

Universidad de Murcia

Curso 2020/21

Grado de Ingeniería Informática, 4º curso

Facultad de Informática

**Compresión Multimedia**

Juan José Morell Fernández

juanjose.morellf@gmail.com

24418612H

Índice

[1. Abstract 3](#_Toc60655516)

[2. Introducción 4](#_Toc60655517)

[3. Metodología 5](#_Toc60655518)

[3.1 Procedimiento de compresión y descompresión 5](#_Toc60655519)

[3.2 Tipos de compresión y Factor de calidad 10](#_Toc60655520)

[3.3 Imágenes y Factores de calidad seleccionados 10](#_Toc60655521)

[3.4 Manual de usuario 13](#_Toc60655522)

[4. Resultados Experimentales 14](#_Toc60655523)

[4.1 Imagen 1 14](#_Toc60655524)

[4.2 Imagen 2 14](#_Toc60655525)

[4.3 Imagen 3 14](#_Toc60655526)

[4.4 Imagen 4 14](#_Toc60655527)

[4.5 Imagen 5 14](#_Toc60655528)

[4.6 Imagen 6 14](#_Toc60655529)

[5. Conclusiones 15](#_Toc60655530)

[6. Bibliografía 16](#_Toc60655531)

# 1. Abstract

# 2. Introducción

Una imagen normalmente está representada por una matriz, donde cada uno de sus valores corresponde a cada pixel y simboliza la intensidad de estos. Si el valor es alto significa que ese pixel tiene un nivel superior de brillo, mientras que si es un valor pequeño el nivel de brillo también lo será. Asimismo, las imágenes con color presentan diferentes canales por cada uno de los componentes de color existente. Los componentes más típicos son el rojo, verde y azul; formando la representación más conocida de las imágenes, RGB.



Figura 1 – Espacio de color RGB

Aunque esta representación de las imágenes es la más común, no es como se almacenan en un disco duro. Por ejemplo, almacenar una imagen con color de 48 megapíxeles conlleva guardar un total de 48 millones de valores por cada uno de los canales, sumando un total de 144 millones de valores. Esto dependiendo del método de almacenamiento puede llegar a ocupar un total de 144 MB. Pero en la realidad, una imagen de esas características llega a ocupar menos de 14 MB. ¿Cómo se consigue lograr esto?

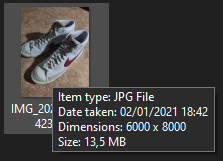


Figura 2 – Imagen de ejemplo, tomada con una cámara de 48 MP.

Justamente esta es la pregunta que hay que conseguir responder en este proyecto. Cuál es el proceso que lleva a una imagen, que supuestamente tendría que ocupar una gran cantidad de megabytes, pueda ocupar tan poco en disco.

La respuesta a esta pregunta es la compresión de imágenes, en este caso JPEG. Esta es una de las extensiones más conocidas mundialmente en las imágenes, y es un formato de compresión de imágenes.

# 3. Metodología

## 3.1 Procedimiento de compresión y descompresión

JPEG es un método de compresión con pérdida, lo que significa que la imagen final que se almacena en nuestro ordenador no tiene toda la información original. Cuanta más información se descarte en el proceso de compresión peor calidad tendrá la imagen comprimida. Este trato entre la calidad y la ocupación de la imagen en disco será uno de los puntos importantes dentro de este proyecto, y será comentado junto a las imágenes que se mostraran en el apartado 3.2.

En general, JPEG consigue una buena relación tamaño/calidad, ya que los datos que va a despreciar van a ser aquellos que no son realmente visibles por la vista humana.

El primer paso de la compresión es la conversión del espacio de color. Para no almacenar la imagen con un tamaño de 114 MB, se cambia el espacio de color RGB, a uno llamado YCbCr. Este se caracteriza por tener un canal que representa la intensidad de luz, y otros dos representan los colores, el azul y el rojo. Esto permite la separación entre la luminancia (Y) y la crominancia (Cb, Cr).

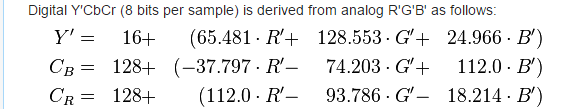


Figura 3 – Transformación de espacio RGB a YCbCr.

La transformación del espacio RGB a YCbCr se trata de una transformación lineal, que puede ser expresada mediante una multiplicación de matrices.

Las imágenes actuales son mucho más sensibles a los cambios de brillo que al de colores, por lo que se reducen los canales de color. Gracias al espacio YCbCr, se puede hacer esto sin problemas mediante el proceso de *downsample*, o también conocida como *chroma subsampling*.

Además de este cambio, otro factor importante a la hora de reducir el tamaño de las imágenes son las características de la vista humana. La idea reside en que los objetos o detalles pequeños que están representado en la imagen son más propensos a pasar desapercibidos que objetos grandes. Este efecto se denomina *sensibilidad de contraste dependiente de la frecuencia.*

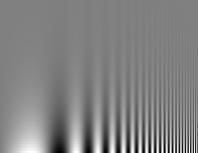


Figura 4 – Efecto de la frecuencia espacial

Un buen ejemplo de este fenómeno sería una imagen con líneas negras donde cada vez se van juntando más hasta que parece que no hay distancia, de izquierda a derecha. Mientras que el contraste va incrementando desde abajo hacia arriba de la imagen. Con esto se puede apreciar una curva en donde se desvanecen las líneas, y se debe a que nuestro ojo es más sensible a las variaciones donde el brillo sea mayor.

Se puede apreciar en la parte superior de la figura 4 que las líneas verticales casi no son visibles. Esto se debe a que se trata de la parte que presenta menor contraste y mayor frecuencia. Estos fenómenos nos dan espacio para poder comprimir más, ya que como se puede observar no detectamos tanta información.

JPEG lleva a cabo esto mediante la división de la imagen en bloques de 8 x 8 y cuantizandolos posteriormente, de forma individual, en una representación del dominio de frecuencias. Cada uno de estos bloques es comparado con un conjunto de 64 patrones de frecuencias donde la frecuencia espacial aumenta desde la esquina superior izquierda hasta la esquina inferior derecha.

Este proceso descompone la imagen en componentes de frecuencia, convirtiendo cada bloque de 8x8, que representaba el nivel de brillo de la imagen, en otro bloque de 8 x 8 que representa la presencia de un componente de frecuencia en particular. Este método se conoce como *discrete cosine transform* o *DCT*.

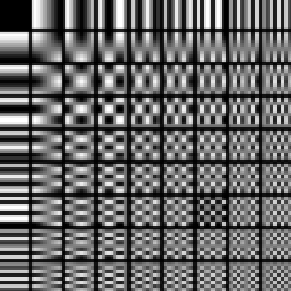


Figura 5 – Conjunto de 64 patrones de frecuencias DCT.

Mediante esta representación se pueden comprimir las frecuencias que son menos visibles al ojo humano, dividiendo la matriz del bloque 8 x 8 entre una matriz constante. Estas frecuencias que son menos visibles son divididas por constantes más grandes, al contrario que en las frecuencias más visibles que son divididas por constantes menores.

En nuestro caso, se utilizan las matrices de cuantización de JPEG por defecto, tanto para luminancia como para crominancia. También son conocidas por el nombre de ITU T.81.

La cuantización en este contexto se trata de redondear el resultado de la división de las matrices a los enteros más cercanos. Mediante el uso de números grandes el resultado de su división se acercará más a cero, lo que resulta en un ratio de compresión mayor, pero reduce también la calidad de la imagen.

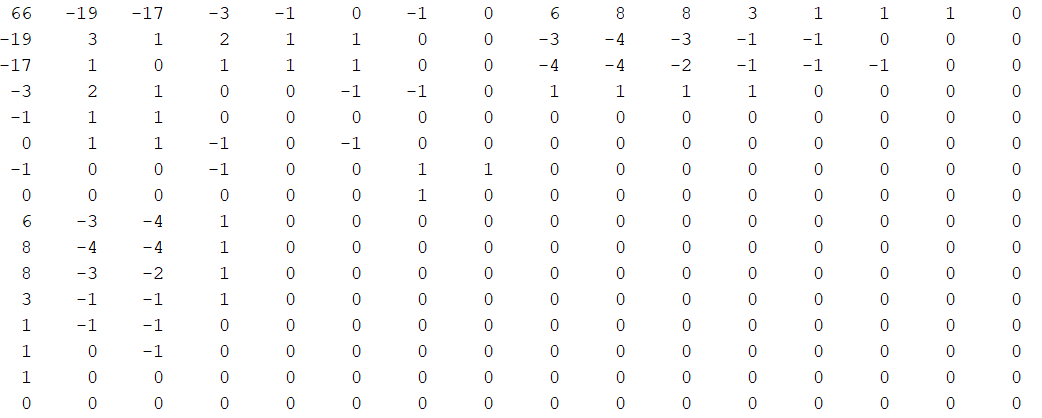


Figura 6 – Matriz tras la cuantización

Tras la cuantización, obtenemos unas matrices donde existen una gran cantidad de ceros en las frecuencias altas. Pero como se puede apreciar en la figura 6, los ceros están intercalados con diferentes números. Para ordenar el contenido de la matriz y poder almacenarlo de una forma mas eficiente se reordenan los coeficientes mediante un recorrido en zigzag desde la esquina superior izquierda a la esquina inferior derecha.

