www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 17(3), 2026, 53-70. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2026.v17i3.1151

A new approach for microcalcification detection using hybrid filtering and DBSCAN segmentation

Ramon O. Guardado-Medina, Arturo García García, Javier Claustro Bobadilla, Luis Gerardo Rodarte Franco
Universidad de Guadalajara, Periférico Norte N° 799 Núcleo Universitario, C. Prol. Belenes, 45100 Zapopan, Jal.
ramon.guardado@cucea.udg.mx, arturo@ciicyt.org, claustro@academicos.udg.mx,
luis.rodarte0791@alumnos.udg.mx

Abstract. In the field of medicine, various methods are available for detecting abnormalities in breast cancer diagnostic procedures. Mammography is the most commonly used method for microcalcification (MC) detection in the early stages. Moreover, the geometric characteristics of MCs evolve through different stages. Several studies have addressed the detection of MCs. In the first step, characteristic enhancement is used to isolate MCs from the whole image. In the final steps, classical clustering techniques are applied. Different studies focus on enhancing characteristics in the initial stage, employing separation techniques to distinguish the present objects from the entire image, and concluding with classical grouping techniques.
Keywords: Breast cancer, microcalcification, Entropy, Morphology, Clustering, DBSCAN.

Article Info
Received November 22, 2025
Accepted January 3, 2026

1 Introduction

Breast cancer is a global health issue that has shown a dramatic increase in prevalence, accompanied by alarmingly high mortality rates. According to (INEGI, 2025), breast cancer is currently the third leading cause of death worldwide, posing a significant public health challenge. In Mexico, mortality rates from malignant breast tumors have exhibited a growing trend from 2019 to 2024, with 16 deaths per 100,000 women aged 20 years and older. This increase highlights the urgency of addressing the disease, particularly in regions with limited access to advanced healthcare technologies. Globally, breast cancer is often categorized as a disease predominantly affecting underdeveloped countries, where 69% of patients succumb to the disease. However, developed countries, commonly referred to as first-world nations, are also significantly impacted, as reported in (OMS, 2025) and (OPS, 2025).

- Research indicates that the number of women diagnosed with breast cancer in certain regions is projected to rise by 46% by 2030, underscoring the growing magnitude of the problem.
- Mortality rates due to breast cancer among women under the age of 65 are disproportionately higher in Latin America (57%) compared to North America (41%), emphasizing the regional disparities in healthcare access and early detection.
- Despite extensive research efforts, the root causes of breast cancer remain elusive, and conclusive findings are yet to emerge. This highlights the need for innovative approaches to combat the disease effectively.

Early detection of microcalcification (MC) clusters in breast tissue is critical for improving survival rates among patients. Radiologists primarily rely on mammography to diagnose breast cancer at an early stage. However, mammograms are inherently noisy images, making it challenging to accurately classify abnormal regions within the breast tissue. This diagnostic complexity can lead to missed detections or false positives, further complicating patient outcomes. Although there are multiple works presented to address this issue, most of them are based on object detection by algorithms of high computational complexity with classical techniques of image processing (Arnau, et al., 2012; Praful, et al., 2014). Some other proposals are based on local feature extractions with a bank of filters to obtain the local description of the MC morphology (Calderon, et al., 2016). Basically, the approach starts with initial training, in order to automatically learn and select the most salient features, Then, they are subsequently applied in a boosted classifier to detect individual MCs. Then the object detection method is extended to identify clusters. On the other hand, some other works use size-adapted segmentation of individual MCs in mammography (Nicolaos, et al., 2010; Pereira, et al., 2014), They are based on scale-space signature estimation which enables robust scale selection for initialization of MCs active contours. The limitations of current methodologies underscore the need for innovative, efficient, and robust techniques for MC detection. By addressing issues such as border effects, computational cost, and noise

sensitivity, a more reliable detection system can be developed. This new approach seeks to enhance image preprocessing to maximize the visibility of MCs while optimizing the performance of clustering algorithms like DBSCAN. Such advancements hold the potential to significantly improve early detection rates and, consequently, patient outcomes worldwide. In summary, breast cancer detection remains a critical challenge in global healthcare. This work aims to contribute to the field by presenting an approach that mitigates the drawbacks of existing methods, paving the way for more effective and accessible diagnostic tools.

In this study, we introduce a hybrid preprocessing pipeline designed to quantitatively improve the separation and detection of microcalcifications (MCs) in mammographic images. The proposed method enhances image contrast and object-to-background distinction by integrating entropy-based filtering with mathematical morphology, generating cleaner and more informative representations prior to clustering. This enhanced preprocessing enables DBSCAN to form more coherent and reliable clusters, reducing both underestimation and over-segmentation—two limitations commonly observed in classical segmentation methods. A quantitative analysis across ten MIAS cases (Table 1) demonstrates that the hybrid filter consistently yields the lowest Non-Uniformity (NU) values among all evaluated techniques. While traditional clustering methods such as k-means, FCM, and sub-segmentation produce NU values ranging from 0.0002 to 0.1902, and DBSCAN alone improves these outcomes with values between 0.0010 and 0.0156, the proposed hybrid DBSCAN-and-filter approach further reduces NU to a range of 0.0005 to 0.0100. This systematic improvement—often reducing NU by 50% or more relative to DBSCAN alone—confirms the robustness and stability of the hybrid preprocessing strategy. These findings support the principal contribution of this work: a contrast-enhanced, morphology-guided preprocessing framework that significantly strengthens DBSCAN's capability to detect homogeneous MC clusters with minimal segmentation noise. By achieving the lowest NU values across all cases, the proposed hybrid pipeline establishes a measurable performance advantage over classical and contemporary clustering techniques, positioning it as a valuable addition to the current literature on microcalcification detection

2. Methods

For the proposed method, the MIAS database (MIAS, 2015) is utilized as the primary source of mammographic data. This database comprises a total of 322 mammograms, each with a high resolution of 1024 × 1024 pixels, providing sufficient detail to analyze the structural characteristics of breast tissue. The MIAS database serves as a standardized dataset, ensuring consistency and reliability in the evaluation of the proposed method. Each image in the database contains detailed annotations regarding the abnormalities present in the breast tissue. These annotations include the precise spatial coordinates of the abnormalities, specified along the x and y axes, which are essential for accurately localizing regions of interest.

Additionally, the database categorizes the abnormalities in the images into distinct classifications, which provide critical information about their nature. For instance, the dataset specifies whether an abnormality is classified as malignant or benign, enabling a clear distinction between potentially life-threatening conditions and non-harmful anomalies. This classification is crucial for evaluating the performance of detection and diagnostic algorithms, as it allows for a more nuanced understanding of the algorithm's ability to differentiate between these two categories.

Furthermore, the MIAS database includes additional metadata that describes the type of abnormality, such as calcifications, circumscribed masses, spiculated masses, or architectural distortions. This level of detail supports comprehensive training and testing of the algorithm, as it facilitates the development of techniques tailored to various types of abnormalities. By leveraging this annotated dataset, the proposed method is better equipped to identify and classify abnormalities with high precision, ultimately contributing to improved diagnostic accuracy and early detection of breast cancer.

This is the main reason for proposing a filtering method based on two complementary approaches. The first one corresponds to a classical filtering approach, which is described here by mathematical morphology, and the other is related to the use of entropy, which compensates for the edge effects produced by the classic algorithms.

