American Journal of Applied Sciences

Review

Textual Figure Processing Method: A Systematic Literature Review

[1,2]Wydyanto Yanto, [1]Norshita Mat Nayan and [1]Riza Sulaiman

[1]Institute of Visual Informatic, Universiti Kebangsaan Malaysia, Selangor, Malaysia
[2]Faculty of Science and Technology, Universitas Binadarma Palembang, Sumatera Selatan, Indonesia

Article history
Received: 21-03-2023
Revised: 02-06-2023
Accepted: 23-09-2023

Corresponding Author:
Wydyanto Yanto
Institute of Visual Informatic, Universiti Kebangsaan Malaysia, Selangor, Malaysia
Email: P93043@siswa.ukm.edu.my

Abstract: A human transaction always involves figured documents that have high practical value and could be screened using Optical Character Recognition (OCR). Optical character recognition is a knowledge to process different types of documents and figures to become data that could be analyzed, stored in files, edited and written for different uses. Lately, researchers used a lot of methods to analyze figures that contain text to be written into a media. This manuscript summarizes the figure processing method using OCR technology between 2014 and 2020 that involves text detection, extraction and segmentation for research guidance that applies collection, synthesis and analysis of research articles on the figure processing method. The source data is taken from electronic data using search protocols via keywords and forward and backward reference searches. This manuscript presents the output and the latest technique on the figure processing method via OCR for the research gap.

Keywords: Optical Character Identification, Figure Processing, Textual Figure, OCR

Introduction

There is a lot of information in a figure and this is very important because figure processing is a very important task. Segmentation is important before analyzing a figure which involves partitioning digital figures into a few segments. The text-embedded figure is large, especially in magazines, advertisements and websites. Text detection studies getting popular since there are a lot of applications for basic, text searching and filing documents. In this informative era, different organizations, complicated systems, indexing, information gathering, automatic classification and multi-scripts develop from time to time. Most of the backward conversion literature limits textual blocks, information letters, formulas, all types of application sheets, etc., practically, documents with poor contrast, blurred, slanted, poor in shape and irregular texture will have heterogeneous blocks such as annotation, printed machine and handwritten, graphic, figure, logo, photo and table structure. This complicates analyzing and detecting documents.

There are a lot of approaches for text and figure segmentation based on region and texture. Region-based approach to every pixel fixes the region and objects. This approach is divided into 2 subcategories: Basic Janani and Sheela (2016); Gundu *et al.* (2014); Grover *et al.* (2009) and connected components. Text region on document figure can be detected using the region base method and texture. This is relatively independent with measurement changes and text orientation but complex in terms of uneven background, e.g., If text touches a graphical object in an original figure and forms a component that is connected to a binary figure. Basically, in an inverted algorithm, the text symbol is stronger than noise, texture background and other graphical items (Janani and Sheela, 2016; Grover *et al.*, 2009; Mathur and Rikhari, 2017). In the top-down technique, the edge binary figure is produced using edge detectors that will be connected by including morphological operators or another algorithm such as run-length smoothing (Grover *et al.*, 2009). The component connected with the figure produced is the candidate text area because each line combines graphical items.

After that, each component decomposed to become a smaller region by analyzing vertical and horizontal proxy profiles. Finally, each heuristic obstacle is labeled as text. Etheredge basic method is rapid and able to detect text in a complex background by having text strings correlated horizontally and vertically. The input other f text figure approach will combine text and non-text or text, figure and background class texture. Most segmentation algorithm uses a classification frame with certain measures to ensure all or most of the pixels in the frame will have the same class (Angadi and Kodabagi, 2009). A classification

DOI: 10.3844/ajassp.2024.41.56

algorithm is used to label every frame in the feature space. Long *et al*. (2021) there are 2 class and 2-method classifications used to group every figure block as text and non-text appropriate with local energy in the wavelet transform domain. By using 3-method grouping on every figure pixel labeled as text, figure, or background according to 9-D feature vector based on Gaussian screening. Major feature statistics and geometrics were proposed for texture segmentation.

Wavelength transformation has become an effective tool in segmentation and classifying texture due to multiresolution characters because the wavelet base feature is interesting. Strong transformation domain for figure modeling marked by their edge. Text segmentation has broad scope but only a few studies focus on it. Text segmentation algorithm has adaptation and learning capability but is time-consuming and training data is satisfying but has practical limitations. To solve the problem, (Khan and Mufti, 2016) give an average procedure and compare it to produce training data from segmented figures manually. Then, the Naïve Bayes Classifier (NBC) was included quickly in the training and application phase. Figure segmentation and edge detection technique studies were proposed (Miri *et al*., 2017) where a new approach segmenting figures into page components (e.g., text, graphic and table). Their approach has 2 main steps: In the first step, an appropriate score from network convolutional for page component category with connected document component (Lu *et al*., 2015). Information extraction method proposed. Automatic smart text extraction identifies and extracts text from different types of figures. This system consists of different levels namely localization, segmentation, extraction and text introduction from figure (Ganesan and Sajiv, 2017). The edge detection method has an application process for the analysis of multiple edge detectors and new trends in edge detection. Edge in processing a figure has discontinuous intensity from one pixel to another. Edge detection is another figure advancement technique to increase the quality of a figure analysis (Karpagam and Manikandan, 2019) to extract text from a figure without limitation as proposed by Angie Venkatesan Karpagam and Mohan Manikan. The MSRA text detection 500 and dataset Street View Text (SVT) were used to test the proposed algorithm. The output is a dataset and localized text entropies to identify blurred text, different font characters to identify blurred text, the different fonts in a text and multilanguage text with different double multiplication from the complex backgrounds (Hoda *et al*., 2017). Systematic Literature Research (SLR) will not only consider the methodology of figure processing which comprises TESK with OCR technology. This manuscript also screens different methods to process figures to become better text.

Methods

As mentioned above, this manuscript aimed to identify and present literature on figure processing using OCR technology based on relevant research and studies. Therefore, the conclusion of this review is: (1) To summarize the accuracy of figure-to-text conversion methods. (2) To find out accuracy shortcomings using advanced observation. (3) To apply figure-based-text processing in identifying new area accuracy. We will follow strategies proposed by Medhat *et al*. (2014). The inclusion, exclusion, searching, selection, data extraction and synthesis criteria are discussed in subsection identification protocols.

Search Protocols

Search protocols were analyzed comprehensively by using philosophical, principled steps from the available systematic literature review (Medhat *et al*., 2014). This protocol identifies search background, search strategies, data extraction, accuracy principles and quality evaluation criteria to choose studies and data analysis. Search protocols focused on highlighting differences between SLR and traditional or narrative search (Hoda *et al*., 2017). This will increase the consistency of the review and reduce bias as the researcher has to be very accurate in presenting the search strategy and criteria.

Inclusive and Exclusive Criteria

This criterion is to ensure the relevancy of the articles used. Inclusive and exclusive criteria include accuracy studies from journals, conferences and a variety of methods used to process figures into text published between 2014 and 2020. A total of 825 articles based on figure processing methods using ORC technology were extracted by using keyword search (Fig. 1) for general search on the selection process. Articles without a clear method were ignored in the article selection procedures. Duplicated and incomplete text was also not included in the studies.

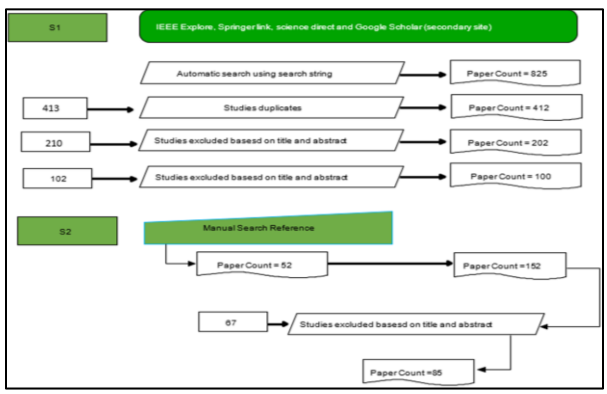


Fig. 1: General search to paper count

Search Strategy

The automatic and manual search of appropriate journals resulted in 825 references. Out of it, only 85 were chosen as reflected by Fig. (1).

The search strategy involves manual and automatic search in Fig. (1). Automatic search helps to identify and acquire main studies for wider perspective and review as recommended (Hoda *et al*., 2017). Then, a manual search was conducted on the obtained studies. A standard database was used to obtain the review article. However, magazines, paperwork, newspapers, books and blogs were ignored because they had verification issues. General keywords are taken from search objectives and volumes used to find articles. The objectives were to maximize the search for relevant articles of interest.

