# Text Recognition for Library Collection in Different Light Conditions

Imam Yuadi [1], A. Robert Sigh [2], Ullin Nihaya [3]

[1] *Department of Information and Library Science, Airlangga University, Surabaya, Indonesia*
[2] *Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur , Tamil Nadu, India*
[3] *Information Management, Tatung University, Taipei, Taiwan*

***Abstract*** **– Book arrangement problems in Indonesia mainly utilize traditional systems. However, it contributes to user frustration and dissatisfaction, leading to difficulties in the retrieval process and hindering book monitoring. Therefore, an efficient arrangement is needed to make it easier for librarians to organize collections at the library. Hence, this study proposes and evaluates Optical Character Recognition (OCR) approaches such as Tesseract-OCR and Efficient and Accurate Scene Text (EAST) by considering various light factors and device applications. An evaluation was performed to determine the effectiveness of the methods for detecting classification numbers using the Character Error Rate (CER) and Word Error Rate (WER). The results obtained can be seen from the lighting and the use of image capture tools that can affect the text detection results. The higher the lighting, the better the text detection results obtained. However, in the extraction tool using a security camera yields a better output than the others. In conclusion, the modernization of book arrangement systems by implementing OCR through security cameras could enhance user experience in the retrieval process.**

***Keywords*** **– Book arrangement, optical character recognition, Tesseract, efficient and accurate scene text, character error rate, word error rate.**

DOI: 10.18421/TEM131-28
https://doi.org/10.18421/TEM131-28

**Corresponding author:** Imam Yuadi,
*Department of Library and Information Science, Airlangga University, Surabaya, Indonesia*
**Email:** imam.yuadi@fisip.unair.ac.id

*Received: 06 September 2023.*
*Revised: 11 December 2023.*
*Accepted: 23 December 2023.*
*Published: 27 February 2024.*



## 1. Introduction

Computer vision is implemented to provide co-computers with the ability to see or process a work environment [1]. Its complexity is inspired by various fields of knowledge such as psychology, neuroscience, physics, robotics, and statistics, computer vision applies detection, classification, and segmentation to accuracies developed to approach human capabilities. The utilization of computer vision is carried out because of its efficiency and effectiveness. Hence, computer vision is often applied for detection, classification, and implementation due to its efficiency and effectiveness when utilized in common problems [2].

Additionally, it could be implemented in the library to avoid book arrangement problems, which frequently occur in Indonesian libraries. For instance, the SMKN 1 Manado library still addresses this problem owing to the lack of classification numbers in their collections. Consequently, it would affect the retrieval process by users and discourage them from coming or using services in the library [3]. Therefore, the arrangement of books plays a very important role in the library ecosystem [4] because it has a functional and linear relationship with user satisfaction [5].

Therefore, some studies have applied the greedy algorithm method, although still lacking because of the value of the book's dependency on the selection [6]. Other studies implemented Urgency, Seriousness Growth (USG), and paired-sample t-test analysis with results that did not significantly affect the provisions for structuring archives [7]. On the other hand, the common book arrangement method that is usually applied in the library or reading room [4], [8] requires a long time and the librarian's accuracy in determining collections. However, the problem faced by libraries in Indonesia is the lack of resources and the cost of managing books.

However, computer vision is needed to overcome these problems.

Optical character recognition, which is in the implementation of computer vision in various studies, has been applied in several fields, such as real-time identification of license plate numbers [9], [10] and recognition text on traffic signs [11], [12]. In several studies, Tesseract and EAST are the most common algorithms. Tesseract has been applied in a great number of studies, such as character recognition of Javanese text [13], digital documents [14], and text captured with a digital camera [15]. The implementation of the EAST algorithm has been used in some studies because of its good performance in natural landscape issues [16], [17]. EAST also has a simple network structure equipped with increased speed and accuracy [18]. Therefore, in this study, the CER and WER were applied to evaluate Tesseract and EAST. CER and WER were implemented as indicators to determine the accuracy of the OCR. Consequently, WER has a higher level of understanding accuracy than the trigram model [19]. Meanwhile, CER and WER have been applied to detect errors and accuracy in children's writing [20]. However, in its implementation, it is often utilized in speech recognition because its use can detect the accuracy of many sentences [21], [22].

Essentially, Tesseract and EAST are used to detect classification numbering in books under various lighting conditions, particularly low lighting. In general, the library has natural and minimal lighting to reduce damage to collections [23]. This is done to reduce the time and cost needed to organize the collection. Hence, the effectiveness of the OCR method for detecting classification numbering is expected to simplify the collection arrangement process because of the importance of structuring the books in the library.

In addition, this article will be divided into several sections, such as Section 2, which contains related work from previous research. Furthermore, Section 3 will present the methods of OCR used in this research. Moreover, section 4 which contains a presentation of the results and a discussion of previous research. Section 5 presents the conclusion of the results.

## 2. Related Work

A library is an institution that provides a collection of written, printed, and recorded library materials as a central source of information that is regulated by a system for community information needs [24]. In the existing system in the library, to meet information and discovery needs, it is necessary to organize collections to make it easier for users to find the information they need. Structuring collections (shelving) is an activity in the library in the form of arranging collections on shelves based on certain provisions [25].

The provisions of the arrangement are applied based on relative or permanent placement and are grouped by class. The classes contained in the book are grouped using classification, which is a systematic grouping of several objects or ideas in a particular class or class [26]. The goal is for the information retrieval system process to be easier for both users and librarians. Information retrieval is the science of searching for information in documents, searching for documents, searching for metadata that describes documents, or searching in databases [27].

