Development of an Algorithm-Based Analysis and Compression Integrated Communication Tracking Management Information System (iCTMIS)

Carlo Jude P. Abuda[1], Ritchell S. Villafuerte[2]
Department of Information Technology, Visayas State University Alangalang, Alangalang, Leyte, Philippines[1]
Department of Information Technology, Eastern Visayas State University, Tacloban City, Philippines[2]

***Abstract*—This study addresses the challenges of administrative tasks and communication tracking at Visayas State University Alangalang (VSUA), highlighting the inefficiencies in the current manual processes. The objective is to develop an Integrated Communication Tracking Management Information System (iCTMIS) that enhances operational efficiency by integrating Optical Character Recognition (OCR) and Lempel-Ziv-Welch (LZW) Lossless and Zlib compression algorithms. By employing a developmental research design and ADDIE model, the system proves that there is an improvement on data analysis and reduces disk space through efficient compression. Significant findings reveal that OCR achieves up to 90% accuracy in text conversion, while LZW compressions substantially deflate data sizes. This was evaluated against ISO 9126 Software Quality Characteristics, the iCTMIS has shown to optimize storage and address VSUA's operational challenges effectively. This research therefore concludes that the systematic integration of advanced algorithmic frameworks in iCTMIS significantly enhances organizational communication and administrative workflows efficiency.**

***Keywords*—Information system; optical character recognition; Lempel-Ziv-welch lossless compression; Zlib compression; communication tracking**

I. Introduction

Effective communication is fundamental to organizational success. Similarly, the need for appropriate software to record, track, and streamline internal and external communications is essential to achieve organizational efficiency [1]. These lay the foundation for streamlined operations that enhance overall efficiency in various transactions and workflows through seamless routing and tracking functions [2][3].

As communication tracking has evolved into a critical component of organizational management, facilitating information flow, leading to well-informed decisions, and improving overall productivity [4]. The roots of communication tracking can be traced back to the early days of administrative processes, where writing and recording correspondence were the primary means of information exchange [5], [6]. Over time, advancements in technology revolutionized communication tracking, with digital systems and algorithms playing a pivotal role in developing these processes [7].

Currently, communication tracking transcends to organizational boundaries, influencing both internal operations and external interaction [8]. Internally, effective communication tracking enhances transparency, accountability, and responsiveness within an organization [9]. Externally, it fosters collaboration and coordination on a global scale, which is essential for institutions with widespread operations or international partnerships. As organizations continue to grapple with the increasing volume and complexity of communication tracking, the need for sophisticated algorithmic frameworks becomes evident [10].

Recent advancements in algorithmic frameworks, such as Optical Character Recognition (OCR) and Lempel-Ziv-Welch (LZW) compression algorithms, have further augmented the capabilities of communication tracking systems [11], [12]. OCR has been proved to be an invaluable attribute in converting scanned documents into machine-readable text, enabling seamless integration into digital tracking systems [13].

Existing research literature supports the assertion that OCR significantly enhances the efficiency of communication tracking, reducing manual data entry and mitigating the risk of errors (e.g., misrouted documents, misspelled communication routing slips, and unreadable remarks) associated with human intervention [14], [15].

Similarly, the LZW compression algorithm contributes to improved communication tracking by reducing the storage space required for digital documents. As documents are archived and indexed, LZW compression ensures efficient storage, making retrieval faster and more economical [16]. Integration of these algorithms into communication tracking is particularly relevant in the Philippine setting, particularly among State University and Colleges (SUCs), such as the Visayas State University Alangalang (VSUA) [17], [18].

VSUA is one of the State Universities and Colleges (SUCs) in the region committed to providing quality education and fostering research and innovation, Visayas State University Alangalang (VSUA) faces notable challenges in administrative responsibilities and factional dynamics. While dedicated to receiving and utilizing data from stakeholders for public services, the administrative processes continue to exhibit bureaucratic tendencies. The institution grapples with several challenges in managing and tracking communication data [19], [20].

As the current manual document tracking system at Visayas State University Alangalang is significantly observed that it hampered by inefficiencies and operational challenges, including the weighty process of tracking routing numbers, the ambiguity in document routing leading to potential misdirection, the frequent misplacement of vital communication letters, and a general reluctance towards embracing technological advancements.

Furthermore, the existing infrastructure struggles with the demands of modern data management, evidenced by issues with system capacity, memory allocation, outdated scanning devices, and software that cannot efficiently handle large volumes of data due to a lack of centralized storage solutions. This decentralized approach not only complicates the reception and recording of documents but also results in physical storage problems, such as the excessive accumulation of paper documents and the consequent risk of damage from pests.

Addressing these challenges through the development of an Analysis-Compression Algorithm-Based Integrated Communication Tracking Management Information System could revolutionize the university's information management system, streamlining processes, enhancing efficiency, and ultimately fostering a more dynamic and responsive educational environment, and by also investigating and obtaining the capabilities of OCR and LZW compression algorithms, the researcher seeks to streamline communication tracking within VSUA. The research is motivated by the recognition that an integrated approach, which combines the strengths and capabilities of both algorithms, has the potential to significantly enhance the efficiency and effectiveness of the existing communication tracking system.

The researcher explored into the complexities of OCR and LZW compression algorithms, exploring their capabilities and potential in the context of communication tracking [21], [22]. The goal of this study to contribute valuable insights and practical solutions that can be adopted not only by VSUA but also by other government agencies facing similar challenges. The research is positioned not only to improve the existing manual system of VSUA Management Information System but also as a testament to the adaptability of algorithmic frameworks in addressing applied organizational complexities [23].

The general objective of the study was to enhance the Communication Tracking Management Information System of Visayas State University Alangalang by implementing analysis-compression algorithms. The specific objectives driving this endeavor encompassed the following focal points that is to reduce the disk space and memory allocation among data and files using LZW compression algorithm. Eliminating noise and converting files-to-text among documents using the OCR analysis algorithm and evaluating the analysis-compression algorithm to its system requirements based on ISO 9126 the Software Quality Characteristics such as functionality, reliability, usability, efficiency, maintenance, and portability.

II. Review of Related Systems and Studies

A. OCR Algorithm for Document Analysis Framework

In the review of related literatures and studies, it is imperative to check and examine existing research and systems that investigate into communication tracking, integrated information management, and algorithm-based solutions within administrative contexts and Integrated communication tracking management information systems aided with a literature review map to identify the sub-themes of each major focal points [24] that are necessary for the Visayas State University Alangalang to streamline administrative processes, enhance transparency, and ensure efficient dissemination of crucial information.

According to Memon (2023) Optical Character Recognition (OCR) has been developed by individual researchers with greater accuracy. The literature has concluded that utilization of OCR frameworks acts differently on different languages due to character style and dataset quality. As it has also been supported by some researchers, they proposed several solutions that is to provide one language or single subset of a language as an input. As the literature shows some practical implications of the said literature, the development of machine learning and deep learning enables accurate recognition of handwritten manuscripts. Towards the development of the, the researcher used several methods in OCR frameworks using the Systematic Literature Review (SLR), some on machine learning techniques, template matching technique, distance (similarity) metrics and Convolutional Neural Networks (CNN). Moreover, the advent of the different techniques performs a better on different languages due to variations in character style and dataset quality [25].

