Systematic Review of Computing Approaches for Breast Cancer Detection Based Computer Aided Diagnosis Using Mammogram Images

Dilovan Asaad Zebari¹, Dheyaa Ahmed Ibrahim², Diyar Qader Zeебariee⁶, Habibollah Haron⁵, Merdin Shamal Salih⁷, Robertas Damaševičius⁸, and Mazin Abed Mohammed⁹

¹Center of Scientific Research and Development, Department of Computer Science, Nawroz University, Duhok, Iraq; ²Communications Engineering Techniques Department, Information Technology Collage, Imam Ja’afar Al-Sadiq University, Baghdad, Iraq; ³Research Center of Duhok Polytechnic University, Duhok, Iraq; ⁴School of Computing, Faculty of Engineering, University Teknologi Malaysia, Johor Bahru, Malaysia; ⁵Faculty of Applied Mathematics, Silesian University of Technology, Gliwice, Poland; ⁶Information Systems Department, College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq

ABSTRACT
Breast cancer is one of the most prevalent types of cancer that plagues females. Mortality from breast cancer could be reduced by diagnosing and identifying it at an early stage. To detect breast cancer, various imaging modalities can be used, such as mammography. Computer-Aided Detection/Diagnosis (CAD) systems can assist an expert radiologist to diagnose breast cancer at an early stage. This paper introduces the findings of a systematic review that seeks to examine the state-of-the-art CAD systems for breast cancer detection. This review is based on 118 publications published in 2018–2021 and retrieved from major scientific publication databases while using a rigorous methodology of a systematic review. We provide a general description and analysis of existing CAD systems that use machine learning methods as well as their current state based on mammogram image modalities and classification methods. This systematic review presents all stages of CAD including pre-processing, segmentation, feature extraction, feature selection, and classification. We identify research gaps and outline recommendations for future research. This systematic review may be helpful for both clinicians, who use CAD systems for early diagnosis of breast cancer, as well as for researchers to find knowledge gaps and create more contributions for breast cancer diagnostics.

Introduction
In 2015, the World Health Organization (WHO) announced that cancer is the second-largest contributor to global deaths. Breast cancer is the leading cause of cancer-related mortalities among women, trailed by colorectal and lung...
cancers (Mohammed et al. 2018) (Obaid et al. 2018). Breast cancer could be effectively diagnosed by employing a medical image examination. Various techniques of medical imaging may be used to examine Infrared Thermography (IRT), microscopic (histological) images, Magnetic Resonance Imaging (MRI), Ultrasound (US), and Digital Mammograms (DMs). To support radiologists in the method of interpreting images and identifying abnormalities, the usage of these modalities renders the process more effective by reducing mortality rates by 30–70%. Utilizing computerized feature extraction and classification that is devised as Computer-Aided Diagnosis (CAD) can become a beneficial technique for physicians in diagnosing and identifying abnormalities (Lahoura et al. 2021).

The primary role of a CAD system is to resolve the challenge of interpreting DMs. The goals of the system include effectively diagnose cancer and correctly interpret DMs. The CAD structures were developed to resolve the reliance of the operator in terms of diagnosis and decrease the cost of medical complementary technology (Mohanty, Senapati, and Lenka 2013). In the analysis on detecting cancer cells by CADs, 80% of the diagnosed cells were able to be detected without CAD, whereas the percent of tested tumor cells that were detected by CADs improved to 90% inside CAD (Horsch et al., 2011). Computerized diagnosis assesses the knowledge which a person or a computer gathers and offers an outcome to decide what kind of lesion is present and whether that is cancerous or not (Zeebaree et al., 2019).

Medical imaging technology with applying CAD-based Machine Learning Techniques (MLTs) is becoming common for cancer diagnosis and detection. To resolve the deficiency and ameliorate the efficiency of the CAD algorithms, the value of representation learning has been highlighted in recent years (Han et al., 2015) (Zeebaree et al., 2019). Deep Learning (DL) is one of representation learning strategies that use the hierarchical representations of image data as features. The main characteristic of DL is that it can take the content and encode it in a high-level of function representation (e.g., vector) without the need for post-processing (LeCun, Bengio, and Hinton 2015).

The main contribution of this review study is to introduce the recently introduced methods in state-of-art that concentrate on various Deep Learning Techniques (DLTs) and Machine Learning Techniques (MLTs) utilized in breast cancer identification based on DMs. The survey seeks to illustrate the issues that remain as to the applicability of DMs in the early detection of breast cancer. This study analyzes the most recent works that have discussed this topic and offers some perspective on current problems. We explore previous works that tackled these challenges, and eventually gives some observations and the potential directions of future study that would be taken to enable more progress. This systematic review is divided in two main parts. The first part introduces the methodology of this research and the CAD methodology with
its steps as well as the ML and DL techniques. The second part of this research presents the review of each phase of the CAD system of the most recent studies.

**Methodology**

The main aim of this study is to identify state-of-the-art studies in the context of CAD systems, especially in the domain of breast cancer identification using DM images also, both Machine Learning (ML) and Deep Learning (DL) techniques as classifiers. To find the answer to the following research questions is the primary purpose of this study:

1. What are DM breast cancer datasets mostly used on CAD systems?
2. What techniques are used for each CAD stage?
3. What challenges that are faced during each stage of CAD?
4. What enhancement techniques are currently applied in the pre-processing stage?
5. What segmentation techniques are applied to derive Region-of-Interest (ROI) in DM images?
6. What type of features are extracted from DM images?
7. What techniques are applied currently to extract features?
8. What techniques are currently implemented to select the most relevant features?
9. What classifiers are currently applied on DM breast cancer-based ML?
10. What DL techniques are recently implemented for identifying breast cancer based on DM images?
11. How they do their classification as benign/malignant, normal/abnormal, benign/malignant/normal, or Breast Imaging-Reporting and Data System (BI-RADS)?
12. What are the evaluation measurements used for the evaluation of the mammogram imaging-based breast cancer CAD systems?

IEEE Xplore, Science Direct (Elsevier), Springer, and other databases were searched. Furthermore, these keywords and sentences were used:

*mammogram breast cancer, mammogram classification, computer-aided diagnosis using mammogram, computer-aided detection using mammogram, CAD-based on mammogram, mammogram pre-processing for breast cancer, breast cancer segmentation using mammogram, breast cancer classification using mammogram, feature extraction technique for mammogram breast cancer, and feature selection technique for mammogram breast cancer.*

Table 1 illustrates the number of articles published in each venue. All publications of this work were investigated and included in (Table) (5 -10) through the years from 2018 to 2020. Only works that have fulfilled the
following inclusion requirements are included: (1) Only breast cancer disease is included; (2) at least one CAD phase is considered; (3) utilized at least one method-based ML or DL as a classifier; (4) only DM modality is utilized; (5) the most popular performance measurement of the performed classifiers is presented; (6) only full published papers are included; (7) published papers between 2018 and 2020 with only one paper in 2021 are included. We excluded non-English papers, surveys, and books. At first, we retrieved 260 research papers, afterward, papers that irrelevant to the inclusion search criteria have been eliminated. Thus, this research includes only 118 papers (44.86%) whereas the rest of 145 papers are not well fitted for the quest criteria, then these papers have been excluded. The flow chart of the publication retrieval process is shown in Figure 1.

Table 1. Published articles per year and journals.

| Year | Journal          | Publications | Year | Journal          | Publications |
|------|------------------|--------------|------|------------------|--------------|
| 2018 | IEEE             | 15           | 2019 | Springer Link    | 20           |
|      | Science Direct   | 12           |      | Others           | 1            |
|      | Springer Link    | 5            | 2020 | Others           | 1            |
|      | Others           | 8            |      | IEEE             | 8            |
| 2019 | IEEE             | 11           |      | Science Direct   | 10           |
|      | Science Direct   | 11           |      | Others           | 7            |

Figure 1. Flow chart-based summarization of publications selection process.
In this systematic review, more than hundreds of publications are reviewed from indexed and referred journals, conference proceedings and papers from main scientific databases such as IEEE Xplore, Web of Science, and Scopus previously mentioned. Scientific literature on mammographic image analysis contains informative and comprehensive studies. This review has been performed based on 12 main question that have been answered during the review process. We provide a survey-based CAD including pre-processing, segmentation, feature extraction and selection, and classification stages using both machine learning and deep learning techniques. Scope and algorithm of each stage has been presented with it is results. In feature extraction the type of extracted feature as well as the technique that has been used in feature extraction have been presented. Moreover, this systematic review presents the classification method, classifying classes and results are addressed. We also provided the contribution of each surveyed paper with used dataset and number of images in evaluation. (Sadoughi et al. 2018) artificial intelligence methods have been used to identify breast cancer utilizing a wide range of image processing methods. The paper provides relevant information, such as references, techniques used, work scopes, datasets, and various performance metrics, for a more comparative analysis between studies. (Oza et al., 2021) discussed about how to identify and classify suspect areas in mammograms using low-level image features, ML algorithms, and DL techniques from the literature utilizing various methods. Bottom-up survey will cover both low-level image analysis and artificial intelligence methods. Readers will be provided with everything they require to get started working on this topic right away after reading this paper. This review has been presented based on four main question including techniques to extract low-level features, machine learning methods used in identifying mistrustful region, deep learning methods in identifying and classifying breast cancer, and public database used in the evaluation of each work. (Jiménez-Gaona, Rodríguez-Álvarez, and Lakshminarayanan 2020) this paper conducts a crucial survey of the existing literature on the use of ultrasound and mammography images in breast tumor diagnosis using DL algorithms. CAD systems, which are using new DL methods to realize breast images automatically and improve the accuracy of radiologists’ diagnoses, are also summarized. Two hundred and fifty research articles were obtained for this review, of which 59 were eligible for further examination after an eligibility process between 2010 and January 2020.

**CAD Method**

Generally, a standard CAD system covers operations encompassing segmenting structures, detecting abnormalities, and extracting characteristics of abnormalities towards classifying the problem. Figure 2 demonstrates algorithms that are commonly implemented in CAD systems (Memon et al., 2021).
Several phases in the block diagram include acquisition of the image, pre-processing, segmenting, extracting features, classifying, and evaluating. During pre-processing, the filter is applied to the image followed by a transformation towards improving the quality of mammograms and reducing noise level. Meanwhile, in segmenting, region-of-interest are separated from the background (Liu et al. 2020). In extracting features, lesions and normal breast tissue that are represented by certain features are taken for evaluations. While classifying step categories extracted features into classes of malignant and benign features. Finally, an algorithm that is proposed will be used to evaluate the classified features exploiting relevant methodologies. The evaluation step is critical as human lives and their well-being highly depend on the results of the assessments (Xi, Shu, and Goubran 2018) (Sajeev, Baijer, and Lee 2018). As such, any evaluation algorithms for CAD systems must consider sensitivity, specificity, and evaluation of positive predictions. Table 2 represents the recent major contributions of various CAD algorithms in the diagnosis of breast cancer infection. It has been illustrated from this systematic review that the proposed works was based on MLTs and DLTs.

Table 3 demonstrates that this research field has provided several widely published articles during the last two decades. An increase in scientific publications can be due to the improved ability of machines, developed methods of extracting features from images to perform image classification, and available more datasets being used in the research. This research performed the search in 2018 to 2020. Forty papers – 34.7% were published in 2018, 42 papers – 36.5% were published in 2019, and 32 papers – 27.8% were published in 2020. The research publication number was approximately the same in 2018 and 2019, whereas it has been decreasing slightly in 2020. Moreover, the most widely utilized datasets of the studies were shown in Table 3. The two most utilized were the Mammographic Image Analysis Society (MIAS) dataset that was used in 68 studies (59.13%) that 39 studies used only MIAS whereas 29 studies used MIAS with another dataset. The Digital Database for Screening
Table 2. Contributions of recent CAD approaches based on DM Images.

| Ref.                          | Contribution                                                                 | Ref.                           | Contribution                                                                 |
|-------------------------------|------------------------------------------------------------------------------|--------------------------------|------------------------------------------------------------------------------|
| (Singh, Singh, and Bhatia 2018) | Retrieval of DM content utilizing wavelet CS-LBP and SOM                     | (Melekooodappattu et al., 2018) | ELM-FOA based on multi-scale features                                        |
| (Berbar, 2018)                | Three feature extraction methods based on WT Build IMEM enhancement technique | (Soulami et al. 2019)          | ELM to extract suspicious region                                              |
| (Suresh, Rao, and Reddy 2018) | CAD based on WBCT + GA-SVM-Mi + SVM                                          | (Matos et al., 2019)           | BoF model based on SIFT, SURF, ORB, and LBP                                   |
| (Salama, Eltass, and Elkamchouchi 2018) | CAD based on CNN                                                                | (Mohanty et al., 2019)          | An efficient hybrid IGWO-ELM for classifying DMs                               |
| (Charan, Khan, and Khurshid 2018) | CAD based on histogram thresholding in segmentation                          | (Yu et al., 2019)              | Binary classification methods based on modifying DNN                           |
| (Dallali et al. 2018)         | CAD based on LTEM based on energy map for feature calculation                | (Liu et al., 2018)             | Method called TWSVML21 for feature selection                                   |
| (Tatikonda, Bhuma, and Samayamantula 2018) | DCT and DFT transformation method with GLCM used in CAD                    | (Gherghout, Tili, and Souici 2019) | Classification method based on neural and queuing network                     |
| (Routray et al., 2018)        | Hybrid system based on labelling and adaptive RG                             | (Gewei et al., 2019)           | An improved non-parametric method-based pixel intensity for BC-detection      |
| (Shastri, Tamrakar, and Ahuja 2018) | DP-HOT and DP-PB-DCT feature descriptors                                    | (Wang et al., 2019)            | CAD based on several deep feature fusion                                      |
| (Goudarzi et al., 2018)       | Integrated CAD method and FrCN for segmentation                              | (Khan et al. 2019)             | MVFF for CAD system based on CNN                                             |
| (Sadad et al., 2018)          | FC-MRG segmentation method and hybrid LBP-GLCM-LPQ feature extraction method | (Al-Salman et al., 2019)       | Improved U-Net architecture to segment masses                                 |
| (AI-Masni et al., 2018)       | Multi-view CNN is Proposed for feature extraction from CC and MLO views     | (Li et al., 2019)              | Multi-view CNN is Proposed for feature extraction from CC and MLO views       |
| (Al-Antari et al., 2018)      | Integrated CAD method and FrCN for segmentation                              | (Sun et al. 2019)              | Multi-view CNN is Proposed for feature extraction from CC and MLO views       |
| (Sadad et al., 2018)          | Hybrid system based on labelling and adaptive RG                             | (Karthiga, Narasimhan, and Usha 2019) | Multi-view CNN is Proposed for feature extraction from CC and MLO views       |
| (Al-Masni et al., 2018)       | CAD method based on YOLO                                                     | (Khan et al. 2019)             | MVFF for CAD system based on CNN                                             |
| (Uthoff et al., 2018)         | Enhanced contourlet and FOA                                                  | (Mohanty et al., 2019)         | WT based enhancement model for segmentation and feature extraction             |
| (Mohanty et al., 2018)        | PM-segmentation based on histogram operation and K-means                    | (Rampunm et al., 2019)         | Contour-based CNN relies on HED for PM-segmentation                            |
| (Mohanty et al., 2019)        | PM-segmentation based on histogram operation and K-means                    | (Rampunm et al., 2019)         | Contour-based CNN relies on HED for PM-segmentation                            |
| (Yousefi et al., 2018)        | PM-segmentation based on histogram operation and K-means                    | (Rampunm et al., 2019)         | Contour-based CNN relies on HED for PM-segmentation                            |

(Continued)
| Ref.                          | Contribution                                                                 | Ref.                  | Contribution                                                                 |
|------------------------------|------------------------------------------------------------------------------|-----------------------|------------------------------------------------------------------------------|
| (Nedra et al., 2018)         | CAD based on robust feature SURF efficient with K-means segmentation         | (Shi et al., 2018)    | Background and PM-segmentation based on pixel clustering                      |
| (Jiao et al., 2018)          | Improved CNN based on parasitic metric learning net                          | (Shinde et al., 2019) | PM-SD based on RG, thresholding, and K-means                                  |
| (Prathíbha et al., 2018)     | CAD utilizing different class classification of CNN                           | (Li et al., 2019)     | Improved DenseNet and inventing a new DenseNet-II neural network models based on DL |
| (Yaşar, Kutbay, and Härdaç, 2018) | CAD based on ANN and complex WT combination                                   | (Tariq et al., 2019)  | CAD based on Global Discriminate Features                                      |
| (Hazanica et al., 2018)      | Method based on seeded RG segmentation                                        | (Rampun et al., 2020) | LSP based on variant LBP, LTP, and LQP                                         |
| (Shen et al., 2018)          | A segmentation method based on GA morphological selection                    | (Mohanty et al., 2020) | KELM is Proposed based WC-SSA for DMs classification                           |
| (Mughal et al., 2018)        | Discrete differentiation operator-based method to segmentation PM             | (Al-Antari, Han, and Kim, 2020) | An integrated CAD based on DL for BC classification                           |
| (Taz et al., 2018)           | A SSEM method to segment PM                                                   | (Muduli et al., 2020)  | MFO-ELM for automated BC detection                                            |
| (Esener et al., 2019)        | RG method with line fitting to detect PM                                      | (Li et al., 2020)     | Bilateral mass detection based on registration and Siamese-Faster-RCNN networks |
| (Mohamed et al., 2018)       | A deep learning-based approach using CNN to build a computerized reader for BI-RADS-based breast density categorization | (Zhang et al., 2020)  | A method called DE-Ada* for mass detection                                    |
| (Xu et al., 2018)            | Residual based CNN method                                                     | (Shen et al., 2020)   | An optimized classifying method based on DBN and SFO                         |
| (Zhu et al., 2018)           | End-to-end system for mass segmentation using FCN+CRF                       | (Christopher et al., 2020) | An enhancement method called NLUMLOGMIN                                       |
| (Ribli et al., 2018)         | Faster R-CNN method                                                          | (Anora, RAI, and Raman, 2020) | DL-method based ensemble transfer learning and NN                             |
| (Singh, Singh, and Bhatia, 2018) | cGAN for mass segmentation                                                  | (Nayak et al., 2020)  | SRVFLE-AE for multi-class abnormalities detection                             |
| (Gao et al., 2018)           | CAD based Shallow-Deep CNN (SD-CNN)                                         | (Rabidas et al., 2020) | A new local texture feature called LPA                                        |
| (Hagos, Mérida, and Teuwen, 2018) | Multi Path input based on CNN for segmentation                               | (Peng et al., 2020)   | A mass detection-method based on DCN and NAS-FPN                              |
| (Teuwen et al., 2018)        | A method of integration-based candidate detector and the classification      | (Patil and Biradar, 2020) | Optimized RG-segmentation based on FC-CSO for BC detection                   |
| (Jung et al., 2018)          | A new mass detection model based on RetinaNet                                | (Zeiser et al., 2020) | CAD based on U-Net for mass segmentation                                      |
| (Samala et al.)              | CNN based on learning multi-stage transfer                                   | (Fanizzi et al., 2020) | CAD based on multiscale texture features                                      |
| (Wang et al., 2018)          | Integration of clinic, CC, MLO view features using DL                        | (Indra et al., 2020)   | MTF based feature extraction and UTVO based classifier                        |
| (Mughal, Muhammad, and Sharif, 2019) | Proposed curve stitching based on thresholding method for PM segmentation | (Ahmed et al., 2020) | Semantic segmentation and classification based on mask RCNN                   |
| (Mohanty et al., 2019)       | Proposed different methods based on 2D-BDWT-GLCM and FOA                     | (Boudraa, Melouah, and Merouani, 2020) | Improved a CAD based super-resolution reconstruction                         |