All permutation was taken into account which involves figure identification using OCR technology, figure processing, figure introduction, figure extraction, etc. After the acquisition of primary data, relevant research was done on the primary data that has inclusive and exclusive criteria. Then, Mendeley was used to keep all related references. Mendeley used to help identify duplicated studies because searching has a lot of basic data. The manual and automatic search was carried out extensively to avoid any overlooked materials by using back-and-forth references. All extracted data were then duplicated into a spreadsheet and repeated processes were done to omit irrelevant studies. Finally, the main post-studies set was added to Mendeley.

Process of Choosing Studies

Throughout the search and observation, about 825 journals were shortlisted based on four sources of presentations.

The toll approach was adopted for the search (Memon *et al*., 2020). About 825 studies were automatically searched in the database searched using relevant keywords (Memon *et al*., 2020). Out of 825, around 413 studies are duplicates and ignored. Exclusive and inclusive criteria were used on title, abstract, keywords and type of publication with the remaining studies 412 studies. About 210 studies were screened resulting in 202. Selection studies managed to leave us with 102 studies. After an automatic search, a manual searching procedure was conducted on 100 studies to cross-check reference and relevancy. A manual search resulted in an additional 52 studies and a prefinal registration was carried out for the total 152 studies. Finally, the 152 studies will be undergoing Quality Assessment Criteria (QAC) for SLR conclusion.

The QAC will identify articles that do not justify the question. After QAC inclusion, 67 studies were excluded leaving 85 primary studies. Figure 1 for a flowchart on the selection process. Table 1 shows the primary study distribution or selection between the different sources of publication before and after the selection process as mentioned Fig. (2) shows the same.

Table 1: Distribution of the database of selected studies before and after applying the selection process

Source	Count before applying selection process	Count before applying selection process
Elsevier	183	22
IEEE	251	24
Springer	261	21
Other	130	18
Total	825	85

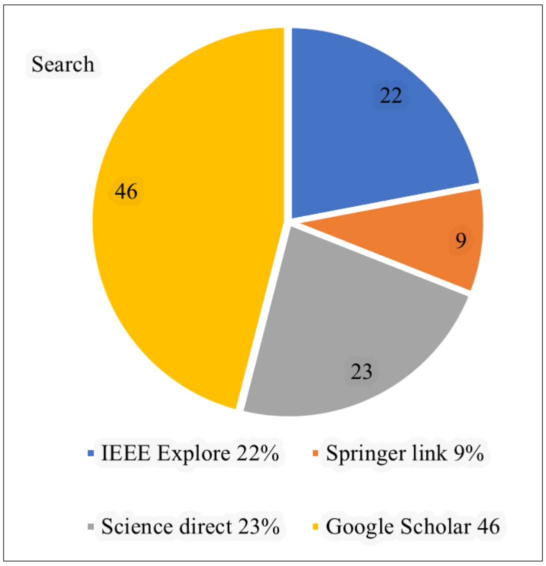


Fig. 2: Source distribution/selected database after selection process

Quality Assessment Criteria (QAC)

Quality Assessment Criteria (QAC) applied to a decision on quality-making principles on entire selected studies (Hoda *et al*., 2017) and listed below. This criterion helps to identify the strength of the conclusion and appropriate and relevant studies. Question on quality evaluation:

1. What is the relevant topic of interest
2. Which studies focussed on the evaluation context
3. Which articles proposed a concise methodology of interest
4. What is the data collection procedure for data of interest
5. Are the data analysis well explained using appropriate examples

A total of 119 studies were selected after treatment with the above credibility analysis. The above 5 quality analysis scheme was inspired by Heidarysafa *et al*. (2020). Evaluation quality was measured based on the score of each QA where a 2 2-point range was imposed i.e., 5-10. Evaluations below 5 are not included in the study which resulted in 96 review articles (Fig. 1) for a general search of the selection process.

Table 2: Extracted meta-data from selected studies

Selection features	Descriptions
Number of study identifications	Exclusive on selected articles
References	Bibliographies, author, volume, etc.
Type of paper	Journal, conference, book chapters
Language	English
Total collection	Collection number
Technique	Figure processing

Data Extraction and Synthesis

A total of 96 meta-data were extracted and Mendeley and MS Excel were used to evaluate them. This procedure will be used to record earlier studies (Medhat *et al*., 2014) data such as study ID (to identify each study), study volume, author, publication year, publication platform (conference, journal, etc.), collection number and accuracy context (a technique used) will be recorded in excel sheets. All studies were selected based on the technique and algorithm proposed by the authors. This enables us to classify appropriate techniques and language. Table 2 shows extracted meta-data from selected studies.

Selected Statistical Studies

In this section, selected statistical studies will be presented based on source, number of collections, temporal view, type of language and type of search methods.

Review of the Source of Publication

In this review, most of the studies were published in well-known journals and conferences. By considering the quality of studies, we believe this review will be used as a reference in finding new trends and triggering advanced studies on handwritten OCR domains. Figure (3) shows the percentage of 49 sources of publication based on study distribution. About 60% were published in journals, 30% were conference articles whereas 5% were each for book chapters and Qtr.

Search Collection

Search collection was obtained from Google Scholar. In total, selectively collected studies showed reliable qualities and are eligible to be included in this review as the authors are active in studies. The result is presented in Fig. (4). In the selected studies, 18 studies have 81-100 collections, 23 studies have 51-80 collections, 19 studies have 31-50 collections and 25 studies have 1-30 collections. The number of collections is predicted to increase in the future as more articles are published in this domain. Figure (4) shows the total selected studies. The numerical value indicates the total collected studies (appropriate value on a bar chart).

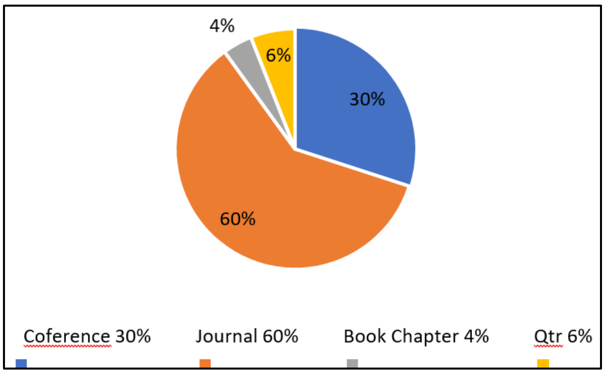


Fig. 3: Publication source based on study distribution

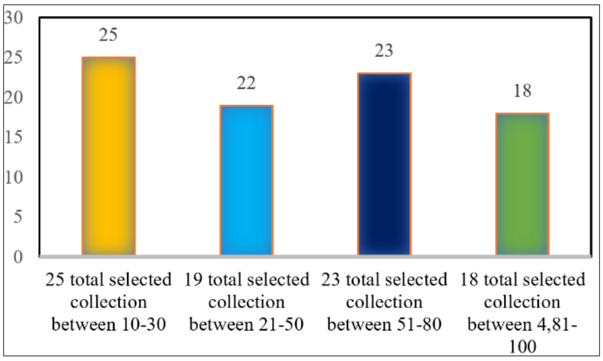


Fig. 4: Total selected studies. The numerical value indicates the total collected studies (appropriate value on a bar chart)

Methods of Image Processing Classification

This manuscript will summarize techniques used to process textual figures based on processing level and type of techniques.

Significance of Pre-Processing in Character Identification

The ability to problem solving which affects the character identification mechanism could be summarised into factors as listed in part 2. Therefore, the pre-processing technique manages to upgrade the figure documents and prepare them for the next character identification steps (Uysal and Gunal, 2014). To achieve a high recognition level, an effective OCR-based system pre-processing algorithm level is needed. This includes accurate advancement of figures, noise removal, figure threshold identification, detection of possible corrective mechanisms, page segmentation, character segmentation, character normalization and morphological techniques.

Pre-Processing Techniques

This technique needs to be applied to colors, grayscale documents, or binary figures that contain text and graphics. In character introduction, most of the

DOI: 10.3844/ajassp.2024.41.56

application uses grayscale or binary figures because processing colored figures needs high computations (Hamad and Mehmet, 2016). The figure might consist of an uneven background or watermark which complicates extraction without pre-processing. Therefore, the aspired output of pre-processing is a binary figure with texts.