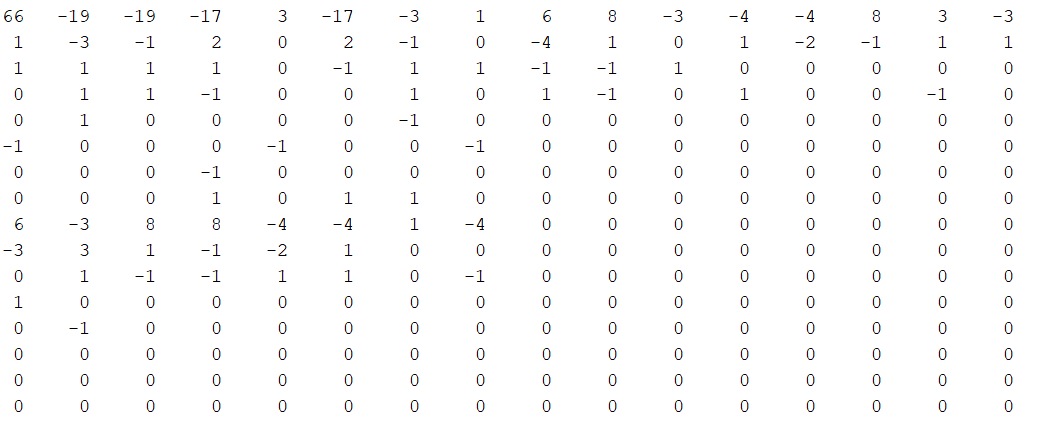


Figura 7 – Matriz tras recorrido en zigzag.

Una vez juntados los ceros, en vez de almacenar la matriz así se puede almacenar su valor junto al número de veces que aparece consecutivamente. Un ejemplo es que, si aparecen cinco ceros seguidos, en vez de almacenar los cinco, se almacena un 5 junto a un 0, representando que existen cinco ceros consecutivos. A esta técnica se le conoce como *Run-Length encoding*.

Otra forma de comprimir esto es teniendo en cuenta la frecuencia de aparición de cada elemento en la matriz, almacenando con más bits aquellos valores con más presencia y menor cantidad al resto. Esto consigue reducir el número de bits por símbolo. Esta técnica es conocida por *Entropy coding* o *Huffman coding*.

Ambas técnicas se recogen en el grupo de codificación sin perdidas, ya que no se pierde nada de la información original a la hora de comprimir. Se centran en almacenar los datos de una manera eficiente aprovechando la redundancia en los datos. Asimismo, estas técnicas también son utilizadas en otros ámbitos, como texto, audio o video.

[EXPLICAR ENTROPIA]

Todo este ejemplo se ha realizado mediante el programa creado tras el desarrollo de la práctica y una imagen en formato *bmp* de tamaño 10 x 10.

Por tanto, el proceso de compresión de una imagen toma el siguiente camino:

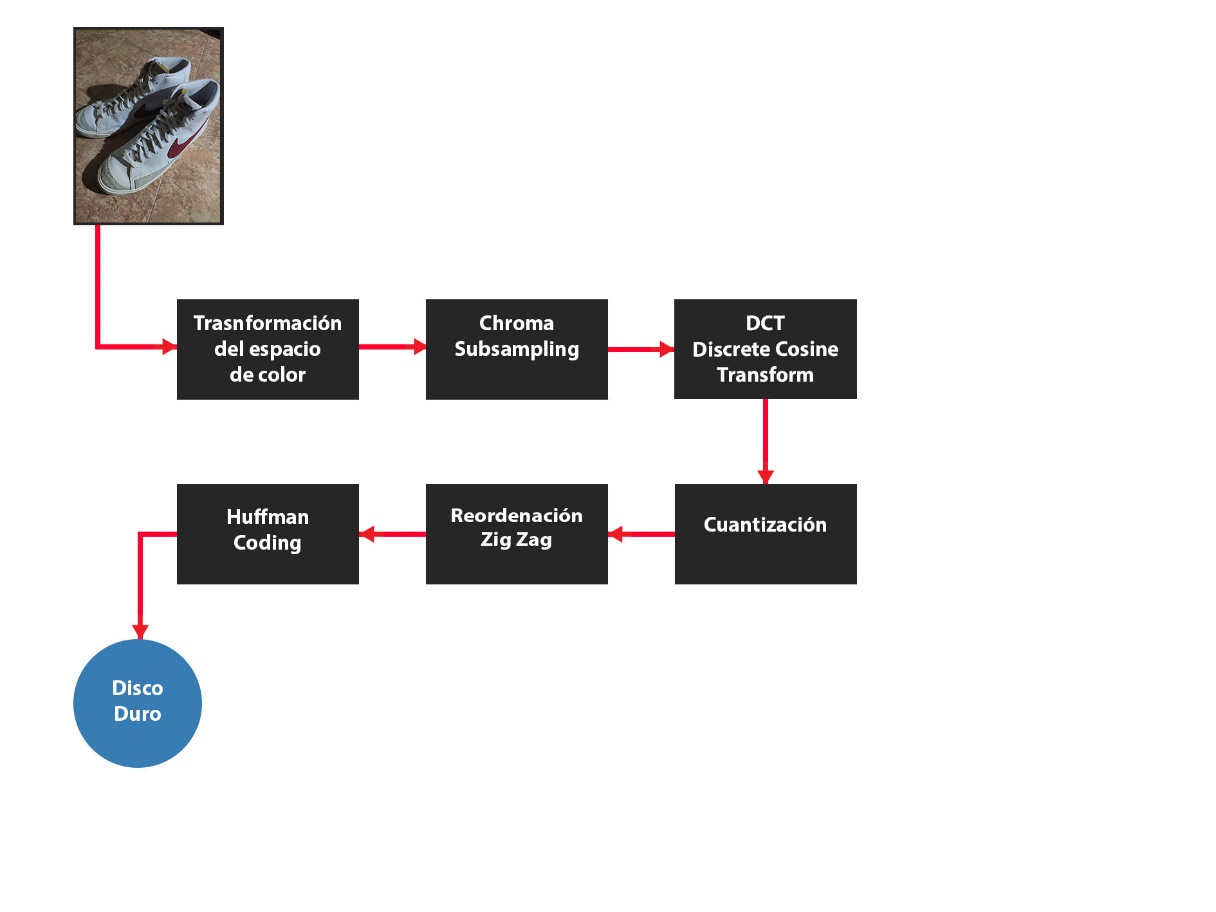


Figura – Esquema de compresión de una imagen

Para la decodificación de la misma imagen comprimida sería el proceso inverso. Debido a que parte de la información original se ha perdido tras los procesos de subsampling o cuantización, la imagen decodificada no va a ser idéntica a la original. Aun así, si la comprensión de la imagen se ha realizado con un correcto factor de calidad no se debería apreciar ninguna diferencia. Cuanto más alto sea dicho factor más notorios serán los efectos de la compresión.



Figura – Comparación entre diferentes compresiones

En esta figura no se podrá apreciar del todo bien el cambio entre imágenes, ya que dicha imagen esta también comprimida en formato jpeg. Si se desean ver las imágenes a su calidad total, se facilitarán las copias originales junto a esta documentación.

## 3.2 Tipos de compresión y Factor de calidad

Dentro del proyecto se desarrollando dos tipos de compresión / descompresión que se nombraran como: default y custom. Las diferencias entre estos son la forma de aplicar *Huffman encoding* durante el proceso de compresión.

El primero de ellos, default, realiza un Huffman encoding guiado por un valor constante, sin tener en cuanta los valores que va a codificar. Mientras que, Huffman encoding custom, trata de mejorar esto teniendo en cuenta la frecuencia de aparición de los datos que se van a codificar.

Por tanto, teóricamente el segundo de los compresores va a ser más eficiente. Esto se tratará en el apartado 4, junto a los datos reales de las pruebas.

Otro aspecto importante es el factor de calidad, parámetro principal de ambas compresiones implementadas, ya que especificar cuanto se va a comprimir una imagen o cuanta calidad se va a perder. Durante el proceso de experimentación se realizarán pruebas con varios valores para este parámetro para conocer cual es el mas eficiente dentro del ratio compresión / calidad.

## 3.3 Imágenes y Factores de calidad seleccionados



Figura -

Imagen 1: Esta imagen presenta a un pájaro junto aun fondo de color verde, y puede presentar un comportamiento interesante a la hora de comprimir un color solido de fondo y un objeto que resalta bastante. Otro fenómeno que tratar sería el cambio brusco entre los colores del pájaro y fondo.



Figura -

Imagen 2: Esta segunda imagen presenta un escenario parecido a la anterior, pero con la particularidad de que el fondo no es homogéneo, sino todo lo contrario. El objetivo con esta imagen es estudiar el comportamiento del proceso de compresión en una imagen donde los elementos tienen un color parecido, pero con gran cantidad de detalles.



Figura -

Imagen 3: La característica principal de esta imagen es la simetría y estudiar cómo se comporta dentro del proceso de compresión, si se trata de algún factor clave a la hora de eliminar redundancia o no.



