Survey on Artificial Intelligence based Image Processing and Image Segmentation analysis

Mei Kurokawa
School of Applied Chemistry, Tongji University, Shanghai, China
Email: ku55582221@hotmail.com

Abstract – This paper critically surveys the aspect of digitalized image processing and segmentation with central focus on artificial intelligence. A digitalized image is composed of numerous elements that must be "analyzed," since better "phrasing", and the "research" done on such elements might disclose a lot of strange data. This data may assist in solving a variety of business challenges, which is one of the numerous end objectives associated with image processing. Digital image analysis of image processing is a collection of methodologies applied to process computerized images. It integrates accomplishing basis assignments such noise minimization and more complex assignments such as image classification, fault diagnostics, emotional response, image fragmentation and image segmentation. Image enhancement applying neural networks has, over the past few decades, been applied widely in the clinical setting due to the advent of advanced computing ecosystems and algorithms. Medicine, industries, police departments, agriculture, defense, finance and security are some of the additional domains where the concept of image processing and segmentation can be applied.

Keywords – Image Processing, Image Segmentation, Image Differentiation, Image Analysis, Artificial Intelligence

I. INTRODUCTION

Image segmentation denotes to the methodology of partitioning images into different divisions known as image entities that minimize the images’ complexities and hence rendering image analysis significantly easier. To classify and divide particular categorization of pixels from images, we utilize different image segmentation methodologies. By fulfilling this, we are providing a label to the pixel and images of similar labels are classified together because there is something common accredited to them. We could establish a limit, draw and divide lines on the most fundamental things in an image from the rest of the less fundamental ones utilizing these tags. We focus on retrieving fundamental elements such as tables, chairs etc. based on the initial image, and therefore all the seats are colored identically. We found instances on the next tab that deal about distinct objects, thus all of the chairs had different colors. Background logical operations in video, segmentation of tumor cells, trace artifacts in satellite photos (such as roads, forests, etc.), traffic handling platforms, facial detection, iris acknowledgment, agribusiness, and medical imaging are just a few of the applications of image processing. Image processing is a technique that involves doing various operations on a picture in order to create a better image or extract usable data from it. Innovative technologies are currently being developed in the disciplines of image analysis, particularly in the field of image segmentation. However, image classification is a time-consuming and difficult task.

Image segmentation is used to separate a picture into meaningful segments with similar traits and properties. The primary goal of segmentation is to simplify a picture, i.e. to represent it in a form that is expressive and easy to analyze. Photograph segmentation is widely regarded among the most important data processing methods [1]. Image differentiation is the process of splitting or grouping an image into parts. It is primarily beneficial for operations such as picture analysis or object recognition since processing the complete picture is costly in these cases. As a result, image segmentation is utilized for additional processing in order to identify image parts. There is a variety of image segmentation techniques that divide the image into many portions based on image qualities such as pixel resolution, color, structure, and so on. The first stage in image analysis is picture segmentation. As seen in Fig 1, the purpose of image fragmentation is to divide an image into numerous parts/segments with comparable traits or properties.

To increase visual recognition and understanding, digital images need image alteration and analysis [2]. Two fundamental motives drive the production of visual information for greater human comprehension and the study of scenic...
elements for autonomous perception. This description encompasses systems ranging from fundamental electronic systems to state-of-the-art parallel processing, as well as computer systems such as desktop computers or workstation. There are increasing aspects of computerized image processing equipment, including hardware and software. While there is specialist imaging equipment available, image processing may also be done on a standard personal computer. Advanced devices, including image analysis equipment or a PCB test technique, are mainly focused with specialized technologies and production phases. The phases are applied in the image fragmentation process (see Fig 1).

