

## GANADERÍA INTELIGENTE, SOSTENIBLE Y EFICIENTE.

Daniel Posada Murillo  
Universidad Eafit  
Colombia  
dposad21@eafit.edu.co

Jaime Uribe  
Universidad Eafit  
Colombia  
jruribem@eafit.edu.co

Simón Marín  
Universidad Eafit  
Colombia  
smaring1@eafit.edu.co

Mauricio Toro  
Universidad Eafit  
Colombia  
mtorobe@eafit.edu.co

### RESUMEN

En la actualidad, el 33% de las proteínas consumidas por los seres humanos provienen de la ganadería convirtiéndose así en un tema sumamente importante a tratar y mejorar de la mano con la tecnología para poder recolectar datos en tiempo real, mejorar así la productividad, detectar patologías con anticipación y hacer de esto algo más amigable con el medio ambiente. Estas tecnologías (GdP) consiguen una optimización de los recursos aumentando el rendimiento, controlando el impacto ambiental y mejorando el bienestar de los animales. Este mejor aprovechamiento de los recursos se lleva a cabo monitoreando la salud y producción, para ello se miden indicadores biométricos de carácter fisiológico y morfológico.

*¿Cuál es el algoritmo propuesto? ¿Qué resultados obtuviste? ¿Cuáles son las conclusiones de este trabajo? El resumen debe tener como máximo 200 palabras. (En este semestre, deberías resumir aquí los tiempos de ejecución, el consumo de memoria, la tasa de compresión y la exactitud).*

### Palabras clave

Algoritmos de compresión, aprendizaje de máquina, aprendizaje profundo, ganadería de precisión, salud animal.

## 1. INTRODUCCIÓN

En la actualidad, la ganadería de precisión es un tema que cada día va convirtiéndose en algo más necesario y útil donde se combinan las ciencias agropecuarias con las tecnologías de la información. Esto es con el fin de aumentar la producción y a su vez, el bienestar del animal recolectando datos en tiempo real que servirán para la temprana detección de patologías, conocer que sucede con el animal, crear parámetros para su habitat, reducir los costos de inversión entre otros

### 1.1. Problema

El problema al cual nos enfrentamos con los algoritmos es que se puede disminuir los tamaños a tal punto que con esto puede reducir la calidad de la imagen y se perdería información, sin embargo, a medida que surgen nuevos algoritmos, la pérdida de calidad de un archivo es menor, para permitir su transmisión y almacenamiento de forma rápida y eficiente.

Cuanto más grande sea el contenido de la información, más tiempo tomará enviarlo a su destino, y en algunos casos, disminuye la cantidad de bytes disponibles para enviar información, por ejemplo

### 1.2 Solución

De acuerdo con la complejidad de la información que vamos a manejar, se decidió usar el algoritmo del vecino más cercano (**k-nearest neighbors (K-nn)**), es un algoritmo de aprendizaje supervisado, es decir, que a partir de un juego de datos inicial su objetivo será el de clasificar correctamente todas las instancias nuevas. El juego de datos típico de este tipo de algoritmos está formado por varios atributos descriptivos y un solo atributo objetivo (también llamado clase).

### 1.3 Estructura del artículo

En lo que sigue, en la Sección 2, presentamos trabajos relacionales con el problema. Más adelante, en la Sección 3, presentamos los conjuntos de datos y los métodos utilizados en esta investigación. En la Sección 4, presentamos el diseño del algoritmo. Después, en la Sección 5, presentamos los resultados. Finalmente, en la Sección 6, discutimos los resultados y proponemos algunas direcciones de trabajo futuras.

## 2. TRABAJOS RELACIONADOS

En lo que sigue, explicamos cuatro trabajos relacionados. en el dominio de la clasificación de la salud animal y la compresión de datos. en el contexto del PLF.

### 3.1 “Precision livestock farming technologies for welfare management in intensive livestock systems”

Las patologías respiratorias son uno de los mayores problemas en los cerdos siendo esta una de las mayores causas de mortalidad de estos. Es por eso por lo que se desarrolló un modelo para clasificar sonidos de los cerdos y detectar a tiempo una tos fuera de lo normal evitando así su propagación.