However, book arrangements cannot be said to be efficient and well-implemented in libraries in Indonesia. This is because the process is time-consuming and lacks resources in most libraries in Indonesia. Computer vision has been used to overcome this problem. Computer vision is an artificial ability to see and change real objects into digital formats [28]. The digital format is displayed in a visual form to help humans work. One of the fields of computer vision that is often used and relevant in the process of structuring books is optical character recognition [29].

Optical character recognition is a technology that digitizes real objects into digital form. Technological developments are rapid in the current digital era. Their use, which is not limited to any object, makes them a popular part of computer vision. OCR is often discussed in research on the detection of real objects in the form of objects or text, as is the case with automatic detection of license plate numbers [30], [31], object detection in image documents [32], handwriting detection [33], identification of ancient manuscripts [34] and many more.

Nevertheless, the discussion about OCR will not be separated from Tesseract OCR. In several studies, Tesseract was used to extract text characters from vehicle registration [35], recognise text with Javanese characters with satisfactory results [13], and use 10 photos of digital documents with the best results on a distance of 30 cm in outdoor conditions [14] or detecting text in images captured by digital cameras on cellphones, where the greater the image quality, the higher the accuracy [15].

However, even so, Tesseract OCR is still not enough to help librarians organize collections on shelves because the library environment is generally made using natural lighting and minimal damage to library materials. Several studies have used text restoration methods to detect text under minimal lighting [36]. Another study used MSER and CNN to detect text under dark lighting [37]. However, several studies have used an efficient and accurate scene text detector for text detection in natural lighting. EAST is a text-detection model that has good speed and performance [17].

This model is used in traffic sign detection [38], web text image detection [16], and natural scenery [18].

## 3. Research Method

This study employs a thorough and precise technique to examine the efficacy of a proposed methodology by evaluating Tesseract and EAST to identify and categorise labels on books. The main goal is to understand the distinct characteristics and labelling scheme by examining a wide range of lighting scenarios and imaging apparatus in detail.

### *3.1. Data Source*

To reach relevant results, careful experimental procedures must be used to collect data. Therefore, a methodical data collection using books and varied lighting settings was implemented. The selection of books included a variety of classification names and numbers; this was done on purpose to test the effectiveness of the suggested strategy for identifying various numerical categorizations. In order to do this, the books were exposed to controlled lighting treatments of varied intensities: 450 lumens, 900 lumens, 1350 lumens, 1800 lumens, and a situation in which there was no natural lighting. By carefully adjusting the lighting, it was possible to thoroughly investigate the method's ability to distinguish between varieties of brightness circumstances. As shown in (Fig. 1), some imaging devices such as a phone camera, a security camera (CCTV), and a webcam, were used to collect data. Every imaging equipment was carefully positioned at a standard distance of 35 cm from the experimental setup. This rigorous planning not only made it possible to collect several datasets, but also made it easier to conduct an empirical evaluation and choose the best imaging tool for the experimental situation. Nevertheless, the study aims to illuminate the complex interplay between different lighting conditions, imaging equipment, and their combined effects on the efficacy of the suggested methodology.

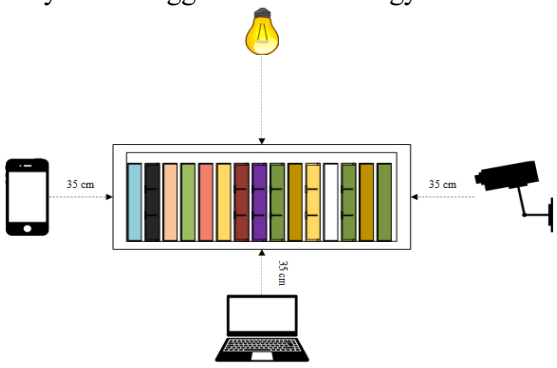



*Figure 1. Lighting and device placement for the experiment*

### *3.2. Methods*

This study's main goal is to identify and carefully examine the classification numbering imprinted on the book's back covers to clarify its unique features and arrangement. In addition to this main purpose, this research project's concurrent objective is to evaluate the effectiveness of optical character recognition (OCR) in the precise identification and extraction of this categorization numbering. Tesseract and EAST are two different yet complementary algorithmic frameworks that must be used in the investigation process to recognize categorization numbering. These mathematical models are cleverly used to challenge text identification within photos and are neatly incorporated into the Python programming environment.

Tesseract and EAST were purposefully chosen as the study's primary algorithms because of their well-known prowess in textual analysis and detection. Tesseract, which is renowned for its prowess in transforming text-containing images into machine-readable text, fits in perfectly with the study's goal of decoding the classification numbers seen on book covers. In addition, the EAST algorithm, known for its accuracy in locating and detecting text inside natural surroundings, adds a crucial component to the research technique. The anticipated discrepancy in results produced by the application of Tesseract and EAST algorithms is essential to this investigation. It is possible to identify discrepancies in their performance features by a careful comparison analysis. The effects of these variations go beyond the scope of algorithmic preferences, providing subtle insights into the advantages and disadvantages inherent in each strategy. Figure 2 serves as the complex interplay between algorithms and picture data by providing a visual explanation.

*Figure 2. Output geometry design*

The primary goal of this study is to evaluate the Tesseract OCR and EAST techniques' accuracy in error detection during the detection stage. The success of optical character recognition and text detection systems depends on this stage. Character Error Rate (CER) and Word Error Rate (WER) are two distinct metrics that are used in the review process. WER evaluates errors at the word level whereas CER assesses faults at the character level.

