A New Aerial Image Segmentation Approach with Statistical Multimodal Markov Fields

Jamal Bouchti[1], Ahmed Bendahmane[2], Adel Asselman[3]
Optique and Photonic Team-Faculty of Sciences, Abdelmalek Essaadi University, Tetuan 93002, Morocco[1, 3]
Department of Computer Science-ENS, Abdelmalek Essaadi University, Tetuan 93002, Morocco[2]

Abstract—Aerial images, captured by drones, satellites, or aircraft, are omnipresent in diverse fields, from mapping and surveillance to precision agriculture. The efficacy of image analysis in these domains hinges on the quality of segmentation, and the precise delineation of objects and regions of interest. In this context, leveraging Markov fields for aerial image segmentation emerges as a promising avenue. The segmentation of aerial images presents a formidable challenge due to the variability in capture conditions, lighting, vegetation, and environmental factors. To meet this challenge, the work proposes an innovative method harnessing the power of Markov fields by integrating a multimodal energy function. This energy function amalgamates key attributes, including color difference measured by the CIEDE2000 metric, texture features, and detected edge information. The CIEDE2000 metric, derived from the CIELab color space, is renowned for its ability to measure color difference more consistently with human perception than conventional metrics. By incorporating this metric into the energy function, the approach enhances sensitivity to subtle color variations crucial for aerial image segmentation. Texture, a vital attribute characterizing regions in aerial images, offers crucial insights into terrain or objects. The method incorporates texture features to refine the separation of homogeneous regions. Contours, playing a fundamental role in segmentation, are identified using an edge detector to pinpoint boundaries between regions of interest. This information is integrated into the energy function, elevating contour consistency and segmentation accuracy. This article comprehensively presents the methodological approach, the conducted experiments, obtained results, and a thorough discussion of the method's advantages and limitations.

Keywords—Image segmentation; multimodal markov fields statistical integration; CIEDE2000 color difference; texture features; edge information

I. INTRODUCTION

Aerial image segmentation is identified as a critical area within the domain of image processing, indispensable for a breadth of applications from environmental monitoring to precision agriculture. The objective of segmenting an image into meaningful regions presents notable challenges due to the diversity and complexity of landscapes captured, variable lighting conditions, and the occurrence of atmospheric phenomena [1]. Therefore, segmentation techniques need to be robust and precise to identify objects and areas of interest effectively [2].

Traditional segmentation approaches, including thresholding methods [3] [4], region growing [5], contour-based techniques [6], and pixel classification [7], are fundamental but exhibit limitations when confronted with the complexity of aerial imagery. For example, thresholding is simple to implement but struggles with intensity variations across images, and region growing demands substantial computational resources and can be compromised by noise. Conversely, deep learning techniques such as convolutional neural networks (CNNs) have advanced the field significantly by facilitating nuanced and precise semantic segmentation, capitalizing on their ability to learn complex features from extensive datasets [8] . Despite the efficiency of these deep learning methods, challenges persist, including the requirement for vast amounts of annotated data for training and a considerable demand for computational power.

Additionally, the selection of an appropriate color space for segmentation remains an unresolved issue, as each space has its own set of benefits and drawbacks. The RGB space, for instance, despite being widely used for display, proves less efficient for segmentation due to the high correlation among its components [9].

In response to these challenges, a novel method based on multimodal Markov fields has been introduced, representing a promising alternative adept at handling the intrinsic diversity and complexity of aerial images. By leveraging the strength of multimodal Markov fields [10], this approach aims to surpass the limitations of both traditional methods and deep learning by integrating multimodal information for more accurate and robust segmentation. This integration enables the capture of spatial dependencies between pixels and subtle variations in texture and color, facilitating detailed segmentation that is finely tuned to the unique challenges of aerial imagery.

The multimodal strategy not only facilitates a clearer distinction between objects and areas of interest but also provides the adaptability required to manage different lighting conditions and atmospheric variances without the need for extensive annotated data sets for training. This approach introduces a sophisticated technique that utilizes the potential of Markov fields through a multimodal energy function. This function integrates several critical attributes, such as color differences measured by the CIEDE2000 metric, texture features, and information from edge detection.

The CIEDE2000 metric, based on CIE Lab color spaces, is recognized for its ability to measure color differences in a manner that aligns more closely with human visual perception compared to traditional metrics. By incorporating this metric,

the method can better account for subtle color variations, which are crucial for the segmentation of aerial images.

Texture, an important characteristic for defining regions in aerial images, provides essential information about the nature of the terrain or objects. Texture features derived from the HSV (Hue, Saturation, Value) color space, known for its reduced sensitivity to lighting variations compared to the RGB space [11], are utilized to improve the separation of homogeneous regions.

Moreover, considering the fundamental role of contours in segmentation, an edge detector is used to identify the boundaries between regions of interest0. This information is incorporated into the energy function to improve the consistency of contours and the accuracy of segmentation.

The paper thoroughly presents the methodological approach, the experiments conducted, the results achieved, and a detailed discussion on the advantages and limitations of the method, opening up new perspectives for the analysis of aerial images (see Fig. 1) in various application domains.

Fig. 1. Aerial image.

