

Universidad de Costa Rica
Facultad de Ingeniería
Escuela de Ciencias de la Computación e Informática

CI-0131 – Diseño de Experimentos

Reporte final:
Análisis del efecto del ruido en algoritmos de compresión de imágenes

por

Josué Amador Rojas - B80435
Maxine Calderón Quirós - C11392
Randy Rojas Pérez - C36937
Queene Zavala Morales - A77201

Ciudad Universitaria Rodrigo Facio

Junio de 2026

Índice general

Índice general	ii
Índice de figuras	iv
Índice de tablas	iv
1 Contexto y propuesta del experimento	1
1.1. Introducción	1
1.2. Planteamiento del problema	1
1.3. Preguntas de investigación	2
1.4. Objetivos	2
1.4.1. Objetivo general	2
1.4.2. Objetivos específicos	2
1.5. Marco teórico	3
1.5.1. Compresión de imágenes sin pérdida: PNG y WebP lossless	3
1.5.2. Ruido gaussiano en imágenes digitales	3
1.5.3. Efecto del ruido sobre la compresión sin pérdida	4
1.6. Enfoque experimental	4
2 Diseño experimental	5
2.1. Tipo de diseño	5
2.1.1. Justificación del bloqueo	5
2.2. Unidad experimental	5
2.3. Variable de respuesta	6
2.4. Factores y niveles	6
2.5. Modelo de datos	6
2.6. Hipótesis	7
2.6.1. Hipótesis para el factor A: Algoritmo de compresión	7
2.6.2. Hipótesis para el factor B: Nivel de ruido	7
2.6.3. Hipótesis para la interacción entre factores A y B	7
2.7. Procedimiento de aleatorización	8

2.8. Datos del experimento	8
3 Análisis exploratorio de datos	9
3.1. Variable de respuesta	9
3.1.1. Estadísticas descriptivas globales	9
3.1.2. Histograma	10
3.1.3. Diagrama de caja	11
3.1.4. Observaciones en el orden de corrida	12
3.2. Variable de respuesta por factor y niveles	12
3.2.1. Factor A: Algoritmo de compresión	12
3.2.2. Factor B: Nivel de ruido gaussiano	13
3.2.3. Diagrama de caja combinado	14
3.3. Gráficos de interacción	16
4 Análisis preliminar	18
4.1. Apreciaciones sobre el factor A	18
4.2. Apreciaciones sobre el factor B	18
4.3. Apreciaciones sobre la interacción	18
4.4. Consideraciones adicionales	19
4.5. Resumen de apreciaciones	19
5 Comprobación de supuestos para el experimento	20
5.1. Comprobación del supuesto de independencia	20
5.1.1. Análisis de los residuales contra el tiempo	20
5.2. Comprobación del supuesto de normalidad	21
5.2.1. Análisis de histograma	21
5.2.2. Análisis de gráfico Q-Q	22
5.2.3. Prueba estadística Shapiro-Wilk	23
5.2.4. Análisis de observaciones atípicas	23
5.3. Comprobación del supuesto de homocedasticidad	24
5.3.1. Análisis de residuales contra valores ajustados	24
5.3.2. Prueba estadística Levene	25
6 Pruebas no paramétricas y análisis de efectos	26
6.1. Modelo no paramétrico mediante ARTool	26
6.2. Reporte del análisis ARTool	26
6.3. Comparaciones múltiples post hoc	27
6.4. Tamaño del efecto	29
6.5. Potencia de la prueba	29
7 Conclusiones	31

Índice de figuras

3.1. Histograma de la TCP. Las líneas discontinua y punteada indican la media (115.80 pp) y la mediana (103.17 pp), respectivamente.	10
3.2. Diagrama de caja de la variable de respuesta. Los puntos corresponden a observaciones individuales superpuestas (<i>jitter</i>).	11
3.3. TCP de cada observación en el orden aleatorizado de corrida, coloreada por algoritmo. La línea discontinua indica la media global (115.80 pp).	12
3.4. Diagrama de caja de la TCP por nivel del factor A (algoritmo de compresión), con observaciones individuales superpuestas.	13
3.5. Diagrama de caja de la TCP por nivel del factor B (nivel de ruido gaussiano), con observaciones individuales superpuestas.	14
3.6. Diagrama de caja de la TCP por combinación de tratamientos (Factor $A \times$ Factor B), con observaciones individuales superpuestas.	15
3.7. Gráfico de interacción: TCP media según el algoritmo para cada nivel de ruido.	16
3.8. Gráfico de interacción: TCP media según el nivel de ruido para cada algoritmo.	17
5.1. Gráfico de residuales contra el orden de corrida	21
5.2. Histograma de residuales	22
5.3. Gráfico Q-Q de los residuales	23
5.4. Imágenes asociadas a las observaciones atípicas identificadas durante el análisis de residuales.	24
5.5. Gráfico de residuales contra valores ajustados	25

Índice de tablas

2.1. Factores, niveles y bloque considerados en el experimento.	6
2.2. Muestra de las primeras observaciones en orden de corrida.	8

3.1. Estadísticas descriptivas globales de la variable de respuesta (TCP, en pp).	9
3.2. Estadísticas descriptivas de la TCP por nivel del factor A (algoritmo).	12
3.3. Estadísticas descriptivas de la TCP por nivel del factor B (nivel de ruido).	13
3.4. Estadísticas descriptivas de la TCP por combinación de tratamientos (pp).	15
6.1. Resultados del análisis mediante ARTool.	27
6.2. Decisión sobre las hipótesis del diseño experimental.	27
6.3. Comparaciones múltiples entre combinaciones de algoritmo y nivel de ruido con ajuste de Tukey.	28
6.4. Comparación múltiple para el factor algoritmo con ajuste de Tukey.	28
6.5. Comparación múltiple para el factor nivel de ruido con ajuste de Tukey.	28
6.6. Tamaño del efecto para los factores y la interacción.	29
6.7. Potencia observada y tamaño de muestra aproximado para alcanzar una potencia de 0.80.	30

Contexto y propuesta del experimento

Este capítulo presenta el contexto general del experimento. Primero se introduce el tema de la compresión de imágenes sin pérdida y el efecto que puede tener el ruido sobre este proceso. Luego se plantea el problema de investigación, las preguntas, los objetivos y el enfoque experimental que se seguirá en el trabajo.

1.1. Introducción

El aumento en la cantidad de datos digitales ha hecho necesario utilizar técnicas eficientes para almacenar y transmitir imágenes. En general, la compresión puede ser con pérdida de información (*lossy*) o sin pérdida de información (*lossless*) [4, 11]. En áreas como la imagenología médica o la astronomía, la información original debe conservarse exactamente, por lo que se requiere compresión sin pérdida [9].

La compresión sin pérdida busca reducir el tamaño de una imagen sin modificar sus píxeles originales. Para lograrlo, aprovecha patrones, repeticiones y relaciones entre píxeles cercanos. Algoritmos como *Run-Length Encoding* (RLE), *Lempel-Ziv-Welch* (LZW), Huffman y la codificación aritmética se basan en esta idea [9, 11]. Cuando una imagen tiene zonas uniformes o patrones repetidos, suele ser más fácil de comprimir. En cambio, cuando la imagen tiene mucha variación, la compresión se vuelve más difícil [6].

Uno de los factores que puede afectar esa regularidad es el ruido digital. En el procesamiento de imágenes, el ruido se entiende como una variación no deseada en la intensidad o el color de los píxeles, que puede aparecer durante la captura o transmisión de la imagen [12]. Al introducir variaciones aleatorias, el ruido puede reducir los patrones que utilizan los algoritmos de compresión sin pérdida, lo que puede aumentar el tamaño del archivo comprimido [1, 6].

En este trabajo se estudia cómo el ruido gaussiano, aplicado en distintos niveles, afecta el tamaño comprimido de imágenes TIFF al utilizar dos algoritmos sin pérdida: PNG y WebP *lossless*. Para ello, se plantea un diseño factorial completo con bloqueo, cuyo detalle se desarrolla en los capítulos siguientes.

1.2. Planteamiento del problema

Los algoritmos de compresión sin pérdida dependen de la presencia de patrones y redundancia en las imágenes. Si una imagen tiene regiones uniformes o píxeles vecinos similares, el algoritmo puede representarla de forma más compacta. Por el contrario, si la imagen contiene muchas variaciones aleatorias, el archivo comprimido tiende a ser más grande [9, 11].

El ruido digital afecta precisamente esta regularidad. Al modificar los valores de los píxeles de forma aleatoria, puede hacer que una imagen sea menos predecible y, por tanto, más difícil de comprimir. En consecuencia, conforme aumenta el nivel de ruido, se espera que también aumente el tamaño del archivo comprimido [5,6].

El problema central de esta investigación es determinar en qué medida el nivel de ruido gaussiano afecta el tamaño resultante del archivo comprimido y si este efecto se comporta de la misma forma para distintos algoritmos sin pérdida. En particular, interesa comparar PNG y WebP lossless, ya que ambos realizan compresión sin pérdida, pero no necesariamente responden igual ante imágenes con ruido.

Además, el contenido visual de la imagen también puede influir en la compresión. Por esta razón, el tipo de contenido se considera como bloque en el diseño experimental, con el fin de controlar esta fuente de variabilidad [13].

En síntesis, se busca estudiar el efecto del algoritmo de compresión, el nivel de ruido y la posible interacción entre ambos sobre el incremento porcentual del tamaño del archivo comprimido. Este análisis permitirá determinar si los factores estudiados influyen de forma significativa en la variable de respuesta.

1.3. Preguntas de investigación

Las siguientes preguntas guían el diseño del experimento y delimitan el alcance del análisis estadístico:

1. ¿Cómo afecta el algoritmo de compresión sin pérdida utilizado (PNG o WebP lossless) al crecimiento porcentual del tamaño del archivo comprimido?
2. ¿Cómo afecta el nivel de ruido gaussiano aplicado a la imagen al crecimiento porcentual del tamaño del archivo comprimido?
3. ¿Existe interacción entre el algoritmo de compresión y el nivel de ruido; es decir, el efecto del ruido cambia según el algoritmo utilizado?

1.4. Objetivos

1.4.1. Objetivo general

Determinar empíricamente el efecto del algoritmo de compresión sin pérdida y del nivel de ruido gaussiano sobre la tasa de crecimiento porcentual del tamaño del archivo comprimido, así como analizar la posible interacción entre ambos factores, mediante un diseño factorial completo con bloqueo aplicado a imágenes digitales en formato TIFF.

1.4.2. Objetivos específicos

1. Comparar la tasa de crecimiento porcentual del tamaño del archivo obtenida al aplicar el algoritmo PNG y el algoritmo WebP lossless sobre un conjunto de imágenes TIFF, para determinar si el tipo de compresor sin pérdida influye significativamente en dicha métrica.