As OCR Framework evolves and adopts to the abrupt development in the field of Information Technology, according to Sahu and Sonkusare (2019) there has been another technique that can be incorporated and partnered with OCR, that is the Magnetic Character Recognition or MCR where there are two frameworks used in recognizing a more complex recognition on specific inputs. The methods discussed in this literature use the OCR used to identify scripts or alphabets in verbal communication primarily used in banking and other industries wherein handwritten text were the primary inputs of the framework [26]. Hence, these conclude and provide a more efficient performance on the OCR with MCR framework thus the researcher of this study suggested that there could be more methods in OCR to be integrated.

As the evolution of OCR to a more sophisticated, it has been already used and applied now to some sectors and institutions wherein according to Karthikeyan (2021) proposes an Internet of Things (IoT)-based library management system using OCR algorithm which then includes a CCTV-based book issuing and returning mechanism. OCR is used to convert text files into audio files for accessibility but not limited to scanning damaged books and converting them into PDF format. This literature has concluded that with OCR being implemented with IoT, book issuing and returning system is more efficient and secure. However, limitations were seen and observed that in in the event of scanning defect with bar code is a challenge in the existing system in the said intuition, which then in the proposed system introduces an effective and time-saving asset tracking and administration system for library using RFID technology [27].

According to Arief et al., (2022) the accuracy rate of document classification achieved at 94% in terms of the document classification origin and subject as CNN methods will

correctly classify the type of character being used as an input. CNN on the other also captures errors that occur in the regular expression's method coming from the original and subject classification as mentioned in the previous literatures. These methods include the utilization of automated hierarchical classification using CNN and regular expression methods, preprocessing with Tesseract OCR and Word2Vec. As this literature concludes, the automatic hierarchical classification method is necessary as this also utilized the classified and analyzed classified documents are stored on hive databases – the Hadoop architecture, wherein the databases are stored and systematized in big data technology [28].

Clearly, this development led to the application of Artificial Intelligence (AI) which have then been used already in OCR frameworks. To the degree of OCR application, AI techniques have improved OCR technologies according to Jain et al. (2023) in the application for general text recognition. OCR models trained on general text struggle with localized or personalized handwritten text. This study aims to create an adaptive framework for OCR models. It develops a digit recognizer using a convolutional neural network. Results show comparable accuracy for localized or personalized handwritten text. The study suggests data augmentation as a solution for scarce and imbalanced data [29].

Another piece of literature that is relevant to the study being conducted is the application OCR with Mobile integration. According to Bisiach & Zabkar (2020), their study compares OCR methods for mobile platforms in prescription label scanning, wherein these methods are pertaining to three methods being evaluated, namely the classic computer vision, standard deep learning, specialized deep learning. To distinguish between these three methods, it has been concluded that Standard Deep Learning (StanDL) (Tesseract 4.1.1) provides the best combination of accuracy, speed, and resource usage. As observed during the implementation and deployment, Tesseract 4.1.1 achieves 76% accuracy, with 10% results being one character away from accurate. Tesseract 4.1.1 achieves 76% accuracy, with 10% results being one character away from accurate. Moreover, 9% of images processed in less than one second, 41% processed in less than 10 seconds. Furthermore, Tesseract 4.1.1 has reasonable resource costs comparable to non-deep learning methods. As the researcher concluded, the application of Tesseract 4.1.1 to OCR Framework had shown a reasonable resource costs comparable to non-deep learning methods [30]. In this study, the methods applied uses classic computer vision techniques, standard deep learning, and specialized deep learning (Tesseract 4.1.1).

In eliminating data noise among files and images, Mande Shen & Hansheng (2019), presents a method to eliminate background images in OCR. As the methodology it provides evidence of enhanced document images and converts color images to gray. Background images are effectively removed without losing text quality. The method improves recognition accuracies in OCR. The researchers have also justified the methods based on the difference in color values of background image pixels. Such uses brightness distortion and chromaticity to enhance contrast. In has shown that the test experiments showed that the output image is clean after preprocessing. Moreover, OCR frameworks perform much better on images with background eliminated or the researcher classified it as document noise. Accordingly, background images can be effectively removed using the method they have used at the rate of 80% to 90% text were recognized and blank pages were eliminated. Moreover, in the readability of documents it is improved after removing background images, in addition the recognition accuracy of OCR are significantly improved. Frameworks also in OCR namely the HANWANG and ABBYY software show significant improvement in OCR performance. OCR Tesseract returns 2% wrong results when given background images without preprocessing as reported, however, in conclusion, all of these three OCR software obtain good results on blend regions (mix of background image and characters) and the algorithm assumes colorful background images, towards the end the researcher of this literature justifies more that there are still gaps in the methods of OCR in addressing black-white backgrounds [15].

The above cited literature provides a complete and comprehensive synthesis on the different gaps, relationships, applications and approaches that may be applied along with the development and deployment of this study specifically on the integration of the existing system of the Commission on Higher Education Regional Office VIII the iCTMIS with OCR Framework in document analysis algorithm in eliminating data noise among files.

B. LZW Algorithm for Document Compression Framework

As the researcher of the study introduces the different applications and uses of compression framework and algorithms, the introduction of LZW compression will be the guided path to look for similarities and useful insights from the literature and research done by other researchers. To provide a context, LZW compression is the second algorithm framework that this study will be using, such that defining LZW is known as a method focused on reducing the size of Tag Image File Format (TIFF) or Graphics Interchange Format (GIF) files [31]. This technique employs a table-based lookup algorithm to eliminate duplicate data, effectively compressing original files into smaller formats. Beyond image files, LZW is also adept at compressing text and PDF files. Rooted in the LZ78 algorithm developed by Abraham Lempel and Jacob Ziv in 1978, LZW compression unfolds as a versatile tool with implications across various data types.

Starting from the different applications and approaches of LZW compressions framework, According to Shah (2019) states that there is an innovative approach to increase the compression ratio of the LZW algorithm. LZW, a widely used lossless compression algorithm for data compression, involves appending frequently encountered string patterns to the dictionary. This selective addition of high probability words reduces the number of bits required. Additionally, this approach has evidently presented an increased compression ratio of the LZW algorithm, reduction in the consumption of exclusive resources and improved data compression techniques for efficient storage and transmission [32]. With these results, this study could implement this technique in compression algorithm that addresses the VSUA problems in indexing and archiving difficulties.

According to their study, it was concluded that an average compression ratio for the LZW algorithm is 42.85% which is more efficient than reduces to 38.55% using modified LZW lossless method and at the same time using also Variable Length Code. However, the compression rate speed did not improve much, indicating the same with the unmodified LZW algorithm. With these findings, this is still useful for the researcher to improve the areas specifically on the dictionary formation during the compression framework and to its ratio of files being compressed. Furthermore, some contributions were also highlighted such as the implementation of variable length code for encoding process, the comparison of data compression performance between LZW algorithm and proposed algorithm and creation of data compression application using Java programming language [33].

In relation to the efficiency of compressed files, a novel algorithm specifies locating patterns in compression using the LZW-compression framework. According to Adldwairi et al., (2019) this novel algorithm for locating patterns in LZW-compressed data, evidently provided an efficient and simple algorithm with superior time complexity, at the same rate it maintains space complexity similarly to the existing algorithms and significant improvement in search time compared to Aho-Corasick algorithm, as it is a scalable algorithm that improves with larger dataset sizes. As the methods discussed, the algorithm comprises a preprocessing phase and a subsequent search phase. It uses a modified version of the generalized suffix tree, a lookup table, a mapping table, and a history tree. The preprocessing phase involves constructing the LZW-AGS and its corresponding mapTable with two naive algorithms or Ukkonen's algorithm can be used for this task given. For the implementation details and practical considerations of the proposed by this approach coming from the researcher, this provides a theoretical evaluation of the algorithm and experimental evaluation of the algorithm [34].