II. NEED FOR IMAGE SEGMENTATION
The idea of separating, splitting, retrieving, labelling, and then utilizing that data to train multiple ML models has solved a lot of commercial issues. Let’s look at what Image Separation can do for you in this area. Image classification is used by a face identification system to recognize employees and allow them to effortlessly register their presence. In the medical business, categorization in image analysis is used to identify illnesses, tumours, and tissue and cell structures from diverse medical imaging provided by radiographs, Magnetic resonance imaging (MRI), endoscope, computed tomography, ultrasound, and other methods. Satellite pictures are analysed to detect distinct trends, items, geographic outlines, soil data,

Fig 1. Image segmentation process
and other features that may be utilized in agriculture, mining, geo-sensing, and other applications. In robots, such as RPA and self-driving automobiles, image segmentation offers a wide range of applications. Hazardous items, hazards, individuals, and situations may be detected using security photos. The approach makes considerable use of image segmentation application programmed in MATLAB, Python or any other programming languages.

A typical case is the television program on the processing plants for food whereby grapes on the conveyor systems are assessed through computerization. In this case, the gathering high-speed images with cameras were strategically positioned, and it was sending commands to a compressor robot to take off rotten, unripe, or otherwise destroyed tomatoes while leaving the great ones to continue to grow. This is a simple, but important, application of image identification, in which the technique was able to extract just the needed elements from a picture, and the system then ranked those pixels as excellent, terrible, or ugly. A seemingly simple solution was having a huge influence on that company, eliminating human error and enhancing productivity. Python, among other traditional languages such as Matlab, C/C++, and others, is commonly used for image segmentation. Image classification in Python has most likely been the most in-demand expertise in the machine learning platform. Section III focuses on the scope of image segmentation.

III. SCOPE OF IMAGE SEGMENTATION

The Approach
When trying to acquire a bird's eye perspective of the Picture Separation responsibilities, one can see a critical step in action: object recognition. Image classification is used in many kinds of applications, from basic to elaborate. As we previously explained, detection is feasible because picture segmentation methods attempt to group like pixels together while separating different pixels. Two ways based on image attributes are used to do this:

Region Approach
The basic idea behind this method is to find comparable pixels in a picture using a threshold, region expanding, area extending, and regions combining. Segmentation and categorization, which find similarities depending on features, are samples of machine learning approaches, which utilize this methodology of identifying the similarity aspect on the unexpected collection of data.

Boundary Approach
This approach represents a reverse on the resemblance discovery approach that considers the disconnections instead of the similarities. Point detection, line detection, and object recognition are some of the sample algorithms for image segmentation, which utilize this approach to identify the edges using various parameters of discontinuity e.g. intensity. The algorithms of image classification are structured using different strategies centred on two strategies. These strategies are utilized based on the image type, which should be handled and evaluate, and fall into three different categories.

Types
Architectural Segments
These methods need that we first understand the architectural data of the picture being scanned. Dimensions, pixel resolution, distribution, graphs, color distributions, and so on may all be included in this. Second, we'll need contextual information about a location we're going to extract from the picture — this segment is all about the determination of the designated targets that are significantly novel to economic concern we are stimulated to address. In these segments of algorithms, a similarity-oriented methodology will be applied.

Stochastic Segmentation
Other than identifying the system of the required segment of the images, the fundamental data required for this class of process is to identify the progressive image pixel of the general medical images. This is vital when handling massive numbers of images whenever there are a lot of confusion about which item inside an object is needed. This method is used in ANN and Computer Training algorithms that employ k-means, for example.

Hybrid Approach
These image segmented techniques, as the name implies, combine structure and probabilistic approaches, i.e., they employ both the architectural and random image data of the picture.
IV. IMAGE SEGMENTATION TECHNIQUES

With respect to image segmentation approaches and the types of processing that need to be incorporated to fulfill a certain ambition, there are methodologies for image segmentation: Artificial Neural Network-based segmentation, Watershed-based method, Clustering-based segmentation, Region-based segmentation, Edge-based segmentation, and Threshold method.

