

Facultad de Ciencias Exactas, Ingeniería y Agrimensura

Tecnicatura Universitaria en Inteligencia Artificial



PROCESAMIENTO DE IMÁGENES I

IA 4.4

Año: 2025

Integrantes:

- Tapia, Fabrizio Joaquín T-3095/3
- Mezzano, Florencia M-7463/2

Fecha de entrega: 22/10/2025

Docentes:

- Gonzalo Sad
- Juan Manuel Calle
- Joaquín Allione

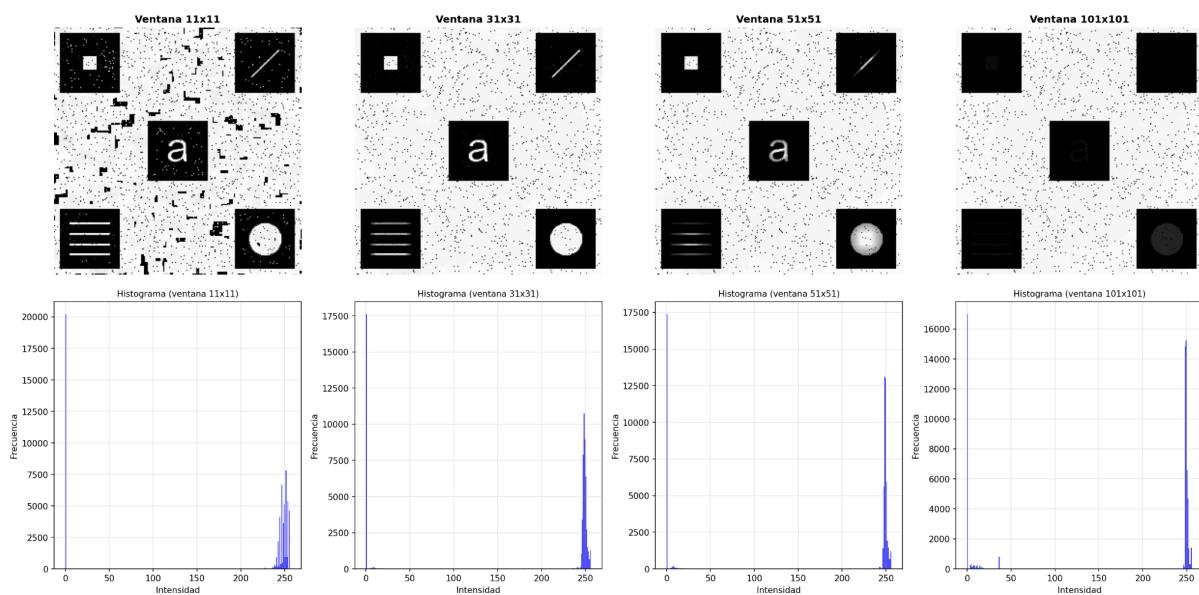
Informe Técnico

Ejercicio 1: Ecualización Local de Histograma

A partir de una imagen con bajo contraste global, con detalles ocultos en zonas muy oscuras o muy claras, el objetivo fue implementar una técnica que mejorara el contraste dinámicamente en pequeñas áreas para revelar esos detalles escondidos ya que la ecualización global no era suficiente porque promedia el contraste de toda la imagen. Por lo que se implementó el algoritmo de ecualización local, utilizando una ventana deslizante que recorre la imagen pixel a pixel aplicando la función de ecualización de histograma de OpenCv (cv2.equalizeHist) sobre la sub-región definida por la ventana en cada paso.