This study addresses the gaps and problems in tracking, routing and moreover in space allocation, a literature provides an idea on address space allocation compression using the LZW framework. According to Safieh & Freudenberger (2019) the space partitioning techniques for parallel dictionary LZW (PDLZW) data compression algorithm. This literature proposes an address space partitioning technique for the PDLZW algorithm. The technique optimizes the compression rate using a Markov model for the data. On the numerical results demonstrated, the improved performance of the proposed partitioning [35].

This research seeks towards the gaps between tracking a literature also guided the researcher to consider the files utilizing and embedding LZW Lossless to Zlib file compressions. According to Yang et al., (2023) and Chirikhin & Ryabko (2019) claimed that Zlib framework library can fifty percent (50%) to seventy-five percent (75%) deflate algorithm for file compression and decompression. Moreover, file compression improves storage efficiency and transfer speed. These compression algorithms can be lossless or lossy, different compression programs and algorithms are used for different file formats also this compression technology has significant benefits in mass data storage and transmission. The benefits of decompression for different file formats need to be studied and evaluated. Limitations were not reported however, existing literatures provided that significant differences in compression performance of different file formats, where some formats have higher compression ratios and significant compression effects and recompression of already compressed formats may result in poor compression, hence, uncompressed formats tend to exhibit high compression ratios and significant compression effects [36], [37].

The literature presented were the different approaches, techniques, and applications of LZW compression algorithm. Moreover, the similarities found were the focal points of the research to align with the VSUA existing manual system such that in tracking of documents that correlates to the transfer rate of files and space allocation that inputs may be the factor to utilize the said algorithm. It suggests that LZW lossless compression, primarily applied to modified versions, achieves an average compression rate of approximately 42.58%, with a focus on text compressions. Considering this, the researcher has introduced a framework that integrates and adapts Zlib data file compression capabilities, with this integration. It is justified by the proven effectiveness of Zlib, showcasing its ability to deflate and compress files at a notable ratio ranging from fifty percent (50%) to seventy-five percent (75%) among files being reported and used as data sets in the conduct of the experiment in the literatures provided.

III. METHODOLOGY

This research applied a developmental research design, in which the intervention happened by integrating the proposed system into the existing manual system of Visayas State University Alangalang that is systematically developed, refined, and evaluated [38]–[40]. Additionally, this research was geared towards a mixed method both qualitative and quantitative research approach and employed a design that performed a complete enumeration of participants, where all dedicated units serving as the end-users were identified and selected in the conducted evaluation of the proposed system.

The research setting for this study is situated within Visayas State University Alangalang, located in Alangalang, Leyte. The participants consist of the university's staff and faculty members who are invited to provide feedback through Google Forms distributed in the Chancellor's Office, the Records and Archives Office, and the Media Information System and Technology Office.

The selection of this setting for the study is strategic, as it aligns with the researcher's role as the head of the Media Information Systems and Technology Office, where the system under investigation is being implemented. This position offers the researcher a unique perspective and facilitates the process of data migration and encoding, ensuring smoother integration, and handling of incoming data.

The identification of the participants in the proposed system were determined on complete enumeration manner where the following units of Visayas State University Alangalang namely the Office of the Records and Archives (ORA), Office of the Chancellor (OOC) and Office of the Media Information Systems and Technology (MIST). In this view, the researcher identified two (2) participants from ORA, four (4) from OOC and four (4)

from MIST, giving the total participants of ten (10) coming from these identified participants.

Moreover, this study employed the ADDIE model [41]–[43], a widely recognized instructional design framework consisting of five stages: Analysis, Design, Development, Implementation, and Evaluation. This approach allowed the researcher to systematically plan, develop, and assess the effectiveness of development duration. In this study, the integration of analysis-compression algorithm to VSUA iCTMIS is developed using the said model by following the stages as follows:

A. Analysis

In the initial phase of this study, a thorough analysis was executed, employing a meticulous approach through complete enumeration, and listing of all the required aspects that the proposed system used and utilized. The process commenced with a systematic identification of the problem at hand from the stakeholders [44], [45] ensuring a clear understanding of the challenges and requirements faced by Visayas State University Alangalang. This step facilitated the subsequent establishment of well-defined goals and objectives crucial for guiding the research direction.

In Fig. 1, a valuable visual representation of the Use Case Diagram was provided by the researcher in which this outlines the different user roles or knowns as system actors, proposed system, objectives or goals and their interactions within the system. The figure aids in understanding the functionalities and features required for each user type or levels, contributing to a more refined and user-centric system design. This also let the researcher extend the problem identified, incorporating detailed planning to address all aspects of the integrated communication tracking management information system.

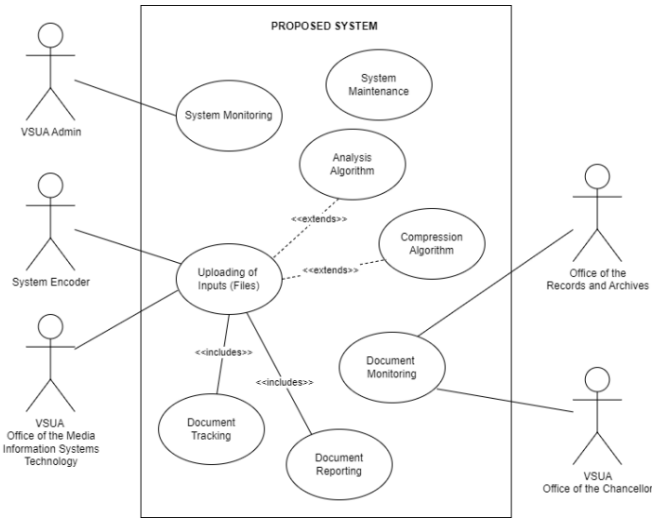

Fig. 1. Proposed system of use case diagram.

Simultaneously, the researcher also carefully selected the programming logic and approaches based on the language chosen for the development of the proposed system. The choice of programming language used is the PHP or Hypertext Preprocessor, along and its corresponding logic, this plays a pivotal role in ensuring the system's efficiency, scalability, and maintainability. For specific instance, the system is required in real-time processing, then researcher opted for a language having strong support for concurrent programming frameworks reviewed and identified in the previous phases.

Afterwards, with various consideration of programming elements complemented with visual representations, which formed a comprehensive approach to system analysis and laying the groundwork for the subsequent phases of development. Furthermore, the researcher also considered factors in ethical considerations, starting from the conduct of preliminary up to the gathering of data undergo to the protocols of Visayas State University Alangalang to take account of necessary actions and approved by the head of the institution coming from the Office of the Chancellor.

Lastly, this thorough process involved not only clarifying the problem but also scrutinizing the intricate details associated with the study goals and objectives. As such, by systematically addressing each component, the researcher ensured that the requirements were not only clearly stated but also precisely identified.

B. Design

During this phase, the researcher thinks of developing an initial prototype of the system. The goal of this stage is to create and identify the overall structure of the proposed system [46]. This includes designing the database, entity relationships, flowchart, data flow, wireframes, style guide, mockups, and the algorithm framework integrations to the existing manual system of VSUA Communication Tracking Management Information System. In Fig. 2 gives the overall methods of what the researcher applied and integrated the procedure in the implementation of the two algorithms, namely the analysis and compression algorithm.