Table 1. Summarization of the merits and demerits of image segmentation techniques

| Technique                  | Definition                                                                 | Merit                                                                                           | Demerit                                                                                          |
|----------------------------|---------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| Threshold technique        | Concentrates on identifying the ultimate value centered on histograms of images to identify the same pixel | Does not necessitate complex pre-processing; it is simplified                                      | Multiple details can be eliminated; threshold errors are most probable                            |
| Edge-based technique       | Centered on the discontinuity identification unlike the detection of similarity | Effective for the images with better contrasts between medical images                           | Is non- suitable for a noisy image                                                               |
| Region-based technique     | Centered on the partitioning of images into homogenous parts               | Operates effective for the images that are noisy; takes user markers for prompt assessment     | Memory-consuming and time-consuming                                                              |
| Traditional Segmentation Algorithms | Subdivides images into “k” homogeneous, mutually-exclusive clusters hence making it easier to obtain an object | Effective approach, fuzzy logic and fundamental for actual-time applications                    | Cost functions for the purpose of minimization can be challenging to determine                   |
| Watershed techniques       | Centered on the topology interpretations of the boundaries presented by an image | The segments gotten are stabilized, detected boundaries are different                            | Gradient calculations for the ridge is complicated                                                |
| Neural Network             | Centered on machine learning methods – Convolutional Neural Network (CNN) | Easy application, does not require following more complex techniques; ready-made repositories present in Python; considers more practical implementations | Training the system for business and custom images is resource consuming, time-consuming and costly |

Threshold Method

This is the simplest common, but effective, method for identifying the needed items in a picture. The frames in a picture are separated depending on brightness by measuring the intensity of each pixel to a threshold level. When the elements in the picture are expected to have higher intensity compared to the backdrop (and undesired elements) of images, the threshold technique is favorable. The thresholds variable T is regarded as a variable at its most basic level. However, given the quantity of distortion (undesired data) in the picture, that technique may be fruitless. As a consequence, we have the option of keeping it constant or changing it dynamically depending on picture parameters to get better outcomes. The following methods of thresholding are based on this: This approach substitutes black or white pixels for pixels in a picture. If the amplitude of Ii, j image and i, j location is beneath T (threshold), whereby pixel is replaced with black; if it is greater, that pixel is replaced with white. This is a multimodal thresholding method.

Otsu Binarization

We utilized an arbitrary number for the threshold value in global thresholding, and it stays constant. How can we determine and assess the validity of the chosen threshold is the main concern here. Trial and error is a less complicated yet ineffective strategy. Take, for example, a picture with two peaks in its histogram (bi-modal picture), one for the backgrounds and one for the picture plane. We may roughly choose a number in the center of those spikes as the barrier value for that picture, similar to Otsu binarization [3]. Simply said, it determines a threshold level for a bi-modal picture automatically from the pixel intensities. Nevertheless, for pictures that are not multimodal, the image size contains
numerous heights, or a single peak or category present has a fundamental variance, this is a disadvantage. Otsu’s Binarization, on the other hand, is frequently utilized in document scanning, eliminating undesirable colours from a statement, pattern matching, and other applications.

Thresholding Adaptive
When a picture contains various foreground and background illumination condition in various actionable localities, a globalized value as the threshold dimension might not be effective in all the situations. We require an adaptive approach, which is capable of altering the threshold for various picture components. In this instance, the approach breaks the images into small segments and identifies the threshold for each of them. As a result, different thresholds are obtained for various areas of the same picture. As a consequence, photos with variable light provide superior outcomes. The threshold value may be calculated automatically by the algorithm. The minimum value might be the average of the linear combinations or the closest location value based on the application of the Gaussian window weight (the window element to identify the regions).

Fig 2. Normal image

Fig 3. Image in Fig 2 segmented using thresholding

Edge Based Segmentation
The activity of detecting a boundary in an image is identified as the corner detection, and it is a fundamental phase in the understanding of image elements. The edge is considered to carry fundamental data and the fundamental features. It significantly diminishes the dimensions of the images to be evaluated and filters the potentially irrelevant datasets, maintaining an concentrating various interruptions in the grey levels, contrasts, saturation, brightness, texture and other variables to locate edges in an image. Additional production steps are required to improve the findings by concatenating all of the sides into edge strands that better match the image’s boundaries. Gradient-based techniques and Gray Diagrams are two of the most used types of edge identification algorithms [4]. These techniques use basic contrast enhancement technicians such as the sobel operator, canny, and Robert’s variable. These technicians assist in identifying edge irregularities and, as a result, identifying edge limitations. The final aim is to achieve at least a slight separation using this method, in which we combine all of the localized side into a novel digitalized image, which typically integrates the edge sequence, which matches the required physical elements or sections of images.
**Region Based Segmentation**

In [5], the method creates segmentation by splitting the picture into distinct elements with comparable features in region oriented segmentation approaches. Simply defined, these elements are a collection of images. Region-based image differentiation algorithms start by looking for certain seed locations in the input picture — either tiny bits or much larger chunks. Then, either additional pixel are added to the seed points, or the seed point is progressively diminished or shrunk to small chunks and merged with other small seed points. As a result, two main strategies based on this strategy exist.

