A Review on Image Processing Techniques for Breast Cancer Analysis

Iqra Rashid, Javeria Naz
COMSATS University Islamabad, Wah Campus, Pakistan, iqrarashid246@gmail.com, javeria126@gmail.com

Abstract
In recent years breast cancer detection has been the most popular research topic in medical image analysis. It is the most common malignancy in women, and men can also be affected. Conferring to the American Cancer Society, in 2019, almost two million new cases were registered, and the death rate was almost 41,000. The death rate can be reduced if the cancer is timely diagnosed. For cancer detection, different modalities are used, like MRI, ultrasound, and mammography. The most common and popular modality is mammography. A mammogram shows breast irregularities that are benign or malignant. In digital mammography, it is not easy to extract accurate breast regions. The main problem in the extraction region of concern is pectoral muscle suppression. The pectoral muscle appears in the breast area. Sometimes it is marked as an area of attention that causes a false positive rate. It is essential to eradicate pectoral muscles from the breast. This manuscript overviews the introduction of basic breast cancer terminologies. The work also analyzes state-of-the-insight imaging procedures used for breast cancer analysis.

Keywords: Breast cancer, Mammogram, Pectoral muscle, Segmentation and Classification

1. Introduction
Women’s widespread malignant disease is Breast Cancer (Amin, J., et al., 2022; Mughal, B., Sharif, M., Muhammad, N., & Saba, T., 2018). It is a very popular kind of tumor after lung cancer. It is usually manifest in women, but men also affect it (Yasmin, M., Sharif, M., & Mohsin, S., 2013). In 2019, almost two million new cases were identified in women, and two thousand new cases were identified in men. Almost 41,760 women and 500 men died from this disease in 2019. Every year almost 4,000 females die due to breast cancer in Pakistan (Menhas, R., & Umer, S., 2015). Breast tumor is a severe disease in women, especially the 40 to 55 year age group. Breast Cancer death rates increase with age. The breast contains fatty tissue intermixed with connective tissue—fewer visible segments, like lobes, ducts, and lymph nodes. A breast contains fifteen to twenty lobes units that also take various units named lobules. Breast parts are associated with narrow tubes named ducts. Connective tissue and muscles support the breast and provide shape to it. Nerves give sense to the breast.
consists of vessels similar to blood vessels and lymph vessels. A core structure is presented in Figure 1 (El-Sharkawy, et al., 2007). Genetic irregularity is the primary cause of breast tumors. Due to genetic irregularities, about 90% of such malignancies are developed in females. Only 5-10% of malignancies are affected by an irregularity inherited from parents.

Cancer happens when variation occurs in a gene that controls cell development. Due to the changes, cells are split and produced in a straightforward method. The unrestrained cells of cancer attack the healthy tissue of the breast and can move to the lymph nodes in the arms. Lymph nodules are the basic path that aid malignancy cells transfer in the rest parts of the body. Mainly, cancer occurs in lobules and ducts of the breast, and cancer generally occurs in the fatty tissue of the breast. Breast cancer can begin from various parts of the breast:

- Mainly cancer arises in ducts that transfer milk to the nipple.
- Certain cancer starts inside the glands which produce milk.
- Some minor cancers arise from breast tissues, and these cancers are not spread to other body parts.

In this research paper, we have highlighted the major symptoms, stages, types, modalities of breast tumors, and current methods proposed by the researchers are discussed in detail. The detection of tumors at the early stages is an important process of the computerized detection of tumors is helping the experts automate the detection process. Many researchers are proposing the artificial intelligence (Masood, S., Sharif, M., Masood, A., Yasmin, M., & Raza, M., 2015; Raza, M., Sharif, M., Yasmin, M., Masood, S., & Mohsin, S., 2012), machine learning (Irum, I., Raza, M., & Sharif, M., 2012) and deep learning-based methods for severe human diseases detection (Sharif, M. I., Li, J. P., Naz, J., & Rashid, I., 2020), such as Covid-19 (Amin, J., et al., 2022), Stomach Tumor Detection (Naz, J., et al., 2021; Naz, J., et al., 2021; Naz, J., et al., 2021), diabetic retinopathy (Amin, J., Sharif, M., Rehman, A., Raza, M., & Mufti, M. R., 2018), brain tumor analysis (Amin, J., Sharif, M., Yasmin, M., & Fernandes, S. L., 2018;