Precision calculations are made possible by the field-established formulas for CER and WER [39], [40]. Combining the two measurements provides a thorough insight of Tesseract OCR and EAST performance, catching minute nuances and having real-world applications. In general, this study improves our comprehension of mistake patterns and advances text detection technology.

$$CER = \frac{S + D + I}{N}$$

The total CER computation is derived by calculating the number of characters with nearly identical text detection results (S), the number of characters that disappear or are missing (D), and the number of insertion characters (I). The total is then divided by the number of characters in the reference text (N). Principally, the lower its value the better the performance. The indicators can be seen in the table below:

*Table 1. CER and WER accuracy indicator*

| Accuracy | CER & WER |
|---|---|
| Good | 1-2% (98-99% accurate) |
| Average | 2-10% (90-98% accurate) |
| Poor | >10% (below 90% accurate) |

On the other hand, WER formulas are nearly identical to CER formulas even though the focus is more on word calculation. However, WER is certain to have a greater value than CER due to the calculation evaluation errors in the detected words. The calculation can be done using the formula below.

$$WER = \frac{S_w + D_w + I_w}{N_w}$$

Therefore, the total of the WER calculation is obtained by calculating how many words have comparable words in the output ($S_w$), how many words are absent ($D_w$), and how many words are inserted ($I_w$). This is then divided by the number of words in the ground truth ($N_w$).

### *3.3. Framework Approaches*

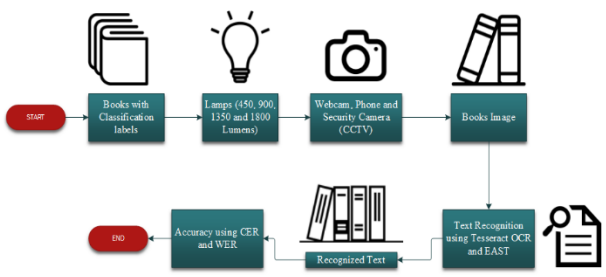



*Figure 3. Research framework*

The Tesseract OCR and EAST methods are used to detect categorization labels attached to books. Figure 3 shows the sequential steps in the detection process. The flowchart shows a thorough process that includes seven distinct steps and culminates with pattern detection. The use of various books, each with a distinct classification number on the spine, is required for the first stage in order to replicate real-world events. Notably, the experiment uses lamps calibrated at 450, 900, 1350, and 1800 lumens that have different luminosity levels, allowing for controlled adjustments in illumination settings. To provide a variety of lighting conditions during the experiment, these lamps are placed carefully atop the book rack. This intentional adjustment of lighting conditions demonstrates thoughtful planning, mimics the diversity found in real-world settings, and strengthens the reliability and relevance of the study's findings.

## 4. Result and Discussion

In this study, several thorough tests were carried out to assess the outcomes of text identification on categorization label images located on book spines in order to determine the efficiency of the suggested method. The introduction of the experimental technique is followed by a display of the text detection results, and the section concludes with a discussion during the last session. Tesseract and EAST, two well-known approaches, were used to picture objects in the form of categorization labels found on the spines of books to evaluate text detection. Each dataset was presented as part of the review process and put through two different experimental scenarios, which are described below.

Scenario-I (SI): given a different behavior, namely by providing different lighting to the book.

Scenario-II (S-II): using different tools to capture the data for providing the best tools to use for detection.

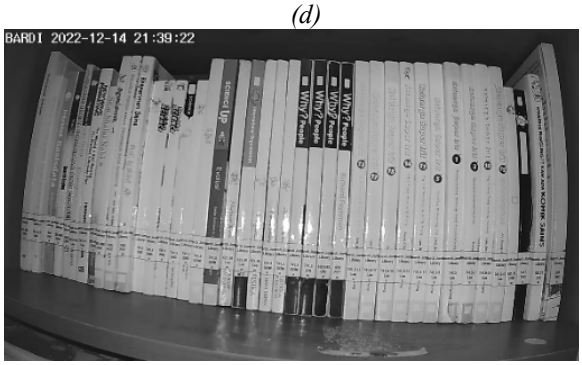
The outcome of the scenario used to create the picture of the bookshelf, which just focuses on various lighting, does not seem to change anything (Fig.4). However, the results of the image without light are different from those of the other images. In essence, a security camera, which includes infrared to enable it to capture items in the dark, can show the books or categorization label in the image whereas a phone or webcam cannot in a situation without lighting. These findings emphasize the critical role that lighting circumstances play in the suggested text identification method's efficacy and emphasize the need to take into account the unique environment and equipment when designing a system for classifying labels on book spines.


*(a)*


*(b)*


*(c)*


*(d)*




*(e)*

*Figure 4. Experiment with various Lighting; (a) 450 Lumens, (b) 900 Lumens, (c) 1350 Lumens, (d)1800 Lumens, and (e) 0 Lumens/ Without Lighting*

On the other hand, in Figure 5 we can see that different camera types, notably phone cameras, security cameras (CCTV), and webcams, produce noticeably different results. These discrepancies are visible not only in the text detection but also in the image quality and noise levels. It is noteworthy that the results from security cameras (CCTV) had better image quality than those from phone cameras and webcams, indicating the potential benefits of using specialized surveillance equipment for this application. In contrast, webcams produced the least accurate text detection and poorest image quality outcomes. This discrepancy highlights how the type of camera might affect how well the suggested method works and highlights how important it is to take imaging equipment into account when designing text detection systems for categorization label identification on book spines.


