الجمهورية الجزائرية الديموقراطية الشعبية

وزارة التعليم العالي و البحث العلمي

University 20 Août 1955 - Skikda

Faculty of Sciences – Computer Science department

Theme

Transformation of Image Graphics to a Table of Values

Directed by:

-INES LASMAR

-IKRAM REHAHLA

Supervised by:

- Dr. MAWLOUD MOSBAH

Reviewed by:

-Dr. BOUCHEHAM BACHIR

Remerciement

On remercie dieu tout le puissant de nous avoir donné la force, la santé, le courage et la patience de pouvoir accomplir ce travail.

Au terme de la rédaction de ce mémoire, ne tenons à remercier notre encadrent : MOSBAH MAWLOUD pour ces précieux conseils et son aide durant toute lapériode du travail.

Un grand merci à toutes nos familles surtout nos parents pour leurs encouragements et leurs suivis avec patience le déroulement de notre projet.

Mr.BOUCHEHAM BACHIR, s'incères remerciement, pour avoir accepté de juger cette modeste contributionet d'avoir consacré du temps à sa lecture.

Nos vifs remerciement vont également aux membres dejury pour l'intérêt qu'ils ont porté à notre projet en acceptant d'examiner notre travail et de l'enrichir par leur proposition.

Merci

Dédicace

Toutes les lettres n'expriment pas complètement comment trouver les mots admiration, tous les mots ne peuvent pas exprimer la gratitude, le respect et la reconnaissance …

En plus, c'est aussi simple que ça.

A mes chers parents

Aucune sincérité ne peut exprimer mon respect,

Mon amour et mon respect éternels dépendent des sacrifices que vous avez consentis pour mon éducation et mon bien-être. Je vous remercie pour les soins et l'amour que vous m'avez témoignés depuis mon enfance et j'espère que vos bénédictions m'accompagneront tout au long de ma vie, même si je le fais. Je ne vous ai jamais suffisamment remboursé.

A mes chers frères

Hana ,Hamza, Amar, je vous souhaitez une vie pleine de bonheur, de réussite, de santé et de bien-être .

Un merci tout particulier à notre unique petit-fils, Ishak, le fauteur de troubles Que Dieu vous protège et vous protège pour vos parents.

A mes chers amis

A mes chères amies Ikram, Amani, Assila et Nessrine que Dieu te protège et te garde pour moi,

À mes amis du master sous le nom de L'Equipe, que Dieu vous accorde la réussite et je vous souhaite un avenir prospère et heureux,

Un merci spécial à ma tante maternelle et ses filles Meriem et Maroua, et ma famille LASMAR chaqun avec son nom,

À mon professeur Mawloud Mosbah,

Un merci tout particulier pour vos efforts et votre aide tout au long du travail

Merci pour tout, je te souhaite le meilleur dans ta vie.

INES

Dédicace

Toutes les lettres n'expriment pas complètement comment trouver les mots admiration, tous les mots ne peuvent pas exprimer la gratitude, le respect et la reconnaissance …

En plus, c'est aussi simple que ça.

A mes chers parents

Aucune sincérité ne peut exprimer mon respect,

Mon amour et mon respect éternels dépendent des sacrifices que vous avez consentis pour mon éducation et mon bien-être. Je vous remercie pour les soins et l'amour que vous m'avez témoignés depuis mon enfance et j'espère que vos bénédictions m'accompagneront tout au long de ma vie, même si je le fais. Je ne vous ai jamais suffisamment remboursé.

A mes chers frères

Iyad,Nour, Malak, je vous souhaitez une vie pleine de bonheur, de réussite, de santé et de bien-être .

Un merci tout particulier à notre unique petit-fils, Mazen, le fauteur de troubles Que Dieu vous protège et vous protège pour vos parents.

A mes chers amis

A mon amie d'enfance Amani, et mes chères amies Ines et Nessrine que Dieu te protège et te garde pour moi,

À mes amis du master sous le nom de L'Equipe, que Dieu vous accorde la réussite et je vous souhaite un avenir prospère et heureux,

Un merci spécial à ma tante paternelle et sa fille Imene, et ma famille Rehahla chaqun avec son nom,

À mon professeur Mawloud Mosbah,

Un merci tout particulier pour vos efforts et votre aide tout au long du travail

Merci pour tout, je te souhaite le meilleur dans ta vie.

IKRAM

Table of Content

List of figures

List of tables

List of abbreviation

RGB: Red, Green, Blue.

HSV: Hue, Saturation, Value.

FCM: Fuzzy C-Means method.

SOFM: Self-Organized Feature Map.

OCR: Optical Character Recognition.

FCN: Fully Convolutional Network.

ANN: Artificial Neural Network.

CNN: Convolutional Neural Network.

GGD: Generalized Gaussian Distribution.

PSNR: Peak Signal Noise Ratio.

MSE: Mean Square Error.

DRD: Distance Reciprocal Distortion.

NRM: Negative Rate Metric.

MPM: Misclassification Penatly Metric.

FOD: First-Order Derivatives.

SOD: Second-Order Derivatives.

GA: Genetic Algorithm.

SVM: Support Vector Machine.

PCA: Principal Component Analysis.

LCE: Local Consistency Error.

BCE: Bidirectional Consistency Error.

EMD: Earth Mover's Distance.

UML: Unified Modelling Language.

Abstract

In this Master dissertation, we deal with an important image processing task which is segmentation through addressing image graphic and trying to shift it to its original table of values. Indeed, for researchers who reads thousands of scientific papers in order to produce their contribution, the task of shifting a graphic image included in *PDF* format documents to its real table of values seems to be a valuable task allowing them to do comparison with real values of other methods and techniques of literature.

Transformation of a graphic image to table of values requires firstly to binarize the image in order to keep only two colours: black and white and discard any other colour belonging to the RGB colour space may be available in the image graphic. Secondly, a segmentation is then necessary in order to extract the essential image graphic features helping to shift them to values.

الملخص

في مذكرة التخرج هذه، نستهدف عملية مهمة من عمليات معالجة الصورة ألا و هي تجزئة الصورة لاستخلاص مختلف المعلومات الموجودة فيها عن طريق محاولة تحويل بيانات مرسومة على شكل مدرجات تكرارية أو منحنيات بيانية إلى جدول قيم عددية بنيت على أساسها هذه البيانات. بالفعل، بالنسبة للباحثين الذين يقرؤون بصفة دورية الآلاف من المقالات بغية المساهمة بكتابة مقالات نوعية، تشكل عملية تحويل بيانات مرسومة موجودة في ملفات قابلة للطباعة إلى جدول قيم عملية بالغة الأهمية تسمح لهم بإجراء مقارنات دقيقة للطرق المقترحة مع ما هو موجود في الأدبيات العلمية.

عملية تحويل بيانات مرسومة إلى جدول قيم عددية يتطلب بداية إجراء عملية تحويل صورة البيانات المرسومة، أين هناك العديد من الألوان، إلى لونين فقط هما الأبيض و الأسود محاولين المحافظة قدر الإمكان على المعلومات الموجودة في الصورة ثم القيام بعملية تجزئة الصورة لاستخراج الخصائص ذات الأهمية البالغة التي بإمكانها أن تعطينا المعلومات التي تساعدنا في استخلاص جدول القيم العددية.

General Introduction

As university under-graduate students, we need to read scientific documents like: papers, articles, theses, and dissertations in order to understand some essential notions and concepts and trying to master the terminology of our field of study. For researchers, belonging to graduation level, they have to read continuously scholars in their scientific field in order to stay in-date with what happens in their field. Moreover, researchers are required to produce themselves higher scientific papers in order to boost science and validate their research. Usually, the scholars read either by us or by researchers are in PDF format. They contain written content and some images of illustration such as curves and histograms. Indeed, researchers who produce scientific scholars prefer to illustrate and present their results in form of curves and histograms to be easy to read by readers. Unfortunately, as researches aiming to produce a paper in the basis of the results included in the published papers, we prefer to find results as tables of values where the results are exact rather than graphical form difficult to extract exact information and doing comparison. In this direction, we aim to design and to implement an automatic application in order to converse and shift a graphical form (either curve or histogram) to a tables of values ready to be used by researchers to do comparison and produce their own publication.

Part 1: State of the Art

Chapter 01: Digital Images

1. Introduction

Digital image as an important multimedia resource has received an important attention in the scientific field especially in pattern recognition and segmentation. In this first chapter of the state of the art, we present some essential and basic definitions related to digital image.

2. Image

An image is a planar reproduction of a scene or an object through painting, sculpture, or an optical instrument. It can be considered as a two-dimensional signal in analogy form (negative, photograph, video, etc.) or digital form (digitized images in various formats, compressed or uncompressed images, or obtained through sensors providing digitized images) [1].

3. Digital Image

A digital image is an image composed of picture elements, also known as pixels, each with finite, discrete quantities of numeric representation for its intensity or gray level that is an output from its two-dimensional functions fed as input by its spatial coordinates denoted with x, y on the x-axis and y-axis, respectively [2]. Depending on whether the image resolution is fixed, it may be of vector or raster type. By itself, the term "digital image" usually refers to raster images or bitmapped images (as opposed to vector images).

4. Image Spaces and Models

There are many spaces and models to represent an image. among of them, we can quote the following [3]:

- **RGB Space**

The RGB (Red, Green, Blue) space is the most commonly used space. This space contains three channels representing the three primary colours of additive synthesis. The values for each channel are integers. These values commonly range from 0 to

255. In reality, the RGB model does not correspond to human colour perception because its colours are correlated[3].

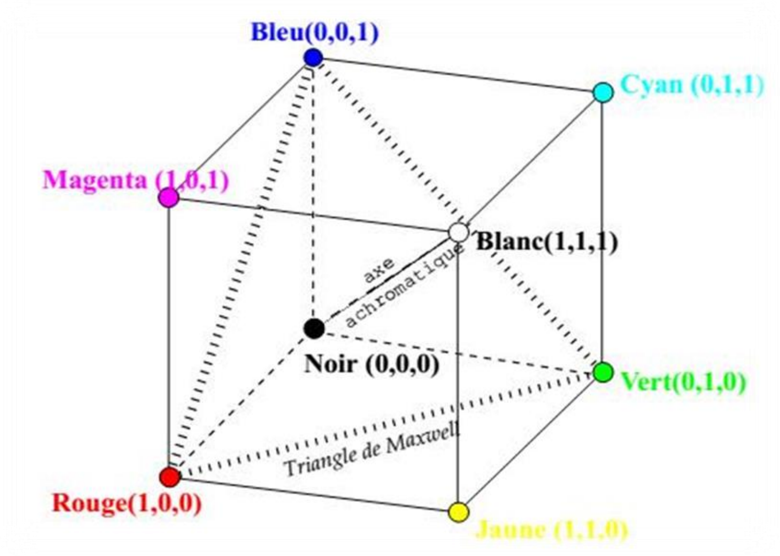

Figure.1 The coordinates of the RGB System [127].

- **Normalized System (r, g, b)**

This system makes it possible to standardize the basic RGB system. The transition from the RGB system to the standardized system (r, g, b) is carried out by dividing the value of each component by the sum of the three components[3]:

$$\begin{cases} r = \dfrac{R}{R+G+B} \\ g = \dfrac{G}{R+G+B} \\ b = \dfrac{B}{R+G+B} \end{cases} \qquad (1)$$

Thus, each normalized colour belongs to the interval [0,1] and the sum of the three new components is equal to 1.

- **XYZ System**

This system was established by the CIR (International Commission on Illumination) to overcome certain disadvantages of the RGB system. This system corresponds to a change of primary colours and is obtained from the RGB system via a pass matrix.

The coefficients of this matrix are determined in relation to an illuminant. In terms of perceptual distance, the XYZ space is not uniform: we perceive more nuances in X and Z than in Y. On the other hand, certain descriptive dimensions of colour: light/dark, pure/faded are not directly accessible. The transition from the RGB system to the XYZ system using the standard illuminant D65 is done by the following matrix calculation[3]:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.430574 & 0.341550 & 0.178325 \\ 0.222015 & 0.706655 & 0.071330 \\ 0.020183 & 0.129553 & 0.939180 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \qquad (2)$$

- **L*u*v System**

The components of this system are calculated from the components of system XYZ as follows:

$$L^* = \begin{cases} 116 \sqrt{\frac{Y}{Y_0}} - 16 \; si \; \frac{Y}{Y_0} > 0.008856 \\ 903.3 \frac{Y}{Y_0} \; si \; \frac{Y}{Y_0} \leq 0.008856 \end{cases} \qquad (3)$$

$$u^* = 13 \, L^*(u' - u'_0) \qquad (4)$$

$$v^* = 13 \, L^*(v' - v'_0) \qquad (5)$$

$$u' = \frac{4X}{X + 15Y + 3Z} \qquad (6)$$

- **HSV System**

The HSV (Hue, Saturation, Value) system is defined by a cylinder that reprents the tent, Saturation and value of colour. The shade H, represented by angle between 0 and 360° , indicates the colour (red, blue, yellow, green, etc.). The saturation S given corresponds to the purity of the colour. The V value gives information on the light intensity indicating whether the colour is light or dark. The components of this system are calculated as follows[3]:

$$V = \max(R, G, B) \qquad (7)$$

$$S = \frac{255(V - min(R, G, B))}{V} \; si \; V \neq 0{,}0 \; sinon \qquad (8)$$

$$H = \begin{cases} \frac{60(G-B)}{S} si\ V = R \\ 180 + \frac{60(B-R)}{S}\ si\ V = G \\ 240 + \frac{60(R-G)}{S}\ si\ V = B \end{cases} \quad (9)$$

- **YC_rC_b System**

This colour system is used in digital television images. It is also used by the old MPEG standards (1,2 and 4). This system is defined by a linear transformation of the RGB space as follows[3]:

$$\begin{cases} Y \\ C_r \\ C_b \end{cases} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.500 & -0.419 & -0.018 \\ -0.169 & -0.331 & 0.500 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (10)$$

5. Image Quantization

Colour quantization is the process of reducing the number of colours in a digital image. The main objective of quantization process is that significant information should be preserved while reducing the colour of an image. In other words, quantization process shouldn't cause significant information loss in the image [4]. We can consider gray-level and binarization as special cases of quantization.

Several colour quantization methods have been proposed in the literature and these can be categorized into two groups: splitting and clustering-based methods. Moreover, there is another categorization regarding the colours considered into quantization process, namely: quantization based on fixed palette [5] and quantization based on dominant colours[6].

- **Splitting Approach**

The splitting methods divide colour spaces into disjoint regions according to some criteria. This process is repeated until getting the desired number of regions. After that, each region is chosen to be the palette colour. Therefore, the set of clusters is constructed by a complete binary tree where each parent has two children [7]. The time consumption of the splitting methods is satisfactory in general. However, it is difficult to find an optimal solution through the splitting decisions[4].

- **Clustering Approach**

The clustering-based methods have the disadvantage of high time consumption and they can be adversely affected from the initial conditions in general when comparing with the splitting methods; however, these methods generally perform better than the splitting methods in terms of finding the optimal solution on account of the fact that they can modify the centroids if it is needed (Omran, 2004). In fact, the clustering-based applications are widely used in some applications such as, spatial and medical image processing[4].

6. Conclusion

In this first chapter, we have given some definitions of some concepts tied to digital image and image processing, namely: colour models and quantization.

Chapter 02: Image Binarization

1. Introduction

Due to the widespread and the relative low price of new technological tools including scanners and cameras that capture images such as smart phones, it is so easy, for everyone, to photograph any scene he/she meets. Although other wired and wireless alternatives for information exchange and communication, capturing the scene with camera seems to be very simple and easy.

Capturing text documents, with cameras, generates what we know as a (print-written) image. These images may be also generated when we scan historical archive documents as hand-written images. There are many challenges addressed especially in handwritten image binarization, such as faint characters, bleed-through, smudges, non-uniform illumination, and large background ink stains [8]. Each of tools for generating image, either a camera or a scanner, has its proper problems. Unfortunately, we cannot treat a image, may be open in an image editor, as the original text document, may be open in a text editor, where we can search, copy, and paste words, phrases, and paragraphs. All the text, in a image, is considered unfortunately as just a set of pixels.

Converting a image to its original text aims to treat, process, and manipulate the text automatically, by the machine, taking profit of operations furnished by any simple text editor and any tool for automatic processing of natural language. The global context of this problem is then a image recognition and understanding.

The rest of the chapter is organized as follows: section 2 presents the approaches and their basic methods of the literature. Section 3 deals with noises may degrade image documents and filters may be used in order to enhance the quality of the image. Section 4 talks about the evaluation process while section 5 shows the experimental results as given in the different considered references.