(Continued)
| Ref.                        | Contribution                                           | Ref.                        | Contribution                                           |
|----------------------------|--------------------------------------------------------|----------------------------|--------------------------------------------------------|
| (Wang et al., 2019)        | MNPNet method for mass segmentation                     | (Agnes et al., 2020)       | Deep MA-CNN for DMs classification                      |
| (Shen et al., 2019)        | cGNA method for mass segmentation                       | (Tavakoli et al., 2019)    | CNN based deep feature for abnormalities detection      |
| (Chen et al., 2019)        | CAD-based feature pool containing various sets of features | (Song, Li, and Wang, 2020) | Different models based on DCNN and XGBoost             |
| (Ting, Tan, and Sim, 2019) | Design effective CNNI-BCC for localize and classification | (Cheng et al., 2020)       | An automated method called SERAN for mass segmentation |
| (Kaur, Singh, and Kaur, 2019) | Automated method based on K-means clustering with MSVM | (Loizidou et al., 2020)    | A method based on temporal subtraction for MCs detection |
| (Das and Das, 2019)        | KFCM based on entropy and brightness features           | (Chen, Wang, and Chen, 2020) | A multi-scale adversarial network based on improved U-Net for mass segmentation |
| (Mabrouk, Aify, and Marzouk, 2019) | CAD using different features for MCCs detection | (Shen et al., 2019)        | A segmentation mixed-supervision guided and ResCU-Net method |
| (Yin et al., 2019)         | PM-segmentation based on improved iterative threshold    | (Shu et al., 2020)         | Pooling structures of CNN for DMs classifying          |
| (Pavan et al., 2019)       | PM-segmentation based on seed AC and Hough transform    | (Zebari et al., 2020)      | A method based an improved threshold and build ML system for background and PM segmentation |
| (Rahimeto et al., 2019)    | Automated PM removal based on CCL method                | (Suganthi et al., 2020)    | A simple AC and intensity-based thresholding method is introduced for PM segmentation |
| (Gong et al., 2019)        | A threshold method based on mixed features for PM-     | (Ali et al., 2020)         | Improved deep Res U-Net for PM segmentation             |
|                            | segmentation                                           |                            |                                                        |
| (Gu et al., 2019)          | A mass segmentation-method based on super pixel         | (Farhan et al., 2020)      | CAD based on analyzing texture features                |
|                            | generation and curve evolution methods                  |                            |                                                        |
| (Yu et al., 2019)          | DL-method based on DenseNet201                          | (Albalawi et al., 2020)    | CNN for DMs classification                             |
| (Zhang and Wang, 2019)     | FCE-Net and MFF method for MC detection                 | (Soleimani and Michaelovich, 2020) | PM-segmentation method based on CNN                        |
| Shayma’a, 2019             | A mass detection-method based on MSER and feature       | (Zhang et al., 2021)       | BDR-CNN-GCN method based on DL for classification      |
|                            | matching                                               |                            |                                                        |
| (Pezeshki et al., 2019)    | Spiculated parts extraction method based on FCM-GA      | (Saffari et al., 2020)     | Proposed CNN method classifies breast cancer into BIRADS categories |
| (Li, Mukundan, and Boyd, 2021) | Proposed ML-LBP feature extraction                      | (Boumaraf et al., 2020)    | Proposed a CAD method to classify DM massed into different BIRADS categories |

BC – Breast Cancer; PM – Pectoral Muscle; MC – Micro Calcification; WT – Wavelet Transform.
| Ref.                                      | Year | Dataset | Total no. of images | Resolution | Ref.                                      | Year | Dataset | Total no. of images | Resolution | Ref.                                      |
|-------------------------------------------|------|---------|---------------------|------------|-------------------------------------------|------|---------|---------------------|------------|-------------------------------------------|
| (Singh, Singh, and Bhatia 2018)           | 2018 | MIAS    | 322                 | 1024*1024  | (Melekoodappattu et al., 2018)            | 2019 | MIAS    | 184                 | N/A        | (Melekoodappattu et al., 2018)            |
| (Berbar, 2018)                            | 2018 | MIAS-BCDR 1024*1024 | 768–291 | 128*128 | (Soulimi et al., 2019)                      | 2019 | MIAS-DDSM | N/A                 | N/A        | N/A | (Soulimi et al., 2019)                      |
| (Suresh, Rao, and Reddy 2018)             | 2018 | MIAS    | 306                 | 1024*1024  | (Junior et al., 2019)                      | 2019 | MIAS-DDSM | 74–621               | N/A        | (Junior et al., 2019)                      |
| (Salama, Eltarrass, and Elkamchouchi 2018) | 2018 | MIAS    | 100                 | 128*128    | (Matos et al., 2019)                       | 2019 | DSDM    | 1155                | N/A        | (Matos et al., 2019)                       |
| (Charan, Khan, and Khurshid 2018)         | 2018 | MIAS    | 322                 | 224*224    | (Mohanty et al., 2019)                      | 2019 | MIAS-DDSM | 314–1500             | 128*128    | (Mohanty et al., 2019)                      |
| (Dallali et al., 2018)                    | 2018 | MIAS    | N/A                 | 1024*1024  | (Yu et al., 2019)                          | 2019 | BCDR    | 406                 | N/A        | (Yu et al., 2019)                          |
| (Tatikonda, Bhuma, and Samayamantula 2018)| 2018 | MIAS    | 126                 | 1024*1024  | (Liu et al., 2018)                         | 2019 | DSDM-INbreast | 466–116-736           | 256*256    | (Liu et al., 2018)                         |
| (Routray et al., 2018)                    | 2018 | MIAS    | 322                 | 1024*1024  | (Gherghout, Tili, and Souici 2019)          | 2019 | MIAS    | 322                 | 1024*1024  | (Gherghout, Tili, and Souici 2019)          |
| (Samant et al., 2018)                     | 2018 | MIAS-DDSM | 150–840 | N/A | (Gewei et al., 2019)                        | 2019 | MIAS    | 156                 | N/A        | (Gewei et al., 2019)                        |
| (Shastri, Tamrakar, and Ahuja 2018)       | 2018 | MIAS-DDSM | 150–2576 | 128*128 | (Wang et al., 2019)                        | 2019 | MIAS    | 400                 | N/A        | (Wang et al., 2019)                        |
| (Goudarzi et al., 2018)                   | 2018 | MIAS    | N/A                 | 1024*1024  | (Khan et al., 2019)                        | 2019 | MIAS-DDSM | 322–2620              | 128*128    | (Khan et al., 2019)                        |
| (Sadad et al., 2018)                      | 2018 | MIAS-DDSM | 360–100 multi     | 250*250    | (Sun et al., 2019)                         | 2019 | MIAS-DDSM | 400                 | 512*512    | (Sun et al., 2019)                         |
| (Al-Masni et al., 2018)                   | 2018 | DDSM    | 600                 | 256*256    | (Li et al., 2019)                          | 2019 | DDSM    | 322                 | 1024*1024  | (Li et al., 2019)                          |
| (Al-Antari et al., 2018)                  | 2018 | INbreast | 112                 | 448*448    | (Karthiga, Narasimhan, and Usha 2019)       | 2019 | MIAS    | 322                 | 1024*1024  | (Karthiga, Narasimhan, and Usha 2019)       |
| (Sadad et al., 2018)                      | 2018 | MIAS-DDSM | 109–72 N/A        | 256*256    | (Zebari et al., 2019)                      | 2019 | MIAS    | 200                 | 1024*1024  | (Zebari et al., 2019)                      |
| (Uthoff et al., 2018)                     | 2018 | DDSM    | 287                 | N/A        | (Iskandar et al., 2019)                    | 2019 | MIAS    | 322                 | 1024*1024  | (Iskandar et al., 2019)                    |
| (Hussain et al., 2018)                    | 2018 | DDSM    | 899                 | N/A        | (Mohamed El et al., 2019)                   | 2019 | MIAS    | 1024*1024             | N/A        | (Mohamed El et al., 2019)                   |
| (Mohamed et al., 2018)                    | 2018 | DDSM    | 250                 | N/A        | (Rahmatika, Handayani, and Setiawan 2019)   | 2019 | MIAS    | 25                  | 1024*1024  | (Rahmatika, Handayani, and Setiawan 2019)   |
| (Mohanty et al., 2019)                    | 2019 | MIAS-DDSM | 314–1500            | 256*256    | (Rampun et al., 2019)                      | 2019 | MIAS-DDSM | 322–100-all              | 256*256    | (Rampun et al., 2019)                      |
| (Yousefi et al., 2019)                    | 2018 | DBT     | 87                  | 1270–2304  | (Rampun et al., 2019)                      | 2019 | MIAS-DDSM | 322–200-all              | 256*256    | (Rampun et al., 2019)                      |
| (Nedra et al., 2018)                      | 2018 | N/A     | N/A                 | N/A        | (Shinde et al., 2019)                      | 2019 | MIAS    | 322                 | 1024*1024  | (Shinde et al., 2019)                      |
| (Prathibha et al., 2018)                  | 2018 | DDSM-MIAS | 600–322 N/A        | 128*128    | (Li et al., 2019)                          | 2019 | FFDM    | 2042                | 1024*1024  | (Li et al., 2019)                          |
| (Yasar, Kutbay, and Hardalac 2018)        | 2018 | DDSM    | N/A                 | 256*256    | (Tantiq et al., 2019)                      | 2019 | MIAS    | 322                 | 1024*1024  | (Tantiq et al., 2019)                      |
| (Hazarika et al., 2018)                   | 2018 | MIAS    | 150                 | 1024*1024  | (Rampun et al., 2020)                      | 2020 | MIAS-INbreast | 272–206              | 1024*1024  | (Rampun et al., 2020)                      |
| (Shen et al., 2018)                       | 2018 | MIAS-DDSM | 322–128-201 multi | 1024*1024  | (Mohanty et al., 2020)                     | 2020 | MIAS-DDSM-BCDR | 314–1500-736           | 128*128    | (Mohanty et al., 2020)                     |
| (Mughal et al., 2018)                     | 2018 | MIAS-DDSM | 322–128-201 multi | N/A        | (Al-Antari, Han, and Kim 2020)              | 2020 | MIAS-DDSM-INbreast | 600–112               | Multi      | (Al-Antari, Han, and Kim 2020)              |

(Continued)
| Ref. | Year | Dataset       | Total no. of images | Resolution | Ref.                     | Year     | Dataset       | Total no. of images | Resolution |
|------|------|---------------|---------------------|------------|-------------------------|----------|---------------|---------------------|------------|
| (Töz et al., 2018) | 2018 | INbreast      | 60                  | 512*512    | Muduli et al., 2020    | 2020     | MIAS-DDSM     | 326–1500            | 1024*1024  |
| (Esener et al., 2019) | 2018 | MIA S        | 322                 | 256*256    | Li et al., 2020        | 2020     | INbreast-TXMD  | 88–2089–200         | Multi      |
| (Mohamed et al. 2018) | 2018 | Private      | 22,000              | 227*227    | Zhang et al., 2020     | 2020     | DDDSM-INbreast | 2781–387            | Multi      |
| (Xu et al. 2018) | 2018 | INbreast      | 410                 | 224*224    | Shen et al., 2020      | 2020     | MIAS          | 322                 | 1024*1024  |
| (Zhu et al. 2018) | 2018 | INbreast-DDSM | 116–158             | N/A        | Christopher et al., 2020 | 2020     | MIAS          | N/A                 | 1024*1024  |
| (Ribli et al. 2018) | 2018 | INbreast      | 115                 | N/A        | Arora, Rai, and Raman 2020 | 2020     | DDDSM         | 1318                | 256*256    |
| (Singh, Singh, and Bhatia 2018) | 2018 | DDDSM-private | 567–194             | N/A        | Nayak et al. 2020      | 2020     | MIAS          | 322                 | 256*256    |
| (Gao et al., 2018) | 2018 | Private- INbreast | 49–89              | 224*224    | Rabidas et al., 2020   | 2020     | DDDSM         | 58–500              | 1024*1024  |
| (Hagos, Mérida, and Teuwen 2018) | 2018 | Private    | 28,294              | 300*300    | Peng et al., 2020      | 2020     | DDDSM-INbreast | 1592–410            | 133*800    |
| (Jung et al., 2018) | 2018 | Private      | 23,405              | N/A        | Patil and Biradar 2020 | 2020     | MIAS          | N/A                 | N/A        |
| (Teuwen et al. 2018) | 2018 | INbreast – private | 410–222             | N/A        | Zeiser et al., 2020    | 2020     | DDDSM         | 7989                | 1024*1024  |
| (Samala et al.) | 2018 | Private – DDDSM | 4039                | 128*128    | Fanizzi et al., 2020   | 2020     | BCDR          | 260                 | N/A        |
| (Wang et al., 2018) | 2018 | BCDR         | 763                 | r*r        | Indra et al., 2020     | 2020     | DDDSM         | 126                 | N/A        |
| (Mughal, Muhammad, and Sharif 2019) | 2019 | MIAS-DDSM    | 322–400             | multi      | Ahmed et al., 2020     | 2020     | MIAS-DDSM     | 322–2620            | 512*512    |
| (Mohanty et al., 2019) | 2020 | MIAS-DDSM    | N/A                 | 128*128    | Boudraa, Melouah, and Merouani 2020 | 2020     | MIAS          | 93                  | 1024*1024  |
| (Wang et al., 2019) | 2019 | INbreast-DDSM | 410–316             | 256*256    | Agnes et al., 2020     | 2020     | MIAS          | 322                 | 192*192    |
| (Shen et al., 2019) | 2019 | INbreast-Private | 112–376             | 256*256    | Tavakoli et al. 2019   | 2020     | MIAS          | 322                 | 1024*1024  |
| (Chen et al., 2019) | 2019 | N/A          | 275                 | N/A        | Song, Li, and Wang 2020 | 2020     | DDDSM         | 1152                | N/A        |
| (Ting, Tan, and Sim 2019) | 2019 | MIAS        | 221                 | 1024*1024  | Cheng et al., 2020     | 2020     | DDDSM         | 400                 | 224*224    |
| (Kaur, Singh, and Kaur 2019) | 2019 | MIAS        | 322                 | N/A        | Loizidou et al. 2020   | 2020     | N/A          | 320                 | 4096*328   |
| (Das and Das 2019) | 2019 | MIAS        | N/A                 | N/A        | Chen, Wang, and Chen 2020 | 2020     | DDDSM-INbreast | 762–107             | 256*256    |
| (Mabrouk, Affy, and Marzouk 2019) | 2019 | MIAS        | 181                 | N/A        | Shen et al., 2019      | 2020     | INbreast      | 112                 | 256*256    |
| (Yin et al., 2019) | 2019 | N/A          | 720                 | N/A        | Shu et al., 2020       | 2020     | CBIS-DDSM-INbreast | 3071–410           | 800*800    |
| (Pavan et al. 2019) | 2019 | Unspecified  | 30                  | N/A        | Zebari et al., 2020    | 2020     | MIAS-BCDR-INbreast | 322–100–200          | 1024*1024  |
| (Rahimeto et al. 2019) | 2019 | MIAS-BGH    | N/A                 | 512*512    | Suganthi et al., 2020  | 2020     | MIAS          | 322                 | 1024*1024  |
| (Gong et al. 2019) | 2019 | MIAS-DDSM    | 209–240–128         | Multi      | Ali et al., 2020       | 2020     | MIAS-INbreast | 322–416             | N/A        |
| Ref.                        | Year | Dataset          | Total no. of images | Resolution       | Ref.                        | Year | Dataset          | Total no. of images | Resolution       |
|-----------------------------|------|------------------|---------------------|------------------|-----------------------------|------|------------------|---------------------|------------------|
| (Gu et al. 2019)            | 2019 | MIAS             | 322                 | 1024*1024        | (Farhan et al., 2020)      | 2020 | MIAS             | 322                 | 30*30            |
| (Yu et al., 2019)           | 2019 | MIAS             | 322                 | N/A              | (Albalawi et al., 2020)    | 2020 | MIAS             | 322                 | 11,024*1024      |
| (Zhang and Wang 2019)       | 2019 | MIAS             | 322                 | 512*512          | (Soleimani and Michailovich, 2020) | 2020 | MIAS-INbreast    | 322–208            | Multi            |
| Shayma’a, 2019              | 2019 | MIAS-DDSM        | 85–100              | Multi            | (Zhang et al. 2021)        | 2021 | MIAS             | 322                 | 1024*1024        |
| (Pezeshki et al., 2019)     | 2019 | MIAS-DDSM        | 58–200              | 1024*1024        | (Saffari et al. 2020)      | 2020 | INBreast         | 410                 | 256*256          |
| (Li, Mukundan, and Boyd 2021)| 2021 | INBreast         | 409                 | 1024*1024        | (Boumaraf et al. 2020)     | 2020 | DDSM             | 500                 | N/A              |
**Table 4. Summary of surveyed studies based pre-processing phase in the literature.**