Thus provides, in the case of the preprocessing of images, a greater enhancement, and conservation of the objects of interest in the image, as is the case of MCs for given examples.

Algorithm 1. Shows the main steps of the proposed method. Furthermore, in the next sections the method is described in-depth.

Algorithm 1 Main steps in the proposed method

- *Get ROI's images of mammograms.*
- *Histogram adjustment, first filter to enhance some characteristics between object and background.*
- *Filter the image using erosion and dilatation to enhance contrast.*
- *Apply entropy to detect the maximum distortion between object and background.*
- *Segment the filtered image with DBSCAN cluster algorithms.*

2.1 Mathematical Morphology

Mathematical morphology technique is used to enhance the small objects, even those that begin to appear in the image. In other words, this technique helps to the next stages to have a better result. Then, the use of morphology over the objects if interest in this proposed method is discussed, followed by the mathematical summary.

Mathematical morphology (Gonzalez, et al., 2008) has had an important development in the applications related to the enhancement of characteristics, separating objects from the background of the same manipulating levels of gray. Mathematical morphology bases all its operations on two operators: erosion and dilatation associated with a structuring element.

The structuring element (ES) is a reference object of small size (some pixels) so that it has a simple geometric structure (point, line, square, circle, disk) associated with the predominant structure or topology in the elements present in the image. The erosion in the images in gray levels is obtained by moving the ES over the image and replacing the central pixel by the minimum of the digital levels covered by the ES. The dilation in the images in gray levels is obtained by moving the ES over the image and replacing the central pixel with the maximum of the digital levels covered by the ES. Aditionally, in this proposed work the erosion technique or opening effect is used. According to (Gonzalez, et al., 2008) Opening result between A by C is the erosion of A by C, followed by dilatation of the result by C. In other words: return N for a set A by structuring element C, denoted $A \circ C$, is defined as:

$$A \circ C = (A \ominus M) \oplus C \tag{1}$$

Or extended form:

$$[f \oplus c](x,y) = \{f(x - s, y - t) + c_N(s,t)\}, \tag{2}$$

where f denotes the domain of the structuring element b and used $(+s, +t)$ in the argument of the function like opening.

2.2 Entropy

After applying mathematical morphology, the entropy is used to find the maximum distortion between object and background. Further, the aim is to separate completely all objects from the background of the image. In other words, the obtained image has enhanced characteristics. As demonstrated by (Zheng, et al., 2015; Mohamed, et al., 2014) entropy is a concept that originally derived from the study of the physics of heat engines. It can be described as a measure of the amount of disorder, in a system in other words: at higher entropy, at higher disorder in the system. Another way to express entropy is to consider the uniform distribution of states in which a system can adapt. A low entropy system occupies a small number of such states, while a high entropy system occupies a large number of states, due to the previous conceptualization as a result we have Shannon's entropy. Shannon's entropy H of an image is defined in:

$$H = -\sum_{g=1}^{M} p_g log_2(p_g) \tag{3}$$

Where M is the number of gray levels and p is the probability associated with gray level g. In the case of an image, these states correspond to the levels of grays in which the individual pixels can be adapted. For example, in an 8-bit pixel, there are 256, $g = 0, 1, ..., M$ of such states. If all of the pixels have the same value, the entropy of the image is zero.

A set of pixels of the image is represented by (Pun, 980; Kapur, et al., 1985):

$$I = \{g \epsilon I / 0 \leq g \leq M\}, \tag{4}$$

and

$$F = \{g \epsilon I / g > t\}, B = \{g \epsilon I / g \leq t\}, \tag{5}$$

where F is the object or foreground in equation and B is a background in equation. In terms of foreground image processing, it is a set of pixels with gray levels above t value and background is the set of pixels with gray levels below t value. Moreover, the estimated probabilities of each pixel g with the quotient between ng and N, is obtained on equation (6)

$$p(g) = \frac{n_g}{N}, \quad g = 0, 1, \ldots, M, \tag{6}$$

where n_g is the number of times the pixel g is appears in the image and N is the total number of pixels. Then, the probabilities of the object and the background according to a threshold t are expressed as follows: P_f defined on the interval $0 \leq g \leq t$ and P_b defined on the interval $t + 1 \leq g \leq M$.

$$P_f(t) = P_f = \sum_{g=0}^{t} p(g) \tag{7}$$

$$P_b(t) = P_b = \sum_{g=t+1}^{M} p(g) \tag{8}$$

As can be seen, the cumulative probability is the sum of the object probability (7) and the background probability (8), and its function is defined as shown in equation (9).

$$P(t) = \sum_{g=0}^{M} p(g) \tag{9}$$

Based on equation (3), and the concept of Shannon's entropy, the entropy for each object and the background are shown in equation (10).

$$H_f(t) = -\sum_{g=0}^{t} p_f(g) log p_f(g) \tag{10}$$

$$H_b(t) = -\sum_{g=t+1}^{M} p_b(g) log p_b(g) \tag{11}$$

Therefore, the sum of Equations 10 and 11 give us the total entropy H of the objects and background of the image, Equation (12).

$$= -\sum_{g=0}^{t} p_f(g) log p_f(g) - \sum_{g=t+1}^{M} p_b(g) log p_b(g) \quad (12)$$
$$= H_f(t) + H_b(t)$$
$$= H(t)$$

The next stage is the enhancement of the image. First, the hybrid filtering of the mathematical morphology with the dilation technique ($A \circ B$) are used to make the objects more visible. Then, the entropy is applied. Finally, an image is obtained without noise and enhancement of the characteristics, separating the background and objects from the image. This process is applied to have a better input for the DBSCAN. Thus, to help DBSCAN technique to obtain better defined clusters.

The hybridization of mathematical morphology and entropy is theoretically justified by the complementary nature of both approaches in characterizing microcalcification structures. Mathematical morphology excels at enhancing geometrical and structural properties—such as shape, connectivity, and local contrast—by isolating fine bright elements that resemble MCs, while simultaneously suppressing diffuse background tissue. Entropy-based measures, in contrast, quantify the local uncertainty and complexity of intensity distributions, highlighting regions with abrupt variations or textural irregularities that typically correspond to microcalcifications. When combined, morphology provides a structural refinement that reduces noise and improves local contrast, and entropy contributes a statistical descriptor that reinforces the distinction between highly informative (high-entropy) MC pixels and homogeneous (low-entropy) background tissue. This dual enhancement has been reported as advantageous in prior work on medical image preprocessing, where geometric and statistical cues jointly improve the detectability of small lesions. In our experiments, this synergy is reflected in the reduction of NU values and in the sharper MC delineation achieved after hybrid filtering, confirming empirically that the integration of morphological operators with entropy-based enhancement provides a more robust foundation for subsequent clustering and density-based segmentation compared to traditional filters.

2.3 DBSCAN

Due to the characteristics of the objects – which are geometrically small and not similar to one another the DBSCAN technique is used to find small objects and to separate it from the background. Whereby, in this section the DBSCAN technique and its contribution to our work is presented. Since clustering algorithms work in the features space, this paper works with the gray level intensity of the pixels, they finish identifying prototypes near to the salient points in the histogram which represents the number of pixels according to gray levels which represents a characteristic in the space of characteristics (Guardado-Medina, et al., 2013).