2. Determinar cómo varía la tasa de crecimiento porcentual al modificar el nivel de ruido gaussiano aplicado a las imágenes entre un nivel bajo y un nivel alto, con el fin de cuantificar el impacto del ruido sobre la compresibilidad de la imagen.
3. Analizar si el efecto del algoritmo de compresión sin pérdida sobre la tasa de crecimiento porcentual depende del nivel de ruido de la imagen, verificando la presencia o ausencia de una interacción estadísticamente significativa entre ambos factores.

1.5. Marco teórico

A continuación se presentan los conceptos necesarios para comprender el experimento: compresión sin pérdida, PNG, WebP lossless, ruido gaussiano y su efecto esperado sobre el tamaño comprimido.

1.5.1. Compresión de imágenes sin pérdida: PNG y WebP lossless

La compresión sin pérdida permite recuperar la imagen original exactamente a partir del archivo comprimido, sin alterar los valores de los píxeles [9]. Para lograrlo, el algoritmo busca patrones o repeticiones dentro de los datos de la imagen y los representa de una forma más compacta.

El formato PNG (*Portable Network Graphics*) es uno de los formatos de compresión sin pérdida más utilizados. Su funcionamiento combina un paso de filtrado de la imagen con un método de compresión llamado DEFLATE, que utiliza LZ77 y codificación Huffman [6]. En términos generales, PNG funciona bien cuando la imagen tiene regiones uniformes o patrones repetidos.

WebP es un formato desarrollado por Google que también permite compresión sin pérdida en su modo *WebP lossless* [2]. Este formato utiliza transformaciones previas y técnicas de codificación para reducir el tamaño del archivo sin perder información. En muchas imágenes, WebP lossless puede obtener archivos de menor tamaño que PNG [2].

Tanto PNG como WebP lossless dependen de que la imagen tenga cierta regularidad. Si la imagen se vuelve más aleatoria, por ejemplo por la presencia de ruido, se espera que ambos algoritmos tengan más dificultad para reducir el tamaño del archivo.

1.5.2. Ruido gaussiano en imágenes digitales

El ruido en imágenes digitales es una alteración no deseada en los valores de intensidad o color de los píxeles. Puede aparecer durante la captura de la imagen, la conversión de la señal o la transmisión de los datos [12]. Existen distintos tipos de ruido, entre ellos el ruido gaussiano, el ruido sal y pimienta y el ruido de Poisson [10].

En este trabajo se utiliza ruido gaussiano. Este tipo de ruido agrega a cada píxel una variación aleatoria que sigue una distribución normal con media cero y varianza σ^2 . Esto puede representarse de la siguiente forma:

$$I_{\text{ruidosa}}(x, y) = I_{\text{original}}(x, y) + \mathcal{N}(0, \sigma^2) \quad (1.1)$$

donde $I_{\text{original}}(x, y)$ es el valor original del píxel, $\mathcal{N}(0, \sigma^2)$ representa el ruido añadido e $I_{\text{ruidosa}}(x, y)$ es el valor resultante. La varianza σ^2 controla la intensidad del ruido: a mayor varianza, mayor es la alteración introducida en la imagen.

Al aumentar el ruido, los valores de los píxeles se vuelven menos regulares. Esto puede hacer que la imagen sea más difícil de comprimir sin pérdida, porque el algoritmo debe conservar exactamente esas variaciones.

1.5.3. Efecto del ruido sobre la compresión sin pérdida

La compresión sin pérdida funciona mejor cuando la imagen contiene patrones, repeticiones o relaciones claras entre píxeles cercanos. Cuando se agrega ruido gaussiano, esos patrones pueden debilitarse, ya que los píxeles vecinos se vuelven menos parecidos entre sí [1, 6]. Como los algoritmos sin pérdida deben conservar todos los valores de la imagen, también deben conservar las variaciones introducidas por el ruido.

Por esta razón, se espera que un mayor nivel de ruido produzca un aumento en el tamaño del archivo comprimido. Este comportamiento ha sido reportado en la literatura: conforme aumenta la variación causada por el ruido, la compresión se vuelve menos eficiente y el archivo comprimido puede acercarse más al tamaño de la imagen original sin comprimir [5, 6].

También es posible que no todos los algoritmos reaccionen igual ante el ruido. Algunos métodos pueden depender más de patrones repetidos, mientras que otros pueden manejar de mejor forma ciertas variaciones en la imagen [9]. Por esta razón, en el experimento no solo interesa estudiar el efecto del nivel de ruido y del algoritmo por separado, sino también si existe una interacción entre ambos factores.

1.6. Enfoque experimental

Con el fin de evaluar el impacto del ruido gaussiano sobre la compresión sin pérdida, se realizó un estudio experimental utilizando imágenes del conjunto de datos RAISE [3].

El experimento compara dos algoritmos de compresión, PNG y WebP lossless, y dos niveles de ruido gaussiano: bajo y alto. La comparación se basa en el cambio porcentual del tamaño del archivo comprimido después de agregar ruido a la imagen.

Dado que el contenido visual puede afectar la compresión, se consideró el tipo de contenido de la imagen como bloque. En particular, se distinguieron imágenes de interior e imágenes de exterior, con el objetivo de controlar esta fuente de variabilidad durante el análisis.

Todas las imágenes fueron procesadas mediante un flujo automatizado implementado en Python utilizando la biblioteca Pillow [7]. Esta biblioteca se usó tanto para agregar el ruido como para guardar las imágenes comprimidas.

El detalle formal del diseño experimental, incluyendo los factores estudiados, la variable de respuesta, el modelo estadístico, las hipótesis y el procedimiento de aleatorización, se presenta en el capítulo 2.

Diseño experimental

2.1. Tipo de diseño

Se empleó un diseño factorial completo 2×2 con bloques y réplicas. Los factores de tratamiento fueron el algoritmo de compresión y el nivel de ruido aplicado a las imágenes. Además, se utilizó como bloque el tipo de contenido de la imagen, con dos niveles: interior y exterior.

Este diseño permite analizar el efecto del algoritmo, el efecto del nivel de ruido y la posible interacción entre ambos factores. El bloque se incluyó para controlar diferencias asociadas al tipo de contenido de la imagen, ya que este puede afectar el comportamiento de la compresión.

Se utilizaron 20 réplicas por cada combinación de tratamientos y bloque, para un total de:

$$2 \times 2 \times 2 \times 20 = 160$$

observaciones experimentales.

2.1.1. Justificación del bloqueo

El contenido visual de una imagen puede influir en su compresión. Por ejemplo, algunas imágenes pueden tener zonas más uniformes, mientras que otras pueden contener más texturas o detalles.

Por esta razón, se incorporó como bloque el tipo de contenido de la escena, separando las imágenes en interiores (*indoor*) y exteriores (*outdoor*). El objetivo del bloqueo no es estudiar este factor como parte principal del experimento, sino controlar parte de la variabilidad que podría introducir.

2.2. Unidad experimental

La unidad experimental corresponde a una imagen digital individual en formato TIFF proveniente del conjunto de datos RAISE [3].

Cada imagen seleccionada se usó una única vez en el experimento y recibió una sola combinación de algoritmo, nivel de ruido y bloque. De esta forma, se evita repetir una misma imagen en varias corridas y se mantiene la independencia entre observaciones.

2.3. Variable de respuesta

La variable de respuesta utilizada es la **Tasa de Crecimiento Porcentual** (TCP), registrada en el conjunto de datos como `response_percent`.

Esta variable mide cuánto aumenta el tamaño del archivo comprimido cuando se agrega ruido a una imagen, en comparación con la misma imagen sin ruido y comprimida con el mismo algoritmo.

La TCP se calcula mediante la ecuación 2.1:

$$\text{TCP} = \frac{b - a}{a} \times 100 \quad (2.1)$$

donde:

- a corresponde al tamaño, en kilobytes (kB), de la imagen original comprimida.
- b corresponde al tamaño, en kilobytes (kB), de la imagen con ruido comprimida.

Valores positivos indican que el ruido aumentó el tamaño del archivo comprimido. Valores cercanos a cero indican que el ruido tuvo poco efecto sobre la compresión.

2.4. Factores y niveles

La tabla 2.1 resume los factores del experimento, sus niveles y el bloque utilizado.

Tabla 2.1: Factores, niveles y bloque considerados en el experimento.

Tipo	Símbolo	Factor	Nivel 1	Nivel 2
Tratamiento	A	Algoritmo de compresión	PNG	WebP (lossless)
Tratamiento	B	Nivel de ruido	Bajo (1 %–5 %)	Alto (45 %–50 %)
Bloque	–	Tipo de contenido de imagen	Indoor (interior)	Outdoor (exterior)

2.5. Modelo de datos

El modelo de efectos para el diseño factorial 2×2 con bloques se muestra en la ecuación 2.2.

$$Y_{ijkl} = \mu + \gamma_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijkl} \quad (2.2)$$

donde:

- Y_{ijkl} representa la TCP observada en la réplica l , dentro del bloque i , con el algoritmo j y el nivel de ruido k .
- μ representa la media general del experimento.

- γ_i representa el efecto del bloque, donde $i = 1, 2$ corresponde a interior y exterior.
- α_j representa el efecto del algoritmo de compresión, donde $j = 1, 2$ corresponde a PNG y WebP lossless.
- β_k representa el efecto del nivel de ruido, donde $k = 1, 2$ corresponde a bajo y alto.
- $(\alpha\beta)_{jk}$ representa la interacción entre el algoritmo y el nivel de ruido.
- ε_{ijkl} representa el error aleatorio de la réplica l , con $l = 1, 2, \dots, 20$.

2.6. Hipótesis

2.6.1. Hipótesis para el factor A: Algoritmo de compresión

$$H_{0A} : \alpha_1 = \alpha_2 = 0$$

$$H_{1A} : \alpha_j \neq 0 \quad \text{para al menos un } j$$

La hipótesis nula indica que el algoritmo de compresión no tiene efecto sobre la variable de respuesta. La hipótesis alternativa indica que al menos un algoritmo sí tiene efecto.

2.6.2. Hipótesis para el factor B: Nivel de ruido

$$H_{0B} : \beta_1 = \beta_2 = 0$$

$$H_{1B} : \beta_k \neq 0 \quad \text{para al menos un } k$$

La hipótesis nula indica que el nivel de ruido no tiene efecto sobre la variable de respuesta. La hipótesis alternativa indica que al menos un nivel de ruido sí tiene efecto.