1.1. Breast Cancer Symptoms

In the primary stages, cancer cannot source any signs. Cancer may be minor to be sensed; however, an irregularity can be seen on a mammogram (Holland, J. H. K., 2019). The first symbol is typically lumped over the breast if cancer can be sensed, which was not earlier. Though, not every lump is a tumor. Breast cancer of all kinds can cause different symptoms. Several signs are the same, but few are different. Some common breast cancers symptoms are:

- The breast lump looks different from other near tissues and has grown recently.
- Pain in breast
- Red, marked skin on your whole breast
- All part of your breast is swollen
- Blood discharge from a nipple
- Peeling of skin on breast or nipple
- A quick, mysterious variation in the form of a mass of breast
- Changes in breast skin
- Swelling below the arm

On the off chance that anyone feels any of these signs, it does not denote you have breast cancer. Such as, pain in the lump may be caused by a non-cancerous cyst (Stephan, P., 2019).

1.2. Breast Cancer Stages

Breast cancer consists of four significant stages (Cadman, B., 2018). Physicians use the TNM staging system that AJCC and UICC introduce to determine staging systems to explain the stages of cancer. Table 1 shows the staging system (Azimi, N., Azar, A., Khan, A., & DeBenedectis, C. M., 2019).

Table 1: TNM staging system (Azimi, N., Azar, A., Khan, A., & DeBenedectis, C. M., 2019)

<table>
<thead>
<tr>
<th>A tumor (T)</th>
<th>Lymph Nodule (N)</th>
<th>Metastasis (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: 0 to 2 centimeters (cm)</td>
<td>N0: Swollen nodes cannot be felt by the Surgeon.</td>
<td>M0: Sample of lump has been detached surgically and verified that it is free from cancer.</td>
</tr>
</tbody>
</table>
T2: 2 to 5 centimeters (cm)  
N1: The specialist can sense various inflammation and thinks nodules are certain (tumorous).  
M1: Nodes consist of malignant cells or micrometastases. The tumor has a discarded cell outside its original place, and cancer may spread to other body parts.

T3: Larger than 5 cm  
N2: Lymph nodes perceive alike, fairly inflamed, unsmooth, and gathered together.  
-  

T4: Tumor of any size which has broken from the skin  
N3: Inflamed nodules are close to the collarbone.  
-  

The meaning of the TNM staging system is following:
• T is a tumor and specifies how much breast tissue is involved.
• Abbreviation of N is nodes and shows that cancer extent to the lymph nodes.
• Abbreviation of M is metastasis, and it shows that cancer extent to the rest body parts.

TNM staging structure usage numbers. These numbers range from 0 to 4, and detail of this range is expressed in the next section. Figure 2 indicates the Breast cancer stages (Abdullah-Al-Wadud, M., et al., 2007).

1.2.1. Stage 0  
It is a non-aggressive type of cancer, and it implies that malignant growth has not spread to other parts and the dangerous cells stay in the bosom where they began developing. Non-invasive cancer is recognized as DCIS, and it resides in ducts. In the initial analysis, stage 0 cancer patients can obtain quick treatment.

1.2.2. Stage 1  
In stage 1, cancerous cells are an assault the nearby muscle, and it is not spread to the other body parts. It has two sub-categories: 1A and 1B.

1.2.3. Stage 2  
Stage 2 cancer is contained in the breast, and this cancer’s growth is near the lymph nodes. It is divided into two sub-categories: i) stage 2A, ii) stage 2B. The differ-
ence in these stages is measured by tumor size.

1.2.4. Stage 3
Stage 3 cancer is an invasive type of cancer, and it is invaded the lymph nodes but does not spread to the other body parts. Stage 3 is considered to be advanced; there is a large number of effective treatment choices.

1.2.5. Stage 4
It is the most advanced stage, also called metastatic cancer. In stage 4, cancer may be recurring in the breast, which is now spread to other body parts.

1.3. Breast Cancer Types
Malignancy is a set of diseases that affect irregular changes and the rise of breast cells. Breast cells are classified into two forms such as i) cancerous cells and ii) non-cancerous cells. Cancerous cells are malignant tumors and divided into infiltrating cancer and in situ cancer, while non-cancerous cells are benign or normal cells. Benign cancer (Benign Breast Conditions, 2019) rises slowly and does not influence the nearby tissues, but cancer (Yamamoto, S., et al., 2019) raises rapidly and ruins tissue. In many cases, the malignant cells formulate a lump called a tumor. Breast cancer looks in several chest parts, like ducts and lobules. Cancerous cells increase unusually in the breast, ultimately scatter in the body if not cured. Breast tumor arises basically in women, though men can also be affected. Two main categories are i) Non-invasive cancer ii) Invasive cancer.