Figura -

Imagen 4: Al igual que la anterior, se intenta seguir estudiando el fenómeno de la simetría. Pero en este caso, al existir cierta semejanza con los patrones de frecuencia presentes en DCT, el objetivo es averiguar que tan bien funciona la compresión con patrón de franjas y diferentes colores.



Figura -

Imagen 5: El objetivo con esta imagen es estudiar el comportamiento del compresor cuando existe un color dominante, pero con detalles.



Figura -

Imagen 6: En esta imagen tiene como objetivo estudiar el comportamiento de la compresión de una imagen que contenga texto, y como trata los contornos de este. Además, presenta vegetación en la parte inferior que añade ruido a la imagen.

## 3.4 Manual de usuario

Para el desarrollo del proyecto se ha realizado uso de un script propio, que permite realizar la comprensión y descompresión, tanto para la versión *default* como *custom*, de una imagen para distintos factores de calidad.

Con esto se facilita el proceso de obtención de las variables MSE y RC, y las graficas que se van a estudiar en el apartado 4.

# 4. Resultados Experimentales

En este apartado se va a realizar el estudio de los resultados obtenidos después de aplicar la compresión y descompresión sobre las imágenes. En este se reflejarán los resultados producidos por los dos tipos de compresión: default y custom, aplicados juntamente con una serie de factores de calidad.

Dichos factores de calidad han sido seleccionados de forma que se pueda estudiar con facilidad los cambios de las relaciones de compresión y el error cuadrático medio. Teniendo en cuenta que un factor de calidad igual a 1 es la mínima compresión posible y 1000 cercano al 100% de compresión se han seleccionado los siguientes factores de calidad para realizar el estudio:

[1, 25, 50, 100, 250, 500, 1000]

Los parámetros que se van a tener en cuenta para el estudio, como se ha comentado anteriormente, son: la relación de compresión y el error cuadrático medio.

La relación de compresión es una medida que representa cuanto es capaz de comprimir un algoritmo de compresión. La definición que se usa en este proyecto es la siguiente:

Donde TO es el tamaño original de la imagen y TC es el tamaño del archivo comprimido. Esta medida se expresa en porcentaje (%).

Asimismo, el error cuadrático medio, MSE, mide la cantidad de error que existe entre dos conjuntos de datos, en este caso imágenes. Para obtener esto se comparan los pixeles de la imagen original con los de la imagen comprimida.

[Formula MSE]

En cada apartado se mostrará una gráfica con la relación entre relación de compresión y MSE, además se añadirá una tabla donde se reflejen los valores de cada uno de ellos en cada factor de calidad. Y, por último, la evolución de las imágenes y la diferencia entre la imagen original y la comprimida con factor de calidad de 1000.

## 4.1 Imagen 1

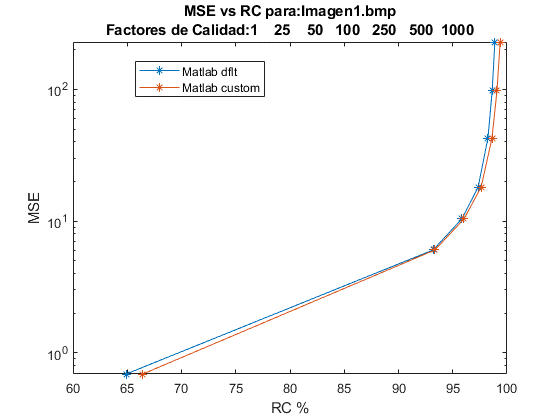
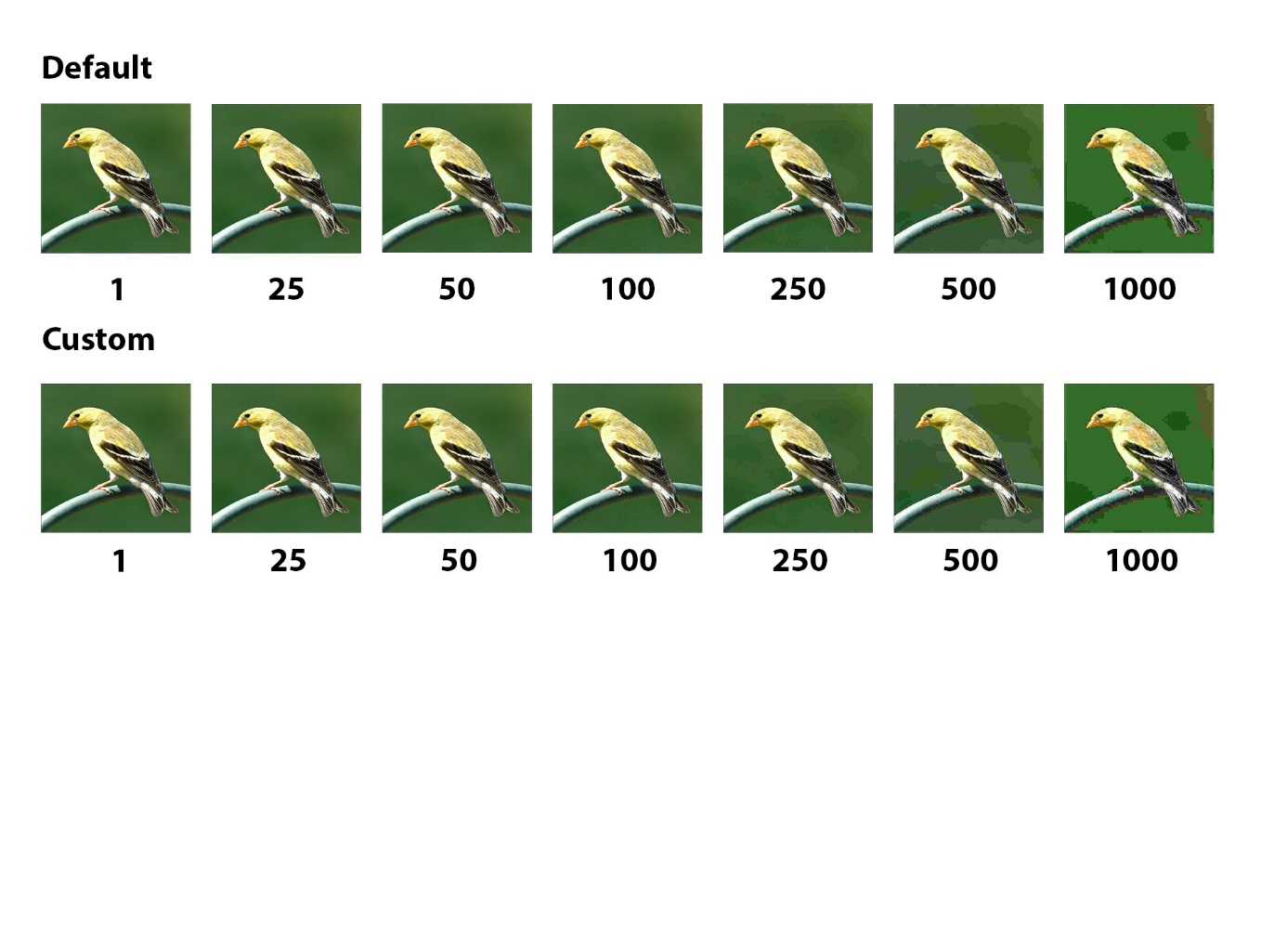