[https://tice.agrocampus-ouest.fr/pluginfile.php/55014/mod\\_resource/content/17/res/u324\\_Berckmans\\_RevSciTechOffIntEpiz.PDF](https://tice.agrocampus-ouest.fr/pluginfile.php/55014/mod_resource/content/17/res/u324_Berckmans_RevSciTechOffIntEpiz.PDF)

**3.2** ¿ganadería de precisión como herramienta para el desarrollo sustentable?

<http://sigloxxi.espam.edu.ec/Ponencias/VIII/I%20FISPAM C/SALA2/EPP-002-2019.pdf>

### 3.3 Escriba un título para el tercer problema

Deberían mencionar el problema que resolvieron, el algoritmo que usaron, la métrica que obtuvieron y la cita de ACM.

### 3.4 Escriba un título para el cuarto problema

Deberían mencionar el problema que resolvieron, el algoritmo que usaron, la métrica que obtuvieron y la cita de ACM.

## 3. MATERIALES Y MÉTODOS

En esta sección, explicamos cómo se recogieron y procesaron los datos y, después, diferentes alternativas de algoritmos de compresión de imágenes para mejorar la clasificación de la salud animal.

### 3.1 Recopilación y procesamiento de datos

Recogimos datos de *Google Images* y *Bing Images* divididos en dos grupos: ganado sano y ganado enfermo. Para el ganado sano, la cadena de búsqueda era "cow". Para el ganado enfermo, la cadena de búsqueda era "cow + sick".

En el siguiente paso, ambos grupos de imágenes fueron transformadas a escala de grises usando Python OpenCV y fueron transformadas en archivos de valores separados por comas (en inglés, CSV). Los conjuntos de datos estaban equilibrados.

El conjunto de datos se dividió en un 70% para entrenamiento y un 30% para pruebas. Los conjuntos de datos están disponibles en <https://github.com/mauriciotoro/ST0245-Eafit/tree/master/proyecto/datasets>.

Por último, utilizando el conjunto de datos de entrenamiento, entrenamos una red neuronal convolucional para la clasificación binaria de imágenes utilizando *Teachable Machine* de Google disponible en <https://teachablemachine.withgoogle.com/train/image>.



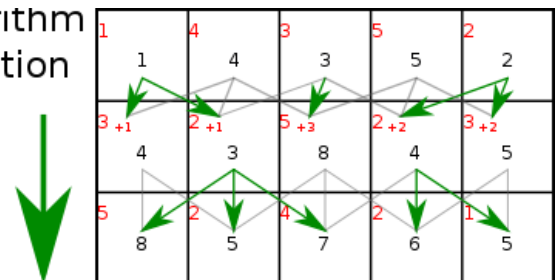
## 3.2 Alternativas de compresión de imágenes con pérdida

### 3.2.1 Seam Carving

Es un algoritmo que nos permite modificar las imágenes sin pérdida de información del contenido.

“En pocas palabras, a cada valor de píxel se le asigna un valor de energía, y luego, en función de este valor de píxel, se utilizan 8 dominios conectados para la programación dinámica para obtener el valor mínimo, y luego, aplicando este algoritmo fila por fila o columna por columna obtendrá una línea de energía, que en realidad es la fase La línea del píxel con el valor de píxel más pequeño de dos filas (columnas) adyacentes se elimina de la imagen original. Cuántas de esas líneas se pueden eliminar depende de la escala que se va a recortar.”

### Algorithm Direction

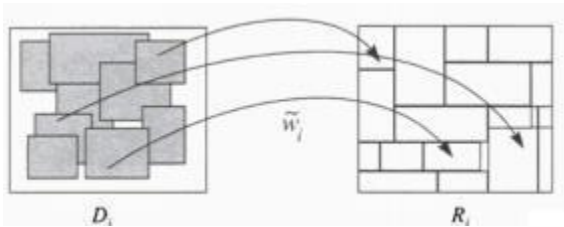


### 3.2.2 Escalada de imágenes

Es el cambio de tamaño de una imagen digital donde al escalar una imagen, se genera una nueva imagen con un número mayor o menor de píxeles

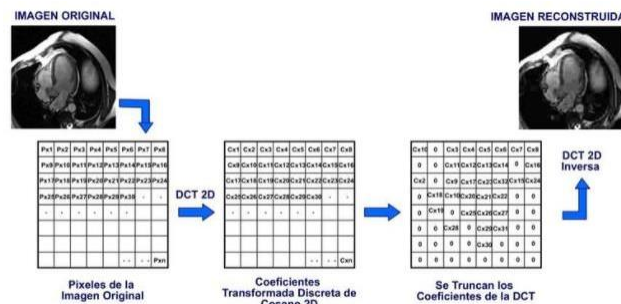
### 3.2.3 compresión fractal

Los algoritmos Fractales se utilizan para describir formas complejas con una reducida cantidad de información mediante sencillos procesos iterativos. Estas técnicas tienen especial interés en el tratamiento de imágenes, tanto en la generación y simulación como en la compresión.