2. Image Binarization: Approaches and Methods

One of the most crucial steps of pre-processing of images subjected to further text recognition is their binarization, which influences significantly obtained OCR results [9]. By definition, the binarization (also known thresholding or foreground-background separation) of an image means labelling its text as text (foreground) or background. Background is, by its self, defined very vaguely, and it may contain unwanted information which may even carry some form of textual data. The motivation of research related with image binarization and recognition is not only the possibility of preserving the cultural heritage and discovering some historical facts but also potential applications of the developed algorithms in some other areas of industry [9]. Image binarization is then an active area of research in computer vision where many algorithms have been proposed throughout the literature. These algorithms can be divided into two major categories: clustering-based methods [10] and threshold-based methods [10] and may be hybrid methods such as [11]. Clustering-based methods usually use model-based features to differentiate text from background and classify it [10]. Fuzzy classification [12] and recursive segmentation in the PCA space [13] are two examples of clustering-based methods. In the threshold-based category, the most famous method, and a pioneering one, is Otsu's method [14], which is based on maximizing the separation between two pre-assumed classes. A previous work [15] classified the methods into six categories: histogram-based methods, clustering-based methods, entropy-based methods, foreground attribute-based methods, spatial binarization methods, and locally adaptive methods. It is worthy to note that there are other methods coming from novel disciplines like optimization such as [15] in which artificial bee colony (ABC) algorithm has been invoked as a clustering algorithm in order to segment the text from the images. The ABC algorithm helps the proposed algorithm to find the best threshold value to get high quality binarized image according to performance evaluation measures.

Consulting literature reveals that there are some works of second position such as: surveys, reviews, and overviews that deal with image binarization [8][14][16][22][23][24][25]. In [8], authors have provided a comprehensive review, discussed, and tested seven types of binarization method on *H-DIBCO12*. As a conclusion, they have given some critics towards the considered techniques of the literature, in the hope to improve performances, like considering traditional methods for contrast enhancement as a post-processing step. In [16], author has considered three major groups for image binarization, namely: threshold-based approach, optimization-based approach, and classification-based approach. Although that authors in [14] did not present any information about evaluation, they have given an essential idea about the assessment step which is indirect evaluation saying that the quality of any binarization method may appear later in the degree of effectiveness of the recognition applications coming later that utilize the output of the binarization as an input. Away from the classical categorization of binarization methods considered in the majority of works, authors in [14] divide binarization methods, based on their chronological appearance, into two categories: foundations and recent works. In [22], authors have presented a comparison of several image binarization techniques according to *H-DIBCO 2013* dataset. In [23], authors have presented a review and a comparison study of some various image binarization techniques according to *DIBCO-2009* and *DIBCO-2010* datasets. In [24], authors have considered four categories for image binarization, namely: (1) global thresholding-based binarization methods, (2) local/adaptive thresholding-based methods, (3) hybrid thresholding-based methods, and (4) machine learning-based methods. In [25], authors have provided a comprehensive review of the field of historical image binarization. They have presented a table which summaries and categorises the various techniques introduced in the last decade. Among these categorizations, we can quote: (1) global threshold-based techniques, (2) local threshold-based techniques, (3) edge-based techniques, (3) image transform-based techniques, (4) mixture model-based techniques, (5) conditional random fields-based techniques, (6) game theory-based techniques, (7) shallow machine learning-based techniques, (8) deep learning-based techniques, (9) supervised tuning-based techniques, and (10) unsupervised tuning-based techniques.

In this work, we consider three approaches for image binarization: (1) image binarization based on clustering, (2) image binarization based on threshold, (3) binarization based on neural networks, and (4) combination of binarization techniques.

2.1 Image Binarization based on Clustering

Clustering and classification are both fundamental tasks in data mining. Classification is used mostly as a supervised learning method, clustering for unsupervised learning. There are some works of literature where clustering (unsupervised classification) algorithms are used for image binarization.

2.1.1 Binarization based on k-means

Some works have considered k-means for image binarization [27]. Indeed, in [27], authors have proposed a k-means based clustering technique for adaptive binarization of degraded images.

2.1.2 Binarization based on Fuzzy Classification

Many works have used fuzzy C-means method (FCM) in the binarization of images such as in [28], [29], [30], [31], [32]. Fuzzy classification is generally considered with local thresholding methods in order to fasten them. Indeed, in [28], authors have proposed a method for fuzzy binarization of digital document claimed suitable for binarization of bulrring or badly illuminated documents and can be easily modified to accommodate any type of spatial characteristics. After training, the neurons of the output competition layer of the considered Self-Organized Feature Map (SOFM) define two bi-level classes. Using content of these classes, fuzzy membership functions are obtained that are next used with the fuzzy C-means (FCM) algorithm in order to reduce the character-blurring problem. In [29], the proposed consists of three parts. In the first part, the image pre-processing operation has been done before the binarization process to enhance image quality through applying both contrast stretching and mean filter. The second part is to apply the binarization algorithm on the image that has undergone an image pre-processing operation. By applying the Fuzzy C-means algorithm to the images, the images are converted to binary images and divided into two components, which is text and background. The last step of the proposed method is to perform Deghost operation to remove "ghost" entities that may have appeared on the image. In [30], authors have presented a fast and competent, yet simple binarization technique that uses a Fuzzy C-means based global thresholding approach, aided by background separation. The proposed method uses a superset of foreground regions to correctly assess background of the image. Background is estimated based on a sliding interpolation window of variable dimension, judged by appraising the nature of text stroke. Ultimately, a global approach is undertaken to binarize the background-separated normalized and enhanced image by clustering the pixels using Fuzzy C-means. In [31], authors have proposed a local threshold binarization method using fast fuzzy C-means clustering. Historical images with non-uniform background, stains, faded ink are first processed by removing the background using in-painting based method. Then fuzzy C-means clustering is used to cluster out the pixels into three main clusters: sure text pixels, sure background pixels and confused pixels which may or may not be labelled as text. In [32], authors have proposed a binarization algorithm, called NFCM, which is a combination of Niblack algorithm and FCM algorithm. They have claimed that their algorithm is good not only at preserving the character stokes, but also at alleviating the ghost artifacts.

2.2 Image Binarization based on Threshold

Image binarization consists commonly to convert a color or gray image into a binary image, where the text and background pixels are marked in black and white respectively [17]. A successful binary result preserves meaningful information while discarding noisy information [18].

Thresholding is a simple yet powerful technique to separate the object of interest from the background [38]. Image thresholding has found various applications in many computer vision and graphics applications such as image analysis. In an image, if the objects are clearly lighter (or darker) than the background it is natural to separate them by setting a threshold [38].

The algorithm of binarization-based threshold is simple. Given a fixed global threshold value T (from 0 to 255), if the intensity value of any pixel of an input image is more than T, the pixel is set to white otherwise it is black. If the source is a colour image, it first has to be converted to gray level.

Unfortunately, a fixed global value for the threshold may degrade drastically the quality of the image, owing to the variation of luminosity over the different regions of the input image generating then many misclassified pixels. This fact leads researchers to think about considering either a dynamic threshold [19], an adaptive threshold as done in [20], or even considering a multiple local thresholds for the same image [17][21]. Due to the importance of thresholding, binarization methods are categorized depending on which principal criteria they consider in calculating the threshold [14]. The core problem, tied to binarization-based threshold, to be asked then, is how to designate the threshold.

Consulting the literature reveals a lot of works to select the binarization threshold for any image in general and image document implementing optical character recognition (OCR) tools in particular. There is the method based on clustering-analysis [26], methods based on entropy [33][34], methods based on image variance [20][35], method based on image contrast [18], methods relying on stroke structural symmetry of strokes [13], and those based on texture features [36]. In [37], authors have conducted a comparison for some binarization methods for historical archive documents.

It is worthy to note that for images with poor and non-uniform illumination, adaptive thresholding is required to separate the objects of interest from the background [38].

2.2.1 Global Thresholding

In global thresholding approach, a single threshold value is determined to be applied to the whole image in order to put aside the pixels under consideration into foreground and background. In the following sub-sections, we quote the basic global-threshold binarization methods.

a) Gradient based Thresholding

In [38], a new approach to create an adaptive threshold surface is proposed to segment an image. The technique is inspired by the Yanowitz's method [39] and is improved upon by the introduction of a simpler and more accurate threshold surface. The method uses gradient based thresholding by constructing a threshold surface. It contains three important phases: first, construct the inverse image $T(i,j)$, second obtaining the k value between [-255, 255], and finally applying binarization to separate object and background.

$$result(i,j) = \begin{cases} 0, (object) & if\ I(i,j) < T(i,j) + k_0 \\ 255, (background) & if\ I(i,j) > T(i,j) + k_0 \end{cases} \quad (11)$$

Where $I(i,j)$ is the intensity of the original image and k_0 is the minimum sum of absolute difference intensity.

b) Otsu Method

In [26], a very simple method with an optimal threshold and straightforward extension to multi-thresholding problem is introduced. The threshold is obtained automatically based on global variance and between-class variance. In the non-uniform image, Otsu assumes the image contains two areas: dark and bright in order to purpose final algorithm. Otsu thresholding is determined then by:

$$k = \frac{\sigma^2 B}{\sigma^2 G} \quad (12)$$

Where: K a threshold value, $\sigma^2 B$ are a global variance of the entire image, and $\sigma^2 G$ between-class variance.

2.2.2 Local Thresholding

In case of local or adaptive thresholding, same threshold is never used throughout the entirety of the image. Rather, the properties of a pixel and its neighbours in a sub-image help to determine the threshold. In the following sub-sections, we give some basic local-threshold binarization methods.

a) Niblack Method

The main purpose of Niblack method [35] is to set the threshold value based on local standard deviation and local mean. The threshold for each pixel is determined by:

$$T(x, y) = m(x, y) + k\delta(x, y) \quad (13)$$

Where: local standard deviation $\delta(x, y)$ and local mean $m(x, y)$ are determined by 80×80 windowing size and standard k value is (-0.2). It is worthy to note that this method does not work correctly if the image suffers from non-uniform illumination.

b) Nick Method

In [40], a sliding window based local thresholding technique, trying to improve the Niblack method by shifting down the binarization threshold, is presented. The threshold is obtained based on the following equation:

$$T(x, y) = m + k\sqrt{\frac{(I^2 - m^2)}{N}} \quad (14)$$

Where: the k factor is similar with Niblack and the windowing size is defined as 15×15, while I and m represent the intensity pixel and mean of grey scale image. N represents the image size.

c) Bradley Method

In [41], authors have presented an improvement and an extension of Wellner's method [42] which is robust to illumination changes within the image. The main idea in Wellner's algorithm is that each pixel is compared to an average of the surrounding pixels while the key idea of the Bradley algorithm is that every image's pixel is set to black if its brightness is T percent lower than the average brightness of surrounding pixels in the window of the specified size, otherwise it is set to white. The default windowing size is 15×15 and T is 10. T is given as follows:

$$T = m\left(1 - \frac{k}{100}\right) \quad (15)$$

d) Bernsen Method

The Bernsen algorithm [43] is based on the estimation of a local threshold value for each pixel. This value is assigned the local threshold value only if the difference between the lowest and the highest grey level value is bigger than threshold k. Otherwise; it is assumed that the window region contains pixels of one class (foreground or background). The default windowing size (w) is 3×3 and k is 15. The final equation as follows:

$$T(x,y) = \frac{Z_{max}+Z_{min}}{2} \quad (16)$$

Where: Z_{min} and Z_{max} are the lowest and highest grey level pixel values.

e) Local Adaptive Thresholding

Local Adaptive Thresholding [44] is a basic simple algorithm to separate the foreground from the background with non-uniform illumination. For each pixel in the image, a threshold has to be calculated. If the pixel value is below the threshold, it is set to the background value, otherwise, it assumes the foreground value. The default local windowing size (w) is 15×15 and local threshold (T) is a 0.05. The threshold is given as follows:

$$T = \frac{max+min}{2} \quad (17)$$

2.3 Binarization based on neural networks

Some works such as [45], [46], [47] [48] have adopted neural networks as an artificial intelligence tool for dealing with image binarization. Indeed, in [45], authors have formulated binarization issue as a pixel classification learning task, to address a multiple image scales, including full resolution, where a fully convolutional network (FCN) architecture has been applied. In [46], a foreground-background separation technique, based on artificial neural networks (ANN), has been applied on old documents with a variety of degradation. The idea has been to train an ANN on a set of pairs of original images and their respective ideal black and white ones relying on global and local information. In [47], authors have considered a back-propagation neural network to directly classify image pixels according to their neighbourhood. In [48], authors have proposed the utilization of convolutional neural networks (CNN) to identify foreground pixels using novel input-generated multichannel images. To create the images, the original source image is decomposed into wavelet sub-bands. Then, the original image is approximated by each sub-band separately, and finally, the multichannel image is constituted by arranging the original source image as the first channel and the approximated image by each sub-band as the remaining channels. Two scenarios are considered, that is, two-channel and four-channel images, and then fed into two types of CNN architectures, namely, single and multiple streams. CNNs used in the architectures are the three popular networks: U-net, SegNet, and DeepLabv3+.

2.4 Combination of Binarization Techniques

Some works of literature have preferred to consider combination of some existed methods [49], [59] of binarization rather than of introducing new binarization technique. Indeed, in [49], authors have built a system that takes advantages of the benefits of a set of selected binarization techniques by combining their results using a Kohonen self-organizing neural network. They claimed that the built system is suitable to classify pixels that have high vagueness. That is, pixels which belong to edges, shadow areas and generally pixels that cannot be easily classified as foreground or background pixels. In [59], authors have presented a technique based on hybrid thresholding combining the advantages of global and local methods. Two stages have been included. In the first stage, global thresholding is applied on the entire image and two different thresholds are determined from which the most of image pixels are classified into foreground or background. In the second stage, the remaining pixels are assigned to foreground or background classes based on local analysis. In this stage, several local thresholding methods are combined and the final binary value of each remaining pixel is chosen as the most probable one.

3. Noises and Filters

3.1 Image Degradation and Noises

Author in [16] have categorized image degradation and noise into two categories: (1) noise from native sources which appears due to aging, quality of paper and ink used, and (2) noise coming from auxiliary sources or external factors like malfunctioning of scanning devices, lighting condition during acquisition, misplacement of the document during scanning. The noise coming from native sources causes issues like:

- ***Ink-bleeding or bleed through***: which appears when both sides of the page are used to write and the ink seeps through one side and spreads over to the other side.
- ***Show through or smear***: which appears because of the ink impression of one side appears on the other side. It often creates dark spots on the page.
- ***Faint text***: it appears in documents written or printed using low quality ink. This is because as the time grows the ink starts to shrink. The quality of paper used may also be the reason for that.
- ***Deterioration of document***: takes place due to aging, poor storage, mishandling, natural calamity and other environmental condition. This type of noise causes presence of dark spot and other artefacts on the images.

The noise coming from auxiliary sources causes problems such as:

- ***Uneven illumination***: which occurs in the light microscopy images as in optical imaging where the incident light decreases drastically along the path.
- ***Contrast variation***: this appears due to the environment under which the image acquisition process is carried out.
- ***Blurring effect***: observed in images are of two types: motion blurring, appearing due to the relative motion of the camera and object or sudden rapid movement of the camera, and out-of-focus blurring, appearing where some of the points are in focus while others are not during acquisition.
- ***Skew***: this appears either due to the misplacement of the document during scanning or due the writing style. This can occur either at the page level or at the region level or both. Generally, for the printed documents, page level skew is found whereas in handwritten documents region level skew is observed mostly.

3.2 Some Filters as Pre-processing, Processing and Post-processing Steps for Binarization

Pre-processing consists of preparing the image that suffers from many problems such as shadow, noisy areas from scanning, document aging, ink blot, non-uniform illumination, ink seepage, smudges, smear, strain, bleed-through, background ink-stain, and fading while post-processing, considered commonly as a final step of the binarization process, consists of eliminating eventual remaining noise such as ghosts, improving the quality of text regions, constructing the broken lines, and filling of possible breaks, gaps or holes.

There are a lot of filters considered as pre-processing or post-processing steps for image binarization operation. In this section, we quote the essential ones that are frequently used.

a) Gabor Filter

In [72], authors have presented a simple method suitable for a software and hardware implementation based on Gabor filters for the binarization of ancient degraded documents. The Gabor filter bank is designed by taking into account the degradation type the document based on the un-blind protocol. First, the image is pre-processed using a Wiener filter to smooth the degradation. Subsequently, the binarization threshold is estimated using texture features, such as the mean and the standard deviation, extracted from the respective original image and the filtered image.