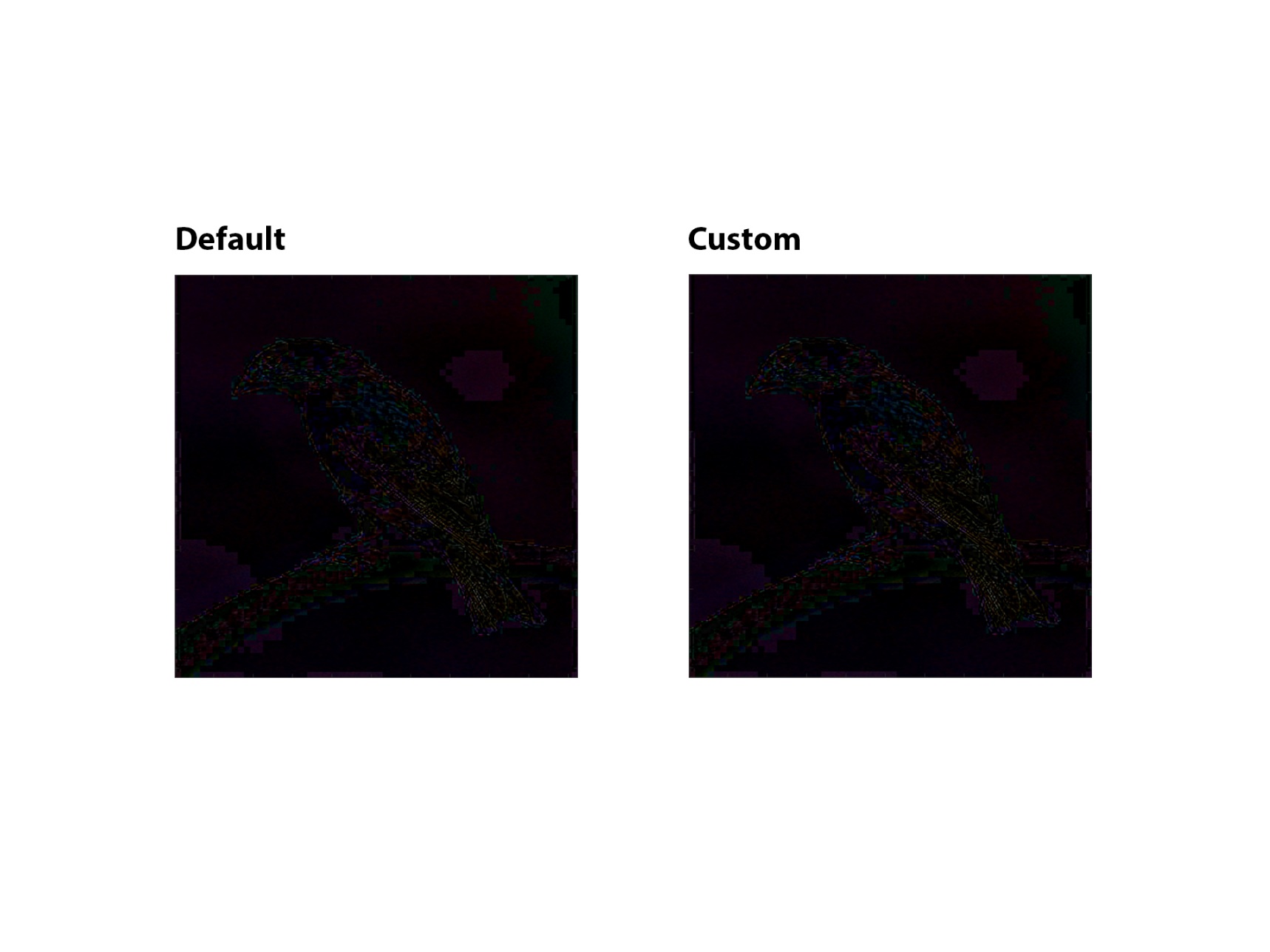
Figura – Gráfica Imagen 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| FC | MSE | DEFAULT | MSE | CUSTOM | RC | DEFAULT | RC | CUSTOM |
| **1** | 0.6902 | 0.6902 | 64.9002 | 66.4337 |
| **25** | 6.0353 | 6.0353 | 93.1874 | 93.3494 |
| **50** | 10.3787 | 10.3787 | 95.8372 | 96.0345 |
| **100** | 18.116 | 18.116 | 97.3429 | 97.6133 |
| **250** | 42.6655 | 42.665 | 98.2421 | 98.6184 |
| **500** | 97.7043 | 97.7043 | 98.6356 | 99.0764 |
| **1000** | 229.0251 | 229.0251 | 98.8801 | 99.3711 |

[Explicación de datos]



[Explicación diferencias junto a los números]



[Explicación diferencia y conclusión]

## 4.2 Imagen 2

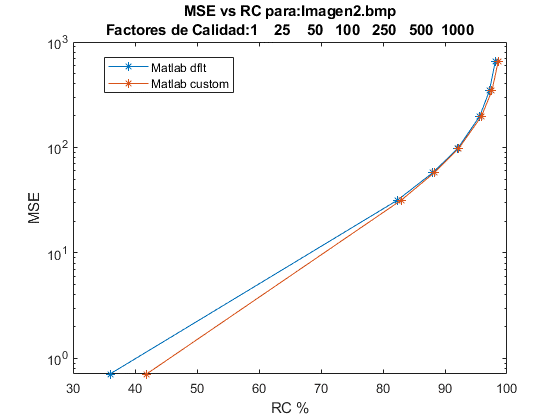
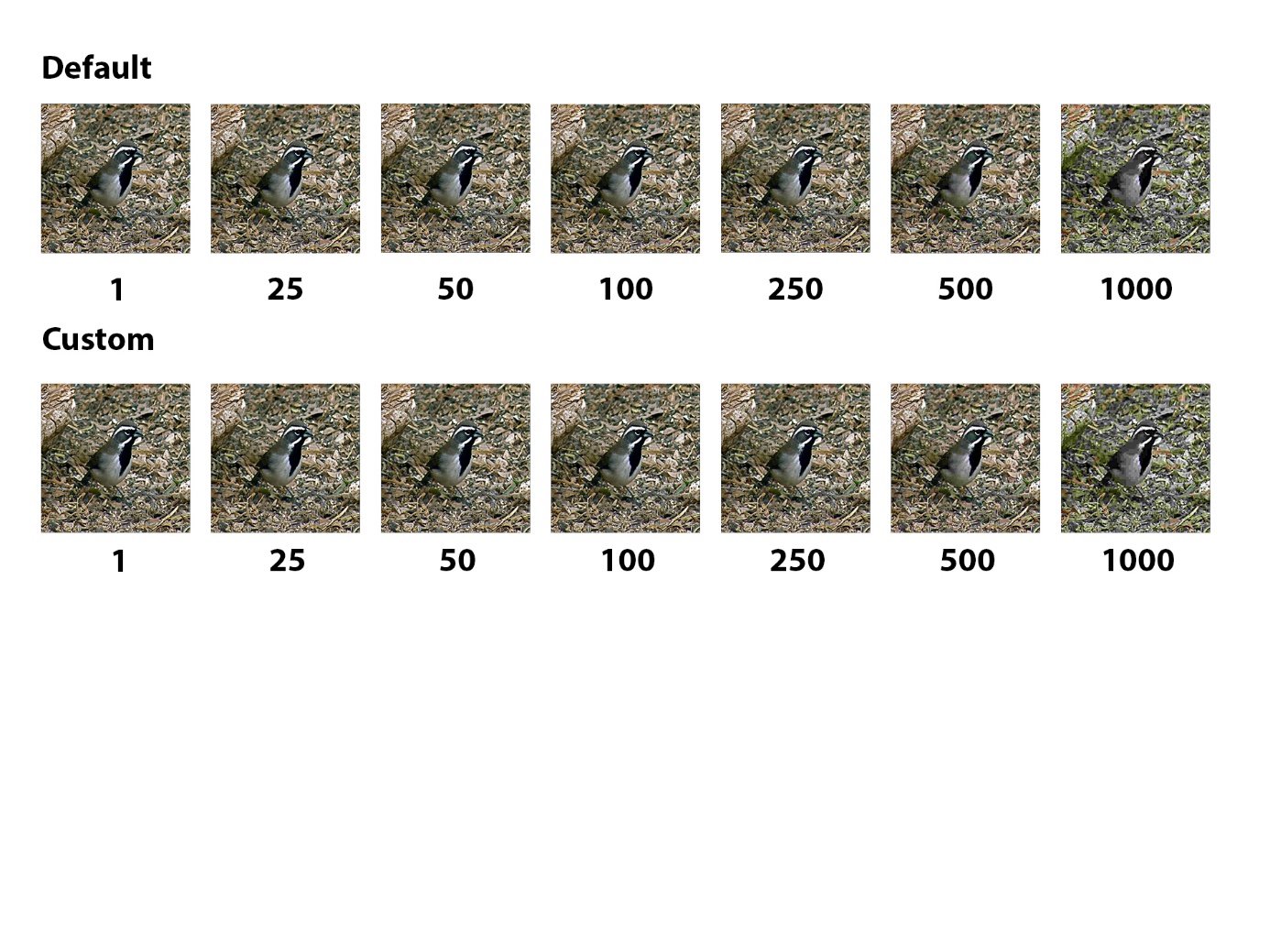