To achieve the aspired output, a few steps need to be carried out. First, figure advancement is needed to remove noise and adjust figure contrast. Then, thresholds will be set up to eliminate background with a scene, watermark, or noise before page segmentation. Now, character segmentation will be carried out to separate characters from each other. Finally, morphological processing will be done to upgrade characters for figures that have thresholds and pre-processing that disrupts the characters and increases the pixels (Poovizhi, 2014; Kaur *et al.*, 2017) The technique above introduces a variety of character identification systems in a few applications which could be applied to OCR at different levels.

Figure Advancement Technique

An appropriate technique used to increase the quality of the figure was used (Abdullah *et al*., 2016; Gupta *et al*., 2016; Gautam *et al*., 2018). The image was presented after noise removal, blurred reduction, contrast intensification and extensive information inclusion.

Spatial Figure Filtering Operations

In processing an image, the filter was used to compress higher frequency on the figure i.e., figure depletion, or lower frequency, increase or detect figure side (Ravi *et al*., 2015). Restoration and upgrading figure explained in spatial and frequency domain i.e., Fourier transformation. However, Fourier transformation complicates substantial calculation in a few cases. Multiplication in the frequency domain is suitable with time convolution in time and spatial domains.

Using a small convolution mask such as 3×3, the distance figures are easy and faster to be transformed or permutated based on Fourier permutation. Therefore, there is a lot of spatial filtering technique is presented in this chapter. The figure taken is influenced by noise and is not ready for analysis (Scheerlinck *et al*., 2019). For a figure to be accepted, its area needs to be compressed or highlighted. The spatial process is classified into point and mask processing. Point processing involves a transformation of individual pixels, independently from different figure pixels. Average operations are used to modify damage on hardcopy figures, i.e., compensating figures. On the other hand, mask processing resulted in pixels with (x, y) coordinates.

Point Processing

Point processing modifies the pixel value of the original figure to an appropriate pixel on a figure advanced using the formulae (1) (Smitha *et al*., 2016):

$$0(x,y)=T\left[I(x,y)\right] \quad (1)$$

where, I (x, y) is the original figure input, O (x, y) is an advanced figure and t presents the transformation between the two figures. The point processing technique includes contrast stretching, global threshold, histogram evenness, log transformation and law power transformation. Mask processing techniques comprise average filter, sharpening filter, local threshold and many more.

Contrast Stretching

At the sensory acquisition level, the contrast level of the figure can vary because of poor brightness or inappropriate arrangement (Negi and Bhandari, 2014). Thus, a figure with manipulated contrast and compensation will complicate figure acquisition. The idea was to change the dynamic range of the grayscale figures. Linear mapping as in Eq. (2) used to stretch the pixel value of a figure with lower and higher contrast via widening the dynamic range of the entire spectrum of a figure from 0- $(L$-1):

$$0(x,y) = 0_1 + \left(\frac{0_2-0_1}{I_2-I_1}\right)[I(x,y) - I_1] \quad (2)$$

where, O_1 corresponds with 0 while O_2 corresponds with the total aspired level i.e., L-1 = 255. The I_1 and I_2 give minimum and maximum ranges for gray inputs. The average processing shape is adapting the brightness of figures by adding bias value, b, onto all pixel's value of figures; where b >0 will increase figure brightness whilst b <0 darkens the figure. Strengthening value can be used as a bias replacement, where the modified pixels of inputs produce brighter output figures.

The value 0< a <1 produces a darker figure and a value is a >1 produces a brighter figure. Equation (3) was obtained by combining bias and gain:

$$0(x,y) = a * I(x,y) + b \quad (3)$$

In this issue, we have to determine the gain and bias value which is complicated practically; therefore, the solution is to design an input figure range (I_1, I_2) to produce an output figure range (O_1, O_2) where O_1 is suitable with O while O_2 suitable with a total level of interest, thus, producing linear Eq. (2).

Global Image Thresholding

Image thresholding is a process of separating information (object) from the background of a figure. Thus, thresholding is commonly included in converted figures regardless of color or grayscale (Lore *et al.*, 2017). Thresholds can be categorized into global and local. The global thresholding method chooses a single threshold value for the entire document figure which depends on the background estimation of the figure intensity histogram because it is estimated as a point processing operation. On

DOI: 10.3844/ajassp.2024.41.56

the other hand, the adaptive local threshold uses a different value for every pixel based on local area information. Hundreds of thresholding algorithms and methods are published in literature and books but only popular methods are discussed here. The global thresholding method is used automatically to reduce grayscale and binary images (Chaubey, 2016).

The figure is assumed to have a 2-pixel class (foreground and background). The threshold method automatically determines the threshold value, T, where the lower pixel value represents the foreground and the upper value represents the background. The average method chooses a range from all input pixels to be used as a threshold value. However, this is not possible if the pixel is not evenly distributed in a figure. The sophisticated approach intensifies the figure pixels of the histogram by using the minimum valley point as the threshold. The histogram approach assumes there are few grayscale values for background and object pixels, but the actual pixel value has a grayscale variation of the surroundings. This is an expensive computation and the figure histogram might not have a clear valley point which complicates choosing the procedure of accurate threshold. A repeated method is an average relative method and does not need specific knowledge (Balabanian *et al*., 2017) as explained below. The repeated procedure is:

Step 1: Choose the initial threshold value (T) which is appropriate for the method such as the average pixel value of the figure
Step 2: Segmented figure included onto an object and background pixels using T. R_1 (background are) consists of the pixel with intensity value $\geq T$ while R_2 (object are) consists of the pixel with intensity $<T$
Step 3: Calculate the average of every area, μ_1 and μ_2 for areas R_1 and R_2.
Step 4: Calculate the new threshold value T as given in Eq. (4)
Step 5: Repeat steps 2-4 using T until a new suitable threshold is achieved

$$T = 1/2(\mu_1 + \mu_2) \tag{4}$$

In literature, a lot of thresholding methods have been published such as Sahoo who differentiate more than 20 global threshold algorithms using even and shape measurement. Differences show Otsu separation class shows the best performances (Zaitoun and Aqel, 2015; Xie *et al*., 2019). On the other hand, change detection evaluation concludes that the Otsu algorithm has lower performance compared to other global methods (Gwet *et al*., 2018) evaluated 4 global techniques and concluded that the Otsu method is the best method among all. Moreover, Fisher compares 15 methods and confirmed Otsu method is preferable in the processing figure (Wang *et al*., 2014). The Otsu method is used widely to change the grayscale figure to become a binary figure before the calculation of the optimum threshold is conducted to separate two classes until minimum compilation widening (intra-class variant). Otsu method looks for a threshold to minimize the intraclass variant as defined in Eq. (5) as a total variated variant from 2 classes (Gurung and Tamang, 2019):

$$\sigma_\omega^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_1^2(t) \tag{5}$$

The ω_1 is a probability obtained from two separated classes using a threshold where t and σ_1^2 are a variant. Otsu shows that there is a minimum intra-class variant which is the same as the maximum inter-class variants as shown in Eq. (6):

$$\sigma_\omega^2(t)\sigma^2 \text{-} \sigma_1^2(t) = \omega_1(t)\omega_2(t)[\mu_1\,(t) - \mu_2(t)]2 \tag{6}$$

This is stated in probability class $\omega\ i$ and significant class μ_1 which could be renewed interactively. The algorithmic steps include: Calculate histogram and probability of every intensity-initialization of ω_1 (0) dan μ_1 (0)-follow all threshold values with $t = 1$ for maximum intensity. -renew ω_1 (0) dan μ_1 (0). Calculate the maximum (σ_b^2(t), suitable with the desired threshold).

Histogram Processing

Histogram processing is used in figure advancement and is useful in figure segmentation and comprehensive processing. Histogram only charted frequency where the grayscale level ranges from 0 (black) to 255 (white). Transferred or processed image has a color threshold or lack contrast (detail) that is able to help machine observation such as segmentation. Histogram evenness and specification (coordination) is a popular method to modify the histogram of a figure with better quality.

Histogram Evenness

Histogram evenness is a global technique where the pixel spectrum is widened (0-255). This upgrades the final figure for human observation and can be used to normalize brightness variation in understanding a figure. This process is averagely applicable to the original level of brightness (j) and new pixel level (k) as given in formula (7):

$$k = \sum_{i=0}^{j} N_i IT \tag{7}$$

where, calculating all pixels in a figure (by integrating histogram) with the same or lesser brightness, (j) and T indicate total pixels (Buchelly *et al*., 2016). On the other hand, histogram evenness is an operation used to get a new figure after histogram specification and modification.

DOI: 10.3844/ajassp.2024.41.56

Histogram Specification (Matching)

Histogram matching is a method used to process a figure with two colors by using a histogram.