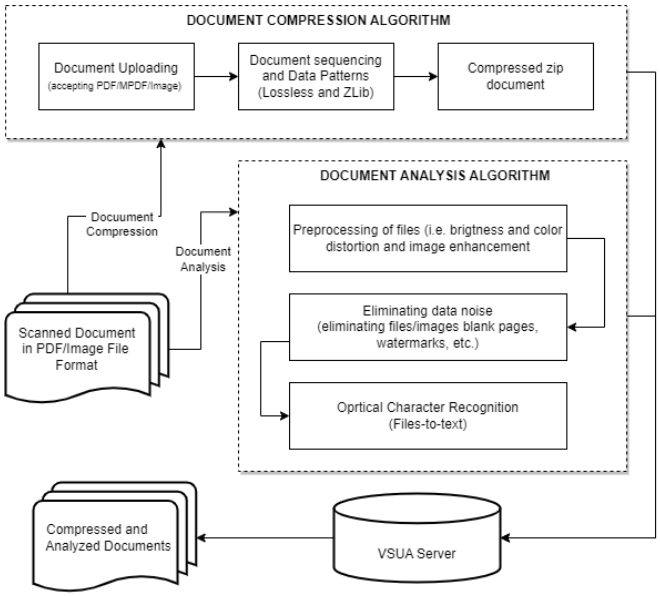

Fig. 2. Implementation of analysis-compression algorithm.

C. Development

In this phase, the development of the analysis-compression framework happens. The proposed system, the framework integration of OCR and LZW is written in PHP programming

language, MySQL for database, and the localhost server installation using XAMPP. In OCR framework two methods were applied in accepting data as input from the stakeholders, that is the acceptance of PDF/MPDF and images type files. Namely the methods applied were using various approaches namely in context of brightness distortion. The brightness distortion referred as a_i is obtained by minimizing by the following the formula:

$$\varphi(\alpha_i) = (p_i - \alpha_i E_i)^2 \quad (1)$$

where a_i represents the pixels strength of brightness with respect to the expected value. For accuracy of the results in *(1)*, a_i must be equal to one (1) if the brightness of the uploaded PDF document is the same to the output in text. Similarly, $a_i < 1$ means the PDF document upload is darker than the expected brightness, otherwise if $a_i > 1$ means it is brighter. An overall mean is computed from the sample data set that the PDF/images documents provided and uploaded are accessible and suitable for document analysis to OCR Tesseract file-to-text recognition. Furthermore, the aspect of color distortion on the other hand, where once the a_i is solved, the chromaticity denoted as CD_i, can now be determined and identified using the RGB color values given the equation below:

$$CD_i = \|p_i - a_i E_i\| \quad (2)$$

The color distortion is defined as the orthogonal distances between the observed color and the expected chromaticity of the file being uploaded and analyzed by the OCR Tesseract. That is using brightness and background subtraction to enhance the recognition of font text and suppress the background.

Lastly, after the framework extracted the brightness and color distortions it proceeds to the image enhancement wherein, image enhancement was the process in the analysis of documents to be specified, it is now the file analysis using OCR to fully maximize and remove data noise found in scanned documents uploaded in PDF/Image format, without altering the text in the foreground, the prior methods were the execution that every pixel must has R, G and B values. This now presumes that the files/images uploaded must have the mentioned three values otherwise, files are then converted using the RGB conversion again. But these values are significant in determining color extracted, process of converting document files and images to text using the OCR Tesseract framework. Thus, the researcher used the formula below in enhancing and converting the three values.

$$p_i = max\{0, min(255, (p_{i-128}) * CD_i + a_i)\} \quad (3)$$

Where in, the end results of these process identify the clean documents and OCR Tesseract successfully converts files from PDF/MPDF and images files to text files.

The researcher also used the modified following frameworks of LZW lossless and ZLib compression algorithm in the process of reducing the files in the manner of encoding the logic as presented in the below snippet Pseudocode 1 (presented in algorithm).

Pseudocode 1: Compression Algorithm

```
Initialize table with single character strings
initial = first input file
WHILE not end of input stream
    C = next input file
    IF initial + C is in the string table
        initial = initial + C
    ELSE
        output the code for initial
        add initial + C to the string table
        initial = C
END WHILE
output code for initial
```

In the application and execution of disk space and memory allocation, it was seen and observed that in the file being uploaded in the system under several logic coding executions following the pseudocode was referred.

In the execution of the algorithm the study used the programming language that is the PHP on its recursive Hypertext Preprocessor using the lossless and ZLib frameworks. Moreover, in Pseudocode 2, which represents source code, the execution of the algorithms in Lossless and Zlib occurs during the deployment of the system.

Pseudocode 2: Modified LZW Algorithm

```
import zlib frameworks
compress_file(input_file_path, compressed_file_path)
DEF  compress(data)
    dictionary = {chr(i): i for i in range(256)}
    current_code = 256
    result = []
    current_str = ""
    FOR char in data:
        initial = initial + C
        current_str += char
        IF current_str not in dictionary:
            result.append(dictionary[current_str[:-1]])
            dictionary[current_str] = current_code
            current_code += 1
            current_str = char
    IF current_str in dictionary:
        result.append(dictionary[current_str])
RETURN result
PRINT("Original Data:", original_data)
PRINT("Compressed Data:", compressed_data)
PRINT("Decompressed Data:", decompressed_data)

FUNCTION compress_file(input_file, output_file)
    data = open(input_file, 'rb').read()
    compressed_data = zlib.compress(data)
    open(output_file, 'wb').write(compressed_data)

input_file_path = 'example.txt'
compressed_file_path = 'example_compressed.zlib'
```

Subsequently, for OCR Algorithm, Pseudocode 3, simplifies the structure by removing unnecessary details and focusing on

the main functions of OCR Tesseract. Each preprocessing step (brightness distortion, color distortion, and image enhancement) is encapsulated in separate functions. Afterwards, the OCR conversion now to files-to-text was done are the cleaning or the preprocessing stages.

Pseudocode 3: OCR Tesseract Code Development

```
SET tesseractPath TO '/path/to/tesseract'
FUNCTION performOCRWithPreprocessing(originalImagePath)
    TRY
        SET brightenedImagePath TO
        applyBrightnessDistortion(originalImagePath, 1.5)
        SET colorDistortedImagePath TO
        applyColorDistortion(brightenedImagePath, 0.8, 1.2, 1.0)
        SET enhancedImagePath TO
        applyImageEnhancement(colorDistortedImagePath)
        RETURN performOCR(enhancedImagePath)
        CATCH Exception e
            PRINT "Error during OCR with preprocessing: " +
            e.getMessage()
        END TRY
    END FUNCTION

    FUNCTION applyBrightnessDistortion(imagePath,
    brightnessFactor)
        // Apply brightness adjustments to the image
        RETURN distortedImagePath
    END FUNCTION

    FUNCTION applyColorDistortion(imagePath, redFactor,
    greenFactor, blueFactor)
        // Apply color distortions to the image
        RETURN distortedImagePath
    END FUNCTION

    FUNCTION applyImageEnhancement(imagePath)
        // Apply image enhancement techniques
        RETURN enhancedImagePath
    END FUNCTION

    FUNCTION performOCR(imagePath)
        // Run OCR using Tesseract on the given image
        RETURN extractedText
    END FUNCTION
RETURN result
SET originalImagePath TO 'input_image.png'
CALL performOCRWithPreprocessing(originalImagePath)
```

D. Implementation

In this phase the actual testing of the proposed system with the integration of analysis-compression algorithm will be delivered to the intended end-users. Primarily, the proposed system will undergo two (2) phases of implementation. In the first phase of implementation, the proposed system was first tested by the programmer and at the same time the researcher along with ten (10) identified participants during the conduct of the Analysis Phase. All the concerns are recorded and afterwards, then satisfied, it proceeds to the second (2) phase where the proposed system is now being migrated to the Visayas State University Alangalang Database Center Server. Through engagement with the intended users, this phase enabled the identification of crucial insights to refine the app's usability, features, and user experience.