### 3.2.4 transformada discreta de cosenos

La imagen se debe dividir en bloques de 8 x 8 pixeles, para aplicar mejor la redundancia de información. El proceso de codificación se realiza con la fórmula principal de la transformada. Como la imagen es una señal bidimensional, la ecuación de la transformada debe adaptarse como una ecuación bidimensional.



## 3.3 Alternativas de compresión de imágenes sin pérdida

### 3.3.1 La transformada de Burrows-Wheeler

Este algoritmo tiene como función el reordenar una cadena de caracteres en series de caracteres similares arrojando como resultado los mismos datos de entrada, distorsionando su ordenamiento; dicha distorsión tiene la capacidad de revertirse.

		$F$	$L$		$L$	$R$	
1	mississippi	1	i mississip p	1	p	1	p i
2	ississippi	2	i ppimissis s	2	s	2	s i
3	ssissippi	3	i ssippimis s	3	s	3	s i
4	ssippimis	4	i ssissippi m	4	m*	4	m* i
5	ssippimiss	5	m ississippi i	5	i	5	i m
6	ssippimissi	6	p imississi p	6	p	6	p p
7	sippimissis	7	p pimississ i	7	i	7	i p
8	ippimississ	8	s ippimissi s	8	s	8	s s
9	ppimississi	9	s issippimi s	9	s	9	s s
10	pimississip	10	s sippimiss i	10	i	10	i s
11	imississippi	11	s sissippi i	11	i	11	i s

(a)

(b)

(c)

(d)

### 3.3.2 LZ77

Este algoritmo pertenece al grupo de comprensores sin Pérdida. Funciona mediante el concepto de Ventana móvil el cual tiene la función de restringir la búsqueda de coincidencias a una parte de los datos de longitud previamente determinada. La ventaja es que el número de bits para indicar referencias se conoce de antemano y no puede crecer.

					s	a	l	s	→ (0,0,s)		
					s	a	l	s	a	→ (0,0,a)	
					s	a	l	s	a	□	→ (0,0,l)

Una vez que la segunda s pasa a ser el carácter futuro inicial, se produce una coincidencia de longitud 2, la cadena sa, y hay que ir 3 cuadros hacia el pasado para encontrarla:

			s	a	l	s	a	u	s	→ (3,2,u)
--	--	--	---	---	---	---	---	---	---	-----------

El resto de la codificación, hasta llegar al fin de los datos sería:

	s	a	l	s	a	u	s	a	l				→ (0,0,s)
s	a	l	s	a	u	s	a	l	a				→ (6,3,a)
a	u	s	a	l	a	d	a						→ (0,0,a)
u	s	a	l	a	d	a							→ (4,1,#)

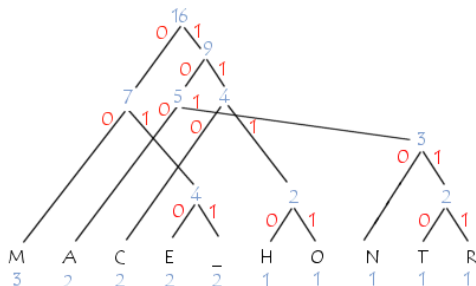
### 3.3.3 codificación Huffman

Es un algoritmo usado para la compresión de datos usando como método la construcción de un árbol binario donde a cada símbolo le sigue un código. Dicho proceso termina cuando solo hay un nodo representando así el fin del árbol.

Para este proceso se calcula la probabilidad que tiene cada carácter para aparecer, luego a los dos menos probables se les asigna un 0 y un 1 respectivamente. Luego se suman las probabilidades de ambos caracteres la cual se lleva a un tercer carácter.

M	A	C	E	_	H	O	N	T	R
3	2	2	2	2	1	1	1	1	1

Éste es el árbol correspondiente:



## 4. DISEÑO E IMPLEMENTACIÓN DE LOS ALGORITMOS

En lo que sigue, explicamos las estructuras de datos y los algoritmos utilizados en este trabajo. Las implementaciones de las estructuras de datos y los algoritmos están disponibles en Github<sup>1</sup>.