1.3.1. Non-Invasive
It is a common category of cancer. In this category, the cancer cell is not extent very quickly. Non-invasive (Akram, M., Iqbal, M., Daniyal, M., & Khan, A. U., 2017) consists of two kinds of cancer that are: i) Lobular Carcinoma in Situ (LCIS) and ii) Ductal Carcinoma In Situ (DCIS).

1.3.1.1. Lobular Carcinoma in Situ (LCIS)
It is a category of cancer that arises in the milk-making glands. Abnormal cells of LCIS (Bahadure, N. B., Ray, A. K., & Thethi, H. P., 2017) are presented in Figure 3. Similarly, DCIS cells are not invading the near tissues.

Fig. 3. Abnormal cells of LCIS (Bahadure, N. B., Ray, A. K., & Thethi, H. P., 2017)
1.3.1.2. Ductal Carcinoma in Situ (DCIS)
It is a category of non-intrusive cancer. Using DCIS, cancer cells are narrowed in the ducts of the breast and do not invade the near breast tissue. Abnormal cells of DCIS in the breast duct are shown in Figure 4 (Ben Rabeh, A., Benzarti, F., & Amiri, H., 2017).

![Ductal Carcinoma In Situ (DCIS)](image)

Fig. 4. Abnormal cells of DCIS (Ben Rabeh, A., Benzarti, F., & Amiri, H., 2017)

1.3.2. Invasive
It is the second category of cancer. In Invasive cancer (Invasive Breast Cancer: Symptoms, Treatments, Prognosis, 2019), the cancer cell is spread very rapidly in another part of the breast body. Four types of Invasive cancer are described below:

1.3.2.1. Invasive Lobular Carcinoma (ILC)
It is abnormal cells. The spread outer walls of lobules create the milk, which is empty in milk ducts, and it is also spread around the breast tissue. Cells of ILC are shown in Figure 5.

![Invasive Lobular Carcinoma](image)

Fig. 5. Cells of ILC

1.3.2.2. Invasive Ductal Carcinoma (IDC)
This sort of cancer starts in milk vessels and then propagates near the breast tissue.
In Figure 6, IDC abnormal cells are shown. Once cancer has invaded tissue outside the milk ducts, it can spread to other nearby organs and tissue.

**Fig. 6. Abnormal Cells of IDC**

1.3.2.3. **Inflammatory Breast Cancer (IBC)**

It is a fast-rising type of bosom cancer where cancer cells penetrate the skin and lymph of the breast. IBC does not make any different tumors that could be felt and separated from the breast. Due to breast cancer, lymph nodes are blocked then signs of this cancer start to show.

1.3.2.4. **Metastatic**

Metastatic is a breast cancer category that invades other organs—common organs (brain, liver, bones, and lungs) where the cancer cells transfer. Bosom cancer cells disperse from the original tumor by the bloodstream and lymphatic system.

1.4. **Breast Cancer Modalities**

The modalities use breast images which are acquired by different methods. For a

**Fig. 7. Breast Cancer Modalities**
better existence and to overcome treatments, different imaging modalities are intro-
duced to identify the disease as soon as possible. Few modalities of imaging are used
for screening, few for indicative purposes, and others for combined assessment.

The most commonly used modalities are mammography, breast ultrasound, ther-
mography, Magnetic Resonance Imaging (MRI), Computed Tomography/ Positron
Emission Tomography (CT/PET), CBE, BSE. In Figure 7, Breast cancer modalities are
presented.

1.5. Cause of Breast Cancer

Breast cancer usually yields no signs in the initial phase once the lump is minor
and cured (Mughal, B., et al., 2018). So, it is too vital for females to monitor suggested
screening for observing breast cancer at the initial phase. Painless swelling is aware of
physical symptoms that may be felt through breast size growth. A few common symp-
toms of growing breast tumors are any diligence variation in the breast, like a lump,
redness of breast skin, and nipple irregularities like sudden discharge or retraction
(Sanuade, O. A., et al., 2021). Significantly, soreness does not show the existence or
the nonexistence of the breast tumor. Screening tests search for cancer at the primary
stage; earlier signs look. Typically, there are three kinds of breast malignancy ways
suggested for initial findings: CBE, BSE, and different imaging methods such as MRI,
ultrasound, mammography, etc. Mammography is used for the detection of the initial
stage of cancer. Initial recognition of cancer is a difficult task. Different modalities
are used for the recognition of breast cancer. The most recommended modality is
mammography. It is problematic to detect abnormal tissue at the initial stage due to
mammography images' low quality. Recently, the researcher introduced different com-
puter vision-based methods for finding breast cancer. A method contains some steps:
preprocessing, artifacts and label removal, pectoral muscle suppression, and image
segmentation.