The F_2 cumulative distribution function is modified based on the histogram of F_1 reference figures. Histogram calculation will be carried out on figures of F_1 and F_2 and the histogram will be modified accordingly. Then the F_2 figure will be matched according to the *CDF* of the reference figure (F_2). The matching output histogram was obtained by matching the *CDF* of F_2 and F_1 by using the nearest value. Finally, the grayscale levels g_1 and g_2, as well as $f_1(g_1) = f_2(g_2)$, were calculated as shown in Fig. 2. Figure 2 is an output from the histogram matching function $M(g_1) = g_2$ (Hongbo and Xia, 2014).

Log Transformation

General log transformation is (8):

$$S = c\,log\left(1+r\right) \tag{8}$$

where, c is a constant and assumes $r \geq 0$. This transformation narrows the score of the grayscale level into an input figure to a wide output range and vice versa (Abdullah and Quteishat, 2015).

Principal Rule of Transformation

The principle rule of transformation has a general shape such as in (9) (Gómez-Aguilar and Atangana, 2018).

$$S = C\left(r+\varepsilon\right)^2 \tag{9}$$

where, c and γ are positive constants and ε is an ignored offset because of display calibration. Therefore, $s = c.r^2$ where $0< \gamma <1$ computes narrow ranges from dark input value to wider range from output value with inverted γ value that is more than 1. This shows that the principal rule of transformation is more flexible than log transformation. However, the log function is able to compress and convert the figure range into a bigger range within the pixel value. Since, the figure undergoes a variety of phases, printing and presentation of appropriate responses need to be adjusted with the exponential power rule, γ. These factors need to be modified till the responding power rule and gamma value is corrected according to the $s = c.r$^$(1/y)$.

Mask Processing

In mask processing (Valero *et al*., 2016), the pixel value is taken from the original and surrounding pixel value. This operation is more expensive than average point processing, yet more reliable. Masker application in the output figure produces similar measurements to inputs.

Smoothing Fil0074er (Low-Pass)

Average filter or average is a simple method of figure smoothing, intuitive and easily included, that is reducing total intensity variant between one pixel and another. This is used to reduce noise in the figure. Generally, the average filter serves as a low-pass frequency filter and therefore, reduces the special intensity in the figure (Kejriwal and Singh, 2016). The compilation of the average filter replaces every pixel value in the figure using the average value from the neighbor including itself.

This has omitted the pixel value that does not belong to the range. Average filtering is commonly known as convolution filter. Just like other convolutions, the base is on masks that represent the shape and surrounding measurement when the sample is calculated from the average. Mostly mask 3×3 m^2 is used as in Fig. (3), though a bigger mask is used (e.g., 5×5, 7×7, 9×9...) the average still becomes complicated. In the figure, a small mask is able to give the same non-identical output after being used multiple times which has the same effect when used on a larger mask. Moreover, the mask element has to be positive and its measurement determines the smoothness. Therefore, the bigger the measurement of a frame the bigger the blurriness because the small object combines with the figure's background.

The average mask coefficient is very important with invertedly weighted pixels as a distance function from the mask center. To reduce blurred figures during the process of evenness, the coefficient value is used as a function to increase the distance from the original point. The strategy used was averagely weighted based on higher center.

Sharpening Filter (High-Pass)

A sharpening filter is used to compress the smallest details in a figure (e.g., the Inversion effect from smoothing). High contrast points can be detected by computing intensity differences in the local area of a figure. Mask quality can be either positive or negative. If the mask is in the area of constant and late grayscale, the convolution output nears (Khorsheed, 2014). When grayscale variates quickly in a compilation, convolution output will increase in number. Generally, points form limits among different objects or background sections (e.g., sides). An example of a smoothing filter is the Laplacian filter defined in Eq. (10) as in the equation:

$$\Delta^2 = f\begin{bmatrix} f\left(x+1,y\right)+f\left(x-1,y\right) \\ +f\left(x,y+1\right)f\left(x,y-1\right)\end{bmatrix} - 4f\left(x,y\right) \tag{10}$$

This implementation could be imposed on all points (x, y) of the figure by combining all figures with spatial masker figures. Figure (4) shows an alternative definition from second digital derivatives that calculates diagonal features and is implemented on a mask in Fig. (4).

DOI: 10.3844/ajassp.2024.41.56

Figure (7) Laplacian filter mask with a derivative operator that sharpens up the figure but nullifies the constant area. Therefore, editing back original figure replaces the back grayscale level tone by using Eq. (11):

$$g(x,y) = f(x,y) + c\left[\Delta^2 f(x,y)\right] \tag{11}$$

where, $f(x, y)$ is the input figure, $g(x, y)$ is the output figure and c is 1 if the positive center mask coefficient or -1 if negative.

Median Filter

A general nonlinear operator used is a median, low-pass specific filter. The median filter chooses the figure area (3×3, 5×5, 7×7, etc.), evaluates all pixel values and replaces the center pixel with the median value (Shrestha, 2014). The median filter complicates convolution. If the considered range contains an even value, the median value will be used. Figure (5) illustrates how the median filter is calculated. A median filter is effective to remove noise impacts such as 'Salt and pepper noise' which is common in black and white pixels.

Maximum Filter

The maximum filter is defined as the maximum pixels in the local area of the figure. Maximum filter included on figures to remove noise and negative outliers (Sanders *et al*., 2017). For instance, a figure with 5 pixels will be maximized into 145 pixels.

Minimum Filter

The minimum filter increased the dark value of a figure. Therefore, the darkest pixels were chosen to be new pixels on the center frame (Eyupoglu, 2017). For example, a figure with a middle pixel of 5 will be replaced with a minimum value of 100.

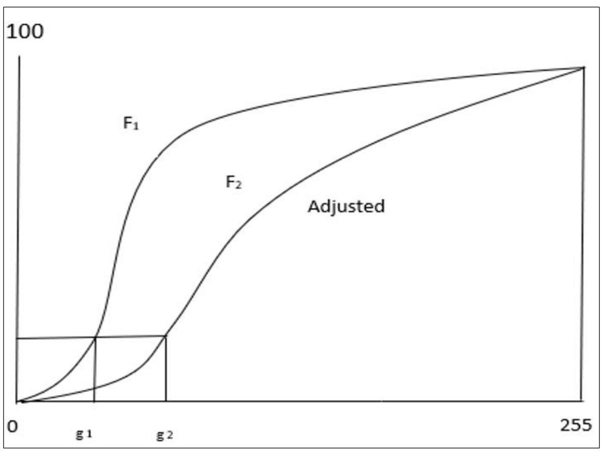


Fig. 5: Cumulative distribution functions for reference and adjusted figure

Range Filter

A range filter is defined as the difference between the maximum and minimum pixel values. For example, a figure with 5 center pixels will be replaced with 45. Local threshold technique used with the figure that has background uneven brightness or complicated background such as a watermark that could be overlapped in a peace document if it has a global threshold. This is because the histogram of figures gives more than two peaks until complicates the global floating technique to separate objects from the background. In this case, local thresholding is the only solution. Local thresholding techniques are advanced in the literature, especially for specific applications and most of the time they do not perform well in different applications. The output exceeds the limit threshold or is reduced in terms of contrast and illumination. From the literature, few surveys have compared different thresholding techniques. Trier and Jain evaluated 11 methods of binarization based on adaptive local sustainability for performance (Ravi and Ashokkumar, 2017).

These techniques were compared using the criteria of the OCR module in identifying handwriting from hydrography. In this evaluation, the Niblack method (Irviantina and Pardosi, 2016) is the best. This observation included for specific application of the hydro graphical figure by using the OCR system. However, as concluded by the author, if the figure used differs with different feature extraction and classification, then, this observation might be inaccurate and other methods can outperform the Niblack method (Granlund and Böhrnsen, 2017) black method calculates the threshold by shifting the frame across the figure and using local mean, μ and standard deviation, σ, for every median pixel of the frame. The threshold limit value for pixels in a fixed circle is an average linear function or circle standard deviation pixel with constant gradient $T(x, y)$ that could be settled to separate objects. Thus, thresholding follows the Eq. (12):

$$T(x,y) = \mu(x,y) + k\sigma(x,y) \tag{12}$$

Surrounding measurement should be small to consider local detail but at the same time large enough to compress noise. The k-value is used to arrange how many printed objects take part in a given object. There are a few methods that introduce a modification to the Niblack method such as Irviantina and Pardosi which suggest the best version of the Niblack method (Irviantina and Pardosi, 2016). Moreover, there are a lot of thresholding methods that is developed based on a different figure. For example, the local thresholding method developed by Boyat uses statistical texture features and MLP-NN to classify background and foreground. This was to characterize surrounding pixel value and property measurements such as smoothness, evenness and variability. The 5 features extracted from the 3×3 frame in this article are average frame pixel, average, standard variation, slanting and entropy. Features are extracted from every pixel in the figure and treated in MLP-NN to classify background (white) and foreground (black). The MLP-NN thresholding method improvises thresholding documents with poor illumination, complex background, or uneven background as could be found in the peace treaty. The MLP-NN threshold limit method is non-application without specific and works widely.