### 4.1 Estructuras de datos

La compresión imágenes es la reducción o compresión de los datos digitales que no resultan necesarios e importantes. Esta compresión permite almacenar mayor número de imágenes al conseguir que los archivos resultantes no ocupen mucho espacio de almacenamiento. Los grandes tamaños de los archivos de imagen han hecho que se desarrollen varios métodos para comprimir datos. La compresión es una parte muy importante en el flujo de trabajo con imágenes digitales ya que sin esta los archivos ocuparían la mayor parte de nuestros sistemas de almacenamiento y su velocidad de transferencia se vería disminuida.

8	2	8	4	8	5
1	8	2	1	8	9
9	6	6	1	6	6
9	7	3	9	2	9
4	2	5	5	6	8
3	5	3	5	8	5

Figura1: Matriz 6x6, que contiene 6 columnas y 6 filas, para un total de 36 elementos

### 4.2 Algoritmos

En este trabajo, proponemos un algoritmo de compresión que es una combinación de un algoritmo de compresión de

imágenes con pérdidas y un algoritmo de compresión de imágenes sin pérdidas. También explicamos cómo funciona la descompresión para el algoritmo propuesto.

#### 4.2.1 Algoritmo de compresión de imágenes con pérdida

La compresión con pérdida descarta información que considera innecesaria consiguiendo una reducción de datos mucho mayor. El principal inconveniente, como su nombre indica, es que la calidad puede verse mermada hasta niveles inaceptables.

Para este tipo de compresión se maneja el algoritmo del vecino más cercano (k-nearest neighbors (K-nn)), este algoritmo consiste en buscar las matrices y seleccionar ciertos píxeles para representar la imagen y colocarlos en una imagen comprimida. Para la descompresión, lo que hace es expandir la imagen comprimida por el mismo factor en que se comprimió, por lo que la imagen descomprimida da como resultado las mismas dimensiones que antes de ser comprimida, pero dado que solo se seleccionan algunos de los píxeles para representar la imagen, hay alguna pérdida de datos después de la descompresión.

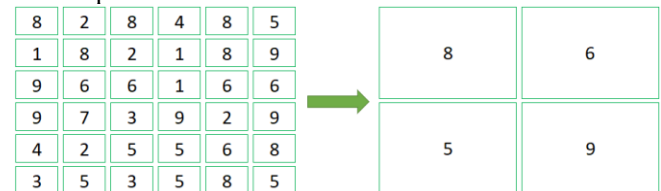


Figura2: Compresión de imagen usando K-nn

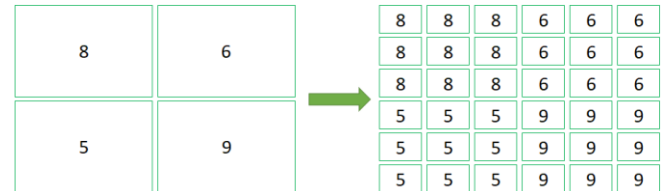
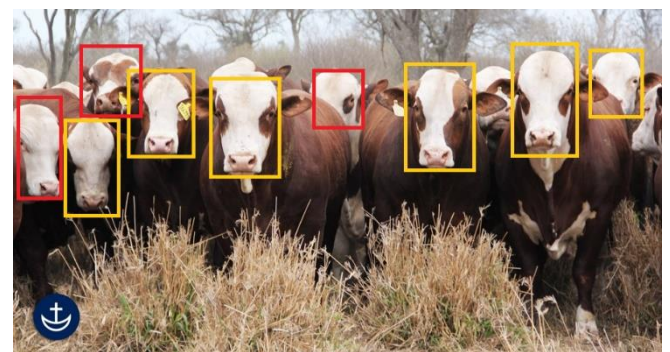


Figura3: Descompresión de imagen usando K-nn



#### 4.2.2 Algoritmo de compresión de imágenes sin pérdida

La compresión sin pérdida se utiliza cuando se quiere conservar la calidad de la imagen. De este modo se organizan los datos de la imagen con mayor eficiencia sin borrar información lo que nos ayuda a tener una buena conversión de la imagen.

La compresión también depende del contenido de la imagen. A mayor detalle menos puede comprimirse. Los dos archivos de imagen que tienen la opción lossless (sin pérdida) más usados son el formato TIFF y el PNG.

<sup>1</sup><http://www.github.com/dposada13/ST0245-001/proyecto/>

La codificación Huffman crea una tabla de códigos con las combinaciones de bits que más se repiten estadísticamente. A estas sentencias se asignan códigos cortos y a otros menos probables se asignan códigos de una longitud mayor, pero provoca problemas al cambiar la imagen original.