Moreover, the conduct of the second phase implementation took two days for the researcher to accomplish. Before the conduct of final and third implementation, the researcher first presented the activity's objectives and secured participants' informed consent. Upon agreement for voluntary involvement, participants must honestly evaluate the system. During this stage, participants will be able to afford ample time to explore the proposed system.

Following this, a Focus Group Discussion (FGD) was conducted to solicit qualitative feedback, aimed at enhancing the system's user interface, usability, features, and overall user experience. Moreover, in this phase the in-charge which is the MIST Office takes full responsibility of the implementation of the system for the installation, recording and creation of accounts pertaining to the usage of the proposed systems. This entails the gathered pertinent data from the analysis phase.

Lastly, the iterative nature of the ADDIE Model allowed the researchers to effectively address user feedback, ensuring that the final system version would cater comprehensively to the needs of all stakeholders involved.

E. Evaluation

The ADDIE model's final phase is evaluation, which aims to assess the effectiveness of the developed system. This phase determines whether the system achieves its intended objectives and benefits its users [47], [48]. The evaluation was conducted through complete enumeration, using the census method, with all intended system users participating as evaluators. These users included the offices of VSU Alangalang, the Office of the Records and Archives (ORA), Office of the Chancellor (OOC), and Office of the Media Information Systems and Technology (MIST).

To evaluate the system, ISO 9126, known as Software Quality Characteristics, provided a simple, reliable tool for classifying and assessing system quality. The evaluation process utilized a 5-point Likert scale to assess various parameters, such as system functionality, reliability, usability, efficiency, maintenance, and portability. The results of the evaluation were recorded in Table I.

TABLE I. EVALUATION TOOL OF THE PROPOSED ALGORITHM

Limit of Scales	Qualitative Interpretation and Description	Qualitative Interpretation Actual Score/Ideal Score
4.21 – 5.00	Strongly Agree(SA)	81 – 100 (Very Good)
3.41 – 4.20	Agree (A)	61 – 80 (Good)
2.61 – 3.40	Neutral (N)	41 – 60 (Enough)
1.81 – 2.60	Disagree (D)	21 – 40 (Not Good)
1.00 – 1.80	Strongly disagree (SD)	0 – 20 (Not Very Good)

Moreover, the percentage scores are included for each scale, indicating the qualitative interpretation values based on actual scores compared to expected and ideal scores. Then, a formula used for this calculation was:

$$p = \frac{\sum actual\ toal\ sore}{\sum ideal\ score\ x\ 100\%} \quad (4)$$

Where p represents the percentage of the weighted score, indicating the acceptance level with a corresponding qualitative interpretation of the proposed system's overall performance based on the six (6) parameters of ISO 9126.

Additionally, the questions from ISO 9126 were customized to suit the needs of the proposed system and tested for reliability using Cronbach's reliability test with JASP. Table II, presents the coefficient values, along with different levels of reliability interpretation.

TABLE II. VALUES AND ITS EQUIVALENT RELIABILITY LEVEL

Coefficient	Reliability Level
More than 0.90	Excellent (E)
0.80 – 0.89	Good (G)
0.70 – 0.79	Acceptable (A)
0.60 – 0.69	Questionable (Q)
0.50 – 0.59	Poor (P)
Less than 0.59	Unacceptable (U)

IV. RESULTS OF THE STUDY

The researcher discussed the significant results, evidence, and findings on the implementation and deployment of the analysis-compression algorithm to iCTMIS which was conducted along with the end-users of the system developed. In the reduction of the disk space and memory allocation among data and files using LZW compression algorithm, the researcher was able to categorize according to classification of documents being accepted namely, the uploaded files are scanned in pure text-based referred as Category 1 documents and the other is a combination of text-based with an attached images referred as Category 2. With these observations said, the researcher provided results using the paired t-test analysis by allowing Wilcoxon's signed ranked test among measures identified as represented in Tables III to VI.

TABLE III. T-TEST SAMPLES OF MEASURE 1 AND 2 (CATEGORY 1)

Measure 1	Measure 2	W	z	df	p
Compressed File size (in KB format)	Original File size (in KB format)	0.000	-5.511		<.001

Note. For all tests, the alternative hypothesis specifies that Compressed File Size (in KB) is less than Original File Size (in KB).

TABLE IV. DESCRIPTIVE STATISTICS REPRESENTATION (CATEGORY 1)

	N	Mean	SD	SE	Coefficient of Variation
Compressed File Size (in KB	40	411.950 (KB)	421.487	66.643	1.023
Original File Size (in KB)	40	1092.389 (KB)	2044.389	323.246	1.870

TABLE V. T-TEST SAMPLES OF MEASURE 1 AND 2 (CATEGORY 2)

Measure 1	Measure 2	W	z	df	p
Compressed File size (in KB format)	Original File size (in KB format)	0.000	-6.624		<.001

Note. For all tests, the alternative hypothesis specifies that Compressed File Size (in KB) is less than Original File Size (in KB).

TABLE VI. DESCRIPTIVE STATISTICS REPRESENTATION (CATEGORY 1)

	N	Mean	SD	SE	Coefficient of Variation
Compressed File Size (in KB	60	907.317 (KB)	1155.287	149.147	1.273
Original File Size (in KB)	60	1113.250 (KB)	1470.836	189.884	1.321

As results presented in the Tables III to VI, the disk space and memory allocation was reduced, in justification with the results from paired-sample t-test was conducted to compare the files between uncompressed and compressed files by integrating the modified LZW Lossless and Zlib compression algorithms. The file unit of measurement being applied is in Kilobyte (KB). Furthermore, the Wilcoxon signed-rank test in JASP was employed [49] to compare these results from compressed file size to original file sizes. The results revealed in Tables III and IV that there is a significant decrease was observed, for Category 1 ($W = 0.000, z = -5.5111, p < 0.001$), and Category 2 ($W = 0.000, z = -6.624, p < 0.001$). This implies that, on average, the file size decreased after file compression of LZW Lossless and Zlib algorithm was employed. The negative z-scores of the categories ($C1 = -5.511; C2 = -6.624$) indicates that, on compressed file sizes in Category 1 ($M1 = 411.950, SD = 421.487$) are significantly smaller than original file sizes ($M2 = 1092.389, SD = 2044.389$) and Category 2 ($M1 = 907.317, SD = 1155.287$) which is also significantly smaller than the original file sizes ($M2 = 1113.250, SD = 1470.836$). Therefore, the analyses suggest that there is strong evidence that the modified LZW lossless and ZLib file compression algorithm effectively deflates the size of files in comparison to their original counterpart's disk and memory allocation.