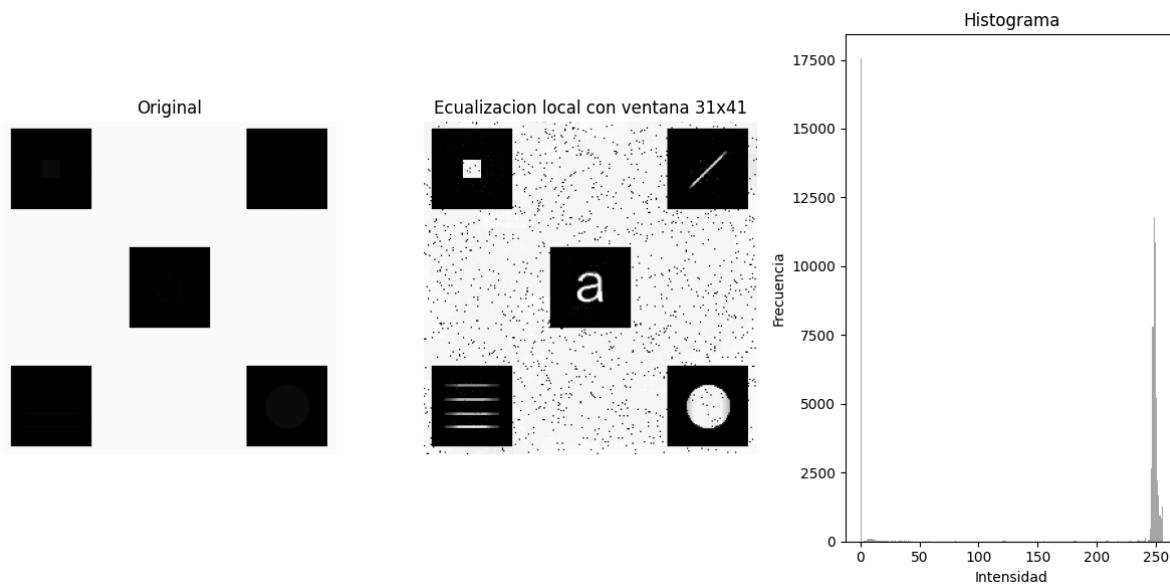
El criterio clave para el rendimiento fue la elección del tamaño de la ventana, ya que influye directamente en la definición del "vecindario local", por lo que se aseguró que las ventanas fueran impares para mantener un píxel central definido. Para el manejo de los bordes de la imagen, donde el píxel carece de un vecindario completo, se aplicó un relleno o padding con replicación de bordes (cv2.BORDER_REPLICATE).

La siguiente figura muestra una comparación de algunos de los resultados obtenidos utilizando diferentes tamaños de ventana cuadrada.



Tamaño de Ventana	Observación / Comentario en el Informe	Explicación Técnica (Criterio)
11x11	La imagen se ve muy ruidosa y artefactada. Se exponen pequeños detalles, pero a costa de amplificar significativamente el ruido de fondo.	Al ser la ventana muy pequeña, la distribución de intensidad del ruido impulsivo (píxeles muy claros o muy oscuros) dentro de la ventana se interpreta como información local, y la ecualización estira el contraste de esas pequeñas variaciones, amplificando el ruido y generando un aspecto granulado (ruido) en el fondo.
31x31 y 51x51	Se alcanza el mejor equilibrio entre realce de detalle y supresión de ruido. Figuras como la 'a' y las barras horizontales muestran bordes claros y transiciones suaves en las zonas de contraste moderado (círculo).	El tamaño mediano de la ventana es lo suficientemente grande como para que la ecualización local promedie la influencia del ruido circundante (actuando como un filtro pasa-bajos sutil para el ruido), pero sigue siendo lo suficientemente pequeña para ajustar el contraste según las características de la región (revelando detalles locales).
101x101	La imagen es más oscura en general. El efecto de realce local se reduce, acercándose al resultado de una ecualización global.	Una ventana demasiado grande abarca un espectro de píxeles tan amplio que su histograma local se asemeja al histograma global de la imagen. Esto debilita la capacidad de ajuste local de la ecualización, perdiendo el beneficio principal de la técnica.

Luego adaptamos el algoritmo a la consigna planteada, que reciba por parámetro una ventana de MxN dimensiones, lo que permite trabajar con estructuras que no son cuadradas (como líneas, texto o texturas orientadas). Al experimentar con la nueva función, encontramos que una ventana de 31x41 (filas x columnas) ofreció resultados particularmente buenos para la imagen.



Tamaño de Ventana	Observación / Comentario en el Informe	Explicación Técnica (Criterio)
31x41 (Filas x Columnas)	Optamos por una ventana rectangular para adaptarnos mejor a los detalles de la imagen. Este tamaño nos dio el mejor equilibrio porque realza los detalles finos ocultos sin amplificar el ruido granulado en exceso.	El criterio fue usar las dos dimensiones de forma diferente: 1. El lado más largo (41 de ancho): Se utiliza para suavizar el ruido y las variaciones lentas. Al abarcar más píxeles, la ecualización "promedia" el ruido de fondo, actuando como un filtro suave. 2. El lado más corto (31 de alto): se mantuvo pequeño para asegurar que el ajuste de contraste siguiera siendo altamente local a lo largo del eje vertical, preservando así la nitidez de los bordes y evitando que la imagen se viera borrosa, logrando un balance óptimo entre reducción de ruido y fidelidad del detalle.

