

Image Data Compression in the Public Reporting System in Lamongan using the Huffman Method and Run Length Encoding

*Mohammad Robihul Mufid, Saniyatul Mawaddah, Arif Basofi, Mochammad Jauhar Ulul Albab
Department of Informatic and Computer Engineering
Politeknik Elektronika Negeri Surabaya, Indonesia
{mufid, saniyatul, ariv}@pens.ac.id,
mjauhar@it.student.pens.ac.id

Nur Syaela Majid, Devvana Arya Pratama, Raisalatun Nuriyah, Nidalifa Choirunnisa
Department of Creative Multimedia Technology
Politeknik Elektronika Negeri Surabaya, Indonesia
{syelamajid, devvanaaryapratama, risalatunn,
nidalifach}@mb.student.pens.ac.id

Abstract— The dry season and the rainy season in the Lamongan area always cause several problems, including water scarcity, flooding, damaged roads, etc. The public report site, which is used to accommodate the aspirations of Lamongan residents who are in trouble, requires image evidence to ensure that there are no false reports. The site has a limit on the size of the file to be uploaded, so if the file size is too large, the upload process cannot be carried out. In this study, an analysis will be carried out to compare what compression method produces photos with the smallest size for further upload on the Lamongan community report site. Huffman compression and Run Length Encoding (RLE) were chosen because the algorithm includes lossless compression where the compressed image will not be damaged even though the size is compressed. From the two methods or algorithms, testing is carried out to find out which algorithm is the best that can produce compressed images with the smallest size. From the experiments conducted, it is known that the RLE method is a better method than Huffman coding. With the RLE method we can compress images up to 93.17%.

Keywords— *public reporting, image compression, RLE, Huffman*

I. INTRODUCTION

Lamongan is one of the areas in East Java which almost every year experiences problems in the dry season and rainy season. These include water scarcity, flooding, damaged roads, crop failure, kasus COVID-19, and so on. In overcoming all the problems in Lamongan, the government innovates to create a public reporting system to accommodate the aspirations of the community so that an appropriate policy can be taken for the community [1, 2].

Public services are important in supporting people's daily activities. Although the government

already has a geographic information system for the distribution of public services in Indonesia [3]. Community participation is needed to assist and improve government performance in providing better public services in national development efforts. Now the government has innovated public services regarding the means of public complaints through LAPOR! or Online People's Aspirations and Complaints Service is an information technology-based system developed by the Presidential Staff Office. This service has become the National Public Service Complaint Management System (SP4N) and has been implemented by all ministries and institutions in Indonesia.

The Lamongan Regency Government has also made innovations related to reporting by creating the LAPOR!

Website (<https://www.lapor.go.id/instansi/pemerintah-kabupaten-lamongan>). Lapor.go.id is a site that can be used by Lamongan residents to report problems, aspirations, criticisms and suggestions related to services in Lamongan Regency. When carrying out the reporting process, it is necessary to upload files as evidence so that the report can be ascertained not to be fake. On the website lapor.go.id there is a maximum size limit for uploaded photos. If the photo file taken is too large, it cannot be uploaded. For that it is necessary to do compression to compress the file size so that it gets smaller and the upload process can be done.

Compression is the process of converting a set of data into a coded form with the aim of saving the need for storage space and time for data transmission. Data compression is done by encoding each character in the message or data by encoding each character in

the message or data with a shorter code. The RLE and Huffman coding methods were chosen because both methods are lossless compression methods, namely compression that does not eliminate the original content of the data, such as removing or changing the original content during compression. The lossless compression method is suitable for compressing images that contain important information that should not be damaged by compression. Because the reporting process requires clear images even though they are compressed, the Huffman coding and RLE methods are very suitable for this research. The image to be uploaded will have a small file size but it is still clear and visible [4].

Actually, there have been several studies examining the RLE and Huffman Coding methods, including Herdianto proposed a study of comparison of the application of compression methods in text data. Where the method used is the RLE and Huffman coding method. Testing is done by compressing text data with a certain number of characters. From the test results it was found concluded that the Huffman method was better than the RLE for compression in text data [4].