II. Theoretical Foundations

Successful segmentation of aerial imagery relies on a solid theoretical foundation, integrating a variety of techniques and measurements. In this section, the essential theoretical underpinnings of the multimodal segmentation approach are explored. Segmenting an image Y involves dividing all the pixels S into homogeneous regions: $S = S_1 \cup S_2 \cup ... \cup S_K$.

The label map (Xs, s∈S) is introduced to represent a partition: pixel $s \in S_j \Leftrightarrow X_s = j$.

The probabilistic modelling approach to the segmentation problem consists of:

- Consider the image $Y = (Y_s)$ and the label map $X = (X_s)$ (to be constructed) as random variables governed by a statistical law π;
- propose a modelling ≡ define such a law π;
- With X and Y linked by the law π, and Y given, reconstruct or estimate X using π and Y.

Note that if the law of image formation F:

$$X = (X_s) \text{ a } Y = (Y_s) = F(X)$$

If it were completely known, the only task would be to invert F! However, such a deterministic function F is unrealistic, because the mechanism of image formation is complex, to say the least, and is marred by noise, i.e. the randomness or handling errors that occur. The probabilistic model approach defines passages by conditional statistical laws. Markov fields are some of the most widely used examples of such laws.

A. Markov Fields in Image Segmentation

Markov fields are a powerful mathematical framework widely used in computer vision [12], particularly for image segmentation [13]. They provide a structured way of modeling the spatial dependencies between pixels in an image. In a segmentation context, Markov fields are used to capture the spatial regularity of regions of interest. More specifically, they model the neighborhood relationships between pixels and facilitate the propagation of information about whether pixels belong to a particular class. The notion of neighborhood is then defined [14], which designates a set of pixels located around a central pixel. Consider a pixel S whose position in the image is given by the coordinates (m, n). Its affix is therefore s = (m, n). A neighborhood of S, denoted V(S), is defined as a set of connected pixels P' defined by:

$$N(i,j) = \left\{ (k,l) \mid 0 < (k-i)^2 + (l-j)^2 < constant \right\}$$

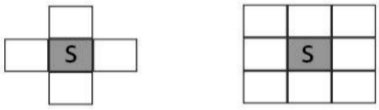

Fig. 2. 4 and 8 neighborhoods.

A clique is any subset A of sites that are mutual neighbors Fig. 2.

Examples of cliques are shown in Fig. 3:

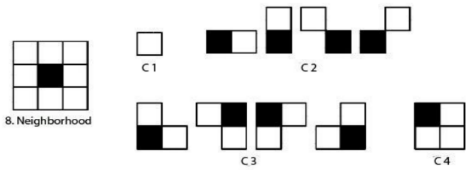

Fig. 3. Cliques for 8 neighborhoods.

1) Gibbs distribution: Gibbs fields are commonly used to model thermodynamic systems in statistical physics. The Gibbs distribution is a central concept in MRFs. This equivalence means that the interaction potential between random variables follows a Gibbs distribution [15]. This makes it possible to describe the interactions between the variables in a coherent way, while maintaining the notion of spatial dependence [16].

2) Hammersley-Clifford theorem: The Hammersley-Clifford theorem is a result in probability theory, mathematical statistics and statistical mechanics that gives the necessary and sufficient conditions under which a strictly positive probability distribution (of events in a probability space) can be represented as events generated by a Markov random field [17].This is the fundamental theorem of random

fields, which states that a probability distribution with strictly positive mass or density satisfies one of the Markov properties with respect to an undirected graph G if and only if it is a Gibbs random field, i.e. its density can be factored over the cliques (or complete subgraphs) of the graph. In other words, this theory states that the probability of a configuration of states depends mainly on the local relationships between the random variables in the field [18].

B. CIEDE2000 Color Difference

The CIEDE2000 metric, derived from CIE Lab color spaces, plays a central role as a color attribute in the approach. Designed to measure color difference more accurately[19], CIEDE2000 takes into account the non-linearities of human perception of color. It subtly captures variations in hue, saturation and luminosity, offering a more robust measurement of color difference than its predecessors [20].

The individual components of this formula are as follows:

$$\Delta E_{00} = \sqrt{(\Delta L'/K_L S_L)^2 + (\Delta C'/K_C S_C)^2 + (\Delta H'/K_H S_H)^2 + RT(\Delta C'/K_C S_C)(\Delta H'/K_H S_H)}$$

Where:

ΔL': Difference in luminance between Lab_1 and Lab_2.

ΔC': Difference in chroma (color intensity) between Lab_1 and Lab_2.

ΔH': Hue difference (hue of the color) between Lab_1 and Lab_2.

The S_L, S_C and S_H components are adjustment factors to take account of non-linearities in the perception of color by the human eye:

S_L: Adjustment factor for luminance.

S_C: Chromaticity adjustment factor.

S_H: Tint adjustment factor.

The k_L, k_C and k_H values are parameters that depend on the luminance of the sample and the color of the average sample.

RT is an additional correction factor.

Integrating the CIEDE2000 metric into the energy function enables more accurate segmentation by considering the subtle nuances of color present in aerial images.