1.6. Problems in Automated Breast Cancer Analysis

In digital mammography, it is challenging to extract the accurate region of the
breast. Mammography of the image contains an area of the breast, pectoral muscle,
artifacts, labels, and markers. Sometimes it is marked as an area of interest that causes
a false-positive ratio (Dhahri, H., et al., 2019). It is important to eradicate pectoral mus-
cles, artifacts, labels, and markers from the breast area. For initial recognition of can-
cer, it is challenging work to improve the quality of an image. Many challenges occur
during the segmentation of the breast area that is listed below (Chan, H. P., Samala, R.
K., & Hadjiiski, L. M., 2019):

- The presence of noise in a mammography image is a basic challenge for breast
  region segmentation.
- The presence of artifacts, labels, and markers in the mammography reduce the
  performance of the image.
- Color information about healthy and unhealthy images are the same.
- If an image is not segmented properly, they affect the performance measures.

The motivation of this manuscript is to provide an insight into the segmentation and
classification of cancer from a mammogram. Existing segmentation approaches have
low accuracy, specificity, and high false-positive rates. Due to the low accuracy of the results, the death rate of women is increased. Different problems occurred in the breast images, such as low visibility of images, poor contrast image, the existence of labels and artifacts, and pectoral muscle in the region of interest.

2. Literature Review on the Existing State of the Art Imaging Methodologies

Mammography is the most recommended and popular modality for the recognition of cancer. In digital mammography, it is a challenging task to extract accurate regions of the chest. It is essential to eradicate pectoral muscles, labels, and artifacts from the mammography. The pectoral muscle is situated in the chest of humans. Typically, pectoral muscles appear in Medio-Lateral Oblique (MLO) X-ray picture [48]. Due to similar pixel values of the pectoral muscle or breast muscle, it is difficult to extract a particular area of the breast. Different researchers used CAD systems (Sasikala, S., Ezhilarasi, M., & Arun Kumar, S., 2020) for segmentation and classification (Saba, T., et al., 2021).

2.1. Image Preprocessing Phase

Pre-processing is an important stage in a computer-aided diagnostic system (Lal, M., et al., 2018; Rehman, M., Iqbal, M., Sharif, M., & Raza, M., 2012; Khan, M. A., et al., 2020; Raza, M., et al., 2018; Sharif, M., Mohsin, S., Jamal, M. J., & Raza, M., 2010, July). It is essential to preprocess images for better results in segmentation and classification (Amin, J., et al., 2022; Nasir, I. M., et al., 2022; Fayyaz, A. M., et al., 2022; Khan, M. A., et al., 2020). Preprocessing of mammogram images (Abdel-Nasser, M., et al., 2017) consists but is not limited to: resizing of mammograms, mammogram image enhancement, segmentation of breast cancer, and determination of orientation and extraction of the pectoral muscle. Image enhancement and denoising algorithms can enhance mammogram images’ detailed information and reduce noise. Appropriate detection of cancer and segmentation leads to accurate segmentation and classification results. Figure 8 indicates preprocessing steps of mammograms.

![Fig. 8. Pre-processing steps of mammogram images](image-url)
2.1.1. Image Denoising

Mammogram images consist of noise, labels, and pectoral muscles, which affect the mammogram images of breast cancer. To remove these issues, the first step is preprocessing. Different researchers used different techniques to remove noise. Gaussian noise removes from the mammogram by using a mean filter (Jen, C. C., & Yu, S. S., 2015). In (Ramani, R., Vanitha, N. S., & Valarmathy, S., 2013), remove noise from mammograms using an adaptive median filter. Adaptive Median filtering is presented to smooth mammogram images without blurring images and preserved edges detailed. Anscombe Transform and Wiener filter (Souto, L. P. M., dos Santos, T. K., & Silva, M. P. S., 2018, July) are implemented for image denoising. In (Bhateja, V., Misra, M., & Urooj, S., 2020), the noise is overcome by using local iterative noise variance estimation. Images attained from a mammogram may contain noises added while capturing the image. Removing noise is still a difficult task. Many filters are introduced to remove the noise from mammogram images. Four different filter techniques are used to remove noise. A hybrid Median Filter (HMF) is one of the best filter techniques (Joseph, A. M., John, M. G., & Dhas, A. S., 2017, March). The ultrasound dataset is used in this research. Ultrasound images consist of speckle noise and Gaussian noise. Wiener and Median filters are applied (Andria, G., et al., 2012). Two algorithms, histogram equalization and median filter, are used to denoise the image. In Histogram equalization, pixels of the image are strained and increase the contrast. The median filter overcomes salt and paper noise (Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P., 2015). Gaussian variable and Laplacian variable are used for noise and edges coefficient.