2.6.3. Hipótesis para la interacción entre factores A y B

$$H_{0AB} : (\alpha\beta)_{jk} = 0 \quad \text{para todo } j, k$$

$$H_{1AB} : (\alpha\beta)_{jk} \neq 0 \quad \text{para al menos un par } (j, k)$$

La hipótesis nula indica que no existe interacción entre el algoritmo y el nivel de ruido. La hipótesis alternativa indica que sí existe interacción entre ambos factores.

2.7. Procedimiento de aleatorización

Para evitar sesgos en el orden de ejecución, las corridas del experimento se realizaron en orden aleatorio. La aleatorización se hizo en R usando una semilla fija igual a 26, de modo que el procedimiento pudiera reproducirse.

El procedimiento fue el siguiente:

1. Filtrar las imágenes elegibles y separarlas en los bloques **indoor** y **outdoor**.
2. Generar las combinaciones de bloque, algoritmo y nivel de ruido.
3. Repetir cada combinación 20 veces, para obtener las 160 corridas del experimento.
4. Mezclar aleatoriamente el orden de las corridas.
5. Asignar a cada corrida una imagen distinta del bloque correspondiente.

2.8. Datos del experimento

Los datos completos del experimento constan de 160 corridas balanceadas y se encuentran en el archivo **results.csv**. Como ejemplo, la tabla 2.2 muestra las primeras 5 observaciones en el orden aleatorio utilizado. Las columnas que no son necesarias para esta sección se omiten.

Tabla 2.2: Muestra de las primeras observaciones en orden de corrida.

Corrida	Nombre de archivo	Bloque	Algoritmo	Nivel de ruido	Respuesta (%)
1	r00d161b7t.TIF	indoor	WebP	high	135.61
2	r0141f0c9t.TIF	indoor	WebP	low	71.26
3	r01511e3et.TIF	indoor	PNG	high	80.81
4	r0225b88bt.TIF	indoor	WebP	high	222.00
5	r007f5116t.TIF	outdoor	PNG	low	90.74
⋮	⋮	⋮	⋮	⋮	⋮

Análisis exploratorio de datos

Este capítulo presenta el análisis exploratorio de las 160 observaciones recolectadas en el experimento factorial 2×2 con bloqueo. El objetivo es observar el comportamiento general de la Tasa de Crecimiento Porcentual (TCP) antes de realizar las pruebas estadísticas. Para ello, se revisan estadísticas descriptivas, diagramas de caja, histogramas, gráficos por factor y gráficos de interacción. Todos los gráficos fueron generados con R [8] a partir del archivo `results.csv`.

3.1. Variable de respuesta

3.1.1. Estadísticas descriptivas globales

La tabla 3.1 resume las estadísticas descriptivas de la TCP considerando las 160 observaciones.

Tabla 3.1: Estadísticas descriptivas globales de la variable de respuesta (TCP, en pp).

Estadístico	Valor (pp)
Observaciones (n)	160
Media ($\bar{Y}$)	115.80
Varianza (s^2)	3 524.74
Desviación estándar (s)	59.37
Mínimo	15.18
Cuartil 1 (Q1)	75.53
Mediana	103.17
Cuartil 3 (Q3)	148.76
Máximo	330.40

La media global fue de 115.80 pp y la mediana fue de 103.17 pp. Como la media es mayor que la mediana, los datos parecen tener una cola hacia la derecha. La desviación estándar fue de 59.37 pp, por lo que existe una variabilidad considerable entre las observaciones. Además, el valor máximo fue de 330.40 pp, bastante mayor que el tercer cuartil, lo cual también sugiere la presencia de valores altos en la cola derecha.

3.1.2. Histograma

La figura 3.1 muestra la distribución de la TCP mediante un histograma.

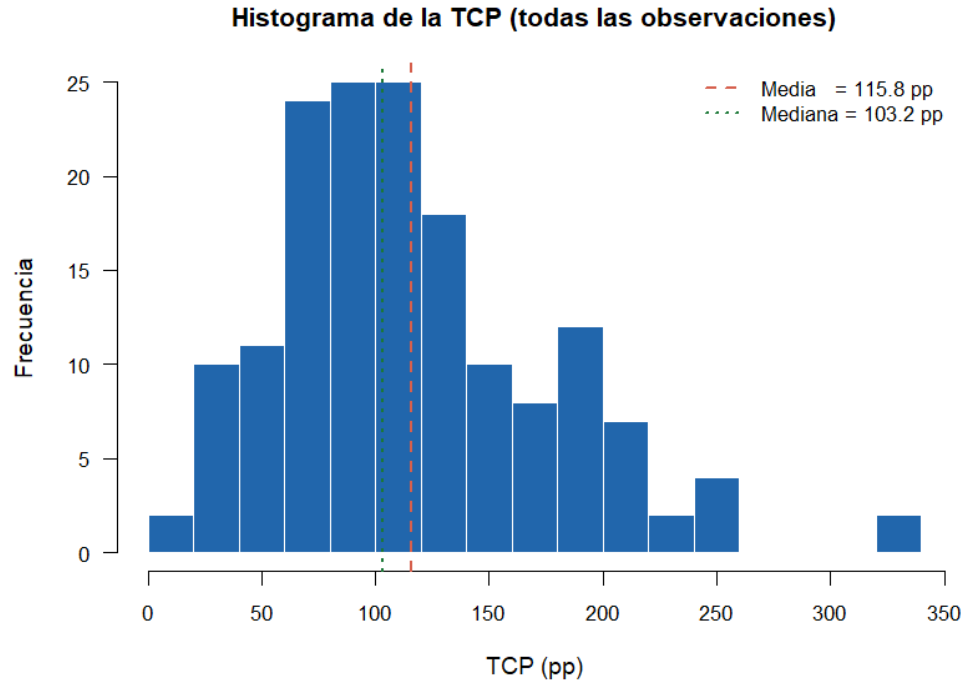


Figura 3.1: Histograma de la TCP. Las líneas discontinua y punteada indican la media (115.80 pp) y la mediana (103.17 pp), respectivamente.

El histograma confirma que la mayoría de observaciones se ubica aproximadamente entre 50 y 150 pp. También se observa una cola hacia la derecha, con algunos valores cercanos a 330 pp. Esto coincide con la diferencia entre la media y la mediana mostrada en la tabla anterior.

3.1.3. Diagrama de caja

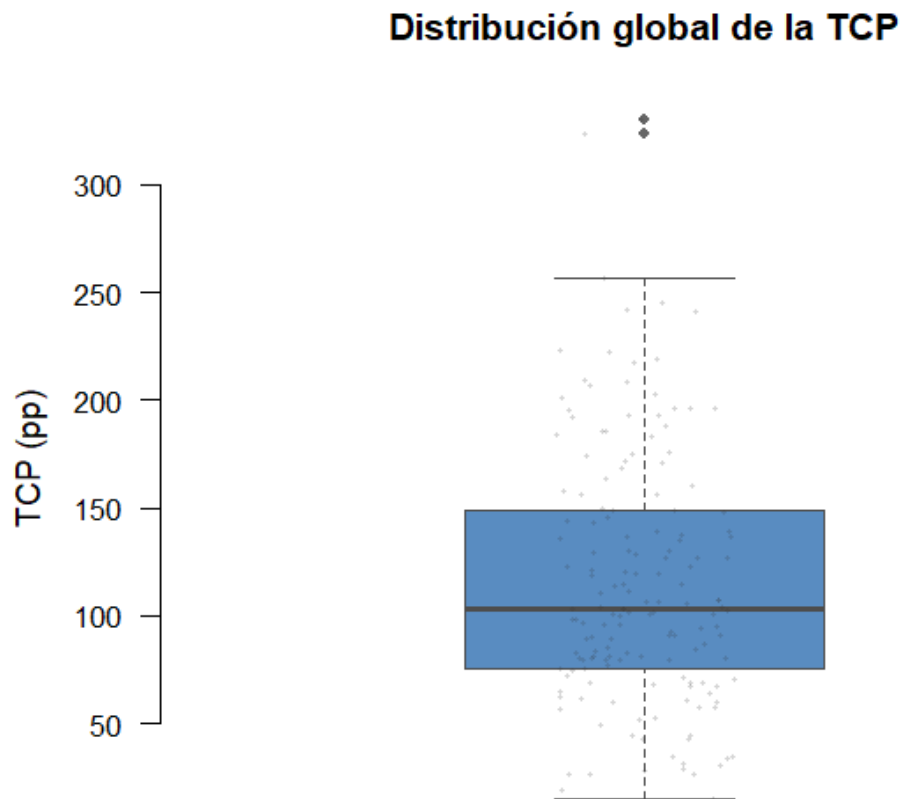


Figura 3.2: Diagrama de caja de la variable de respuesta. Los puntos corresponden a observaciones individuales superpuestas (*jitter*).

El diagrama de caja de la figura 3.2 muestra una distribución asimétrica. El bigote superior es más largo que el inferior y se observan valores atípicos por encima de aproximadamente 270 pp. Esto refuerza lo observado en el histograma: la TCP presenta algunos valores altos que elevan la media.

3.1.4. Observaciones en el orden de corrida

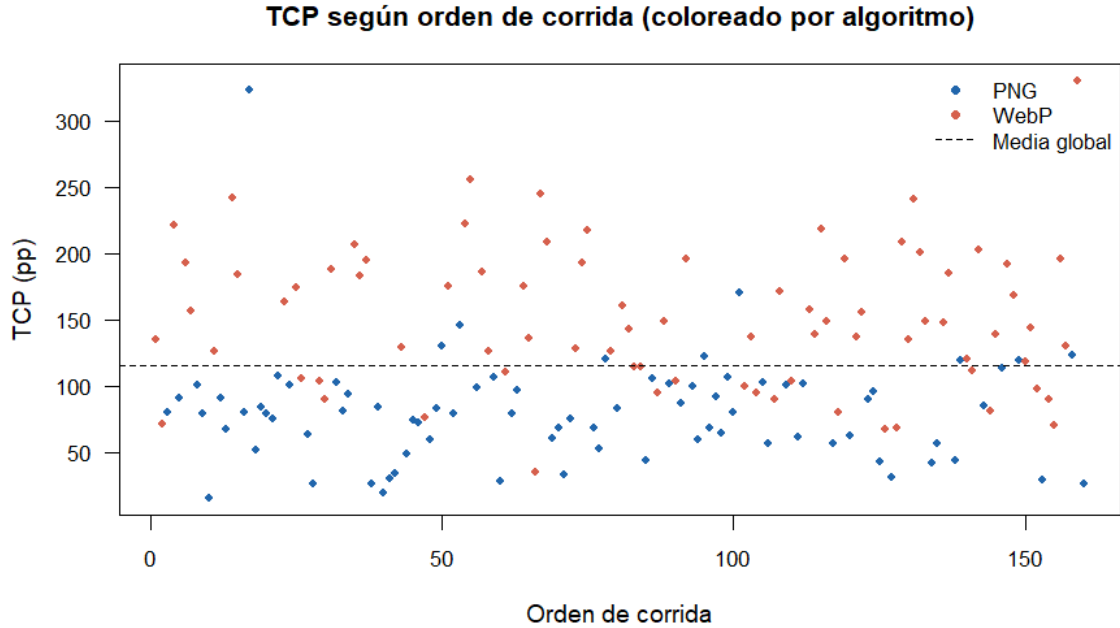


Figura 3.3: TCP de cada observación en el orden aleatorizado de corrida, coloreada por algoritmo. La línea discontinua indica la media global (115.80 pp).