C. Texture as a Segmentation Attribute

Texture is an essential element for characterizing regions of interest in aerial images. It represents the repetition of patterns or structures and can provide crucial information about the nature of the terrain or objects.

1) Co-occurrence matrix in the HSV Space: The co-occurrence matrix, also known as the correlation matrix, is a powerful image processing technique that quantifies the spatial relationships of grey levels or pixel values in an image. The co-occurrence matrices contain a very large amount of information and are therefore difficult to manipulate. For this reason, fourteen indices (defined by Haralick) [21] which correspond to descriptive characteristics of textures can be calculated from these matrices. In the context of the study, An innovative approach is taken using the HSV (Hue, Saturation, Value) color space. Specifically, The focus is on the hue (Hue) and intensity (Value) components. Hue represents color tone, while intensity captures luminance. The aim is to exploit these two components to assess the homogeneity and correlation of textures in aerial images.

a) Homogeneity: The more frequently the same pair of pixels is found, the higher this index becomes, for example in a uniform image, or a texture that is periodic in the direction of translation Fig. 4.

$$Homogeneity = \sum_{i=1}^{q-1}\sum_{j=1}^{q-1}\frac{P(i,j)}{1+|i-j|}$$

Fig. 4. Homogeneity result.

b) Correlation: describes the correlations between the rows and columns of the cooccurrence matrix Fig. 5.

$$Correlation = \sum_{i=1}^{q-1}\sum_{j=1}^{q-1}\frac{ijP(i,j)-u_i u_j}{\sigma_i \sigma_j}$$

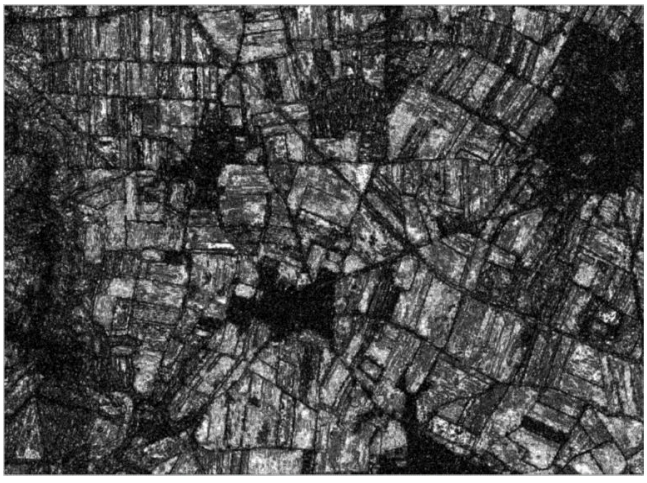

Fig. 5. Correlation result.

The co-occurrence matrix in HSV space allows us to analyze how hue and intensity values co-vary within a local image window. This approach is essential for detecting regions with similar textures based on variations in hue and intensity. More formally, the co-occurrence matrix tells us the joint probability of observing a pair of hue and intensity levels in each neighborhood. This information is then used to calculate texture measures such as homogeneity and correlation, which are incorporated into the Markov field model to improve aerial image segmentation. These characteristics make it easier to distinguish homogeneous regions, enhancing the quality of the segmentation.

D. Role of Contour Detectors

Contours play a fundamental role in the segmentation of aerial images. Precise delineation of regions of interest depends largely on edge detection.

Several methods have been developed to accomplish this task [22], each with its own advantages and disadvantages as shown in Table I.

TABLE I. SEVERAL METHODS FOR EDGE DETECTION, ADVANTAGES, AND DISADVANTAGES

Method	Advantages	Disadvantages
Directional Derivatives[23]	-Simple implementation; -Effective for sharp contours	-Sensitive to noise; -Reaction to varying brightness
Edge Detection Filters	-Flexibility in adjusting filters; -Easily extendable to color image	-Sensitive to noise; -Limited response to diagonal contours; -Excessive smoothing on curved contours
Laplacian Operators[24]	-Sharp edge detection; -Robust to variable lighting; -Capable of detecting fine contours	-Sensitive to noise; -Computationally intensive; -Limited response to subtle details
Hough Transform[25]	-Robust to noise presence; -Can detect non-linear contours	-High computational cost; -Sensitive to discontinuities; -Parameter tuning required
Canny Edge Detector[26]	-Good detection of sharp contours; -Effective noise suppression; -Accurate localization	-Sensitive to parameter settings; -Computationally intensive; -May be sensitive to weak contours

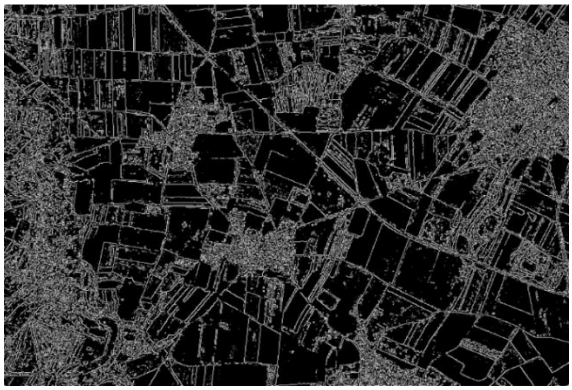

Fig. 6. Edge detection result.