Noise Removal

Advanced technology produced figure acquisition with an upgraded version. While modern technology enables noise, reduction related to electro-optic to an ignorable level, there are still a few noises that could not be reduced (Jundale and Hegadi, 2015). The figure obtained from the modern sensors can be contaminated with noise sources. Stochastic variants opposite deterministic distortion such as shade and less focussed were used as references. There is a lot of noise related to electronic capture or light sources such as the photon, thermal, electronic on-chip and quantization. Most noise can be removed with a capture sensor or CCD camera. Noise removal at pre-processing gives a substantial increase in reliability and stiffness from feature extraction and OCR system introduction. General manifestation from noise in binary figure takes the isolated pixels, salt and pepper, or speckle noise. Thus, the process of removing a type of noise called filling where every pixel isolated 'island' sand and pepper filled with 'ocean' around it (Barlas *et al*., 2014). Grayscale figures or median filters and low-pass filters such as average blur filters or Gaussian proved to remove isolated pixel noise. A Gaussian blur and average filter are the better choice to give a smooth texture to the figure. On the other hand, periodic noise that manifests itself as an explosion such as an impulse is obvious in the Fourier spectrum screened using notch filtering. Butterworth notch transfer function from sequence n, vu H given by Eq. (13):

$$H(\mu,v) = \frac{1}{1\left[\frac{D^2}{D_1(\mu,v)D_2(\mu,v)}\right]} \quad (13)$$

where:

$$D_2(\mu,v) = \left[\left(\mu - \frac{M}{2} - \mu_0\right)^2 + \left(v - N/2 - v_0\right)^2\right]^{1/2} \quad (14)$$

And:

$$D_2(\mu,v) = \left[\left(\mu - \frac{M}{2} - \mu_0\right)^2 + \left(v - N/2 - v_0\right)^2\right]^{1/2} \quad (15)$$

where, (μ_0, v_0) and symmetry (-μ_0, -v_0) is the notch location and D is the digits, Eqs. (14-15). Filter fixed with frequency length with the square center.

Slanted Detection or Correction

Due to input rotation and sensitivity, a lot of figure analysis methods have to undergo figure correction (Nilsson and Lindahl, 2016). Slanting detection techniques are roughly classified into proxy profile analysis, hough transformation, connected component, classification and line correlation technique. Adam and Paquet researched 25 different methods of detecting slanted documents. This includes hough transform analysis, proxy profile, feature point distribution and feature analysis with orientation sensitivity. The survey shows a large part of the technique reported a 0.1 accuracy level, proving a stronger requirement for further studies as well as focussing on verifying the strengths and weakness of individual algorithms. Thus, there is a new specific application as proposed by Sawant and Chougule (2015) with a gravitational baseline (Binmakhashen and Mahmoud, 2019). Hence, a slanted or correction technique is used on the figures.

Page Segmentation

After upgrading the figure, noise removal detection/slanting correction, the next step in mixed content figure or composite figure by page segmentation which separates halftone image, line and graphic. The attractive output should be figured with text. Therefore, page segmentation will be done by classifying the document into 3 categories namely top-down, bottom-up and hybrid. The top-down segmentation method divides the large region into smaller subregions. Segmentation will be stopped when the criteria are fulfilled and the end segmented background is obtained. On the other hand, the bottom-up method starts by classifying interested pixels and combining bigger blocks to become larger blocks or connected components as characters before grouped to form words, lines, or textual blocks. A hybrid method is a combination of a top-down and bottom-up strategy. Run-length Smearing Algorithm is the favorite method of the top-down. This is used on a binary figure (setting 1 for the white pixel and 0 for the black pixel) by connecting both black pixels that are on the limit. This method included line-by-line or column-by-column and then both of the outputs were combined in OR logic before the smoothing threshold was used to produce the final segmentation.

From the RLSA output, a black text block and figure will be produced. Finally, a statistic classifier is used to classify blocks (Medhat *et al*., 2014). A bottom-up algorithm is an *X*-*Y* recursive method also known as profile proxy cutting, with presented document assumption with a tree-shaped block and nested length square (Smitha *et al*., 2016) The *X*-*Y* recursive cutting resolves the documented figure to become a block set with square length, without detailing that does not define cuttings. On the other hand, the hybrid method is the Kruatrachue and Suthaphan segmentation approach involves a top-down extraction method with a block detection method and a multicolumn segmentation method (Angulo and Velasco-Forero, 2014). Block column-based segmentation extracts edge-to-edge algorithm modification by using 32×32 frame pixels until paragraph characters can be obtained. Above are only a few examples and methods developed for document segmentation. Detection and slanted algorithms at the pre-processing level are important to ensure performance (Alzoubi *et al*., 2023). At largely explained technique takes time and is not effective in document processing layouts with complicated geometry. Specifically, the top-down approach can only process an average document that has a certain format or appropriate information. Failure to process documents with complicated geometric structures. This research targets the binary or grayscale figure with an even background (Karthick *et al*., 2019). A lot of figures are transferred from technical journals and magazines which have a certain format. Document segmentation on grayscale figures has complicated background or unevenness due to complications and figure threshing. Therefore, the main technique is directed for a specific application with a specific format which is prone to fail when the parameter is not suitable. Alginate uses local threshold MLP-NN to limit figures with even backgrounds and includes RLSA with modified parameters to classify mixed figures onto text, line, halftone figures and graphics (Sinecen, 2016).

Characters Segmentation

Character segmentation is an important pre-processing step, especially in cursive scripts such as Arab, Urdu and other related characters (Lou *et al*., 2014). Therefore, there is a lot of advanced technology for character segmentation and most of them have a specific script that is dissimilar to other scripts. Though printed, handwritten documents still need character segmentation. For instance, Latin characters are easy to print after segmentation of histogram profiles horizontally and vertically. However, small and serif fonts can introduce touch that could complicate the next steps further.

Normalization of Figure Measurement

Character segmentation prepares isolated characters for the feature extraction phase (Medjahed, 2015). Consequently, normalized and measured isolated characters were determined empirically or experimentally. This depends on the application, feature extraction and classification techniques. Then, extracted characters are treated with the same measurement to ensure the evenness of the data (Roynard *et al*., 2019).

Morphological Processing

Segmentation of output causes loss of pixels and produces holes in a section of a figure; exhibited by holes present in characters where few pixels are removed up to the threshold (Savitha *et al*., 2014). Larger holes will separate the characters into two or more. On the other hand, inversion is also possible as segmentation can combine together with separated objects and complicate character separation; interpretation is complicated when it involves solid objects. However, this problem can be overcome by morphological filtering. The techniques used include erosion and widening, opening and closing, deciphering, depletion and skeletonization which are only applicable to binary figures (Raid *et al*., 2014; Louis, 2016).

Erosion and Dilation

Erosion and dilation are morphological operations that will add to and reduce the measurement of an object and are very important at the pre-processing level (Nayak *et al*., 2015). Erosion will make objects look smaller by removing and eroding side pixels; However, widening makes larger objects by adding side pixels (Chatbri *et al*., 2016). There are two general techniques for erosion or widening i.e., threshold technique and closing. The threshold observes a nearby pixel and changes its status if the total nearby different pixels exceed a threshold. The baseline is if the total null pixel around the ranging pixel exceeds a threshold parameter and the pixel is fixed to null. Figures (6-7) shows erosion square output by using three threshold values (Alaei *et al*., 2016).