En resumen, la ecualización local de histograma demostró ser útil para revelar detalles invisibles con una elección adecuada del tamaño de ventana, aunque sufre limitaciones de rendimiento y sensibilidad al contenido de la imagen.

Ejercicio 2: Validación Automática de Formularios

El objetivo principal de este ejercicio fue desarrollar un sistema capaz de validar automáticamente formularios escaneados, replicando el proceso de revisión manual.

Detección y Segmentación Estructural

El principal reto inicial fue detectar y segmentar correctamente la estructura (celdas). Las variaciones en el grosor de las líneas del formulario o la presencia de trazos de escritura similares a líneas verticales complicaron la división precisa.

Para resolverlo, se empleó la técnica de proyección de píxeles con umbralización adaptativa para localizar las líneas horizontales. El criterio se basó en que las líneas estructurales, al ser continuas, acumulan una suma de píxeles considerablemente mayor que el texto. Posteriormente, para distinguir las líneas divisorias verticales de los trazos de texto, se aplicó un filtrado por altura: una componente vertical dentro de una fila que ocupara al menos el 70% de la altura de la celda se consideró una línea estructural.

FORMULARIO A		
Nombre y apellido		
Edad		
Mail		
Legajo		
	Si	No
Pregunta 1		
Pregunta 2		
Pregunta 3		
Comentarios		

Formulario Original

FORMULARIO A		
Nombre y apellido	JUAN PEREZ	
Edad	45	
Mail	JUAN_PEREZ@GMAIL.COM	
Legajo	P-3205/1	
	Si	No
Pregunta 1	X	
Pregunta 2		X
Pregunta 3	X	
Comentarios	ESTE ES MI COMENTARIO.	

Detección de líneas horizontales

Esta figura ilustra la primera etapa del algoritmo: la detección de las líneas divisorias horizontales mediante la técnica de proyección de píxeles.

- **Criterio Aplicado:** Las filas que acumulan una alta suma de píxeles negros (superando el umbral del 50% del máximo) son identificadas como límites de la celda.
- **Resultado:** Las líneas punteadas marcan las posiciones centrales detectadas. Esta detección es esencial, ya que define el límite vertical de cada fila para la posterior segmentación horizontal.

FORMULARIO A		
Nombre y apellido	JUAN PEREZ	
Edad	45	
Mail	JUAN_PEREZ@GMAIL.COM	
Legajo	P-3205/1	
	Si	No
Pregunta 1	X	
Pregunta 2		X
Pregunta 3	X	
Comentarios	ESTE ES MI COMENTARIO.	

Detección de columnas estructurales largas

Esta imagen muestra las columnas divisorias verticales que delimitan los campos de datos: (Columna 1: Etiquetas, Columna 2: Respuestas).

- **Procedimiento:** Estas líneas se detectaron mediante una proyección de píxeles vertical aplicada a toda la altura del formulario.
- **Criterio Aplicado:** Se seleccionan sólo aquellas líneas que se extienden verticalmente por la mayor parte del formulario (sin incluir las líneas cortas de las preguntas SI/NO). Estas son las columnas estructurales principales.

FORMULARIO A		
Nombre y apellido	JUAN PEREZ	
Edad	45	
Mail	JUAN_PEREZ@GMAIL.COM	
Legajo	P-3205/1	
	Si	No
Pregunta 1	X	
Pregunta 2		X
Pregunta 3	X	
Comentarios	ESTE ES MI COMENTARIO.	

Detección de todas las columnas

Esta figura muestra la detección completa de todas las columnas estructurales del formulario.

- **Diferencia Clave:** A diferencia de la imagen anterior, esta incluye las columnas internas más cortas (las que separan "Sí" de "No" y sus límites en el bloque de preguntas).
- **Criterio Aplicado (Filtro de Altura):** Para identificar las líneas divisorias dentro del área de las preguntas, se aplicó un filtro de altura que solo considera como línea estructural a aquella componente vertical que ocupa al menos el 70% de la altura de la fila. Esto permite aislar correctamente las sub-celdas del bloque SI/NO, mientras descarta picos cortos que podrían ser trazos de texto.

Mail	JUAN_PEREZ@GMAIL.COM
------	----------------------

Extracción de celda (ejemplo)

Esta imagen ilustra la extracción final de una celda de interés, el resultado del proceso combinado de segmentación horizontal y vertical.

- **Procedimiento:** Una vez que las líneas horizontales y verticales han sido detectadas, la celda de la dirección "Mail" es recortada con precisión de la imagen de entrada.
- **Propósito:** Este recorte aísla completamente el campo para que las etapas posteriores del algoritmo (conteo de caracteres, conteo de palabras y validación

semántica/estructural) puedan operar sobre el contenido sin interferencia de las líneas divisorias o campos vecinos.