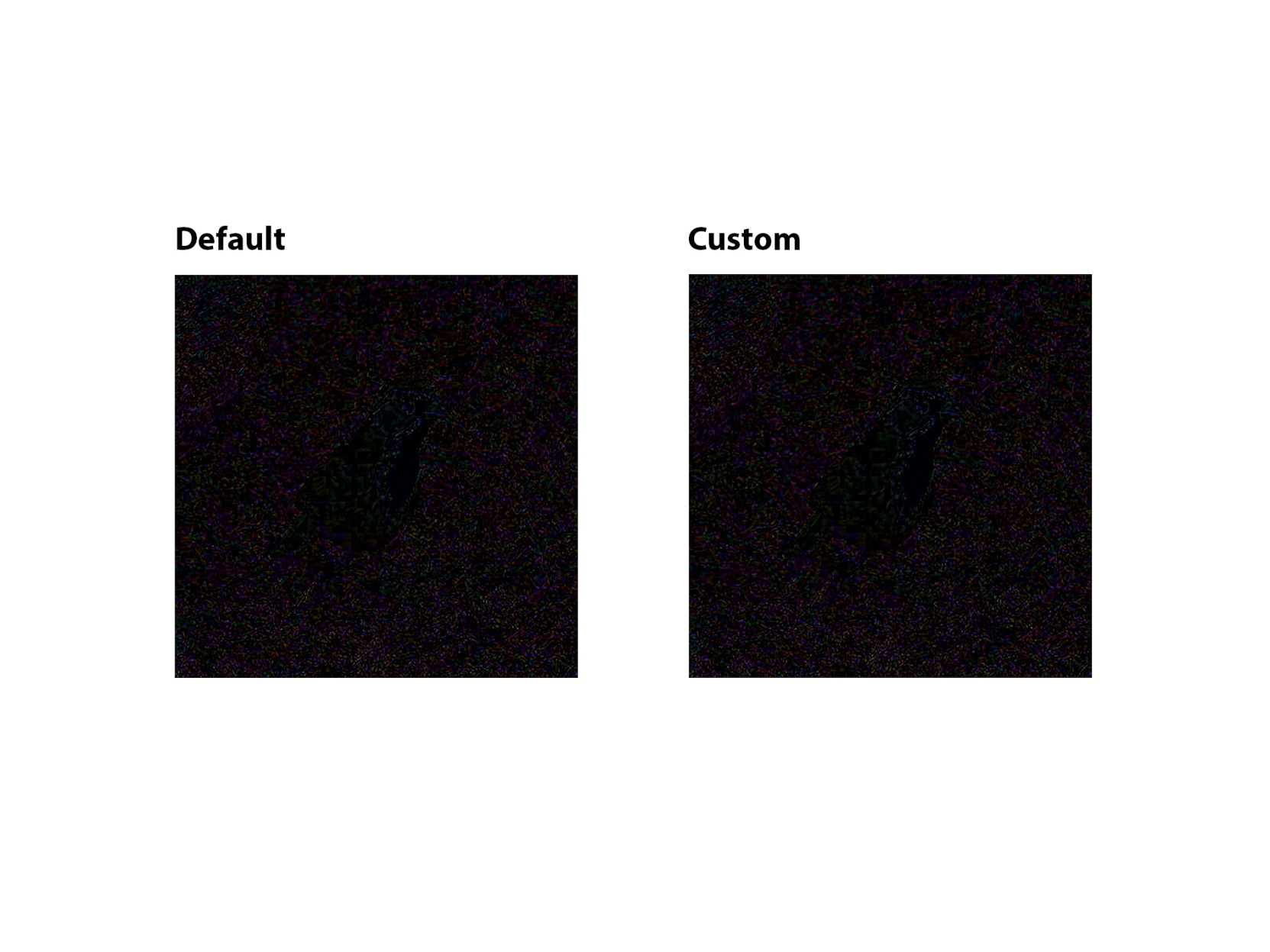
Figura – Gráfica Imagen 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| FC | MSE | DEFAULT | MSE | CUSTOM | RC | DEFAULT | RC | CUSTOM |
| 1 | 0.7111 | 0.7111 | 35.9714 | 41.8698 |
| 25 | 31.7015 | 31.7015 | 82.3525 | 82.9575 |
| 50 | 57.7807 | 57.7807 | 88.0428 | 88.2856 |
| 100 | 97.6127 | 97.6127 | 92.0297 | 92.1962 |
| 250 | 198.3503 | 198.3503 | 95.6089 | 95.8995 |
| 500 | 346.0193 | 346.0193 | 97.1755 | 97.5721 |
| 1000 | 655.3356 | 655.3356 | 98.1176 | 98.5897 |

[Explicación de datos]



[Explicación diferencias junto a los números]



[Explicación diferencia y conclusión]

## 4.3 Imagen 3

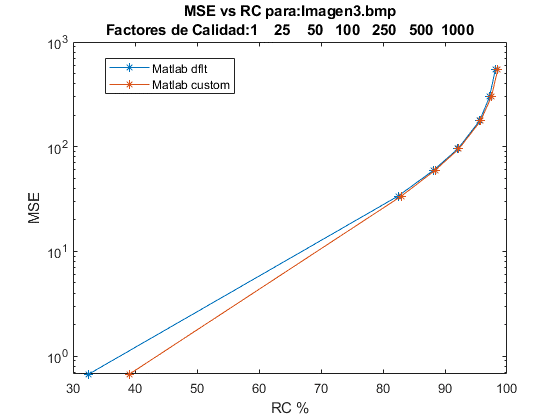
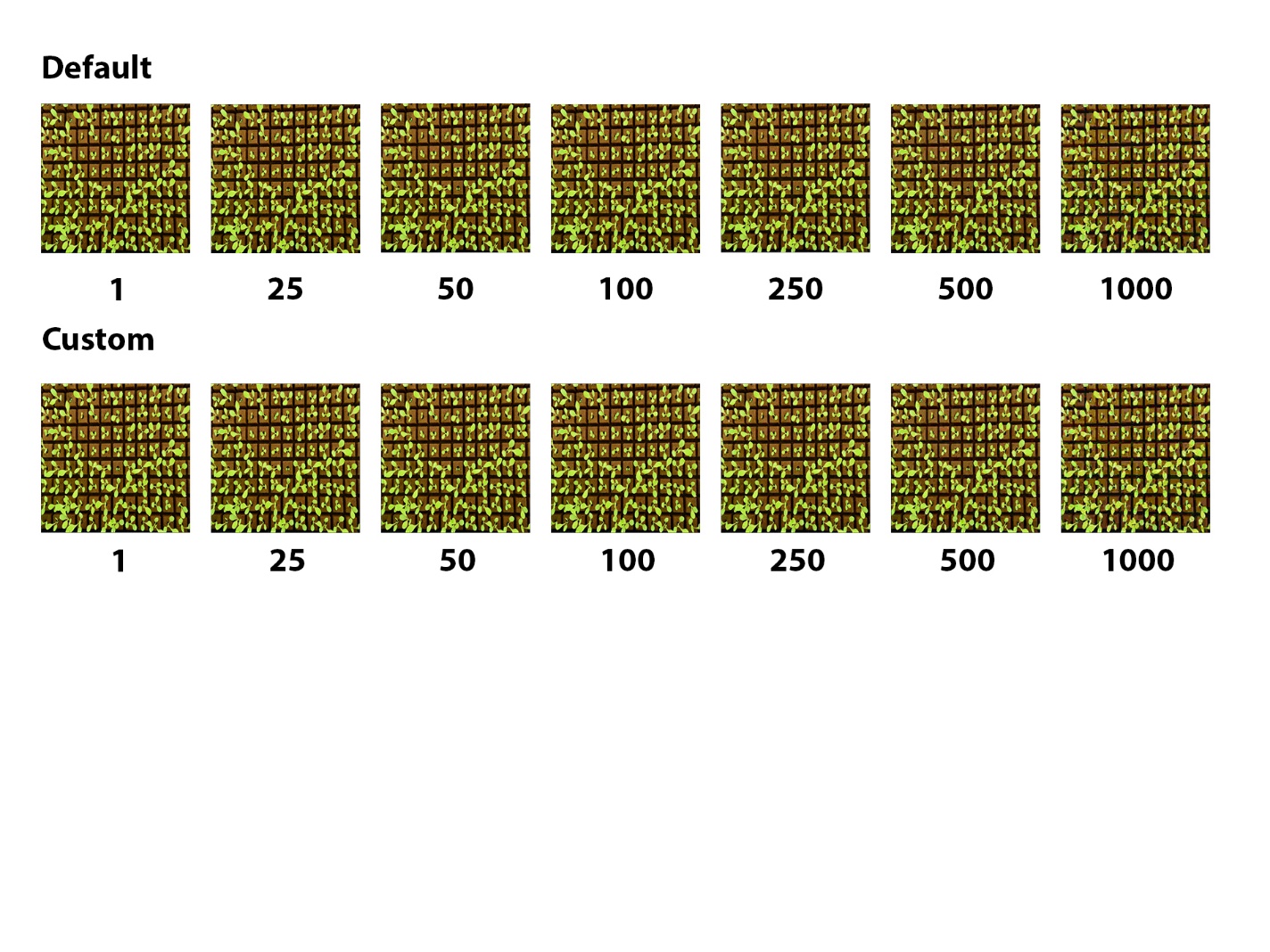