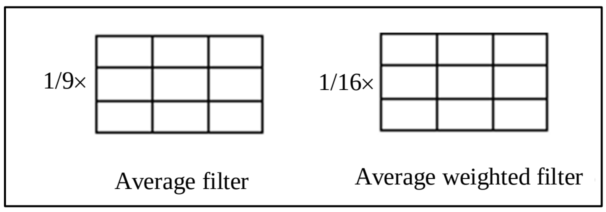
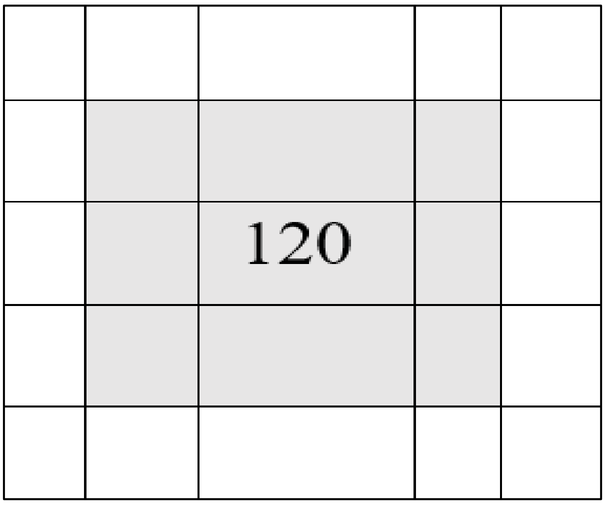


Fig. 6: Shows the 3×3 average filter used in the average weighted filter

123	127	150	120	100
119	115	134	121	120
111	120	122	125	150
11	119	145	100	200
110	120	120	130	150

(a)



(b)

Fig. 7: (a) Figure inputs; (b) Filtered figure uses median filters which exhibit only center pixels. Pixel values sorted from the shaded area are: (100, 115, 119, 120, 121, 122, 125, 134 dan 145), giving median value 121 on the output figure

The widening process shows erosion inversion. This calculates the pixel value next to null pixels. If the calculation exceeds threshold parameters, null pixels are adjusted to the pixel value. Widening on Fig. (8) uses two threshold values.

Figure (9) widening output (a) Internal (b) Using 2 thresholds.

Masking techniques use an $n \times n$ array (3×3, 5×5, etc.). The 1 s and 0 s on the output figure erode and widen inputs. Using masks, erosion direction or widening can be managed. Some widely used square maskers are 3×3, 5×5, 7×7, etc., while others are still usable (Wick and Puppe, 2018; Nguyen and Nakagawa, 2016). Mask measured 3×3 from the different directions are shown.

This adjusts the next mask on the figure until the middle array is null. If one of the 1 s coefficients in the overlapping mask is shared with the white pixel (255) on the figure, then the value will be adjusted to null. Vertical mask erosion removes the upper and borderline from an object, horizontal mask removes the left and right column while horizontal-vertical mask removes pixels from all angles. To conclude, widening causes objects to grow in measurement because every pixel value with maximum value with frame $n \times n$ around pixels. The process can be repeated to create larger effects. However, the erosion works the same way except the object is reduced because every pixel is replaced with a minimum value with an $n \times n$ frame around the pixel (Zhang *et al*., 2018).

Opening and Closing

Opening and closing are morphological operators originating from the base erosion operator and dilation that are used in the binary figure. Base effects of opening such as erosion more prone to remove the front pixel from a side area of a background pixel. In general, this is less damaging than erosion. In so many cases, closing is similar to widening because prone to the widening of the limit of the background area in the figure but less damaging original form limit. Opening nearby object space, releasing touching objects and widening holes in an object. Figure (8) shows 2 objects combined by a thread; an opening is used to remove the thread and separate the 2 objects. Therefore, by eroding the object two times, the thread is removed. In this issue, widening will enlarge the 2 objects to return to the original measurement but will not repeat threads (Sultan *et al*., 2017).

Opening can widen the desirable hole in an object; and involve one or more erosion followed by a widening process. Closing combines the covering of a spoilt object and an undesirable hole. Figure (9) shows two objects that should combine to form a line and Fig. (10) shows how holes have been covered in an object.

Opening and closing operators work well but can produce undesirable output by combining objects that could not be combined while opening can widen the hole and cause the picture to break. The answers involve specific opening and closing to avoid problems. For additional information refer to (Babin *et al*., 2018; Babu and Sunitha, 2011).

(a)

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(b)

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	255	255	0	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	0	255	255	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Fig. 8: Output erosion of long square using 3 thresholds

(a)

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(b)

0	0	0	0	0	0	0	0
0	0	0	255	255	0	0	0
0	0	255	255	255	255	0	0
0	255	255	255	255	255	255	0
0	255	255	255	255	255	255	0
0	0	255	255	255	255	0	0
0	0	0	255	255	0	0	0
0	0	0	0	0	0	0	0

Fig. 9: Widening output; (a) Internal (b) Using 2 thresholds

vertical mask	horizontal mask	horizontal and vertical masks	
0 1 0	0 0 0	0 1 0	1 1 1
0 1 0	1 1 1	1 1 1	1 1 1
0 1 0	0 0 0	0 1 0	1 1 1

Fig. 10: Shows widening output by using horizontal masks

DOI: 10.3844/ajassp.2024.41.56

Outline

The outline is a type of edge detection that is only functional for binary figures yet able to produce more reliable figures than a regular edge detector. Deciphering binary figures is easy with erosion and widening (Jayasree and Narayanan, 2015). Deciphering an object internally, as well as scraping objects will reduce the originally scraped in Fig. (11). Figure (12) shows the effect of deciphering objects externally and reducing the original figure by figure widening. An exterior outline is understandable where it widens an object by increasing the pixel value and reducing inputs from larger objects.

Depleting and Skeletonization

Skeletonization is a process of reducing background area in binary figures which largely affects upgrading and original connectivity area by removing background pixels. The frame is used as a bitangent curve location which makes a desirable background area of consideration which could be illustrated in the form of the long straight. There are two base techniques used to produce object frames: Base depletion and average bar transformation. Depletion is a morphological operation used to eliminate background pixels, chosen from binary figures such as erosion and opening.

Depletion is a process of data reduction that erodes an object till becoming one pixel, producing an object framework that is able to recognize object-like characters. Figure (13) shows the depletion of character E to produce thinner characters. Depletion will be installed on binary figures to produce output binary figures. Depletion erodes objects multiple times (without breaking them) until their width is 1 pixel. On the other hand, average bar transformation centralized all points of an object to form a center line.

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	255	255	255	255	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(a)

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	255	255	255	255	255	255	0
0	255	255	255	255	255	255	0
0	255	255	255	255	255	255	0
0	255	255	255	255	255	255	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(b)

Fig. 11: Widening output; (a) Using horizontal mask; (b) Erosion mask are inversion from widening

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	255	255	255	0	255	255	0
0	255	255	255	0	255	255	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(a)

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	255	255	255	255	255	255	0
0	255	255	255	255	255	255	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(b)

Fig. 12: Output after undesirable hole covering on the object to form a line

0	0	0	0	0	0	0
0	255	255	255	255	255	0
0	255	255	255	255	255	0
0	255	255	0	255	255	0
0	255	255	0	255	255	0
0	255	255	255	255	255	0
0	255	255	255	255	255	0
0	0	0	0	0	0	0

(a)

0	0	0	0	0	0	0
0	255	255	255	255	255	0
0	255	255	255	255	255	0
0	255	255	255	255	255	0
0	255	255	255	255	255	0
0	255	255	255	255	255	0
0	255	255	255	255	255	0
0	0	0	0	0	0	0

(b)

Fig. 13: Output of undesirable hole covering on object

Centre bar transformation is similar to Euclidean distance measurement where minimal distancing is used to evaluate the internal and external points of an object. This is clear when a figure contains a different object with various brightness among the area of a picture. Hence, appropriate brightness needs to be selected either manually or automatically.

Conclusion

A character identification system was introduced as a pre-processing technique using figured documents. These are important for future online character identification used on cellular, video-based-text extraction, peace treaty information extraction and historical document processing. The aim of the observation was to ensure the accuracy and efficacy of information extraction in real time. There are a lot of methods and techniques were developed, but a few problems still occur. Thus, a serious solution is vital. Most of the available pre-processing techniques are application-specific and not compatible with other applications. So, this gives the idea that every technique of pre-processing differs according to the quality of figures. Manipulation or advancement techniques are not compulsory for the entire figure since not all parts of the figure have noise or contrast variation. Therefore, advancing only part of the original figure will be more influential in many situations since a figure that contains different objects may have different brightness levels regardless of automatic or manual search of local details. In conclusion, pre-processing is an important level in analyzing the accuracy of automatic documents.

Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Wydyanto Yanto: Responsible for the compilation and collection of materials.

Norshita Mat Nayan: Provided expert guidance in the preparation of journal manuscripts.