Análisis de Contenido (Caracteres y Palabras)

Después de segmentar la estructura, se procedió a cuantificar los caracteres y palabras en cada campo.

- **Conteo de Caracteres:** Se empleó el etiquetado de Componentes Conectadas con conectividad-8 para identificar cada trazo de escritura como un objeto. Se aplicó un filtro por área mínima (10 píxeles) para eliminar ruido y detectar celdas vacías.
- **Separación de Palabras:** Para distinguir los espacios entre letras de los espacios entre palabras, se utilizó la distancia relativa entre componentes conectadas. El umbral de espacio se calibró empíricamente a 0.6 veces el ancho promedio de los caracteres; cualquier separación superior a este valor se interpretó como una nueva palabra.

Clasificación y Validación Lógica

- **Validación SI/NO:** La validación de preguntas binarias se realizó mediante lógica XOR. El criterio fue doble: asegurar que solo una celda estuviera marcada (exclusividad) y que esta celda contuviera una única marca, es decir, un solo carácter.

Clasificación de Formulario: Se simplificó la clasificación del formulario midiendo el área absoluta de la letra del encabezado (A, B o C) a través de un proceso de calibración empírica.

Formulario A		
FORMULARIO A		
Nombre y apellido	JUAN PEREZ	
Edad	45	
Mail	JUAN_PEREZ@GMAIL.COM	
Legajo	P-3205/1	
	Si	No
Pregunta 1	X	
Pregunta 2		X
Pregunta 3	X	
Comentarios	ESTE ES MI COMENTARIO.	

Extracción Área del Tipo de Formulario

Resultados y Limitaciones

Se identificaron problemas de robustez, como:

- **Fusión de Caracteres:** A pesar de usar conectividad-8 y el filtro por área mínima, la proximidad o el contacto físico de ciertos caracteres (ej., 'LA') hizo que el sistema los interpretara como una única componente conectada, llevando a un conteo erróneo.



- **Sensibilidad a Parámetros:** El rendimiento del sistema dependió en gran medida de la calidad del escaneo y de la calibración precisa de los parámetros de umbral, área mínima y factor de separación de palabras.

En líneas generales, se logró un sistema que valide la estructura y los campos según las reglas predefinidas.

Conclusiones Generales

El desarrollo de los ejercicios integró conceptos fundamentales del procesamiento de imágenes, abarcando desde el mejoramiento de contraste (ecualización) hasta la extracción de atributos (segmentación y conectividad) y la validación estructural.

El trabajo reforzó la importancia crítica del pre-procesamiento y la experimentación con parámetros. El éxito de las soluciones dependió directamente de dos pilares conceptuales:

Criterios de Diseño y Justificación de Parámetros

Se comprobó que técnicas que a priori parecen simples pueden ser altamente efectivas si se ajustan a un criterio bien definido:

- **Segmentación Estructural:** La proyección de píxeles demostró ser un criterio robusto para aislar elementos basados en su extensión (líneas y. texto).
- **Control de Calidad (E1):** La elección del tamaño de la ventana MxN en la ecualización local se convirtió en un criterio de diseño esencial, ya que determina el balance entre la amplificación del ruido (ventanas pequeñas) y la pérdida de localidad (ventanas grandes).
- **Análisis de Contenido (E2):** El análisis de componentes conectadas permitió extraer atributos de bajo nivel (área, posición). Esto fue crucial para calibrar dinámicamente el umbral de separación de palabras, basándose en la distancia relativa entre caracteres .

El Balance Fundamental de PDI

Ambos ejercicios evidenciaron la necesidad constante de negociar tres factores clave en cualquier solución de procesamiento digital:

- **Precisión vs. Ruido:** En la ecualización local, la búsqueda de mayor detalle aumentó intrínsecamente la sensibilidad al ruido.
- **Eficiencia vs. Detalle:** La implementación por ventana deslizante del Ejercicio 1 impuso una limitación de rendimiento, demostrando que el análisis fino del detalle tiene un alto costo computacional.
- **Robustez vs. Sensibilidad:** La validación de formularios demostró una alta dependencia de la calidad de escaneo y una fragilidad inherente en la clasificación (ej. por área del encabezado) ante la falta de validaciones semánticas.

En conclusión, el desarrollo del trabajo práctico logró con éxito integrar y aplicar conceptos fundamentales del procesamiento de imágenes: ecualización, segmentación, conectividad y validación estructural.