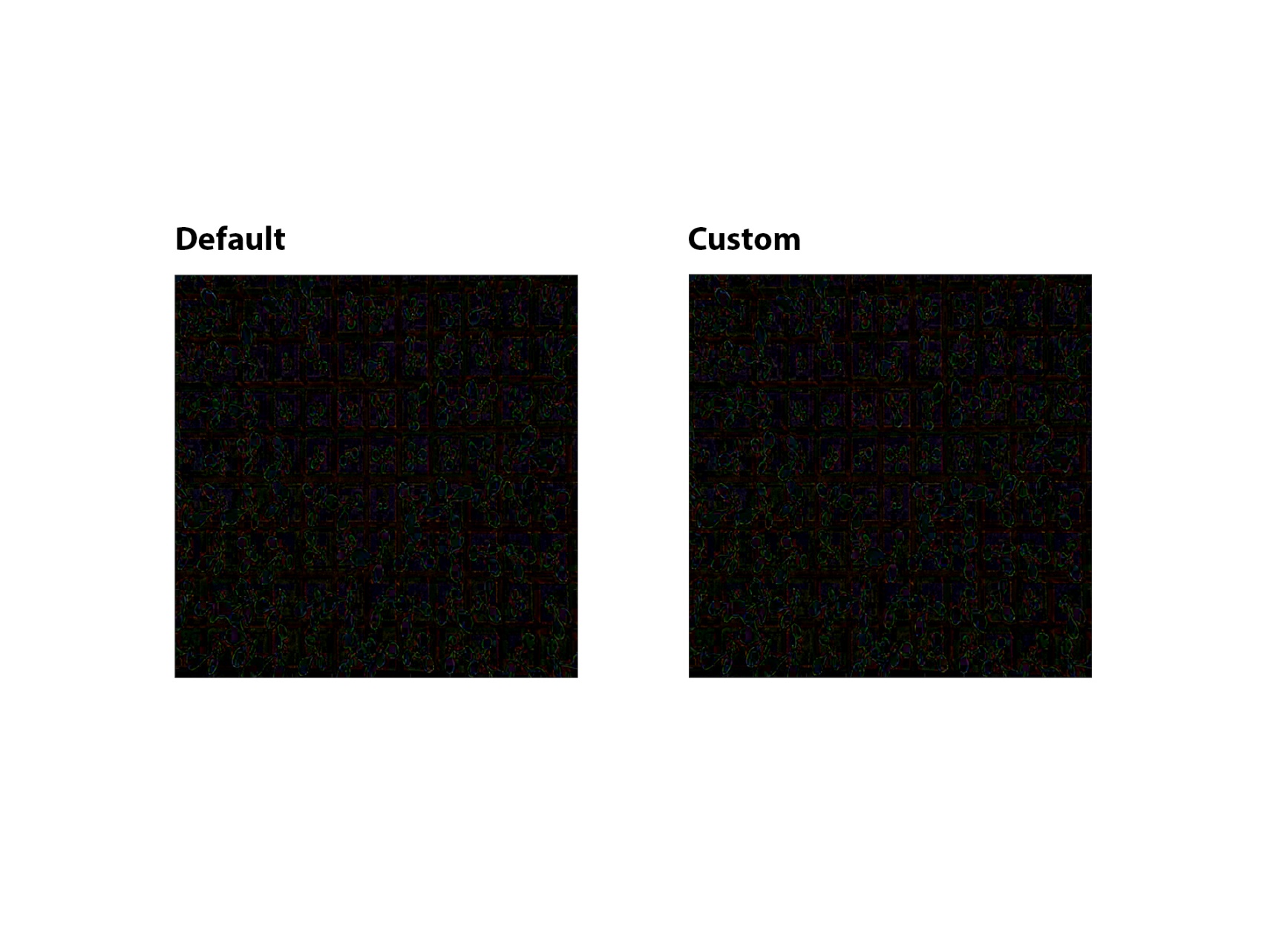
Figura – Gráfica Imagen 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| FC | MSE | DEFAULT | MSE | CUSTOM | RC | DEFAULT | RC | CUSTOM |
| 1 | 0.6735 | 0.6735 | 32.5640 | 39.1390 |
| 25 | 33.7043 | 33.7043 | 82.4212 | 82.96 |
| 50 | 59.4832 | 59.4832 | 88.1384 | 88.3939 |
| 100 | 95.1046 | 95.1046 | 92.0545 | 92.2226 |
| 250 | 176.2319 | 176.2319 | 95.5596 | 95.7405 |
| 500 | 298.8728 | 298.8728 | 97.1813 | 97.4738 |
| 1000 | 549.6785 | 549.6785 | 98.1355 | 98.525 |

[Explicación de datos]



[Explicación diferencias junto a los números]



[Explicación diferencia y conclusión]

## 4.4 Imagen 4

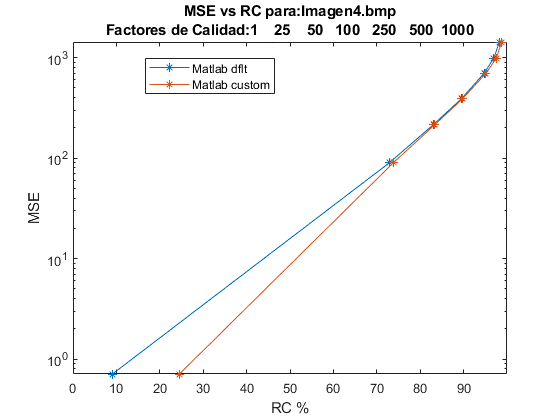
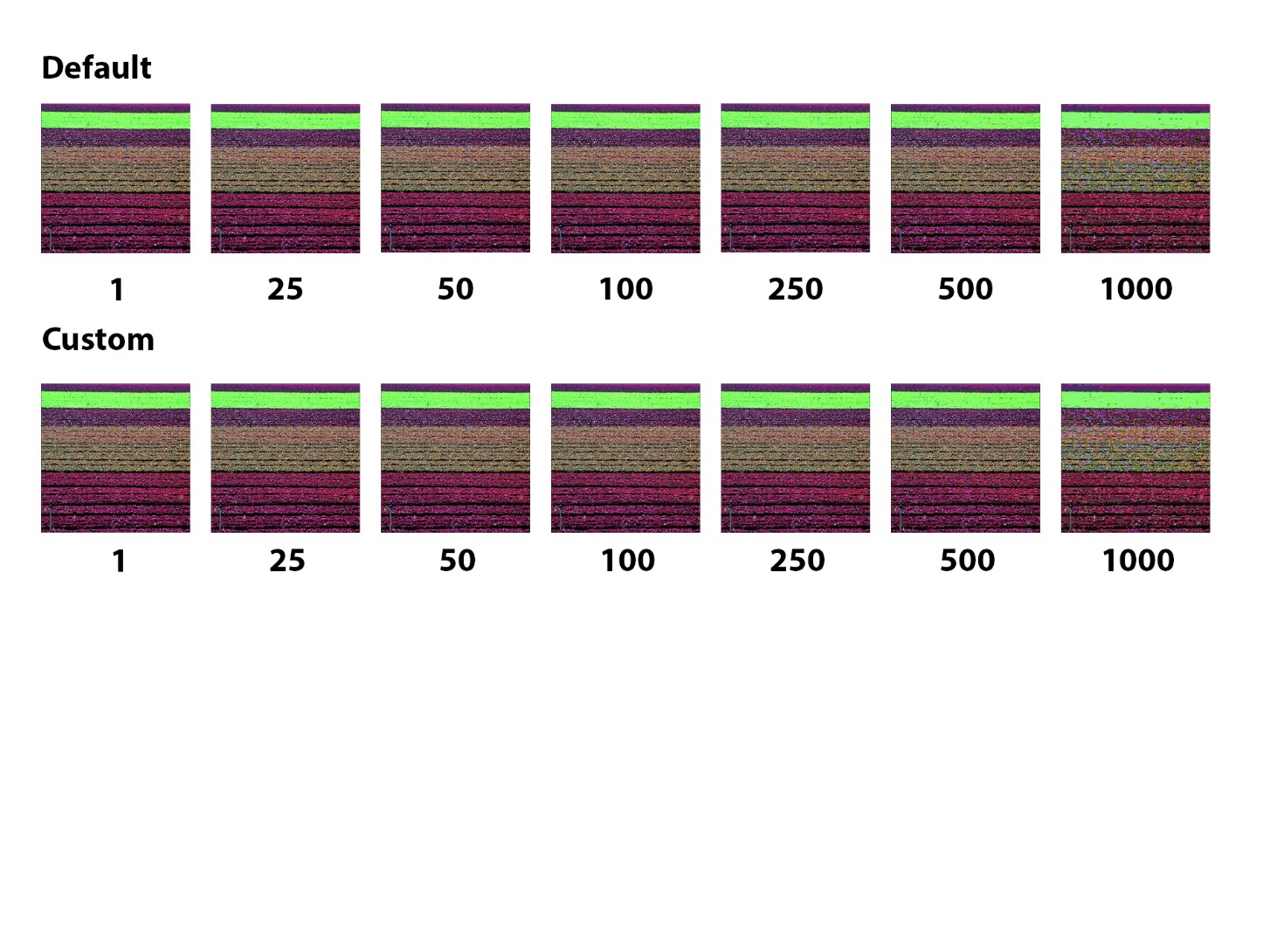