En la figura 3.3 no se observa una tendencia clara a lo largo del orden de corrida. Los valores altos y bajos aparecen distribuidos en distintas partes de la secuencia, por lo que no parece haber un patrón asociado al orden en que se ejecutaron las pruebas. Además, los puntos de PNG y WebP aparecen mezclados, lo cual es consistente con la aleatorización del experimento.

3.2. Variable de respuesta por factor y niveles

3.2.1. Factor A: Algoritmo de compresión

La tabla 3.2 presenta las estadísticas descriptivas de la TCP para cada algoritmo.

Tabla 3.2: Estadísticas descriptivas de la TCP por nivel del factor A (algoritmo).

Nivel	Media	Var.	s	Mín.	Q1	Med.	Q3	Máx.
PNG	80.38	1 715.32	41.42	15.18	57.18	79.81	100.78	323.23
WebP	151.21	2 838.89	53.28	34.78	111.29	145.71	192.25	330.40

WebP presentó una TCP media de 151.21 pp, mientras que PNG presentó una media de 80.38 pp. Por lo tanto, la diferencia entre ambos algoritmos fue de 70.83 pp. En la figura 3.4 también se observa que los valores de WebP tienden a ser mayores que los de PNG, aunque hay algunos traslapes entre los grupos.

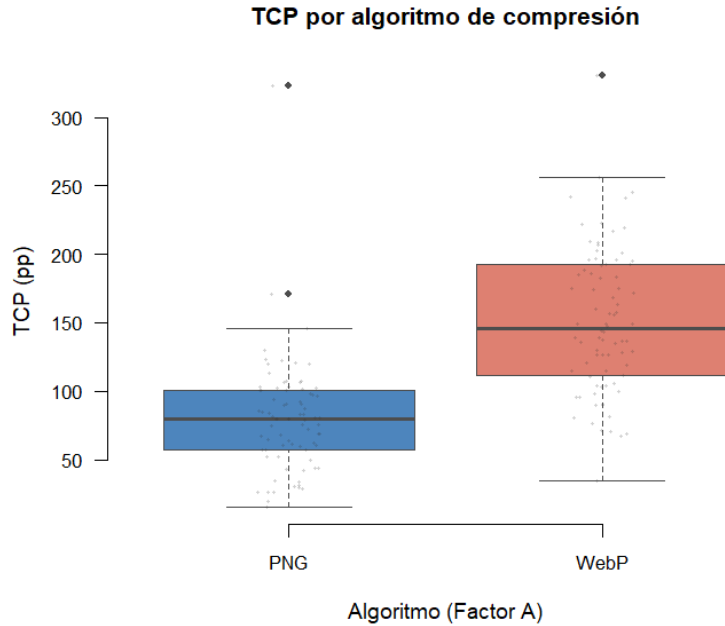


Figura 3.4: Diagrama de caja de la TCP por nivel del factor A (algoritmo de compresión), con observaciones individuales superpuestas.

3.2.2. Factor B: Nivel de ruido gaussiano

La tabla 3.3 presenta las estadísticas descriptivas por nivel de ruido.

Tabla 3.3: Estadísticas descriptivas de la TCP por nivel del factor B (nivel de ruido).

Nivel	Media	Var.	s	Mín.	Q1	Med.	Q3	Máx.
Alto (<i>high</i>)	132.97	3 690.50	60.75	26.25	89.73	120.60	183.54	323.23
Bajo (<i>low</i>)	98.62	2 806.43	52.98	15.18	63.67	89.75	130.23	330.40

El nivel alto de ruido presentó una TCP media de 132.97 pp, mientras que el nivel bajo presentó una media de 98.62 pp. La diferencia entre ambos niveles fue de 34.35 pp. En la figura 3.5 se observa que el nivel alto tiende a producir valores mayores, aunque las distribuciones de ambos niveles se traslapan más que en el caso del algoritmo.

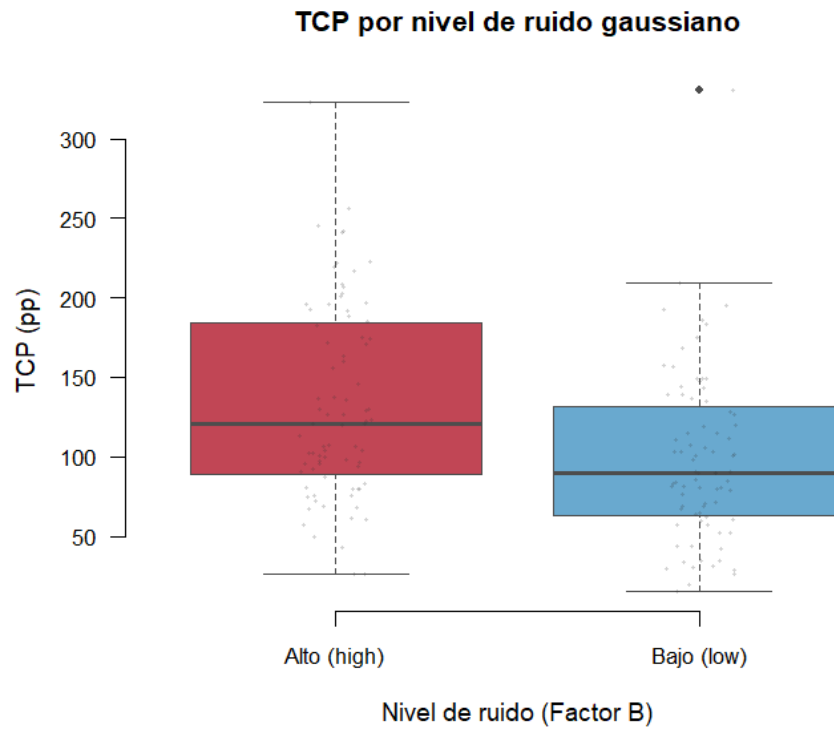


Figura 3.5: Diagrama de caja de la TCP por nivel del factor B (nivel de ruido gaussiano), con observaciones individuales superpuestas.

3.2.3. Diagrama de caja combinado

La figura 3.6 y la tabla 3.4 presentan los resultados para las cuatro combinaciones de algoritmo y nivel de ruido.

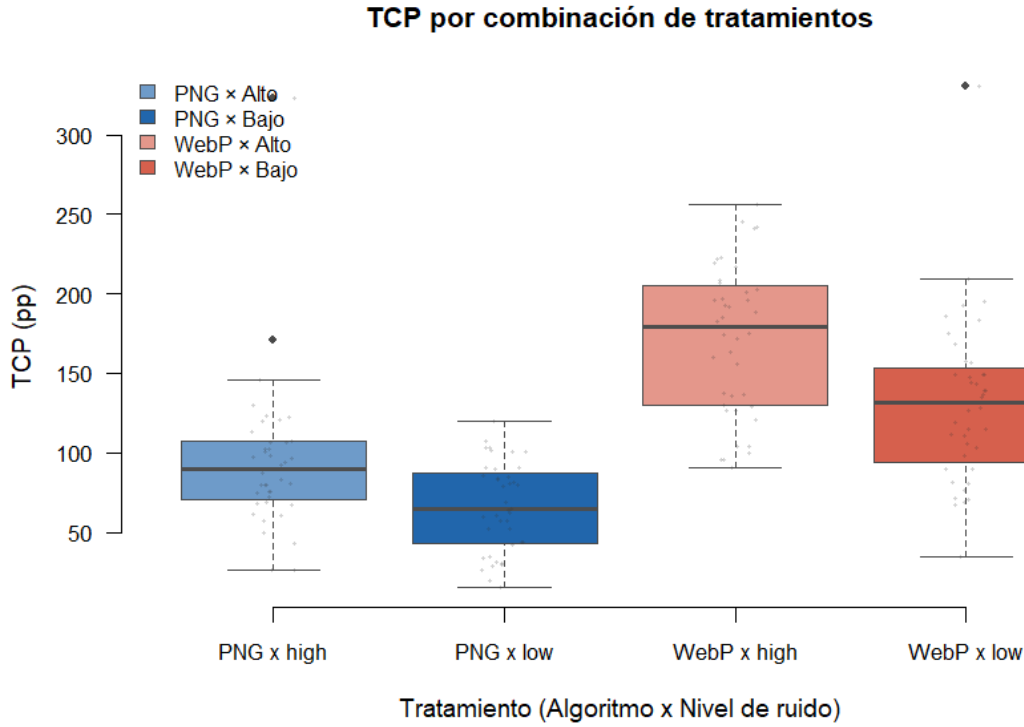


Figura 3.6: Diagrama de caja de la TCP por combinación de tratamientos (Factor $A \times$ Factor B), con observaciones individuales superpuestas.

Tabla 3.4: Estadísticas descriptivas de la TCP por combinación de tratamientos (pp).

Alg.	Ruido	Media	s	Mín.	Med.	Máx.
PNG	Alto	94.73	47.57	26.25	89.83	323.23
PNG	Bajo	66.04	28.09	15.18	64.32	119.60
WebP	Alto	171.20	47.04	90.58	178.94	256.26
WebP	Bajo	131.21	52.13	34.78	131.91	330.40

La combinación WebP con nivel alto de ruido presentó la mayor media, con 171.20 pp. En cambio, PNG con nivel bajo de ruido presentó la menor media, con 66.04 pp. También se observa que PNG con nivel bajo tuvo la menor desviación estándar ($s = 28.09$ pp), mientras que WebP con nivel bajo presentó la mayor desviación estándar ($s = 52.13$ pp) y el valor máximo del experimento.

3.3. Gráficos de interacción

Las figuras 3.7 y 3.8 muestran las medias de las combinaciones de tratamientos. Estos gráficos permiten observar si las líneas son aproximadamente paralelas o si existe algún indicio visual de interacción.

