**A Review on Pre-processing Techniques for Digital Mammography Images**

**ABSTRACT**

Mammograms are the **soft X-rays kind of imaging technique used for the detection of any lesions or cysts in breasts.** Digital mammograms have many kinds of artifacts that affect the accuracy of the detection of tumor tissues in the automated Computer Aided Detection (CAD) system for mammograms. Preprocessing helps to remove such artifacts is an important step. Image preprocessing is used to maintain image efficiency in mammogram images there are many artifacts need to be removed like labels, patient name, muscle part, etc. and enhance the region of interest which helps for efficient segmentation and detection of tumor. **The basic objective of this study is to evaluate and discuss different techniques and approaches proposed in order to enhance the breast cancer images and an efficient preprocessing technique for mammography. It aims to find the existing preprocessing techniques for mammography images and discuss the techniques used and their advantages.**

**PREPROCESSING**



**Types of artifacts and noises observed in mammogram image**

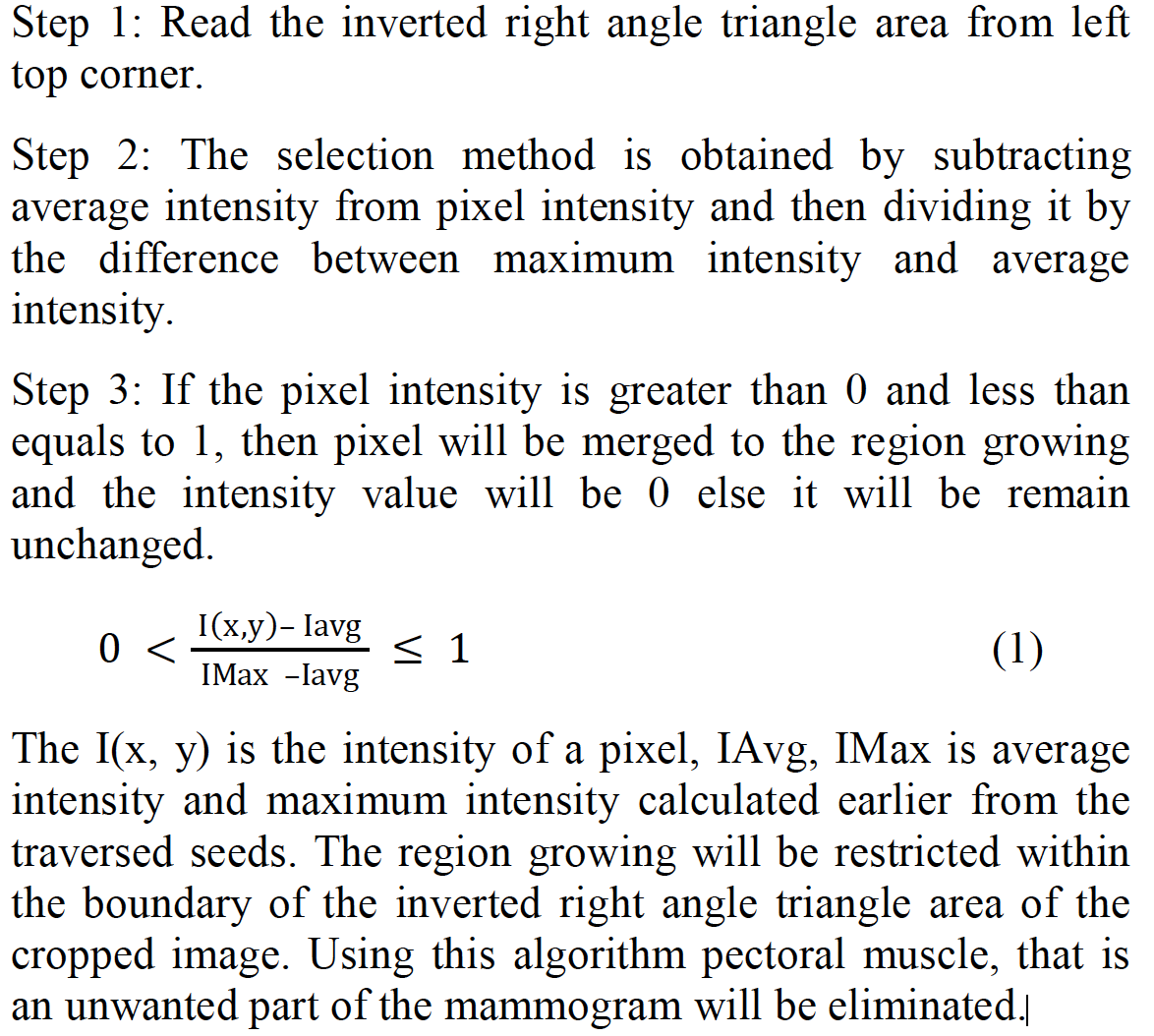


**Active Contour Method**

Active contour technique incorporates in two stages, first it removes hurdles like labels, patient name, scanning and taping artifacts using an **automatic algorithm based on thresholding**. **Applying active contour and automatic stopping algorithm it generates the contour that contains the boundary of the muscle part.** It extracts the muscle part binary image from the contour. Finally, combine with the muscle part binary image and the original image it obtains the required image and the pectoral muscle.

**Seeded Region Growing Technique**

Initially, a contrast enhancement by using the contrast limited **CLAHE** technique [21]. Then defining **the rectangle to 24 separate the muscle part of the ROI** and finally suppress the muscle part using modified **Seeded Region Growing (SRG) algorithm**. The muscle part have a rather higher intensity compared to the remainder of the breast tissue and seem in upper left corner of MLO view of the mammogram (the orientation of right breast mammograms are flipped horizontally to left). **After contrast enhancement applied SRG algorithm on the extracted rectangle, to suppress the pectoral muscle. Region growing is a procedure that groups pixels or sub regions into larger regions based on predefined criteria [3, 8, 10].