In the second objective of the study, eliminating noise and converting files-to-text among document types using the OCR Analysis has proven highly effective as presented in Table VII.

TABLE VII. DOCUMENT DISTORTIONS AND IMAGE ENHANCEMENT

Sample Size (N)	Category of Documents	Mean
40	Category 1: Text-based	0.9485
60	Category 2: Text-images based	0.9526
OVERALL MEAN (N = 100)		**0.9505**

In Table VII, the results reveal a commendable performance across distinct file categories, as exemplified by the following mean scores: In Category 1, encompassing purely text-based files, the OCR algorithm achieved a mean score of 0.9485. Similarly, in Category 2, which involves a combination of image and text files, the algorithm demonstrated robust efficacy with a mean score of 0.9526. These findings underscore the algorithm's

versatility and reliability in converting images to text are highlighted by a mean score greater than 0.50 but less than one (1). This indicates that documents with these scores are suitable for conversion, and the algorithm effectively eliminates noise in the data sets, making it a valuable tool for the document types, and enhancing overall document clarity. Moreover, in eliminating noise among accepted documents by the OCR Tesseract algorithm, the following approaches brightness and color distortion and image enhancement were considered.

Furthermore, in the evaluation of the study, the researcher adopted ISO 9126 or known as Software Quality Characteristics, which focused on the aspects of systems functionality, reliability, usability, efficiency, maintenance, and portability. In the conduct of evaluation, the questions were adopted and modified based on ISO 9126 Software Quality Characteristics Metrics. The questions were categorized into six (6) measures namely functionality, reliability, usability, efficiency, maintenance, and portability.

To also determine the point of scaling for every question, a five-point Likert scale was used for the respondents to avoid confusion on answering and to provide an accurate comparison for every question given in the evaluation.

The results from the evaluation conducted, a formula is presented below that was used for computing the mean of every category of the evaluation. A limit of scale was used as an indicator that helped determine the qualitative descriptions. The researcher used the following formula:

$$\bar{x} = \sum fw/n \tag{5}$$

In computing the mean, where $\bar{x}$ is the computed mean, $\sum fw$ is the sum of all the scores in the set and n is the total numbers of respondents. Additionally, since the researcher used Cronbach's Alpha for consistency, or reliability, of a set of survey evaluations conducted, the researcher used the following formula:

$$\alpha = \frac{N * \bar{c}}{\bar{v} + (N-1) * \bar{c}} \tag{6}$$

where N is the number of items, $\bar{c}$ is the mean of covariance between items and $\bar{v}$ is the mean item variance. As presented, the researcher presents the Table VIII, a tabulated presentation in relation to the conduct of evaluation adapted and modified through ISO 9126 in its six (6) measures with the used statistical analysis measures that the evaluation Cronbach's Alpha validity and reliability test that in the measures of functionality and usability provided a good level of reliability and to the measures in reliability, efficiency, maintenance and portability was concluded acceptable in validity and reliable.

For the aspect of ISO 9126, the overall percentage provided a 91% percent of its overall weighted mean that the figures presented that the usability, efficiency, maintenance, and portability manifested an above 90% operative performance while functionality and reliability manifested an above 85% which still indicates an operative performance during the conduct of evaluation.

TABLE VIII. RESULTS OF THE EVALUATION

Algorithm Evaluation	Mean	Cronbach's Alpha	Weighted Mean (%)	Interpretation
Functionality	4.383	0.839	88%	SA / G
Reliability	4.361	0.715	87%	SA / A
Usability	4.583	0.824	92%	SA / G
Efficiency	4.583	0.707	92%	SA / A
Maintenance	4.625	0.755	93%	SA / A
Portability	4.667	0.764	93%	SA / A
Overall	4.534	0.767	91%	SA / A

V. CONCLUSION

The research aimed to improve the communication tracking management system at Visayas State University Alangalang. This was achieved by the objectives set such as, implementing analysis-compression algorithms, specifically focusing on reducing disk space and memory allocation using LZW Lossless and Zlib compressions, as well as eliminating data noise and converting files to text using OCR analysis. Additionally, the study evaluated the proposed system using a modified version of the ISO 9126 as an evaluation tool, focusing on the six (6) measures functionality, reliability, usability, efficiency, maintenance, and portability. Participants were selected based on specific criteria relevant to the research goals. The research design and methodology employed was developmental research design and the ADDIE Model. Moreover, the instrument utilized is the ISO 9126, which underwent a reliability test to ensure consistent results during the evaluation of the proposed system.

As objectives discussed, this study has achieved and provided significant highlights. Firstly, it successfully reduced disk space and memory allocation among data and files through the implementation of the LZW compression algorithm – Lossless and ZLib. The utilization of these LZW compressions not only significantly decreased data sizes but also adhered to the study's first objective. This reduction in data size is further substantiated by the statistical analysis tools employed, which demonstrate that the achieved compression figures are statistically significant and provide acceptable results.

Secondly, the study effectively eliminated data noise and converted files to text using the OCR analysis algorithm. The OCR analysis played a crucial role in achieving this objective, ensuring that the data is converted to a usable, accurate and reliable output from files into text format. Again, statistical analysis computed mean was used to justify the effectiveness of the OCR algorithm in eliminating noise and converting files, further supporting the study's objectives.

Lastly, the study conducted a comprehensive evaluation of the analysis-compression algorithm against the system requirements based on ISO 9126 Software Quality Characteristics, which encompass functionality, reliability, usability, efficiency, maintenance, and portability. The results of this evaluation revealed that the algorithm performed commendably across all these parameters of ISO 9126. Statistical analysis tools were employed to provide thorough evidence of the algorithm's performance in meeting these quality

characteristics, reinforcing the alignment with the study's third objective.

The integration of algorithms in the analysis-compression frameworks significantly improves the management information system for communication tracking at Visayas State University Alangalang. By the incorporation of algorithms, particularly the modified LZW Lossless and ZLib frameworks, it provided a notable significant decrease and efficient performance in the data set file sizes, thereby optimizing disk space and memory allocation which mainly addresses the challenges that VSU Alangalang in adapting technology innovations that with this development the use of algorithms in analysis-compression showed a significant improvements in the document and communication tracking of information systems.

Furthermore, the application of OCR Tesseract for eliminating noise and converting files to text proves to be highly effective. The algorithm achieves a significant ninety-five (95%) success rate in converting files to text, employing techniques such as brightness and color distortion adjustments, along with image enhancement methods wherein mean score among the data sets showed a greater than 0.50 but less than 1.0, that indicates that documents with these scores proves that are suitable for conversion framework, and the algorithm effectively eliminates noise in the data sets.

In terms of system software quality characteristics, the proposed system exhibits very good remarks with a ninety-one percent (91%) weighted mean score that indicates and proves the quality characteristics of the proposed system in the measures and performance across various dimensions in functionality, reliability, usability, efficiency, efficiency, maintenance, and portability, indicating a high level of quality and effectiveness in its operation.

VI. Recommendations

It is recommended that, the end-users of the system, such as the Visayas State University Alangalang Staff/Faculty and Personnel and Office of the Records and Archives (ORA), should upload only PDF/MPDF and image files containing pure text, excluding images. Additionally, this proposed system it is also recommended that this should be expanded to the other Offices such as the Office of the two (2) Colleges which also struggles with the Information Systems that should replace the existing manual management of documents.