**Growing Region**

It represents the bottom-up approach whereby we begin with a small group of pixels and gradually accumulate or merge them depending on pre-determined resemblance restrictions [6]. The region growth method begins by selecting a random seed image in the picture and comparing it to its neighbors. If surrounding pixels have a match or resemblance, they are joined to the original seed pixel, expanding the diameter of the zone. When we hit saturation and the expansion of that area can no longer continue, the program selects a new seed pixel, which must not belong to any of the presently existing sections, and repeats the process. Using region-growing algorithms, it is common to accomplish successful segmentation that closely matches the observed edges. However, if the system allows an area to develop entirely before attempting other
seeds, the classification is frequently biased in favor of the parts, which had initially been segmented. To effectively counteract this process, many algorithms begin with human application of resemblances initially; no one area is permitted to entirely dominate and expand and numerous regions may grow at the same time. Area growth, like thresholding, is a pixel-based technique, but the main distinction is that thresholding removes a huge region based on comparable pixels from different locations of the images; however, the region-growth tends to retrieve neighboring pixels. For noisy pictures where detecting the edges is challenging, region growth approaches are preferred.

**Splitting and Merging Regions**

Breaking and combining based segment methods divide a picture by using two fundamental approaches – region dividing and section integrating – in tandem. Splitting is the process of repeatedly separating a picture into sections with similar properties, whereas merging is the process of uniting neighboring regions that are related in some way. Whereas region development, a region split supports the entire input picture as the corporate interest area. Then it tries to match a specified set of variables or pre-defined similarity restrictions, and it selects all pixel regions that satisfy the requirements. In contrast to the area expansion algorithm, this is a divided and rule strategy [7]. After executing the split procedure, we will have different aspect of similarly market location distributed on the image pixel, considering that an ultimate portion will integrate the dispersed clusters of adjoining parts with comparable and identical elements. To complete the whole process, it is fundamental to execute the process of merging that examines nearby areas after each split and, if necessary, combines them commonly used evaluation degrees. Split-merge methods are a kind of split-merge technique.

![Fig 7. Normal Image](image)

![Fig 8. Fig 1 image segmented using region-based](image)

**Clustering Based Segmentation Methods**

In [8], segmentation methods, unlike Categorization techniques, are unregulated, meaning the user does not have a pre-defined gathering of the features, groups and classes. Segmentation methods help in the process of extracting fundamental, secret knowledge from data, such as frameworks, strands, and groups, which are often unidentified from a heuristic standpoint. The picture is segmented into regions or discontinuous sets of images with similar features using clustering algorithms. The data components are divided into clusters using fundamental Data Clustering characteristics, with components in the same cluster being more comparable to one another than ones in other clusters. In the clustering-based methodologies offered, some of the most effective clustering methods are applied: fuzzy c-mean, enhanced k-means, k-
means, and enhanced fuzzy c-mean methods. Due to its computation and simple effectiveness, k-means groupings are selected as a popular approach. The Improved K-means method may reduce the amount of iterations that a k-means algorithm typically requires. The FCM technique enables data points (in this example, pixels) to be classified into numerous classes with variable membership values. Enhanced FCM overcomes the slower processing speed of an FCM. We may use numerous analytical measures such as the Silhouette Factor, rand index (RI), and others to assess the integrity of the sections created by clustering-based ML algorithms.

**Clustering using the k-means method**

K-means is one of the major unsupervised learning methods for clustering issues [9]. The method employs a straightforward and straightforward approach to classifying a given picture using a predetermined number of clusters. The method begins here, when the picture space is partitioned into k pixels, each indicating a cluster centroid. Each item is now allocated to a culture based on its proximity to the cluster. The coordinates now shift and are redistributed after all of the images are allocated to all of the clusters. This process is repeated until the centroids are unable to move any more. We have regions of the picture divided into "K" groups with certain degrees of similarity when this method converges.