A. Phani Sridhar who proposes an idea to compression with using the LZW method in medical records of patients who are increasing. Utilization of compression of this data is carried out to carry out the efficiency of storing medical data that increases sharply. The process of implementation on this research is done using Python and Hadoop. Results of testing of medical data compression using this LZW is better than RLE and Huffman [5, 6].

M. A. Hussin conducts a study of comparison of lossy data compression methods and lossless to be done on a 2D signal data. The Lossy method used is the Discrete Wavelet Transform (DWT) method and for lossless methods using the Huffman method. Based on the results of testing from the evaluation of the compression ratio, it is better to get Huffman, but for the results of the 2D signal output better by using the Huffman method [7].

While this paper discusses the comparison of the Huffman and RLE algorithms in uploading images on the public reporting system in Lamongan to make it easier for the community to express their aspirations.

II. OVERVIEW OF HUFFMAN CODING

Huffman Coding is a compression algorithm that has an output that will not lose data from the previous data. This algorithm runs by selecting a representative of each symbol or character to be used as a prefix. Huffman Coding has been widely applied to be used as a compression method in various multimedia data

because of its lossless nature (not losing the original data).

Huffman Coding Algorithm is used to compress multimedia data so that the data is lower than the previous size. The algorithm steps can be seen in the case of AABCDE word compression. Where the steps for completing the word compression begin with.

1. Select each symbol or character from the word and count the frequency, for example the word AABCDE consists of symbols or characters A(2), B(1), C(1), D(1), E(1).
2. Then select from the characters that have the smallest frequency, then combine them in a set.
3. Repeat step 2 until no symbols can be included in the set.
4. Then do the formation of the Huffman tree.
5. Then do the coding on the Huffman tree.

III. OVERVIEW OF RUN LENGTH ENCODING

The RLE algorithm or Run Length Encoding is an encoding that encodes a series of data that has the same value into one data value along with the number of occurrences. The number or length of the number of consecutive pixels that have the same data value is called run length, while encoding for compression needs is called run length encoding.

The RLE algorithm uses a space approach. This algorithm is used to compress text that has groups of pixels with the same gray level. This method converts an entire line of text into a single run line, then calculates the run-length for each successive gray level.

Most RLE methods are used on data containing uniform characters. Therefore, if the algorithm is used universally, it is necessary to group or transform similar characters/symbols. The steps for compressing data using Run Length Encoding are as follows:

- Run-length encoding (RLE) is a very simple data compression method
- This compression technique with RLE is useful for data that has a lot in common and the data is close together, for example text or graphics such as icons or line drawings that have a lot of similar patterns [10].

Image compression can be done using certain applications. In this study, we chose MATLAB R2015a as a tool or image compression medium.

IV. SYSTEM DESIGN

Here, we will discuss the system design, how the image compression process using the Huffman coding method and RLE using Matlab is carried out until the upload process is carried out on the lapor.go.id website. The following is a series of flows carried out:

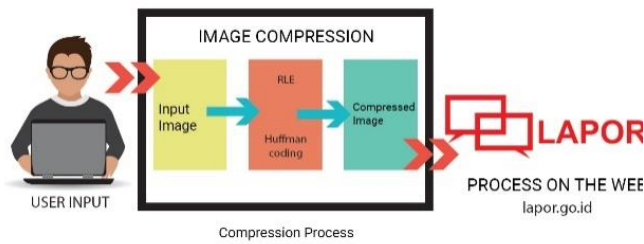


Fig. 1. System Design

4.1 User Input

This is the stage where when the reporter (the community) sees that there is a violation committed by the agency in its service, then the reporter must take a photo of the incident and collect it as evidence of a violation or complaint from the community to the Lamongan Regency Government. to be uploaded on the website lapor.go.id. The data collected is in the form of images, documents or videos with a maximum size of 2MB. The following is a series of user input processes:

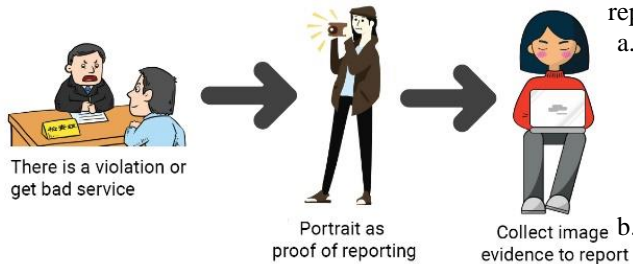


Fig. 2. User Input Process

In our research, we will compress 10 images of reporting evidence, each of which is compressed using the Huffman Coding and Run Length Encoding methods. Compression will be done using the MATLAB R2015a software to ensure the size of the file to be uploaded is no more than 2MB.