Furthermore, a new protocol, namely un-blind protocol, is proposed for estimating the standard deviation according to the degradation type for setting the optimal parameters of the Gabor filter such as the central frequency and the number of angles.

b) Generalized Gaussian Distribution

In [9], authors have confirmed the use of GGD (Generalized Gaussian Distribution) histogram approximation with the use of the Monte Carlo method for pre-processing of degraded images before binarization and further analysis. Indeed, through extracting the parameters of GGD, distortions may be modelled and removed, enhancing the quality of input data for further thresholding and text recognition. Due to an efficient use of the Monte Carlo method, the overall processing time has been shortened and the proposed approach-based GGD may be further combined with other binarization algorithms not considered in [9].

c) Wienner Filter

In [50] authors have used adaptive Wienner Filter, based on local statistics estimated from a local neighbourhood around each pixel, as a pre-processing step to de-noise the image. The use of a low-pass Wienner filter [51] has proved efficient for the elimination of noisy areas, smoothing of back-ground texture as well as contrast enhancement between background and text areas. The Wienner filter is commonly used in filtering theory for image restoration.

d) Shrink and Swell Filtering

In [50] and [73] authors have considered Shrink and Swell filtering. A Shrink filter is used to remove noise from the background through converting a foreground pixel surrounded by background pixels to a background pixel while a Swell filter is used to fill possible breaks, gaps or holes in the foreground via assuming a background pixel among foreground pixels as a hole in the object to be filled up with the foreground pixel value.

e) Local Entropy Filter

In [74], authors have considered the local entropy filter as one of the pre-processing steps for adaptive image binarization of unevenly illuminated images subjected to further optical text recognition.

f) Median Filter

In [75], Gaussian and median filtering are used in order to improve the final binarized output. Gaussian filter is used for further enhance the output and median filter is applied to remove noises.

g) Bilateral Filter

In [76], authors have considered bilateral filter which first introduced by Aurich and Weule [77] under the name “nonlinear Gaussian Filter”. It was later rediscovered by Tomasi and Manduchi [78] who called it the “bilateral filter”. The bilateral filter is a technique to smoothen images while preserving their edges. The weight assigned to each neighbour decreases with both the distance values among pixels of the image plane (the spatial domain S) and the distance on the intensity axis (the range domain R).

h) Kalman Filter

In [79], authors have proposed a two-step approach, called KFBin, for the binarization of images based on the Kalman filtering technique. According to [80], kalman filter is probably the most widely used technique in statistical signal processing, that processes data samples recursively in a very effective manner. Kalman filter is used in the literature for image de-noising, image reconstruction, video de-noising, and object tracking. In [79], authors have used kalman filter to map any multi-channel input images to two outputs images, whereby one output contains the enhanced foreground component and the other contains the background component.

i) Swell Filter

In [81], authors have used swell filter, as a post-processing step, to scan the entire binary image and each white pixel is examined.

Unfortunately and to the best of our knowledge, some interesting filters, such as guided filtering [82] and Bayesian filter [83], have not been well considered, up to now, for image binarization. Guided filtering, as bilateral filter, shares the property of edge-preserving smoothing with a better behaviour near the edge [84]. To note that in [55] and [57], where many methods of the literature have been compared, many filters such as median filter, Gaussian filter, Guided filter, Bilateral filter, and Wiener filter have been considered by the methods made in competition.

j) Curvelet Transform

In [85], the Curvelet transform and Otsu’s method were combined to binarize the non-uniform illuminated images. The non-uniform illumination image was decomposed by the Curvelet and the Curvelet coefficients were enhanced by non-linear functions. The reconstructed image was processed by Otsu’s binarization method.

k) Contrast Stretching Filter

Contrast stretching (often called normalization) is a simple image enhancement technique that attempts to improve the contrast in an image by stretching the range of intensity values it contains to span a desired range values. In [15], authors have considered, as a first phase, stretching the intensity level of the image by contrast stretching filter and removing the noise by image cleaning algorithm.

4) Evaluation of Binarization

In this section, we deal with evaluation in binarization through presenting the different metrics considered in the literature as well as the various considered datasets.

4.1 Metrics

There exist a lot of metrics to evaluate the performance of a binarization system. In this section, we present the essential metrics may be used in the literature.

a) Precision

Precision value represents the true pixels binarized divided by the total pixels binarized. It is given as follows:

$$precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (18)$$

Where: TruePositive means the set of pixels that are foreground in both ground truth (GT) and the binarized image. FlasePositive denotes the set of pixels identified as foreground in the binarized image but are background in the ground truth image.

b) Recall

Recall value represents the true pixels binarized divided by the total number of true pixels. It is given as follows:

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (19)$$

Where: TruePositive means the set of pixels that are foreground in both ground truth and binarized image. FalseNegative denotes the set of pixels identified as background in binarized image but they are foreground in ground truth image.

c) F-measure

F-measure is the harmonic measure of Precision and Recall. Its value should be high for better results.

$$F_\beta = (1+\beta^2) * \frac{Precision*Recall}{\beta^2*Precision+Recall} \quad (20)$$

With $\beta = 1$, the standard F-score is obtained as follows:

$$F_1 = \frac{2*Precision*Recall}{(Precision+Recall)} \quad (21)$$

d) Pseudo F-measure

The pseudo F-measure formula is the same of that F-measure with the exception that the ground truth is skeletonised and the pixels in the skeleton are used as foreground pixels in the calculation of pseudo F-measure.

e) PSNR

PSNR (Peak Signal Noise Ratio) is used to check the similarity between two images. It is used for images having noise. PSNR is given as follows:

$$PSNR = 10log\left(\frac{C^2}{MSE}\right) \quad (22)$$

Where:

$$MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(I(x,y)-I'(x,y)\right)^2}{MN} \quad (23)$$

C is a constant and MSE (Mean Square Error) depicts the difference between the distorted image and the original one. PSNR is a measure of how close is an image to another. The higher the value of PSNR is, the higher the similarity between them is.

f) DRD

The Distance Reciprocal Distortion Metric (DRD) has been used to measure the visual distortion in binary images [52]. It properly correlates with the human visual perception and it measures the distortion for all the S flipped pixels as follows:

$$DRD = \frac{\sum_{k=1}^{S} DRD_k}{NUBN} \quad (24)$$

Where NUBN is the number of the non-uniform (not all black or white pixels) 8x8 blocks in the GT image, and DRD_k is the distortion of the k-th flipped pixel that is calculated using a 5x5 normalized weight matrix W_{Nm} as defined in [47]. DRD_k equals to the weighted sum of the pixels in the 5x5 block of the GT that differ from the centered k^{th} flipped pixel at (x, y) in the binarization result image B.

$$DRD_k = \sum_{i=-2}^{2} \sum_{j=-2}^{2} |GT_k(i,j) - B_k(x,y)| \times W_{Nm}(i,j) \quad (25)$$

g) Pseudo Recall

The pseudo Recall formula is the same of that Recall with the exception that the ground truth is skeletonised and the pixels in the skeleton are used as foreground pixels in the calculation of pseudo Recall.

h) NRM

NRM (Negative Rate Metric) is calculated using the unmatched pixels between the binarized image and the ground truth image. It is given as follows:

$$NRM = \frac{R_{FN} + R_{FP}}{2} \quad (26)$$

Where:

$R_{FN} = \frac{FN}{FN+TP}$ and $R_{FP} = \frac{FP}{FP+TN}$

R_{FN} is the rate of false negatives and R_{FP} is the rate of false positives. Lesser the value of NRM is, better the result of binarization is.

i) MPM

MPM (Misclassification Penalty Metric) is used to evaluate how the binarized image constitutes the contour of the ground truth image. It is given as follows:

$$MPM = \frac{MP_{FN} + MP_{FP}}{2} \quad (27)$$

Where:

$MP_{FN} = \frac{\sum_{i=1}^{N_{FN}} d_{FN}^{i}}{D}$ and $MP_{FP} = \frac{\sum_{j=1}^{N_{FP}} d_{FP}^{j}}{D}$

d_{FN}^{i} and d_{FP}^{i} represent respectively the i^{th} false negative and false positive pixel. Normalization factor is given by D, the sum of overall pixel-to-contour distances of the GT object. Smaller the value of MPM better is the quality of the algorithm.

j) Accuracy

$$Accuracy = \frac{TruePositive+TrueNegative}{TruePositive+TrueNegative+FalsePositive+FalseNegative} \quad (28)$$

k) Specificity

$$Specificity = \frac{TrueNegative}{TrueNegative+FalsePositive} \quad (29)$$

l) Mis-classification Error (ME)

$$mis = \frac{No\ of\ incorrectly\ predicted\ samples}{Total\ Number\ of\ Images} * 100 \quad (30)$$

m) Sensitivity

$$Sensitivity = \frac{TruePositive}{TruePositive+FalseNegative} * 100 \quad (31)$$

3.2 Datasets

As depicted in Table below, we consider here 12 datasets. For more information about image datasets, readers should to ask references [86] and [87].

Table 1 The considered datasets

Image Dataset	Number of images	Its web site to be downloaded
DIBCO2011	16	http://utopia.duth.gr/~ipratika/DIBCO2011/benchmark
H-DIBCO2012	14	http://utopia.duth.gr/~ipratika/HDIBCO2012/benchmark
DIBCO2013	16	http://utopia.duth.gr/~ipratika/DIBCO2013/benchmark
HDIBCO2014	10	http://users.iit.demokritos.gr/~bgat/HDIBCO2014/benchmark
DIBCO201	10	http://vc.ee.duth.gr/h-dibco2016/benchmark
DIBCO2017	20	http://vc.ee.duth.gr/dibco2017/benchmark/
DIBCO2018	10	http://vc.ee.duth.gr/h-dibco2018/benchmark/
Harvard Dataset	813	http://ocp.hul.harvard.edu/ihp/manuscripts.html

4 Experimental Results and Discussions

In this section, the experimental results and some associated discussions are given.

5.1 Experimental Results

In this sub-section, we give the results given in the different considered references. Table 2 yields the different considered datasets (12 datasets) and the associated references that use these datasets. Table 3 provides the best algorithm of the literature, with its associated reference, over the different datasets and the different metrics.

Table 2 the considered references over the considered datasets.

Dataset	References	Number of references
DIBCO 2009	[9], [10], [53], [54], [55], [56], [64], [66], [67], [68], [17], [70], [72]	14
DIBCO 2010	[9], [56], [63], [67], [17], [70], [72]	8
DIBCO 2011	[9], [56], [67], [17], [70], [71], [72]	8
DIBCO 2012	[9], [59], [62], [17], [70], [72]	6
DIBCO 2013	[57], [17], [71], [72]	4
DIBCO 2014	[9], [61], [17], [71]	4
DIBCO 2016	[9], [60], [65], [69], [17], [71], [8]	7
DIBCO 2017	[9], [57], [65]	3
DIBCO 2018	[57], [65]	2
DIBCO 2019	[58]	1
Synthesis	[59]	1
Harvard dataset	[15]	1

Table 3 The best algorithms of the literature over the different datasets and the different metrics

<table>
<tr><td>Dataset</td><td>The best algorithm</td><td>The reference</td><td>The considered metric</td><td>Number of considered methods in the considered reference</td></tr>
<tr><td rowspan="2">DIBCO 2019</td><td>Non-parametric binarization method based on ensemble of clustering algorithms (10b in [58])</td><td>[84] Considered in [58]</td><td>F-measure
Pseudo F-measure</td><td rowspan="2">24</td></tr>
<tr><td>LADDERNET (2a in [58])</td><td>inspired from [88] and considered in [58]</td><td>PSNR
DRD</td></tr>
<tr><td>DIBCO 2018</td><td>Proposed Local+Global</td><td>[65]</td><td>F-measure
Pseudo F-measure
PSNR</td><td>7</td></tr>
</table>

			DRD	
DIBCO 2017	Proposed Local+Global	[65]	F-measure PSNR DRD	
DIBCO 2016	Proposed Local+Global	[65]	F-measure Pseudo F-measure PSNR	
	Hierarchical deep supervised network	[71]	DRD	15
DIBCO 2014	Structural Symmetry of Strokes	[17]	F-measure	10
	Hierarchical Deep Supervised Network	[71]	Pseudo F-measure	16
	Hierarchical Deep Supervised Network with the global prediction	[71]	PSNR DRD	
DIBCO 2013	Hierarchical deep supervised network	[71]	F-measure	18
	Structural Symmetry of Strokes	[17]	Pseudo F-measure	10
	Entry 3	[67] Considered in [71]	PSNR	18
	Wolf	[89] Considered in [72]	DRD	6
DIBCO 2012	Lelore	[90] Considered in [17]	F-measure	9
	Structural Symmetry of Strokes	[17]	Pseudo F-measure	9
	Nick after combination of several thresolding techniques	[40] Considered in [59] with adaptation	PSNR	11
	Wolf	[89] Considered in [72]	DRD	6
DIBCO 2011	Oracle	Considered in [67]	F-measure	5
	Structural Symmetry of	[17]	Pseudo F-measure	10

	Strokes			
	Lelore	[90] Considered in [17]	PSNR	
	Wolf	[89] Considered in [72]	DRD	6
DIBCO 2010	Oracle	Considered in [67]	F-measure	5
	Numerated by 2: Adaptive binarization with removal of false objects and accurate local region-based active contour	Considered in [63]	Pseudo F-measure	17
	Adaptive Image Contrast Method	[56]	PSNR	8
	Howe	[91] Considered in [17]	DRD	9
DIBCO 2009	Howe	[91] Considered in [17]	F-measure	10
	Lelore	[90] Considered in [17]	Pseudo F-measure	
	Howe	[91] Considered in [17]	PSNR DRD	
Harvard dataset	ABCO (Artificial Bee Colony Optimizer)	Considered in [15]	PSNR NRM DRD MPM	4

5.2 Discussions

According Table 2 and Table 3, resulted from the results given in the considered references, we can give the following points:

- ✓ There are two kinds of references: a reference as a competition where a large number of methods of the literature are experimented and a reference where a novel algorithm is introduced and is compared with some limited number of methods of the literature.
- ✓ *DIBCO 2009* dataset (the oldest database) is the most considered one among all the used databases with 14 references.

- ✓ The greatest number of methods utilized in the considered references is 24 methods experimented in [59] on *DIBCO 2019* followed respectively by [72] and [64] with 18 and 17 methods for each reference.
- ✓ *Hierarchical deep supervised network* method is the best one against other 17 methods over 4 metrics, namely: *F-measure* (on *DIBCO 2013*), *Pseudo F-measure*, *PSNR*, and *DRD* (on *DIBCO 2014*).
- ✓ *Wolf* method seems to be good according to *DRD* metric at least over *DIBCO 2011*, *DIBCO 2012*, and *DIBCO 2013* and at least against over 5 methods of the literature.
- ✓ Lelore and Howe are among the best methods as given in Table 3 where their names figure over three to four datasets and over three to four metrics.

5 Conclusion

In this paper, we have given an experimental comparison study of the different techniques of image binarization in terms of effectiveness considering a large number of metrics and datasets. In addition, an overview of the various approaches, metrics, types of noises, and the considered datasets. According to the comparative study, it is clear that the machine learning approach (or neural network approach) seems globally to be the best approach although the consuming time of learning required. As a future work, we plan to conduct another comparison study in terms of efficiency which goes beyond the scope of this paper.

Chapter 03: Image Segmentation

1. Introduction

With the proliferation of computer technology, the demand for digital image processing has surged, finding extensive applications across various sectors such as organizations, businesses, and medicine. Image segmentation serves as a crucial tool for analyzing images to extract pertinent information. Over time, numerous algorithms and techniques have been developed in the realm of image segmentation, making it a prominent task within machine vision. Machine vision empowers computers to perceive real-world problems akin to humans and respond accordingly to solve them. Hence, it becomes imperative to devise techniques applicable to image segmentation. The emergence of modern segmentation methods like instance, semantic, and panoptic segmentation has propelled the evolution of machine vision[92].

2. Image Segmentation

Image segmentation is the process of partitioning an image into group of pixels grounded on some homogeneous features like colour or intensity to extract some meaningful information [93], [94]. Image segmentation as a pre-processing phase is a part of almost all computer vision system in the real world complex applications, ranging from object extraction to medical images, satellite images, video and traffic surveillance system etc [95], [96], [97], [98], [99], [100]. Steps for segmentation shown in Figure 2.