As mentioned in the preceding chapter, the fuzzy C Means k-means algorithm enables for separating and categorizing pixels in a picture that have specific levels of resemblance. The fact that the clusters and their individuals are continuously incompatible is one of the most remarkable aspects of k-means. The data sets, or pixels in our example, may be grouped in many clusters using a Fuzzy C Implies clustering algorithm. To put it another way, a set of pixels may belong to many clusters or groups, but their degrees of interconnection within each group might vary. The FCM method is coupled with an optimisation problem, and the reduction of this mechanism is required for the process to converge. We have regions inside the picture split into "C" sections at the completion of this method, where the component pixels within a category exhibit certain degrees of likeness, as well as a level of linkage with other organizations.

**Methods Using Watersheds**

Watersheds are a ridge method, which is likewise a region-based technique, and is focused on topographical interpretations. For different elements of a picture, we explore the geographic environment comparison with hills and valleys. The gray numbers of the individual pixels – known as the integral image – clearly quantify the slope and height of the aforementioned terrain. The hydrological transformation disintegrates a picture into sections known as "catchment areas" using this 3D model, which is often used for Earth sceneries. A watershed basin is made up of all images whose sharpest decline in gray levels ends at the local minimal level. In a nutshell, the algorithm treats a pixel as the “elevation” (surface tomography) and therefore begins with user-defined landmark [10]. The programs therefore form basins that are minima spots, and the catchment is inundated from the marker to watersheds lines. Watersheds divide basins in this area. As a result of the pixels given to each area or watershed, the image is deconstructed.

**Method Using Artificial Neural Networks**

Image Classification is the term used to describe the process of employing neural networks to segment images [11]. It use artificial intelligence to analyze and recognize visual elements such as products, faces, text, handwritten writing, and so on. Because of its ability to recognize and handle high-definition visual data, Artificial Neural Networks are utilized particularly for this procedure. A picture is either a collection of matrices (color tagged polygons) or a raster, depending on the technique employed (a table representing pixels with values for the color). The matrix or graphics are segmented into smaller components, which show the real object and characteristics that make up a picture. Object detection algorithms can rationally examine these structures by selecting the most significant portions and then arranging the information using feature engineering and categorization methods. These algorithms are employed in many different sectors and professions. Users' search needs and browsing history are used by the e-commerce business to provide relevant items to them. It's used in the industrial business to identify anomalies, damaged goods, and worker protections, among other things. For visually challenged and speech disabled pupils, image recognition is well-known in learning and skills. Though training programmes for Neural Nets takes time, the final findings have been quite encouraging, and their implementation has been very effective [12].

**V. PRE-REQUISITES AND APPLICATION**

Image analysis has been deployed in a number of computer languages, including Java, Matplotlib, C++, and others. Python provides strong libraries to execute multiple approaches used in Image Classification because of its essential characteristics of flexibility, varied implementation, and usage in computer science stack, computer vision, and supervised learning.
Scikit-image, OpenCV, Mahotas, Pillow, matplotlib, SimpleITK, and other Python modules are well-known for implementing image analysis in particular and picture classification in specific. Using Python libraries is a simpler method of deployment that does not need any difficult prerequisites prior to deployment - save for a basic understanding of Python programs and pandas, of course. Certain fundamental abilities are required to have better influence over the dark box libraries employed in this procedure. Deep learning, together with statistical inference, is among the most important criteria for the architecture in general, as well as for image analysis division. Understanding of linear programming and differential calculus, and the mathematical expression enhances the control over image identification pre-process methods.

The equipment and materials of neural networks, mostly those centered on deep learning, is required for ANN image analysis and division in image detection and analysis approaches. A rudimentary understanding of modern signals handling is necessary for noise reduction, recognizing contours, and creating histogram – a process, which is pre-applied for a number of elements described previously.