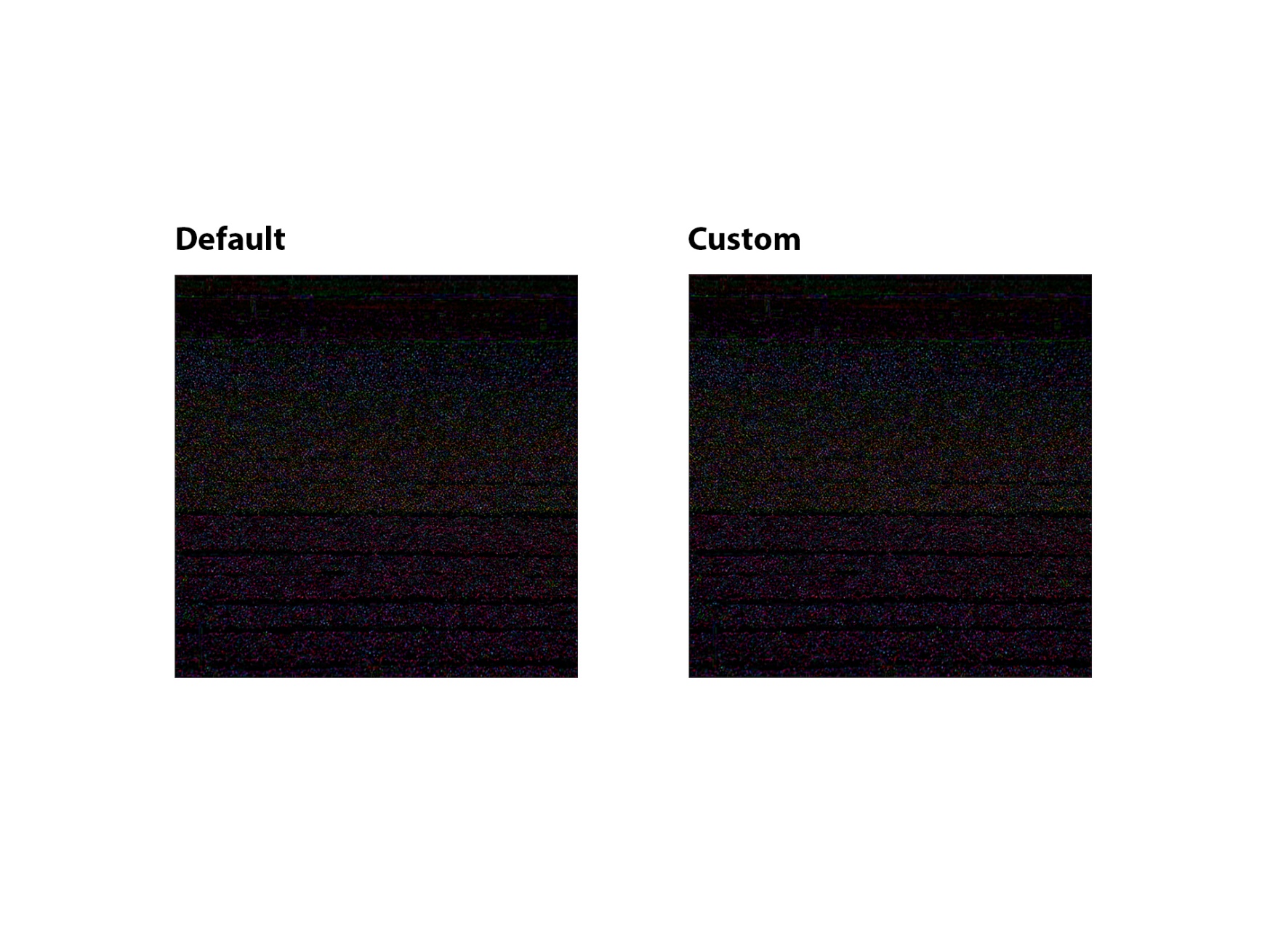
Figura – Gráfica Imagen 4

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| FC | MSE | DEFAULT | MSE | CUSTOM | RC | DEFAULT | RC | CUSTOM |
| 1 | 0.7126 | 0.7126 | 9.0890 | 24.5258 |
| 25 | 89.5999 | 89.5999 | 72.8602 | 73.7744 |
| 50 | 213.6284 | 213.6284 | 82.9538 | 83.2954 |
| 100 | 388.3704 | 388.3703 | 89.4751 | 89.7534 |
| 250 | 697.9605 | 697.9605 | 94.7237 | 95.1610 |
| 500 | 983.8421 | 983.8421 | 96.9412 | 97.5145 |
| 1000 | 1437.11 | 1437.11 | 98.2586 | 98.8088 |

[Explicación de datos]



[Explicación diferencias junto a los números]



[Explicación diferencia y conclusión]

## 4.5 Imagen 5

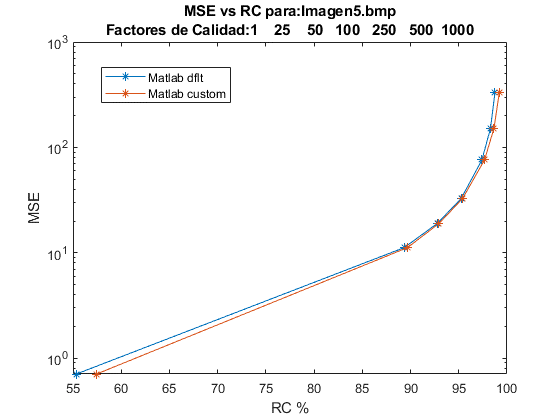
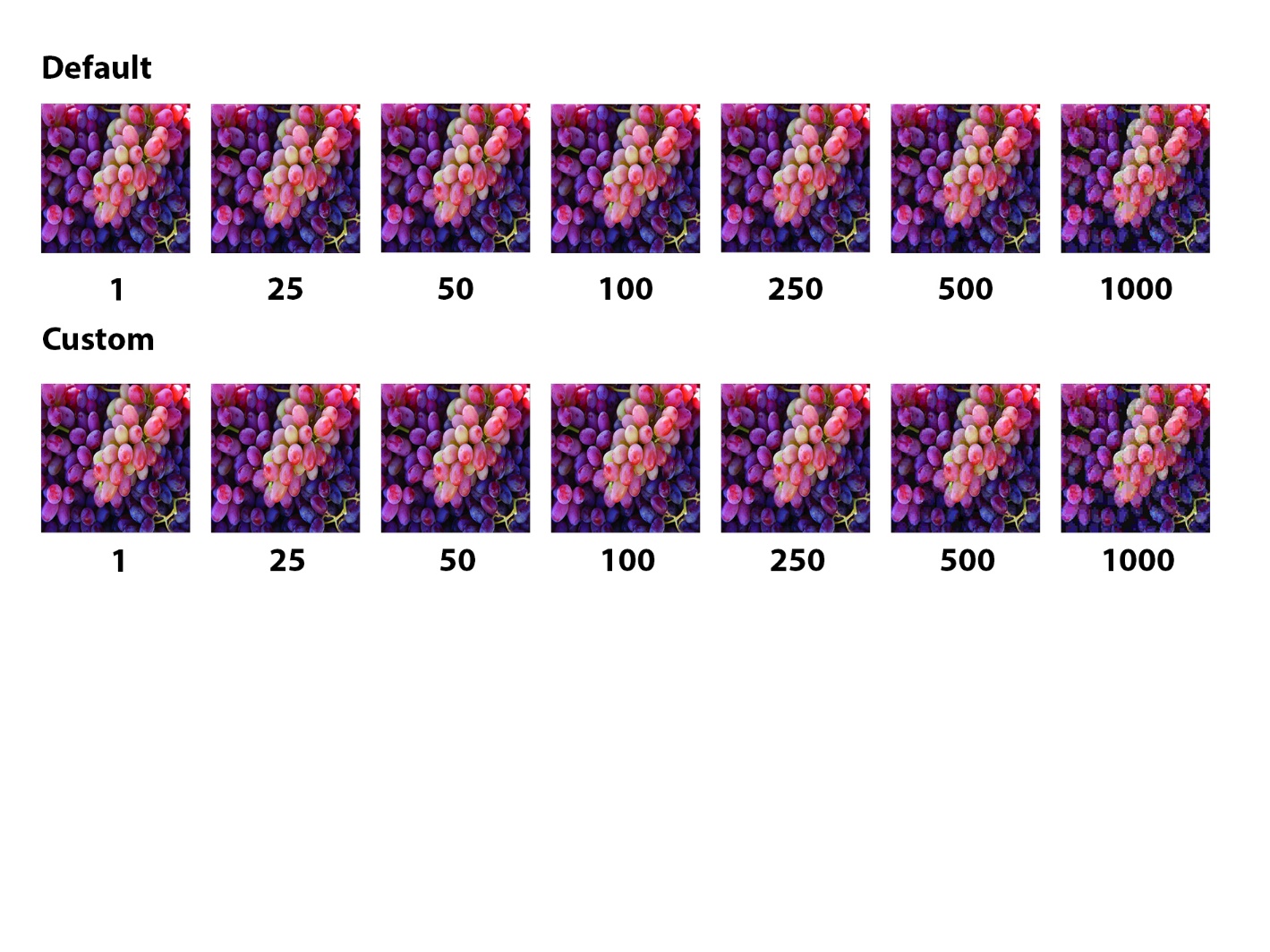