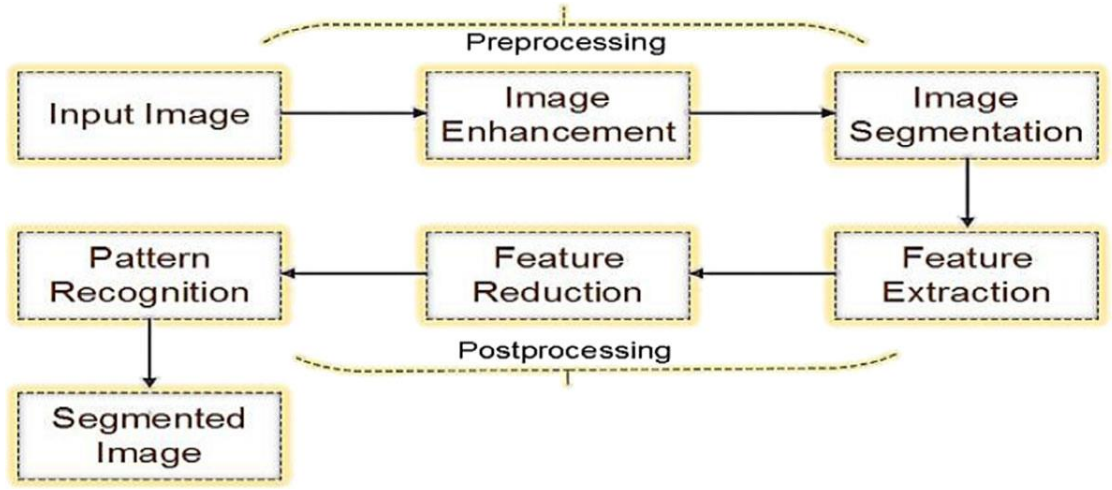

Figure 2. Segmentation steps [101].

3. Approaches and Techniques

Several segmentation methods exist, with specific examples outlined in Figure 3, while comprehensive details regarding each method are elaborated upon in the next sections.

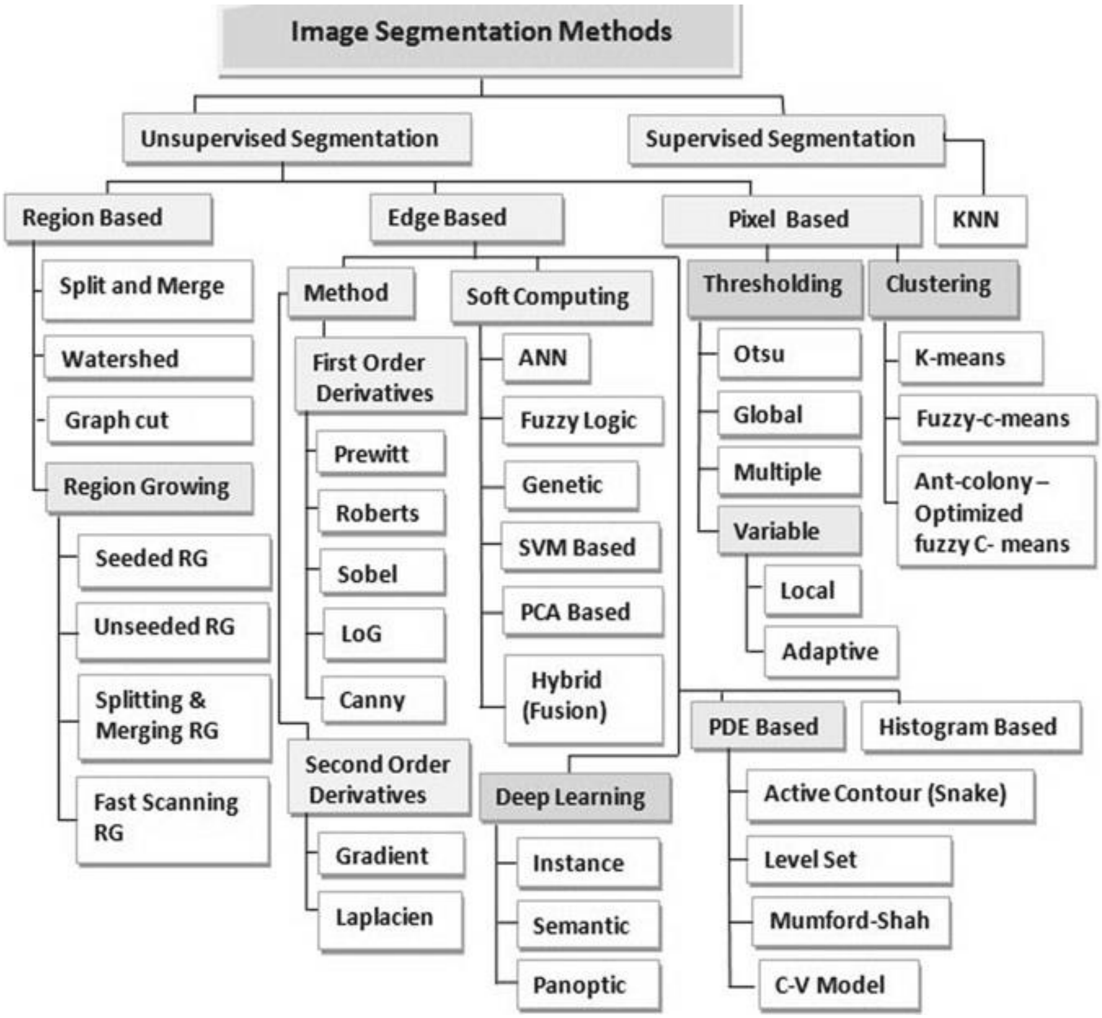

Figure 3. Segmentation techniques [102].

Broadly, there are two main approaches [102]: unsupervised image segmentation, and supervised image segmentation.

3.1 Unsupervised Image Segmentation

It does not need any previous knowledge about the image and it does not require training data, hence it is also called nearest neighbour segmentation.

3.1.1 Region-Based Segmentation Method

In this method every pixel of an image is being checked in incremental fashion and it does require using adequate thresholding method. Types of region-based segmentation are discussed as follows:

- **Split and Merge**

This method is applied to the whole images. Split region is a top-down process and in this method instead of selecting kernel points, image can be partitioned into a set of random unconnected regions and then combine the regions [103]. Advantage: It is an easy method. Disadvantage*:* It may lead to over segmentation.

- **Watershed Segmentation**

It is a morphological gradient-based segmentation that partitions segmentation areas drained by different river system. It is a very effective and accepted segmentation technique and it can offer more precise segmentation with little computational resources [95].Watershed is excellent to handle objects with blurry edges and uniformed background [104]. Advantage: No seed is required and the processed regions are connected and it can also find best possible boundaries [105]. Its range is very good. Disadvantage: It is sensitive to noise and in-homogeneity [104].

- **Graph Cut**

It has got very much concentration as it uses both boundary and regional information. Moreover, graphcut-based method is effective because it can get globally optimal result, and it is also applicable to segment dimensional image. There are three types of method; that is, interactive-based, speed up-based and shape prior-based graph cut [105]. Advantage: It is efficient to the natural image without any prior information. Disadvantage: Memory utilization increases as the image size increases.

- **Region Growing**

In this approach of the segmentation, initial seed pixel is identified in the image and then the region is growing based on the pixel value. It is typical sequential region segmentation, and its fundamental thought is to have similar character of the pixels together to shape a region[106]. It has four types, that is, seeded, unseeded, region splitting and merging and fast scanning. Advantage: Various criteria can be applied at once and it provides best outcome with a smaller amount noise. Disadvantages: It may produce over segmentation image due to noise.

3.1.2 Edge-Based or Boundary-Based

This type of method transforms an image into edge images by altering the gray value in the images. An object is made up of various part of diverse color levels, and lack of continuity shows the presence of edges. An edge is the local changes of image intensity value and an edge appears in the boundary between two parts of the image [106]. It has various types as discussed below.

- **First-Order Derivatives (FOD)**

Derivative is defined as differences and it is used to detect abrupt and local changes to the images, so FOD and second-order derivatives (SOD) are suitable for this purpose. FOD is
good to select the strongest edges by thresholding the gradient magnitude. Advantage: FOD uses simple pixel differences for calculating changes in gray intensity. Disadvantages: It is very responsive to noise and generate very thicker edges. FOD can be represented using the following Eq. (32):

$$df/dx = f^{'}(x) = f(x+1) - f(x) \quad (32)$$

Prewitt filter, Roberts Filter, Sobel Filter, LoG (Laplacian of a Gaussian) Filter, Canny filter, and so on are types of FOD filters.

- **Second-Order Derivatives (SOD)**

It is more refined methods for automatized edge detection. It does boost fine details (including noise) as compared to FOD [107]. In both the steps, the ramp edge that is performed in second derivative has opposite sign (positive to negative or negative to positive) as it transits inside and outside of an edge, and this dualedge effect is an important characteristic of SOD [107]. Sign of SOD is also applied to find out whether an edge is transition from dark to light (positive second derivatives) or from light to dark (*negative second derivatives*) where the sign is observed as we move into the edge [107]. SOD can be represented using the following Eq. (33):

$$d^2x/dx^2 = f'(x) = f(x+1) + f(x-1) - 2f(x) \quad (33)$$

- **Gradient Operator**

This operator identifies edges after searching for the minimum and maximum value in first derivatives of an image [108]. Advantage: It has a robust average response in area of ramps and steps transition than does the Laplacian [107]. Disadvantage: The response of this operator to fine details and noise is less than the Laplacian and can be lesser further after smoothing gradient with an average filter [107].

Laplacian Operator

It is a two-dimensional measure of second derivative of an image. It detects regions having quick intensity variation and therefore most often used for edge detection. Zero crossing is the point where the Laplacian changes sign (where value passes through zero) to find edges[108].

- **Soft-Computing-Based Segmentation**

Artificial Neural Network (ANN)

It is a processing system which is inspired by the human neurons. ANN is made up of using various layers where the first layer is input layer in which input is given using
matrix and the last layer is output layer and in between layers are called hidden layers. There are various activation functions, like sigmoid, binary, rectified linear unit (ReLU) and so on. ANN learns through activation function, and the learning rate must be between [0, 1] and the error is suggested to be below 0.1. *Advantage:* It works very well in some field where other segmentation method does not. It performs very well in face recognition. *Disadvantage:* It needs so much training data and time to train the ANN model.

Fuzzy Logic

L A Zadeh invented fuzzy set theory in 1965 to deal with imprecise information. Let Z be a collection of objects and genetic object of Z represented by z; like, Z = {z}. Fuzzy set A in Z is illustrated by membership function, μA(z), that is attached to each element of Z and is a real number in the interval of 0 and 1. The value of μA*(z)* in z represents the ranking of membership of element in A [107]. A fuzzy set is formally defined as given by Eq. (34).

$$A = \{z, \mu A(z) \mid z \in Z\} \quad (34)$$

Genetic Algorithm (GA)

GA is eventually applied for solving the optimization problem and it selects the initial population and then it repetitively changes a population of individual. In every step, it chooses population randomly from the existing population to be used as parents and then parents produce the children; used for the next generation; and this process continues until we get the population repeated. GA segments an image by using an optimization function without using any threshold values and it can produce more accurate results than basic segmentation methods and it provides faster convergence to the best possible solution [109]. Genetic algorithm involves three steps, that is, *selection*, *crossover* and mutation.

- **Support Vector Machine (SVM)-Based Segmentation**

SVM is derived from the support vector classifier. It is derived from enlarging the feature space using kernels. The kernel approach is just a well organized computational strategy for applying a nonlinear boundary between classes [110]. Both of the classifications (binary and multiclass) are used in SVM [110].

- **Principal Component Analysis (PCA)-Based Segmentation**

PCA is mainly used for dimension reduction of an object and used for filtering of different features and overthrow by analyzing the principal component. Karim T. F. et al. (2010) [111] used PCA-based segmentation for face recognition and it has been observed that they got good result instead of having some shortcoming of the system [111].

- **Hybrid (Fusion)-Based Segmentation**

Hybrid means combination of two or more than two techniques for the image segmentation, like PCA has been used with K-means clustering for doing the segmentation, and histogram has been used with K-means cluster for segmentation [112].

Figure 4. a Image **b** after applying instance segmentation [113]. **c** Image **d** after applying semantic segmentation [114]. **e** After applying panoptic segmentation [115]

- **Deep Learning-Based Segmentation**
  - **Instance Segmentation:** Instance segmentation is a futuristic segmentation and its objective is to discover particular object in an image and to create a mask around the desired objects shown in Figure 4b [113]. It produces mask in place of a bounding box [113]. Instance segmentation does not aim to label every pixel in the image unlike semantic segmentation, which tries to assign label to every pixel in image [113]. In Figure 4 one sheep is different from the other sheep and both have been assigned different colours.
  - **Semantic Segmentation:** This type of segmentation is important for the task where the analysis of image is very much important. The basic idea is to associate each and every pixel of an image with a class label (like car, road, ocean or flower) [114]. It assigns same colour label to the objects belonging to the same class, as shown in Figure 4d. Semantic segmentation is used in self-driving cars, industrial inspection, medical imaging and so on.
  - **Panoptic Segmentation:** Panoptic segmentation combines both instance and semantic segmentation. Here the basic idea is to merge different modules of instance and semantic segmentation, but it introduces more challenges (see Figure 4e) [115].
  - **Histogram-Based Segmentation:** It is fundamentally a region growing method, where histogram features are used for growing. Histogram feature calculation is most important for algorithms and it also saves time of computation [116]. It is applied on digital image intensity level between 0 and L–1.
  - **PDE-Based Image Segmentation:** PDE is based on mathematical equation. PDE uses two or more independent variables, an unknown function that is dependent on variables and partial derivation of the unknown function with respect to the independent variables. Methods based on PDE are: Active contour model, Mumford Shah model, Level Set and C–V model.

3.1.2 Pixel (Point)-Based Segmentation: It is a formal technique of segmentation and is done based on the pixel information, like gray-level value and so on.

- **Thresholding-Based Segmentation:** It is a popular segmentation technique; it separates foreground from the background, where objects are seen very clear [117]. Using thresholding, one can extract the object from the background

[107]. It has four types, that is, Otsu, global, multiple, variable thresholding and so on.

- **Clustering:** Clustering is a method of separating an image into different subgroups based on some similarity. Clustering is of two types: *supervised* and *unsupervised clustering.* K-means is one of the most frequently used algorithms.
  - **Fuzzy-c Means (FCM):** FCM is the most frequently used unsupervised method and it is mostly used in medical imaging [118]. It allows each data to pertain to more than two clusters [118]. *Advantage:* It takes the advantage of fuzzy logic to form clusters, so it is suitable for segmenting the complex images like medical image and so on.
  - **Ant-Colony Optimized Fuzzy-c Means:** This is the modified version of fuzzy-c means clustering. Here, number of center and center value are obtained by ant-colony optimization and then the fuzzy-c means is used for classifying the remote sensing image. Visual range of ant is small, so in the beginning search is blind. However, the ant leaves pheromone on the path [119].

3.2 Supervised methods

These are methods in which the classes are known a priori before performing the identification operation of the image elements. They require a training phase on the representative sample to learn the characteristics of each class and another phase to decide the membership of a pixel to one class or another.

Among these methods, we can mention: Bayesian segmentation, Markov field segmentation [120], neural networks[121], etc.

4. Evaluation

Every scientific work should be evaluated in order to measure the quality and the performance of the considered or the introduced methods and algorithms. Even in segmentation, the process of evaluation is important to judge the quality of the considered tools. Broadly, there are two approaches of evaluation: manual evaluation established by expert humans and automatic evaluation done by the machine. The first one is subjective while the second one is objective.

In order to establish evaluation, we have to run our segmentation model on a set of images whose we know in prior the ideal output. This set of images is known as a benchmark or a dataset. Moreover, in order to make comparison between the output of the model and the desired output (included in the benchmark), we need to use a metric which relies on a feature of some features.

In the following sub-sections, we details the different benchmarks and the various metrics we can find in literature to evaluate the proposed methods and algorithms.

4.1 Datasets

Consulting the literature reveals that there are many kinds of images: there are 2D images considered usually in medical for diagnosis or 3D images what we call range images (or robotic images).

Table 4 some datasets with their description.

Name of the Dataset	Its address	Description
Peng and Varshney	https://doi.org/10.1016/j.cviu.2014.11.004	took the segmentation algorithm Mean Square Error (MSE) as an example and developed a systematic method for evaluating the lower bound in the statistical estimation framework.
Kitti	http://www.cvlibs.net/datasets/kitti/	Suite of vision components for an autonomous driving platform. The object detection dataset contains 7481 training images annotated with 3D bounding boxes. A full description of the annotations can be

		found in the readme of the object development kit readme on the Kitti homepage. Pixel-level annotation of a subset of images from the dataset is provided by Ref.
Flores FC, Lotufo RdA	https://doi.org/10.1109/SIBGRAPI.2008.22	Benchmark for quantitative evaluation of assisted object segmentation Methods to image sequences.