On the other hand, clustering algorithms are unsupervised learning methods for classification (Bezdek, 1999; Jain, et al. 1998). They are based on the identification of groups, although no all information is available on the optimal number of groups. Hence, the goal of clustering algorithms is to partition a database, where data belonging to a group must be as similar or close as possible, whereas these must be as dissimilar as possible data belonging to the other groups. According to (Ester, et al., 1996), The goal is to find the corresponding pixel density of an object by separating it from the background of the image. in other words, the object is enhanced with the DBSCAN. In this algorithm, a set of points in some space to be clustered. To DBSCAN clustering, the points are classified as core points, (density) reachable points and outliers, as follows:

- A point p is a core point if at least minPts points are within distance ϵ (ϵ is the maximum radius of the neighborhood from p) of it (including p). Those points are said to be directly reachable from p.
- A point q is directly reachable from p if point q is within distance ϵ from point p and p must be a core point.
- A point q is reachable from p if there is a path p_1, ..., p_n with $p_1 = p$ and $p_n = q$, where each p_{i+1} is directly
- reachable from p_i (all the points on the path must be core points, with the possible exception of q).
- All points not reachable from any other point are outliers.

Now if p is a core point, then it forms a cluster together with all points (core or non-core) that are reachable from it. Each cluster contains at least one core point; non-core points can be part of a cluster, but they form its "edge", since they cannot be used to reach more points.

In the algorithm general version of DBSCAN, is given omitting details of data types and generation of additional information about clusters. Algorithms that work based on density result in clusters depends on the dense regions of objects in the data space that are separated by low-density regions. Therefore, this type of method is very useful for filtering noise and finding clusters. Furthermore, most partitioning methods perform the clustering process based on the distance between two objects. Thus, the techniques of clustering have

the problem of designing algorithms that can measure the quality of the partitioning of the clusters according to the similarity and difference between the data (Kohavi, R. 1995).

Algorithm 2 DBSCAN Algorithm

Input: P (Set of Points), $\mathfrak{D}$ (Distance function), ϵ (Radius), ϵP (Min points)
Output: Labels for points and number of clusters (C)
$C \leftarrow 0$
for $p \in$ P do
  if $label\,(p) \neq UNDEFINED$ **then**
  continue
  end if
  $N \leftarrow RangeQuery$(P,$\mathfrak{D}$, p, ϵ) ◁ Get neighbors of p
  if $|N| < \epsilon P$ **then**
    $label\,(p) \leftarrow NOISE$
    continue
  end if
  $C \leftarrow C + 1$
  $label\,(p) \leftarrow C$
  $S \leftarrow N \setminus \{p\}$ ◁ Seed set
  for $q \in S$ **do**
    if $label\,(q) = NOISE$ **then**
      $label\,(q) \leftarrow C$
    end if
    if $label\,(q) \neq UNDEFINED$ **then**
      continue
    end if
    $label\,(q) \leftarrow C$
    $N \leftarrow RangeQuery$(P,$\mathfrak{D}$, q, ϵ) ◁ Get neighbors of q
    if $|N| \geq \epsilon P$ **then**
      $S \leftarrow S \cup N$ ◁ Add new neighbors to seed set
    end if
  end for
end for
return $label$,C

Algorithm 3 Range Query Function

Input: P (Set of Points), $\mathfrak{D}$ (Distance function), Q (Query point), ϵ (Radius)
Output: N (List of neighbors)
$N \leftarrow \{\}$ ◁ Initialize empty list of neigh
for $p \in$ P **do**
  if $\mathfrak{D}(Q, p) \leq \epsilon$ **then**
    $N \leftarrow N \cup \{p\}$ ◁ Add point p to neighbors
  end if
end for
return N

2.4 Our approach

2.4.1 Overview

In this section, we present a novel approach for the detection and classification of microcalcifications (MCs) in medical imaging. The primary objective of this approach is to enhance the identification and characterization of MCs by leveraging preprocessing techniques to amplify their distinctive features. This step is essential for assisting the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm in accurately identifying MCs.

2.4.2 Preprocessing Pipeline

The preprocessing stage focuses on improving the quality of the input data to maximize the separability and visibility of features belonging to MCs. This is achieved through several steps:

2.4.2.1 Contrast Adjustment

- The first step in the preprocessing pipeline is the adjustment of the image to enhance contrast. This is accomplished by saturating 1% of the lowest and highest pixel values in the background and the object respectively. By doing so, the dynamic range of pixel intensities is effectively increased, making MCs more distinguishable from the surrounding background.
- This contrast enhancement step is crucial as it provides the foundation for subsequent processes by improving the visibility of MC regions in the output image.

2.4.2.2 Morphological Dilation

- The next step involves the application of a dilation operation using a specific kernel, as defined in Equation
- This step aims to expand the regions around MCs, effectively increasing the size of the areas where MCs are likely to be present.
- The dilation process creates a new image where the regions of interest are more pronounced, facilitating further analysis.

2.4.2.3 Entropy Filtering and Normalization

- Following dilation, an entropy-based process ($EntropyImage$) is applied to the resulting image, as described by Equation 12. This process evaluates the local complexity or randomness within the image, emphasizing the areas with higher information content, which often correspond to MCs.
- The resultant image undergoes normalization to standardize pixel intensity values, ensuring consistency across the dataset and improving the reliability of the subsequent clustering step.

$$\vec{x}_k = (x, y), \tag{13}$$

In this context, x and y represent the pixel coordinates within the image, where each pixel has a corresponding value, $p_{x,y}$, which in this case is a grayscale intensity level. The index k refers to the position in the vector representation that uniquely identifies the pixel located at (x,y). This transformation of the image into a vectorize format allows it to be interpreted as a "cloud of points" in a multidimensional space. The spatial and intensity information encoded in this representation is crucial for clustering algorithms such as DBSCAN, as it groups nearby points into clusters based on their proximity.

By treating the image as a cloud of points, DBSCAN can effectively identify regions of interest. Points that are spatially close and share similar grayscale values are grouped into the same cluster, which aligns well with the nature of microcalcifications (MCs) that tend to occur in localized patterns. Once DBSCAN is applied, the output includes clusters corresponding to potential MCs and other regions in the image. Algorithm 6 outlines the procedure for extracting these regions from the output. The size of each cluster serves as a critical indicator; small clusters are likely to correspond to MCs due to their compact nature, whereas larger clusters and unclassified pixels are typically associated with background elements. This distinction is vital for isolating MCs from the surrounding tissue.

2.4.3 Region Non-Uniformity (NU)

The NU measure (Zhang, et al., 1996) allows to set a reference about the quality of image segmentation, with values within range (0, 1). The measure is defined as:

$$NU = \frac{P * \sigma_P^2}{R * \sigma^2}; \qquad NU : 0 < NU < 1,$$

where P and R are the numbers of MCs and total number pixels in the segmented images, respectively. And, $\sigma P2$ and $\sigma 2$ are the variance of grey-scale values in the microcalcification space and the total variance in the image, respectively. In other words, the NU verifies if the object is formed by elements with common characteristics referring to their class, which allows establishing a relationship of similarity and uniformity among the MCs.