| Ref. | Technique | Scope | Evaluation results (%) |
|------|-----------|-------|-------------------------|
| (Singh, Singh, and Bhatia 2018) | CCL, morphological operations, adaptive K-means, adaptive median filter | Labels and artifact suppression, Pectoral muscle removal, and noise reduction | PT 10 |
| (Berbar., 2018) | Contrast stretching | Contrast ROI | PT 10 |
| (Suresh, Rao, and Reddy 2018) | Intelligibility mammogram enhancement method | Image enhancement | PSNR = 9.103MSE = 8.54 |
| (Salama, Eltrass, and Elkanachouchi 2018) | 2-D median filter and connected component labelling | Remove noise and artifact sources also to suppress the pectoral muscle | PT 10 |
| (Dallali et al. 2018) | Histogram equalization and adaptive limited contrast and thresholding | Contrast enhancement and pectoral muscle removal | MSE = 0.125374 RMSE = 0.354082 PSNR = 4.508960 SSIM = 0.206977 |
| (Tatikonda, Bhumia, and Samayamantula 2018) | Median filter and CLAHE | Improve image quality | PT 10 |
| (Shastri, Tamrakar, and Ahuja 2018) | Normalization and TS-CLAHE | Image enhancement | PT 10 |
| (Goudarzi et al., 2018) | Thresholding, shrinkwrap, wavelet, HE, and median filter | Pectoral muscle, labels, and noise removal also improving the image contrast | PT 10 |
| (Yousefi et al., 2018) | Nonlinear Anscombe transformation, adaptive wiener filter, Hough transform | Noise and pectoral muscle removal also | PT 10 |
| (Esener et al., 2019) | Median filter | Noise reduction | PT 8 |
| (Mughal, Muhammad, and Sharif 2019) | optimized Bayesian non-local means filter (OBNLM) | Unwanted noise removal | Sn = 96.6, Sn = 96.4 |
| (Kaur, Singh, and Kaur 2019) | Median filter and morphological operations | Reduce image redundancy to enhance the fineness of mammogram image | PT 10 |
| (Mabrouk, Affy, and Marzouk 2019) | Full-Scale Histogram Stretching (FSHS), Histogram Equalization (HE), Morphological, WT | | PT 10 |
| (Rahimeto et al. 2019) | Wiener filter and Otsu’s thresholding | Noise and tags removal | PT 8 |
| (Gong et al. 2019) | Otsu threshold | Remove unwanted area | PT 10 |
| (Yu et al., 2019) | Morphological operations and threshold | Remove noise and region of breast extraction | PT 10 |
| Shayma’a, 2019 | Median filter and SEBHE | Noise removal and image enhancement | PT 10 |
| (Pezeshki et al., 2019) | CLAHE | enhance the significant features of the mass smoothing, sharpening, noise removal and edge detection | PT 10 |
| (Melekkodappattu et al., 2018) | Wiener filter, CLAHE | | PT 10 |
| (Soulami et al. 2019) | low threshold, labelling, 2D – median filter | Artifacts, background, and noise removal | PT 10 |
| (Matos et al., 2019) | Logarithmic transformation anisotropic diffusion filter | Image enhancement | PT 10 |
| (Gherghout, Tilli, and Souici 2019) | Adaptive mean filter, algorithm | Reduce noise and preserve edges | PT 10 |
| (Wang et al., 2019) | Top-hat and bottom-hat transforms, morphological and curvelet transforms | avoid the impact of noise, enhancement | PT 10 |
| (Karthiga, Narasimhan, and Usha 2019) | Thresholding and, Weiner filter and CLAHE filters | Contrast enhancement, sharpen and wrapping | PT 10 |
| (AlsAlman et al., 2019) | | | PT 10 |

*(Continued)*
Table 4. (Continued).

| Ref. | Technique | Scope | Evaluation results (%) |
|------|-----------|-------|-------------------------|
| (Dabass et al., 2019) | CLAHE and entropy based Intuitionistic Fuzzy Method | Contrast enhancement | Entropy = 4.8868 PSNR = 22.4759 |
| (El-Sokkary et al., 2019) | Threshold and method of (Memon et al., 2021) | Artifact and pectoral muscle removal | PT 10 |
| (Rahmatika, Handayani, and Setiawan 2019) | Median filter | Noise reduction | PT 8 |
| (Zebari et al. 2019) | Wavelet transform | Image enhancement for segmentation and feature extraction | Segmentation: Ac = 90.5 Feature extraction: PSNR = 69.95 |
| (Rampun et al. 2020) | Active counter, restricted contour growing incorporating edge information, and median filter multi-threshold peripheral equalization | Breast region segmentation and Noise reduction | PT 10 |
| (Al-Antari, Han, and Kim 2020) | Histogram equalization | enhance the peripheral regions of breast images | PT 10 |
| (Shen et al. 2020) | Wang-Mendel | breast image noise removal | PT 10 |
| (Christopher et al., 2020) | NLUML0GMIN | Mammogram enhancement improve upon the contrast and dynamic range of an image | EME = 3.89, AME = 23.92, SDME = 49.36 |
| (Arora, Rai, and Raman 2020) | Histogram equalization | | PT 10 |
| (Patil and Biradar 2020) | Median filter | Noise elimination | PT 10 |
| (Zeiser et al., 2020) | CLAHE | removal of irrelevant information | PT 8 |
| (Indra et al., 2020) | Adaptive median filter | Remove speckle, salt, and pepper noises | PT 10 |
| (Ahmed et al. 2020) | Binarization, Median filter, HE, morphological operations, savitzky golay filter, masking, and histogram equalization | Artifact, noise, and pectoral muscle removal | PT 8 |
| (Agnes et al. 2020) | Median filter, global thresholding, morphological operations, and single seeded region growing | Noise reduction, background, and pectoral muscle removal | PT 10 |
| (Tavakoli et al. 2019) | Otsu’s thresholding CLAHE | eliminates irrelevant areas from the image and enhances the contrast | PT 10 |
| (Cheng et al. 2020) | Gamma transformation and OTSU | Enhance details of image and extract breast region | PT 10 |
| (Loizidou et al. 2020) | Border removal function, Otsu’s thresholding, Demons | Background and pectoral muscle removal, image registration | PT 10 |
| (Zebari et al. 2020) (Ali et al., 2020) | Wavelet transform Gaussian, median filters, and CLAHE | Highlight breast region Noise reduction and image sharpening | PT 8 PT 8 |
| (Farhan et al., 2020) | CLAHE | Image enhancement | PT 10 |
| (Albalawi et al., 2020) | Wiener filter | Noise reduction | PT 10 |
| (Li, Mukundan, and Boyd 2021) | LBP | ROI | PT 10 |
| (Boumaraf et al. 2020) | Histogram equalization | Image enhancement | PT 10 |

PT 8 means the results are presented in Table 5, and PT 10 means results are presented in Table 7. CLAHE – contrast limited adaptive histogram equalization.
| Ref. | Technique | Scope | Evaluation results (%) |
|------|-----------|-------|------------------------|
| (Singh, Singh, and Bhatia 2018) | Threshold based seeded region growing | ROI extraction | PT 10 |
| (Salama, Eltrass, and Elkamchouchi 2018) | Watershed | ROI extraction | PT 10 |
| (Charan, Khan, and Khurshid 2018) | Morphological closing operation and masking | ROI extraction | PT 10 |
| (Dallali et al. 2018) | Histogram thresholding | Mass detection | None |
| (Samant et al., 2018) | Otsu's thresholding | Remove unwanted labels | PT 10 |
| (Sapate et al. 2018) | Automatic seed selection, adaptive fuzzy region growing, region merging | Identifying the suspicious region | None |
| (Al-Masni et al., 2018) | Deep learning-based YOLO | ROI extraction | PT 10 |
| (Al-Antari et al. 2018) | Deep learning-based YOLO and FrCN | To segment the mass | Ac = 92.97, Sens = 92.72, Sp = 93.21, Dice = 92.69, Jac = 86.37, AUC = 92.97, MCC = 85.93 |
| (Sadad et al. 2018) | Fuzzy C-Means (FCM) and region-growing algorithm called FCMRG | Tumor segmentation | PT 10 |
| (Uthoff et al., 2018) | Otsu's thresholding and region growing | Lesion segmentation | PT 10 |
| (Yousefi et al., 2018) | Level set | ROI selection | PT 10 |
| (Nedra et al., 2018) | K-means | Breast tissues segmentation | PT 10 |
| (Hazarika et al., 2018) | Region growing | Pectoral muscle removal | Ac = 92 |
| (Shen et al. 2018) | Polynomial fitting/Curve Estimation, genetic algorithm, and morphological selection | Pectoral muscle removal | MIAS: FP = 2.03, FN = 6.9, Jac = 91.25, Dice = 94.96 |
| (Mughal et al., 2018) | Convex hull | Pectoral muscle removal | MIAS: FP = 0.99, FN = 5.67, FFDM: FP = 0.98, FN = 5.66, Sn = 95.6 |
| (Tuz et al., 2018) | Geometrical properties | Pectoral muscle segmentation | Ac = 94.4, Sn = 89.62, Sp = 99.99 |
| (Esener et al., 2019) | Region growing | Pectoral muscle segmentation | Inbreast: Dice = 90.97 FDCM: Dice = 91.3 |
| (Zhu et al. 2018) | FCN+ CRF | Lesion segmentation | Inbreast: Dice = 90.97 DDM: Dice = 91.3 |
| (Singh, Singh, and Bhatia 2018) | cGAN-Unet | Lesion segmentation | DDM: Ac = 0.97, dice = 0.94, Jac = 0.89, Sn = 0.92, Sp = 0.98, Private: Ac = 0.95, dice = 0.86, Jac = 0.76, Sn = 0.85, Sp = 0.97 |
| (Mughal, Muhammad, and Sharif 2019) | Curve stitching and adaptive hysteresis thresholding (CSAHT) | Separation of breast region and internal details of breast parenchyma from background | PT 10 |
| (Wang et al., 2019) | MNPNet | Breast mass segmentation | Inbreast: Dice = 91.1, DDM: Dice = 91.69 |

(Continued)
| Ref. | Technique | Scope | Evaluation results (%) |
|------|-----------|-------|-------------------------|
| (Shinde et al., 2019) | Improved U-net by combining conditional generative adversarial network (cGAN) and original dataset | Breast mass segmentation | INbreast: Ac = 92, Sn = 90.57, Sp = 93.09, Jac = 78.47, Dice = 87.58, MCC = 82.14 Private: Ac = 88.82, Sn = 95.61, Sp = 83.41, Jac = 79.35, Dice = 88.2, MCC = 78.8 |
| (Das and Das 2019) | Kernel based fuzzy c-means (FCM) | Detection of masses | |
| (Mabrourk, Affy, and Marzouk 2019) | Local threshold and Otsu method | MCC extraction | PT 10 |
| (Yin et al., 2019) | Active counter | Pectoral muscle removal | Ac = 94.6, Dice = 0.986 Jac = 0.92 |
| (Pavan et al. 2019) | Active counter | Pectoral muscle removal | Ac = 98.62, IoU = 0.8362, RMSE = 0.1033 |
| (Rahimeto et al. 2019) | Otsu’s multi-thresholding technique and CCL | Automatic pectoral muscle removal segment the breast glandular tissue | PT 10 |
| (Gong et al. 2019) | New threshold method | | |
| (Gu et al. 2019) | Superpixel generation based SLIC and DBSCAN, and curve evolution method | Breast mass segmentation | TP = 86.76, FP = 13.24, SI = 86.33, DSC = 90.12 |
| Shayma’a, 2019 | MSER detector-based SURF and features matching | Breast cancer mass detection | DDSM: Ac = 96 MIAS: Ac = 96.47 |
| (Pezeshki et al., 2019) | FCM | Tumor segmentation | PT 10 |
| (Melekoodappattu et al., 2018) | Gray level and global thresholding | Background and pectoral muscle region segmentation | PT 10 |
| (Soulimi et al., 2019) | EML | To segment breast area | PT 10 |
| (Junior et al. 2019) | Combination of MeanShift and Fast Scanning non-parametric method and level-set function detection | Lesion detection | MIAS: Dr = 97.3, FP = 0.89 DDSM: Dr = 91.63, FP = 0.86 Ac = 94.937 |
| (Gherghout, Tili, and Souici 2019) | Adaptive mass region detection | Tumor detection | PT 10 |
| (Wang et al., 2019) | Extract breast mass region | | |
| (Li et al., 2019) | combines densely connected U-Net with attention gates (AGs) | Breast mass segmentation | Ac = 78.38, Sn = 77.89, Sp = 87.62, F1-score = 82.24 |
| (AlSalman et al., 2019) | k-means | Segment ROI | PT 10 |
| (El-Sokkary et al., 2019) | PSO and GMM | ROI segmentation | PT 10 |
| (Rahmatika, Handayani, and Setiawan 2019) | Histogram operation and k-means clustering | Breast tissue segmentation | |
| (Rampunn et al., 2019) | CNN with modified HED | Pectoral muscle segmentation | MIAS: Ac = 99.3, Sn = 98.2, Sp = 99.5, Jac = 94.6, Dice = 97.5; BCDR: Ac = 99.6, Sn = 95.2, Sp = 99.8, Jac = 92.6, Dice = 95.6; INbreast: Ac = 99.9, Sn = 99.6, Sp = 99.6, Jac = 96.9, Dice = 98.8 |
| (Shi et al., 2018) | Pixel-wise clustering | Pectoral muscle segmentation | MIAS: Ac = 97.08, Jac = 94.89, Dice = 96.4; BCDR: Ac = 97.38, Jac = 95.6, Dice = 97.6; INbreast: Ac = 97.91, Jac = 96.22, Dice = 97.66 |
| (Shinde et al., 2019) | Machine learning | Pectoral muscle segmentation | Ac = 93.71 |