4.2 Compression Process

Image compression is done using MATLAB R2015a software by entering scripts and reporting evidence images that have previously been collected in the user input process. The following is a MATLAB R2015a display that will be used for the compression process. The compression process for both Huffman Coding and RLE using MATLAB R2015a is shown by a flow chart in figure 3.

1. Open Matlab R2015a Software.
2. Click New script and enter a compression script (Huffman Coding or RLE).
3. In the script, also included the image file to be compressed.
4. Then click RUN in the editor.
5. Wait a while until the results of the compression size appear in the command window.

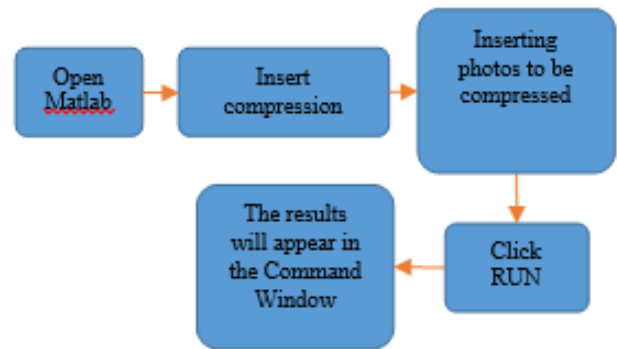


Fig. 3. Flowchart Compression

4.3 Verification Process on the website lapor.go.id

After the compression process has been successfully carried out, the compressed image will be uploaded to the lapr.go.id site. Here is the initial view of the website lapor.go id. Figure 4 show about reporting scheme. The reporting site itself has a reporting scheme.

- a. Report Input (REPORTING USER)

The process begins with community reporting related to complaints or aspirations clearly and completely (accompanied by evidence in the form of images, documents or videos up to 2MB).
- b. Report Verification (LAPOR.GO.ID)

Reports that have been submitted will then be verified within 3 days by the authorized agency.
- c. Follow-up Process (PEMDA)

Within 5 days, the authorized agency (PEMDA) will follow up and reply to incoming reports.
- d. Give Feedback (REPORTER)

Then the reporter can give a response related to the reply given by the relevant agency.
- e. Completed

The report is then followed up until the case is confirmed to be completed.

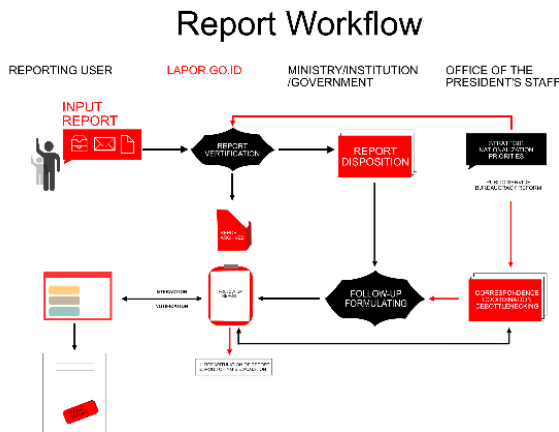


Fig. 4. Schematic of LAPOR

V. PERFORMANCE EVALUATION

In the reporting process on the lapor.go.id page, uploading photos as evidence, has a capacity limit of 2MB. Meanwhile, not all photos that will be uploaded have a capacity of 2MB, but exceeds the specified capacity. Therefore, a compression process is needed on the photos to be uploaded, so that they can be accepted by the system.