Algorithm 4 Proposed Algorithm

```
I: Osvaldo, Arturo, Andres, Humberto (I)
2: I ← Image
3: I_A ← Image Ad just (I, a, b, c, d, gamma)
4: K ← Ones Matrix                                              ◁ K
5: nI ← 1                                                       ◁ number of iterat
6: I_D ← Dilate (I_A, K, nI)
7: I_E ← Entropy Image ( I_D )                                  ◁ (Eq. 12)
8: I_T ← Normalize255(I_D + I_E)                                ◁ normalize new image
9: I_T ← 255 * I_T/MaxValue (I_T)                               ◁ normalize
10: P ← Image2VectorXY (I_T)   ◁ converting the image to a kind of vector ImageAd just representation
11: ϵ ← 5//
12: ϵ_S ← Minimum number of samples
13: D ← distance function
14: db ← DBSCAN(P, D, ϵ, ϵ_S)
15: maxE ← Maximum number of elements in the cluster to be considered
16: I_I ← GetInterestRegions (db, maxE)   ◁ getting the regions that could be a damaged tissue
17: return I_I
```

Algorithm 5 Image to a Vector function

```
1: Image2VectorXY (I)
2: I ← Gray-Scale Image to be transformed
3: nCols ← GetNumCols (I)                                   ◁getting image number of columns
4: nRows ← GetNumRows(I)                                    ◁getting image number of rows
5: X ← ZeroMatrix(nCols * nRows, 3).   ◁generating the space for the vector representation of the image
6: for i from 0 to nRows – 1 do
7:     for k from 0 to nCols – 1 do
8:         X[i * nRows + k] [0] ← i
9:         X[i * nRows + k] [1] ← k
10:        X[i * nRows + k] [2] ← I[i][k]
11:    end for
12: end for
13: return X
```

Algorithm 6 Get Interest Regions From DBSCAN result, function

```
1:  GetInterestRegions(db, maxE)
2:  nC ← db.nClasses                          ◁ number of the classes given by DBSCAN
3:  vCId ← db.vClassId                        ◁ id of the belonging class of each element given by the DBSCAN
4:  IV Indx ← {}                              ◁creating a empty list variable to keep the indic
5:  for from cl = 0 to nC do
6:      vIndxClass ← GET_VECTOR_INDEX(vCId == cl)   ◁getting the index of the same class elemen
7:      nElemInClass ← COUNT(vIndxClass).           ◁counting the elements of the same cl
8:      if cl < maxE then
9:          // it is a cluster of interest
10:         lVIndx ← lVIndex ∪ {vIndxClass}
11:     end if
12: end for
13: return lVIndx
```

3. Results

In this section is presented the results obtained by the proposed method. First, the original image (Figure 1), labeled "mdb170_roi", of the MIAS database is taken. This image contains an Architectural distortion that is defined as an alteration of the breast parenchyma without the coexistence of a mass, which may be due to malignant lesions. Presents a high probability to become cancer in the future.

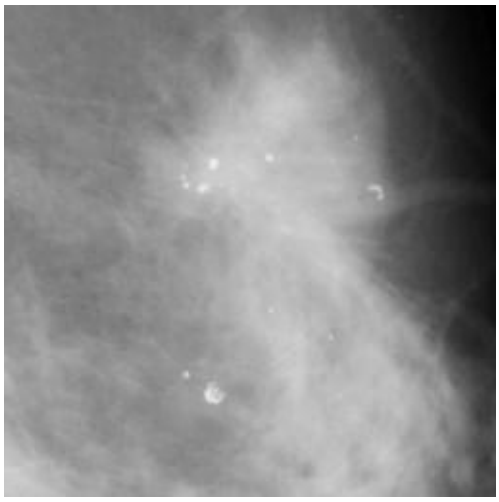

Figure 1: Original Image. From database MIAS, mdb170_roi

As it can be seen in (Figure 1), when an identification method is applied a most common problem are false positives due to the abnormality's gray levels. Furthermore, these can change during the process of their growth. Thus, they can be confused with the level of gray that exists in the tissue of the breast.

In many applications, the technique of mathematical morphology is a critical tool applied to images, particularly after the segmentation process. Its primary goal is to reconstruct features or characteristics that may have been lost during earlier preprocessing steps. This technique is especially useful for enhancing the structural integrity of segmented objects in an image, making it possible to refine their

shapes and boundaries. However, in grayscale images, the application of mathematical morphology presents unique challenges, such as the consideration of edge effects that can lead to unintended pixel loss or distortion.

To effectively separate the object from the background in grayscale images, the process begins with an adjustment of intensity values in the original image. This step enhances the contrast between the object and its surroundings, making it easier to apply morphological operations. Subsequently, morphological filtering is employed to refine the object's structure. Among the various morphological operations, dilation plays a central role in this methodology. Dilation enlarges the object by adding pixels to its boundaries, which is crucial for emphasizing regions of interest.

Morphological operations rely on structural elements, which are predefined shapes that determine how pixels are added or removed during the filtering process. In this study, several flat structural elements were evaluated, including Disk, Square, Rectangle, Octagon, and Diamond. Each structural element interacts differently with the image, leading to varying degrees of pixel preservation and enhancement.

Among the structural elements tested, the Disk demonstrated superior performance in preserving pixel integrity, particularly around the edges of objects. This is due to the continuous contact of the Disk element with the object's boundary during the dilation process, minimizing the edge effects that often lead to pixel loss. Unlike other structural elements, which may produce uneven results or exacerbate edge-related distortions, the Disk consistently maintains the object's shape and size.

Further experimentation revealed that the Disk element with a parameter size of 15 yielded the best results. This configuration effectively expanded the object while maintaining its structural integrity. Figure 2 illustrates the effect of using the Disk element in the dilation process, highlighting the morphological enhancements achieved.

Despite its advantages, the morphological filtering process introduces a notable attenuation effect on pixel intensities, particularly at the edges of the object. This phenomenon occurs because the dilation process involves interpolating and smoothing pixel values, which can reduce intensity contrast at the object's boundaries. While this attenuation may slightly alter the visual appearance of the object's edges, it does not compromise the overall structural accuracy. Instead, it helps to smooth irregularities and noise, further enhancing the object's delineation from the background.

The use of the Disk structural element in morphological filtering offers a robust solution for edge-preserving dilation in grayscale images. Its ability to maintain continuous contact with the object's boundary ensures minimal pixel loss, making it an ideal choice for applications that require high precision. The parameter tuning (Disk size of 15) allows for tailored adjustments to suit specific imaging requirements, ensuring that the dilation process effectively enhances the object without introducing significant distortions or artifacts.

The mathematical morphology technique, particularly when using the Disk structural element, proves to be an invaluable tool for refining the representation of objects in grayscale images. By addressing edge effects and enhancing object boundaries, this approach significantly improves the quality and reliability of image analysis, paving the way for more accurate segmentation and detection processes in applications such as medical imaging.