(Continued)
| Ref.                          | Technique                                      | Scope                                      | Evaluation results (%)                                                                 |
|------------------------------|-----------------------------------------------|--------------------------------------------|----------------------------------------------------------------------------------------|
| (Al-Antari, Han, and Kim 2020) | Deep learning YOLO                             | detection of suspicious breast lesions     | DDM: Ac = 99.17, MCC = 98.36, Dice = 99.28; INbreast: Ac = 97.27, MCC = 93.93, Dice = 98.02 |
| (Li et al. 2020)             | Combining self-supervised learning network and Siamese-Faster-RCNN | bilateral mass detection                  | INbreast: TP = 0.88, FP = 1.12; BCPKUPH: TP = 0.85, FP = 1.86; TXMD: TP = 0.85, FP = 2.70 |
| (Shen et al. 2020)           | Otsu thresholding and mathematical morphology  | Separate useful part of image              |                                                                                       |
| (Peng et al. 2020)           | Faster R-CNN (DCN C3–C5 NAS-FPN OHEM)         | Mass detection                             | DDM: TPR = 0.9345; INbreast: TPR = 0.9554                                               |
| (Patil and Biradar 2020)      | Optimized region growing based on FC-CSO       | Tumor segmentation                         | Ac = 0.98, Sn = 0.59, Sp = 0.99, Pre = 0.99, F1-score = 0.74, MCC = 0.76                 |
| (Zeiser et al., 2020)        | Data augmentation and U-Net model              | Mass diagnosis                             | Ac = 85.95, Sn = 92.32, Sp = 80.47, Dice = 79.39, AUC = 86.40                           |
| (Indra et al., 2020)         | Multi scale invariant threshold                | detecting cancer                           | PT 10                                                                                   |
| (Ahmed et al. 2020)          | DeepLab RCNN                                   | Mass segmentation                          | Ac = 0.98, pre = 0.8                                                                     |
| (Cheng et al. 2020)          | Spatial Enhanced Rotation Aware Network (SERAN) | Breast mass segmentation                   | Ac = 99.84, Sn = 87.7, Sp = 99.9, IOU = 73.95, Dice = 84.3                             |
| (Chen, Wang, and Chen 2020)  | Improved U-Net                                  | Breast mass segmentation                    | DDM: Ac = 0.9981, Sn = 0.8523, Sp = 0.9986, Dice = 0.8216                              |
| (Shen et al., 2019)          | MS-ResCU-Net                                   | Simultaneous segmentation and classification| Ac = 94.16, Sn = 93.11, Sp = 95.02, Dice = 91.78, Jac = 85.13, MCC = 87.22, AUC = 94.57|
| (Zebari et al. 2020)         | New threshold technique based on texture features Machine learning based on HOG and NN | BS segmentation                            | BS: Ac = 99.31, Sn = 99.54, Sp = 99.41, Jac = 98.67, Dice = 99.14, PM: Ac = 98.64, Sn = 98.25, Sp = 99.63, Jac = 96, Dice = 98.5 |
| (Suganthi et al. 2020)       | Contrast enhancement and intensity-based thresholding | Breast region segmentation                | Ac = 92.55                                                                              |
| (Ali et al., 2020)           | Fully convolutional network                    | Pectoral muscle segmentation               | MIAS: Ac = 96, Dice = 94.5                                                              |
| (Albalawi et al., 2020)      | K-means clustering                             | Mass segmentation                          | INbreast: Ac = 95, Dice = 94 PT 10                                                       |
| (Soleimani and Michailovich, 2020) | CNN                                         | PM segmentation                            | Dice = 97.22, Ac = 99.64                                                                |
| (Saffari et al. 2020)        | conditional Generative Adversarial Networks (cGAN) network  | Breast tissue segmentation                | Ac = 98, Dice = 88, Jaccard index = 78                                                  |
| (Boumaraf et al. 2020)       | Region growing                                | ROI segmentation                           | PT10                                                                                     |
| Ref. | Feature | Feature extraction | Feature selection |
|------|---------|--------------------|-------------------|
| (Singh, Singh, and Bhatia 2018) | Texture | LBP, CS-LBP, WLBP, and WCS-LBP | N/A |
| (Berbar., 2018) | Texture and statistical | ST-GLCM, wavelet-CT1, and wavelet-CT2 | N/A |
| (Salama, Eltrass, and Elkamchouchi 2018) | Texture, Statistical, and shape | GLCM and improved WBCT | GA+SVM+MI and PSO+SVM+MI |
| (Tatikonda, Bhuma, and Samayamantula 2018) | Texture | Combination of HOG and GLCM | N/A |
| (Routray et al., 2018) | Texture | Laws Texture Energy Measure (LTEM) | N/A |
| (Samant et al., 2018) | 22 Texture | GLCM | N/A |
| (Shastri, Tamarakar, and Ahuja 2018) | Texture | Combination of HOG with Gabor filter (HOT) and PB-DCT | N/A |
| (Goudarzi et al., 2018) | Geometric and texture | Compactness, entropy, mean, and smoothness | N/A |
| (Satap et al. 2018) | Geometric and texture | N/A | N/A |
| (Al-Masni et al., 2018) | Deep feature | CNN | N/A |
| (Sadad et al. 2018) | Texture | Hybrid LB-GLCM+LPQ | mRMR |
| (Uthoff et al., 2018) | 13 Histogram, texture, 18 shape | GLRL, GLSZ, NGSTD, LTEM | k-melodies clustering, IO |
| (Hussain et al., 2018) | Texture | Morphological, SIFT, and EFDs | N/A |
| (Mohamed et al. 2018) | 50 texture, shape | GLCM, GLRLM, wavelet | Two sample T-test with PVE |
| (Mohanty et al., 2019) | Texture | Contourlet transform | Forest optimization |
| (Yousef et al., 2018) | Statistical, texture, gray level, morphological | Hand-crafted | N/A |
| (Nedra et al., 2018) | Texture | SURF and BoW | N/A |
| (Mohanty et al., 2019) | 480 Texture | 2D-BDWT and GLCM | PCA + FOA |
| (Chen et al., 2019) | 59 Shape and density | FFT features, DCT features and WT features | PSO |
| (Mabrouk, Affify, and Marzouk 2019) | Shape, texture, and invariant moment | Morphological and GLCM | Fisher score |
| (Gong et al. 2019) | Texture and statistical | GLCM | N/A |
| (Pezeshki et al., 2019) | 34 intensity histograms, texture, margin and shape | FD, GLCM, LBP | GA |
| (Melekooodappattu et al., 2018) | Texture | SURF, Gabor filter, GLCM | GWO |
| (Soula et al., 2019) | Shape | N/A | N/A |
| (Matos et al., 2019) | Texture | SIFT, SURF, ORB, LBP, SIFT+LBP | BOF |
| (Mohanty et al., 2019) | Texture | Discrete Tchebichef transform (DTT) | PCA and LDA |
| (Liu et al., 2018) | Texture and geometric | GLCM | TWSVML21 |
| (Gherghoul, Tili, and Souici 2019) | Texture | GLCM, GLRLM | RELIEF and MMR |
| (Wang et al., 2019) | Deep learning, morphological, texture, and density | CNN, GLCM | N/A |
| (Karthiga, Narasimhan, and Usha 2019) | 14 textural | GLCM | N/A |
| (Al-Salman et al., 2019) | 22 statistical | DWT and GLCM | N/A |
| (El-Sokkary et al., 2019) | Texture and shape | GLCM | N/A |
| (Tariq et al. 2019) | 20 textural | GLCM | N/A |
| (Rampun et al., 2020) | Texture | Local septenary patterns (LSP) | Dominant patterns |
| (Mohanty et al., 2020) | Shannon entropy, Tsallis entropy, Renyi entropy, and energy | Block-based discrete wavelet packet transforms (BDWPT) | Principal component analysis (PCA) |
| (Muduli et al., 2020) | Lifting wavelet transform (LWT) | SIFT, GIST, HOG, LBP, ResNet, DenseNet, and VGG | PCA + LDA |
| (Zhang et al., 2020) | Texture, shape, and deep learning | N/A |
| (Shen et al. 2020) | Statistical textural | DWT + GLCM | N/A |

(Continued)
Table 6. (Continued).

| Ref.                  | Feature                          | Feature extraction               | Feature selection |
|-----------------------|----------------------------------|-----------------------------------|-------------------|
| (Arora, Rai, and Raman 2020) | Deep learning                    | AlexNet, VGG16, GoogLeNet, ResNet18, and Inception ResNet | N/A               |
| (Rabidas et al., 2020)  | Texture                          | Local Photometric Attributes (LPA) | Stepwise logistic regression |
| (Patil and Biradar 2020)  | Texture                          | GLCM and GLRM                    | N/A               |
| (Fanizzi et al., 2020)  | Texture                          | Haar wavelet decompositions       | Embedded and filter |
| (Indra et al., 2020)   | Texture                          | MTF based matrix vectors          | N/A               |
| (Boudraa, Melouah, and Merouani 2020) | Statistical texture               | GLCM, GLRLM                      | N/A               |
| (Tavakoli et al. 2019) | Deep learning                    | CNN                               | N/A               |
| (Song, Li, and Wang 2020) | Deep learning and texture         | DCNN, GLCM, GOG                   | N/A               |
| (Loizidou et al. 2020) | Shape, intensity, texture        | FOS, GLCM                         | t-test, MANOVA    |
| (Farhan et al., 2020)  | Texture                          | LBP, HOG, and GLCM                | N/A               |
| (Albalawi et al., 2020) | Deep                            | CNN                               | N/A               |
| (Li, Mukundan, and Boyd 2021) | texture                         | M-FD + MLBP                       | PCA               |
| (Boumaraf et al. 2020) | handcrafted                      | Shape, density, margin            | GA                |

N/A – not available.

Mammography (DDSM) was cited in 45 papers (40%), whereas 12 studies used only DDSM, and 33 studies used DDSM with another dataset. These databases are most popular not only because they included a large set of images but also because they permitted free usage of such images provided the licenses are respected. For INbreast dataset, 23 studies (20%) used to evaluate their study, where only 5 studies used only INbreast, while 18 studies used INbreast with another dataset. Only eight studies (7%) used Braxet Cancer Digital Repository (BCDR) dataset. Some research used private datasets and databases, such as those supplied by the Alberta Program for Early Detection of Breast Cancer and the database given by the University of Chicago. Private datasets seem to surface less often in the studies relative to public ones, so it is more challenging to get access to them. Only seven publications (6%) utilized 100 or less images in the training phase to perform the testing phase. Moreover, 12 publications (11%) utilized between 101 and 200 images, 44 publications (38.26%) used between 200 and 500 images and 42 papers (36.52%) used 500 or more images in their performance evaluation. Furthermore, 11 publications (10%) did not determine the utilized image number. 68.69% of the publications utilize 200 or more than 200 images.

Generally, the CAD method includes segmented systems, the identification of anomalies, and the extraction of their characteristics for the corresponding classification. CAD systems typically reach four main phases. The first phase of pre-processing involves improving the contrast and tuning out the noise to prepare the dataset images for the following phases through a set of image pre-processing operations as illustrated in Table 4. The second phase is the segmentation allows the system to extract features more easily from ROI as illustrated in Table 5. The third phase is the feature extraction and selection
### Table 7. Summary of surveyed studies for classification based on ML and DL techniques.

| Ref. | ML/DL | Technique | Class | Scope | Evaluation results (%) |
|------|-------|-----------|-------|-------|------------------------|
| (Singh, Singh, and Bhatia 2018) | ML | SOM | N/AB | Content retrieval system | Precision = 79.61 |
| (Berbar., 2018) | ML | SVM | N/AB | Breast mass classification | DDSM: Ac = 98.69, Sn = 98.82, AUC = 0.98637 |
| (Salama, Eltass, and Elkamchouchi 2018) | DL | CNN | N/AB | Early detection of cancer | MIAS: Ac = 97.89, Sn = 96.12, AUC = 0.8769 |
| (Charan, Khan, and Khurshid 2018) | ML | Linear SVM, kernel SVM, and KNN | B/M | Breast cancer diagnosis | Ac = 97.5, Sn = 93.75, Sp = 100 |
| (Tatikonda, Bhuma, and Samayamantula 2018) | ML | IONN, CT, SVM, ANN, LDA, and NB | B/M | Benning and malignant classification | Ac = 99.11, Sn = 98.11, Ap = 100 |
| (Routray et al., 2018) | ML | ANN | B/M/N | Breast cancer detection | Ac = 94.4, Sn = 90.9, Sp = 99.99 |
| (Samant et al., 2018) | ML | SVM and KNN | B/M/N | detect and classify tumours | Ac = 93.89 |
| (Shastry, Tamrakar, and Ahuja 2018) | ML | SVM | N/AB | Breast cancer classification into normal and abnormal also further classification into benign and malignant | MIAS: Ac = 97.33, Sn = 97.18, Sp = 98.11, AUC = 100 |
| (Al-Masni et al., 2018) | ML | Combination of fuzzy, discretization and bagging | N/AB | Breast cancer classification | DDSM: Ac = 84.84, Sn = 87.18, Sp = 97.93, AUC = 92.2 |
| (Goudarzi et al., 2018) | ML | SVM and KNN | B/M | Early detection is the important key to reduce breast cancer mortality rate | MIAS: 100 all metrics |
| (Al-Antari et al., 2018) | DL | FC-NNs | B/M | Classification of breast masses | DDSM: Ac = 93.1, Sn = 95, Sp = 88, F-score = 93, MCC = 85 |
| (Sadad et al. 2018) | ML | DT, LDA, SVM, KNN, Ensemble, Logistic regression | B/M | Classification of breast masses | Ac = 97, Sn = 100, Sp = 94, AUC = 96.45 |

(Continued)
Table 7. (Continued).

| Ref.                        | ML/ DL | Technique                          | Class | Scope                                           | Evaluation results (%) |
|-----------------------------|--------|------------------------------------|-------|------------------------------------------------|------------------------|
| (Uthoff et al., 2018)       | ML     | ANN                                | B/M   | Breast mammogram lesion classification          | Ac = 96.2, Sn = 97.6,  |
|                             |        |                                    |       |                                                 | Sp = 95.2, AUC = 0.971 |
| (Hussain et al., 2018)      | ML     | SVM, DT, Bayesian                   | N/AB  | Automated BC detection                          | Sn = 1, Sp = 1, AUC = 1, |
|                             |        |                                    |       |                                                 | PPV = 1NPV = 1,        |
|                             |        |                                    |       |                                                 | FPR = 0                |
| (Mohamed et al. 2018)       | ML     | ANN, SVM, KNN                       | B/M   | Automatic mass classification                    | Ac = 98.9, Sn = 100, Sp = 97.8 |
| (Mohanty et al, 2019)       | ML     | SVM, KNN, Naïve-Bayes, C4.5         | N/AB  | Classify suspicious regions                     | N/AB for both dataset: Ac = 100, Sn = 1,  |
|                             |        |                                    |       |                                                 | Sp = 1, MCC = 1        |
|                             |        |                                    |       |                                                 | MIAS B/M: Ac = 98.74, Sn = 0.97, Sp = 1,  |
|                             |        |                                    |       |                                                 | MCC = 0.97             |
| (Yousefi et al., 2018)      | DL     | MI-RF                              | B/M   | Mass detection in DBT                           | Ac = 99, Sn = 99.49,  |
|                             |        |                                    |       |                                                 | Sp = 95.53             |
| (Nedra et al., 2018)        | ML     | SVM                                | B/M   | Classification of breast abnormalities          | Ac = 86.81, Sn = 86.6,  |
|                             |        |                                    |       |                                                 | Sp = 87.5, AUC = 0.87  |
| (Jiao et al. 2018)          | DL     | Deep CNN and parasitic metric learning network | B/M   | Distinguish Malign cases from Benign ones        | DDSM: Ac = 97.4        |
|                             |        |                                    |       |                                                 | MIAS: Ac = 96.7        |
|                             |        |                                    |       |                                                 | Ac = 85.4              |
| (Prathibha et al., 2018)    | DL     | CNN-Bandelet, CNN-ORT II, and CNN-Tetrolet | B/M/N | Mammogram classification                        | Ac = 94.79             |
| (Yaşar, Kutbay, and Hardalç 2018) | ML   | ANN+ complex wavelet transforms   | F/D/G | Tissue density classification                   |                        |
| (Mohamed et al. 2018)       | DL     | CNN (AlexNet) (Transfer learning)  | BI-RADS | Breast density estimation                        | AUC = 0.9882           |
| (Xu et al. 2018)            | DL     | CNN                                | BI-RADS | Breast density estimation                        | Ac = 92.63             |
| (Ribli et al., 2018)        | DL     | Faster R-CNN                        |       | Detection and classification                     | AUC = 0.95             |
| (Gao et al., 2018)          | DL     | SD-CNN                             | B/M   | Lesion classification                           | Ac = 09, AUC = 0.92    |
| (Hagos, Mérida, and Teuwen 2018) | DL | Multi-input CNN                   | B/M   | Lesion detection and classification              | AUC = 0.93             |
| (Jung et al., 2018)         | DL     | Fast R-CNN and mask R-CNN with ResNet | B/M   | Lesion detection and classification              | Sn = 0.97, FP = 3.56 per image |
| (Teuwen et al. 2018)        | DL     | RetinaNet                          | B/M   | Mass detection and classification                | Ac = 0.98, FP = 1.3 per image |
| (Samala et al.)             | DL     | Multistage finetuned CNN           | B/M   | Classification performance on varying sample sizes | AUC = 0.91             |
| (Wang et al., 2018)         | DL     | CNN and LTSM                       | B/M   | Classification of breast masses using contextual information | AUC = 0.89             |
| (Mohanty et al, 2019)       | DL     | Deep SRVFL-AE                      | N/AB  | Mass classification                             | Ac = 94.79             |
| (Chen et al. 2019)          | ML     | SVM                                | B/M   | Predict the likelihood of case being malignant  | Sn = 81, Sp = 77        |
| (Ting, Tan, and Sim 2019)   | DL     | CNNI-BCC                           | B/M/N | Improves the breast cancer lesion classification | Ac = 90.5, Sn = 89.47, Sp = 90.71 |

(Continued)
| Ref.                                                                 | ML/DL | Technique                                  | Class  | Scope                                                                 | Evaluation results (%) |
|--------------------------------------------------------------------|-------|--------------------------------------------|--------|----------------------------------------------------------------------|------------------------|
| (Kaur, Singh, and Kaur 2019)                                       | ML    | K-mean clustering + MSVM/decision tree     | B/M/N  | Detection and validation of automated mammogram breast cancer       | Ac = 95.6, Sn = 98.6, Sp = 97.6, ROC = 0.99 |
| (Kaur, Singh, and Kaur 2019)                                       | ML    | AN, KNN, SVM                               | B/M/N  | micro calcifications cancer                                          | Ac = 96, Sn = 98, Sp = 94 |
| (Gong et al. 2019)                                                 | ML    | SVM                                        | N/AB   | Identify microcalcification clusters and detect breast cancer        | Ac = 96, Sn = 98, Sp = 94 |
| (Yu et al., 2019)                                                  | DL    | DenseNet201-c                              | N/AB   | Diagnosis of breast abnormality                                      | Ac = 96, Sn = 98, Sp = 94 |
| (Zhang and Wang 2019)                                              | DL    | FCE-Net + MFF                             | N/AB   | Identification of malignancy and benignity                          | Ac = 96, Sn = 98, Sp = 94 |
| (Pezeshki et al., 2019)                                           | ML    | SVM                                        | B/M    | Tumor classification                                                | Ac = 96, Sn = 98, Sp = 94 |
| (Melekoodappattu et al., 2018)                                     | ML    | ELM-FOA                                    | B/M    | Automatic Micro Calcification Detection                              | Ac = 99.04,            |
| (Soulam et al., 2019)                                              | ML    | SVM                                        | B/M    | Detection abnormalities                                             | Ac = 99.65, Sn = 99.24, Pre = 99.99,  F1-score = 99.99, |
| (Matos et al., 2019)                                               | ML    | SVM, AdaBoost, RF                          | B/M    | discrimination of malignancy and benignity                          | Ac = 99.65, Sn = 99.24, Pre = 99.99,  F1-score = 99.99, |
| (Mohanty et al., 2019)                                             | ML    | LGWO-ELM                                   | N/AB   | Automated and accurate classification                                | Ac = 99.65, Sn = 99.24, Pre = 99.99,  F1-score = 99.99, |
| (Yu et al., 2019)                                                  | DL    | GoogLeNet, AlexNet, CNN2, CNN3, SVM        | B/M    | Differentiation of breast lesions                                    | Ac = 99.65, Sn = 99.24, Pre = 99.99,  F1-score = 99.99, |
| (Liu et al., 2018)                                                 | ML    | TWSVML21                                   | B/M    | Mass classification                                                 | Ac = 99.65, Sn = 99.24, Pre = 99.99,  F1-score = 99.99, |
| (Gherghout, Tili, and Souici 2019)                                 | ML    | BPNN                                       | N/AB   | Classification of breast mass                                       | Ac = 86.5, Sn = 85.1, Sp = 88.02, AUC = 0.923 |
| (Khan et al. 2019)                                                 | DL    | ELM, SVM                                   | B/M    | Breast cancer classification                                         | Ac = 86.5, Sn = 85.1, Sp = 88.02, AUC = 0.923 |
| (Khan et al. 2019)                                                 | DL    | MVFF based CNN                             | N/AB   | Mammogram classification                                             | Ac = 93.73, Sn = 96.31, Sp = 90.47, AUC = 0.934 |
| (Sun et al. 2019)                                                  | DL    | MVMDCNN-LOSS                               | B/M    | Mammogram classification                                             | Ac = 77.66, Sn = 81.82, Sp = 72.02, AUC = 0.769 |

