csv('Imagen2.txt', header = None)
print (archivo2 0)
archivo3_0 = pd.read_csv('Imagen3.txt', header = None)
print (archivo3_0)
archivo4_0 = pd.read_csv('Imagen4.txt', header = None)
print (archivo4_0)
archivo5_0 = pd.read_csv('Imagen5.txt', header = None)
print (archivo5_0)
archivo6_0 = pd.read_csv('Imagen6.txt', header = None)
print (archivo6 0)
archivo7_0 = pd.read_csv('Imagen7.txt', header = None)
print (archivo7 0)
archivo8_0 = pd.read_csv('Imagen8.txt', header = None)
print (archivo8 0)
archivo9_0 = pd.read_csv('Imagen9.txt', header = None)
print (archivo9 0)
archivo10_0 = pd.read_csv('Imagen10.txt', header = None)
print (archivo10_0)
#Primer Apartado
archivo1=archivo1_0
archivo2=archivo2 0
archivo3=archivo3 C
```

archivor3-eachivor3\_0
archivor3-eachivor3\_0
[9][9]-143 [9][2]-123 [9][3]-99 [9][4]-86 [9][5]-88 [9][6]-97 [9][7]-99 [9][8]-91 [9][9]-74 [9][10]-66 [9][1]-71 [9][12]-88 [9][3]-79 [9][3]-7

### CONCLUSIONES

# **Aragon Gonzalez Francisco Javier**

Los fundamentos de programación son importantes para nuestra carrera, por que básicamente son las bases para elaborar programas o software más complejo. Así mismo nos sirve para poder elaborar una I.A. bastante elaborada, es decir que contenga redes neuronales, machine learning, deep learning, entre otras, para poder así, satisfacer ciertas necesidades de las personas e incluso de ciertas empresas. Este proyecto que hemos realizado bien pueden ser las bases para una inteligencia artificial, que pueda procesar los pixeles de una imagen, o de algún video, para combinar esos pixeles y llegar a algún resultado esperado.

### Castillo Montoya Jessica Alejandra

A lo largo de la carrera es importante entender los fundamentos de la programación para poder elevar los niveles de uso e implementación de la misma, una base bien definida termina por dar frutos en cuanto a la comunicación con tecnologías de IA simples, complejas o novedosas. A su vez proyectos como este representan un desafío para cosas importantes como apuntadores o manejo de archivos, que al no ser bien implementados, los programas serían muy complejos, casi imposibles de resolver.

## Soto Alvarez Alejandro

La programación es importante ya que nos enfocamos en toda la carrera a un entorno lenguaje máquina y es esencial para poder dar instrucciones a la máquina del trabajo a realizar para desarrollar software para diferentes instancias, este proyecto es muy importante para aprender y poder hacer comparaciones de archivos a través de su formato y valor numérico interpretado por el lenguaje máquina

### Torres López Marco Antonio

La programación en inteligencia artificial es muy importante ya que esta nos dará las bases para inicializar en nuestra carrera, además con la IA se pueden lograr grandes cosas como resolver problemas a gran escala que una persona podría tardar años en poder resolverlo, y este tipo de proyectos son esenciales para poder adentrarte al mundo de los archivos ya sea imagenes, videos, gifs y entre otros.

# REFERENCIAS BIBLIOGRÁFICAS

- -DiarioTI (26/02/2016)." IBM presenta APIs avanzadas de Watson en la nube"<a href="https://diarioti.com/ibm-presenta-apis-avanzadas-de-watson-en-la-nube/92931">https://diarioti.com/ibm-presenta-apis-avanzadas-de-watson-en-la-nube/92931</a>.
- Arcos Sánchez Alberto. (05/04/2019). El peligro de no interpretar las predicciones de tus modelos.BBVA, "https://www.bbva.com/es/el-peligro-de-no-interpretar-las-predicciones-de-tus-modelos
- Arcos Sánchez Alberto. (05/04/2019). El peligro de no interpretar las predicciones de tus modelos.BBVA, "https://www.bbva.com/es/el-peligro-de-no-interpretar-las-predicciones-de-tus-modelos/".}
- PDI CURSOS visto en: <u>https://www.cs.buap.mx/~iolmos/pdi/Sesion12\_OpEntrelmagenes.pdf</u>
- CSV, D. (2020, 04 09). Dot CSV. Retrieved from https://www.youtube.com/watch?v=0dzzHxzYh-s