A single edge detector may be limited in its ability to capture the diversity of existing edges.

This is why a combination of detectors is used in the approach, each bringing its own specific expertise to highlight certain types of contours to identify the boundaries between objects and structures present in the image.

The information extracted by these edge detectors Fig. 6 is incorporated into the Markov field energy function, which promotes edge coherence between pixels and, as a result, more robust and accurate segmentation.

III. METHODOLOGY

In this section, the methodology developed for the segmentation of aerial images using Markov fields with a multimodal energy function is described in detail.

A. Multimodal Approach to the Energy Function

The segmentation approach is based on a multimodal energy function, designed to capture various key features of aerial imagery simultaneously. This energy function integrates color difference based on the CIEDE2000 metric, texture features, and detected edge information.

The aim of this approach is to improve the consistency and accuracy of segmentation by taking advantage of several key attributes.

B. Combining Attributes in the Energy Function

1) CIEDE2000 color difference: The CIEDE2000 metric is integrated into the energy function as a measure of pixel similarity. It encourages the grouping of pixels that share similar color characteristics, while taking subtle color nuances into account.

2) Texture: Texture characteristics are extracted from aerial images and used to assess the textural coherence of regions. This component of the energy function distinguishes homogeneous regions from textured areas, contributing to more accurate segmentation.

3) Contour detector: Detected contour information is incorporated to encourage contour consistency in segmentation. This component aims to ensure that the boundaries of the regions of interest are well defined.

C. Comparative Analysis of Models

In the quest for the most effective method for segmenting aerial images, a range of models was evaluated, each possessing unique characteristics and capabilities Table II. Central to this analysis were Markov Random Fields (MRF), known for their robust modeling of spatial interactions, alongside Conditional Random Fields (CRF), Deep Learning techniques, Graph Cut optimizations, and the Watershed algorithm. The choice of model significantly impacts the quality of segmentation, particularly in complex scenarios such as aerial imagery, where accuracy, detail, and computational efficiency are paramount. A comprehensive comparison is provided below, highlighting the strengths and weaknesses of these models:

After thorough consideration of the advantages and limitations of each model, the decision to utilize Markov Random Fields (MRF) for the segmentation of aerial images was driven by the model's exceptional ability to handle spatial complexities and the rich textural and contour information inherent in aerial imagery. Despite the computational demands, the flexibility and robustness of MRFs, particularly when combined with an efficient optimization algorithm like Iterated Conditional Modes (ICM), offer a sophisticated balance between detail accuracy and processing efficiency. This makes MRF an ideal choice for studies aiming at high-quality segmentation of aerial images where precision and reliability are crucial.

TABLE II. VARIOUS IMAGE SEGMENTATION METHODS, ADVANTAGES, AND DISADVANTAGES

Model	Characteristic	Advantages	Disadvantages
MRF	Generative model emphasizing spatial interactions among pixels.	- Models spatial dependencies effectively. - Robust to local variations. - Suited for images with complex textures and structures.	- Computationally intensive. - Energy function definition demands precision.
CRF	Conditional model focusing on pixel label dependency on observed data.	- Integrates global and local information. - Modelling of conditional dependencies is flexible.	- High computational complexity for inference. - Feature and parameter selection is critical.
Deep Learning	Data-driven approach using neural networks for feature extraction and segmentation.	- Capable of learning from large data sets. - Demonstrates excellent performance across diverse tasks.	- Requires extensively annotated data sets. - Interpretability and control over decisions are limited.
Graph Cut	Optimization model based on graph theory, aiming to minimize a cost function for segmentation.	- Captures global image properties effectively. - Yields precise and clean segmentations.	- Initialization sensitivity. - May over-segment highly textured images.
Watershed	Morphological model that segments images based on gradient analysis, treating images as topological surfaces.	- Intuitive and straightforward to implement. - Effective at outlining object boundaries in high-contrast images.	- Prone to over-segmentation in noisy contexts. - Frequently necessitates post-processing for optimal segmentation.

D. Markov Field Model

The Markov field model is the underlying structure of the segmentation method. It is used to model the spatial relationships between pixels and to propagate information about whether pixels belong to a particular class. The MRF model is employed to describe the spatial dependency of pixels and attributes within the image. This allows us to efficiently exploit the multimodal information embedded in the energy function.

The Energy function incorporates color difference, texture and contour detector attributes, enabling a comprehensive approach to aerial image segmentation that is sensitive to subtle variations in color, texture, and shape:

The energy function E is defined as the sum of three terms:

$$E(I,S)=\alpha\ E_{color}(I,S) + \beta\ E_{texture}(I,S) + \lambda\ E_{contour}(I,S)$$

Where:

α, β, λ are weighting coefficients to control the influence of each term of the function.

I represent the input image.

S is the map of segmentation labels, where each pixel is associated with a class (object or background).

Each term in the energy function is defined as follows:

1) CIEDE2000 Color Difference Term: Ecolor (I,S) measures the color difference between pixels in the same region (class) in the segmented image I_S using the CIEDE2000 metric. It encourages color consistency within each region:

$$E_{\text{Color}}(I, S) = \sum_r \sum_{p \in r} \Delta E00(I_p, \mu_r)$$

The average μr essentially represents the average color of region r in CIELab space. It is calculated by traversing all the pixels that belong to the region reconverting its color components (L, a, b) in CIELab space, summing them to obtain three sums: ΣL, Σa and Σb, then devising each sum by the number of pixels N in the region r to obtain the average components μr of the region r.