Riza Sulaiman: Supplied essential data for the journal articles.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Abdullah, A., Palash, W., Rahman, A., Islam, K., & Alim, A. (2016). Digital image processing analysis using Matlab. *American Journal of Engineering Research (AJER)*, *5*(12), 2320-0936. https://www.ajer.org/

Abdullah, N. A., & Quteishat, A. M. (2015). Wheat seeds classification using multi-layer perceptron artificial neural network. *International Journal of Electronics Communication and Computer Engineering*, *6*(2), 306-309. https://www.ijecce.org/administrator/components/comjresearch/files/publications/IJECCE_3391_Final.pdf

Alaei, A., Nagabhushan, P., Pal, U., & Kimura, F. (2016). An efficient skew estimation technique for scanned documents: An application of piece-wise painting algorithm. *Journal of Pattern Recognition Research*, *1*, 1-14. http://www.jprr.org/

Alzoubi, I. H. E., Sahuri, S. N. S., & Ismail, H. (2023). Teacher and parents persuasion of use of picture exchange communication system to improve attention span for autistic students in abu dhabi autism center. *Al-Hikmah: International Journal of Islamic Studies and Human Sciences*, *6*(1), 53-69. https://doi.org/10.46722/hikmah.v6i1.353

Angadi, S. A., & Kodabagi, M. M. (2009). A texture based methodology for text region extraction from low resolution natural scene images. *International Journal of Image Processing (IJIP)*, *3*(5), 229.

Angulo, J., & Velasco-Forero, S. (2014). Morphological processing of univariate Gaussian distribution-valued images based on Poincaré upper-half plane representation. *Geometric Theory of Information*, 331-366. https://doi.org/10.1007/978-3-319-05317-2_12

Balabanian, F., Sant'Ana da Silva, E., & Pedrini, H. (2017). Image thresholding improved by global optimization methods. *Applied Artificial Intelligence*, *31*(3), 197-208. https://doi.org/10.1080/08839514.2017.1300050

Barlas, P., Adam, S., Chatelain, C., & Paquet, T. (2014). A typed and handwritten text block segmentation system for heterogeneous and complex documents. *In 2014 11th IAPR International Workshop on Document Analysis Systems*, (46-50). IEEE. https://doi.org/10.1109/DAS.2014.39

Babin, D., Pižurica, A., Velicki, L., Matić, V., Galić, I., Leventić, H., ... & Philips, W. (2018). Skeletonization method for vessel delineation of arteriovenous malformation. *Computers in Biology and Medicine*, *93*, 93-105. https://doi.org/10.1016/j.compbiomed.2017.12.011

Babu, K. R., & Sunitha, K. (2011). A new fuzzy Gaussian noise removal method for gray-scale images. *IJCSIT*, *2*(1), 504-511.

Binmakhashen, G. M., & Mahmoud, S. A. (2019). Document layout analysis: A comprehensive survey. *ACM Computing Surveys (CSUR)*, *52*(6), 1-36.

Buchelly, F. J., Mayorca, D., Ballarin, V., & Pastore, J. (2016). Digital image processing techniques applied to pressure analysis and morphological features extraction in footprints. *In Journal of Physics: Conference Series*, *705*(1), 012020. IOP Publishing. https://doi.org/10.1088/1742-6596/705/1/012020

Chaubey, A. K. (2016). Comparison of the local and global thresholding methods in image segmentation. *World Journal of Research and Review*, *2*(1), 1-4. ISSN: 2455-3956.

Chatbri, H., Kameyama, K., & Kwan, P. (2016). A comparative study using contours and skeletons as shape representations for binary image matching. *Pattern Recognition Letters*, *76*, 59-66. https://doi.org/10.1016/j.patrec.2015.04.007

Eyupoglu, C. (2017). Implementation of Bernsen's locally adaptive binarization method for gray scale images. *The Online Journal of Science and Technology*, *7*(2), 68-72. https://www.tojsat.net/journals/tojsat/articles/v07i02/v07i02-10.pdf

Gautam, S., Saxena, P.T., & Trivedi, P. (2018). Low Contrast Color Image Enhancement by Using GLCE with Contrast Stretching. https://api.semanticscholar.org/CorpusID:212479599

Ganesan, P., & Sajiv, G. (2017). A comprehensive study of edge detection for image processing applications. *In 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICIIECS.2017.8275968

Gómez-Aguilar, J. F., & Atangana, A. (2018). Fractional derivatives with the power-law and the Mittag-Leffler kernel applied to the nonlinear Baggs-Freedman model. *Fractal and Fractional*, *2*(1), 10. https://doi.org/10.3390/fractalfract2010010

Granlund, O., & Böhrnsen, K. (2017). Improving character recognition by thresholding natural images (Dissertation). https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-208899

Grover, S., Arora, K., & Mitra, S. K. (2009). Text extraction from document images using edge information. *In 2009 Annual IEEE India Conference*, (pp. 1-4). IEEE. https://doi.org/10.1109/INDCON.2009.5409409

Gundu, N. K., Jadhav, S. M., Kulkarni, T. S., & Kumbhar, A. S. (2014). Text extraction from image and displaying its related information. *International Journal of Scientific and Research Publications*.

Gupta, A. K., Chauhan, S. S., & Shrivastava, M. (2016). Low contrast image enhancement technique by using fuzzy method. *International Journal of Engineering Research and General Science*, *4*(2), 518-526. https://pnrsolution.org/Datacenter/Vol4/Issue2/71.pdf

Gurung, A., & Tamang, S. L. (2019). Image segmentation using multi-threshold technique by histogram sampling. *arxiv preprint arxiv:1909.05084*. https://doi.org/10.48550/arXiv.1909.05084

Gwet, D. L. L., Otesteanu, M., Libouga, I. O., Bitjoka, L., & Popa, G. D. (2018). A review on image segmentation techniques and performance measures. *International Journal of Computer and Information Engineering*, *12*(12), 1107-1117. https://www.isical.ac.in/~sankar/paper/PR_93_NRP_SKP.PDF

Hamad, K., & Mehmet, K. (2016). A detailed analysis of optical character recognition technology. *International Journal of Applied Mathematics Electronics and Computers*, (Special Issue-1), 244-249. https://doi.org/10.18100/ijamec.270374

Heidarysafa, M., Reed, J., Kowsari, K., Leviton, A. C. R., Warren, J. I., & Brown, D. E. (2020). From videos to urls: A multi-browser guide to extract user's behavior with optical character recognition. *In Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference* (*CVC*), *1*(1), (pp. 503-514). Springer International Publishing. https://doi.org/10.1007/978-3-030-17795-9_37

Hoda, R., Salleh, N., Grundy, J., & Tee, H. M. (2017). Systematic literature reviews in agile software development: A tertiary study. *Information and Software Technology*, *85*, 60-70. https://doi.org/10.1016/j.infsof.2017.01.007

Hongbo, Y., & Xia, H. (2014). Histogram modification using grey-level co-occurrence matrix for image contrast enhancement. *IET Image Processing*, *8*(12), 782-793. https://doi.org/10.1049/iet-ipr.2013.0657

Irviantina, S., & Pardosi, I. (2016). Salt and Pepper Noise Removal dengan Spatial Median Filter dan Adaptive Noise Reduction. *Jurnal SIFO Mikroskil*, *17*(2), 127-136. https://doi.org/10.55601/jsm.v17i2.348

Janani, R., & Sheela, T. (2016). Automatic text extraction system for complex images. ISSN: 2278-8727.

Jayasree, M., & Narayanan, N. K. (2015, April). An efficient mixed noise removal technique from gray scale images using noisy pixel modification technique. *In 2015 International Conference on Communications and Signal Processing* (*ICCSP*), (pp. 0336-0339). IEEE. https://doi.org/10.1109/ICCSP.2015.7322901

Jundale, T. A., & Hegadi, R. S. (2015). Skew detection and correction of Devanagari script using Hough transform. *Procedia Computer Science*, *45*, 305-311. https://doi.org/10.1016/j.procs.2015.03.147

Karpagam, A. V., & Manikandan, M. (2019). Text extraction from natural scene images using Renyi entropy. *The Journal of Engineering*, *2019*(8), 5397-5406. https://doi.org/10.1049/joe.2018.5160

Kaur, A., Dhir, R., & Lehal, G. S. (2017). A survey on camera-captured scene text detection and extraction: towards Gurmukhi script. *International Journal of Multimedia Information Retrieval*, *6*, 115-142. https://doi.org/10.1007/s13735-016-0116-5

Karthick, K., Ravindrakumar, K. B., Francis, R., & Ilankannan, S. (2019). Steps involved in text recognition and recent research in OCR; a study. *International Journal of Recent Technology and Engineering*, *8*(1), 2277-3878.