*(a)*

*(b)*

*(c)*

*Figure 5. Experiment based on devices; (a) Security Camera (CCTV), (b) Phone Camera, and (c) Webcam*

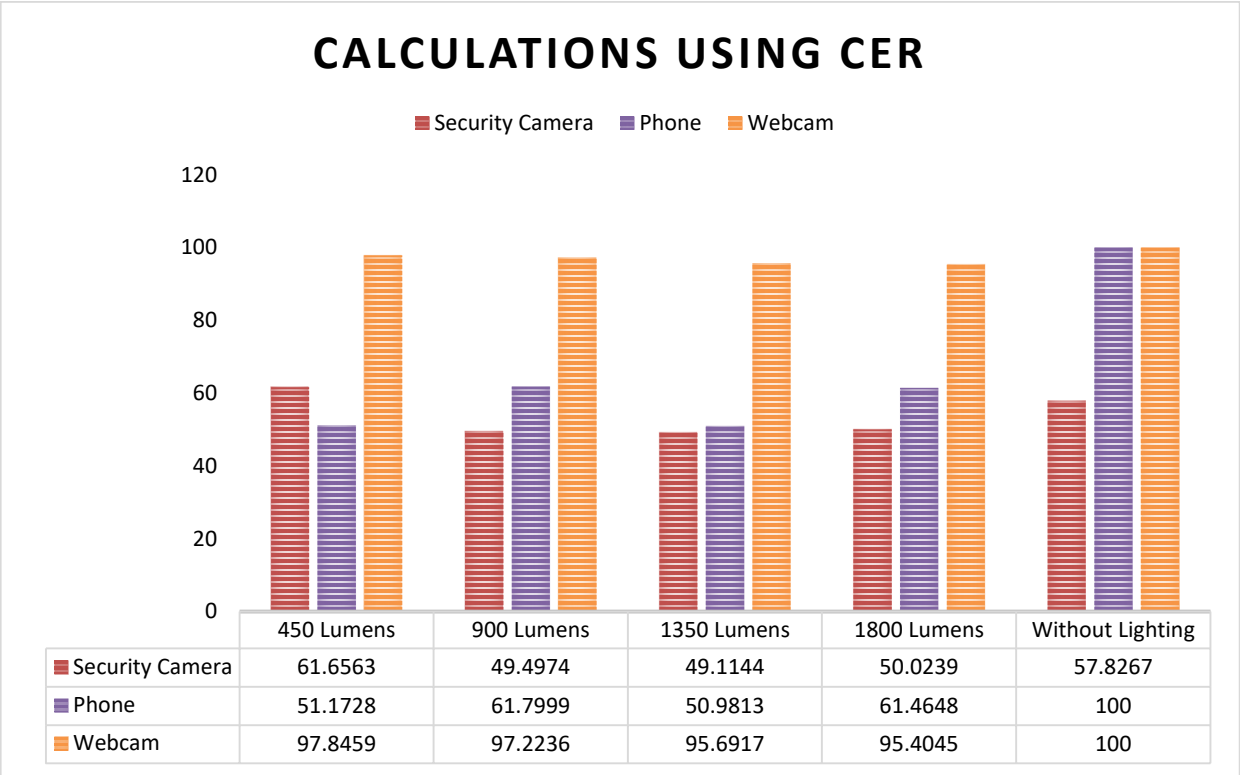
Furthermore, a complex assessment is required to properly evaluate the two scenarios that are applied to the developed method and the objects to which it is applied, respectively. Performance is measured using Character Error Rate (CER) and Word Error Rate (WER), in which the results of text identification are compared to text references, also known as ground truth. Ground truth in this study refers to text that has been taken from book images and classification labels. It is significant to note that the wording in the reference materials is not perfect, primarily because of folds or obstructions that hide some of the text on the classification labels. We took a practical approach in light of these difficulties, concentrating mostly on the classification labels themselves. As a result, when we evaluated the effectiveness of the procedure after the trial, we solely used the categorization numbers as the reference text. This specialized technique of defining ground truth takes into consideration the inherent difficulties in text extraction from categorization labels with information that is obscured or only partially visible, enabling a more accurate evaluation of the method's effectiveness in real-world applications.

The crucial significance that ground truth plays in determining the accuracy of the approach must be emphasized in the context of using character error rrate (CER) and word error rate (WER) as performance measures. The importance of ground truth is made clear by the fact that it is essential to the computation of metrics like precision, recall, and F-measure, all of which depend on comparisons to the ground truth references. The goal of this study is to provide more light on the intricacies of text-based mistake detection (Fig. 6). The graphical depiction makes it clear that while some texts can be detected textually accurately, the localization aspect is still not ideal, which therefore lowers the precision rating.


*Figure 6. Text detection results*

The resulting findings are systematically computed in accordance with the following formula for assessing the effectiveness of text detection systems. As a result, we have combined the findings of text detection evaluations carried out using both Tesseract and EAST data, with an emphasis on Character Error Rate (CER) and Word Error Rate (WER) as the main evaluation metrics. The CER computation for text detection on book photos produced results that could be classified as "bad," as seen in (Fig. 7). In particular, the CER values were higher than 10, which showed that there were considerable text detection errors. Upon closer inspection, it becomes clear that the ambient lighting conditions in the room have a significant impact on the text detection ability. Notably, better detection outcomes were seen in settings with a light level of 1350 lumens. On the other hand, results with 450 lumens of lighting and no ambient light were shown to produce the worst results. However, it is noteworthy that the security camera's infrared capabilities showed improved text identification performance even in circumstances with only 450 lumens of lighting, outperforming that of typical lighting conditions. This emphasizes the additional advantage of using infrared technology to achieve reliable text identification, especially in low-light situations.