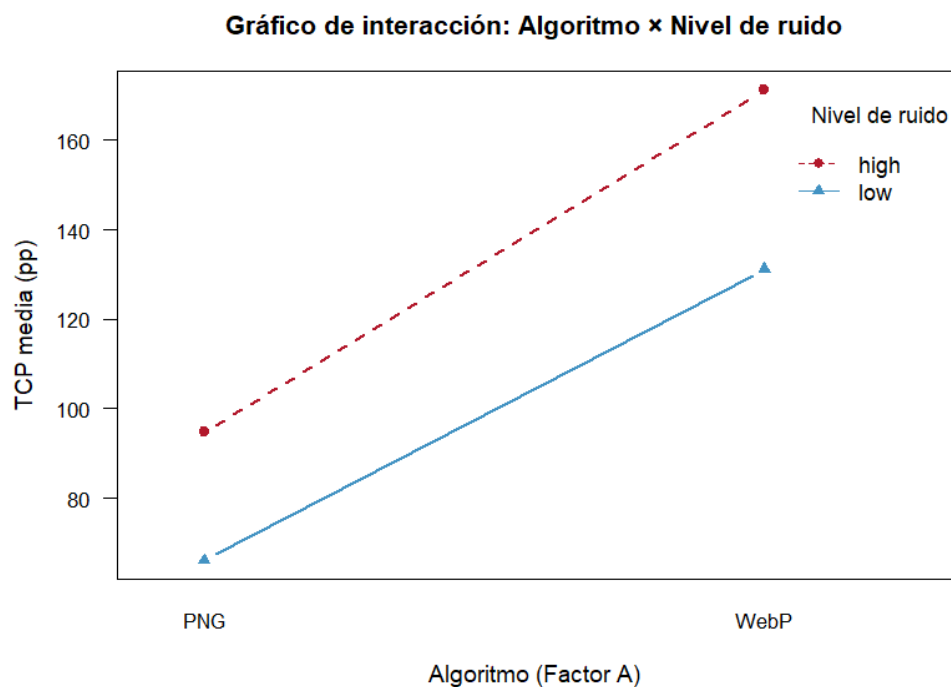


Figura 3.7: Gráfico de interacción: TCP media según el algoritmo para cada nivel de ruido.

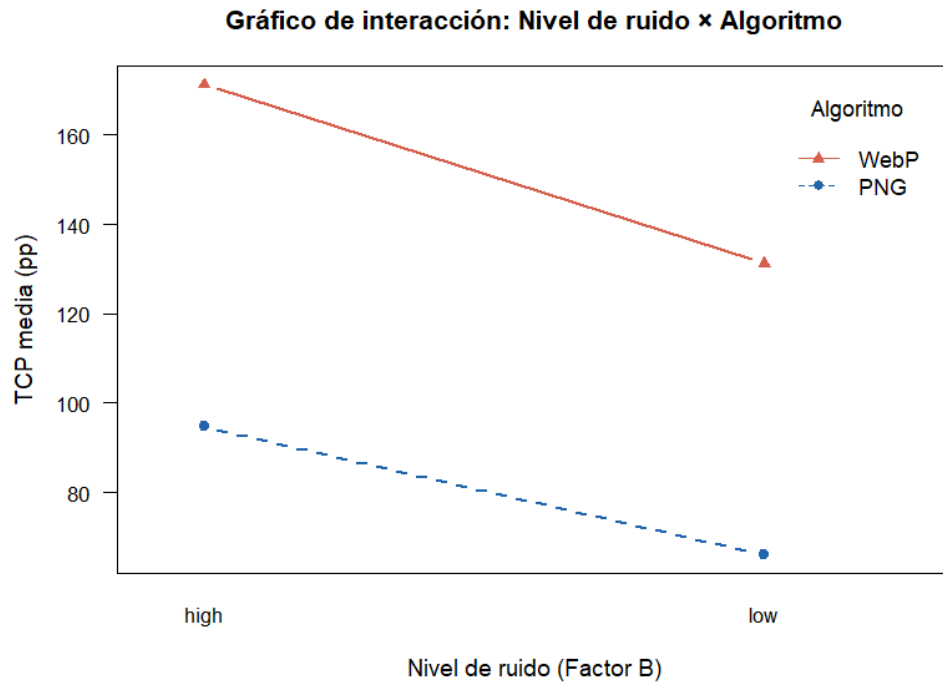


Figura 3.8: Gráfico de interacción: TCP media según el nivel de ruido para cada algoritmo.

En ambos gráficos se observa que las líneas son bastante paralelas. Esto sugiere que no hay una interacción marcada entre los factores, es decir, el efecto del nivel de ruido parece mantenerse de forma similar para PNG y WebP. Para PNG, pasar de nivel bajo a nivel alto aumenta la media en 28.69 pp, mientras que para WebP el aumento es de 39.99 pp. Aunque el aumento es ligeramente mayor en WebP, el patrón general es parecido en ambos algoritmos. Por lo tanto, visualmente no se observa evidencia clara de interacción entre el algoritmo de compresión y el nivel de ruido.

Análisis preliminar

El análisis exploratorio del capítulo anterior permite observar el comportamiento general de la variable de respuesta antes de realizar las pruebas estadísticas. En este capítulo se resumen las principales apreciaciones sobre cada factor y sobre la posible interacción entre ellos.

4.1. Apreciaciones sobre el factor A

Con base en la figura 3.4 y la tabla 3.2, el algoritmo de compresión parece tener un efecto claro sobre la Tasa de Crecimiento Porcentual. La TCP media de WebP lossless fue de 151.21 pp, mientras que la de PNG fue de 80.38 pp. Esto representa una diferencia de 70.83 puntos porcentuales.

Además, el tercer cuartil de PNG (100.78 pp) se encuentra por debajo de la mediana de WebP (145.71 pp). Esto muestra una separación visible entre ambos grupos. En general, PNG presentó incrementos porcentuales menores que WebP lossless, por lo que se espera que el efecto del factor algoritmo sea significativo en el análisis formal.

4.2. Apreciaciones sobre el factor B

Con base en la figura 3.5 y la tabla 3.3, el nivel de ruido también parece influir en la TCP. El nivel alto de ruido presentó una media de 132.97 pp, mientras que el nivel bajo presentó una media de 98.62 pp. La diferencia entre ambos niveles fue de 34.35 puntos porcentuales.

En el diagrama de caja se observa que las distribuciones de ambos niveles se superponen. Aun así, la mediana del nivel alto (120.60 pp) es mayor que la del nivel bajo (89.75 pp). Por esta razón, se espera que el nivel de ruido también tenga un efecto significativo, aunque visualmente parece menor que el efecto del algoritmo.

4.3. Apreciaciones sobre la interacción

Las figuras 3.7 y 3.8 permiten revisar visualmente si existe interacción entre el algoritmo de compresión y el nivel de ruido. En ambos gráficos se observa que las líneas son bastante paralelas, lo cual sugiere que no hay una interacción marcada entre los factores.

Al pasar de ruido bajo a ruido alto, PNG aumenta de 66.04 pp a 94.73 pp, es decir, una diferencia de 28.69 pp. En el caso de WebP, la media aumenta de 131.21 pp a 171.20 pp, para una diferencia de 39.99 pp. Aunque el aumento es un poco mayor en WebP, el patrón general es parecido en ambos algoritmos.

Por lo tanto, visualmente no se observa evidencia clara de interacción entre el algoritmo de compresión y el nivel de ruido. Esta apreciación deberá confirmarse posteriormente con la prueba estadística correspondiente.

4.4. Consideraciones adicionales

El análisis exploratorio mostró algunos aspectos que deben revisarse en el análisis de supuestos:

1. Posible diferencia de varianzas: las varianzas no parecen iguales en todos los grupos. Por ejemplo, la combinación WebP $\times$ bajo presentó la mayor variabilidad ($s^2 = 2717.59$, aproximadamente), mientras que PNG $\times$ bajo presentó la menor ($s^2 = 789.06$). Por esto, será necesario aplicar una prueba formal de homogeneidad de varianzas.
2. Forma de la distribución: el histograma (figura 3.3) y el diagrama de caja (figura 3.2) muestran una cola hacia la derecha, con valores que llegan hasta 330.40 pp. Además, la media (115.80 pp) es mayor que la mediana (103.17 pp). Esto deberá revisarse posteriormente mediante el análisis de normalidad.
3. Valores atípicos: en el diagrama de caja global (figura 3.2) se observan algunos valores altos, especialmente por encima de 270 pp. Estos valores corresponden a la combinación WebP $\times$ bajo. Será importante revisar si estos valores afectan los supuestos del modelo.
4. Aleatorización: la figura 3.3 no muestra un patrón claro asociado al orden de corrida. Esto sugiere que la aleatorización fue adecuada.

4.5. Resumen de apreciaciones

El análisis exploratorio sugiere que el algoritmo de compresión y el nivel de ruido tienen efecto sobre la Tasa de Crecimiento Porcentual. En general, PNG presentó valores menores que WebP lossless, y el nivel alto de ruido produjo valores mayores que el nivel bajo.

En cuanto a la interacción, los gráficos muestran líneas bastante paralelas. Por esta razón, visualmente no se observa una interacción marcada entre el algoritmo y el nivel de ruido. Se espera que esta apreciación sea confirmada en el análisis formal.

Finalmente, también se observaron posibles diferencias de varianza, una cola hacia la derecha y algunos valores atípicos. Estos aspectos deberán revisarse en la validación de supuestos antes de interpretar de forma definitiva los resultados del análisis.

Comprobación de supuestos para el experimento

Antes de realizar el análisis estadístico del experimento, es necesario verificar los supuestos asociados a los métodos de inferencia considerados. En particular, se evaluarán los supuestos de independencia, normalidad y igualdad de varianzas (homocedasticidad) a partir de los residuales del modelo ajustado. La comprobación de estos supuestos permitirá determinar si es apropiado utilizar un análisis paramétrico, como ANOVA, o si resulta más conveniente recurrir a una alternativa no paramétrica. De esta manera, se garantiza que las conclusiones obtenidas estén respaldadas por una metodología consistente con las características observadas en los datos.

Con el fin de verificar los supuestos estadísticos, se ajustó preliminarmente el modelo factorial con bloques y se calcularon sus residuales. Estos residuales fueron utilizados para evaluar la independencia, normalidad y homogeneidad de varianzas, dado que representan estimaciones de los errores experimentales.

5.1. Comprobación del supuesto de independencia

La independencia de las observaciones es un supuesto fundamental en el análisis de experimentos, ya que garantiza que el valor observado en una unidad experimental no influya sobre los demás. En este estudio, dicho supuesto se evaluará mediante el análisis gráfico de los residuales en función del orden de ejecución del experimento, con el fin de identificar posibles patrones, tendencias o correlaciones temporales.

5.1.1. Análisis de los residuales contra el tiempo

En la figura 5.1 podemos observar que no se aprecian patrones aparentes tras graficar los residuales contra el orden en que se ejecutaron las corridas para obtener las observaciones. Esto coincide también con la figura 3.3, donde realizamos un gráfico similar sobre las observaciones.