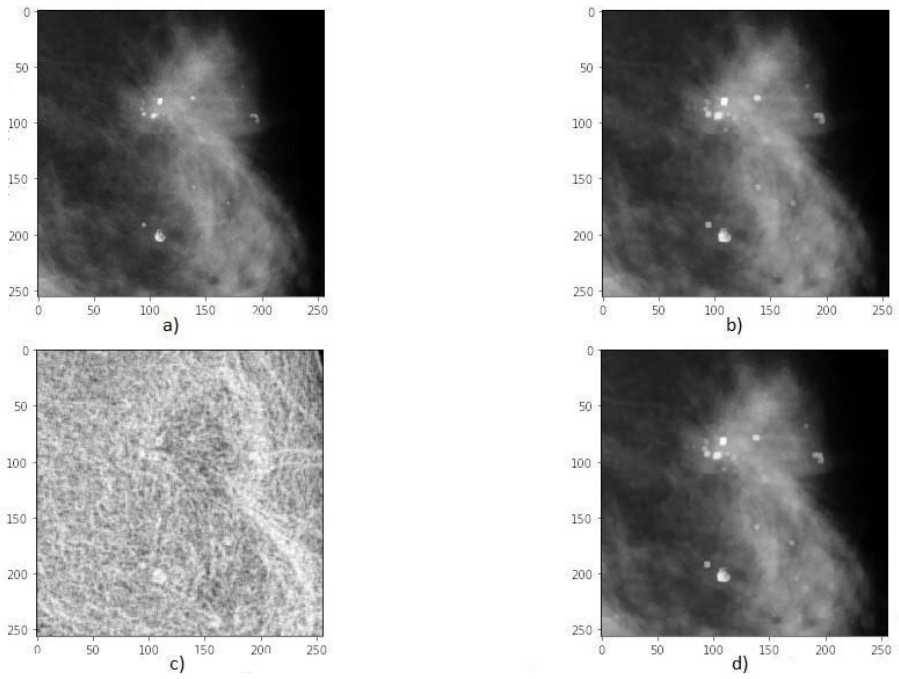

Figure 2: Preprocessing. a) Histogram b) Adjustment, c) Morphological and entropy filter, d) Hybrid filtering.

The edge effects introduced by certain filtering techniques can cause a partial or even total loss of the object's edges, leading to false negatives in detection processes. This challenge necessitates a method to present the object without noise or errors resulting from these edge effects. To address this issue, a visualization improvement technique is applied, leveraging the concept of entropy in two dimensions. This approach is based on Shannon's Information Theory, where entropy is interpreted as a measure of uncertainty.

In the context of this study, the entropy technique is particularly effective in analyzing the transition zone between the object and the image background. Entropy relies on probability distributions, and when all possible outcomes are equally probable, the entropy reaches its maximum, indicating a high degree of unpredictability. For the purposes of this work, entropy is utilized to capture and highlight the subtle transitions within the object. This process is applied individually to the regions of interest, as illustrated in (Figure 2), enhancing the visualization of the object's edges and ensuring better differentiation from the background.

One-dimensional entropy measures the difference between two probability distributions. Lower entropy implies greater similarity between the two sources. However, this approach does not consider the spatial dependencies of pixel values in the image, which limits its effectiveness for complex image analysis tasks. To overcome this limitation, this work introduces the use of second-order entropy, which takes spatial relationships into account. By incorporating spatial dependencies, second-order entropy provides a more robust method for distinguishing between the object and the background, leading to more accurate segmentation results.

The final stage of preprocessing involves the collaborative use of morphological filtering and entropy-based filtering. This hybrid approach integrates the strengths of both methods to enhance the object's features while minimizing noise and artifacts. Specifically:

Morphology and Entropy Collaboration: This step involves averaging the results of morphological and entropy filtering to create a third image that combines the benefits of both techniques. This hybrid image captures both structural integrity and intensity-based transitions, improving the overall clarity of the object.

Frequency and Entropy Collaboration: Similarly, the results of frequency-based filtering and entropy are averaged to produce another hybrid image. This method ensures that the object's features are preserved across both intensity and frequency domains.

These hybrid filtering techniques result in an output image that is free of noise and exhibits enhanced object characteristics relative to the background. This process is visually demonstrated in (Figure 2), showcasing the effectiveness of the hybrid filtering approach.

Once the preprocessing is complete, the resulting clean image serves as the input for the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. As described earlier, DBSCAN identifies regions of high density and expands clusters based on these dense areas. This clustering method is particularly well-suited for data that contains clusters of similar density, as is the case with microcalcifications (MCs) in mammographic images.

The preprocessing steps significantly improve the separability of the object from the background, enabling DBSCAN to more effectively highlight MC clusters. By applying the characteristics of DBSCAN to the preprocessed images, the method achieves a complete separation of the object from the image background, ensuring that the MCs are distinctly identified and highlighted. This comprehensive approach not only enhances the accuracy of detection but also ensures that the MCs are prominently featured, facilitating better diagnostic outcomes.

Table 1. No-Uniformity with Hybrid filter.

	k-means [12]	FCM [3]	Sub Segmentation [20]	DBSCAN [5]	DBSCAN and Filter
mdb058	0.0432	0.0326	0.0022	0.0010	0.0005
mdb170	0.0208	0.0193	0.0193	0.0150	0.0075
mdb188	0.0286	0.0286	0.0029	0.0025	0.0015
mdb204	0.0165	0.0134	0.0134	0.0100	0.0050
mdb209	0.0123	0.0181	0.0002	0.0023	0.0008
mdb219	0.1092	0.1389	0.0335	0.0156	0.0050
mdb223	0.0106	0.1902	0.0540	0.0321	0.0100
mdb227	0.0387	0.0276	0.0003	0.0060	0.0020
mdb248	0.0289	0.0434	0.0058	0.0058	0.0020
mdb249	0.0569	0.0569	0.0044	0.0042	0.0015

In (Figure 3), the Estimated Number of Clusters panel illustrates the clusters generated by applying the proposed method to the original image (mdb170_roi), resulting in a total of 32 clusters. These clusters were obtained by setting the parameters $eps = 5$ (defining the maximum radius of the neighborhood), $minPts = 7$ (minimum number of points to form a cluster), and using the Euclidean distance as the metric. These parameter values were chosen to optimize the clustering process for this specific dataset, balancing the sensitivity to detect microcalcifications (MCs) with the robustness to avoid over-segmentation. The same parameter configuration was consistently applied across all other images analyzed to ensure uniformity in results and to validate the method's adaptability to varying conditions.

The Image with Hybrid Filter panel in (Figure 3) showcases the outcome of combining entropy-based filtering and mathematical morphology. This hybrid filtering step is critical for preparing the image for clustering, as it removes noise and enhances the features of interest, particularly in the transition zones between the object (MCs) and the background. By integrating entropy, which measures uncertainty, and morphological techniques, which preserve structural integrity, the hybrid filter produces a clean and enhanced image. This step ensures that the boundaries of the MCs are well-defined and that the contrast between clusters and the background is maximized, making it easier for clustering algorithms to identify and group relevant regions.

The Segmented Image panel in (Figure 3) presents the result of applying the DBSCAN clustering technique to the hybrid-filtered image. DBSCAN's ability to identify clusters based on density is well-suited for this application, as it efficiently separates regions of interest from noise. The clustering process successfully isolates the MCs, with each cluster representing a distinct group of high-density points that share similar characteristics. The output not only highlights the MCs but also minimizes false positives by treating areas of low density or unclassified pixels as background.

The proposed method demonstrates a high degree of robustness in handling noise and enhancing features, as evidenced by the consistent results across all analyzed images. The combination of hybrid filtering and DBSCAN clustering ensures that the method can adapt to

various imaging conditions while maintaining accuracy. By producing clean, noise-free images with well-defined features, the preprocessing steps enable DBSCAN to perform optimally, resulting in precise cluster identification.