(Continued)
| Ref.                              | ML/DL | Technique                        | Class | Scope                             | Evaluation results (%) |
|----------------------------------|-------|----------------------------------|-------|-----------------------------------|------------------------|
| (Karthiga, Narasimhan, and Usha 2019) | ML    | SVM                              | N/AB  | Breast cancer diagnosis          | Ac = 98                |
| (AlSalman et al., 2019)          | ML    | ANN                              | B/M/N | Breast cancer diagnosis          | Ac = 96.5625          |
| (El-Sokkary et al., 2019)        | ML    | Non-linear SVM                   | B/M/N | Breast cancer detection and diagnosis | MIAS: Ac = 100, Sn = 1, Sp = 1, MCC = 1, F-score = 1, AUC = 1 |
| (Mohanty et al, 2019)            | ML    | SVM, KNN, C4.5                   | N/AB  | DM identification                | MIAS: Ac = 100, Sn = 1, Sp = 1, MCC = 1, F-score = 1, AUC = 1 |
| (Li et al., 2019)                | DL    | DenseNet neural network          | B/M   | DM classification                 | DDSM: same results as MIAS with SVM and KNN |
| (Tariq et al., 2019)             | ML    | ANN                              | B/M   | Breast cancer classification      | Ac = 99.4, Sn = 99.58, SP = 99.37 |
| (Rampun et al. 2020)             | ML    | SVM                              | BI    | Breast density classification     | Ac = 83.3, Ac = 80.5   |
| (Mohanty et al, 2020)            | ML    | WC-SSA-KELM                      | N/AB  | correctly classify digital mammograms | MIAS: Ac = 99.62, AUC = 0.99, Sn = 0.99, Sp = 0.97, MCC = 0.98, F-measure = 0.99 |
| (Ali-Antari, Han, and Kim 2020)  | DL    | Regular feedforward CNN, ResNet-SO, and InceptionResNet-V2 | B/M   | improve the diagnostic performance of breast lesions | MIAS: Ac = 100, AUC = 1 |
| (Muduli et al., 2020)            | ML    | MFO-ELM                          | N/AB  | Classification of breast masses  | DDSM: Ac = 99.76, AUC = 0.9958 |
| (Zhang et al., 2020)             | DL    | AdaBoost                          | B/M   | Breast mass classification        | DDSM: Ac = 98.8, AUC = 0.9927, InBreast: Ac = 97.3, Sn = 72.4, SP = 98.38 |
| (Shen et al, 2020)               | DL    | Deep Belief Network (DBN)        |       | automatic diagnosis of breast cancer | Ac = 91.5, Sn = 94.1, SP = 72.4 |

(Continued)
| Ref. | ML/DL | Technique | Class | Scope | Evaluation results (%) |
|------|-------|-----------|-------|-------|------------------------|
| (Arora, Rai, and Raman 2020) | ML | NN | B/M | Automatic classification | Ac = 0.88, AUC = 0.88 |
| (Rabidas et al., 2020) | ML | SVM, RF, FLDA | B/M | Characterization of mammographic masses | Mias: ROC = 0.94, Ac = 86.90 |
| (Patil and Biradar 2020) | DL | CNN + recurrent neural network (RNN) = CRNN | B/M/N | Breast cancer detection | Ac = 0.90, Sn = 0.92, Sp = 0.89, Pre = 0.78, F1-score = 0.84, MCC = 0.78 |
| (Fanizzi et al., 2020) | ML | Binary RF | N/AB | Breast microcalcification diagnosis | Ac = 97.31, AUC = 98.16 |
| (Indra et al., 2020) | ML | UTVO, ANN, ANN+PSO | B/M | Diagnosis system for healthcare applications | Ac = 97.12, Sn = 90.06, Sp = 99.52 |
| (Bouddraa, Melouah, and Merouani 2020) | ML | Simple logistic | B/M | Automatic differentiation | Ac = 96.7, Sn = 100, Sp = 94.7 |
| (Agnes et al. 2020) | DL | Multiscale All Convolutional Neural Network (MA-CNN) | B/M/N | Mammogram classification | Ac = 96.47, Sn = 96, Sp = 96, AUC = 0.99 |
| (Tavakoli et al. 2019) | DL | Fully connected layer with sigmoid function | N/AB | Detect and diagnose breast masses | Ac = 94.68, Sn = 93.33, Sp = 95.31, AUC = 0.95 |
| (Song, Li, and Wang 2020) | ML | SVM, XGBoost | B/M/N | Mass classification | Ac = 92.80 |
| (Loizidou et al. 2020) | ML | LDA, KNN, NB, SVM, DT, EDT | B/M | Micro-calcification detection and classification | Ac = 99.55, Sn = 98.82, Sp = 99.72 |
| (Shen et al., 2019) | DL | MS-ResCU-Net | B/M | Simultaneous segmentations and classification | Ac = 94.12, Sn = 97.56, Sp = 88.89, Pre = 93.02, F1-score = 95.24, AUC = 96.16 |
| (Shu et al. 2020) | DL | RGP and GGP | B/M | Mammographic diagnosis | CBIS-DDSM: Ac = 0.767, AUC = 0.823, p-value = 0.009 |
| (Farhan et al., 2020) | ML | SVM, LR, and KNN | B/M | Breast cancer detection | INbreast: Ac = 0.922, AUC = 0.924, p-value = 0.011 |
| (Albalawi et al., 2020) | DL | CNN | N/B/M | Cancer classification | Ac = 97.14, Sn = 96.52, Sp = 98.88 |
| (Zhang et al., 2021) | DL | Net-5 | N/AB | Cancer detection | Ac = 96.1, Sn = 96.2, Sp = 96 |
| (Saffari et al., 2020) | DL | CNN | BIRADS | Cancer classification | Pre = 97.85, Sn = 97.85, Sp = 99.28 |
| (Li, Mukundan, and Boyd 2021) | ML | SVM | BIRADS | Cancer classification | Ac = 84.6, AUC = 95.3 |
| (Boumaraf et al. 2020) | ML | BPN | BIRADS | Mass classification | Ac = 84.5, PPV = 84.4, NPV = 94.8, MCC = 79.3 |
teach the system to detect the same suspicious features that are assessed by radiologists. Features that have been selected can distinguish between benign and malignant regions to reduce errors of classification. Despite considerable effort, no consensus has been reached as of yet about those functions, which are needed, as illustrated in Table 6. The last phase of the CAD system which is considered as the CAD heart is classification. It is a data mining operation that is an effective means of finding and extracting trends from broad datasets using various methods of ML and DL.

**Pre-processing (Enhancement)**

In the data processing procedure for image processing, pre-processing is regarded as critical. The ultimate goal is to enhance the quality of the images produced. A pre-processing step is used in image processing techniques to either improve image quality by suppressing unwanted distortions or to improve image features before any further processing is performed (Zebari et al. 2019). The success of subsequent image processing steps, such as segmentation, feature extraction, feature selection, and classification is highly dependent on the accuracy of pre-processing. Inhomogeneity, low contrast, and unidentified noise are all common characteristics of clinical images that necessitate pre-processing. Pre-processing can help suppress these problems...
in medical images where they affect analysis. Many techniques are used in pre-processing, such as manual correction and mathematical operations, noise removal and enhancement (George et al., 2017).

In this systematic review, 45 of the 107 studies using DMs in the first phase were pre-processed to improve the following phases of the 107 studies on breast cancer that were surveyed. DM’s pre-processing stage is compared among recent publications in Table 4. The pre-processing phase was used by some publications, but evaluation was done in a later phase, as shown by the segmentation results in this paper. The pre-processing techniques used by most DMs consist of three stages. Remove radiopaque artifacts and labels by denoising the mammogram, enhancing the contrast, and applying these techniques. Median, Gaussian, Morphological and Wiener filters are commonly used for denoising DMs. Many publications use contrast enhancing algorithms such as contrast stretching, histogram equalization, contrast limited adaptive histogram equalization, logarithmic contrast enhancement, and exponential contrast enhancement, among others Exponential Contrast Enhancement (ECE). These algorithms are used to enhance the DMs so that specific ROIs or microcalcifications or masses visible in the image can be displayed more clearly. Whopping 46 papers (40%) of the papers in this sample had some form of pre-processing done. This filter has the highest rate of use for denoising DMs in the literature with 14 papers (30.04%), while the Contrast Limited Adaptive Histogram Equalization (CLAHE) filter has the highest rate of use for improving contrast with nine papers (19.56%). Additionally, the pre-processing phase is used to narrow down the ROI by

Figure 4. CAD pipelines based on ML and DL models.
eliminating regions with artifacts, noise, or pectoral muscle. The detection of ROI is made possible by the thresholding technique, which removes artifacts, background, and noise from images of the pectoral muscle (11% – 23.91%).

**Segmentation**

The process of segmentation involves splitting an image into several areas that share common characteristics including contrast, brightness, texture, color, and grey level. Segmentation aims to perform manipulation of an image’s representation towards easier analysis and improved meaningful content (Sharma and Preet 2016). Each segmented area is allocated with pixels from an image. During the enhancing process of an image, segmentation typically comes after pre-processing. The primary purpose of executing image segmentation is not to produce an image with higher quality, rather the step is carried out to delineate and discover observable structures and regions of focus (Zebari et al. 2020).

Segmentation can be broadly categorized into two image intensity characteristics, namely discontinuity and similarity. Similarity divides an image into several areas based on similarity, dependent on pre-set criteria. Meanwhile, discontinuity refers to dividing an image according to rapid intensity fluctuations (Patil and Deore 2013). Figure 3 illustrates primary segmentation types that have been widely utilized in the segmentation of medical images.

![Diagram](image)

**Figure 5.** Medical Image processing using CAD Based on ML techniques.
Threshold Based Segmentation
Pixel-based segmentation technique falls under the sub-category of segmentation techniques (Patil and Deore 2013). The pixel-based technique is considered the most rudimentary image segmentation technique attributed to the simplicity of its implementation concept. Despite this, the technique is effective in segmenting images containing bright objects that are surrounded by a dark background. In the pixel-based technique, thresholding is used to calculate the value where an object should be separated from the background. Thresholding may be subdivided into two, namely, local thresholding and global thresholding (Zebari et al. 2020). Thresholding via global thresholding exploits global information. As abnormalities appear lighter than tissues around them, thresholding is thus a viable solution to perform separation of objects from background in segmentation. Local thresholding is also labeled as adaptive thresholding. In operations, adaptive thresholding dynamically alters the values of thresholding, conditional on local properties of an image’s sub-regions. Specifically, the division of an image into regions is followed by a determination of a threshold value that is contingent on the properties of local pixels in a specific region of interest (Triyani et al. 2016). Heuristic optimisation methods can be used to perform thresholding (Kadry et al. 2021).

Region-Based Segmentation
Similarity-based segmentation divides an image into several regions depending on criteria of similarity that have been pre-set. The technique begins either with an individual pixel or a cluster of pixels, which are also known as seeds. Through this technique, neighboring seeds are examined, and subsequently, only seeds that meet the criteria of similarity for a structure would be considered for inclusion (Zeebaree et al., 2019a). Similarity can be described based on an image’s edges and/or intensities. Reiteration of examination of seeds that meet a set of pre-set criteria is ended when no new pixels are included in a structure of interest. A primary distinctive feature of this technique is its ability to perform segmentation of similar regions and generating relevant regions (Sadad et al. 2018).

Machine Learning-Based Segmentation
One of the most potent techniques in automating analysis and segmenting medical images is machine learning. The technique can perform learning on complex relationships from empirical data to derive decisions accurately (Liu et al., 2014). Machine learning-based techniques for segmentation may be further classified into supervised and non-supervised techniques. Supervised machine learning primarily thrives in performing a different set of tasks via only altering the training set. Segmentation training data are labelled automatically by grouping identical pixels under unsupervised learning (Gordillo et al., 2013).
**Edge Based Segmentation**

Segmentation based on edges is the most widely utilized technique for detecting edges, such as boundaries that are responsible for delineating different regions. Edge-based segmentation operates based on discovering dissimilarities of pixels towards determining nearby boundaries that correspond to objects within an image (Gupta and Anand 2017). The technique achieves a fast computation and is operable without needing historical information about an image’s content (Thanh et al. 2020). Furthermore, the technique is designed, such that it is highly perceptive to substantial fluctuations in grey level values and in a way that allows it to independently evaluate whether an edge falls within an edge or otherwise (Liu et al. 2020). This technique is effective in overcoming the consequence of size changes in the segmented object that is caused by the incompatible thresholding strategy utilized in segmenting an image.

**Deep Learning Based Segmentation**

DL-based image segmentation techniques have achieved good results in the field of image segmentation with artificial intelligence’s rapidly developing. Deep learning has some benefits in segmentation accuracy and speed over traditional machine learning and computer vision methods. This can help doctors verify the size of tumors and quantify the treatment effect before and after using deep learning to segment medical images. This reduces the amount of work that doctors have to do by a great deal (Liu et al. 2021).

Despite the fact that traditional image segmentation methods no longer hold a candle to the cutting-edge deep learning-based segmentation methods currently in use, the concepts still hold value. For example, the presented threshold-based image segmentation algorithm, the region-based image segmentation technique, and the edge detection-based segmentation method (He et al. 2017). To segment an image, these techniques draw on expertise in digital image processing and mathematics. It is easy to calculate and quick to segment, but there is no way to insure the segmentation is accurate down to the last detail. Deep learning models for image segmentation have made significant progress recently. The accuracy of their segmentation has outperformed that of conventional techniques. Image semantic segmentation was first effectively implemented with a fully convolutional network. This was the first time convolutional neural networks were used for image segmentation, and it was a breakthrough (Lin et al. 2017). Researchers proposed the use of full convolutional networks, which were developed by the authors. In addition to these, there are a number of segmentation networks that excel at processing fine edges, including U-Net, Mask R-CNN, RefineNet, and DeconvNet.

Based on the literature review of segmentation techniques for DMs of breast cancer, several segmentation methods typically utilized by various researchers such as neural networks, level set, watershed algorithm, clustering,
thresholding, hybrid techniques, etc. as shown in Table 5. It is shown that the surveyed papers introduced efficient automated CAD systems for the identification of breast cancer. From this systematic review it is observed that (59 papers – 51.3%) performed segmentation methods in CAD systems. The researchers used an adaptive thresholding method to segment the DMs of breast cancer. This method will also aid in distinguishing between the various forms of the tumors, e.g., benign and malignant. Based on the surveyed papers (8 papers – 13.55%) used thresholding technique to segment ROI from DMs. Clustering is a mathematical study from unsupervised learning, this technique deals with discovering a hidden structure from an unlabeled data set. Since clusters are divided from each other by regions of the comparatively low density of point, clusters define as “continuum-like regions of this space,” or areas surrounded by space that have a high density of points, which are separated from other high-density functions by low-density regions of the point. Accurate and efficient techniques to detect ROI in DMs based clustering were presented, 5 papers – 8% from the surveyed studies used clustering methods. Similarly, the surveyed papers used edge detection-based segmentation methods to segment ROI from DMs. Moreover, Table 5 showed that (8 papers – 13.55%) of researchers were introduced different automatic computing system based on region-based segmentation as well as hybrid techniques to extract ROI from DMs to improve a classification method which could predict breast cancer. Furthermore, recently DLTs were used widely in image processing fields, from our surveyed papers it has been investigated that DLTs were used widely in the segmentation of DMs (15 papers – 25.42%). Eventually, (10 papers – 16.94%) used other segmentation techniques to segment DMs for further processing.

**Feature Extraction**

Image processing tasks regularly involve a large corpus, which consumes a significant amount of time and is less practical for the task of efficiently classifying objects from background in segmentation. One strategy to reduce computation time is to perform the transformation of input data by reducing the number of feature vectors. The process of transforming the input data is known as feature extraction. Feature vectors typically hold related information and are exploited as input vectors in classification tasks. Classification of features could be performed based on shape, texture, and color (Tatikonda, Bhuma, and Samayamantula 2018). As seen on mammogram DMs of the individual body, various organs and tissues have very various texture detail. Texture has traditionally been a significant diagnostic function since texture analysis is a good method for lesion identification and disease diagnosis. Computerized feature extraction from mammography images is the most promising strategy to be used in performing breast cancer diagnosis. This is
attributed to faster analysis and higher accuracy in diagnosing possible signs of breast cancer. The features hold vital information about digital images that are useful in analyzing images. Primary criteria which have been utilized to discriminate malignant and benign masses include shape and texture (Goudarzi et al., 2018) (Chaieb and Kalti 2019).