Compression testing was carried out on 10 sample images that would be used for uploading to the lapor.go.id page and then an analysis was carried out to compare Huffman Coding and RLE which of the two methods is better at compressing images. From the Huffman Coding and RLE compression experiments, the data obtained in graphical form is as follows:

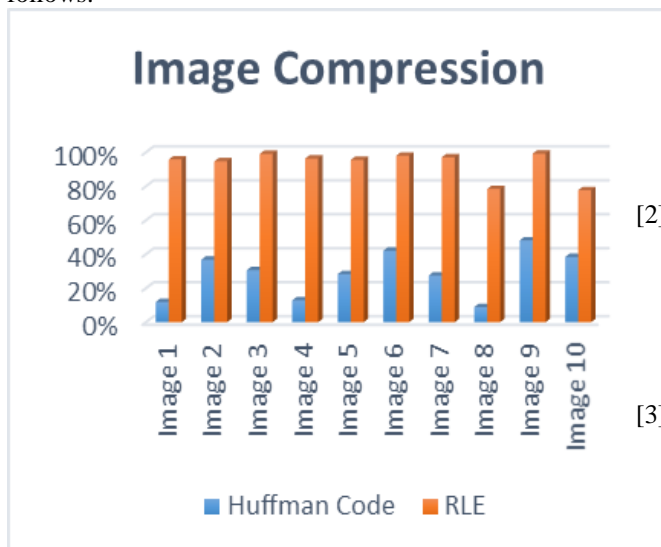


Fig. 5. Graph of Image Compression

From the data above, it can be seen that from 10 sample images compressed using Huffman Coding

can be reduced to 28.91% and images compressed using RLE can be reduced to 93.17%.

Based on the experiments conducted, when compared between the Huffman and RLE methods, the best method, which produces the smallest compression output, is the RLE method. With the RLE method, images can be compressed up to 93.17%.

From the experimental results obtained, it can be concluded that to compress image data in the public reporting system in Lamongan, it is better to use the RLE method as the compression method. Because the RLE method can compress image data to the maximum without affecting the quality of the image.

VI. CONCLUSION

Lamongan has a public reporting system that is used to accommodate people's aspirations when problems occur. In the reporting process, evidence is needed to ensure the authenticity of the report. There is a maximum limit for the size of the file to be uploaded, so compression is necessary. Based on the experiments we have done to compare the compression method between Huffman coding and RLE, we found that the best method in terms of reducing file size is RLE. Images can be compressed up to 93.17% without reducing image quality. For further research, the author recommends to try using the different compression algorithm to images for get best result compression.

REFERENCES

- [1] D. K. Sari, I. B. Riharjo, M. P. Priyadi, "Accountability of Cash-Intensive on Village Funds (Case study in Kendal Village, Sekaran District, Lamongan Regency)," Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences, Vol. 4, No.1, pp. 1385-1396, 2021.
- [2] Khusnul Khotimah, Mohammad Robihul Mufid, Arif Basofi, Saniyatul Mawaddah, Nurul Fuad, "Risk diagnosis and mitigation system of covid-19 using expert system and web scraping", International Electronics Symposium (IES), pp. 577 - 583, 2020.
- [3] M. R. Mufid, N. R. K. S. Putri, A. Fariza, "Fuzzy Logic and Exponential Smoothing for Mapping Implementation of Dengue Haemorrhagic Fever in Surabaya," International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), pp. 372-377, 2018.
- [4] Herdianto, Herdianto, "Perbandingan Metode RLE dan Huffman dalam Kompresi Data Teks," Jurnal Ilmiah Core IT: Community

- Research Information Technology Vol. 9, no. 1, 2021.
- [5] A. Phani Sridhar, P. V. Lakshmi, "An Efficient Lossless Medical Data Compression using LZW compression for Optimal Cloud Data Storage," *Annals of the Romanian Society for Cell Biology* Vol. 25, no. 6, pp. 17144-17160, 2021.
- [6] M. R. Mufid, A. Basofi, I. Syarif, F. Sanjaya, "Estimated Vehicle Fuel Calculation Based on Google Map Realtime Distance," *International Electronics Symposium (IES)*, pp. 354-358, 2019.
- [7] Mumtaz Anwar Hussin, Farhana Ahmad Poad, Ariffuddin Joret, "A Comparative Study on the Performance of DWT and Huffman Compression Technique on a 2D Signal," *Journal of Electronic Voltage and Application* vol. 2, no. 1, pp. 11-19, 2021.