To optimize processing, a region graph is constructed with the number of pixels N and the mean μr, which is updated each time a pixel is added to a region.

$\Delta E00\ (I_p\ ,\ \mu_r)$ is the CIEDE2000 color difference between pixel Ip and the average color of region r (μ_r). The lower ΔE00 is, the more similar the color of pixel Ip is to the average color μr of region r.

2) Texture term: $E_{texture}(I, S)$ evaluates the texture in each segmented region. Texture measurements will be used based on the co-occurrence matrix calculated from the variation of the Hue and intensity attributes of the pixel color in HSV space with respect to the average of the region and neighborhood to which it belongs, to promote the homogeneity of textures within each class:

$$E_{\text{texture}}(I, S) = \sum_r \sum_{p \in r} 1 - H(I_p, \mu_r)$$

$H(I_p\ ,\ \mu_r)$ is a measure of the homogeneity of the texture of pixel Ip with respect to the region μ_r to which it belongs, The higher H is, the more homogeneous the texture.

The homogeneity measure from the co-occurrence matrix is already a normalized value between 0 and 1, where 0

represents minimum homogeneity (maximum variability) and 1 represents maximum homogeneity (no variability). This is the reason the homogeneity value is subtracted from 1, to minimize the energy function the more homogeneous the region becomes.

The steps below will be followed to calculate H (I_p , μ_r):

1) Calculate the co-occurrence matrix for the region r: For each pixel Ip in region r, examine the Hue and intensity attributes of the neighboring pixels (we'll use neighborhood 8) in region r. Create the co-occurrence matrix, which records the frequency of pairs for each attribute and is generally symmetrical.

2) Normalize the co-occurrence matrix: Each element of the co-occurrence matrix is divided by the sum of all the elements of the matrix to normalize the values in the range 0 to 1. This step produces a co-occurrence probability matrix.

3) Calculate homogeneity: The standardized matrix is used to calculate homogeneity, which is a measure of the inverse of the variation in the attributes used.

The formula used to calculate the homogeneity $H(I_p,\mu_r)$ is as follows:

$$H(I_p,\mu_r)=\sum_{i,j} 1+|i-j|2P(i,j)$$

P(i,j) is the probability of co-occurrence of chromaticity's i and j in the normalized matrix.

|i-j| is the difference between chromaticity i and j.

4) Average homogeneity: Once the homogeneity has been calculated for each pixel Ip in region r, these values can be averaged to obtain an overall measure of the homogeneity of the texture in region r.

This measure will be used as a component of the energy function to encourage texture consistency within each segmented region. The higher the homogeneity, the more uniform the texture is, and vice versa.

3) Contour term: $E_{contour}(I, S)$ encourages contour consistency, Firstly, Edge detectors will be used to identify the edge locations in the image, and then an edge map will be built where the marked pixels or regions correspond to the edge locations. This map will contain binary values (edge or non-edge).

The following function is defined:

$$E_{\text{Contour}}(\text{I}, \text{S}) = \sum_r \sum_{pq \in r} D_{pq} \cdot \left|\text{S}_\text{p} , \text{S}_\text{q}\right|$$

D_{pq} is a factor based on contour detection between pixels Ip and Iq. It is calculated by comparing the contour values of neighboring pixels p and q in the contour map. If pixels p and q are neighbors and one is on the contour while the other is not, this indicates a label discontinuity along the contour:

If pixel p is on the contour (high contour value) and pixel q is not on the contour (low contour value), or vice versa, then D_{pq} is defined as a high penalty factor, D_{pq} = 1 (to strongly penalize label discontinuity).

If the two pixels p and q are both on the contour (or both outside the contour), it is defined as a low penalty factor, D_{pq} = 0 (so as not to penalize label consistency).

The expression "$|S_p - S_q|$" is a term which measures in absolute value the difference in labels (S_p and S_q) of pixels p and q within the same region of the segmentation and which penalizes label discontinuities to encourage their coherence within each region.

If "S_p" and "Sq" are the same (i.e. neighboring pixels have the same label), then "$|S_p - S_q|$" is zero. This means that there is no penalty for label consistency, as the labels are already the same.

On the other hand, if "S_p" and "Sq" are different (i.e. neighboring pixels have different labels), then "$|S_p$ - Sq|" is greater than zero. This means that there is a penalty for label discontinuity within the same region. This penalty encourages the model to assign similar labels to neighboring pixels in the same region, thereby promoting the consistency of the segmentation.

This energy model integrates the three attributes (color difference, texture, and contours) to promote the coherence of the segmentation regions by taking into account the visual and structural characteristics of the pixels. Segmentation is achieved by minimizing this energy function using the Iterated Conditional Modes (ICM) optimization algorithm.

4) ICM algorithm for segmentation optimization: The segmentation is optimized using the Iterated Conditional Modes (ICM) algorithm. This is an efficient iterative algorithm that iterates through the set of pixels taking into account spatial dependencies and the multimodal energy function and seeks to find the best pixel label configuration that minimizes the energy function E(I,S) and corresponds to the most accurate segmentation [27].

