

Estructura de datos aplicada para la compresión de datos enfocado al ámbito de la ganadería.

Daniel Melguizo Roldan
Universidad Eafit
Colombia
dmelguizor@eafit.edu.co

Holmer Ortega Gomez
Universidad Eafit
Colombia
hortegag@eafit.edu.co

RESUMEN

El objetivo de este informe es digitalizar los datos que encuentran físicos en los cuadernos o sistemas de información de los zootecnistas. Mediante la compresión de los datos se busca una reducción considerable del consumo de energía empleado para estas tareas y el desarrollar las actividades de la ganadería de precisión de manera eficiente.

Se usó el algoritmo de compresión de imágenes JPG, comprime de manera muy eficaz las imágenes, logrando comprimirla de 4 a 5 veces su tamaño original y conservando una calidad aceptable de esta, en cuanto al tiempo de ejecución también es bastante apto y consume poca memoria.

Palabras clave

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

1. INTRODUCCIÓN

En estos días la salud animal en el ganado es muy importante ya que este es el consumo de la mayoría de las personas, aprovechando los recursos y métodos que se tiene para la compresión de datos y buscando una manera que se ahorre tiempo y dinero para el análisis del ganado, (porque actualmente algunas ganaderas optan por guardar los datos en cuaderno y no en bases de datos tecnológicas) de esta manera identificando cuando el ganado no se encuentre en condiciones para el consumo.

Para esto es muy eficiente el uso de un algoritmo que mediante la compresión de imágenes pueda aprender a identificar y brindar alternativas para que el ganado pase de no ser apto a ser idóneo para el consumo.

2. Problema

En la actualidad existen muchas ganaderas que no usan los sistemas tecnológicos para mejorar su producción agrícola de esta forma estas industrias tienden a tener muchas pérdidas millonarias. Se busca en la Ganadería de precisión, a través de la tecnología plantear soluciones enfocados a sistemas de producción agrícola dándole prioridad a las necesidades de los animales brindando una atención óptima para mejorar la calidad de producción con un objetivo de largo plazo que es suministrar un consumo adecuado de alimentos para la población mundial del futuro.

2.1. Solución

Mediante el uso del algoritmo de compresión de

imágenes (Nearest) se almacenarán estas imágenes comprimidas en un sistema de reconocimiento de salud animal, mediante este algoritmo se piensa tener un alto

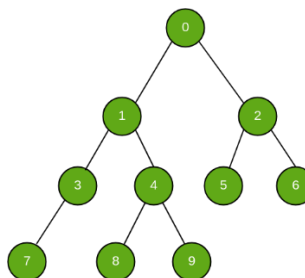
ahorro en memoria y energía a la hora de guardar imágenes y clasificarlas, notificando una alerta cuando se observe algo fuera de lo común.

3. TRABAJOS RELACIONADOS

3.1 Alternativas de compresión de imágenes con pérdida

3.1.1 Codificación triangular de árbol binario

El esquema de codificación se basa en la descomposición recursiva del dominio de la imagen en triángulos rectángulos dispuestos en un árbol binario. El método es atractivo debido a su codificación rápida, $O(n \log n)$, y decodificación, $O(n \log n)$, donde n es el número de píxeles, y porque es fácil de implementar y paralelizar. Los estudios experimentales indican que BTTC produce imágenes de calidad satisfactoria desde un punto de vista subjetivo y objetivo. Una ventaja de BTTC sobre JPEG es su menor tiempo de ejecución.



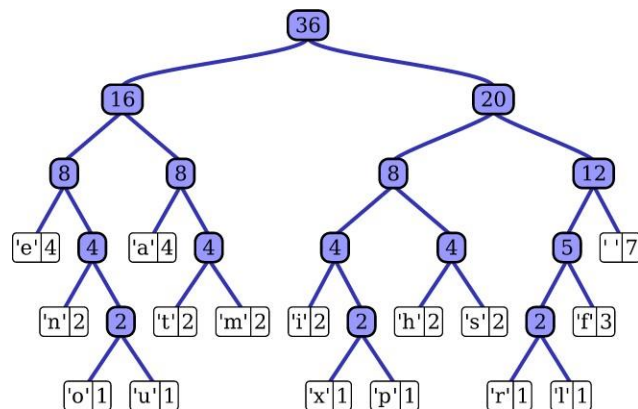
<https://www.geeksforgeeks.org/number-of-isosceles-triangles-in-a-binary-tree/>

3.2 Alternativas de compresión de imágenes sin pérdida

3.2.1 Algoritmo de Huffman.

Establecen 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

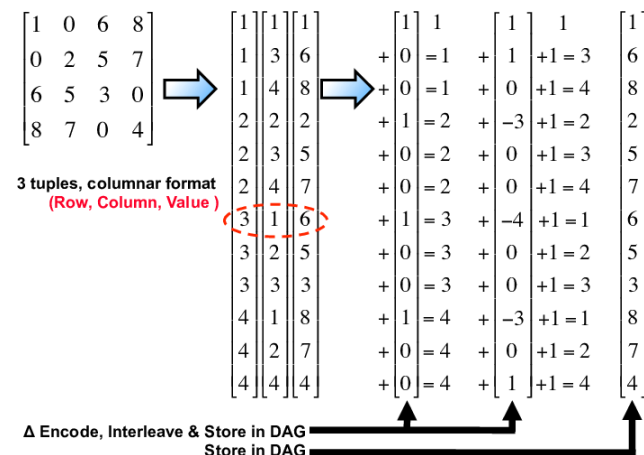
asignan códigos de una longitud mayor, pero provoca problemas al cambiar la imagen original.