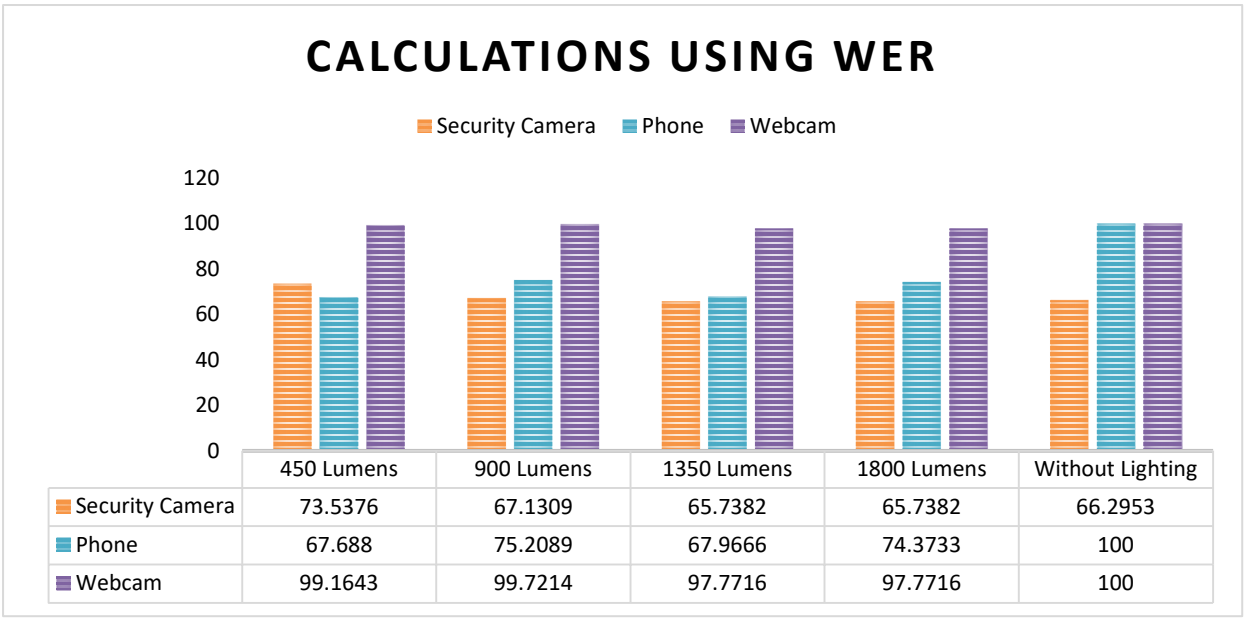
In contrast, results from the examination of text detection on book photographs using the word error rate (WER) metric show an even more difficult situation than those found with the character error rate (CER) measure (Fig. 8). In this situation, the WER computations produce results that are between 60 and 100, highlighting a substantially higher level of text identification error. A comprehensive analysis of the WER calculation results reveals that the effectiveness of the room's illumination has a significant impact on the results. Once more, text detection performs best in lighting conditions with a luminance of 1350 lumens. The worst outcomes, on the other hand, are constantly seen in situations with 900 lumens of brightness and when there is no ambient light present, which is especially noticeable when utilizing phone cameras and webcams. It is interesting to note that the security camera, utilizing its infrared capabilities, maintains superior text detection ability even when compared to scenarios with 450 lumens and 900 lumens of brightness, reiterating its robustness in low-light situations. These results highlight the crucial interaction between lighting and text detection precision and highlight the benefits of specialist tools like security cameras for this particular application.




| | 450 Lumens | 900 Lumens | 1350 Lumens | 1800 Lumens | Without Lighting |
|---|---|---|---|---|---|
| Security Camera | 61.6563 | 49.4974 | 49.1144 | 50.0239 | 57.8267 |
| Phone | 51.1728 | 61.7999 | 50.9813 | 61.4648 | 100 |
| Webcam | 97.8459 | 97.2236 | 95.6917 | 95.4045 | 100 |

*Figure 7. The results of calculations using CER*

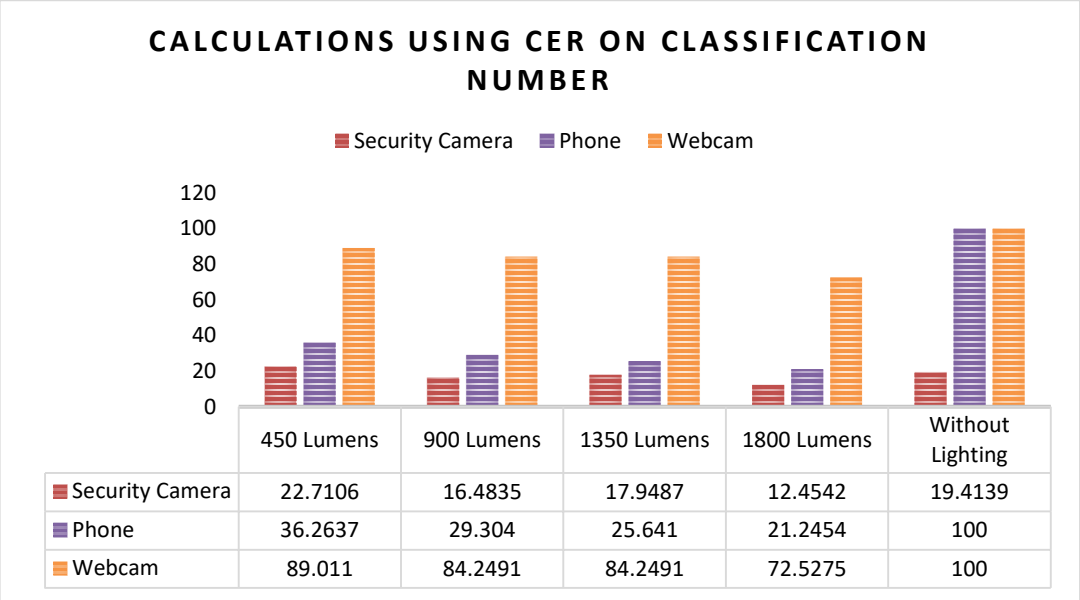
An alternative strategy is followed by restricting the focus of text detection solely to the classification numbering portion that contains book labeling information because the acquired results are still less than ideal. This change departs from the previous ground truth definition and necessitates a commensurate adjustment in the ground truth implementation. This more sophisticated method is intended to solve the articular difficulties presented by the categorization numbering section, which could have unique qualities and complexity in comparison to the rest of the text on the book spine. We aim to improve the precision of text detection in this crucial area of the classification label byadjusting the ground truth to be in line with this more specific purpose.