4.2 Metrics

Evaluation of image segmentation differs considerably from the binary foreground/background segmentation evaluation problem, in that the correctness of the two class boundary localization is not the only quantity to be measured [123]. This derives from the presence of an arbitrary number of regions in both the reference segmentation and the segmentation to be evaluated. An evaluation metric is desired to take into account the following effects [123]:

- **Over-segmentation.** A region of the reference is represented by two or more regions in the examined segmentation.
- **Under-segmentation.** Two or more regions of the reference are represented by a single region in the examined segmentation.

Under-segmentation is considered to be as a much more serious problem as it is easier to recover true segments through a merging process after over segmentation rather than trying to split an heterogeneous region [123]. One desirable property of a good evaluation measure is to accommodate refinement only in regions that human segmenters could find ambiguous and to penalize differences in refinements elsewhere [123]. In addition to being tolerant to refinement, any evaluation measure should also be robust to noise along region boundaries and tolerant to different number of segments in each partition [123].

4.2.1 Region-Based Evaluation

The region-based scheme evaluates the segmentation accuracy in the number of regions, the locations and the sizes. A region-based evaluation between two segmented images can be defined as the total amount of differences between corresponding regions[123].

- **Hamming Distance**

Huang and Dom [124] introduced the concept of directional Hamming distance between two segmentations, denoted by $S_1 \Rightarrow S_2$. Let S and R be two segmentations. They began by establishing the correspondence between each region of S with a region of R such that $s_i \cap r_i$ is maximized. The directional Hamming distance from S to R is defined as:

$$D_H(S \Rightarrow R) = \sum_{r_i \in R} \sum_{s_k \neq s_j, s_k \cap r_i \neq 0} | r_i \cap s_k \qquad (35)$$

where |·| denote the size of a set. Therefore, *DH* ($S \Rightarrow R$) is the total area under the intersections between all $ri \in R$ and their non-maximal intersected regions from *S*. A region based evaluation measure based on normalized Hamming distance is defined as

$$p = 1 - \frac{D_H(S \Rightarrow R) + D_H(R \Rightarrow S)}{2 \times | S |} \qquad (36)$$

where |*S*| is the image size and $p \in [0, 1]$. The smaller the degree of mismatch, the closer the *p* is to one.

- **Local Consistency Error**

To compensate for the difference in granularity while comparing segmentations, many measures allow label refinement uniformly through the image. D. Martin's thesis [125] proposed an error measure to quantify the consistency between image segmentations of differing granularities - *Local Consistency Error* (LCE) that allows labelling refinement between segmentation and ground truth.

$$LCE(S, R, p_i) = \frac{1}{N} \sum_i min\{E(S, R, p_i), E(R, S, p_i)\} \qquad (37)$$

where *E* (*S*, *R*, *p*) measures the degree to which two segmentations agree at pixel *p*, and *N* is the size of region where pixel *p* belongs.

Note that the LCE is an error measure, with a score 0 meaning no error and a score 1 meaning maximum error. Since LCE is tolerant to refinement, it is only meaningful if the two segmentations have similar number of segments. As observed by Martin in [125], there are two segmentations that give zero error for LCE - one pixel per segment, and one segment for the whole image.

- **Bidirectional Consistency Error**

To overcome the problem of degenerate segmentations, Martin adapted the LCE formula and proposed a measure that penalizes dissimilarity between segmentations proportional to the degree of region overlap. If we replace the pixelwise minimum with a maximum, we get a measure that does not tolerate refinement at all. The *Bidirectional Consistency Error* (BCE) is defined as:

$$BCE(\,S,R,p_i) = \frac{1}{N}\sum_{i} max\,\{(S,R,p_i),E(R,S,p_i)\}. \qquad (38)$$

- **Partition Distance Measure**

Cardoso and Corte-Real [126] proposed a new discrepancy measure - *partition distance* $(dsym)$ defined as: "given two partitions *P* and *Q* of *S*, the partition distance is the minimum number of elements that must be deleted from *S*, so that the two induced partitions (*P* and *Q* restricted to the remaining elements) are identical". $dsym\ (Q,P) = 0$ means that no points need to be removed from *S* to make the partitions equal, i.e., when $Q = P$.

4.2.2 Boundary-Based Evaluation

Boundary-based approach evaluates segmentation in terms of both localization and shape accuracy of extracted regions boundaries[123].

- **Distance Distribution Signatures**

Huang and Dom in [124] presented a boundary performance evaluation scheme based on the distance between distribution signatures that represent boundary pints of two segmentation masks. Let *BS* represent the boundary point set derived from the segmentation and *BR* the boundary ground truth. A distance distribution signature from the set *BS* to the set *BR* of boundary points, denoted $DB(Bs,BR)$, is a discrete function whose distribution characterizes the discrepancy, measure in distance, from *BS* to *BR*. Define the distance from *x* in set *BS* to *BR* as the minimum absolute distance from all the points in *BR*, *d* (*x*, *BR*) = min *{dE* (*x, y*)*}* , ∀*y* ∈ *BR*, where *dE* denotes the Euclidean distance between points *x* and *y*. The discrepancy between *BS* and *BR* is described by the shape of the signature, which is commonly measured by its mean and standard deviation. As a rule, *DB* (*Bs, BR*) with a near-zero mean and a small standard deviation indicates high between segmentation masks. Since these measures are not normalized, we cannot determine which segmentation is the most desirable. We introduce a modification to the distance distribution signature of Huang and Dom, in order to normalize the result between 0 and 1. Doing *d* (*x, BR*) = min *{dE* (*x, y*), *c},* where the *c* value sets an upper limit for the error, the two boundary distances could be combined in a framework similar to the one presented in Eq. (39):

$$b = 1 - \frac{D_B(D_S, B_R) + D_B(B_R, B_S)}{c \times (| R | + | S |)} \qquad (39)$$

where |*R*| and |*S*| are the number of boundary points in reference mask and segmented mask, respectively.

- **Precision-Recall Measures**

Martin in his thesis [128], propose the use of *precision* and *recall* values to characterize the agreement between the oriented boundary edge elements (termed *edgels*) of region boundaries of two segmentations. Given two segmentations, *S* and *R*, where *S* is the result of segmentation and *R* is the ground truth, precision is proportional to the fraction of edgels from *S* that matches with the ground truth *R*, and recall is proportional to the fraction of edgels from *R* for which a suitable match was found in *S*. Precision measure is defined as follows:

$$Precision = \frac{Matched\ (S, R)}{| S |} \quad Recall = \frac{Matched\ (R, S)}{| R |} \qquad (40)$$

where |*S*| and |*R*| are the total amount of boundary pixels. In probabilistic terms, precision is the probability that the result is valid, and recall is the probability that the ground truth data was detected. A low recall value is typically the result of under-segmentation and indicates failure to capture salient image structure. Precision is low when there is significant over segmentation, or when a large number of boundary pixels have greater localization errors than some threshold δmax. Precision and recall measures have been used in the information retrieval systems for a long time [129]. However, the interpretation of the precision and recall for evaluation of segmentation are a little different from the evaluation of retrieval systems. In retrieval, the aim is to get a high precision for all values of recall. However in image segmentation, the aim is to get both high precision and high recall. The two statistics may be distilled into a single figure of merit:

$$F = \frac{PR}{\alpha R + (1-\alpha)P}, \qquad (41)$$

where α determines the relative importance of each term. Following [128], α is selected as 0.5, expressing no preference for either. The main advantage of using precision and recall for the evaluation of segmentation results is that we can compare not only the segmentations produced by different algorithms, but also the results produced by the same algorithm using different input parameters. However, since these measures are not tolerant to refinement, it is possible for two segmentations that are perfect mutual refinements of each other to have very low precision and recall scores.

- **Earth Mover's Distance**

The concept of using the Earth Mover's Distance (EMD) to measure perceptual similarity between images was first explored by Peleg *et al.* in [130] for the purpose of measuring distance between two grey-scale images. More recently EMD has been used for image retrieval [131].

EMD evaluates dissimilarity between two distributions or *signatures* in some feature space where a distance measure between single features is given. The EMD between two distributions is given by the minimal sum of costs incurred to move all the individual points between the signatures. Let $P = \{(p1, wp1), ..., (pm, wpm)\}$ be the first signature with m pixels, where pi is the pixel representative and wpi is the weight of the pixel; the second

signature with n pixels is represented by $Q = \{(q1, wq1), ..., (qn, wqn)\}$; and $D = [dij]$ the distance matrix where dij is the distance between two contour points' image coordinates pi and qj . The flow fij is the amount of weight moved from pi to qj . The EMD is defined as the work normalized by the total flow fij , that minimizes the overall cost:

$$EMD(P,Q) = \frac{\sum_i \sum_j f_{ij} d_{ij}}{\sum_i \sum_j fij} \qquad (42)$$

As pointed by Rubner *et al* [11], if two weighted point sets have unequal total weights, EMD is not a true metric. It is desirable for robust matching to allow point sets with varying total weights and cardinalities. In order to embed two sets of contour features with different total weights, we simulate equal weights by adding the appropriate number of points, to the lower weight set, with a penalty of maximal distance. Since normalizing signatures, with the same total weight do not affect their EMD, we made $i,j\ fij = 1$. Equation (42) becomes,

$$EMD(P,Q) = \sum_i \sum_j f_{ij} d_{ij} \qquad (43)$$

subject to the following constraints: $f_{ij} \geq 0, \sum_j f_{ij} = w_{pi}\ and\ \sum_i f_{ij} = w_{qj}$. As a measure of distance for the EMD ground distance we use

$$d_{ij} = 1 - e^{-\frac{||p_i - q_j||}{\alpha}} \qquad (44)$$

where $pi - qj$ is the Euclidean distance between pi and qj and α is used in order to accept some deformation resulted from manual segmentation of ground truth. The exponential map limits the effect of large distances, which otherwise dominate the result.

5. Conclusion

Image segmentation stands as a pivotal and foundational stage in digital image processing. It is evident that numerous segmentation methods emerge regularly, yet none can be universally deemed superior. Therefore, the choice of segmentation method greatly hinges on the specific problem at hand. Recently, we observe the advent of contemporary segmentation techniques, such as semantic and panoptic segmentation.

Part 2: Our Contribution

Chapter 04: Requirement Analysis and Design

1. Introduction

In this chapter, we deal with analysis requirement and design of our system with focusing on the different steps of the development process through providing progressively more details and presenting the functional architecture as well as the various diagrams allowing to well define our future prototype.

Our aim is to design and implement an automatic system helping to firstly extract a graphical image from a scientific PDF scholar and proceeding to convert it to its original table of values.

We start firstly this chapter with presenting the motivation of our system another to pass, after that, to analysis requirement as well as architectural and detail design.

2. Requirement analysis

2.1 Our approach

Our aim is to design and implement a system which is able to convert and transform a graphical image in the form of histogram or curve extracted from a scientific *PDF* scholar. Our system starts then with extracting this graphical from the entire document then transforms it from multi-colours image to the binarised form considering only two colours, namely: black and white. Binarisation goes here beyond the scope of our work. Therefore, we just implement one method from those existing in the literature without giving more interest to this task.

Our work belongs then to image segmentation field but the regions or the objects to be recognized are special here. Indeed, we have to designate the top levels of the different histograms and the absis axis. After that, we proceed to compute the number of pixels from the absis axis and the different top levels.

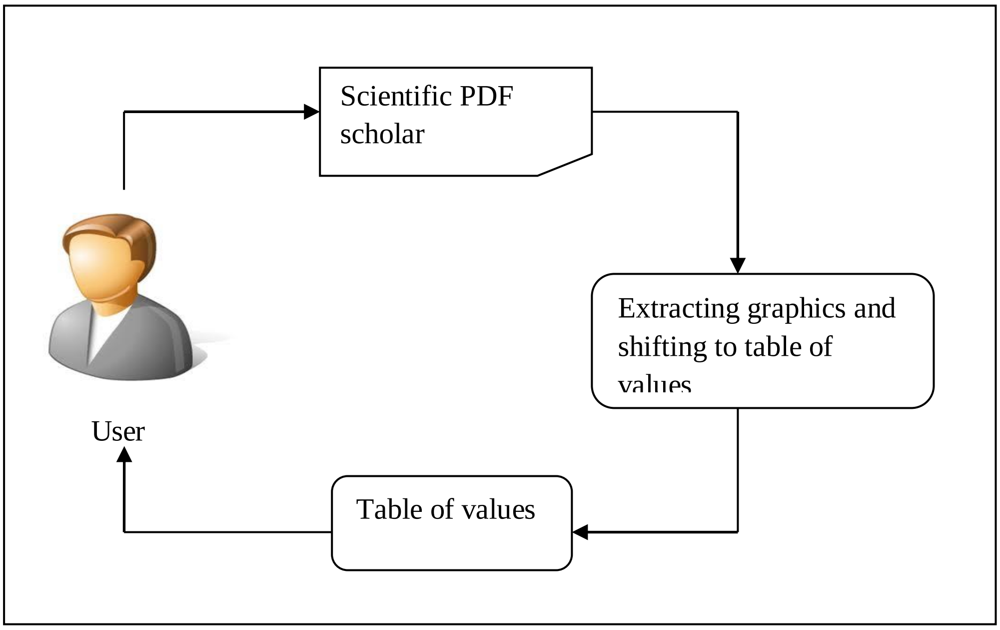

Figure 5. General Architectural of our System.

2.2 Motivation

The idea of our work comes from the observation that although the graphics (histograms and curves) included in scientific documents are clear from human being point of view, they are unfortunately ambiguous to be exploited in other publications. Indeed, it is difficult for us as researchers to read exactly the implicit values included in the graphics. An application helping to extract the exact table of values from the graphics seems to be necessary.

Our application is implemented then to answer this matter. As depicted in Fig.6 , the execution of our application is as follows: firstly, the user submits his/her PDF scientific paper to our system. The system proceeds to extract the graphical image through discarding transcription. This graphical image is binarized, as pre-processing step, considering one of some binarization methods of the literature. This binarized image is explored pixel by pixel from left to right and from top down. The purpose is to find the bin width of the graphic. This one is the number of black pixels in the same line (straightforward black pixels) and usually repeated along the graphic. Once the bin width is recognized, we start to recognize the frequency of the different bars in the histogram. These ones are the first straight black pixels between two straight white pixels (we consider here edge segmentation). They are coloured as red. The next purpose now is to recognize the absis axis. We explore the graphic image from left to right and from bottom up. We look for the Red colour. Once we localize the first Red pixel (belonging to the little frequency). For this one, we start to look for the absis axis. It is when we find the first black pixel. Absis axis should be coloured also in Red. Every pixel in black should be shifted to Red. As such, we get a graphic image with three colours: Red, black and White. Only two modified colours especially the red. Now the graphic image is Ready. We start to measure the different bars (each one from its frequency to the absis axis straightforward). Now, we look for the vertical axis. We explore the graphic image from left to right and from top to down, we look for a pixel coloured in Red. The first Red pixel is the beginning of the vertical absis vertically. The next purpose now is to find a number in order to convert the measures (from the top level to axis) from pixels to decimal values.