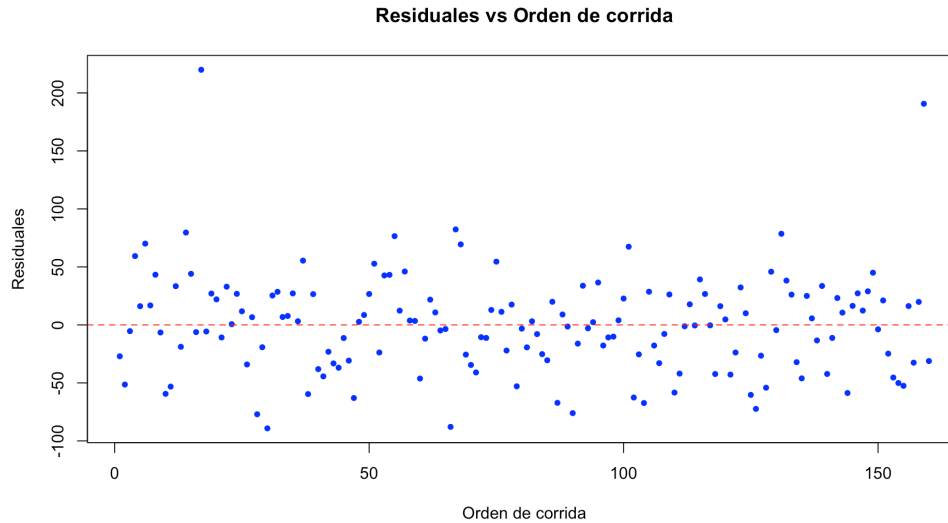


Figura 5.1: Gráfico de residuales contra el orden de corrida

Así, base en el proceso de aleatorización aplicado dentro de los bloques durante el diseño experimental en el capítulo 2 y en la inspección visual del gráfico de residuales en función del orden de corrida, no se observaron patrones sistemáticos, tendencias ni estructuras que sugieran dependencia entre las observaciones. Por lo tanto, **se considera razonable asumir que el supuesto de independencia se cumple para los datos analizados.**

5.2. Comprobación del supuesto de normalidad

La normalidad de los residuales es uno de los supuestos fundamentales de los métodos paramétricos de análisis, ya que garantiza la validez de las pruebas de significancia y de los intervalos de confianza asociados al modelo. Debido a que los errores experimentales no son observables directamente, este supuesto se evaluará a partir de los residuales obtenidos del modelo ajustado. Para ello, se emplearán herramientas gráficas y una pruebas estadística que permitan determinar si la distribución de los residuales es compatible con una distribución normal.

5.2.1. Análisis de histograma

En la figura 5.2, el histograma muestra una concentración importante de observaciones alrededor de cero, lo que inicialmente podría sugerir una distribución aproximadamente simétrica. Sin embargo, también se observa una cola más extensa hacia los valores positivos y la presencia de algunos residuales extremos, particularmente por encima de 150, lo que indica una posible asimetría a la derecha. Estos valores atípicos pueden afectar la forma de la distribución y alejarla del comportamiento normal.

Este comportamiento es muy similar a lo observado en la figura 3.1, donde se grafica el histograma de las observaciones.

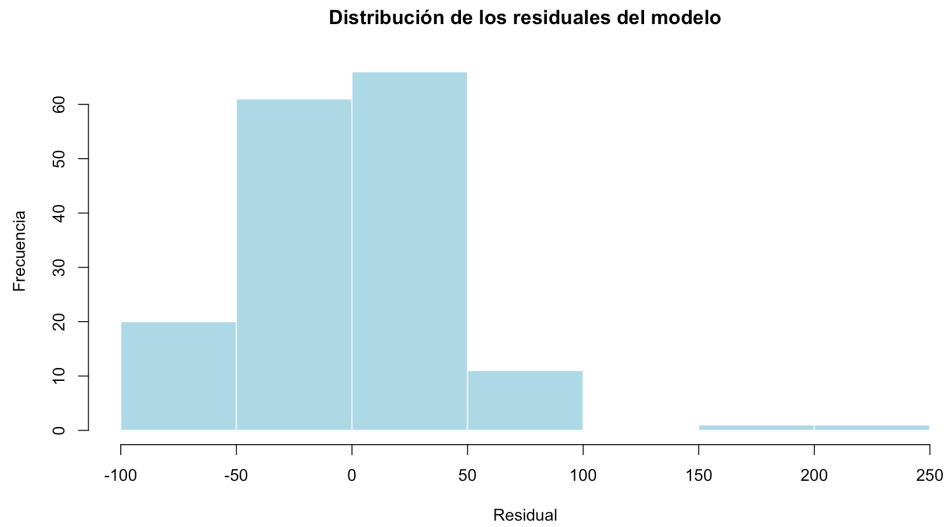


Figura 5.2: Histograma de residuales

Aunque el histograma proporciona una primera aproximación visual, por sí solo no permite concluir de manera definitiva si el supuesto de normalidad se cumple, por lo que resulta necesario complementarlo con herramientas más sensibles, como el gráfico Q-Q.

5.2.2. Análisis de gráfico Q-Q

En la figura 5.3, el gráfico Q-Q muestra que la mayoría de los residuales se alinean razonablemente con la recta de referencia en la región central de la distribución, lo que sugiere un comportamiento cercano a la normalidad para una gran parte de las observaciones. Sin embargo, se observan desviaciones importantes en las colas, especialmente en la cola superior, donde varios puntos se alejan considerablemente de la línea teórica debido a la presencia de residuales positivos extremos.

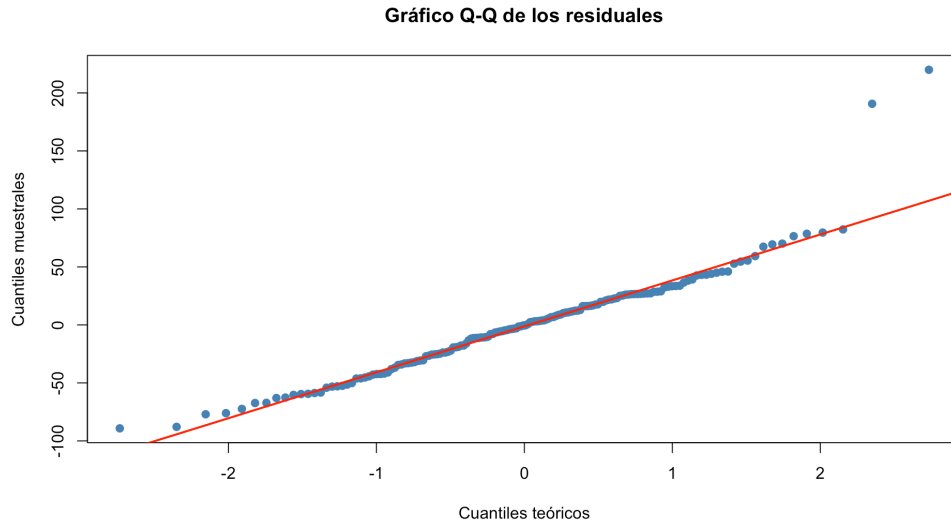


Figura 5.3: Gráfico Q-Q de los residuales

Aunque el histograma y el gráfico Q-Q proporcionan una evaluación visual de la normalidad, resulta conveniente complementar este análisis mediante una prueba estadística formal. Para ello, utilizamos la prueba Shapiro-Wilk.

5.2.3. Prueba estadística Shapiro-Wilk

La prueba de Shapiro-Wilk produjo un estadístico ($W = 0.9265$) y un valor ($p = 2.753 \times 10^{-7}$). Dado que el valor p es considerablemente menor que el nivel de significancia $\alpha = 0.05$, se rechaza la hipótesis nula de normalidad. Por lo tanto, **existe evidencia estadísticamente significativa para concluir que los residuales no siguen una distribución normal**. Este resultado es consistente con lo observado previamente en el histograma y en el gráfico Q-Q, donde se identificaron desviaciones respecto al comportamiento esperado bajo normalidad, especialmente en las colas de la distribución.

5.2.4. Análisis de observaciones atípicas

Con el fin de investigar las posibles causas de la falta de normalidad observada en los residuales, se identificaron las observaciones con los mayores residuales positivos y se realizó una inspección visual de las imágenes asociadas. Se encontró que ambas corresponden a escenas exteriores con una alta proporción de regiones homogéneas y de tonos claros, principalmente nieve, cielo nublado o superficies con muy poca textura. Las imágenes resultantes de estas observaciones se pueden ver en la figura 5.4.

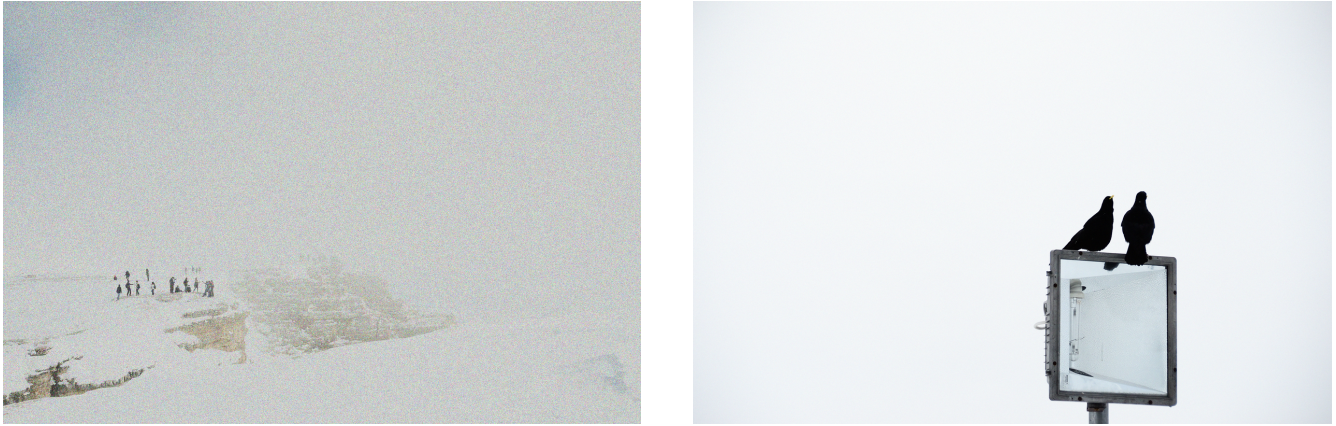


Figura 5.4: Imágenes asociadas a las observaciones atípicas identificadas durante el análisis de residuales.