The selection of *eps* = 5 and *minPts* = 7, along with the use of the Euclidean distance metric, plays a crucial role in achieving accurate clustering. These parameters determine the size and density of the clusters, ensuring that small, compact regions such as MCs are captured effectively. The parameter configuration also strikes a balance between detecting individual MCs and avoiding over-segmentation, which could lead to false positives.

As shown in (Figure 3), the DBSCAN clustering technique effectively separates the MCs from the surrounding background, providing a clear and detailed segmentation of the regions of interest. This comprehensive separation is a critical step in the analysis, as it lays the foundation for further diagnostic processes. By isolating the MCs with high precision, the method facilitates accurate assessment and classification, which are essential for early breast cancer detection.

In conclusion, (Figure 3) encapsulates the success of the proposed method in enhancing image quality and enabling effective clustering. The integration of entropy and morphological filtering, followed by DBSCAN, provides a powerful approach to identifying and segmenting MCs with high accuracy. This methodology not only demonstrates robustness across different images but also highlights the potential for application in real-world diagnostic scenarios.

The NU values reported for the analyzed images help explain why certain cases exhibit stronger improvements after applying the hybrid filtering and DBSCAN approach. Images such as mdb058, mdb188, mdb209, mdb227 and mdb249 show substantial NU reduction often below 0.005 indicating that the hybrid filter successfully enhances local contrast and structural continuity around microcalcifications, enabling DBSCAN (with *eps = 5 and minPts = 7*) to form compact, well-defined clusters without fragmentation. These parameter settings are particularly effective when MCs present small but locally dense structures, allowing DBSCAN to isolate them while suppressing noise. In contrast, more complex cases such as mdb219 and mdb223 retain higher NU values across all traditional methods, reflecting intrinsic tissue heterogeneity; although the hybrid-filtered DBSCAN still achieves notable improvements, the fixed ε and minPts must contend with irregular intensity distributions that limit further gains. A broader comparative discussion against k-Means and FCM further highlights these differences: while both conventional algorithms struggle with sensitivity to intensity variation and tend to produce fragmented or overly smooth cluster boundaries, the hybrid-DBSCAN pipeline consistently yields lower NU values and more stable segmentation outcomes. This contextualization underscores the strengths of the proposed method while clarifying its limitations in highly heterogeneous mammograms.

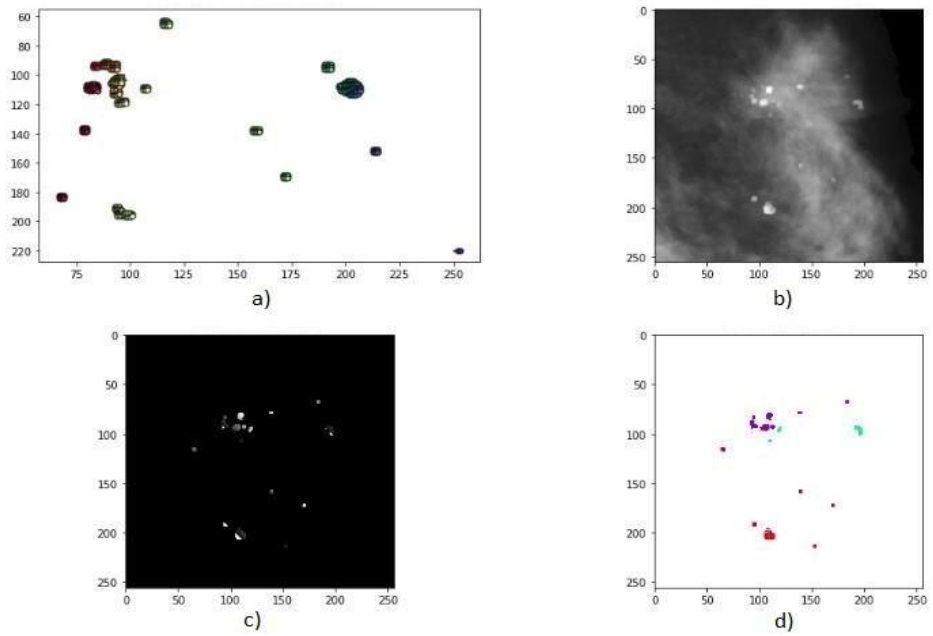

Figure 3: Clustering. a) Estimated number of clusters, b) Hybrid filter, c) Segmented image d) Macrocalcification.

Thus, the proposed method effectively separates and detects MCs, enabling their clear identification within mammographic images. Once the MCs are isolated, each cluster is labeled and visually distinguished from the image background, ensuring a clear representation of the regions of interest. The proposed technique, rooted in the DBSCAN clustering method, is evaluated against other well-established clustering algorithms, including kMeans, Fuzzy C-Means (FCM), and Possibilistic Fuzzy C-Means (PFCM), through a comparative sub-segmentation analysis. Importantly, all these methods operate on a preprocessed image generated using the hybrid filtering approach that combines mathematical morphology and entropy. This preprocessing step significantly enhances the quality of the input data, providing a noise-free and feature-enhanced image that forms the foundation for clustering.

In this approach, only the gray-level feature from the filtered image I is used as the basis for clustering. This means the feature space corresponds to the histogram of the image, where pixel intensity values are represented as a distribution. The expected classes or regions of interest within the mammograms are limited to three distinct categories: the background, the tissue, and the MCs. This simplification of the clustering task ensures a focused and efficient analysis of the regions critical for breast cancer diagnosis.

Partitional clustering algorithms such as k-Means and FCM work by defining prototypes or cluster centers that adapt to the areas with the highest pixel density. These prototypes tend to converge near the peaks of the histogram, where pixel intensities are most concentrated. This behavior makes them suitable for segmenting larger regions, such as the background or tissue, which occupy a significant portion of the image and form prominent peaks in the histogram. However, MCs, being small and localized objects, are represented by a much smaller number of pixels. This sparsity presents a challenge for partitional algorithms, as their prototypes may not adequately capture the small, dispersed clusters corresponding to MCs.

To address this, the number of clusters must be significantly increased to achieve a more stable grouping of MCs. By increasing the number of clusters, the algorithms can better differentiate the subtle intensity variations that define the MC regions. For this study, a value of 5 clusters is used for both k-Means and FCM methods, as suggested by the literature (Ojeda-Magaña, et al., 2009), to optimize their performance for detecting MCs. This configuration ensures a finer granularity in segmentation, allowing the small but clinically significant MC regions to be more effectively isolated.

The use of DBSCAN, in contrast to partitional algorithms, offers distinct advantages for this task. DBSCAN is a density-based clustering method that identifies clusters based on the density of data points, making it particularly suitable for identifying sparse but high-density regions such as MCs. Unlike k-Means and FCM, which rely on the number of clusters being predefined, DBSCAN automatically determines the number of clusters based on the density distribution of the data. This characteristic enables DBSCAN to adapt dynamically to the varying distributions of MCs in different images, providing a more flexible and robust solution.

Furthermore, the hybrid filtering applied before clustering plays a crucial role in enhancing the effectiveness of all these methods. By reducing noise and emphasizing the boundaries and features of MCs, the preprocessing step ensures that the input data is optimized for clustering. This preparation not only improves the performance of DBSCAN but also enhances the results of k-Means, FCM, and PFCM, as the preprocessed image provides a clearer and more consistent representation of the regions of interest.