**Texture Features**

Among the most essential characteristics considered for distinguishing ROI or artifacts in the image is the texture feature. The estimation of most of the textural features is performed utilizing values of gray level from the entire image or the ROIs only. During this accelerated phase within cancerous tumors, there is the development of a growing number of nuclei in cancerous tissue. Therefore, it is possible to distinguish various stages of cancer with the aid of texture characteristics (Sajeev, Baiger, and Lee 2018) (Saleck, ElMoutaouakkil, and Mouçouf 2017). An explanation of such characteristics involves resemblance, variance, curvature, comparison, etc. Features of texture may be categorized into two including frequent and statistical features. Statistical features utilized in this study comprised five classes, namely, First-Order Statistics (FOS), Gray-Level Run-Length Matrices (GLRLM), Gray-Level Difference Matrices (GLDM), Gray-Level Co-occurrence Matrices (GLCM), and Tamura features (Chaieb and Kalti 2019). Frequency features are a texture that is transformed into the frequency domain, which does not involve an image’s spatial domain. Two structural transformation techniques are studied including 2D wavelet transform and Gabor transform (Bagchi et al. 2020).

Feature selection is a technique used to reduce the dimension of data, which is widely utilized in the areas of data mining, statistics, pattern recognition, and machine learning. In operations, the technique reduces a set of features into a subset of important features that are dependent on certain criteria. Typically, a set of features consumes a large dimensionality space attributed to large variations of abnormalities and normal tissues (Mohanty et al., 2019) (Tubishat et al., 2020). Thus, it becomes necessary to remove features deemed insignificant and perform selection on features that are deemed most promising to be used to discriminate tumors from a set of all features. This comes with its inherent challenge to select features that are capable of uplifting accuracy while at the same time can improve searching time (Shastri, Tamrakar, and Ahuja 2018) (Kou et al. 2020).

**Morphological Features**

Geometric features have also been termed as shape or morphological features. The features take after the shapes of regions of interest (Vikhe and Thool 2018). Analyzing geometric features of suspected lesions that are identified from views in mammograms meticulously is useful, as this may be able to
positively envisage the probability of abnormality and substantiate subsequent necessity to conduct a biopsy. Along with density, lesion’s margin, size, and shape are critical in defining the probability of a lesion falling either under a malignant tumor or a benign mass category (Sapate et al. 2018).

**Intensity Features**

Intensity characteristics exclude from the voxel depiction of ROI. Although several visualizations are built upon the local features (median, mode, and variance), typically ROI visualizations are built upon the intensity-based features (Mohamed et al. 2018). Regardless of the data or the likelihood class, the values of gray-scale values inside an ROI are represented by a statistical model. The histogram of the intensities helps describes the structure of the area, the details of each pixel, and other suspicious characteristics (Berbar., 2018). These features and properties help detect and define the ROI. In two dimensions, an image is a function that maps the spatial coordinates x and y into a value $f(x, y)$ that represents the image’s gray level intensity at that point. An image is a function in two dimensions. A digital image is one in which $x$, $y$, and $f(x, y)$ are all discrete and finite quantities. Each pixel in a digital image has a specific position and gray intensity value, and together they make up a digital image. The spatial domain refers to the area covered by an image’s coordinates (Massafrā et al. 2021). In general, statistical features may be produced from the histogram of an image, such as, mean, variance, skewness, kurtosis, entropy, and capacity (Kaushal et al., 2019) (Pashoutan, Shokouhi, and Pashoutan 2017).

**Deep Features**

Machine learning has a connection to the problem of learning from input data samples because of the unified rule base that are used in it. This method includes analytical, statistical, and mathematical methods instead of explicitly programming the machines to learn from the training data. In the improvement of computer-aided breast cancer identification methods, machine learning techniques such as SVM, naive Bayes, artificial neural networks (ANNs), and set classifiers were becoming popular (Oza et al., 2021). Machine learning algorithms begin with the extraction of image features. Image features are frequently defined using arrays or descriptors, which training processes can then make use of. Choosing the right features is critical for training accuracy. Due to a variety of issues, the traditional machine learning paradigm has evolved into deep learning. Deep learning is more general than conventional machine learning because it focuses on mechanisms for drawing inferences from data and achieves higher generalization levels. One of the most influential deep learning networks is the so-called CNN, which has convolutional layers (Pillai et al., 2019) (Oza et al., 2021). To the contrary of traditional machine learning approaches, deep learning techniques do not require feature
extraction steps because of the large number of inner layers that extract features as they pass through layer-embedded operators. By studying thousands of images during the training process (Sechopoulos, Teuwen, and Mann 2020), DL-based algorithms learn what an abnormal mass looks like instead of inserting data on its shape, size, pattern as well as other features.

Table 6 observed that various researchers utilized various methods for feature extraction purposes. Many researchers used texture features (26 papers – 53.06%) as classifier input and obtained good results. GLCM is a method that is mostly utilized to extract texture features based on the surveyed papers (20 papers – 40.81%). Shape features are terminology used to characterize the shape of masses such as circularity, convexity or concavity indexes, spiculation index, perimeter, and more. Cancerous masses are more irregular and spiculated whereas healthy ones are rounder and more oval. Due to this reality, shape features are commonly utilized as identifiers in mass classification. This consistency includes the use of an appropriate segmentation method that can extract the ROI from unwanted regions. The most widely utilized methods for feature extraction in DMs are texture and morphological methods. Therefore, the combination between both features texture and morphological is regarded as the best method. Seven papers – 14.28% have used the integration of texture and morphological features. Moreover, DL is also used to extract features (7 papers – 14.28%) as an input to the classifier.

Breast Cancer Diagnosis

The most advanced sense of a human being is vision, but sometimes, the human vision is limited in it is capacity to process images. Therefore, through the concept of image processing and ML, computerized systems can acquire information about a problem that the human vision cannot acquire (Yadav and Jadhav 2020). This means that sometimes computerized systems are required in cases whereby the human vision is limited and cannot distinguish a problem. Analysis of medical images for instance X-rays, ultrasound (Irfan et al. 2021), thermal (Rajinikanth et al. 2021) images and scanners can help in radiologic diagnosis (Saxena et al., 2020). Figure 4 presents steps involved in a CAD system using the ML and DL techniques. The pre-processing and segmentation stages can be used for both ML and DL.

Machine learning techniques and image processing have made great contributions to the area of medicine through the digitalization of medical images, which allows the analysis and investigation of phenomena using a computer. The basic capability of ML is that it can discover new models without learning much about the underlying structures (Gardezi et al. 2019). This sort of research can extract complicated knowledge from noise or other details with a great deal of success. As the usage of statistical models for expert systems eliminates subjective assessments, these models provide excellent insight into
the clinical analysis of provided diseases (Singh, Singh, and Bhatia 2018) (Asri et al. 2016). The ML techniques may be used to find the breast lesion trends since these algorithms are used in the processing and forecasting of medical images. Therefore, much research has also utilized various machine learning methods in the prediction and diagnosis of breast cancer. Figure 5 shows a step of CAD system-based machine learning in medical image processing.

Deep learning strategies are representation-learning methods that consist of complex yet basic components and are utilized to change the representation at one stage into a more complex presentation at marginally more intellectual stages. The incredibly Deep Neural Network (DNN) framework made it capable of high-level inference and advanced artificial intelligence functions (Murtaza et al., 2019). DL paradigms provide new opportunities in the area of biomedical informatics due of its features for instance excellent results, end-to-end learning model with integrated learning feature, capacity to manage complex and multi-modal data and so on. DL methods have been utilized in the productive classification and interpretation of DMs of breast cancer (Zheng et al., 2020).

DL differs from ML because it addresses data in the method in a certain way, it is described a bit differently. Whilst Artificial Neural Networks (ANN) are employed to replicate the convolutions of the nociceptor neuron, ML approaches are based on certain standardized knowledge regarding the data that they operate upon. Unlike supervised learning, which is the process of learning a mapping function input to an output based on previously seen input-output pairs, unsupervised learning is not characterized by minimal human control and may be characterized as a kind of ML that occurs when a machine looks for unknown trends in data without prior labeling (Dembrower et al., 2020) (Sharma and Mehra 2020) (Hussein, 2012) (Kim-Soon, Abdulmaged, and Mostafa 2021).

When performing classification on suspected lesions, the goal is to identify those with a high likelihood of being correctly identified and the lowest risk of leading to diagnostic errors. Textural and geometric features’ values are utilized to proceed with classification, as elaborated earlier (Sapate et al. 2018). In this section, general classification techniques that are utilized to differentiate between the types or subtypes of cancers are briefly described. In essence, two learning algorithms are commonly widely used in the task of classifying tumors namely supervised and unsupervised algorithms. Most of the CAD systems for breast cancer detection from mammogram images used ML techniques to classify cancer subtypes. Several supervised and unsupervised techniques were used: Support Vector Machine (SVM), K-nearest Neighbor (KNN), Neural Network (NN), Naive-Bayes (NB), C4.5, Decision Tree (DT), Linear Discriminant Analysis (LDA), Ensemble, Logistic regression, ANN, Bayesian, Multilayer Perceptron (MLP), Self-Organizing Map (SOM), Neuro-Fuzzy System (ANFIS), Probabilistic Neural Network (PNN),
Fully Connected Neural Networks (FC-NNs), Multiple-Instance Random Forest (MI-RF), and Convolutional Neural Network (CNN) on breast cancer databases to compare the performance of those algorithms. The surveyed papers have used different techniques as classifiers from two main groups including MLTs and DLTs. From Table 7 it is shown that (44 papers – 57.14%) used MLTs whereas (33 papers – 42.85%) used DLTs. We categorized the analyzed studies based on the technique used to discriminate breast masses. We extracted the techniques they used in each paper, the classes used in the classification, the scope of the study, and the results they achieved. From the 118 papers analyzed in this study, 80 papers presented in Table 7, 35% (27 papers), 11.68% (9 papers), 14.28% (11 papers), 24.67% (19 papers) used SVM, KNN, ANN, and CNN as a single classifier to distinguish mammographic masses, respectively. We analyzed 34 papers (44.15%) that used more than one method to classify mammographic masses. Some of these studies proposed a hybrid classifier that combined different methods, while other studies examined different classifiers for classification. The studies (Melekoodappattu et al., 2018) (Mohanty et al., 2019) (Mohanty et al., 2020) (Muduli et al., 2020) (Patil and Biradar 2020) (Indra et al., 2020) (Kaur, Singh, and Kaur 2019) (Zhang and Wang 2019) created a hybrid classifier based on using different classifiers, an overview of papers that used one or more than one technique is given in Table 7. SVM has a higher rate of use whereas KNN and ANN have a lower rate.

Typically, the classification process is binary, i.e., benign and malignant (46.75% – 36 papers). However, (12 papers – 15%) papers used the class normal and abnormal, and (12 papers – 15%) used three classes (benign, malignant, and normal). Moreover, we showed that (10 papers – 12.5%) used multi-classes in the classification while at the first step, the classification has been done into normal and abnormal then the abnormal has classified into benign and malignant. Also, some studies (8 papers – 10%) also used BI-RADS classes (2, 3, 4, and 5) for classification. In terms of results, accuracy was reported in 69 papers (94%). Most of the surveyed papers (38 papers – 52%) presented their performance evaluation based on accuracy, sensitivity, and specificity, while (36 papers – 49%) used AUC in their evaluation.

**Discussion**

In this paper, various techniques employed in different stages of the CAD system to diagnose breast cancer using DMs images have been discussed. Pre-processing is the initial step in digital image analysis which is performed after the image acquisition. It plays an important role in diagnosing the biological tissues captured in an image by refining the quality of the image without destroying the important features. The current study shows that most of the
researchers use a median filter to reduce noise as well as CLAHE as a contrast enhancement technique. Several surveyed papers used pre-processing methods to segment pectoral muscle, artifacts, and image background in DMs.

To classify breast cancer into different classes, feature extraction is essential. Textural and morphological features were used for early diagnostics of DMs of the breast. The textural features can aid in the grading of the cancerous tissue. GLCM technique has a superior rate in using feature extraction technique based on the surveyed papers. In classification, both MLTs and DLTs are used to classify extracted features into different classes. As per the surveyed papers, SVM has the maximum rate in using as a classifier from MLTs whereas CNN has a higher rate from DLTs. SVM can recognize non-linear boundaries between classes in feature space and have many kernels to be used. They also can deal well with overfitting, particularly in the high-dimensional feature space.

We epitomize the recommendations as well as review the guidelines on how to boost the efficiency of breast cancer diagnosis and classification utilizing DMs. During the survey of this SR, it is noticed that most of the publications utilized datasets from one database only. Moreover, the pre-processing stage is a crucial stage to improve the performance of further stages, wherein most publications do not utilize any method of this stage; e.g., CLAHE to ameliorate the contrast of DMs, to smooth the DMs based on unsharp masking method, and to reduce noise from the image using noise reduction filters. Furthermore, to increase the generalization and reduce the overfitting of the system, both augmentation and drop-out are recommended to utilize. For mathematically practical it is preferred to utilize better image quality or full resolution whereas many researchers reduced image resolution. Another problem according to the dataset is that utilizing only one database or format during the evaluation. The classifier would have an easier time dealing with this, whereas DMs from different databases and the use of both formats Screen-Film Mammography (SFM) and Full-Field Digital Mammography (FFDM) together would be problematic.

Further, some recurring issues have been noticed in some of the surveyed publications. The issue outlined here is the challenge in contrasting the sensibility and specificity of a report that presents only the Area Under the Curve (AUC) with another that presents only the sensibility and specificity. This challenge in the study fields renders it challenging to figure out the literature in this research domain. Another supposed issue with this analysis is that the researchers do not equate the findings by the classifier with the results that are collected by the clinician for the reasons of whether the classifier is more reliable. The next issue we noticed was the fact that in several publications the approach utilized during the experiments is not explicit or was not present e.g., k-fold cross-validation, a left one out technique, a holdout technique, and so on. Over the above, one standard repository which is
generated along with the ground truth of the images is needed to test and verify the segmentation results; thus, it helps in successful diagnosis. Despite this, it is suggested to provide uniform open-access image datasets that include images from various image modalities for the same case to endorse the dependence on more than one image modality in the classification role and merge the details from several views. CAD systems enable to provide results relying on various perspectives related to various image modalities.

**Conclusion**

The results of this systematic review can help to support inventive research efforts for improving automated CAD systems to help the medical research community in the identification of breast cancer at an early stage. Current MLTs have utilized various image modalities in CAD systems for breast cancer detection. The basic components of the CAD system for breast cancer diagnosis are based on DMs including the pre-processing, segmentation, feature extraction, feature selection, and classification stages. Recent trends have been analyzed for pre-processing techniques that show that it needs more quality of the image before segmentation or feature extraction phases. To explore new developments regarding segmentation and classification methods, this analysis examined the influences of CAD schemes. The research reveals that the potential CAD method can be independent of the magnification factor and dataset. ML classifiers based on DL that were built by adding several layers in the framework become more computationally challenging as the number of layers increase. For the conventional methods, it is rather complicated to compare to DL. It also needs a massive number of datasets for training. However, the data augmentations that come from the assistance of numerous deep learning algorithms have contributed to delivering more consistent and accurate performance. While there are some effective approaches in the literature, there is also a potential to explore more efficient strategies in future work to aid with breast cancer detection at an early level. We hope that this study will guide the breast tissue research community to continue to improve their methods of diagnosing breast cancer.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

The author(s) reported that there is no funding associated with the work featured in this article.
References

Agnes, A., J. Anitha, S. I. A. Pandian, and J. D. Peter, et al. 2020. Classification of mammogram images using multiscale all convolutional neural network (MA-CNN). *Journal of medical systems* 44 (1):1–9. doi:10.1007/s10916-019-1494-z.

Ahmed, L., M. M. Iqbal, H. Aldabbas, S. Khalid, Y. Saleem, and S. Saeed, et al. 2020. Images data practices for semantic segmentation of breast cancer using deep neural network. *Journal of Ambient Intelligence and Humanized Computing* 1–17. https://doi.org/10.1007/s12652-020-01680-1

Al-Antari, M. A., M. A. Al-Masni, M.-T. Choi, S.-M. Han, and T.-S. Kim, et al. 2018. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *International journal of medical informatics* 117:44–54. doi:10.1016/j.ijmmedinf.2018.06.003.

Al-Antari, M. A., S.-M. Han, and T.-S. Kim. 2020. Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. *Computer methods and programs in biomedicine* 196:105584. doi:10.1016/j.cmpb.2020.105584.

AlSalman, H., and N. Almutairi. 2019. IDSS: An Intelligent Decision Support System for Breast Cancer Diagnosis. In 2019 2nd International Conference on Computer Applications & Information Security, 1–6, Riyadh, Saudi Arabia.

Arafa, E.-S. A., A. H. Asad, and H. A. Hefny. 2019. Machine Learning Algorithms for Breast Cancer CADx System in the Mammography. In 2019 15th International Computer Engineering Conference (ICENCO), 210–15, Giza, Egypt.

Arora, D. S., R. Vig, and M. Hanmandlu. 2019. Mammogram image enhancement using entropy and CLAHE based intuitionistic fuzzy method. In 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 24–29.

Arora, R., P. K. Rai, and B. Raman. 2020. Deep feature-based automatic classification of mammograms. *Medical & biological engineering & computing*, 58(6):1199-1211.