Moreover, the compression framework demonstrated a slight two to three (2-3%) reduction in the size of communication letters that include images. In contrast to pure text documents exhibited a more substantial compression ranging from 50-80% compared to their original sizes. 3) To enhance optimization, it is suggested that there should be an implementing a file categorization system based on criteria such as the number of pages and the presence of images. Additionally, the proposed approach involves classifying the quality of uploaded images based on resolution, aligning with specifications derived from various scanner equipment from the different officers/units for the desired output images.

Regarding the OCR Tesseract framework, a uniform template for communication letters is recommended to enhance recognition of handwritten letters from stakeholders. Furthermore, if the institution decides to upload the system to the cloud, it may pose complexities in installation and configuration, with associated costs for VSUA.

To address the used Cronbach reliability test has provided valuable insights into the software quality characteristics, as per the ISO 9126 standards. The focus of this report is to provide targeted recommendations for enhancing specific measures and sub-characteristics of usability, maintenance, and portability.

Future work could include enhancing software usability through improved documentation, interactive tutorials, streamlined user interfaces, and enhanced visual appeal, optimizing maintenance by adopting modular coding practices, conducting thorough testing, and ensuring system stability, and improve portability by adopting standardized coding practices for adaptability across platforms and providing clear guidelines for seamless system replacement. Lastly, as to the proposed system, since VSU Alangalang is still on gradually moving towards digitization it is also recommended that this system undergoes the Unified Theory of Acceptance and Use of Technology a kind of information systems modeling specifically by using the concept of Performance Expectancy that may help even more in redefining and the level of acceptance it may as the system accepts data from the different Office of the VSUA.

References

[1] N. Nabiha, S. Najah, and R. Sakrani, “Communication Barriers in Work Environment : Understanding Impact and Challenges,” vol. 13, no. 11, pp. 1489–1503, 2023, doi: 10.6007/IJARBSS/v13-i11/19498.

[2] A. Zuhri et al., “Business Process Innovations For Courier Service Sector : Case Study In J & T Dungun,” J. Technol. Oper. Manag., vol. 18, no. 1, pp. 80–88, 2023.

[3] C. J. P. Abuda and R. S. Villafuerte, “Development of an Algorithm-Based Analysis-Compression Integrated Communication Tracking Management Information System (iCTMIS),” 2024 IEEE Open Conf. Electr. Electron. Inf. Sci. eStream 2024 - Proc., 2024, doi: 10.1109/eStream61684.2024.10542580.

[4] S. Song, “Virtual Reality Interactive Method and Device Based on Wireless Communication Tracking,” Wirel. Commun. Mob. Comput., vol. 2021, no. March 2014, 2021, doi: 10.1155/2021/6876974.

[5] S. Sathyavenkateshwaren and S. Malathi, “Humanoid Robot: A Survey on Communication, Tracking and Voice recognition,” Proc. 3rd Int. Conf. Inven. Comput. Technol. ICICT 2018, pp. 555–560, 2019, doi: 10.1109/ICICT43934.2018.9034329.

[6] M. N. Farin, “Acceptability and Usability of Quick Response Code for on Line Document Tracking in a Higher Education Institution in the Philippines,” Int. J. Multidiscip. Res. Anal., vol. 05, no. 01, pp. 211–219, 2022, doi: 10.47191/ijmra/v5-i1-26.

[7] M. Zhang, W. Li, Z. Wang, B. Li, and X. Ran, “A RFID-based material tracking information system,” Proc. IEEE Int. Conf. Autom. Logist. ICAL 2007, pp. 2922–2926, 2019, doi: 10.1109/ICAL.2007.4339081.

[8] K. X. Tee, M. T. Chew, and S. Demidenko, “An intelligent warehouse stock management and tracking system based on silicon identification technology and 1-wire network communication,” Proc. - 2011 6th IEEE Int. Symp. Electron. Des. Test Appl. DELTA 2011, vol. 02, no. 13, pp. 110–115, 2020, doi: 10.1109/DELTA.2011.62.

[9] R. G. Luciano, G. M. Alcantara, and R. Bauat, “Design and Development of Alumni Tracking System for Public and Private HEIs,” Int. J. Sci. Technol. Res., vol. 9, no. 06, pp. 12–19, 2020.

[10] C. X. Wang and F. O. Mormah, "Digital Technology for a Borderless World: Innovative Educators in Practice," TechTrends, vol. 67, no. 3, pp. 475–476, 2023, doi: 10.1007/s11528-023-00854-w.

[11] S. Srivastava, A. Verma, and S. Sharma, "Optical Character Recognition Techniques: A Review," in 2022 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), 2022, pp. 1–6. doi: 10.1109/SCEECS54111.2022.9740911.

[12] S. G. Mohammed, S. S. Abdul-Jabbar, and F. G. Mohammed, "Art Image Compression Based on Lossless LZW Hashing Ciphering Algorithm," J. Phys. Conf. Ser., vol. 2114, no. 1, 2021, doi: 10.1088/1742-6596/2114/1/012080.

[13] W. Zhu, N. Sokhandan, G. Yang, S. Martin, and S. Sathyanarayana, "DocBed: A Multi-Stage OCR Solution for Documents with Complex Layouts," Proc. 36th AAAI Conf. Artif. Intell. AAAI 2022, vol. 36, pp. 12643–12649, 2022, doi: 10.1609/aaai.v36i11.21539.

[14] J. B. Tumasis, "Web based management information system with optical character recognition technology for a phlippine accounting firm," South Asian J. Eng. Technol., 2022.

[15] Mande Shen and Hansheng Lei, "Improving OCR Performance with Background Image Elimination," 2019 12th Int. Conf. Fuzzy Syst. Knowl. Discov. FSKD 2015, pp. 1566–1570, 2019.

[16] Y.-Z. Zhang, C.-A. Chen, J.-S. Zhang, and J.-W. Wang, "VLSI Design of Near-Lossless Image Compression using Improved LZW," in 2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2023, pp. 887–891. doi: 10.1109/APSIPAASC58517.2023.10317200.

[17] Y. Fedkin et al., "Development and evaluation of the effectiveness of the integration gateway for the interaction of the learning management system with external systems and services of state information systems," Eastern-European J. Enterp. Technol., vol. 3, no. 2 (117), pp. 30–38, 2022, doi: 10.15587/1729-4061.2022.258089.

[18] H. Tolle, T. S. Putri, and I. Aknuranda, "Information Management and Information System Analysis to Support the Achievement of University Performance Agreements with the Government," J. Sist. Inf., vol. 17, no. 1, pp. 30–43, 2021, doi: 10.21609/jsi.v17i1.989.

[19] S. I. Lagas and J. D. Isip, "Challenges to Digital Services in Philippine Academic Libraries," vol. 43, no. 1, pp. 27–38, 2023.

[20] P. R. Vessels, I. A. Journal, T. Innovations, and V. No, "A Centralized Document Processing and Support System (CeDoPSS) for Innovatus : A Journal on Computing Technology Innovations Innovatus : Special Issue on Digital Transformation," no. 5, pp. 1–3, 2022.

[21] R. Abu Khurma, I. Aljarah, A. Sharieh, M. Abd Elaziz, R. Damaševičius, and T. Krilavičius, "A Review of the Modification Strategies of the Nature Inspired Algorithms for Feature Selection Problem," Mathematics, vol. 10, no. 3, 2022, doi: 10.3390/math10030464.