** The basic approach is to start with a seed point and from this it grows into regions by appending to each seed those neighboring pixels that have properties similar to that seed. Selection of the seed depends on the nature of the problem. **A problem in the seeded region growing algorithm is the formulation of stopping rules.** Basically, growing a region should stop when no more pixels satisfy the criteria for inclusion in that region. The steps followed in the seeded region growing algorithm are

****

**Morphological Operations**

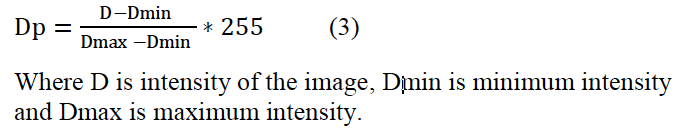
**Morphological operations** are basic and efficient preproprocessing techniques for mammography. The basic morphological operations are **dilatation, erosion, opening and closing** [16]. **Erosion is used to fill the gaps or holes in the mammography then dilated gradient mask shows the cells nicely but still some holes remains in the interior of the tumor. To eliminate these holes light structure is connected to image border [1]. In image the structuring element disc is applied with radius 5 starting with dilation followed by erosion, opening, closing and ends with Top hat and Bottom hat transforms [2]. It was originally defined for binary images, later on extended to grayscale images, and then for a complete lattices [5]. Artifacts can also be removed by morphological open operation followed by reconstruction operation [8].**

**Curve Fitting­**

The curve fitting techniques involve **computing matrix** and derive it in order to determine the highest inflection point at breast tissues pectoral muscle interface. By referring to equation (2) slop of breast equation was obtained at a specific calculated point [3]. **Mb = 3aI2 + 2bI + c** (2) Where, **I is highest variation point. Through this slope the inverted right angle was found that was the pectoral muscle.**