Asri, H., H. Mousannif, H. Al Moatassime, and T. Noel, et al. 2016. Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Computer Science* 83:1064–69. doi:10.1016/j.procs.2016.04.224.

Aziz, H. W., S. Saeed, S. Rathore, and M. Rafique, et al. 2018. Automated breast cancer detection using machine learning techniques by extracting different feature extracting strategies. In 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), New York, NY, USA, 327–31.

Bagchi, S., K. G. Tay, A. Huong, and S. K. Debnath, et al. 2020. Image processing and machine learning techniques used in computer-aided detection system for mammogram screening-A review. *International Journal of Electrical and Computer Engineering* 10 (3):2336.

Berbar, M. A. 2018. Hybrid methods for feature extraction for breast masses classification. *Egyptian informatics journal* 19 (1):63–73. doi:10.1016/j.eij.2017.08.001.

Boudraa, S., A. Melouah, and H. F. Merouani. 2020. Improving mass discrimination in mammogram-CAD system using texture information and super-resolution reconstruction. *Evolving Systems* 11 (4):697–706. doi:10.1016/j.s12530-019-09322-4.
Boumaraf, S., X. Liu, C. Ferkous, and X. Ma, et al. 2020. A new computer-aided diagnosis system with modified genetic feature selection for bi-RADS classification of breast masses in mammograms. *BioMed Research International* 2020. doi: 10.1155/2020/7695207.

Chaieb, R., and K. Kalti. 2019. Feature subset selection for classification of malignant and benign breast masses in digital mammography. *Pattern Analysis and Applications* 22 (3):803–29. doi:10.1007/s10044-018-0760-x.

Charan, S., M. J. Khan, and K. Khurshid. 2018. Breast cancer detection in mammograms using convolutional neural network. In 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 1–5.

Chen, L. C., S. Wang, and P. Chen. 2020. A Novel Multi-Scale Adversarial Networks for Precise Segmentation of X-Ray Breast Mass. IEEE Access, 8, 103772–81.

Chen, X., A. Zargari, A. B. Hollingsworth, H. Liu, B. Zheng, and Y. Qiu, et al. 2019. Applying a new quantitative image analysis scheme based on global mammographic features to assist diagnosis of breast cancer. *Computer methods and programs in biomedicine* 179:104995. doi:10.1016/j.cmpb.2019.104995.

Cheng, Y., Y. Gao, L. Xie, X. Xie, and W. Lin, et al. 2020. Spatial Enhanced Rotation Aware Network for Breast Mass Segmentation in Digital Mammogram. IEEE Access. DOI: 10.1109/ACCESS.2020.2978009

Dallalı, A., A. Slimen, S. El Khediri, and Y. Khemili, et al. 2018. Detection of lesion in mammograms. In 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), Hammamet. Tunisia, 479–83.

Das, P., and A. Das. 2019. A fast and automated segmentation method for detection of masses using folded kernel based fuzzy c-means clustering algorithm. *Applied Soft Computing* 85:105775. doi:10.1016/j.asoc.2019.105775.

Dash, M. R., and B. Majhi. 2020. Automated breast cancer detection in digital mammograms: A moth flame optimization based ELM approach. *Biomedical Signal Processing and Control* 59:101912. doi:10.1016/j.bspc.2020.101912.

Esener, İ. I., S. Ergin, and T. Yüksel. 2018. A novel multistage system for the detection and removal of pectoral muscles in mammograms. *Turkish Journal of Electrical Engineering & Computer Sciences* 26 (1):35–49. doi:10.3906/elk-1703-272.

Farhan, A. H., and M. Y. Kamil. 2020. Texture Analysis of Breast Cancer via LBP, HOG, and GLCM techniques. In IOP Conference Series: Materials Science and Engineering, IOP Publishing, 928, 072098. 7. DOI: 10.1088/1757-899X/928/7/072098.

Gao. 2018. SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis. *Computerized Medical Imaging and Graphics* 70: 53–62. doi:10.1016/j.compmedimag.2018.09.004.

Gardezi, S. J. S., A. Elazab, B. Lei, and T. Wang, et al. 2019. Breast cancer detection and diagnosis using mammographic data: Systematic review. *Journal of medical Internet research* 21 (7):e14464. doi:10.2196/14464.

George, M. J., and D. A. S. Dhas. 2017. Preprocessing filters for mammogram images: a review. In 2017 Conference on Emerging Devices and Smart Systems (ICEDSS), Mallasamudram, India, 1–7.

Geweid, G. G., and M. A. Abdallah. 2019. A novel approach for breast cancer investigation and recognition using M-level set-based optimization functions. IEEE Access, 7, 136343–57.

Gherghout, Y., Y. Tili, and L. Souici. 2019. Classification of breast mass in mammography using anisotropic diffusion filter by selecting and aggregating morphological and textural features. *Evolving Systems* 12, 273–302.

Gong, X., Z. Yang, D. Wang, Y. Qi, Y. Guo, and Y. Ma, et al. 2019. Breast density analysis based on glandular tissue segmentation and mixed feature extraction. *Multimedia Tools and Applications* 78 (22):31185–214. doi:10.1007/s11042-019-07917-2.
Gordillo, N., E. Montseny, and P. Sobrevilla. 2013. State of the art survey on MRI brain tumor segmentation. *Magnetic resonance imaging* 31 (8):1426–38. doi:10.1016/j.mri.2013.05.002.

Gou, S. C., J. Wang, and F.-Y. Wang. 2019. Simultaneous segmentation and classification of mass region from mammograms using a mixed-supervision guided deep model. IEEE Signal Processing Letters, 27, 196–200.

Goudarzi, M., and K. Maghooli. 2018. Extraction of fuzzy rules at different concept levels related to image features of mammography for diagnosis of breast cancer. *Biocybernetics and Biomedical Engineering* 38 (4):1004–14. doi:10.1016/j.bbe.2018.09.002.

Gu, S., Y. Chen, F. Sheng, T. Zhan, and Y. Chen, et al. 2019. A novel method for breast mass segmentation: from superpixel to subpixel segmentation. *Machine Vision and Applications* 30 (7–8):1111–22. doi:10.1007/s00138-019-01020-0.

Gupta, D., and R. Anand. 2017. A hybrid edge-based segmentation approach for ultrasound medical images. *Biomedical Signal Processing and Control* 31:116–26. doi:10.1016/j.bspc.2016.06.012.

Hagos, Y. B., A. G. Mérida, and J. Teuwen. 2018. Improving breast cancer detection using symmetry information with deep learning. In Danail Stoyanov et al. *Image Analysis for Moving Organ, Breast, and Thoracic Images*, 90–97. Springer.

Han, F. 2015. Texture feature analysis for computer-aided diagnosis on pulmonary nodules. *Journal of digital imaging* 28 (1):99–115. doi:10.1007/s10278-014-9718-8.

Hazarika, M., and L. B. Mahanta. 2018. A novel region growing based method to remove pectoral muscle from MLO mammogram images. In *Advances in Electronics, Communication and Computing*, 307–16. Singapore: Springer.

He, K., G. Gkioxari, P. Dollár, and R. Girshick, et al. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 2961–69. 22–29 October 2017.

Hussein, J., M. A. Burhanuddin, M. A. Mohammed, M. Elhoseny, and B. Garcia-Zapirain, et al. 2021. Fully automatic segmentation of gynaecological abnormality using a new viola–jones model. *Computers, Materials & Continua* 66 (3):3161–82. doi:10.3390/cmc.2021.012691.

Indra, M. M. 2020. Multilevel Tetrolet transform based breast cancer classifier and diagnosis system for healthcare applications. *Journal of Ambient Intelligence and Humanized Computing* 12, 3969–3978.

Irfan, R., A. A. Almazroii, H. T. Rauf, R. Damaševičius, E. A. Nasr, and A. E. Abdelgawad, et al. 2021. Dilated semantic segmentation for breast ultrasonic lesion detection using parallel feature fusion. *Diagnostics* 11 (7). doi:10.3390/diagnostics11071212.

Jiao, Z., X. Gao, Y. Wang, and J. Li, et al. 2018. A parasitic metric learning net for breast mass classification based on mammography. *Pattern Recognition* 75:292–301. doi:10.1016/j.patcog.2017.07.008.

Jiménez-Gaona, Y., M. J. Rodríguez-Álvarez, and V. Lakshminarayanan. 2020. Deep-Learning-Based Computer-Aided Systems for Breast Cancer Imaging: A Critical Review. *Applied Sciences* 10 (22):8298. doi:10.3390/app10228298.

Jung. 2018. Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network. *PloS one* 13(9):e0203355. doi:10.1371/journal.pone.0203355.

Junior, G. B., S. V. Da Rocha, J. D. de Almeida, A. C. de Paiva, A. C. Silva, and M. Gattass, et al. 2019. Breast cancer detection in mammography using spatial diversity, geostatistics, and concave geometry. *Multimedia Tools and Applications* 78 (10):13005–31. doi:10.1007/s11042-018-6259-z.

Kadry, S., R. Damasevicius, D. Taniar, V. Rajinikanth, and I. A. Lawal, et al. 2021. Extraction of tumour in breast MRI using joint thresholding and segmentation - A study. 2021 IEEE 7th International Conference on Bio Signals, Images and Instrumentation, ICBSII 2021, Chennai, India. 10.1109/ICBSII51839.2021.9445152.
Karthiga, R., K. Narasimhan, and G. Usha. 2019. Breast Cancer Diagnosis using Curvelet and Regional Features. In 2019 International Conference on Computer Communication and Informatics (ICCCI) Coimbatore, India, 1–5.

Kaur, P., G. Singh, and P. Kaur. 2019. Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. Informatics in Medicine Unlocked 16:100151. doi:10.1016/j.imu.2019.01.001.

Khan, H. N., A. R. Shahid, B. Raza, A. H. Dar, and H. Alquhayz, et al. 2019. Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. IEEE Access, 7, 165724–33.

Kim-Soon, N., A. I. Abdulmaged, and S. A. Mostafa. 2021. A framework for analyzing the relationships between cancer patient satisfaction, nurse care, patient attitude, and nurse attitude in healthcare systems. J Ambient Intell Human Comput. doi:10.1007/s12652-020-02888-x.

Kou, G., P. Yang, Y. Peng, F. Xiao, Y. Chen, and F. E. Alsaadi, et al. 2020. Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods. Applied Soft Computing 86:105836.

Lahoura, V., H. Singh, A. Aggarwal, B. Sharma, M. A. Mohammed, R. Damaševičius, S. Kadry, and K. Cengiz, et al. 2021. Cloud Computing-Based Framework for Breast Cancer Diagnosis Using Extreme Learning Machine. Diagnostics 11 (2):241. doi:10.3390/diagnostics11020241.

LeCun, Y., Y. Bengio, and G. Hinton. 2015. Deep learning. Nature 521 (7553):436–44. doi:10.1038/nature14539.

Li, H., R. Mukundan, and S. Boyd. 2021. Novel Texture Feature Descriptors Based on Multi-Fractal Analysis and LBP for Classifying Breast Density in Mammograms. Journal of Imaging 7 (10):205. doi:10.3390/jimaging7100205.

Li, H., S. Zhuang, D.-A. Li, J. Zhao, and Y. Ma, et al. 2019. Benign and malignant classification of mammogram images based on deep learning. Biomedical Signal Processing and Control 51:347–54. doi:10.1016/j.bspc.2019.02.017.

Li, L. M., J. Wang, F. Wu, T. Liu, and Y. Pan, et al. 2014. A survey of MRI-based brain tumor segmentation methods. Tsinghua science and technology 19 (6):578–95. doi:10.1109/TST.2014.6961028.

Li, S., M. Dong, G. Du, and X. Mu, et al. 2019. Attention dense-u-net for automatic breast mass segmentation in digital mammogram. IEEE Access, 7, 59037–47.

Li, Y., L. Zhang, H. Chen, and L. Cheng, et al. 2020. Mass detection in mammograms by bilateral analysis using convolution neural network. Computer Methods and Programs in Biomedicine 195:105518. doi:10.1016/j.cmpb.2020.105518.

Lin, G., A. Milan, C. Shen, and I. Reid, et al. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 1925–34. 21–26 July 2017.

Liu, Q., Z. Liu, S. Yong, K. Jia, and N. Razmjooy, et al. 2020. Computer-aided breast cancer diagnosis based on image segmentation and interval analysis. Automatika 61 (3):496–506. doi:10.1080/00051144.2020.1785784.

Liu, X., L. Song, S. Liu, and Y. Zhang, et al. 2021. A review of deep-learning-based medical image segmentation methods. Sustainability 13 (3):1224. doi:10.3390/su13031224.

Liu, X., T. Zhu, L. Zhai, and J. Liu, et al. 2019. Mass classification of benign and malignant with a new twin support vector machine joint |l_ {2, 1}|-norm. International Journal of Machine Learning and Cybernetics 10 (1):155–71. doi:10.1007/s13042-017-0706-4.

Loizidou, K., G. Skouroumouni, C. Nikolaou, and C. Pitris, et al. 2020. An automated breast microcalcification detection and classification technique using temporal subtraction of mammograms. IEEE Access, 8, 52785–95.
Mabrouk, M. S., H. M. Aify, and S. Y. Marzouk. 2019. Fully automated computer-aided diagnosis system for micro calcifications cancer based on improved mammographic image techniques. Ain Shams Engineering Journal 10 (3):517–27. doi:10.1016/j.aesej.2019.01.009.

Massafra, R., S. Bove, V. Lorusso, A. Biafora, M. C. Comes, V. Didonna, and D. La Forgia, et al. 2021. Radiomic Feature Reduction Approach to Predict Breast Cancer by Contrast-Enhanced Spectral Mammography Images. Diagnostics 11 (4):684. doi:10.3390/diagnostics11040684.

Matos, E. F. 2019. Diagnosis of breast tissue in mammography images based local feature descriptors. Multimedia Tools and Applications 78 (10):12961–86. doi:10.1007/s11042-018-6390-x.

Melekoodappattu, J. G., and P. S. Subbian. 2019. A hybridized ELM for automatic micro calcification detection in mammogram images based on multi-scale features. Journal of medical systems 43 (7):1–12. doi:10.1007/s10916-019-1316-3.

Memon, A. R., E. Wang, J. Hu, J. Egger, and X. Chen, et al. 2020. A review on computer-aided design and manufacturing of patient-specific maxillofacial implants. Expert review of medical devices 17 (4):345–56. doi:10.1080/17434440.2020.1736040.

Mohamed, A. A., W. A. Berg, H. Peng, Y. Luo, R. C. Jankowitz, and S. Wu, et al. 2018. A deep learning method for classifying mammographic breast density categories. Medical physics 45 (1):314–21. doi:10.1002/mp.12683.

Mohamed, B. A., and N. M. Salem. 2018. Automatic classification of masses from digital mammograms. In 2018 35th National Radio Science Conference (NRSC), Cairo, Egypt, 495–502.

Mohammed, M. A., B. Al-Khateeb, A. N. Rashid, D. A. Ibrahim, M. K. Abd Ghani, and S. A. Mostafa, et al. 2018. Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. Computers & Electrical Engineering 70:871–82. doi:10.1016/j.compeleceng.2018.01.033.

Mohanty, A. K., M. R. Senapati, and S. K. Lenka. 2013. An improved data mining technique for classification and detection of breast cancer from mammograms. Neural computing & applications 22 (S1):303–10. doi:10.1007/s00521-012-0834-4.

Mohanty, F., S. Rup, B. Dash, B. Majhi, and M. Swamy, et al. 2019a. A computer-aided diagnosis system using Tchebichef features and improved grey wolf optimized extreme learning machine. Applied Intelligence 49 (3):983–1001. doi:10.1007/s10489-018-1294-z.

Mohanty, F., S. Rup, B. Dash, B. Majhi, and M. Swamy, et al. 2019b. Digital mammogram classification using 2D-BDWT and GLCM features with FOA-based feature selection approach. Neural computing & applications 1–15. doi:10.1007/s00521-019-04095-y.

Mohanty, F., S. Rup, B. Dash, B. Majhi, and M. Swamy, et al. 2019c. Mammogram classification using contourlet features with forest optimization-based feature selection approach. Multimedia Tools and Applications 78 (10):12805–34. doi:10.1007/s11042-018-5804-0.

Mughal, A., N. Muhammad, and M. Sharif. 2019. Adaptive hysteresis thresholding segmentation technique for localizing the breast masses in the curve stitching domain. International journal of medical informatics 126:26–34. doi:10.1016/j.ijmmedinf.2019.02.001.

Muhammad, M. N., M. Sharif, A. Rehman, and T. Saba, et al. 2018. Removal of pectoral muscle based on topographic map and shape-shifting silhouette. BMC cancer 18 (1):1–14.

Murtaza, G. 2019. Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. Artificial Intelligence Review 53, 1655–1720.

Nayak, D. R., R. Dash, B. Majhi, R. B. Pachori, and Y. Zhang, et al. 2020. A deep stacked random vector functional link network autoencoder for diagnosis of brain abnormalities and breast cancer. Biomedical Signal Processing and Control 58:101860. doi:10.1016/j.bspc.2020.101860.
Obaid, O. I., M. A. Mohammed, M. K. A. Ghani, A. Mostafa, and F. Taha, et al. 2018. Evaluating the performance of machine learning techniques in the classification of Wisconsin Breast Cancer. *International Journal of Engineering & Technology* 7 (4.36):160–66.

Oza, P., P. Sharma, and S. Patel 2021. Machine Learning Applications for Computer-Aided Medical Diagnostics. In Proceedings of the Second International Conference on Computing, Communications, and Cyber-Security, Ghaziabad, India, Singapore: Springer, 377–92. 3–4 October.