However, an iterative algorithm without a stopping condition could continue to iterate indefinitely. Introducing this stopping condition saves computation time and resources by stopping the algorithm once convergence criteria are satisfied.

Here Table III shows some common stopping conditions for ICM:

TABLE III. COMMON STOPPING CONDITIONS, ADVANTAGES, AND DISADVANTAGES

Stopping Criterion	Brief Description	Advantages	Disadvantages
Energy Convergence	Stops the algorithm when the energy converges, i.e., it ceases to decrease significantly.	- Can lead to rapid convergence.	- Sensitive to local energy minima.
Maximum Number of Iterations	Halts the algorithm after a fixed number of iterations.	- Precise control over execution time.	- May not converge if the number is too low.
Label Stagnation	Stops the algorithm when labels no longer change between successive iterations.	- Saves computation time.	- May lead to suboptimal segmentation.
Local Convergence	Halts the algorithm if labels locally converge around certain pixels.	- Accelerates local convergence.	- Risk of premature convergence.
Cross-Validation	Uses cross-validation to estimate model performance and stops when performance stabilizes.	- Suitable for avoiding overfitting.	- Can be computationally expensive.
Maximum Execution Time	Stops the algorithm after a predefined execution time.	- Controls overall execution time.	- May lead to suboptimal convergence.
Segmentation Quality Criterion	Stops the algorithm based on a specific measure of segmentation quality achieved.	- Directly optimizes segmentation quality.	- Depends on a subjective measure of quality.

In the method, a combination of several of these conditions will be used to ensure that the algorithm stops appropriately. More specifically, the global energy convergence criterion combined with the execution time will be applied to prevent the algorithm from running in an infinite loop.

Algorithm 1:

Input
S: Source image
α, β, λ: Ponderation parameters
ε: Convergence threshold
τ: Maximum execution time

//Initialization

Read the source image S
Convert S to HSV space S_{HSV}
Convert S to Lab space S_{Lab}
Calculate the contour map
Compute Homogeneity and correlation for each pixel x ∈ S
Initialize the LabelMap with Labels {1,2,.....,SHighxSWidth}

// ***Label Propagation***

For each pixel (x in S):
  If (P1 and P2 are Neighborhood Pixels And $P1_{indH} = P2_{indH} = 1$) then
  $P1_{Label} = P2_{Label} = \min(P2_{Label}, P2_{Label})$
  End
End

// ***Energy Minimization***

For each pixel (x in S):
  // ***Compute Energy Function***
  Calculate $E(I, S) = \alpha E_{Color}(I, S) + \beta E_{texture}(I, S) + \lambda E_{contour}(I, S)$
  //Update Label
  Evaluate E(I, S)
  Update the label X_{Label}

//Convergence Check

While ($\|E(I, S)_{New} - E(I, S)_{Old}\| \geq \varepsilon$ And Execution time < τ) do
  For each pixel (x in S):
  //Update label based on energy minimization
  X_{Label} = argmin_Label E (I, S) for all possible labels
  $E(I, S)_{Old} = E(I, S)_{New}$
  End
End

//Output

Display the final segmented image

IV. RESULT

In this phase, the algorithm begins by reading the source image S Fig. 2. Subsequently, the conversion of the image S to the HSV and Lab color spaces is performed, providing a suitable representation for color and brightness analysis. The contour map is calculated to capture significant variations in the image. Simultaneously, measures of homogeneity and correlation are computed for each pixel, laying the foundation for the initial label assignment Fig. 7.

Fig. 7. Original Image used in the experimental test.

The label propagation phase is initiated to establish initial relationships between neighboring pixels. For each pixel x in the source image S, a neighborhood analysis is conducted to examine adjacent pixels, namely P1 and P2. If both exhibit homogeneity ($P1_{indH} = P2_{indH} = 1$), their labels are adjusted to ensure coherence. This step aims to create an initial label assignment that considers the homogeneity characteristics within local pixel neighborhoods, setting the groundwork for subsequent energy minimization and label refinement Fig. 8.

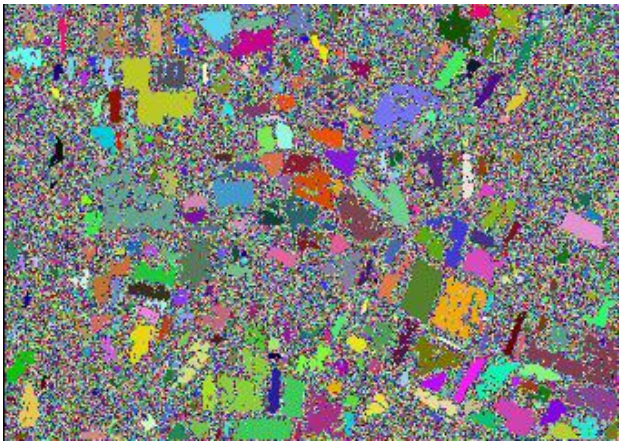

Fig. 8. Class and label assignment after neighborhood label update.