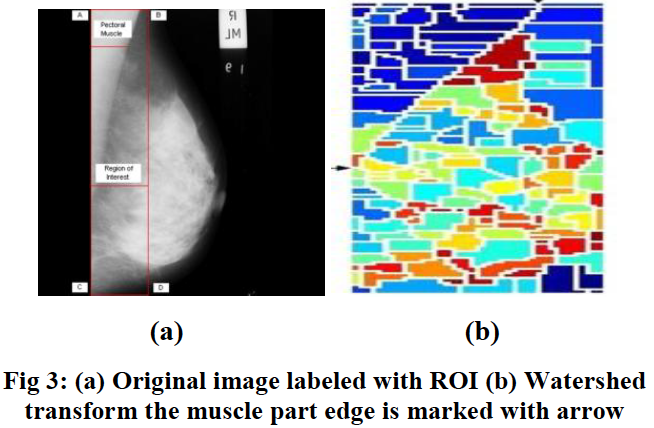
**Image Cropping and Pruning**

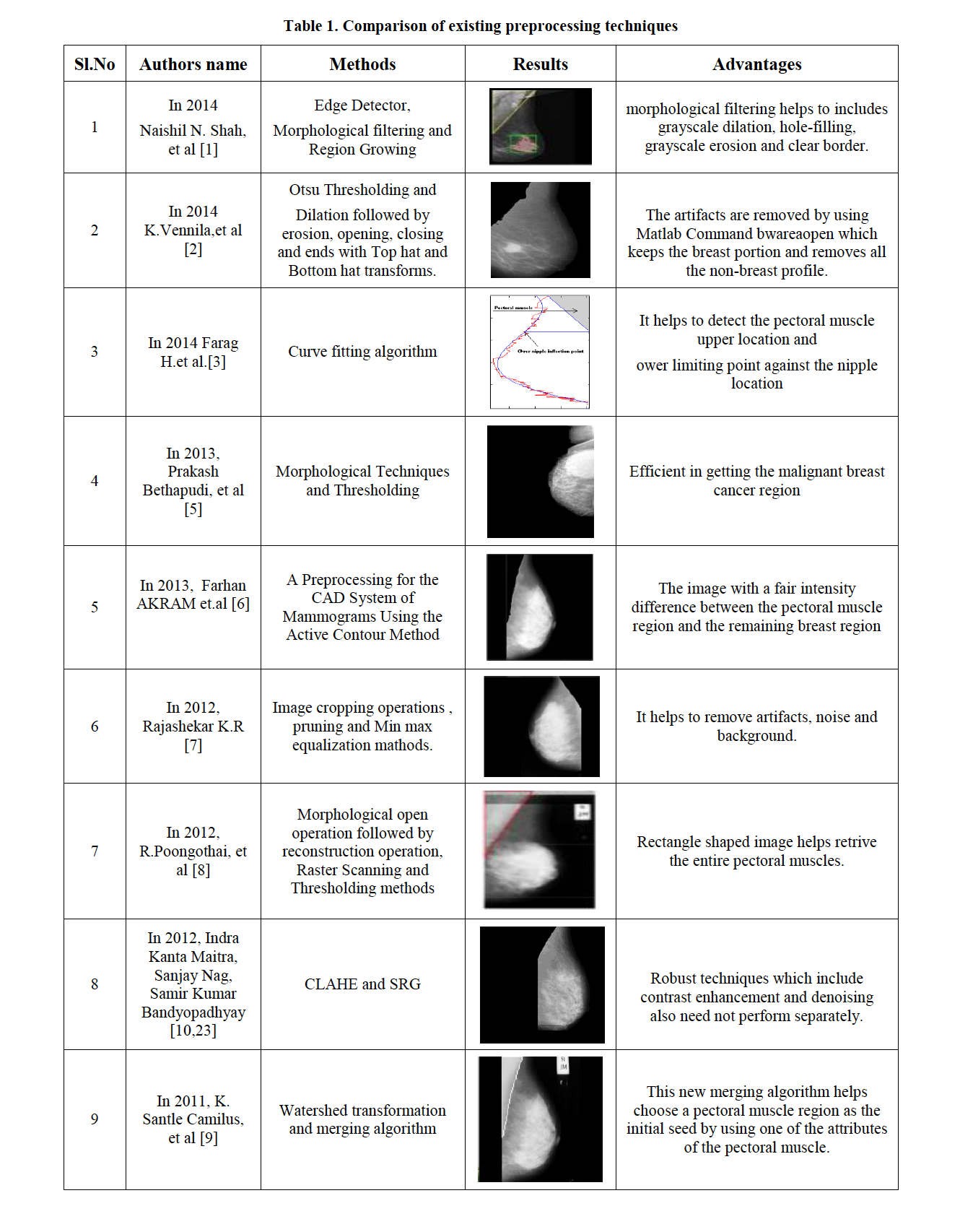
The spatial domain to reduce the effect **of higher brightness and lower darkness** in the image [7]. **Min-Max techniques helps assign lower darkness value is 0 which is min and higher brightness value is 255 which max**. The intermediate pixel values are changed according to equation



**Watershed Transformation**

When the gradient of the ROI of a mammogram [9] was treated with the watershed transform [16, 28, 30], the results, given in Figure 3, showed a **strong indication of the presence of the pectoral muscle boundary with a set of properties: (i) There is a unique continuous watershed line (hereafter it is called the watershed line of interest) which starts from the top and ends at the left of the image [19].** The width, extending from the left-most pixel of each row (starting from the top row) to its current position, is gradually decreasing and becomes zero when it reaches the left-most position at the end. (ii) The watershed line of interest encloses a triangular region covering the left top region of the image. **It has a curved shape. (iii) The pectoral muscle is oversegmented; this is caused by several irrelevant regional minima within the pectoral muscle**. The irrelevant regional minima may appear due to noise, local variations, etc.





**The region growing techniques and CLAHE respectively will be the best suitable techniques for this process.**