In summary, the performance of DBSCAN in detecting microcalcifications is closely tied to the selection of its parameter's *eps* and minPts, and different configurations reveal how sensitive the algorithm is to local density variations within mammographic tissue. In preliminary exploratory tests, smaller neighborhood radii (e.g., *eps = 3–4*) increased the algorithm's selectivity, successfully isolating very compact MC structures but also producing fragmented clusters or misclassifying faint calcifications as noise. Conversely, larger radii (e.g., *eps = 6–8*) tended to merge adjacent bright regions or incorporate background tissue, reducing segmentation precision in images with heterogeneous texture. A similar pattern was observed with *minPts*: lower values (e.g., *minPts = 4–5*) favored the detection of isolated or faint MCs but occasionally introduced false positives in high-contrast backgrounds, while higher values (e.g., *minPts = 8–10*) improved robustness by forming only well-supported clusters, at the cost of missing subtle calcifications. The adopted configuration (eps = 5, minPts = 7) represents a balance identified through these observations, offering stable performance across different cases. These reflections highlight that the robustness of the proposed method arises from the interplay between parameter settings and the structural properties of each image, emphasizing the importance of tuning *eps and minPts* when dealing with varying image complexities and tissue heterogeneity.

3.1 Discussion

The objective of this study is to perform a detailed comparative analysis of the segmentation results of mammographic images using three different clustering algorithms, aiming to assess their effectiveness in isolating microcalcifications (MCs). A key aspect of this comparison is the establishment of a quantitative metric that directly correlates with the quality of segmentation results. For this purpose, the Region Non-Uniformity (NU) measure has been proposed, which quantifies the homogeneity and size of the overall image and, more importantly, the segmented regions, particularly those containing MCs. The NU metric takes values within the interval $(0, 1)$, where lower values indicate higher uniformity and better segmentation quality.

Quantitative Analysis Using NU

Table 1 summarizes the NU values obtained for each segmented mammogram and for all evaluated algorithms, providing a detailed view of the homogeneity and compactness of the microcalcification (MC) regions. In addition to presenting the clustering results of k-Means, Fuzzy C-Means (FCM), and Sub-Segmentation, the table also includes the performance of DBSCAN and DBSCAN combined with filtering. This comprehensive comparison allows a direct quantitative assessment of segmentation quality across methods.

Although the NU metric was introduced as a measure of regional homogeneity, a deeper examination of its sensitivity is essential to fully interpret the values reported in Table 1. NU varies consistently with the degree of separation achieved between MCs and surrounding mammary tissue, making it not only a numerical descriptor but also a direct indicator of segmentation quality. The metric strongly penalizes both over-segmentation (fragmented MC regions) and under-segmentation (mixing with background tissue), causing NU to increase when cluster boundaries are imprecise and to approach zero when MCs are well isolated and compact.

In practical terms, NU values near 0.00 denote excellent segmentation, with clean and uniform MC regions; values between 0.01 and 0.05 reflect moderate but acceptable segmentation; and values above 0.05 indicate poor separation, often associated with low contrast or algorithmic limitations. The NU values observed in Table 1 clearly exhibit this behavior. Sub-Segmentation consistently produces NU values within the lowest range, confirming its ability to isolate MCs accurately. Conversely, k-Means and FCM yield higher NU values in several images, revealing their sensitivity to noise and intensity variation. DBSCAN shows greater stability, and its performance is further enhanced when preprocessing filters are incorporated, resulting in the lowest NU values across all methods.

Overall, the variability of NU across the different algorithms provides a practical and interpretable assessment of segmentation quality. By observing how the NU metric decreases or increases for a given image depending on the method used, it becomes evident which

algorithms generate well-defined MC boundaries and which one's struggle with homogeneity or density inconsistencies. This makes NU a robust and discriminative quantitative tool for comparative evaluation in mammographic image segmentation.

Performance of Clustering Algorithms

The k-Means algorithm, while widely used due to its simplicity and computational efficiency for many tasks, encounters significant challenges in this context. Achieving optimal clustering with k-Means requires multiple executions, each involving a different number of clusters, and numerous iterations to ensure convergence. This iterative process results in a high computational cost, making it less practical for tasks requiring rapid analysis. Additionally, k-Means is prone to falling into local minima, which can lead to incorrect classifications. These limitations increase the likelihood of False Positive (FP) or False Negative (FN) outcomes during the clustering process, reducing its reliability for segmenting MCs. Furthermore, the NU values for k-Means tend to be relatively high, reflecting poor segmentation quality due to the algorithm's inability to effectively differentiate the fine details of MC regions.

The FCM algorithm offers an improvement over k-Means by incorporating a degree of fuzziness, allowing pixels to belong to multiple clusters with varying degrees of membership. This flexibility reduces the number of iterations required for convergence, resulting in lower computational costs compared to k-Means. However, while FCM demonstrates better performance, its NU values remain high, indicating suboptimal segmentation quality. To achieve a significant reduction in NU, adjustments to the algorithm's parameters, such as increasing the number of segmented regions or fine-tuning the typicality threshold, are necessary. These adjustments, while effective, require careful calibration and may still fail to address the inherent challenges of segmenting MCs with high precision.

Sub-Segmentation

Among the three methods, Sub-Segmentation demonstrates the most promising results. This approach effectively separates MCs from the surrounding mammary tissue, achieving lower NU values compared to k-Means and FCM. The reduced NU values reflect improved segmentation quality, with the MC regions being more accurately isolated from the background and tissue. Sub-Segmentation benefits from its ability to focus on smaller, localized regions, which aligns well with the characteristics of MCs. By addressing the limitations of the other methods, Sub-Segmentation provides a robust solution for analyzing mammographic images and identifying regions of interest with high precision.

Implications for Mammogram Segmentation

Comparative analysis highlights the importance of selecting an appropriate clustering algorithm for the segmentation of mammographic images. The high NU values observed for k-Means and FCM emphasize the need for advanced techniques capable of addressing the unique challenges posed by MCs, such as their small size and sparse distribution. Sub-Segmentation, with its ability to achieve lower NU values, emerges as a more reliable method, particularly for applications requiring accurate identification of MC regions. the results of this study underscore the critical role of quantitative metrics like NU in evaluating segmentation quality. While k-Means and FCM offer certain advantages, their limitations in this specific application render them less effective compared to Sub-Segmentation. The findings pave the way for further exploration of hybrid approaches and parameter optimization to enhance the performance of segmentation algorithms in medical imaging.

4 Conclusions

As it can be observed, the segmentation quality can be significantly enhanced by incorporating filtering techniques during the preprocessing stage of the image analysis. These filters play a crucial role in improving the contrast of the image, creating a clearer distinction between objects of interest, such as microcalcifications (MCs), and the background. This enhanced separation is pivotal for effective segmentation, as it allows clustering algorithms to more accurately identify and delineate regions of interest. As emphasized, the identification of MCs becomes a less challenging task when there is a pronounced contrast between the objects and the background, and the proposed algorithm demonstrates highly satisfactory results in such scenarios.

