|  |  |  |
| --- | --- | --- |
| Resultado de imagen para ipn logo png | 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**

Este proyecto se llevará acabo para poder satisfacer la necesidad de cierta empresa para trabajar en evaluar las operaciones básicas de un agente inteligente clasificador de datos.

**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. Hoy en día tenemos ya al alcance de nuestras manos sistemas más precisos que los propios humanos, en las tareas de clasificación y detección en imágenes.

Según un reciente estudio 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.

**OBJETIVOS**

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 color y en blanco y negro a través de la grayscale para obtener dichas matrices con valores.

**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.

**JUSTIFICACIÓN**

Es importante tener herramientas de este estilo que nos permitan entender más del lenguaje máquina al lenguaje humano, además de la comprensibilidad 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.

**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.

Explicación a consultar:

<https://www.cs.buap.mx/~iolmos/pdi/Sesion12_OpEntreImagenes.pdf>

**INTERPRETACIÓN**

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/>”.

Los modelos de aprendizaje han demostrado ser una herramienta potente para solucionar problemas. Hoy en día esta herramienta se utiliza en una gran cantidad 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.

**CASO: DETECCIÓN DE DAÑOS EN COCHES**

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/>”.

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 superpixeles de la imagen de entrada, va encendiendo y apagando los pixeles para comprobar en la predicción final.

**RECONOCIMIENTO DE IMÁGENES E INTELIGENCIA EMOCIONAL**

# DiarioTI (26/02/2016).” IBM presenta APIs avanzadas de Watson en la nube”

<https://diarioti.com/ibm-presenta-apis-avanzadas-de-watson-en-la-nube/92931>.

BM 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.