In this phase, the system's energy is minimized for each pixel x. The energy function E (I,S) is calculated by combining contributions from CIEDE2000 color difference, texture, and contours. This energy is used to update labels, promoting pixel coherence within the context of the entire image.

The convergence check loop is introduced to iterate through label updates until satisfactory convergence is achieved or the specified maximum execution time (τ) is exceeded. In each iteration, labels are updated using the ICM approach, where each pixel adjusts its label to minimize local energy. This step continues until the energy difference between consecutive iterations falls below a threshold ε, indicating satisfactory convergence Fig. 9.

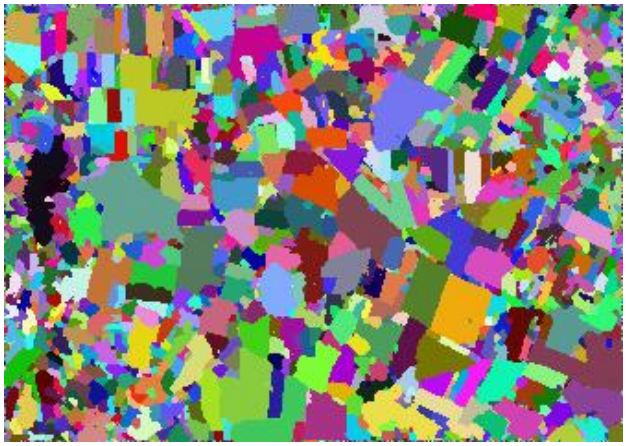

Fig. 9. Final labeling result.

Finally, the algorithm leads to the presentation of the final labeling result, displaying the segmented image. The optimized labels obtained (see Fig. 10) after the algorithm's convergence reflect the successful segmentation of the original image based on color, texture, and contour criteria.

Fig. 10. Segmented image.

V. Conclusion

The image segmentation approach based on Markov field methodology (MRF) and exploiting the attributes of color difference, texture and edge detection was subjected to an exhaustive evaluation. The results obtained demonstrate the robustness of the method in accurately delineating the contours of complex objects within images.

The CIEDE2000 color difference measurement was particularly effective at capturing subtle variations in color, ensuring accurate segmentation even under changing lighting conditions. The incorporation of texture information has enhanced the method's ability to discriminate between homogeneous but textured regions, improving segment consistency.

At the same time, the use of edge detectors, such as the Canny operator, has made it possible to highlight the boundaries between objects, improving the sharpness and overall accuracy of the segmentation.

To quantitatively evaluate the performance of the approach, commonly used metrics such as precision were utilized, recall and F-measure. The results demonstrated competitive performance.

With existing methods, highlighting the ability of the model to produce segmentations faithful to the real contours of objects in a variety of images.

In addition, in-depth visual analyses have been carried out, highlighting the ability of the method to handle complex cases such as the presence of fine structures, objects with blurred edges, and significant texture variations. These qualitative observations confirm the relevance of the approach in various applications, from computer vision to medical image analysis.

In conclusion, the results obtained support the validity and effectiveness of the Markov field-based image segmentation approach, demonstrating its potential for a variety of applications requiring accurate and robust segmentation. Ongoing improvements and future extensions to this methodology promise to further enhance its versatility and applicability in a variety of contexts.

REFERENCES

[1] M. Hossain et D. Chen, « Segmentation for Object-based Image Analysis (OBIA): A Review of Algorithms and Challenges from Remote Sensing Perspective », ISPRS Journal of Photogrammetry and Remote Sensing, vol. 150, p. 115-134, févr. 2019, doi: 10.1016/j.isprsjprs.2019.02.009.

[2] I. Kotaridis et M. Lazaridou, « Remote sensing image segmentation advances: A meta-analysis », ISPRS Journal of Photogrammetry and Remote Sensing, vol. 173, p. 309-322, mars 2021, doi: 10.1016/j.isprsjprs.2021.01.020.

[3] M. Sezgin et B. Sankur, « Survey over image thresholding techniques and quantitative performance evaluation », JEI, vol. 13, no 1, p. 146-165, janv. 2004, doi: 10.1117/1.1631315.

[4] S. Pare, A. Kumar, G. K. Singh, et V. Bajaj, « Image Segmentation Using Multilevel Thresholding: A Research Review », Iran J Sci Technol Trans Electr Eng, vol. 44, no 1, p. 1-29, mars 2020, doi: 10.1007/s40998-019-00251-1.

[5] E. S. Biratu, F. Schwenker, T. G. Debelee, S. R. Kebede, W. G. Negera, et H. T. Molla, « Enhanced Region Growing for Brain Tumor MR Image Segmentation », Journal of Imaging, vol. 7, no 2, Art. no 2, févr. 2021, doi: 10.3390/jimaging7020022.

[6] S. Bandyopadhyay, S. Das, et A. Datta, « Comparative Study and Development of Two Contour-Based Image Segmentation Techniques for Coronal Hole Detection in Solar Images », Sol Phys, vol. 295, no 8, p. 110, août 2020, doi: 10.1007/s11207-020-01674-4.