Kejriwal, L., & Singh, I. (2016). A hybrid filtering approach of digital video stabilization for UAV using kalman and low pass filter. *Procedia Computer Science*, *93*, 359-366. https://doi.org/10.1016/j.procs.2016.07.221

Khan, M. F., & Mufti, N. (2016). Comparison of various edge detection filters for ANPR. *In 2016 6th International Conference on Innovative Computing Technology* (*INTECH*), (pp. 306-309). IEEE. https://doi.org/10.1109/INTECH.2016.7845061

Khorsheed, O. K. (2014). Produce low-pass and high-pass image filter in java. *International Journal of Advances in Engineering and Technology*, *7*(3), 712. https://www.proquest.com/docview/1550522151?sourcetype=Scholarly%20Journals

Lore, K. G., Akintayo, A., & Sarkar, S. (2017). LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, *61*, 650-662. https://doi.org/10.1016/j.patcog.2016.06.008

Lou, S., Jiang, X., & Scott, P. J. (2014). Applications of morphological operations in surface metrology and dimensional metrology. *In Journal of Physics: Conference Series*, *483* (1), 012020. IOP Publishing. https://doi.org/10.1088/1742-6596/483/1/012020

Louis, L. (2016). Working principle of Arduino and u sing it. *International Journal of Control, Automation, Communication and Systems (IJCACS)*, *1*(2), 21-29. https://doi.org/10.5121/ijcacs.2016.1203

Long, S., He, X., & Yao, C. (2021). Scene text detection and recognition: The deep learning era. *International Journal of Computer Vision*, *129*(1), 161-184. https://doi.org/10.1007/s11263-020-01369-0

Lu, S., Chen, T., Tian, S., Lim, J. H., & Tan, C. L. (2015). Scene text extraction based on edges and support vector regression. *International Journal on Document Analysis and Recognition (IJDAR)*, *18*, 125-135. https://doi.org/10.1007/s10032-015-0237-z

Mathur, G., & Rikhari, M. S. (2017). Text detection in document images: Highlight on using FAST algorithm. *International Journal of Advanced Engineering Research and Science*, *4*(3), 237108. https://dx.doi.org/10.22161/ijaers.4.3.43

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093-1113. https://doi.org/10.1016/j.asej.2014.04.011

Memon, J., Sami, M., Khan, R. A., & Uddin, M. (2020). Handwritten Optical Character Recognition (OCR): A comprehensive Systematic Literature Review (SLR). *IEEE Access*, *8*, 142642-142668. https://doi.org/10.1109/ACCESS.2020.3012542

Miri, M. S., Abràmoff, M. D., Kwon, Y. H., Sonka, M., & Garvin, M. K. (2017). A machine-learning graph-based approach for 3D segmentation of Bruch's membrane opening from glaucomatous SD-OCT volumes. *Medical Image Analysis*, *39*, 206-217. https://doi.org/10.1016/j.media.2017.04.007

Medjahed, S. A. (2015). A comparative study of feature extraction methods in images classification. *International Journal of Image, Graphics and Signal Processing*, *7*(3), 16. https://doi.org/10.5815/ijigsp.2015.03.03

Nayak, A. A., Venugopala, P. S., Sarojadevi, H., & Chiplunkar, N. N. (2015). An approach to improvise canny edge detection using morphological filters. *International Journal of Computer Applications*, *116*(9).

Negi, S. S., & Bhandari, Y. S. (2014). A hybrid approach to image enhancement using contrast stretching on image sharpening and the analysis of various cases arising using histogram. *In International Conference on Recent Advances and Innovations in Engineering* (*ICRAIE-2014*), (pp. *1*-6). IEEE. https://doi.org/10.1109/ICRAIE.2014.6909232

Nilsson, S., & Lindahl, M. (2016). A literature review to understand the requirements specification's role when developing integrated product service offerings. *Procedia CIRP*, *47*, 150-155. https://doi.org/10.1016/j.procir.2016.03.225

Nguyen, K. C., & Nakagawa, M. (2016). Text-line and character segmentation for offline recognition of handwritten Japanese text. *IEICE Technical Report*, *115*(517), 53-58.

Poovizhi, P. (2014). A study on preprocessing techniques for the character recognition. *International Journal of Open Information Technologies*, *2*(12), 21-24.

Ravi, P., & Ashokkumar, A. (2017). Analysis of various image processing techniques. *International Journal of Advanced Networking and Applications*, *8*(5), 86-89.

Ravi, H., Subramanyam, A. V., & Emmanuel, S. (2015). Spatial domain quantization noise based image filtering detection. *In 2015 IEEE International Conference on Image Processing (ICIP)* (1180-1184). IEEE. https://doi.org/10.1109/ICIP.2015.7350986

Raid, A. M., Khedr, W. M., El-Dosuky, M. A., & Aoud, M. (2014). Image restoration based on morphological operations. *International Journal of Computer Science, Engineering and Information Technology*, *4*(3), 9-21.

Roynard, M., Carlinet, E., & Géraud, T. (2019). An image processing library in modern C++: Getting simplicity and efficiency with generic programming. *In Reproducible Research in Pattern Recognition: Second International Workshop, RRPR 2018, Beijing, China, August 20, 2018, Revised Selected Papers 2* (121-137). Springer International Publishing. https://doi.org/10.1007/978-3-030-23987-9_12

Savitha, G., Venugopal, P. S., & Chiplunkar, N. (2014). An approach for object detection in android device. *In 2014 5th International Conference on Signal and Image Processing* (pp. 9-14). IEEE. https://doi.org/10.1109/ICSIP.2014.6

Sawant, A. S., & Chougule, D. G. (2015). Notice of Removal: Script independent text pre-processing and segmentation for OCR. *In 2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO)* (1-5). IEEE. https://doi.org/10.1109/EESCO.2015.7253643

Sanders, E. D., Ramos Jr, A. S., & Paulino, G. H. (2017). A maximum filter for the ground structure method: An optimization tool to harness multiple structural designs. *Engineering Structures*, *151*, 235-252. https://doi.org/10.1016/j.engstruct.2017.07.064

Scheerlinck, C., Barnes, N., & Mahony, R. (2019). Asynchronous spatial image convolutions for event cameras. *IEEE Robotics and Automation Letters*, *4*(2), 816-822.

Shrestha, S. (2014). Image denoising using new adaptive based median filters. *Arxiv Preprint Arxiv:1410.2175*. https://doi.org/10.48550/arXiv.1410.2175

Sinecen, M. (2016). Digital image processing with MATLAB. *Applications from Engineering with Matlab Concepts*, *1*. ISBN-10: 9789535124597.

Smitha, M. L., Antony, P. J., & Sachin, D. N. (2016). Document image analysis using ImageMagick and tesseract-ocr. *International Advanced Research Journal in Science, Engineering and Technology (IARJSET), 3,* 108-112. ISSN: 2393-8021.

Sultan, A., M., Alturkistani, D., M., Sultan, C., A., Sagr, N., A., Bargawi, H., A., Natto, K., K., ... & Alghamdi, B., H. (2017). Journal Homepage: -www. journalijar. com. ISSN: 2320-5407.

Uysal, A. K., & Gunal, S. (2014). The impact of preprocessing on text classification. *Information Processing and Management*, *50*(1), 104-112. https://doi.org/10.1016/j.ipm.2013.08.006

Valero, S., Morin, D., Inglada, J., Sepulcre, G., Arias, M., Hagolle, O., ... & Koetz, B. (2016). Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions. *Remote Sensing*, *8*(1), 55. https://doi.org/10.3390/rs8010055

Wang, L., Yao, H. W., Wu, Z. X., & Liang, D. (2014). The application of image processing in the electrical fire physical evidence identification. *Procedia Engineering*, *71*, 57-64. https://doi.org/10.1016/j.proeng.2014.04.008

Wick, C., & Puppe, F. (2018). Fully convolutional neural networks for page segmentation of historical document images. *In 2018 13th IAPR International Workshop on Document Analysis Systems* (*DAS*), (pp. 287-292). IEEE. https://doi.org/10.1109/DAS.2018.39

Xie, D. H., Lu, M., Xie, Y. F., Liu, D., & Li, X. (2019). A fast threshold segmentation method for froth image base on the pixel distribution characteristic. *PloS One*, *14*(1), e0210411. https://doi.org/10.1371/journal.pone.0210411

Zaitoun, N. M., & Aqel, M. J. (2015). Survey on image segmentation techniques. *Procedia Computer Science*, *65*, 797-806. https://doi.org/10.1016/j.procs.2015.09.027

Zhang, X., Sugano, Y., & Bulling, A. (2018). Revisiting data normalization for appearance-based gaze estimation. *In Proceedings of the 2018 ACM Symposium on Eye Tracking Research and Applications*, (pp. 1-9). https://doi.org/10.1145/3204493.3204548