El problema que presenta es que un diccionario de claves único tiene resultados muy diferentes en distintos originales.

El funcionamiento de Huffman es:

- Contar cuantas veces aparece cada carácter
- Crear una lista enlazada con la información de caracteres y frecuencias.
- Ordenar la lista de menor a mayor en función de la frecuencia.
- Convertir cada elemento de la lista en un árbol.
- Fusionar todos estos árboles en uno único, para hacerlo se sigue el siguiente proceso, mientras la lista de árboles contenga más de un elemento:
  - o Con los dos primeros árboles formar un nuevo árbol, cada uno de los árboles originales en una rama.
  - o Sumar las frecuencias de cada rama en el nuevo elemento árbol.
  - o Insertar el nuevo árbol en el lugar adecuado de la lista según la suma de frecuencias obtenida.
- Para asignar el nuevo código binario de cada carácter sólo hay que seguir el camino adecuado a través del árbol. Si se toma una rama cero, se añade un cero al código, si se toma una rama uno, se añade un uno.
- (Cero a la izquierda y Uno a la derecha)

### 4.3 Análisis de la complejidad de los algoritmos

Explique, con sus propias palabras, el análisis del peor caso usando la notación O. ¿Cómo calculó tales complejidades. Por favor, explique brevemente.

Algoritmo	La complejidad del tiempo
Compresión	$O(N^2 * M^2)$
Descompresión	$O(N^3 * M * 2^N)$

**Tabla 2:** Complejidad temporal de los algoritmos de compresión y descompresión de imágenes. *(Por favor, explique qué significan N y M en este problema).*



Algoritmo	Complejidad de la memoria
Compresión	$O(N \cdot M \cdot 2N)$
Descompresión	$O(2M \cdot 2N)$

**Tabla 3:** Complejidad de memoria de los algoritmos de compresión y descompresión de imágenes. *(Por favor, explique qué significan N y M en este problema).*

#### 4.4 Criterios de diseño del algoritmo

Explica por qué el algoritmo fue diseñado de esa manera. Use un criterio objetivo. Los criterios objetivos se basan en la eficiencia, que se mide en términos de tiempo y consumo de memoria. Ejemplos de criterios no objetivos son: "Estaba enfermo", "fue la primera estructura de datos que encontré en Internet", "lo hice el último día antes del plazo", etc. Recuerde: Este es el 40% de la calificación del proyecto.

### 5. RESULTADOS

#### 5.1 Evaluación del modelo

En esta sección, presentamos algunas métricas para evaluar el modelo. La exactitud es la relación entre el número de predicciones correctas y el número total de muestras de entrada. La precisión es la proporción de estudiantes exitosos identificados correctamente por el modelo a estudiantes exitosos identificados por el modelo. Por último, sensibilidad es la proporción de estudiantes exitosos identificados correctamente por el modelo a estudiantes exitosos en el conjunto de datos.

##### 5.1.1 Evaluación del conjunto de datos de entrenamiento

A continuación presentamos las métricas de evaluación del conjunto de datos de entrenamiento en la Tabla 3.

	Conjunto de datos de entrenamiento
Precisión	0.02
Precisión	0.03
Recordar	0.01

**Tabla 3.** Evaluación del modelo de clasificación de imágenes con el conjunto de datos de entrenamiento.

##### 5.1.2 Evaluación del conjuntos de datos de prueba

A continuación presentamos las métricas de evaluación del conjunto de datos de prueba, en la Tabla 4, sin compresión y, en la Tabla 5, con compresión.

	Conjunto de datos de prueba
Exactitud	0.01
Precisión	0.012
Sensibilidad	0.013

**Tabla 4.** Evaluación del modelo de clasificación de imágenes, con el conjunto de datos de prueba, sin compresión.

	Conjunto de datos de prueba
Exactitud	0.001
Precisión	0.0012
Sensibilidad	0.0013

**Tabla 5.** Evaluación del modelo de clasificación de imágenes, con el conjunto de datos de prueba, con compresión.

#### 5.2 Tiempos de ejecución

En lo que sigue explicamos la relación entre el tiempo promedio de ejecución y el tamaño promedio de las imágenes del conjunto de datos completo, en la Tabla 6.

Calcular el tiempo de ejecución de cada imagen en Github. Informar del tiempo medio de ejecución vs. el tamaño medio del archivo.