After that, the shrink function is combined at consecutive scaling, and then the wavelet transform is put on (Scharcanski, J., & Jung, C. R., 2006). To overcome the effects of mammograms, dual-tree contourlet transform (DCT) is used. DCT has shift-invariant, directionality, and anisotropy (Dong, M., et al., 2015, December). Different linear and nonlinear filter approaches are applied to mammogram images to denoising the image. A Wiener filter is a suitable approach for denoise the image (Makandar, A., & Halalli, B., 2015). Wavelet and curvelet transform is used with thresholding techniques to suppress noise from mammogram images. In one work, wavelet transform is implemented to remove noise from images (Makandar, A., & Halalli, B., 2015). Various Image denoising and contrast enhancement methods are presented in Table 1.

2.1.2. Image Enhancement

To enhance visualization of images for better understanding is called image enhancement (Rehman, M., Sharif, M., & Raza, M., 2014; Irum, I., Shahid, M. A., Sharif, M., & Raza, M., 2015; Shah, G. A., et al., 2015; Irum, I., Sharif, M., Yasmin, M., Raza, M., & Azam, F., 2014; Irum, I., Sharif, M., Raza, M., & Yasmin, M., 2014). This method highlights the details and alters information from the image (Sharif, M., Irum, I., Yasmin, M., & Raza, M., 2017). To get a better result in medical imagining (Naz, J., et al., 2021; Shahzad, A., et al., 2021), different techniques (Maini, R., & Aggarwal, H., 2010; Singh, G., & Mittal, A., 2014) are used for image enhancement. Different researchers have used different techniques to improve mammography images edges and information in breast cancer detection. Several approaches are presented for enhancing the image without isolating details of the image. In (Akila, K., Jayashree, L. S., & Vasuki,
A., 2015), different contrast enhancement methods are implemented on low contrast mammograms for enhancing images. Recursive Mean-Separate Histogram Equalization (RMSHE) is the best enhancement technique to enhance mammograms. A novel approach is used to enhance the low contrast mammographic image. The fuzzy-based approach is used for the enhancement of images (Angadi, S. A., & Kodabagi, M. M., 2013, August). A new Guided Image Filter technique is used in (Houben, G., Fujita, S., Takahashi, K., & Fujii, T., 2019) for detail enhancement and edge smoothness of mammographic images. The guided filter used the edge-preserving operator. Guided filtering is the nonlinear filter that not just plane low gradient area it also reserves robust boundaries. It contains the edge-preserving as matched to bilateral filters. CLAHE is presented for enhancement of a mammogram. The wavelet transform technique is used in (Yousefi, P., 2015, November) for mammography contrast enhancement. Wavelet transform has enhanced the resolution of images. Decomposition of mammograms has enhanced the images.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Method</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee, S., et al., 2019</td>
<td>2019</td>
<td>FNLM</td>
<td>Denoising</td>
</tr>
<tr>
<td>Bhatnagar, S., &amp; Gupta, R., 2019</td>
<td>2019</td>
<td>DWT</td>
<td>Denoising</td>
</tr>
<tr>
<td>Mayer, A., 2019, October</td>
<td>2019</td>
<td>CLC-NLM</td>
<td>Denoising</td>
</tr>
<tr>
<td>Chan, N. H., Hasikin, K., &amp; Kadri, N. A., 2019, March</td>
<td>2019</td>
<td>FRB</td>
<td>Contrast Enhancement</td>
</tr>
<tr>
<td>Dabass, J., et al., 2019, March</td>
<td>2019</td>
<td>Intuitionistic Fuzzy</td>
<td>Contrast Enhancement</td>
</tr>
</tbody>
</table>

2.2. Image Orientation

Flipping of the mammogram image is called orientation, and the density value determines whether mammogram images are Left Mediolateral Oblique (L-MLO) images or the Right Mediolateral Oblique (R-MLO) images. Different researchers used different methods for the direction of the mammogram images. The orientation of the mammogram image is shown in Figure 9 (Bandyopadhyay, S. K., 2018).
2.3. Image Binarization