Oza, P., P. Sharma, S. Patel, and A. Bruno, et al. 2021. A Bottom-Up Review of Image Analysis Methods for Suspicious Region Detection in Mammograms. *Journal of Imaging* 7 (9):190. doi:10.3390/jimaging7090190.

Pashoutan, S., S. B. Shokouhi, and M. Pashoutan. 2017. Automatic breast tumor classification using a level set method and feature extraction in mammography. In 2017 24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME), Tehran, Iran, 1–6.

Patil, D. D., and S. G. Deore. 2013. Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing* 2 (1):22–27.

Patil, R. S., and N. Biradar. 2020. Automated mammogram breast cancer detection using the optimized combination of convolutional and recurrent neural network. *Evolutionary Intelligence* 14, 1459–1474.

Pavan, A. L., A. Vacavant, A. F. Alves, A. P. Trindade, and D. R. de Pina, et al. 2019. Automatic identification and extraction of pectoral muscle in digital mammography. In World Congress on Medical Physics and Biomedical Engineering 2018, Springer, Prague, Czech Republic, 151–54.

Peng, J., C. Bao, C. Hu, X. Wang, W. Jian, and W. Liu, et al. 2020. Automated mammographic mass detection using deformable convolution and multiscale features. *Medical & biological engineering & computing* 58 (7):1405–17. doi:10.1007/s11517-020-02170-4.

Pillai, R., P. Oza, and P. Sharma 2020. Review of machine learning techniques in health care. In Proceedings of the ICRIC 2019, Jammu, India, Cham, Switzerland: Springer, 103–11. 8–9 March 2019.

Prathibha, G., and B. C. Mohan. 2018. Mammograms Classification Using Multiresolution Transforms and Convolution Neural Networks. In 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 1–7.

Preet, S. B. 2016. Classification of mammogram images by using CNN classifier. In 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 2743–49.

Rabidas, R., and W. Arif. 2020. Characterization of mammographic masses based on local photometric attributes. *Multimedia Tools and Applications* 79 (29–30):21967–85. doi:10.1007/s11042-020-08959-7.

Rahimeto, S., T. G. Debelee, D. Yohannes, and F. Schwenker, et al. 2019. Automatic pectoral muscle removal in mammograms. *Evolving Systems* 12, 519–526.

Rahmatika, A., A. Handayani, and A. W. Setiawan. 2019. Automated Segmentation of Breast Tissue and Pectoral Muscle in Digital Mammography. In 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT), Yogyakarta, Indonesia, 397–401.

Rajinikanth, V., S. Kadry, D. Taniar, R. Damasevicius, and H. T. Rauf, et al. 2021. Breast-cancer detection using thermal images with marine-predators-algorithm selected features. 2021 IEEE 7th International Conference on Bio Signals, Images and Instrumentation, ICBSII 2021, Chennai, India. 10.1109/ICBSII51839.2021.9445166.
Rampun, A. 2019. Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network. *Medical image analysis* 57:1–17. doi:10.1016/j.media.2019.06.007.

Rampun, A., P. J. Morrow, B. W. Scotney, and H. Wang, et al. 2020. Breast density classification in mammograms: An investigation of encoding techniques in binary-based local patterns. *Computers in Biology and Medicine* 122:103842. doi:10.1016/j.compbiomed.2020.103842.

Rastgarpour, P. M., A. Sharifi, and S. Yazdani. 2019. Extraction of spiculated parts of mammogram tumors to improve accuracy of classification. *Multimedia Tools and Applications* 78 (14):19979–20003. doi:10.1007/s11042-019-7185-4.

Routray, N. P. R. 2018. Textural feature based classification of mammogram images using ANN. In 2018 9th International Conference on Computing, Communication and Networking Technologies, Bengaluru, India, 1–6.

Rup, M. S., B. Dash, B. Majhi, and M. Swamy. 2020. An improved scheme for digital mammogram classification using weighted chaotic salp swarm algorithm-based kernel extreme learning machine. *Applied Soft Computing* 91:106266. doi:10.1016/j.asoc.2020.106266.

Sadad, T., A. Munir, T. Saba, and A. Hussain, et al. 2018. Fuzzy C-means and region growing based classification of tumor from mammograms using hybrid texture feature. *Journal of computational science* 29:34–45. doi:10.1016/j.jocs.2018.09.015.

Sadoughi, F., Z. Kazemy, F. Hamedan, L. Owji, M. Rahamanikatigari, and T. T. Azadboni, et al. 2018. Artificial intelligence methods for the diagnosis of breast cancer by image processing: A review. *Breast Cancer Targets Ther* 10:219. doi:10.2147/BCTT.S175311.

Saffari, N., H. A. Rashwan, M. Abdel-Nasser, V. Kumar Singh, M. Arenas, E. Mangina, and D. Puig, et al. 2020. Fully Automated Breast Density Segmentation and Classification Using Deep Learning. *Diagnostics* 10 (11):988. doi:10.3390/diagnostics10110988.

Sajeev, S., M. Baijer, and G. Lee. 2018. Superpixel texture analysis for classification of breast masses in dense background. *IET Computer Vision* 12 (6):779–86. doi:10.1049/iet-cvi.2017.0586.

Salama, M. S., A. S. Eltrass, and H. M. Elkamchouchi. 2018. An improved approach for computer-aided diagnosis of breast cancer in digital mammography. In 2018 IEEE international symposium on medical measurements and applications (MeMeA), Rome, Italy, 1–5.

Saleck, M., A. ElMoutaouakkil, and M. Mouçouf. 2017. Tumor detection in mammography images using fuzzy C-means and GLCM texture features. In 2017 14th International Conference on Computer Graphics, Imaging and Visualization, Marrakesh, Morocco, 122–25.

Samant, N., and P. Sonar. 2018. Mammogram classification in transform domain. In 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 56–62.

Sapate, S. G., A. Mahajan, S. N. Talbar, N. Sable, S. Desai, and M. Thakur, et al. 2018. Radiomics based detection and characterization of suspicious lesions on full field digital mammograms. *Computer methods and programs in biomedicine* 163:1–20. doi:10.1016/j.cmpb.2018.05.017.

Saxena, S., and M. Gyanchandani. 2020. Machine learning methods for computer-aided breast cancer diagnosis using histopathology: a narrative review. *Journal of medical imaging and radiation sciences* 51 (1):182–93. doi:10.1016/j.jmir.2019.11.001.

Sechopoulos, I., J. Teuwen, and R. Mann. 2020. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Semin Cancer Biol* 72:214-225. doi: 10.1016/j.semcancer.2020.06.002.

Sharma, S., and R. Mehra. 2020. Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight. *Journal of digital imaging* 33 (3):632–54. doi:10.1007/s10278-019-00307-y.
Shastri, A. A., D. Tamrakar, and K. Ahuja. 2018. Density-wise two stage mammogram classification using texture exploiting descriptors. *Expert Systems with Applications* 99:71–82. doi:10.1016/j.eswa.2018.01.024.

Shayma’a, A. H., M. S. Sayed, M. I. Abdalla, and M. A. Rashwan, et al. 2019. Detection of breast cancer mass using MSER detector and features matching. *Multimedia Tools and Applications* 78 (14):20239–62. doi:10.1007/s11422-019-7358-1.

Shen, L., M. He, N. Shen, N. Yousefi, C. Wang, and G. Liu, et al. 2020. Optimal breast tumor diagnosis using discrete wavelet transform and deep belief network based on improved sunflower optimization method. *Biomedical Signal Processing and Control* 60:101953. doi:10.1016/j.bspc.2020.101953.

Shen, R., K. Yan, F. Xiao, J. Chang, C. Jiang, and K. Zhou, et al. 2018. Automatic pectoral muscle region segmentation in mammograms using genetic algorithm and morphological selection. *Journal of digital imaging* 31 (5):680–91. doi:10.1007/s10278-018-0068-9.

Shinde, V., and B. T. Rao. 2019. Novel approach to segment the pectoral muscle in the mammograms. In *Cognitive Informatics and Soft Computing*, 227–37. Springer.

Shoaib, N. M., and S. Gattoufi. 2018. Detection and classification of the breast abnormalities in Digital Mammograms via Linear Support Vector Machine. In 4th Middle East Conference on Biomedical Engineering (MECBME), Tunis, Tunisia, 141–46.

Shu, X., L. Zhang, Z. Wang, Q. Lv, and Z. Yi, et al. 2020. Deep neural networks with region-based pooling structures for mammographic image classification. IEEE transactions on medical imaging, 39, 2246–55. 6.

Simon, C. P. 2020. A Novel Approach for Mammogram Enhancement using Nonlinear Unsharp Masking and L0 Gradient Minimization. *Procedia Computer Science* 171:1848–57. doi:10.1016/j.procs.2020.04.198.

Singh, R., R. Singh, and A. Bhatia. 2018. Sentiment analysis using Machine Learning technique to predict outbreaks and epidemics. *Int J Adv Sci Res* 3 (2):19–24.

Singh, V. K. 2018. Conditional generative adversarial and convolutional networks for X-ray breast mass segmentation and shape classification. In International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer,Granada, Spain, 833–40.

Singh, V. P., and R. Srivastava. 2018. Automated and effective content-based mammogram retrieval using wavelet based CS-LBP feature and self-organizing map. *Biocybernetics and Biomedical Engineering* 38 (1):90–105. doi:10.1016/j.bbe.2017.09.003.

Soleimani, O. V. 2020. Michailovich, On Segmentation of Pectoral Muscle in Digital Mammograms by Means of Deep Learning. *IEEE Access*, 8, 204173–82.

Song, R., T. Li, and Y. Wang. 2020. Mammographic Classification Based on XGBoost and DCNN With Multi Features. *IEEE Access*, 8, 75011–21.

Soulami, K. B., M. N. Saidi, B. Honnit, C. Anibou, and A. Tamtaoui, et al. 2019. Detection of breast abnormalities in digital mammograms using the electromagnetism-like algorithm. *Multimedia Tools and Applications* 78 (10):12835–63. doi:10.1007/s11422-018-5934-4.

Suganithi, S. J., V. G., M. Parvathy, and C. D. Devi, et al. 2020. Pectoral Muscle Segmentation in Mammograms. *Biomed Pharmacol* 13 (3):1357–65. doi:10.13005/bpj/2005.

Sun, L., J. Wang, Z. Hu, Y. Xu, and Z. Cui, et al. 2019. Multi-view convolutional neural networks for mammographic image classification. *IEEE Access*, 7, 126273–82.

Suresh, R., A. N. Rao, and B. E. Reddy. 2018. Improving the mammogram images by intelligibility mammogram enhancement method. In 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 1011–15.

Tariq, N., B. Abid, K. A. Qadeer, I. Hashim, Z. Ali, and I. Khosa, et al. 2019. Breast cancer classification using global discriminate features in mammographic images. *Breast Cancer* 10:2.
Tatikonda, K. C., C. M. Bhuma, and S. K. Samayamanthula. 2018. The analysis of digital mammograms using HOG and GLCM features. In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 1–7.

Tavakoli, N., M. Karimi, A. Norouzi, N. Karimi, S. Samavi, and S. R. Soroushmehr, et al. 2019. Detection of abnormalities in mammograms using deep features. Journal of Ambient Intelligence and Humanized Computing 1–13. https://doi.org/10.1007/s12652-019-01639-x

Teuwen, J., S. van de Leemput, A. Gubern-Mérida, A. Rodriguez-Ruiz, R. Mann, and B. E. Bejnordi, et al. 2018. Soft tissue lesion detection in mammography using deep neural networks for object detection. Proceedings of the 1st Conference on Medical Imaging with Deep Learning; MIDL’18; July 4-6, 2018; Amsterdam, The Netherlands. 2018. pp. 1–9.

Thanh, D. N., N. N. Hien, V. S. Prasath, and N. H. Hai, et al. 2020. Automatic initial boundary generation methods based on edge detectors for the level set function of the Chan-Vese segmentation model and applications in biomedical image processing. In Satapathy S., Bhatheja V., Nguyen B., Nguyen N., Le DN. (eds) Frontiers in Intelligent Computing: Theory and Applications. Advances in Intelligent Systems and Computing, vol 1014. Springer, Singapore. https://doi.org/10.1007/978-981-13-9920-6_18

Ting, F. F., Y. J. Tan, and K. S. Sim. 2019. Convolutional neural network improvement for breast cancer classification. Expert Systems with Applications 120:103–15. doi:10.1016/j.eswa.2018.11.008

Toz, P. E. 2018. A single sided edge marking method for detecting pectoral muscle in digital mammograms, Engineering. Technology and Applied Science Research 8 (1):2367–73. doi:10.48084/etasr.1719

Triyani, Y., H. A. Nugroho, M. Rahmawaty, I. Ardiyanto, and L. Choridah, et al. 2016. Performance analysis of image segmentation for breast ultrasound images. In 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 1–6.

Uthoff, J. C. S. 2018. Information theory optimization based feature selection in breast mammography lesion classification. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 817–21.

Vikhe, P., and V. Thool. 2018. Morphological operation and scaled rényi entropy based approach for masses detection in mammograms. Multimedia Tools and Applications 77 (18):23777–802. doi:10.1007/s11042-018-5681-6

Wang, H. et al., 2018. Breast mass classification via deeply integrating the contextual information from multi-view data. Pattern Recognition 80: 42–52. doi:10.1016/j.patcog.2018.02.026

Wang, R. et al., 2019a. Multi-level nested pyramid network for mass segmentation in mammograms. Neurocomputing 363: 313–20. doi:10.1016/j.neucom.2019.06.045

Wang, Z. 2019b. Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features. IEEE Access, 7, 105146–58.

Xi, P., C. Shu, and R. Goubran. 2018. Abnormality detection in mammography using deep convolutional neural networks. In 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rome, Italy, 1–6.

Xu, J., C. Li, Y. Zhou, L. Mou, H. Zheng, and S. Wang, et al. 2018. Classifying mammographic breast density by residual learning. arXiv preprint arXiv:1809.10241

Yadav, S. S., and S. M. Jadhav. 2020. Thermal infrared imaging based breast cancer diagnosis using machine learning techniques. Multimedia Tools and Applications 1–19. https://doi.org/10.1007/s11042-020-09600-3

Yan, Y. S., C. Song, and B. Zheng. 2019. A robust method for segmenting pectoral muscle in mediolateral oblique (MLO) mammograms. International journal of computer assisted radiology and surgery 14 (2):237–48. doi:10.1007/s11548-018-1867-7.
Yaşar, H., U. Kutbay, and F. Hardalaç. 2018. A new combined system using ANN and complex wavelet transform for tissue density classification in mammography images. In 2018 4th International Conference on Computer and Technology Applications (ICCTA), Istanbul, Turkey,179–83.

Yousefi M., Krzyżak A., and Suen C. Y. 2018. Mass detection in digital breast tomosynthesis data using convolutional neural networks and multiple instance learning. Computers in biology and medicine, 96, 283–293.

Yu, S., L. Liu, Z. Wang, G. Dai, and Y. Xie. 2019. Transferring deep neural networks for the differentiation of mammographic breast lesions, Science China. Technological Sciences 62 (3):441–47. doi:10.1007/s11431-017-9317-3.

Yu, X., N. Zeng, S. Liu, and Y.-D. Zhang. 2019. Utilization of DenseNet201 for diagnosis of breast abnormality. Machine Vision and Applications 30 (7–8):1135–44. doi:10.1007/s00138-019-01042-8.

Zebari, A., H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, et al. 2019. Enhance the Mammogram Images for Both Segmentation and Feature Extraction Using Wavelet Transform. In 2019 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Iraq, 100–05.

Zebari, D. A., D. Q. Zeebaree, A. M. Abdulazeez, H. Haron, and H. N. A. Hamed, et al. 2020. Improved Threshold Based and Trainable Fully Automated Segmentation for Breast Cancer Boundary and Pectoral Muscle in Mammogram Images. IEEE Access, 8, 203097–116.

Zeebaree, D. Q., H. Haron, A. M. Abdulazeez, and D. A. Zebari, et al. 2019a. Machine learning and region growing for breast cancer segmentation. In 2019 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Iraq, 88–93.

Zeebaree, D. Q., H. Haron, A. M. Abdulazeez, and D. A. Zebari, et al. 2019b. Trainable model based on new uniform LBP feature to identify the risk of the breast cancer. In 2019 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Kurdistan region, Iraq, 106–11.

Zeiser, F. A. 2020. Segmentation of Masses on Mammograms Using Data Augmentation and Deep Learning. Journal of digital imaging 33(4): 858–868.

Zhang, X., and Z. Wang. 2019. A microcalcification cluster detection method based on deep learning and multi-scale feature fusion. The Journal of Supercomputing 75 (9):5808–30. doi:10.1007/s11227-019-02867-w.

Zhang, Y.-D., S. C. Satapathy, D. S. Guttery, J. M. Górriz, and S.-H. Wang, et al. 2021. Improved Breast Cancer Classification Through Combining Graph Convolutional Network and Convolutional Neural Network. Information Processing & Management 58 (2):102439. doi:10.1016/j.ipm.2020.102439.

Zhang. 2020. DE-Ada*: A novel model for breast mass classification using cross-modal pathological semantic mining and organic integration of multi-feature fusions. Information Sciences 539: 461–86. doi:10.1016/j.ins.2020.05.080.

Zheng, X. 2020. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. Nature communications 11 (1):1–9. doi:10.1038/s41467-019-13993-7.

Zhong, S. J., A. Rampun, and H. Wang. 2018. A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms. Computers in biology and medicine 96:178–88. doi:10.1016/j.compbiomed.2018.03.011.

Zhu, W., X. Xiang, T. D. Tran, G. D. Hager, and X. Xie, et al. 2018. Adversarial deep structured nets for mass segmentation from mammograms. In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018), Washington, DC, USA, 847–50.