Este tipo de imágenes suele presentar una alta compresibilidad debido a la uniformidad de sus píxeles. Sin embargo, al introducir ruido, dicha uniformidad se ve alterada, lo que puede provocar aumentos considerablemente mayores en el tamaño de los archivos comprimidos en comparación con imágenes que contienen más detalles y variaciones naturales. Como resultado, estas observaciones generan valores de respuesta significativamente superiores al resto de la muestra.

Dado que no se encontró evidencia de errores de medición, procesamiento o registro de los datos, estas observaciones no fueron eliminadas del análisis y se consideraron parte de la variabilidad natural del fenómeno estudiado. Por lo tanto, se concluye que las desviaciones respecto a la normalidad están asociadas a características propias de ciertas imágenes y no a errores experimentales.

5.3. Comprobación del supuesto de homocedasticidad

La homogeneidad de varianzas, o homocedasticidad, es otro de los supuestos fundamentales de los métodos paramétricos de análisis. Este supuesto establece que la variabilidad de los errores debe mantenerse aproximadamente constante entre los diferentes tratamientos y combinaciones de factores evaluadas. Su verificación es importante, ya que diferencias importantes en las varianzas pueden afectar la validez de las pruebas de significancia y las conclusiones derivadas del modelo. Para evaluar este supuesto se analizará la dispersión de los residuales mediante herramientas gráficas y pruebas estadísticas.

5.3.1. Análisis de residuales contra valores ajustados

En la figura 5.5, el gráfico de residuales contra valores ajustados muestra que los residuales se distribuyen alrededor de cero sin presentar un patrón claro. En general, la dispersión de los puntos es similar entre los distintos valores ajustados, aunque se observan algunos residuales extremos y ciertas diferencias en la amplitud de la dispersión entre grupos. A simple vista no se aprecia una violación evidente del supuesto de homogeneidad de varianzas; sin embargo, estas diferencias sugieren que es conveniente realizar una prueba estadística formal, como la prueba de Levene, para evaluar este supuesto con mayor precisión.

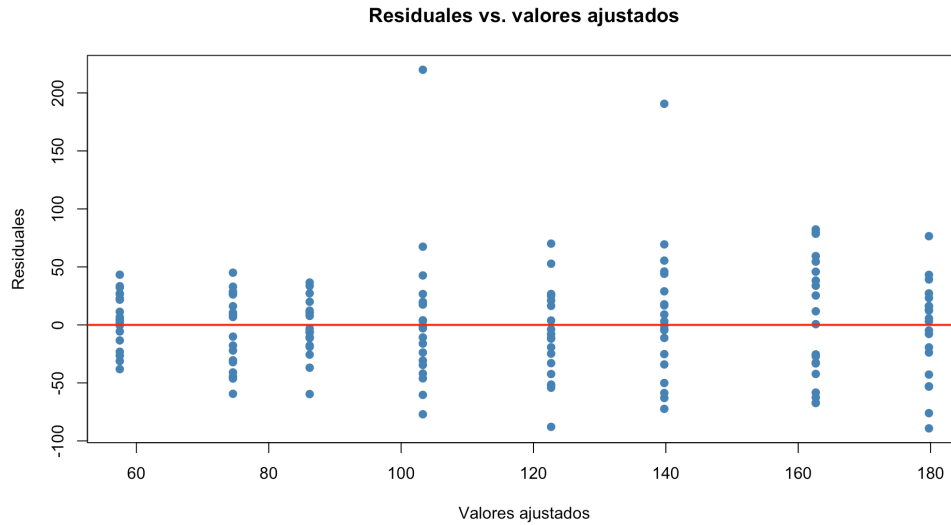


Figura 5.5: Gráfico de residuales contra valores ajustados

5.3.2. Prueba estadística Levene

La prueba de Levene produjo un valor $p = 0.0796$. Dado que este valor es mayor que el nivel de significancia de 0.05, no se rechaza la hipótesis nula de igualdad de varianzas entre los grupos. Por lo tanto, no se encontró evidencia estadísticamente significativa que indique diferencias importantes en la variabilidad de los residuales. Este resultado es consistente con lo observado en el gráfico de residuales versus valores ajustados, donde no se apreciaba un patrón claro de aumento o disminución de la dispersión. En consecuencia, **el supuesto de homogeneidad de varianzas puede considerarse razonablemente satisfecho para los datos analizados.**

Pruebas no paramétricas y análisis de efectos

En el capítulo 5 se verificaron los supuestos requeridos para aplicar un análisis de varianza paramétrico. A partir de dicha revisión, se consideró razonable asumir independencia entre las observaciones y homogeneidad de varianzas. Sin embargo, la prueba de Shapiro-Wilk aplicada sobre los residuales produjo un estadístico $W = 0.9265$ y un valor $p = 2.753 \times 10^{-7}$. Por lo tanto, con un nivel de significancia $\alpha = 0.05$, se rechazó la hipótesis nula de normalidad. Además, el histograma y el gráfico Q-Q mostraron desviaciones en las colas.

Debido a esta violación del supuesto de normalidad, no se utilizó el ANOVA paramétrico como prueba principal. En su lugar, se aplicó la transformación de rangos alineados mediante ARTool. Esta prueba permite analizar los factores del diseño experimental y su interacción sin depender directamente del supuesto de normalidad de los residuales.

6.1. Modelo no paramétrico mediante ARTool

El modelo utilizado conserva la misma estructura general del diseño experimental: se consideran los factores `algorithm` y `noise_level`, su interacción, y el bloque `content_block`. En R, el modelo se ajustó de la siguiente manera:

```
response_percent ~ algorithm * noise_level + (1 | content_block).
```

De esta forma, el análisis permite evaluar las hipótesis planteadas en la sección 2.6: la hipótesis asociada al algoritmo de compresión, la hipótesis asociada al nivel de ruido y la hipótesis asociada a la interacción entre ambos factores.

6.2. Reporte del análisis ARTool

La tabla 6.1 presenta los resultados obtenidos mediante ARTool. Para cada efecto, se compara el valor p con el nivel de significancia $\alpha = 0.05$. A partir de esta comparación se decide si se rechaza o no la hipótesis nula correspondiente, según las hipótesis planteadas en la sección 2.6.

Tabla 6.1: Resultados del análisis mediante ARTool.

Efecto	F	gl	gl residuales	valor p
Algoritmo	142.3229	1	155	$< 2.22 \times 10^{-16}$
Nivel de ruido	30.7968	1	155	1.2124×10^{-7}
Algoritmo $\times$ nivel de ruido	2.6136	1	155	0.10799

Para el factor A, algoritmo de compresión, el valor p fue menor que 0.05. Por lo tanto, se rechaza H_{0A} y se acepta H_{1A} . Esto indica que el algoritmo de compresión sí tiene un efecto significativo sobre la variable de respuesta.

Para el factor B, nivel de ruido, el valor p también fue menor que 0.05. Por lo tanto, se rechaza H_{0B} y se acepta H_{1B} . Esto indica que el nivel de ruido sí tiene un efecto significativo sobre la variable de respuesta.

En cambio, para la interacción entre los factores A y B, el valor $p = 0.10799$ fue mayor que 0.05. Por lo tanto, no se rechaza H_{0AB} . Esto indica que no hay evidencia estadística suficiente para afirmar que exista interacción entre el algoritmo de compresión y el nivel de ruido.

La tabla 6.2 resume las decisiones tomadas para cada una de las hipótesis.

Tabla 6.2: Decisión sobre las hipótesis del diseño experimental.

Efecto	Hipótesis nula	Decisión
Algoritmo	H_{0A}	Se rechaza
Nivel de ruido	H_{0B}	Se rechaza
Algoritmo $\times$ nivel de ruido	H_{0AB}	No se rechaza

En resumen, los resultados muestran que el algoritmo de compresión y el nivel de ruido influyen significativamente en el porcentaje de incremento del tamaño comprimido. Sin embargo, no se encontró evidencia suficiente para afirmar que el efecto de un factor dependa del otro.

6.3. Comparaciones múltiples post hoc

Después del análisis mediante ARTool, se realizaron comparaciones múltiples con ajuste de Tukey. Estas comparaciones permiten identificar entre cuáles niveles o combinaciones de tratamientos se presentan diferencias estadísticamente significativas.

Primero se compararon las combinaciones entre algoritmo y nivel de ruido. Los resultados se muestran en la tabla 6.3.

Tabla 6.3: Comparaciones múltiples entre combinaciones de algoritmo y nivel de ruido con ajuste de Tukey.

Contraste	Estimación	EE	t	valor p
PNG, alto – PNG, bajo	24.6	7.16	3.431	0.0043
PNG, alto – WebP, alto	-62.6	7.16	-8.752	< 0.0001
PNG, alto – WebP, bajo	-33.9	7.16	-4.741	< 0.0001
PNG, bajo – WebP, alto	-87.2	7.16	-12.182	< 0.0001
PNG, bajo – WebP, bajo	-58.5	7.16	-8.172	< 0.0001
WebP, alto – WebP, bajo	28.7	7.16	4.011	0.0005

De acuerdo con la tabla 6.3, todos los contrastes entre combinaciones presentaron valores p menores que 0.05. Por lo tanto, se observan diferencias significativas entre las combinaciones comparadas. En los contrastes donde aparece WebP frente a PNG, las estimaciones negativas indican que WebP presentó valores mayores de la variable de respuesta. Además, al comparar los niveles de ruido dentro de un mismo algoritmo, el nivel alto presentó valores mayores que el nivel bajo.

También se realizó la comparación para el factor algoritmo. El resultado se muestra en la tabla 6.4.

Tabla 6.4: Comparación múltiple para el factor algoritmo con ajuste de Tukey.

Contraste	Estimación	EE	t	valor p
PNG – WebP	-63.2	5.30	-11.930	< 0.0001

En este caso, el contraste entre PNG y WebP fue significativo, ya que el valor p fue menor que 0.05. La estimación negativa para PNG – WebP indica que WebP obtuvo un porcentaje de incremento mayor que PNG. Este resultado es consistente con la decisión de rechazar H_{0A} .

Finalmente, se realizó la comparación para el factor nivel de ruido. El resultado se presenta en la tabla 6.5.

Tabla 6.5: Comparación múltiple para el factor nivel de ruido con ajuste de Tukey.

Contraste	Estimación	EE	t	valor p
Alto – bajo	37.0	6.67	5.549	< 0.0001

El contraste entre el nivel de ruido alto y el nivel bajo también fue significativo. Como la estimación es positiva, se concluye que el nivel de ruido alto produjo un porcentaje de incremento mayor que el nivel de ruido bajo. Este resultado es consistente con la decisión de rechazar H_{0B} .

En resumen, las comparaciones post hoc respaldan lo observado en el análisis ARTool: existen diferencias significativas asociadas al algoritmo de compresión y al nivel de ruido. Aunque las combinaciones de tratamientos muestran diferencias entre sí, la prueba global no encontró evidencia suficiente para rechazar H_{0AB} .