| | 450 Lumens | 900 Lumens | 1350 Lumens | 1800 Lumens | Without Lighting |
|---|---|---|---|---|---|
| Security Camera | 73.5376 | 67.1309 | 65.7382 | 65.7382 | 66.2953 |
| Phone | 67.688 | 75.2089 | 67.9666 | 74.3733 | 100 |
| Webcam | 99.1643 | 99.7214 | 97.7716 | 97.7716 | 100 |

*Figure 4. The results of calculations using WER*

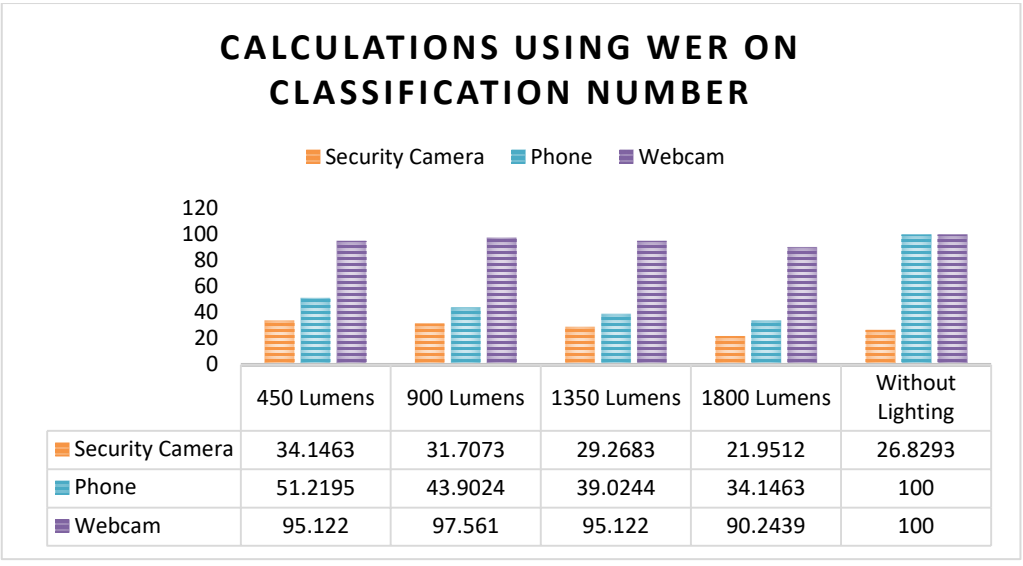
The results have improved noticeably by focusing only on the classification numbering segment however they have not yet attained a level of complete satisfaction. Meanwhile, these improved results do show a noticeable improvement, with performance indicators often falling between 10 and 100. Character error rate (CER) calculations (Fig. 9) make it clear that the best outcomes were consistently seen in settings with 1800 lumens of lightin. The least desirable results, on the other hand, were consistently connected with conditions described by 450 lumens of lighting and circumstances where no ambient lighting was used, which was especially apparent when utilizing phone cameras and webcams. These findings show that lighting conditions are crucial for improving text identification accuracy even when the classification numbering section is the main emphasis.




| | 450 Lumens | 900 Lumens | 1350 Lumens | 1800 Lumens | Without Lighting |
|---|---|---|---|---|---|
| Security Camera | 22.7106 | 16.4835 | 17.9487 | 12.4542 | 19.4139 |
| Phone | 36.2637 | 29.304 | 25.641 | 21.2454 | 100 |
| Webcam | 89.011 | 84.2491 | 84.2491 | 72.5275 | 100 |

*Figure 5. The results of calculations using CER*

The optimaland suboptimal detection scenarios nevertheless correspond with those found in the CER calculations when using the word error rate (WER) computation (Fig. 10). Overall, taking into account text detection in the categorization numbering part, the best results are invariably connected to illumination conditions of 1800 lumens. This finding emphasizes the connection between better text detection outcomes and higher illumination levels. It is important to note that using a security camera consistently produces better results, regardless of the lighting situation, especially when taking into account the convex camera type. In contrast, webcam performance constantly falls short of that of other imaging technologies. As a result of their subpar performance in contrast to other equipment options, it is advised against using webcams for textbook detection in a variety of lighting situations.




| | 450 Lumens | 900 Lumens | 1350 Lumens | 1800 Lumens | Without Lighting |
|---|---|---|---|---|---|
| Security Camera | 34.1463 | 31.7073 | 29.2683 | 21.9512 | 26.8293 |
| Phone | 51.2195 | 43.9024 | 39.0244 | 34.1463 | 100 |
| Webcam | 95.122 | 97.561 | 95.122 | 90.2439 | 100 |

*Figure 6. The results of calculations using WER*

Based on the experiments and analyses that have been implemented, the results show that the approach using computer vision text detection can be applied in detecting classification numbering in the book. The implementation of Tesseract and EAST as a method that is considered better than using only Tesseract [17] in minimum lighting [38] is expected to bring satisfactory results. However, the results obtained in the text detection on the book and the numbering of the classification are still not at a satisfactory stage.