Conversion of a grayscale image in black and white is called binarization of image. A threshold value is used in image Binarization (Garg, N., & Garg, N., 2013). In this process, all pixel values specify and define threshold values. If a value is smaller than a particular value, the image is black; otherwise white. In breast cancer detection, different methods are used to binarize the image. Adaptive local thresholding is used in (Anitha, J., Peter, J. D., & Pandian, S. I. A., 2017) for binarization of mammogram images. Bright local spots are binarized using the adaptive thresholding method. Two algorithms, otsu and Kittler’s method, have been applied for binarization (Pragathi, J., & Patil, H. T., 2013). In the Otsu threshold, the gray level image decreases to a binary image and is also used in Kittler’s method for binarization. In (Saidin, N., et al., 2013), thresholding is used for image binarization. Mammogram and Binarized images are described in Figure 10.

2.4. Artifacts and Label Removal

Artifacts consist of labels and markers that reduce the quality of mammographic images. Due to artifacts, it is difficult to detect the tumor. Different radiologists identify the artifact that can help detect cancer (Odle, T. G., 2015). The monostatic artifacts suppression procedure is used to eradicate early-stage labels from the multistate ra-
The entropy-based adaptive technique is implemented to combine signals by related artifacts and then eradicate artifacts from each group individually by a hybrid artifacts exclusion algorithm (Elahi, M. A., et al., 2017). The reference wavelet is based on an average of artifacts in all channels. Subtract the reference waveform from the recorded signal to remove the artifacts (Elahi, M. A., et al., 2013). Morphological operation (Bajaj, V., et al., 2019) is used for the suppression of artifacts. It is the most widely used method for removing isolated pixels from mammographic images. A novel approach, an automated iterative perturbation correction algorithm, is presented in (Uddin, K. S., & Zhu, Q., 2019) to remove the artifacts from the images. It is based on the visual wavelengths' structural similarity (SSIM) index. Artifacts in mammograms affect the segmentation of images. A weighted median filter is used to eradicate artifacts from images (Taifi, K., et al., 2020). Artifacts decrease the quality of images; to overcome this problem (Ling, Q., et al., 2018) patch-based frequency signal filtering technique is applied to remove the artifacts. Mammogram images consist of different types of labels like low-intensity and high-intensity labels. The blob or bounding box analysis technique is presented in (Shah, N. N., Ratanpara, T. V., & Bhensdadia, C. K., 2014) for removing the low-intensity label. In (Moghbel, M., et al., 2020; Arefan, D., et al., 2015) researcher used the threshold technique and morphological operations to eliminate labels and artifacts. Various existing Artifact and Label Removal Phase techniques are given in table 3.

Table 3. Various Techniques of Artifact and Label Removal Phase

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Method</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joseph, A. J., &amp; Pournami, P. N.</td>
<td>2021</td>
<td>Thresholding, associated element split for artifact, and label removal</td>
<td>Multifractal theory-based breast tissue portrayal</td>
</tr>
<tr>
<td>Sarah Siham Fadhil et al.</td>
<td>2021</td>
<td>Region Growing, OTSU, Split Orientation Local Thresholding (SOLTH)</td>
<td>Noise Reduction, Artifact Removal, Pectoral Muscles Detection</td>
</tr>
<tr>
<td>Bandyopadhyay, S. K.</td>
<td>2019</td>
<td>flipdim</td>
<td>Orientation</td>
</tr>
<tr>
<td>Al-Khalidi, F. Q., Alkindy, B., &amp; Abbas, T.</td>
<td>2019</td>
<td>Threshold</td>
<td>Binarization</td>
</tr>
<tr>
<td>Lukashenko et al.</td>
<td>2019</td>
<td>Quantization, threshold</td>
<td>Binarization</td>
</tr>
<tr>
<td>Tavakoli, N., et al.</td>
<td>2019</td>
<td>Otsu threshold</td>
<td>Binarization</td>
</tr>
<tr>
<td>Suri, J. S., Sun, Y., &amp; Janer, R.</td>
<td>2019</td>
<td>Adaptive localized threshold</td>
<td>Binarization</td>
</tr>
</tbody>
</table>

2.5. Pectoral Muscles Suppression

Muscles that connect the bones of the shoulder and arm with the front of the chest are called pectoral muscles. Generally, pectoral muscle seems in mammogram images' mediolateral-oblique (MLO) fragments. The pectoral muscle is a core section
of mammograms that contains important information and influences the segmentation and classification method, which is the source of a high false-positive rate, so it is obligatory to suppress the pectoral muscle. Various methods are used for the removal of pectoral muscles. Some techniques are described in Table 4. Many techniques are failed due to variations in pectoral muscle. Each pectoral muscle is different in size, intensity, and shape. Hough transformation techniques are based on distance transformation.