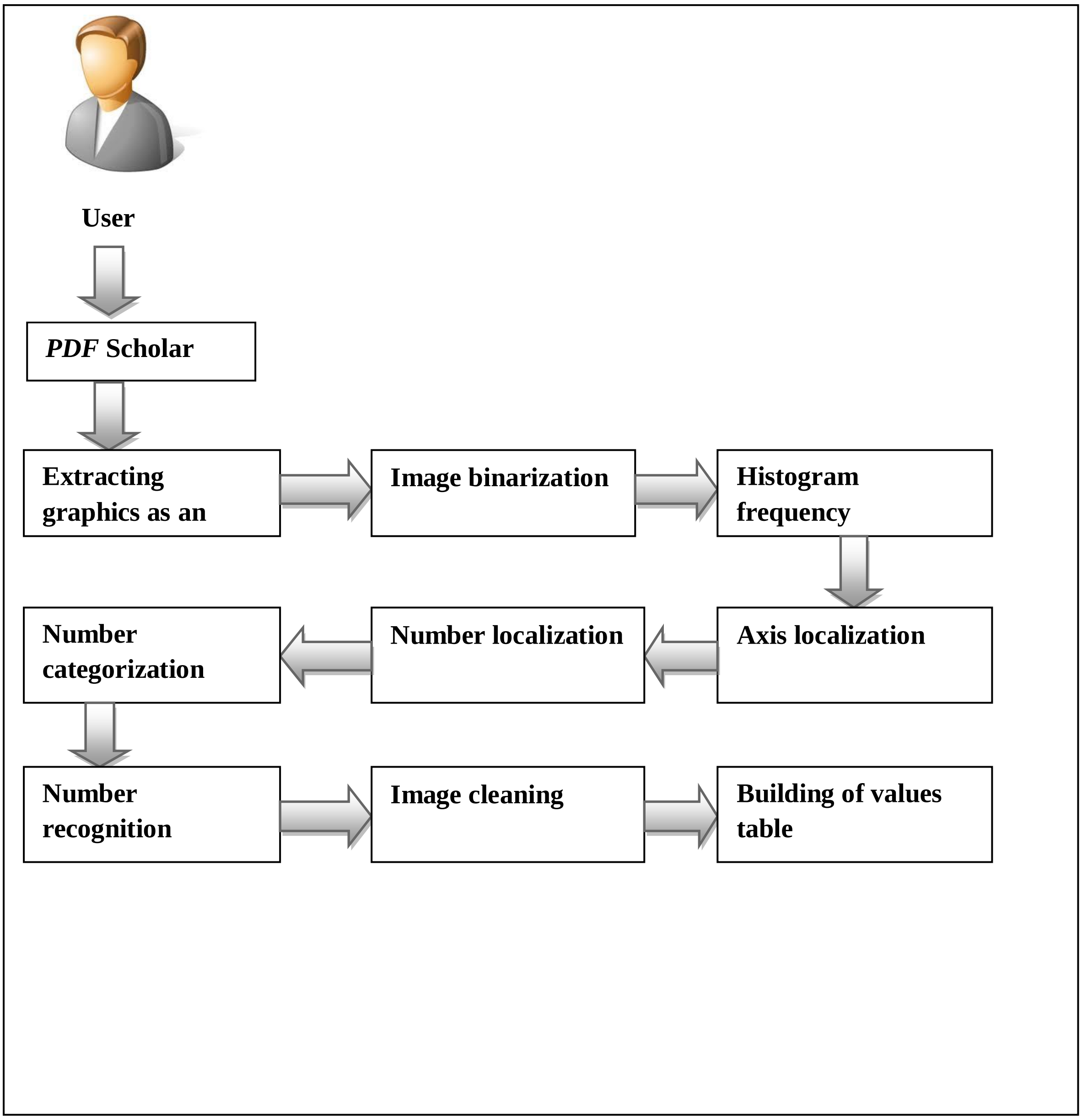

Figure 6. Functional architecture of our system.

3.Conclusion

In this chapter, we have given our approach and motivation with designation of the general architectural and the functional architecture of our system.

Chapter 05: Implementation

1. Introduction

In this chapter, we present the implementation of our system.

2. Development environment

2.1 Java

Java is a versatile, object-oriented programming language and computing platform first released by Sun Microsystems in 1995 and now owned by Oracle Corporation. It is renowned for its portability, performance, and wide adoption across various types of applications, from enterprise servers to mobile devices. Key characteristics of Java include:

- **Object-Oriented:** Java is based on the principles of object-oriented programming, which helps in organizing and structuring complex software systems using objects and classes.

- **Platform Independence:** Java achieves platform independence through the use of bytecode. Java programs are compiled into bytecode, which can run on any device equipped with a Java Virtual Machine (JVM), making the "write once, run anywhere" capability possible.

- **Robust and Secure:** Java includes features that ensure reliable and secure applications, such as strong memory management, exception handling, and a security manager that defines access rules for classes.

- **Simple and Familiar Syntax:** Java's syntax is intended to be easy to understand and similar to C and C++, which helps programmers quickly learn and adopt the language.

- **Multithreading:** Java supports multithreading, allowing developers to create applications that can perform multiple tasks simultaneously, thus improving performance for complex applications.

- **Automatic Memory Management:** Java uses an automatic garbage collection mechanism to manage memory, reducing the risk of memory leaks and other related issues.

- **Extensive Standard Library:** Java comes with a vast standard library that includes a wide range of classes and interfaces for various tasks such as data structures, networking, file I/O, graphical user interface (GUI) development, and more.

- **Dynamic and Extensible:** Java supports dynamic loading of classes, allowing new code to be added at runtime. Its modular nature and the availability of numerous third-party libraries and frameworks enhance its extensibility.

- **High Performance:** Java's Just-In-Time (JIT) compiler optimizes byte-code into native machine code at runtime, improving the performance of Java applications.

Java is widely used for developing web applications, mobile applications (notably Android apps), enterprise software, scientific applications, and large-scale systems due to its reliability, scalability, and extensive ecosystem of tools and frameworks.

Java's popularity grew rapidly, fuelled by its versatility and suitability for a wide range of applications. It became a dominant force in enterprise software development, powering numerous server-side applications and back-end systems. Additionally, Java applets were once widely used for web development, although their usage has declined in favour of other technologies like JavaScript.

Today, Java remains one of the most popular programming languages globally, with vast ecosystem of libraries, frameworks, and tools supporting its development. It continues to evolve, with ongoing updates and improvements to meet the needs of modern software development.

2.2 NetBeans

NetBeans is an open-source integrated development environment (IDE) primarily used for developing applications in Java, although it also supports other programming languages such as PHP, C/C++, HTML5, and JavaScript. Initially developed by Sun Microsystems and now under the stewardship of the Apache Software Foundation, NetBeans offers numerous features to facilitate software development. Here are some key aspects of NetBeans:

- **Multi-language Support:** While NetBeans is particularly well-suited for Java, it supports several other programming languages, making it versatile for different types of development.

- **Powerful Code Editor:** NetBeans provides an intelligent code editor with features like code completion, syntax highlighting, code refactoring, and integrated documentation, helping developers write code faster and with fewer errors.

- **Integrated Debugger:** The IDE offers a comprehensive graphical debugger that allows developers to test and fix their code efficiently by setting breakpoints, monitoring variables, and stepping through the program's execution.

- **Project Management Tools:** NetBeans provides robust tools for managing projects, including support for version control systems like Git, Mercurial, and Subversion, as well as tools for automating the build and deployment of applications with Apache Maven and Gradle.

- **Web Development:** It integrates tools for web development, including editors for HTML, CSS, and JavaScript, as well as popular frameworks like AngularJS, Node.js, and Bootstrap.

- **Graphical User Interface (GUI):** NetBeans includes a graphical designer for Swing and JavaFX, allowing developers to create rich and interactive user interfaces visually by dragging and dropping components.

- **Extensibility:** Through a module system, NetBeans can be extended with many plugins available in the community, enabling developers to add new features and enhance their development environment according to their needs.

- **Documentation and Tutorials:** NetBeans is well-documented, with an abundance of tutorials, guides, and support forums available online to help developers learn and solve problems.

- **Performance and Productivity:** Designed to be fast and responsive, NetBeans helps developers maintain high productivity, even in large-scale projects.

- **Free and Open Source:** As an open-source project under the Apache Software Foundation, NetBeans is free to use and distribute, also benefiting from contributions from developers worldwide.

NetBeans is a preferred choice for developers looking for a comprehensive and flexible IDE for creating desktop, mobile, and web applications, especially when working with Java and related technologies.

3. Some Screen-Shots for our Application

in this section, w present the different screen-shots of our prototype.

Figure 7 depicts the original image taken here as an example for presenting screenshots.

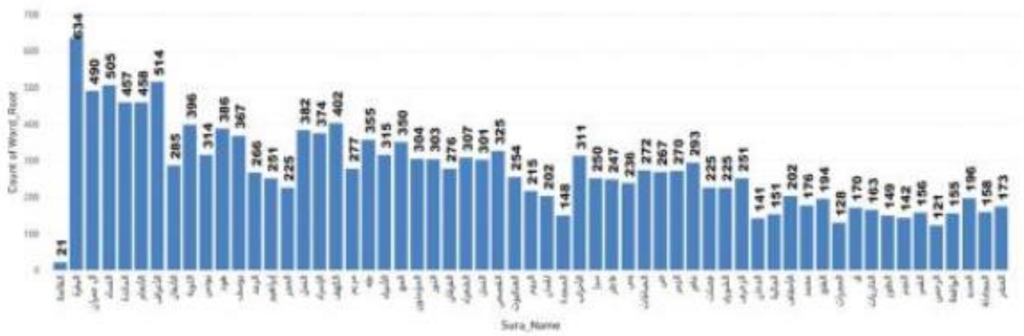

Figure 7. The Original Image.

Figure 8 present the result of the binarization as the first pre-processing step.

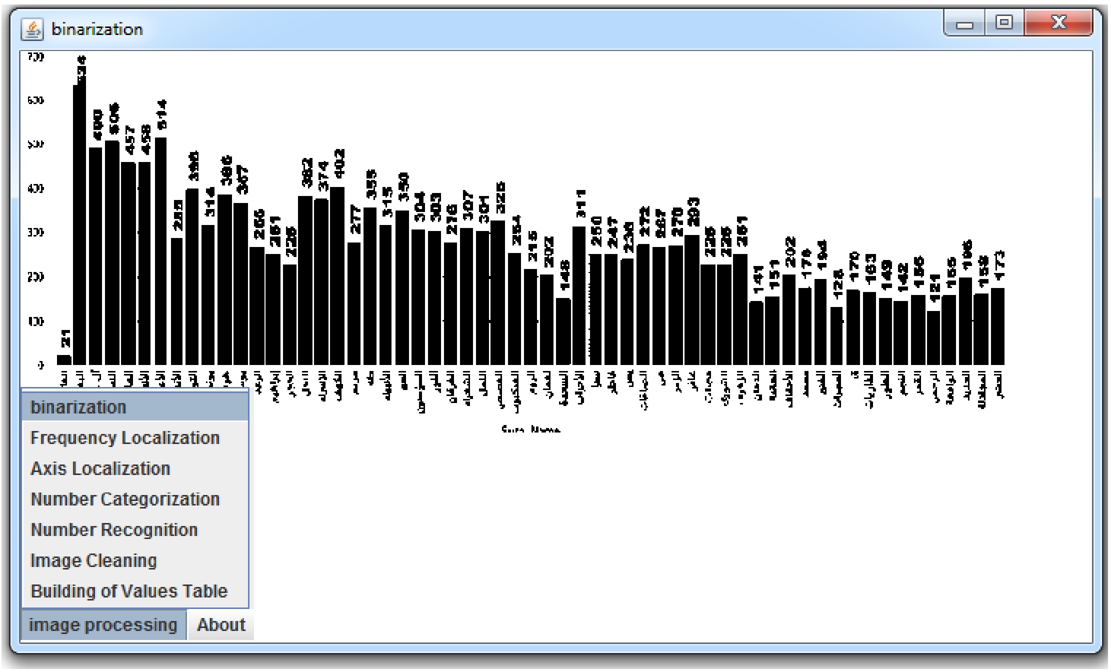

Figure8. The Results for Image binarization.

Figure 9 presents the image after designating the frequency of the different bins constituting the histogram.

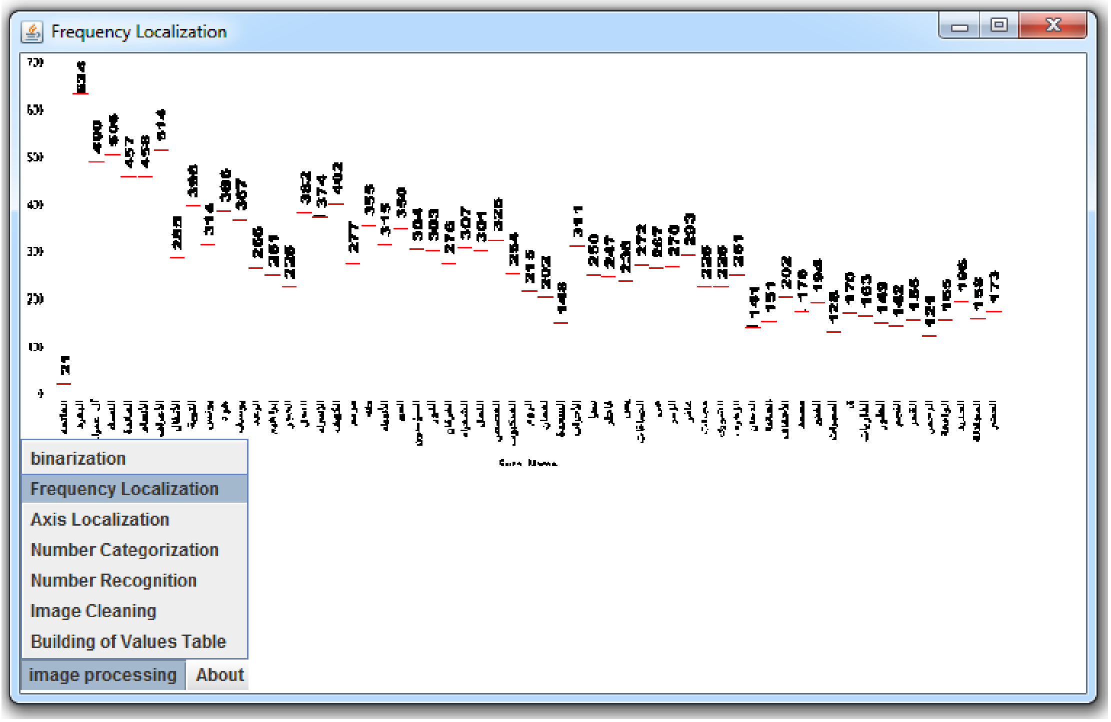

Figure.9 Frequency Localization.

Figure 10 gives the image after the designating of the x-axis and the y-axis.

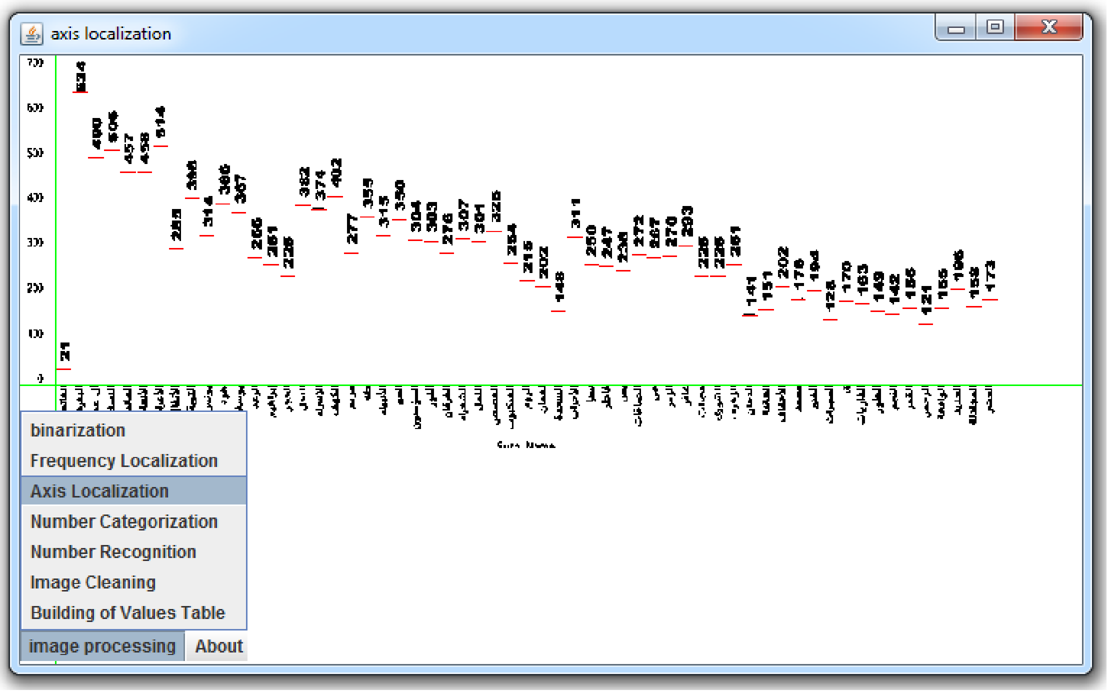

Figure.10 Axis localization.

Figure 11 depicts the image after the cleaning task.