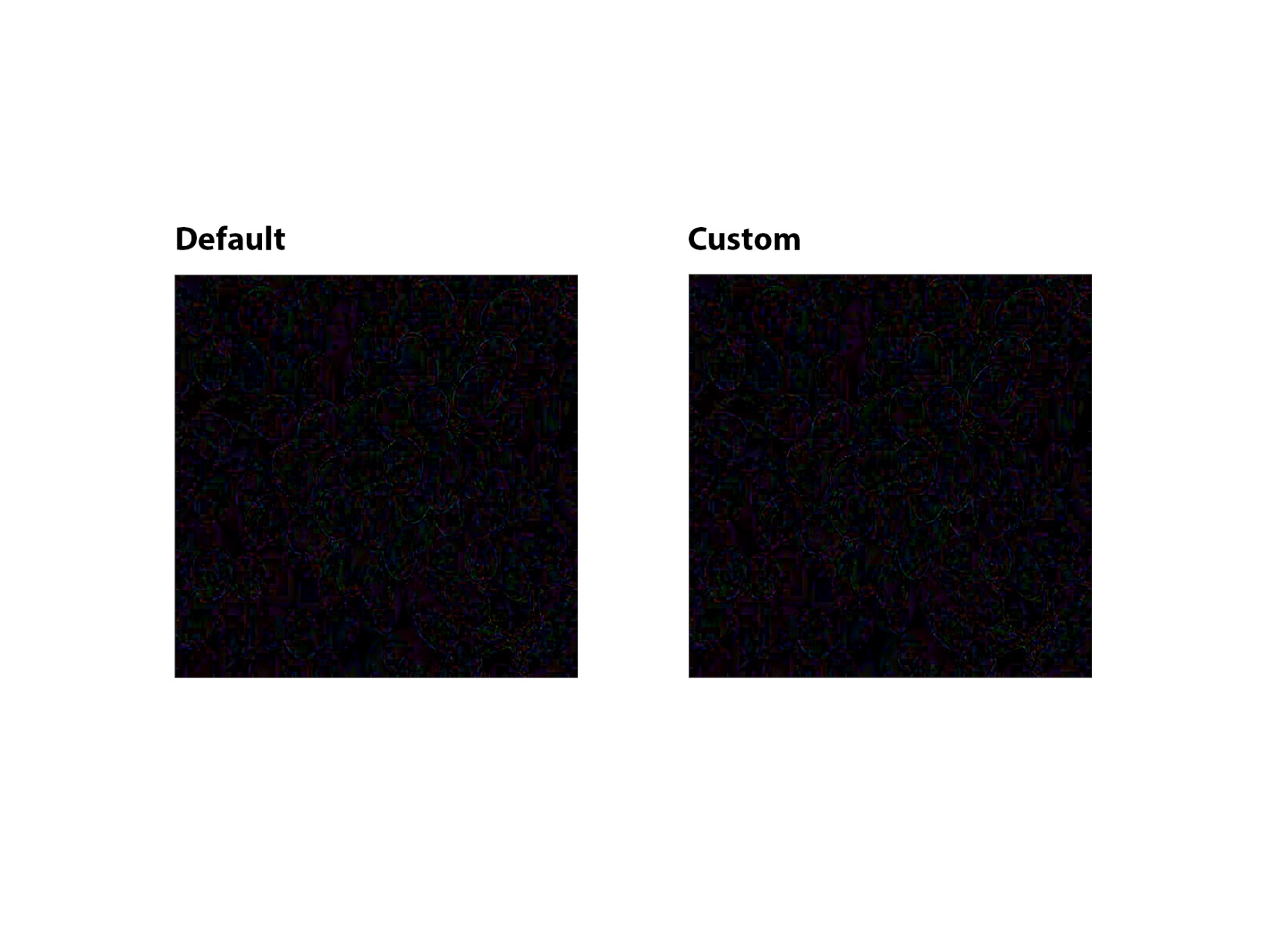
Figura – Gráfica Imagen 5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| FC | MSE | DEFAULT | MSE | CUSTOM | RC | DEFAULT | RC | CUSTOM |
| 1 | 0.7067 | 0.7067 | 55.3466 | 57.4507 |
| 25 | 11.1792 | 11.1792 | 89.3327 | 89.6772 |
| 50 | 18.9244 | 18.9244 | 92.7570 | 92.9402 |
| 100 | 32.6870 | 32.6870 | 95.2429 | 95.3936 |
| 250 | 76.0380 | 76.0380 | 97.4025 | 97.6717 |
| 500 | 151.9227 | 151.9227 | 98.2911 | 98.6517 |
| 1000 | 332.5231 | 332.5231 | 98.7543 | 99.2047 |

[Explicación de datos]



[Explicación diferencias junto a los números]



[Explicación diferencia y conclusión]

## 4.6 Imagen 6

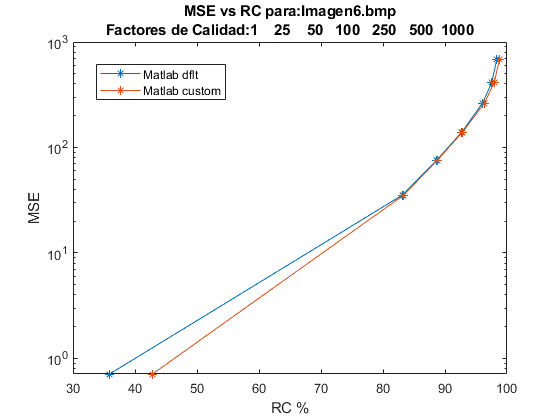


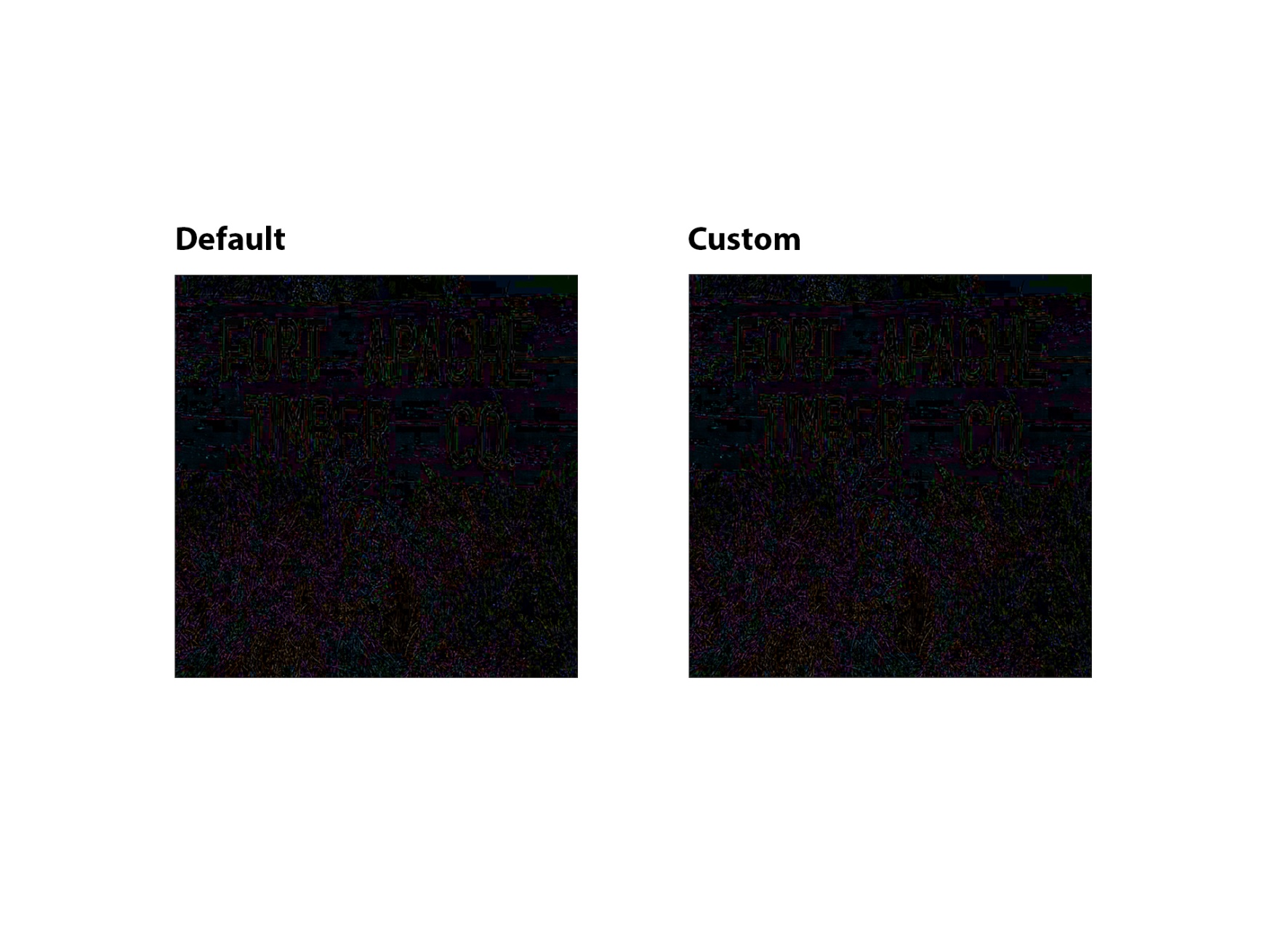
Figura – Gráfica Imagen 6

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| FC | MSE | DEFAULT | MSE | CUSTOM | RC | DEFAULT | RC | CUSTOM |
| 1 | 0.7119 | 0.7119 | 35.9031 | 42.8234 |
| 25 | 35.3456 | 35.3456 | 83.1041 | 83.3303 |
| 50 | 75.6001 | 75.6001 | 88.6372 | 88.7547 |
| 100 | 138.7864 | 138.7864 | 92.6629 | 92.7883 |
| 250 | 263.8015 | 263.8015 | 96.0627 | 96.3578 |
| 500 | 409.4427 | 409.4427 | 97.5164 | 97.9139 |
| 1000 | 687.5086 | 687.5086 | 98.3653 | 98.8039 |

[Explicación de datos]



[Explicación diferencias junto a los números]



[Explicación diferencia y conclusión]

# 5. Conclusiones

# 6. Bibliografía