On the other hand, if the filtering process results in an image with high homogeneity, the clustering behavior of algorithms like k-means, Fuzzy C-Means (FCM), and sub-segmentation changes notably. Specifically, the number of regions identified by these methods tends to increase significantly, potentially complicating the segmentation process. This increase in regions requires careful observation of changes within the MC regions, especially as the NU (Non-Uniformity) value decreases. NU value serves as an indicator of the segmentation quality, and monitoring it provides valuable insights into the performance of the algorithms.

For the DBSCAN algorithm, which operates based on density, the filter's impact on MC identification is similarly crucial. Proper threshold adjustment in DBSCAN plays a vital role in fine-tuning the algorithm, enabling it to balance between underestimation and overestimation of MC regions. By setting an appropriate threshold, the algorithm can minimize errors, ensuring that the MCs are neither overlooked nor exaggerated in size. This capability underscores the importance of parameter optimization in achieving reliable segmentation outcomes.

Additionally, the inherently small proportion of pixels belonging to MCs compared to the total pixels in the image significantly influences the NU value. When segmentation is performed correctly, the NU value approaches zero, indicating a high level of homogeneity and accurate object identification. This low NU value is a direct measure of the algorithm's success in isolating MC regions with minimal overlap or misclassification. Across all cases, the best object identification is observed when the NU value is near zero, signifying precise segmentation and effective discrimination of MCs from the surrounding tissue and background.

In conclusion, the integration of preprocessing filters and the careful adjustment of algorithmic parameters collectively enhance the segmentation process. By improving image contrast and enabling more accurate clustering, these methods ensure that MCs are effectively identified and segmented. The low NU values achieved reflect the robustness of the proposed approach and its ability to deliver high-quality results, particularly in the challenging task of microcalcification detection in mammographic images.

Acknowledgments

To Robert, for the bagels and explaining CMYK and color spaces.

References

Agrawal, P., Vatsa, M., & Singh, R. (2014). Saliency based mass detection from screening mammograms. *Signal Processing, 99,* 29–47. https://www.sciencedirect.com/science/article/abs/pii/S0165168413004982

Arikidis, N. S., Karahaliou, A., Skiadopoulos, S. G., Korfiatis, P., Likaki, E. A., Panayiotakis, G., & Costaridou, L. (2010). Size-adapted microcalcification segmentation in mammography utilizing scale-space signatures. *Computerized Medical Imaging and Graphics, 34(6),* 487–493. https://doi.org/10.1016/j.compmedimag.2009.12.009

Arnau Oliver, A., Torrent, A., Lladó, X., Tortajada, M., Tortajada, L., Sentís, M., Freixenet, J., & Zwiggelaar, R. (2012). Automatic microcalcification and cluster detection for digital and digitised mammograms. *Knowledge-Based Systems, 28,* 68–75. https://doi.org/10.1016/j.knosys.2011.11.021

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms.* Plenum Press. https://doi.org/10.1007/978-1-4757-0450-1

Bezdek, J. C. (1999). *Fuzzy models and algorithms for pattern recognition and image processing.* Kluwer Academic Publishers.

Calderón, J. M., & Febles, G. (2016). Actualidad para el diagnóstico del carcinoma ductal in situ: correlación radiopatológica. *Revista Chilena de Radiología, 17,* 57–65. https://www.redalyc.org/pdf/3716/371650379010.pdf

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining* (pp. 226–231). AAAI Press. https://file.biolab.si/papers/1996-DBSCAN-KDD.pdf

Gonzalez, R. C., & Woods, R. E. (2008). Morphological image processing. In *Digital Image Processing* (3rd ed., pp. 633–639). Pearson Education.

Guardado Medina, R. O., Ojeda Magaña, B., Quintanilla Domínguez, J., Ruelas, R., & Andina, D. (2013). Quality of microcalcification segmentation in mammograms. In *Proceedings of SOCO'13-CISIS'13-ICEUTE'13* (pp. 208–229). Salamanca, Spain.

INEGI. (2025). Statistics, world day against cancer. https://www.beta.inegi.org.mx

Jain, A. K., Murty, M. N., & Flynn, P. J. (1998). *Algorithms for clustering data.* Prentice Hall.

Kapur, J. N., Sahoo, P. K., & Wong, A. K. C. (1985). A new method for grey-level picture thresholding using the entropy of histogram. *Computer Vision, Graphics, and Image Processing, 29(3),* 273–285. https://doi.org/10.1016/0734-189X(85)90039-9

Kohavi, R. (1995). *Wrappers for performance enhancement and obvious decision graphs* (Doctoral dissertation, Stanford University).

MacQueen, J. (1976). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297). University of California Press.

Mammographic Image Analysis Society (MIAS). (2015). MIAS digital mammogram database. https://peipa.essex.ac.uk/info/mias.html

Mohamed, H., Mabrouk, M. S., & Sharawy, A. (2014). Computer aided detection system for microcalcification in digital mammograms. *Computer Methods and Programs in Biomedicine, 113(2),* 226–235. https://doi.org/10.1016/j.cmpb.2013.11.003

Oliveira, F. S., Carvalho Filho, A. O., Silva, A. C., Paiva, A. C., & Gattass, M. (2015). Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM. *Computers in Biology and Medicine, 60,* 42–53. https://doi.org/10.1016/j.compbiomed.2015.02.002

Ojeda Magaña, B., Quintanilla Domínguez, J., Ruelas, R., & Andina, D. (2009). Marges sub-segmentation with the PFCM clustering algorithm. In *Proceedings of the 7th IEEE International Conference on Industrial Informatics* (pp. 499–503). IEEE.

OMS. (2025). Cáncer de mama: prevención y control. Organización Mundial de la Salud. https://www.who.int/topics/cancer/breastcancer/es/

OPS. (2025). Organización Panamericana de la Salud (OPS). https://www.paho.org

Pal, N. R., Pal, K., Keller, J. M., & Bezdek, J. C. (2005). A possibilistic fuzzy c-means clustering algorithm. *IEEE Transactions on Fuzzy Systems, 13(4),* 317–370. https://doi.org/10.1109/TFUZZ.2005.851354

Patricia, L. (2016). Breast calcifications: Description and classification according to BI-RADS. *Revista Chilena de Radiología, 22,* 80–91.

Pereira, D. C., Ramos, R. P., & Nascimento, M. Z. (2014). Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. *Computer Methods and Programs in Biomedicine, 114(1),* 88–101. https://doi.org/10.1016/j.cmpb.2013.11.021

Pun, T. (1980). A new method for gray-level picture threshold using the entropy histogram. *Signal Processing, 2(3),* 223–237. https://doi.org/10.1016/0165-1684(80)90050-8

Zhang, Y., & Gerbrands, J. J. (1996). A survey on evaluation methods for image segmentation. *Pattern Recognition, 29(8),* 1335–1346. https://doi.org/10.1016/0031-3203(95)00150-5

Zheng, Q., Lu, Z., Zhang, M., Xu, L., Ma, H., Song, S., … He, T. (2015). Automatic segmentation of myocardium from black-blood MR images using entropy and local neighborhood information. *Journal of Magnetic Resonance Imaging, 41(3),* 706–713. https://doi.org/10.1002/jmri.24521

A Abbreviations

MC: Microcalcification

DBSCAN: Density-Based Spatial Clustering of Applications with Noise

FCM: Fuzzy C-Means

A.1 Availabillity of data and Materials

The image data set used in this work is available in the mini-MIAS mammograms database (http://peipa.essex.ac.uk/info/mias.html)