6.4. Tamaño del efecto

Además de revisar la significancia estadística de cada efecto, se calculó el tamaño del efecto mediante el cuadrado parcial (η_p^2). Esta medida permite estimar qué tan grande es el efecto de cada factor o interacción dentro del análisis realizado.

Para calcularlo se utilizó la siguiente fórmula:

$$\eta_p^2 = \frac{F \cdot gl_{\text{efecto}}}{F \cdot gl_{\text{efecto}} + gl_{\text{residual}}}. \quad (6.1)$$

La escala utilizada para interpretar el valor de η_p^2 fue la siguiente:

- alrededor de 0.01: efecto pequeño,
- alrededor de 0.06: efecto mediano,
- alrededor de 0.14 o más: efecto grande.

También se calculó f^2 de Cohen, ya que este valor se utiliza posteriormente para el análisis de potencia:

$$f^2 = \frac{\eta_p^2}{1 - \eta_p^2}.$$

Los resultados se presentan en la tabla 6.6.

Tabla 6.6: Tamaño del efecto para los factores y la interacción.

Efecto	η_p^2	f^2 de Cohen	Interpretación
Algoritmo	0.4787	0.9182	Grande
Nivel de ruido	0.1658	0.1987	Grande
Algoritmo $\times$ nivel de ruido	0.0166	0.0169	Pequeño

De acuerdo con la escala anterior, el algoritmo presentó un tamaño del efecto grande, con $\eta_p^2 = 0.4787$. Esto concuerda con el rechazo de H_{0A} y confirma que el algoritmo tuvo un efecto importante sobre la variable de respuesta.

El nivel de ruido también presentó un tamaño del efecto grande, con $\eta_p^2 = 0.1658$, aunque menor que el del algoritmo. Esto concuerda con el rechazo de H_{0B} .

Por último, la interacción entre algoritmo y nivel de ruido presentó un tamaño del efecto pequeño, con $\eta_p^2 = 0.0166$. Este resultado es consistente con la decisión de no rechazar H_{0AB} .

6.5. Potencia de la prueba

La potencia de la prueba se calculó para cada efecto utilizando f^2 de Cohen, el nivel de significancia $\alpha = 0.05$ y los grados de libertad correspondientes. Además, se estimó la cantidad aproximada de observaciones necesarias para alcanzar una potencia de 0.80.

Como el diseño experimental tiene cuatro combinaciones principales entre algoritmo y nivel de ruido, los tamaños de muestra requeridos se redondearon hacia arriba para mantener un diseño balanceado. La tabla 6.7 muestra la potencia observada, la cantidad aproximada de observaciones totales necesarias y la cantidad aproximada de observaciones por combinación de algoritmo y nivel de ruido.

Tabla 6.7: Potencia observada y tamaño de muestra aproximado para alcanzar una potencia de 0.80.

Efecto	Potencia observada	Obs. totales	Obs. por combinación
Algoritmo	1.0000	16	4
Nivel de ruido	0.9998	48	12
Algoritmo $\times$ nivel de ruido	0.3660	472	118

La potencia observada para el efecto del algoritmo fue prácticamente igual a 1, mientras que para el nivel de ruido fue 0.9998. Esto indica que el experimento tuvo una potencia muy alta para detectar los efectos principales observados. Esto es consistente con el rechazo de H_{0A} y H_{0B} .

En cambio, la potencia observada para la interacción fue de 0.3660, la cual se encuentra por debajo del valor usual de referencia de 0.80. Esto indica que el experimento tuvo una capacidad limitada para detectar una interacción del tamaño observado.

Para alcanzar una potencia de 0.80 con una interacción de esta magnitud, se necesitarían aproximadamente 472 observaciones totales, equivalentes a unas 118 observaciones por combinación de algoritmo y nivel de ruido. Como el experimento actual cuenta con 160 observaciones totales, equivalentes a 20 observaciones por combinación, se concluye que el tamaño de muestra fue suficiente para detectar los efectos principales, pero no necesariamente para evaluar la interacción entre los factores. Por esta razón, la decisión de no rechazar H_{0AB} debe interpretarse junto con la potencia baja observada para la interacción.

Conclusiones

En este trabajo se estudió el efecto del algoritmo de compresión y del nivel de ruido gaussiano sobre la Tasa de Crecimiento Porcentual (TCP) del tamaño de archivos comprimidos sin pérdida. Para esto se utilizó un diseño factorial completo con dos factores: algoritmo de compresión, con niveles PNG y WebP lossless, y nivel de ruido, con niveles bajo y alto. Además, se consideró el tipo de contenido de la imagen como bloque, con el fin de controlar parte de la variabilidad asociada a las imágenes utilizadas.

Los resultados obtenidos muestran que el algoritmo de compresión sí tuvo un efecto importante sobre la variable de respuesta. En el análisis exploratorio, WebP lossless presentó una TCP media de 151.21 pp, mientras que PNG presentó una media de 80.38 pp. Esta diferencia también fue confirmada mediante el análisis ARTool, donde el efecto del algoritmo resultó estadísticamente significativo, con un valor $p < 0.05$. Por lo tanto, se concluye que el algoritmo utilizado influye en el incremento porcentual del tamaño comprimido cuando se introduce ruido en la imagen.

También se encontró evidencia de que el nivel de ruido afecta la TCP. El nivel alto de ruido presentó una media de 132.97 pp, mientras que el nivel bajo presentó una media de 98.62 pp. Esta diferencia fue significativa en el análisis ARTool, por lo que se concluye que un mayor nivel de ruido tiende a producir un mayor incremento porcentual en el tamaño del archivo comprimido. Este resultado concuerda con lo esperado, ya que el ruido introduce variaciones en los píxeles y reduce la regularidad que aprovechan los algoritmos de compresión sin pérdida.

En cuanto a la interacción entre algoritmo y nivel de ruido, los gráficos de interacción mostraron líneas bastante paralelas. Esto sugería, desde el análisis exploratorio, que no había una interacción marcada entre ambos factores. El análisis ARTool confirmó esta apreciación, ya que la interacción no resultó estadísticamente significativa al nivel de significancia de 0.05. Por esta razón, no se encontró evidencia suficiente para afirmar que el efecto del ruido dependa del algoritmo utilizado, o que el efecto del algoritmo dependa del nivel de ruido.

Las comparaciones múltiples permitieron detallar las diferencias entre niveles y combinaciones de tratamientos. En general, WebP lossless presentó valores mayores de TCP que PNG, y el nivel de ruido alto presentó valores mayores que el nivel bajo. Estos resultados respaldan las conclusiones obtenidas en el análisis principal: los efectos más relevantes corresponden al algoritmo de compresión y al nivel de ruido, no a la interacción entre ambos.

La validación de supuestos mostró que no se cumplió el supuesto de normalidad de los residuales. Por esta razón, se utilizó ARTool como alternativa para realizar el análisis de los factores y su interacción. Aunque se había ajustado un ANOVA de forma preliminar, las conclusiones finales se basaron en ARTool, ya que esta prueba fue más adecuada según los supuestos observados.

El cálculo del tamaño del efecto también apoyó estas conclusiones. El algoritmo presentó un tamaño del efecto grande, con $\eta_p^2 = 0.4787$, y el nivel de ruido también presentó un tamaño del efecto grande, con $\eta_p^2 = 0.1658$. En cambio, la interacción presentó un tamaño del efecto pequeño, con $\eta_p^2 = 0.0166$. Esto indica que los factores principales tuvieron una influencia clara sobre la TCP, mientras que la interacción tuvo un efecto bajo.

Finalmente, el análisis de potencia mostró que el experimento tuvo una potencia muy alta para detectar los efectos principales del algoritmo y del nivel de ruido. Sin embargo, la potencia para la interacción fue baja. Para detectar una interacción pequeña como la observada se requeriría una cantidad mayor de observaciones. Por lo tanto, el diseño utilizado fue suficiente para responder las preguntas principales sobre los efectos del algoritmo y del ruido, pero no necesariamente para detectar interacciones pequeñas entre ambos factores.

En síntesis, el experimento permite concluir que tanto el algoritmo de compresión como el nivel de ruido gaussiano influyen significativamente en el crecimiento porcentual del tamaño comprimido. WebP lossless mostró incrementos mayores que PNG, y el nivel alto de ruido produjo incrementos mayores que el nivel bajo. No se encontró evidencia significativa de interacción entre los factores, por lo que las conclusiones principales del estudio se centran en los efectos individuales del algoritmo y del nivel de ruido.

Bibliografía

- [1] Osama K. Al-Shaykh and Russell M. Mersereau. Lossy compression of noisy images. *IEEE Transactions on Image Processing*, 7(12):1641–1652, 1998.
- [2] Cloudflare, Inc. What is the WebP image format?, 2024.
- [3] Duc-Tien Dang-Nguyen, Cecilia Pasquini, Valentina Conotter, and Giulia Boato. Raise: A raw images dataset for digital image forensics. In *Proceedings of the 6th ACM Multimedia Systems Conference*, 2015.
- [4] Mandeep Kaur and Pooja Choudhary. Lossless and lossy image compression techniques: A review. *International Journal of Computer Applications*, 142(6):1–5, 2016.
- [5] Vladimir Lukin, Sergey Krivenko, Oleksii Rubel, and Karen Egiazarian. Analysis of lossless compression of noisy images. *Journal of Imaging*, 10(1):14, 2024.
- [6] William D. Pence, R. Seaman, and R. L. White. Lossless astronomical image compression and the effects of noise. *Publications of the Astronomical Society of the Pacific*, 121(878):414–427, 2009.
- [7] Pillow Contributors. Pillow (PIL fork) documentation, 2024. Disponible en: <https://pillow.readthedocs.io/>.
- [8] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023.
- [9] Md. Mahmudur Rahman and Mohammed Hamada. Lossless image compression techniques: A state-of-the-art survey. *Symmetry*, 11(10):1274, 2019.
- [10] Sanaullah Rajvor, Sumair Aziz, Muhammad Salman Fareed, and Malak Ahmed Ansari. A comprehensive study of image degradation and restoration techniques. *Multimedia Tools and Applications*, 81:7807–7850, 2022.
- [11] S. R. Subramanya. Image compression technique. *IEEE Potentials*, 20(1):19–23, 2001.
- [12] Anil Kumar Verma and Imran Ali. A comparative study of various types of image noise and efficient noise removal techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(10):617–622, 2013.
- [13] Aleksey Zemliachenko, Vladimir Lukin, Sergey Krivenko, Jaakko Astola, and Karen Egiazarian. Still image/video frame lossy compression providing a desired visual quality. *EURASIP Journal on Image and Video Processing*, 2016(1):1–14, 2016.