**Table 4. Various Pectoral Muscle Suppression and Segmentation**

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Dataset</th>
<th>Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Divyashree, B. V., et al.</td>
<td>2022</td>
<td>Mini-MIAS</td>
<td>Granular Computing</td>
</tr>
<tr>
<td>Sagheer, A. M., &amp; Veisi, H.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gómez, K. A. H., et al.</td>
<td>2021</td>
<td>Mini-MIAS and UTP</td>
<td>Region-growing segmentation and polynomial contour fitting</td>
</tr>
<tr>
<td>Rahman et al.</td>
<td>2020</td>
<td>RIDER Breast MRI Dataset</td>
<td>OTSU, Thresholding, Holes Filling</td>
</tr>
<tr>
<td>Sakai, A., et al.</td>
<td>2020</td>
<td>Aoyama Hospital dataset, + public data</td>
<td>Normalization, SVM, Naïve Bayes, Random Forest</td>
</tr>
<tr>
<td>Rampun et al.</td>
<td>2019</td>
<td>MIAS, BCDR, IN-breast</td>
<td>Contour based CNN</td>
</tr>
</tbody>
</table>

This technique divides the image into sub-regions (Rodriguez-Ruiz, A., et al., 2019). A modified tracking algorithm is presented in (Sreedevi, S., & Sherly, E., 2015) to identify the pectoral muscle accurately. The pectoral muscle suppression is essential for recognizing microcalcification because it contains comparable pixel amounts like lesions which influence the consequences of automatic recognition. The random Sample Consensus (RANSAC) algorithm and the morphological process are presented in (Yoon, W. B., et al., 2016) for pectoral muscle suppression. A novel process is proposed to eliminate annotations usually initiated in mammograms. An adaptive algorithm is presented for automatically pectoral muscle suppression (Majeed, T. F., Al-Jawad, N., & Sellahewa, H., 2013, September). Computerized pectoral muscle exclusion from Medio-Lateral Oblique view mammograms is an essential step for the mammography processing techniques. The automatic detection of a breast cancer occurrence in the pectoral muscle gives false-positive results. In the pectoral muscle
segmentation, three existing methods (Shinde, V., & Thirumala Rao, B., 2019), region growing, K-mean clustering, and thresholding, have been implemented, and the machine learning algorithm was implemented.

2.6. Mass Segmentation

Divide the image into different segments is known as segmentation (Shahzad, A., Sharif, M., Raza, M., & Hussain, K., 2008). Mass dissection of the breast is a vital phase in the computer-aided diagnosis (CAD) system for achieving accurate results with the depletion of false presumptions. In breast mass segmentation, mammographic images are used, and it is a difficult task due to some problems like irregular shapes of boundaries, different intensity ranges of images and breast tissues, Labels, artifacts, and the existence of pectoral. These challenges confused radiologists regarding how they could find the area of interest. The most vital information of the image is contour and shape since it gives essential info about the extent ability of mass. The most common segmentation technique is thresholding. Some techniques for mass segmentation are described in Table 5. For the segmentation, rough entropy-based granular computing is used in (Roselin, R., & Thangavel, K., 2012, March). Morphological watershed transformation is presented in (Sharma, J., & Sharma, S., 2011) for mass segmentation. Adaptive hysteresis thresholding is presented in (Mughal, B., Muhammad, N., & Sharif, M., 2019) for mass segmentation (Singh, V. K., et al., 2020).

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Dataset</th>
<th>Segmentation Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarangi, S., Rath, N. P., &amp; Sahoo, H. K.</td>
<td>2021</td>
<td>MIAS</td>
<td>Legendre neural network with a single layer with optimal Threshold</td>
</tr>
<tr>
<td>Zhou, K., Li, W., &amp; Zhao, D.</td>
<td>2021</td>
<td>Mini-MIAS, DDSM</td>
<td>Deep learning based region extraction, contrast improvement by applying CLAHE and segmentation using Deeplab v3+</td>
</tr>
<tr>
<td>Li, Y., Zhang, L., Chen, H., &amp; Cheng, L.</td>
<td>2020</td>
<td>INbreast, and two private dataset (BCPKUPH) and TXMD</td>
<td>Self-supervised learning Network, and Siamese Faster RCNN</td>
</tr>
<tr>
<td>Radhi, E. A., &amp; Kamil, M. Y.</td>
<td>2021</td>
<td>Mini-MIAS</td>
<td>Chan-Vese, Active Contour</td>
</tr>
<tr>
<td>Caballo, M., et al.</td>
<td>2020</td>
<td>Dataset collected from 69 Patients</td>
<td>generative adversarial network (GAN), UNet, multivariate analysis of variance (MANOVA) and inter class correlation (ICC)</td>
</tr>
<tr>
<td>Wang, R., et al.</td>
<td>2019</td>
<td>INBreast</td>
<td>MNPNet</td>
</tr>
</tbody>
</table>
2.7. Tumor Classification