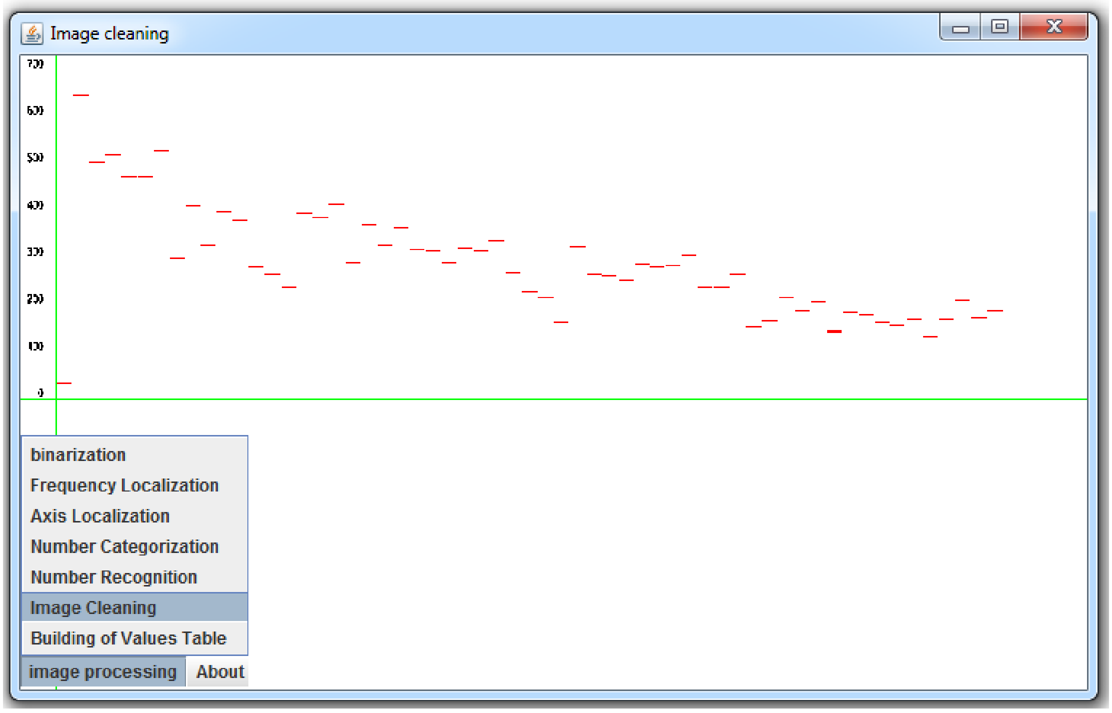

Figure.11 Image cleaning.

Figure 12 shows the image after the localization and the categorization of the high frequency number.

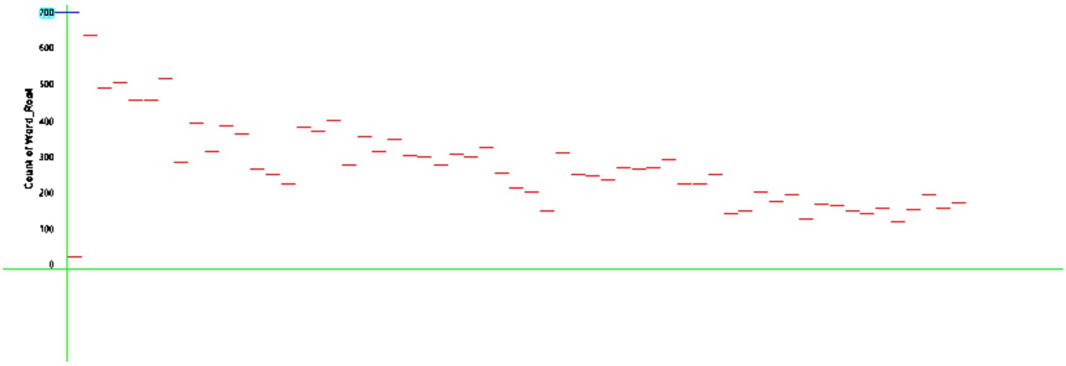

Figure.12 localization and categorization of the high frequency number.

Table 5 presents the values of the different bins of the histogram in terms of pixels and in terms of real values.

Table 5 The results of values in terms of pixels and real values.

The Frequency of the bin	The value in Pixels	The real value
1	9	31.81
2	172	637.03
3	133	492.59
4	137	507.40
5	125	462.96
6	125	462.96
7	140	518.51
8	79	292.59
9	108	400
10	87	322.22

11	**106**	**392.59**
12	**100**	**370.37**
13	**74**	**274.07**
14	**70**	**259.25**
15	**63**	**233.33**
16	**105**	**388.88**
17	**102**	**377.77**
18	**110**	**407.40**
19	**77**	**285.18**
20	**98**	**362.96**
21	**87**	**322.22**
22	**96**	**355.55**
23	**84**	**311.11**
24	**83**	**307.40**
25	**77**	**285.18**
26	**85**	**314.81**
27	**83**	**307.40**
28	**90**	**333.33**
29	**71**	**262.96**
30	**60**	**222.22**
31	**57**	**211.11**
32	**43**	**159.25**
33	**86**	**318.51**
34	**70**	**259.25**
35	**69**	**255.55**
36	**66**	**244.44**
37	**75**	**277.77**
38	**74**	**274.07**
39	**75**	**277.77**
40	**81**	**300**
41	**63**	**233.33**
42	**63**	**233.33**
43	**70**	**259.25**
44	**41**	**151.85**
45	**43**	**159.25**
46	**57**	**211.11**
47	**50**	**185.18**
48	**55**	**203.70**
49	**37**	**137.03**
50	**48**	**177.77**
51	**47**	**174.07**
52	**43**	**159.25**
53	**41**	**151.85**
54	**45**	**166.66**
55	**35**	**129.62**
56	**44**	**162.96**
57	**55**	**203.70**
58	**45**	**166.66**
59	**49**	**181.48**

4.Conclusion

In this chapter, we presented the implementation of our system. We have given some definitions of some concepts namely: Java, NetBeans. Then, we have presented some screen-shots for our application.

General conclusion

In this Master dissertation, we have dealt with a recognition issue through trying to transform a graphic (a histogram) to its associated table of values. We have considered some pre-processing, processing, and post-processing steps: binarization, frequency localization, axis localization, number localization, number categorization, number recognition, image cleaning, and building of the table of values. As a contribution, we have implemented all the steps else binarization with own proper algorithms. However, our considered segmentation, although its originality, belongs to the edge segmentation approach.

The initial purpose of the work was transforming Graphic in general either a curve or a histogram. As a perspective then, we aim to generate the table of values from a curve where we think that the complexity of the problem may be higher.

References

[1] Meghrbi, O & Ahmed Saadi, S (2009) Qantification de la couverture nuageuse en utilisant des images ciel-total.

[2] Gonzalez, Rafael (2018). Digital image processing. New York, NY: Pearson. ISBN 978-0-13-335672-4. OCLC 966609831 .

[3] Mawloud, M. (2017). *Mesures de distances dans le contexte de la recherche d'images par le contenu (CBIR)* (Doctoral dissertation, Université du 20 Août 1955).

[4] Ozturk, C., Hancer, E., & Karaboga, D. (2014). Color image quantization: a short review and an application with artificial bee colony algorithm. *Informatica*, *25*(3), 485-503.

[5] Orchard, M. T., & Bouman, C. A. (1991). Color quantization of images. *IEEE transactions on signal processing*, *39*(12), 2677-2690.

[6] Atsalakis, A., & Papamarkos, N. (2006). Color reduction and estimation of the number of dominant colors by using a self-growing and self-organized neural gas. *Engineering Applications of Artificial Intelligence*, *19*(7), 769-786.

[7] Braquelaire, J. P., & Brun, L. (1997). Comparison and optimization of methods of color image quantization. *IEEE transactions on image processing*, *6*(7), 1048-1052.

[8] Mustafa, W. A., & Kader, M. M. M. A. (2018, June). Binarization of images: A comprehensive review. In *Journal of Physics: Conference Series* (Vol. 1019, No. 1, p. 012023). IOP Publishing.

[9] Krupinski, R., Lech, P., Teclaw, M., & Okarma, K. (2019). Binarization of degraded images with generalized Gaussian distribution. In *Computational Science-ICCS 2019: 19^{th} International Conference, Faro, Portugal, June 12-14, 2019, Proceedings, Part V 19 (pp. 177-190). Springer International Publishing.*

[10] Moghaddam, R. F., & Cheriet, M. (2010). A multi-scale framework for adaptive binarization of degraded images. *Pattern Recognition*, 43(6), 2186-2198.

[11] Kwon, S. H. (2004). Threshold selection based on cluster analysis. *Pattern Recognition Letters*, 25(9), 1045-1050.

[12] Papamarkos, N. (2001, November). A technique for fuzzy document binarization. In Proceedings of the 2001 ACM Symposium on Document Engineering (pp. 152-156).

[13] Drira, F. (2006, April). Towards restoring historic documents degraded over time. In *Second International Conference on Image Analysis for Libraries* (*DIAL'06*) (pp. 8-357). IEEE.

[14] Chaki, N., Shaikh, S. H., Saeed, K., Chaki, N., Shaikh, S. H., & Saeed, K. (2014). A comprehensive survey on image binarization techniques (pp. 5-15). *Springer India*.

[15] Abd Elfattah, M., Hassanien, A. E., Mostafa, A., Ali, A. F., Amin, K. M., & Mohamed, S. (2015, December). Artificial bee colony optimizer for historical Arabic manuscript images binarization. In *2015 11th International Computer Engineering Conference (ICENCO)* (pp. 251-255). IEEE.

[16] Bhowmik, S. (2023). Image Binarization. In Document Layout Analysis (pp. 11-30). *Singapore: Springer Nature Singapore*.

[17] Jia, F., Shi, C., He, K., Wang, C., & Xiao, B. (2018). Degraded image binarization using structural symmetry of strokes. *Pattern Recognition*, 74, 225-240.

[18] Wu, Y., Natarajan, P., Rawls, S., & AbdAlmageed, W. (2016, September). Learning image binarization from data. In *2006 IIEEE International Conference on Image Processing* (*ICIP*) (pp. 3763-3767). IEEE.

[19] Bernsen, J. (1986). Dynamic thresholding of gray-level images [A]. *In Proceedings of 8th International Conference on Pattern Recognition*.

[20] Sauvola, J., & Pietikainen, M. (2000). Adaptive image binarization. *Pattern recognition*, 33(2), 225-236.

[21] Kim, I. J. (2004, October). Multi-window binarization of camera image for document recognition. In *Ninth International Workshop on Frontiers in Handwriting Recognition* (pp. 323-327). IEEE.

[22] Mustafa, W. A., Aziz, H., Khairunizam, W., Ibrahim, Z., Shahriman, A. B., & Razlan, Z. M. (2018, August). Review of different binarization approaches on degraded images. In *2018 International Conference on Computational Approach in Smart Systems Design and Applications* (*ICASSDA*) (pp. 1-8). IEEE.

[23] Chauhan, S., Sharma, E., & Doegar, A. (2016, September). Binarization techniques for degraded images-A review. In *2016 5th International Conference on Reliability, Infocom Technologies and Optimization* (*Trends and Future Directions*) (*ICRITO*) (pp. 163-166). IEEE.

[24] Sulaiman, A., Omar, K., & Nasrudin, M. F. (2019). Degraded historical document binarization: A review on issues, challenges, techniques, and future directions. *Journal of Imaging*, 5(4), 48.

[25] Tensmeyer, C., & Martinez, T. (2020). Historical image binarization: A review. *SN Computer Science*, 1(3), 173.

[26] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on Systems, man, and cybernetics*, 9(1), 62-66.

[27] Jana, P., Ghosh, S., Bera, S. K., & Sarkar, R. (2017, December). Handwritten image binarization: An adaptive k-means based approach. In *2017 IEEE Cacutta Conference (CALCON)* (pp. 226-230). IEEE.

[28] Papamarkos, N. (2001, November). A technique for fuzzy document binarization. In *proceedings of the 2001 ACM Symposium on Document engineering* (pp. 152-156).

[29] Mustafa, W. A., Aziz, H., Khairunizam, W., Zunaidi, I., Razlan, Z. M., & Shahriman, A. B. (2019, June). Images binarization using hybrid combination of fuzzy C-means and Deghost method. In *IOP Conference Series: Materials Science and Engineering* (Vol. 557, No. 1, p. 012012). IOP Publishing.

[30] Jana, P., Ghosh, S., Sarkar, R., & Nasipuri, M. (2017, December). A fuzzy C-means based approach towards efficient image binarization. In *2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR)* (pp. 1-6). IEEE.

[31] Mondal, T., Coustaty, M., Gomez-Kramer, P., & Ogier, J. M. (2019, September). Learning free image binarization based on fast fuzzy c-means clustering. In *2019*

International Conference on Document Analysis and Recognition (ICDAR) (pp. 1384-1389). IEEE.

[32] Tong, L. J., Chen, K., Zhang, Y., Fu, X. L., & Duan, J. Y. (2009, October). Image binarization based on NFCM. In *2009 2nd International Congress on Image and Signal Processing* (pp. 1-5). IEEE

[33] Johannsen, G. (1982). A threshold selection method using information measures. In *International Conference on Pattern Recognition* (*ICPR'82)*, 140-143.

[34] Kapur, J. N., Sahoo, P. K., & Wong, A. K. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Computer vision, graphics, and image processing*, 29(3), 273-285.

[35] Niblack, W. (1985). An introduction to digital image processing. *Strandberg Publishing Company.*

[36] Liu, Y., & Srihari, S. N. (1997). Image binarization based on texture features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5), 540-544

[37] He, J., Do, Q. D. M., Downton, A. C., & Kim, J. (2005, August). A comparison of binarization methods for historical archive documents. In *Eighth International Conference on Document Analysis and Recognition* (*ICDAR'05*) (pp. 538-542). IEEE.

[38] Yazid, H., & Arof, H. (2013). Gradient based adaptive thresholding. Journal of Visual Communication and Image Representation, 24(7), 926-936.

[39] Yanowitz, S. D., & Bruckstein, A. M. (1989). A new method for image segmentation. *Computer vision, Graphics, and Image Processing*, 46(1), 82-95.

[40] Khurshid, K., Siddiqi, I., Faure, C., & Vincent, N. (2009, January). Comparison of Niblack inspired binarization methods for ancient documents. In *Document Recognition and Retrieval XVI* (Vol. 7247, pp. 267-275). SPIE.

[41] Bradley, D., & Roth, G. (2007). Adaptive thresholding using the integral image. *Journal of graphics tools*, 12(2), 13-21.

[42] Pierre, D. (1993). WELLNER Adaptive Thresholding for the Digital Desk. *Technical Report EPC*-1993-110.

[43] Bernsen, J. (1986). Dynamic Thresholding of Grey-Level Images, [in:] *Proc. of the 8th Int. Conf. on Pattern Recognition.*

[44] E Woods, R., & C Gonzalez, R. (2008). Digital image processing

[45] Tensmeyer, C., & Martinez, T. (2017, November). Image binarization with fully convolutional neural networks. In *2017 14th IAPR International Conference on document analysis and recognition (ICDAR)* (Vol. 1, pp. 99-104). IEEE.

[46] Kefali, A., Sari, T., & Bahi, H. (2014). Foreground-background separation by feed-forward neural networks in old manuscripts. *Informatica*, 38(4).

[47] Sari, T., Kefali, A., & Bahi, H. (2012, September). An MLP for binarizing images of old manuscripts. In *2012 International Conference on Frontiers in Handwriting Recognition* (pp. 247-251). IEEE.

[48] Akbari, Y., AI-Maadeed, S., & Adam, K. (2020). Binarization of degraded images using convolutional neural networks and wavelet-based multichannel images. *IEEE Access*, 8, 153517-153534.

[49] Badekas, E., & Papamarkos, N. (2007). Optimal combination of document binarization techniques using a self-organizing map neural network. *Engineering Applications of Artificial Intelligence*, 20(1), 11-24.

[50] Gatos, B., Pratikakis, I., &Perantonis, S. J. (2006). Adaptive degraded image binarization. 39(3), 317-327.

[51] Jain, A. K. (1989). Fundamentals of digital image processing. *Prentice-Hall, Inc.*

[52] Lu, H., Kot, A. C., & Shi, Y. Q. (2004). Distance-reciprocal distortion measure for binary images. *IEEE Signal Processing Letters*, 11(2), 228-231.

[53] Moghaddam, R. F., & Cheriet, M. (2012). AdOtsu: An adaptive and parameterless generalization of Otsu’s method for image binarization. *Pattern Recognition*, 45(6), 2419-2431.

[54] Lu, S., B., & Tan, C. L. (2010). Image binarization using background estimation and stroke edges. *International Journal on Document Analysis and Recognition (IJDAR)*, 13, 303-314.

[55] Su, B., Lu, S., & Tan, C. L. (2010, June). Binarization of historical images using the local maximum and minimum. In *proceedings of the 9th IAPR International Workshop on Document Analysis Systems* (pp. 159-166).

[56] Su, B., Lu, S., & Tan, C. L. (2012). Robust image binarization technique for degraded images. *IEEE Transactions on Image Processing*, 22(4), 1408-1417.