Nevertheless, the findings can be observed by analyzing the illumination and selecting the appropriate image capture tool, as these factors could impact the text detection outcomes. In image detection of the books, 1350 lumens lighting has the best detection results compared to the others. Meanwhile, the 1800 lumens lighting has the best impact on detecting classification numbers compared to the others. This means the brighter the lighting, the better the detection results obtained. On the other hand, the use of a security camera has better results than a webcam. Based on these results, the use of Tesseract and EAST text detection provides quite good results when compared to methods without computer vision. Wibisono [6] uses the greedy algorithm method but in its application, it still has drawbacks because the selection function is affected by the value of the books being managed. Meanwhile, Herawan [7] carried out the Urgency, Serousness, Growth (USG), and Paired-Samples T-Test analysis methods for structuring inactive archives with results that did not significantly affect the provisions for structuring archives. The conventional or traditional practice of organizing books in libraries often consumes a significant amount of time [4], [8].

## 5. Conclusion

In this study, the use of computer vision, namely text detection was implemented to deal with book structuring problems, which so far have consumed more resources and time. In its implementation, text detection is applied by considering several scenarios, namely in the aspect of lighting and the use of different tools. With the use of text detection with Tesseract and EAST, can be utilized to make it convenient to detect the classification numbering. Furthermore, from experiments of the lighting scenario and the tools used, it was determined that the outcome would be better if it had more light. For taking pictures, security cameras demonstrably outperform webcams, which deliver the lowest quality results among the compared devices. It would be concluded that the use of a webcam should be avoided to prevent the worst result of text detection.

The implementation of it especially for detecting the classification number could provide a promising solution to the challenges associated with book structuring. Traditionally resource-intensive and time-consuming, the application of text detection, particularly utilizing Tesseract and EAST algorithms, has shown significant potential in efficiently extracting and interpreting classification numbering.

**Acknowledgement**

*The authors would like to thank Airlangga University for their support in the Satu JRS internal research funding scheme 2022 for this research.*

**References:**

[1]. Learned-Miller, E. G. (2011). *Introduction to Computer Vision & Robotics.* Deparment of Computer Science University of Massachusetts, 2–8.

[2]. Neethu, N.J., , & Anoop, B.K. (2015). Role of Computer Vision in Automatic Inspection Systems. *International Journal of Computer Applications, 123*(13), 28–31. Doi: 10.5120/ijca2015905603.

[3]. Ariyani, L. P. S. (2017). Manajemen Perpustakaan Sekolah pada SMA dan SMK Negeri Se-Kota Singaraja: Permasalahan dan Solusi Pemecahannya. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi Dan Kearsipan, 2*(2), 123. Doi: 10.14710/lenpust.v2i2.13479

[4]. Rustiarini, N. W., & Dewi, N. K. C. (2021). Penataan Perpustakaan Desa untuk Meningkatkan Literasi Membaca. *LOSARI: Jurnal Pengabdian Kepada Masyarakat*, *3*(1), 8-15.

[5]. Nalole, S. R., Golung, A. M., & Pasoreh, Y. (2019). Pengaruh Penataan Koleksi Terhadap Proses Temu Kembali Informasi Bagi Mahasiswa di Upt Perpustakaan Universitas Negeri Gorontalo. *Acta Diurna Komunikasi, 1*(3), 1–12.

[6]. Wibisono, O. (2011). Greedy Algorithm in Books Arrangement Problem. *Makalah IF3051 Strategi Algoritma, 1*.

[7]. Herawan, L. (2020). Pemilihan Metode Penataan Arsip Inaktif Konvensional di Records Center Arsip Nasional Republik Indonesia. *Khazanah: Jurnal Pengembangan Kearsipan, 13*(2), 132. Doi: 10.22146/khazanah.56509.

[8]. Yudiati, R., Annisa, A., & Ananda, A. (2022). Penataan Literasi dan Pengadaan Reading Corner di Paud Al-Hasan. *Jurnal SOLMA, 11*(1), 183–189.

[9]. Patel, C., Patel, A., & Patel, D. (2012). Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study. *International Journal of Computer Applications, 55*(10), 50–56. Doi: 10.5120/8794-2784.

[10]. Patel, C., Shah, D., & Patel, A. (2013). Automatic Number Plate Recognition System (ANPR): A Survey. *International Journal of Computer Applications, 69*(9), 21–33. Doi: 10.5120/11871-7665.

[11]. Greenhalgh, J., & Mirmehdi, M. (2014). Recognizing text-based traffic signs. *IEEE Transactions on Intelligent Transportation Systems*, *16*(3), 1360-1369.

[12]. Reina, A. V., Sastre, R. L., Arroyo, S. L., & Jiménez, P. G. (2006). Adaptive traffic road sign panels text extraction. In *Proceedings of the 5th WSEAS International Conference on Signal Processing, Robotics and Automation*, 295-300. World Scientific and Engineering Academy and Society (WSEAS).

[13]. Abdul Robby, G., Tandra, A., Susanto, I., Harefa, J., & Chowanda, A. (2019a). Implementation of optical character recognition using tesseract with the javanese script target in android application. *Procedia Computer Science, 157*, 499–505. Doi: 10.1016/j.procs.2019.09.006.

[14]. Rahmawati, A. N., Wibowo, S. A., & Sunarya, U. (2021). Analisis Performansi Sistem Optical Character Recognition (ocr) Pada Dokumen Digital Menggunakan Metode Tesseract. *eProceedings of Engineering*, *8*(5).

[15]. Hengaju, U., & Bal, B. K. (2023). Improving the Recognition Accuracy of Tesseract-OCR Engine on Nepali Text Images via Preprocessing. *Advancement in Image Processing and Pattern Recognition*, *3*, 40-52.

[16]. Xiang, Y., & Luo, F. (2020). Multi-type web image text detection based on the improved EAST algorithm. *Journal of Physics: Conference Series, 1544*(1). Doi: 10.1088/1742-6596/1544/1/012115.

[17]. Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., & Liang, J. (2017). East: an efficient and accurate scene text detector. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 5551-5560.