classification is an essential assignment of image analysis (Hameed, M., et al., 2012). In this step, mammograms are categorized into two groups which are normal and abnormal. Breast tumor classification is vital for assessing tumors and suggesting treatment according to their classes. Different researchers used different techniques for tumor classification. Various classification techniques are shown in Table 6. In (Setiawan, A. S., Wesley, J., & Purnama, Y., 2015) Artificial Neural Network (ANN) is implemented to catalog tumors. The classification was performed in two stages: firstly, the dataset was categorized into normal and abnormal classes then abnormal images were subdivided into benign and malignant. In (Rodríguez-López, V., & Cruz-Barbosa, R., 2015, June) matched Bayesian network models' performance for confine benignity and malignancy of the tumor. Adaptive De-convolutional Networks have been used in (Litjens, G., et al., 2017) for tumor classification. The probabilistic Neural Network (PNN) and the radial function are implemented in (Hamad, Y. A., Simonov, K., & Naeem, M. B., 2018, November) for the automatic tumor classification.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Dataset</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vyshnavi, V., Vijayan, D., &amp; Lavanya, R.</td>
<td>2021</td>
<td>MIAS</td>
<td>Breast Imaging-Reporting and Data System (BI-RADS)</td>
</tr>
<tr>
<td>Lbachir, I. A., Daoudi, I., &amp; Tallal, S.</td>
<td>2021</td>
<td>MIAS, CBIS-DDSM</td>
<td>HRAK algorithm, SVM</td>
</tr>
<tr>
<td>Li, H., Mukundan, R., &amp; Boyd, S.</td>
<td>2021</td>
<td>FFDM, IN-Breast</td>
<td>Multifractal features and LBP features</td>
</tr>
<tr>
<td>Sakai et al.</td>
<td>2020</td>
<td>Aoyama Hospital data, public data</td>
<td>Normalization, SVM, Naïve Bayes, Random Forest</td>
</tr>
<tr>
<td>Tsochatzidis, L., Costaridou, L., &amp; Pratikakis, I.</td>
<td>2019</td>
<td>DDSM</td>
<td>CNN</td>
</tr>
<tr>
<td>Boldbaatar, E. A., Lin, L. Y., &amp; Lin, C. M.</td>
<td>2019</td>
<td>WBCD</td>
<td>RWENN</td>
</tr>
</tbody>
</table>

2.8. Breast Cancer Datasets

Various breast cancer datasets are available on the internet, and most of the datasets are not publicly available. The most common and open accessed datasets are Mini-MIAS and DDSM (Kaur, P., Singh, G., & Kaur, P., 2019; Lee, R. S., et al., 2017). The mini-MIAS dataset comprises 322 images. The dataset comprises three kinds of images that are Normal, Benign (B), and Malignant (M), and related tissues are Normal Fatty Tissue (F), Fatty-Glandular Tissue (G), and Dense-Glandular Tissue (D). In Figure 2.8, some sample images are presented:
3. Conclusion

In this work, a review of methods for finding and categorization of cancer from mammograms is presented. The review consists of five pipeline processes of breast cancer analysis, including tumor classification, preprocessing, artifacts and labels removal, pectoral muscle suppression, mass segmentation, and tumor classification. This work aims to give an insight to researchers that will lead them to build novel methods for automated segmentation and classification of a mammogram. The paper introduces basic breast cancer terminologies, then the problems of breast cancer CAD systems development are described. The paper has then proceeded to the state of the insight methods for breast cancer analysis.

References


mogram images. *Technology, 10*(02), 96-105.


Iqra Rashid, et al.


grams with depth-first search algorithm towards breast density classification. *Biocybernetics and Biomedical Engineering, 41*(3), 1224-1241.


Saba, T., et al. (2021). Suspicious activity recognition using proposed deep
L4-branched-ActionNet with entropy coded ant colony system optimization. *IEEE Access, 9*, 89181-89197.


Submitted: 31.08.2021
Accepted: 23.11.2021