

Toolkit pipelines de transformación de imagen

El software “Toolkit pipelines de transformación de imagen”, es una aplicación web producto de investigación, el cual realiza la aplicación de múltiples técnicas de procesamiento de imágenes y genera todas las posibles combinaciones de esta técnicas para la transformación y optimización de la imagen fuente para etapas superiores de procesamiento, para este caso para el reconocimiento de texto.

Descripción de funcionalidades

- **Cargue o selección de imagen fuente:** Funcionalidad que permite el ingreso al sistema de la imagen original, la cual será procesada digitalmente.

Pipeline de optimización de imagen

Cargar o seleccionar imagen

No file chosen

Imagenes totales: 1



- **Aplicación individual de 28 técnicas de procesamiento digital:** Posterior al cargue de la imagen fuente, el software aplicará a la imagen original, 28 técnicas de procesamiento digital, luego se observa el resultado del procesamiento junto al nombre de la técnica aplicada.

Pipeline de optimización de imagen

Lista de transformaciones

Original



Análisis OCR

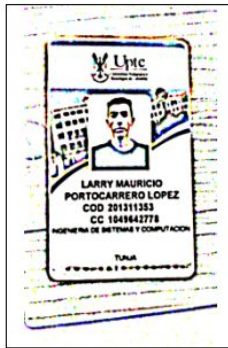
LARRY MAURICIO PORTOCARRERO LOPEZ

Analizar

Transformaciones totales: 28



☐ Binarization



☐ Clean imagemagic



☐ Contrast adjust



☐ Crop morphology

- **Composición de técnicas para generar pipelines de procesamiento digital:**
Cada técnica individual puede ser seleccionada para la generación de un pipeline, denominados pipelines a las posibles combinaciones de un conjunto de técnicas.



☒ Remove noise



☒ Solarize



☒ Standardize



☒ Text binarization

Transformaciones seleccionadas

- text_binarization
- standardize
- solarize
- remove_noise

Procesar

Luego se observa el resultado de cada pipeline.

Pipeline de optimización de imagen

Lista de pipelines

Original



Análisis OCR

LARRY MAURICIO PORTOCARRERO LOPEZ

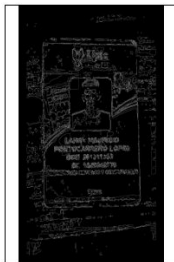
Analizar

Pipelines totales: 34



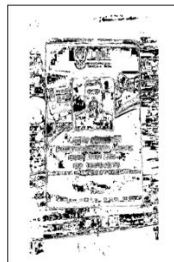
1

[Ver pasos](#)



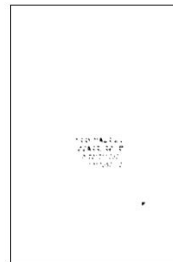
2

[Ver pasos](#)



3

[Ver pasos](#)



4

[Ver pasos](#)

- **Detalle de pasos aplicados sobre la imagen original de cada pipeline:** Cada resultado de pipeline podrá detallar la composición de técnicas usadas para generar este pipeline.

Pipeline de optimización de imagen

Lista de pasos

Original



Análisis OCR

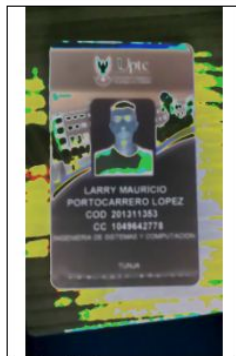
LARRY MAURICIO PORTOCARRERO LOPEZ

Analizar

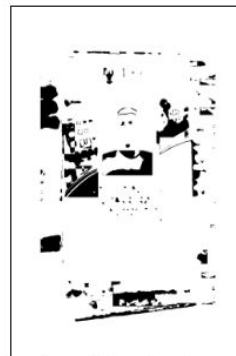
Pasos totales: 3



1)remove noise



2)solarize



3)text binarization

- **Análisis de OCR (Reconocimiento de texto):** Como se observa en las anteriores funcionalidades está presente la función de **Análisis OCR**, la cual realiza la comparación de texto original contra el texto detectado en cada transformación. Se obtiene el porcentaje de similitud útil para la comparación de resultados y eficiencia de la generación, aplicación y combinación de técnicas.

Composición del software - Módulos

Iteración

Se hace uso de la librería, incluida en python, *itertools*, la cual provee funciones de generación de combinaciones y permutaciones útiles para la obtención de todas las posibilidades de la lista de transformaciones seleccionadas previamente. Además hace una gestión adecuada de la memoria, usando solo espacio de memoria para el objeto en proceso.

Gestión de memoria

Módulo que surge con la necesidad de limitar y controlar el uso de memoria temporal. Al tener $3.0488834461171384e+29$ permutaciones y 268435456 elementos del conjunto potencia para el máximo generado de pipelines de las 28 técnicas se hace necesario limitar el uso del recurso de memoria basado en los recursos del equipo donde se ejecuta la generación automática de pipelines.

Gestión de archivos de imagen

Módulo de lectura, transformación binaria y almacenamiento de las imágenes recibidas, procesadas y resultantes a través del procesamiento de imágenes digitales. Para este módulo se hace uso de la librería numpy y PIL que permiten la obtención de datos binarios a partir de una imagen, utilizados como entrada de cada una de las implementaciones de técnicas de transformación.

Transformaciones

Módulo de recopilación de implementaciones de código abierto, cada una de las técnicas está descrita por una función parametrizada. Este módulo permite la adición de nuevas técnicas manteniendo el formato de entrada y salida (image

np.ndarray) .

OCR

Implementación de las librerías Python-tesseract y Fuzzywuzzy. Python-tesseract es una herramienta de reconocimiento óptico de caracteres (OCR) para python, contenedor para el motor Tesseract-OCR de Google. Fuzzy Wuzzy, librería de coincidencia difusa de strings, basada en la teoría de Levenshtein Distance para calcular las diferencias entre secuencias. Este módulo se parametriza con la ruta de la imagen y el texto original, retorna el texto detectado junto con el porcentaje de similitud de los textos.

Servidor Web

Desarrollado con el framework Flask, provee interfaz para uso de la herramienta desde un navegador. Incluye las funciones de selección o subida de una imagen para su transformación, selección de técnicas para la generación automática de pipelines, visualización de resultados de cada una de las pipelines, para cada pipeline es posible ver la secuencia de pasos de la transformación y finalmente la comparativa del porcentaje de similitud de cada una de las técnicas e imagen original con el proceso de reconocimiento de texto.