	Tiempo promedio de ejecución (s)	Tamaño promedio del archivo (MB)
Compresión	100.2 s	12.4 MB
Descompresión	800.1 s	12.4 MB

**Tabla 6:** Tiempo de ejecución de los algoritmos *(Por favor, escriba el nombre de los algoritmos, por ejemplo, tallado de costuras y LZ77)* para diferentes imágenes en el conjunto de datos.

#### 5.3 Consumo de memoria

Presentamos el consumo de memoria de los algoritmos de compresión y descompresión en la Tabla 7.

	Consumo promedio de memoria (MB)	Tamaño promedio del archivo (MB)

Compresión	634 MB	3.12 MB
Descompresión	9 MB	878.12 MB

**Tabla 7:** Consumo promedio de memoria de todas las imágenes del conjunto de datos, tanto para la compresión como para la descompresión.

Para medir el consumo de memoria, deberían usar un generador de perfiles. Uno muy bueno para Java es

VisualVM, desarrollado por Oracle, <http://docs.oracle.com/javase/7/docs/technotes/guides/visualvm/profiler.html>. Para Python, usa el *C profiler*.

### 5.3 Tasa de compresión

Presentamos los resultados de la tasa de compresión del algoritmo en la Tabla 8.

	<i>Ganado sano</i>	<i>Ganado enfermo</i>
Tasa de compresión promedio	1:23	1:34

**Tabla 8:** Promedio redondeado de la tasa de compresión de todas las imágenes de ganado sano y ganado enfermo.

## 6. DISCUSIÓN DE LOS RESULTADOS

Explique los resultados obtenidos. ¿Son la exactitud, la precisión y la sensibilidad apropiadas para este problema? ¿El modelo está sobreajustado? ¿Es apropiado el consumo de memoria y el consumo de tiempo? ¿Es la relación de compresión apropiada? ¿Cambia la compresión significativamente la exactitud con el conjunto de datos de la prueba? (En este semestre, según los resultados, ¿puede la compresión mejorar la clasificación de la salud animal en el contexto del PLF? )

### 6.1 Trabajos futuros

Responda ¿qué le gustaría mejorar en el futuro? ¿Cómo le gustaría mejorar su algoritmo y su implementación? ¿Qué tal usar la transformación de coseno discreto o la compresión con ondeletas a futuro?

### RECONOCIMIENTOS

Identifique el tipo de reconocimiento que quiere escribir: para una persona o para una institución. Considere las siguientes pautas: 1. El nombre del profesor no se menciona porque es un autor. 2. No debe mencionar sitios web de autores de artículos que no ha contactado. 3. Debe mencionar estudiantes, profesores de otros cursos que le hayan ayudado.

Como ejemplo: Esta investigación fue apoyada/parcialmente apoyada por [Nombre de la Fundación, Donante, Beca].

Agradecemos la asistencia con [técnica particular, metodología] a [Nombre Apellido, cargo, nombre de la institución] por los comentarios que mejoraron enormemente el manuscrito o la codificación del algoritmo.

### REFERENCIAS

D. Berckmans, Precision livestock farming technologies for welfare management in intensive livestock systems,

[https://tice.agrocampus-ouest.fr/pluginfile.php/55014/mod\\_resource/content/17/res/u324\\_Berckmans\\_RevSciTechOffIntEpiz.PDF](https://tice.agrocampus-ouest.fr/pluginfile.php/55014/mod_resource/content/17/res/u324_Berckmans_RevSciTechOffIntEpiz.PDF)

Ernesto, Hurtado - Fátima, Arteaga-Chávez, ¿GANADERÍA DE PRECISIÓN COMO HERRAMIENTA PARA EL DESARROLLO SUSTENTABLE?, 2019

<http://sigloxxi.espm.edu.ec/Ponencias/VIII/I%20FISPAM C/SALA2/EPP-002-2019.pdf>

Gloria Puetamán Guerrero - Hernán Salazar Escobar, Compresión de imágenes usando wavelets

[https://repository.eafit.edu.co/bitstream/handle/10784/155/Gloria\\_PuetamanGuerrero\\_2007.pdf?sequence=1](https://repository.eafit.edu.co/bitstream/handle/10784/155/Gloria_PuetamanGuerrero_2007.pdf?sequence=1)

Fernando Chamizo, Resumen Codificación y compresión.

<http://matematicas.uam.es/~fernando.chamizo/asignaturas/model1415/entrop.pdf>

<https://www.fotonostra.com/fotografia/compresion.htm>