[18]. Yang, L., Ergu, D., Cai, Y., Liu, F., & Ma, B. (2022). A review of natural scene text detection methods. *Procedia Computer Science, 199*, 1458–1465. Doi: 10.1016/j.procs.2022.01.185.

[19]. Wang, Y. Y., Acero, A., & Chelba, C. (2003). Is word error rate a good indicator for spoken language understanding accuracy. *2003 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2003*, 577–582. Doi: 10.1109/ASRU.2003.1318504.

[20]. Das, D. (2012). Handwriting Characters Recognition Errors and Recognizing Errors of Children on the Tablet PC. *International Journal of Science, Environment and Technology, 1*(1), 36–40.

[21]. Saputra, I. K. E., Suhartono, E., & ... (2020). Efisiensi Steganografi Audio Untuk Meyisipkan Pesan Teks Dengan Rekonstruksi IRLS Dan Metode LSB. *EProceeding, 7*(2), 3576–3582.

[22]. Engwall, O., Lopes, J., & Cumbal, R. (2022). Is a Wizard-of-Oz Required for Robot-Led Conversation Practice in a Second Language*? International Journal of Social Robotics, 14*(4),1067–1085. Doi: 10.1007/s12369-021-00849-8.

[23]. Mumpuni, P. W., Widayat, R., & Aryani, S. M. (2017). Pencahayaan Alami Pada Ruang Baca Perpustakaan Umum Kota Surabaya. Vitruvian - *Jurnal Arsitektur, Bangunan, & Lingkungan, 6*, 71–78.

[24]. Supriyanto, W., & Muhsin, A. (2008). *Teknologi informasi perpustakaan*. Kanisius.

[25]. Lasa, H. (1990). *Kamus Istilah Perpustakaan*. Kanisius.

[26]. Hamakonda, T. P. (1983). *Katalogisasi Sebuah Pengantar.* Jakarta: Proyek Pengembangan Lembaga Pendidikan Tenaga Kependidikan Dirjen Dikti Depdikbud.

[27]. Rahma, E. (2018). *Akses dan Layanan Perpustakaan: teori dan aplikasi*. Kencana.

[28]. Purno, A., & Wibowo, W. (2016). Implementasi Teknik Computer Vision Dengan Metode Colored Markers Trajectory Secara Real Time. *Jurnal Teknik Informatika, 8*(1), 45–48.

[29]. Abdul Robby, G., Tandra, A., Susanto, I., Harefa, J., & Chowanda, A. (2019b). Implementation of optical character recognition using tesseract with the javanese script target in android application. *Procedia Computer Science, 157*, 499–505. Doi: 10.1016/j.procs.2019.09.006.

[30]. Onim, M. S. H., Nyeem, H., Roy, K., Hasan, M., Ishmam, A., Akif, M. A. H., & Ovi, T. B. (2022). BLPnet: A new DNN model and Bengali OCR engine for Automatic Licence Plate Recognition. *Array, 15*. Doi: 10.1016/j.array.2022.100244.

[31]. González-Cepeda, J., Ramajo, Á., & Armingol, J. M. (2022). Intelligent Video Surveillance Systems for Vehicle Identification Based on Multinet Architecture. *Information (Switzerland), 13*(7). Doi: 10.3390/info13070325.

[32]. Naik, S., Hashmi, K. A., Pagani, A., Liwicki, M., Stricker, D., & Afzal, M. Z. (2022). Investigating Attention Mechanism for Page Object Detection in Document Images. *Applied Sciences (Switzerland), 12*(15). Doi: 10.3390/app12157486.

[33]. Niharmine, L., Outtaj, B., & Azouaoui, A. (2022). Tifinagh handwritten character recognition using optimized convolutional neural network. *International Journal of Electrical and Computer Engineering, 12*(4), 4164–4171. Doi: 10.11591/ijece.v12i4.pp4164-4171.

[34]. Munivel, M., & Felix Enigo, V. S. (2022). Optical Character Recognition for Printed Tamizhi Documents using Deep Neural Networks. *DESIDOC Journal of Library and Information Technology, 42*(4), 227–233. . Doi: 10.14429/djlit.42.4.17742.

[35]. Thammarak, K., Kongkla, P., Sirisathitkul, Y., & Intakosum, S. (2022). Comparative analysis of Tesseract and Google Cloud Vision for Thai vehicle registration certificate. *International Journal of Electrical and Computer Engineering, 12*(2), 1849–1858. Doi: 10.11591/ijece.v12i2.pp1849-1858.

[36]. Hsu, P. H., Lin, C. T., Ng, C. C., Kew, J. L., Tan, M. Y., Lai, S. H., ... & Zach, C. (2022, August). Extremely low-light image enhancement with scene text restoration. In *2022 26th International Conference on Pattern Recognition (ICPR)*, 317-323. IEEE.

[37]. Xue, M., Shivakumara, P., Zhang, C., Xiao, Y., Lu, T., Pal, U., Lopresti, D., & Yang, Z. (2021). Arbitrarily-Oriented Text Detection in Low Light Natural Scene Images. *IEEE Transactions on Multimedia, 23*, 2706–2720. D
Doi: 10.1109/TMM.2020.3015037.

[38]. Lu, M., Mou, Y., Chen, C. L., & Tang, Q. (2021). An efficient text detection model for street signs. *Applied Sciences (Switzerland), 11*(13).
Doi: 10.3390/app11135962.

[39]. Leung, K. (2021). Evaluate OCR Output quality with character error rate (CER) and word error rate (WER). In *Towards Data Sciene*.

[40]. Park, Y., Patwardhan, S., Visweswariah, K., & Gates, S. C. (2008). An empirical analysis of word error rate and keyword error rate. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH,* September, 2070–2073.
Doi: 10.21437/interspeech.2008-537.