[57] Pratikakis, I., Zagoris, K., Kaddas, P., & Gatos, B. (2018). ICFHR2018 Competition on handwritten image binarization contest (H-DIBCO 2018). In *International Conference on Frontiers in Handwritting Recognition* (pp. 1-1).

[58] Pratikakis, I., Zagoris, K., Karagiannis, S., Tsochatzidis, L., Mondal, T., Wang, X, … & Li, M. (2019, November). *ICDAR 2019 Competition on image binarization*. In *ICDAR*.

[59] Sari, T., Kefali, A., & Bahi, H. (2014). Text extraction from historical images by the combination of several thresholding techniques. *Advances in Multimedia*, 2014, 11-11.

[60] Pratikakis, I., Zagoris, K., Barlas, G., & Gatos, B. (2016, October). ICFHR2016 handwritten image binarization contest (H-DIBCO 2016). In *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (pp. 619-623). IEEE.

[61] Ntirogiannis, K., Gatos, B., & Pratikakis, I. (2014, September). ICFHR2014 competition on handwritten image binarization (H-DIBCO 2014). In *2014 14th International Conference on Frontiers in Handwriting Recognition (pp. 809-813)*. IEEE.

[62] Patrikakis, I., Gatos, B., & Ntirogiannis, K. (2012, September). ICFHR 2012 competition on handwritten image binarization (H-DIBCO 2012). In 2012 *International Conference on Frontiers in Handwriting Recognition* (pp. 817-822). IEEE.

[63] Pratikakis, I., Gatos, B., & Ntirogiannis, K. (2010, November). H-DIBCO 2010-handwritten image binarization competition. In *2010 12th International Conference on Frontiers in Handwriting Recognition* (pp. 727-732). IEEE.

[64] Gatos, B., Ntirogiannis, K., & Pratikakis, I. (2009, July). ICDAR 2009 image binarization contest (DIBCO 2009). In *2009 10th International Conference on document analysis and recognition* (pp. 1375-1382). IEEE.

[65] Huang, X., Li, L., Liu, R., Xu, C., & Ye, M. (2020). Binarization of degraded images with global-local U-Nets. *Optik*, 203, 164025.

[66] Shi, J., Ray, N., & Zhang, H. (2012). Shape based local thresholding for binarization of images. *Pattern Recognition Letters*, 33(1), 24-32.

[67] Howe, N. R. (2013). Document binarization with automatic parameter tuning. International Journal on Document Analysis and Recognition (IJDAR), 16, 247-258.

[68] Su, B., Lu, S., & Tan, C. L. (2011, September). Combination of image binarization techniques. In *2011 International Conference on Document Analysis and Recognition* (pp. 22-26). IEEE.

[69] Mustafa, W. A., & Kader, M. M. M. A. (2018, June). Binarization of image using optimum threshold modification. In *Journal of Physics: Conference Series* (Vol. 1019, No. 1,p. 012022). *IOP Publishing.*

[70] Lu, D., Huang, X., & Sui, L. (2018). Binarization of degraded images based on contrast enhancement. *International Journal on Document Analysis and Recognition (IJDAR)*, 21, 123-135.

[71] Vo, Q. N., Kim, S. H., Yang, H. J., & Lee, G. (2018). Binarization of degraded images based on hierarchical deep supervised network. *Pattern Recognition*, 74, 568-586.

[72] Sehad, A., Chibani, Y., Hedjam, R., & Cheriet, M. (2019). Gabor filter-based texture for ancient degraded image binarization. *Pattern Analysis and Applications*, 22, 1-22.

[73] Biswas, B., Bhattacharya, U., & Chaudhuri, B. B. (2014, August). A global-to-local approach to binarization of degraded images. In *2014 22nd International Conference on Pattern Recognition* (pp. 3008-3013). IEEE.

[74] Michalak, H., & Okarma, K. (2019). Improvement of image binarization methods using image preprocessing with local entropy filtering for alphanumerical character recognition purposes. *Entropy*, 21(6), 562.

[75] Ziaei Nafchi, H., Farrahi Moghaddam, R., & Cheriet, M. (2013). Historical document binarization based on phase information of images. In *Computer Vision-ACCV 2012 Workshops: ACCV 2012 International Workshops, Daejeon, Korea*, November 5-6, 2012, Revised Selected Papers, Part II 11 (pp. 1-12). Springer Berlin Heidelberg.

[76] Almeida, M., Lines, R. D., Bernadino, R., Jesus, D., & Lima, B. (2018). A new binarization algorithm for historical documents. *Journal of Imaging*, 4(2), 27.

[77] Aurich, V., & Weule, J. (1995). Non-linear Gaussian filters performing edge preserving diffusion. In Musterekennung 1995: Verstehen akustischer und visueller Informationen (pp. 538-545). *Springer Berlin Heidelberg*.

[78] Tomasi, C., & Manduchi, R. (1998, January). Bilateral filtering for gray and color images. In *Sixth International Conference on Computer Vision* (IEEE Cat. No. 98CH36271) (pp. 839-846). IEEE.

[79] Rahiche, A., & Cheriet, M. (2019). KFBin: Kalman filter-based approach for image binarization. In *Image Analysis and Recognition: 16th International Conference, ICIAR 2019*, Waterloo, ON, Canada, August 27-29, 2019, Proceedings, Part I 16 (pp. 150-161). Springer International Publishing.

[80] Chang, C. I., & Chang, C. I. (2017). Discrete-time Kalman filtering for hyperspectral processing. *Real-Time Recursive Hyperspectral Sample and Band Processing: Algorithm Architecture and Implementation*, 49-71.

[81] Chen, Y., & Wang, L. (2017). Broken and degraded images binarization. *Neurocomputing*, 237, 272-280.

[82] He, K., Sun, J., & Tang, X. (2012). Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, 35(6), 1397-1409.

[83] Sarkka, S., & Svensson, L. (2023). Bayesian filtering and smooting (Vol. 17). *Cambridge university press*.

[84] Bera, S. K., Ghosh, S., Bhowmik, S., Sarkar, R., & Nasipuri, M. (2021). A non-parametric binarization method based on ensemble of clustering algorithms. *Multimedia Tools and Applications*, 80, 7653-7673.

[85] Wen, J., Li, S., & Sun, J. (2013). A new binarization method for non-uniform illuminated images. *Pattern recognition*, 46(6), 1670-1690.

[86] Nikolaidou, K., Seuret, M., Mokayed, H., & Liwicki, M. (2022). A survey of historical image datasets. *International Journal on Document Analysis and Recognition* (*IJDAR*), 25(4), 305-338.

[87] Mustafa, W. A., Khairunizam, W., Zunaidi, I., Razlan, Z. M., & Shahriman, A. B. (2019, June). A comprehensive review on image (DIBCO) database. In *IOP Conference Series: Materials Science and Engineering* (Vol. 557, No. 1, p. 012006). IOP Publishing.

[88] Zhuang, J. (2018). LadderNet: Multi-path networks based on U-Net for medical image segmentation. *arXiv preprint arXiv: 1810.07810*.

[89] Wolf, C., Jolion, J. M., & Chassaing, F. (2002, August). Text localization, enhancement and binarization in multimedia documents. In 2002 International Conference on Pattern Recognition (Vol. 2, pp. 1037-1040). IEEE.

[90] Lelore, T., & Bouchara, F. (2013). FAIR: a fast algorithm for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 2039-2048.

[91] Howe, N. R. (2011, September). A laplacian energy for document binarization. In *2011 International Conference on Document Analysis and Recognition* (pp. 6-10). IEEE.

[92] Jaiswal, S., & Pandey, M. K. (2020). A review on image segmentation. *Rising Threats in Expert Applications and Solutions: Proceedings of FICR-TEAS 2020*, 233-240.

[93] Balla-Arabé, S., Gao, X., & Wang, B. (2013). A fast and robust level set method for image segmentation using fuzzy clustering and lattice Boltzmann method. *IEEE transactions on cybernetics*, *43*(3), 910-920.

[94] Chun, D. N., & Yang, H. S. (1996). Robust image segmentation using genetic algorithm with a fuzzy measure. *Pattern recognition*, *29*(7), 1195-1211.

[95] Agrawal, S., Panda, R., & Dora, L. (2014). A study on fuzzy clustering for magnetic resonance brain image segmentation using soft computing approaches. *Applied Soft Computing*, *24,* 522-533.

[96] Chamalis, T., & Likas, A. (2017, April). Region merging for image segmentation based on unimodality tests. In *2017 3rd international conference on control, automation and robotics (ICCAR)* (pp. 381-384). IEEE.

[97] Deng, W. Q., Li, X. M., Gao, X., & Zhang, C. M. (2016). A modified fuzzy C-means algorithm for brain MR image segmentation and bias field correction. *Journal of computer science and technology*, *31*(3), 501-511.

[98] Riomoros, I., Guijarro, M., Pajares, G., Herrera, P. J., Burgos-Artizzu, X. P., & Ribeiro, A. (2010, December). Automatic image segmentation of greenness in crop fields. In *2010 international conference of soft computing and pattern recognition* (pp. 462-467). IEEE.

[99] Saito, S., Li, T., & Li, H. (2016). Real-time facial segmentation and performance capture from rgb input. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VIII 14* (pp. 244-261). Springer International Publishing.

[100] Yardimci, A. (2009). Soft computing in medicine. *Applied soft computing*, *9*(3), 1029-1043.

[101] Chouhan, S. S., Kaul, A., & Singh, U. P. (2018). Soft computing approaches for image segmentation: a survey. *Multimedia Tools and Applications*, *77*(21), 28483-28537.

[102] Jaiswal, S., & Pandey, M. K. (2020). A review on image segmentation. *Rising Threats in Expert Applications and Solutions: Proceedings of FICR-TEAS 2020*, 233-240.

[103] Senthilnath, J., Omkar, S. N., Mani, V., Tejovanth, N., Diwakar, P. G., & Shenoy, A. B. (2012). Hierarchical clustering algorithm for land cover mapping using satellite images. *IEEE journal of selected topics in applied earth observations and remote sensing*, *5*(3), 762-768.

[104] Zhang, M., Zhang, L., & Cheng, H. D. (2010). A neutrosophic approach to image segmentation based on watershed method. *Signal processing*, *90*(5), 1510-1517.

[105] Yi, F., & Moon, I. (2012, May). Image segmentation: A survey of graph-cut methods. In *2012 international conference on systems and informatics (ICSAI2012)* (pp. 1936-1941). IEEE.

[106] Wani, M. A., & Batchelor, B. G. (1994). Edge-region-based segmentation of range images. *IEEE transactions on pattern analysis and machine intelligence*, *16*(3), 314-319.

[107] R.C. Gonzalez, Richard, *Digital Image Processing*, 3rd edn. (Hardcover, 2007)

[108] Alazzawi, A. (2015). Edge detection-application of (first and second) order derivative in image processing: communication. *Diyala Journal of Engineering Sciences*, 430-440.

[109] Kale, A., Yadav, H., & Jain, A. (2014). A review: image segmentation using genetic algorithm. *Int. J. Sci. Eng. Res*, *5*(2), 455-458.

[110] . M. Peixeiro, Introduction to Support Vector Machine (2019). Home page https://towardsdatascience.com/introduction-to-support-vector-machine-svm4671e2cf3755. Last accessed 21 Oct 2019

[111] Karim, T. F., Lipu, M. S. H., Rahman, M. L., & Sultana, F. (2010, July). Face recognition using PCA-based method. In *2010 IEEE International Conference on Advanced Management Science (ICAMS 2010)* (Vol. 3, pp. 158-162). IEEE.

[112] Mignotte, M. (2008). Segmentation by fusion of histogram-based $ k $-means clusters in different color spaces. *IEEE Transactions on image processing*, *17*(5), 780-787.

[113] LNCS Homepage, https://www.analyticsvidhya.com/blog/2019/04/introduction-imagesegmentation-techniques-python/. Last accessed 17 Oct 2019

[114] LNCS Homepage, https://in.mathworks.com/help/vision/ug/getting-started-with-semantic segmentation-using-deep-learning.html. Last accessed 17 Oct 2019

[115] Kirillov, A., He, K., Girshick, R., Rother, C., & Dollár, P. (2019). Panoptic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9404-9413).

[116] Gonzalez, J. P., & Ozguner, U. (2000, October). Lane detection using histogram-based segmentation and decision trees. In *ITSC2000. 2000 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 00TH8493)* (pp. 346-351). IEEE.

[117] Orlando Tobias, Seara, Rui: image segmentation by histogram thresholding using fuzzy sets. IEEE Trans. Image Process. **11**, 1457–1465 (2002)

[118] Christ, M. J., & Parvathi, R. M. S. (2011, April). Fuzzy c-means algorithm for medical image segmentation. In *2011 3rd International conference on electronics computer technology* (Vol. 4, pp. 33-36). IEEE.

[119] Yan, J. (2014). Remote sensing image segmentation based on ant colony optimized fuzzy C-means clustering. *Journal of Chemical and Pharmaceutical Research*, *6*(6), 2675-2679.

[120] Held, K., Kops, E. R., Krause, B. J., Wells, W. M., Kikinis, R., & Muller-Gartner, H. W. (1997). Markov random field segmentation of brain MR images. *IEEE transactions on medical imaging*, *16*(6), 878-886.

[121] Raff, U., Scherzinger, A. L., Vargas, P. F., & Simon, J. H. (1994). Quantitation of grey matter, white matter, and cerebrospinal fluid from spin-echo magnetic resonance images using an artificial neural network technique. *Medical Physics*, *21*(12), 1933-1942.

[122] Chaabane, L., & Abdelouahab, M. (2011). Evaluation of the segmentation by multispectral fusion approach with adaptive operators: Application to medical images. *International Journal of Advanced Computer Science and Applications*, *2*(9).

[123] Monteiro, F. C., & Campilho, A. C. (2006, September). Performance evaluation of image segmentation. In *International Conference Image Analysis and Recognition* (pp. 248-259). Berlin, Heidelberg: Springer Berlin Heidelberg.

[124] Huang, Q., & Dom, B. (1995, October). Quantitative methods of evaluating image segmentation. In *Proceedings., international conference on image processing* (Vol. 3, pp. 53-56). IEEE.

[125] Martin, D. R. (2002). *An empirical approach to grouping and segmentation*. University of California, Berkeley.

[126] Cardoso, J. S., & Corte-Real, L. (2005). Toward a generic evaluation of image segmentation. *IEEE Transactions on Image Processing*, *14*(11), 1773-1782.

[127] Gouiffès, M. (2005). *Apports de la Couleur et des Modèles de Réflexion pour l'Extraction et le Suivi de Primitives* (Doctoral dissertation, Poitiers).

[128] Martin, D. R. (2002). *An empirical approach to grouping and segmentation*. University of California, Berkeley.

[129] Raghavan, V., Bollmann, P., & Jung, G. S. (1989). A critical investigation of recall and precision as measures of retrieval system performance. *ACM Transactions on Information Systems (TOIS)*, *7*(3), 205-229.

[130] Peleg, S., Werman, M., & Rom, H. (1989). A unified approach to the change of resolution: Space and gray-level. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(7), 739-742.

[131] Rubner, Y., Tomasi, C., & Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. *International journal of computer vision*, *40*, 99-121.

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليـــــم العالــــــي و البحث العلمـــــي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université 20 Août 1955 - Skikda
Faculté des Sciences
Bibliothèque de la faculté

جـامعة 20 أوت 1955- سكيكدة
كـــليــة العلـــوم
مكتبـة الكليـة

سكيكدة في / /2024

بطاقة معلومات خاصة بمذكرة التخرج

اسم و لقب الطالب :

رقم التسجيل :

*لصمو يتاس .... *2019 36003799 ....
*رحاحلة اكرام .... *2019 36003790 ....
*........................ *........................
*........................ *........................

اسم و لقب المشرف على المذكرة : مولود مصباح

عنوان المذكرة :

Tranformation of Imagegraphic to a table of values

القسم : اعلام آلي

المستوى : الثانية ماستر

التخصص : R.S.D

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليـــــم العالـــــي و البحث العلمـــي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université 20 Août 1955- skikda-

Faculté des Sciences

Département d'Informatique

جـامعة 20 أوت 1955 - سكيكدة

كــــلية العلـــوم

قسم الاعلام الالي

Université 20 Août 1955 SKIKDA

الرقم : ...../ ق/ ا/ أ/ ك ع/2024

Autorisation de Dépôt de Mémoire de Master

Je soussigné: Dr. Mawloud Mosbah

Certifie que l'étudiant(e) : Ines Lasmar et Ikram Rebiha

Spécialité : RSD

Ayant soutenu le projet intitulé : Transformation of Imag Graphic to a table of values

A apporté les corrections nécessaires sur son manuscrit de Master

Signature de l'encadreur

Mawloud Mosbah