[22] M. B. Batan, J. K. D. Treceñe, J. R. N. Delos Santos, and R. R. Paler, "Assessment of Competencies in Technology Operation and Concepts among Teachers in a Philippine State University," Eur. J. Educ. Pedagog., vol. 3, no. 3, pp. 306–309, 2022, doi: 10.24018/ejedu.2022.3.3.389.

[23] H. De Bruijn, M. Warnier, and M. Janssen, "The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making," Gov. Inf. Q., vol. 39, no. 2, p. 101666, 2022, doi: https://doi.org/10.1016/j.giq.2021.101666.

[24] A. Klarin and Y. Suseno, "An Integrative Literature Review of Social Entrepreneurship Research: Mapping the Literature and Future Research Directions," Bus. Soc., vol. 62, no. 3, pp. 565–611, 2023, doi: 10.1177/00076503221101611.

[25] J. Memon, M. Sami, R. A. Khan, and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," IEEE Access, vol. 8, pp. 142642–142668, 2020, doi: 10.1109/ACCESS.2020.3012542.

[26] N. Sahu and M. Sonkusare, "A Study on Optical Character," vol. 4, no. 1, pp. 1–14, 2019, doi: 10.5121/ijcsitce.2019.4101.

[27] D. Karthikeyan, V. P. Arumbu, K. Surendhirababu, K. Selvakumar, and P. Divya, "Sophisticated and modernized library running system with OCR algorithm using IoT," vol. 24, no. 3, pp. 1680–1691, 2021, doi: 10.11591/ijeecs.v24.i3.pp1680-1691.

[28] R. Arief, A. B. Mutiara, T. M. Kusuma, and Hustinawaty, "Automated hierarchical classification of scanned documents using convolutional neural network and regular expression," Int. J. Electr. Comput. Eng., vol. 12, no. 1, pp. 1018–1029, 2022, doi: 10.11591/ijece.v12i1.pp1018-1029.

[29] P. H. Jain, V. Kumar, J. Samuel, S. Singh, A. Mannepalli, and R. Anderson, "Artificially Intelligent Readers: An Adaptive Framework for Original Handwritten Numerical Digits Recognition with OCR Methods," Inf., vol. 14, no. 6, 2023, doi: 10.3390/info14060305.

[30] J. Bisiach and M. Zabkar, "Evaluating Methods for Optical Character Recognition on a Mobile Platform: comparing standard computer vision techniques with deep learning in the context of scanning prescription medince labels," p. 0, 2020.

[31] R. Awati, "What is LZW compression and how does it work? – TechTarget Definition."

[32] S. Shah, "A New Approach to Increase Visual Performance," CRST Eur., vol. 10, no. 10, pp. 7–9, 2019.

[33] R. Maulunida and A. Solichin, "Optimization of LZW Compression Algorithm With Modification of Dictionary Formation," IJCCS (Indonesian J. Comput. Cybern. Syst., vol. 12, no. 1, p. 73, 2020, doi: 10.22146/ijccs.28707.

[34] M. Aldwairi, A. Y. Hamzah, and M. Jarrah, "MultiPLZW: A novel multiple pattern matching search in LZW-compressed data," Comput. Commun., vol. 145, no. June, pp. 126–136, 2019, doi: 10.1016/j.comcom.2019.06.011.

[35] M. Safieh and J. Freudenberger, "Address space partitioning for the parallel dictionary LZW data compression algorithm," 2019 16th Can. Work. Inf. Theory, CWIT 2019, pp. 0–5, 2019, doi: 10.1109/CWIT.2019.8929928.

[36] H. Yang, G. Qin, and Y. Hu, "Compression Performance Analysis of Different File Formats," vol. 1, 2023.

[37] K. S. Chirikhin and B. Y. Ryabko, "Application of data compression techniques to time series forecasting," pp. 2–6, 2019, [Online]. Available: http://arxiv.org/abs/1904.03825

[38] F. S. Taruc, T. A. S. Martin, C. N. P. Olipas, and R. T. Alegado, "Docu-Go: The Development and Assessment of a Web-Based Barangay Document Requesting System," Int. J. Inf. Technol. Comput. Eng., no. 34, pp. 40–49, 2023, doi: 10.55529/ijitc.34.40.49.

[39] T. Prihandini, "Interactive Mobile Technologies," Int. J. Interact. Mob. Technol., vol. 17, no. 15, pp. 135–154, 2023.

[40] E. Y. Oh and D. Song, "Developmental research on an interactive application for language speaking practice using speech recognition technology," Educ. Technol. Res. Dev., vol. 69, no. 2, pp. 861–884, 2021, doi: 10.1007/s11423-020-09910-1.

[41] R. Rizal, D. Rusdiana, W. Setiawan, and P. Siahaan, "Development of a problem-based learning management system-supported smartphone (PBLMS3) application using the ADDIE model to improve digital literacy," Int. J. Learn. Teach. Educ. Res., vol. 20, no. 11, pp. 115–131, 2021, doi: 10.26803/ijlter.20.11.7.

[42] J. P. Guevarra, A. M. Ongkeko, C. A. T. Antonio, A. N. C. Bermudez, and P. H. Fernandez Marcelo, "The Application of the ADDIE Model and the Training Cycle in the Development, Implementation and Evaluation of Training Program on Data Use for Decision-making among End-users of Electronic Health Information System in Geographically Isolated and Disadvan," Acta Med. Philipp., vol. 55, no. 4, pp. 398–405, 2021.

[43] C. M. Budoya, M. M. Kissaka, and J. S. Mtebe, "Instructional Design Enabled Agile Method Using ADDIE Model and Feature Driven Development Method," Int. J. Educ. Dev. Using Inf. Commun. Technol., vol. 15, p. 35, 2019.

[44] "ADDIE: 5 Steps To Effective Training Courses | LearnUpon," LearnUpon Blog, 2023.

[45] A. J. L. Piloton, M. Dela Cerna, and R. Reyna, "Development of Records Tracking Management System with QR Code," Int. J. Multidiscip. Res., vol. 5, no. 4, pp. 0–14, 2023, doi: 10.36948/ijfmr.2023.v05i04.5508.

[46] M. A. Stapa and N. Mohammad, "The Use of Addie Model for Designing Blended Learning Application at Vocational Colleges in Malaysia," Asia-Pacific J. Inf. Technol. Multimed., vol. 08, pp. 49–62, 2019, doi: 10.17576/apjitm-2019-0801-05.

[47] D. P. D. O. Godeiro, M. L. R. Dantas, M. D. S. Celestino, and D. C. Da Silva, "Application of Importance and Performance Matrix to Assess the Quality of Services Provided by Business Incubators," REGEPE - Rev. Empreendedorismo e Gestão Pequenas Empres., vol. 7, no. 3, pp. 01–29, 2018, doi: 10.14211/regepe.v7i3.704.

[48] H. T. Amijaya, R. Ramlan, and G. N. Fajar, "Quality Measurement of Accurate-5 Accounting Software Using the Iso 9126 Model," J. Comput. Bisnis, vol. 17, no. 1, pp. 32–41, 2023, doi: 10.56447/jcb.v17i1.181.

[49] M. Kapse, B. Akhil, N. Elangovan, V. Sharma, and K. Rajagopal, "A Comparative Study of Pollution Levels in Major Cities of India During Covid-19 in India," Australas. Accounting, Bus. Financ. J., vol. 17, no. 1, pp. 247–255, 2023, doi: 10.14453/aabfj.v17i1.16.