<http://ramon-gzz.blogspot.com/2013/04/codificacion-huffman.html>

3.2.2 Delta encoding

Una función Delta hace referencia a la diferencia entre dos instancias de determinado objeto, en el caso de los archivos a respaldar, una función Delta para estos se refiere a la diferencia que existe dentro de un mismo archivo en dos versiones distintas. El algoritmo Delta encoding tiene como objetivo obtener solamente los bytes que han sido modificados desde la última versión del archivo, permitiendo reducir considerablemente el tamaño de este y lograr una optimización del uso de la red al momento de realizar un respaldo. Una vez que se tiene comprimido un archivo, se puede obtener el archivo original teniendo la versión de referencia del archivo y el archivo generado por el algoritmo de Delta encoding.

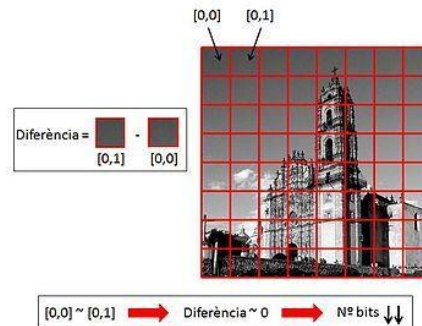


https://www.researchgate.net/figure/HCOO-Format-Delta-Encoding-and-Interleaving-Scheme_fig1_254462652

3.2.3 DPCM (Differential PCM)

DPCM Codifica la diferencia entre un valor y el anterior. Los sistemas DPCM codifican la diferencia entre un valor y el anterior. En codificación de imagen, se utiliza para codificar los coeficientes DC de la Transformada Coseno Discreta (Discrete Cosine Transform - DCT). Estas presentan el nivel de gris medio del bloque y, a menudo, el nivel medio de gris

de un bloque será similar al nivel medio de gris del bloque anterior. Por lo tanto la diferencia entre dos bloques consecutivos es mínima. Si por ejemplo codificamos una imagen donde hay una gran parte de cielo, entonces enviaremos números más pequeños, ya que la diferencia con el anterior valor será casi 0. Si no utilizáramos la codificación DPCM los valores serían mucho más grandes, la cual equivale a más bits en binario.



https://es.wikipedia.org/wiki/Codificaci%C3%B3n_de_diferencias

4. Diseño e Implementación de algoritmo.

4.1 Estructura de datos.

Usando la librería pandas, se implementan unos métodos que permitan la lectura de los datos que contienen los diferentes archivos csv, ya sea ganado sano o enfermo.

Para la compresión de imágenes con pérdida, se usa la librería PIL (pillow), que permiten la manipulación y guardado de los diferentes formatos de imágenes.



4.2 Algoritmo

Con Perdida

Se carga la imagen a comprimir, se crea una variable con el nombre que tendrá la imagen comprimida, y se almacena en el sistema con los pixeles y el peso del archivo reducido.

Original:

Tamaño: 983 KB (1.006.857 bytes)

1800 x 1799



Comprimida:

Tamaño: 59,2 KB (60.700 bytes)

1800 x 1799



4.3 Análisis de la complejidad de los algoritmos

La complejidad de nuestro algoritmo de compresión es $O(n)$ ya que los elementos a comprimir están divididos en carpetas y estos forman parte de una lista, no revisa una misma carpeta más de una vez, por eso no aumenta exponencial y lo hace de manera constante

La complejidad en memoria se mantiene igual a la complejidad en el tiempo

4.4 Criterios de diseño del algoritmo

Se realizó este algoritmo porque se puede usar de manera fácil, rápida y eficaz, decidimos realizarlo ordenándolo por carpetas ya que facilita su uso y comprensión.

5. RESULTADOS

5.1 Tasa de compresión

Comprime de manera muy eficaz las imágenes, logrando comprimirla de 4 a más veces su tamaño original y conservando una calidad aceptable de esta.

5.2 Tiempos de Ejecución

	<i>Tiempo promedio de ejecución (s)</i>	<i>Tamaño promedio del archivo (MB)</i>	<i>Tamaño del archive comprimido (KB)</i>
<i>Compresion</i>	0,22 s	1,95 MB	413KB

5.3 CONCLUSIÓN

Este algoritmo de compresión de imágenes permite de manera rápida y sencilla almacenar una gran cantidad de imágenes en sus respectivas carpetas, mediante la compresión se logra disminuir de 5 a más veces el tamaño en memoria que tendrían estas respectivas imágenes, además se logra mantener una resolución aceptable perdiendo detalles insignificantes. Se logró entender los conceptos de compresión y aplicarlos en una IOT de reconocimiento de ganado enfermo y sano.

REFERENCIAS.

Harry Alexander Vite Cevallos, & Oliverio Napoleón Vargas González. (2018). Ganadería de precisión en la provincia de El Oro. Diagnostico situacional. Espirales: Revista Multidisciplinaria de Investigación, 2(17). <https://doi-org.ezproxy.eafit.edu.co/10.31876/re.v2i17.263>

Stefanova, M., 2019. Precision poultry farming: software architecture framework and online zootechnical diary for monitoring and collaborating on hens' health. Commun. Comput. Informat. Sci. 953, 191–205. <https://www.springerprofessional.de/en/precision-poultry-farming-software-architecture-framework-and-on/16464166>

Wikipedia. Compresion de imagen. Retrieved January 1 2021.https://es.wikipedia.org/wiki/Compresi%C3%B3n_de_imagen

Wikipedia. Codificacion entropica. Retrieved October 24 2019.https://es.wikipedia.org/wiki/Codificaci%C3%B3n_entr%C3%B3pica