[7] R. Zhou et al., « Weakly Supervised Semantic Segmentation in Aerial Imagery via Explicit Pixel-Level Constraints », IEEE Transactions on Geoscience and Remote Sensing, vol. 60, p. 1-17, 2022, doi: 10.1109/TGRS.2022.3224477.

[8] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, et D. Terzopoulos, « Image Segmentation Using Deep Learning: A Survey », IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no 7, p. 3523-3542, juill. 2022, doi: 10.1109/TPAMI.2021.3059968.

[9] S. B et A. P, « Effect of Different Color Spaces on Deep Image Segmentation », in 2021 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), déc. 2021, p. 1-4. doi: 10.1109/WIECON-ECE54711.2021.9829655.

[10] Y. Li et al., « A Comprehensive Review of Markov Random Field and Conditional Random Field Approaches in Pathology Image Analysis », Arch Computat Methods Eng, vol. 29, no 1, p. 609-639, janv. 2022, doi: 10.1007/s11831-021-09591-w.

[11] J. Li, K. Feng, J. Yu, et H. Gu, « River extraction of color remote sensing image based on HSV and shape detection », in Seventh Symposium on Novel Photoelectronic Detection Technology and Applications, SPIE, mars 2021, p. 1594-1601. doi: 10.1117/12.2587284.

[12] S. Y. Chen, H. Tong, et C. Cattani, « Markov Models for Image Labeling », Mathematical Problems in Engineering, vol. 2012, p. e814356, août 2011, doi: 10.1155/2012/814356.

[13] Z. Kato, « Markov Random Fields in Image Segmentation », Foundations and Trends® in Signal Processing, vol. 5, no 1-2, Art. no 1-2, 2011, doi: 10.1561/2000000035.

[14] V. V. Mottl, A. B. Blinov, A. V. Kopylov, et A. A. Kostin, « Optimization Techniques on Pixel Neighborhood Graphs for Image Processing », in Graph Based Representations in Pattern Recognition, J.-M. Jolion et W. G. Kropatsch, Éd., in Computing Supplement. Vienna: Springer, 1998, p. 135-145. doi: 10.1007/978-3-7091-6487-7_14.

[15] H. Derin et H. Elliott, « Modeling and segmentation of noisy and textured images using gibbs random fields », IEEE Trans Pattern Anal Mach Intell, vol. 9, no 1, p. 39-55, janv. 1987, doi: 10.1109/tpami.1987.4767871.

[16] S. Geman et D. Geman, « Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images », IEEE Transactions on pattern analysis and machine intelligence, no 6, Art. no 6, 1984.

[17] P. Clifford et J. M. Hammersley, « Markov fields on finite graphs and lattices », 1971, Consulté le: 25 février 2024. [En ligne]. Disponible sur: https://ora.ox.ac.uk/objects/uuid:4ea849da-1511-4578-bb88-6a8d02f457a6.

[18] S. Dachian et B. Nahapetian, « On Gibbsianness of Random Fields ». arXiv, 12 septembre 2007. doi: 10.48550/arXiv.math/0609688.

[19] R. He, K. Xiao, M. Pointer, M. Melgosa, et Y. Bressler, « Optimizing Parametric Factors in CIELAB and CIEDE2000 Color-Difference Formulas for 3D-Printed Spherical Objects », Materials, vol. 15, no 12, Art. no 12, janv. 2022, doi: 10.3390/ma15124055.

[20] M. Gomez-Polo, M. Portillo, M. Luengo, P. Vicente, P. Galindo, et M. María, « A comparison of the CIELab and CIEDE2000 color difference formulas », The Journal of prosthetic dentistry, vol. 115, sept. 2015, doi: 10.1016/j.prosdent.2015.07.001.

[21] R. M. Haralick, K. Shanmugam, et I. Dinstein, « Textural Features for Image Classification », IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-3, no 6, p. 610-621, nov. 1973, doi: 10.1109/TSMC.1973.4309314.

[22] R. Sun et al., « Survey of Image Edge Detection », Frontiers in Signal Processing, vol. 2, 2022, Consulté le: 25 février 2024. [En ligne]. Disponible sur: https://www.frontiersin.org/articles/10.3389/frsip.2022.826967.

[23] F. Mokhtarian et F. Mohanna, « Performance evaluation of corner detectors using consistency and accuracy measures », Computer Vision and Image Understanding, vol. 102, no 1, p. 81-94, avr. 2006, doi: 10.1016/j.cviu.2005.11.001.

[24] X. Wang, « Laplacian Operator-Based Edge Detectors », IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no 5, p. 886-890, mai 2007, doi: 10.1109/TPAMI.2007.1027.

[25] L. Chandrasekar et G. Durga, « Implementation of Hough Transform for image processing applications », in 2014 International Conference on Communication and Signal Processing, avr. 2014, p. 843-847. doi: 10.1109/ICCSP.2014.6949962.

[26] W. McIlhagga, « The Canny Edge Detector Revisited », Int J Comput Vis, vol. 91, no 3, p. 251-261, févr. 2011, doi: 10.1007/s11263-010-0392-0.

[27] J. Sublime, Y. Bennani, et A. Cornuéjols, « A Compactness-based Iterated Conditional Modes Algorithm For Very High Resolution Satellite Images Segmentation », janv. 2015.