VI. CONCLUSION

Image differentiation is the division of a digitized picture into several parts in computerized computer vision and image processing (pixels, which are also referred to as the image object). The main aim of image classification is to make images simpler and intelligible to evaluate by changing or simplifying its presentation. In many instances, image classification is applied to identify the constraints and components (such as curves and lines) in images. Image classification, to put it another way, is the method of giving a label to each image in the picture such that objects with same identity share particular properties. Segmentation process is a potential set of Neural Learning abilities since it plays an essential part in diagnostic imaging, and many groups are working to develop an efficient program for proactive detection using medical imagery. One of the most well-known uses is in cancerous cells detection methods, where Image Classification proved crucial in detecting sick tissues and organs from imaging more quickly, allowing doctors to treat patients sooner. The production sector increasingly heavily depends on machine vision systems to discover abnormalities that might otherwise go undetected by human eyes, resulting in increased improved efficiency. The ability to do image segmentation using Python is in high demand, and there is plenty of training provided. Both classic image processing methods and Artificial Neural implementation must be well understood. Python allows for a clear approach that requires little code and commitment.

References

[1]. R. Bergman and H. Nachlieli, “Perceptual Segmentation: Combining Image Segmentation With Object Tagging”, IEEE Transactions on Image Processing, vol. 20, no. 6, pp. 1668-1681, 2011. Available: 10.1109/tip.2010.2088970.
[2]. K. Huang and T. Tan, “Vs-star: a visual interpretation system for visual surveillance”, Pattern Recognition Letters, vol. 31, no. 14, pp. 2265-2285, 2010. Available: 10.1016/j.patrec.2010.05.029.
[3]. P. LECH, “Binarization of document images using the modified local-global Otsu and Kapur algorithms”, PrzeglID Elektrotechniczny, vol. 1, no. 2, pp. 73-76, 2015. Available: 10.15199/2815.2015.02.18.
[4]. I. M. B. M. N. P. S. P. S., "Development Of Edge Detection Algorithms For The Identification Of Regional Wall Motion Abnormalities In A Canine Model", Anesthesiology, vol. 81, no., p. A546, 1994. Available: 10.1097/00000542-199409001-00545.
[5]. L. Zhengdong, "A Garment Image Segmentation Method Based on Salient Region and JSEG", Journal of Software, vol. 10, no. 11, pp. 1274-1282, 2015. Available: 10.17706/jsw.10.11.1274-1282.
[6]. J. Bonito, “A Bottom-Up Approach to Examining Group-Level Communication Patterns: A Multilevel Latent Profile Analysis of Functional Group Interaction”, Human Communication Research, vol. 45, no. 2, pp. 202-225, 2019. Available: 10.1093/hcrr/hby020.
[7]. R. Feiz and M. Rezghi, “A splitting method for total least squares color image restoration problem”, Journal of Visual Communication and Image Representation, vol. 46, pp. 48-57, 2017. Available: 10.1016/j.jvci.2017.03.001.
[8]. K. Bhuvaneswari and P. Geetha, "Semantic Segmentation and Categorization of Brain MRI Images for Gioma Grading", Journal of Medical Imaging and Health Informatics, vol. 4, no. 4, pp. 554-566, 2014. Available: 10.1166/jmih.2014.1286.
[9]. M. Lopez de Prado and M. Lewis, "Detection of False Investment Strategies Using Unsupervised Learning Methods", SSRN Electronic Journal, 2018. Available: 10.2139/ssrn.3167017.
[10]. L. B. B. W. W. M. M., "Image Segmentation: A Watershed Transformation Algorithm", Image Analysis & Stereology, vol. 28, no. 2, p. 93, 2011. Available: 10.5566/ias.v28.p93-102.
[11]. S. Yan, S. Sayad and S. Balke, "Image quality in image classification: Adaptive image quality modification with adaptive classification", Computers & Chemical Engineering, vol. 33, no. 2, pp. 429-435, 2009. Available: 10.1016/j.compchemeng.2008.10.017.
[12]. Haldaroi and A. Ramu, “An Intelligent-Based Wavelet Classifier for Accurate Prediction of Breast Cancer,” Intelligent Multidimensional Data and Image Processing, pp. 306–319.