LITERATURE REVIEW

[Document subtitle]

Dr. B. LAKSHMANAN (ASSOCIATE PROFESSOR)
V.RAJA SUBRAMANIAN & K.VIJAYA GOKUL

1. **A NOVEL BREAST CANCER DETECTION ARCHITECTURE BASED ON A CNN-CBR SYSTEM FOR MAMMOGRAM CLASSIFICATION**

Dataset Used:

-CBIS-DDSM dataset (Curated Breast Imaging Subset Digital Database for Screening Mammography)

-The breast cancer detection approach presented in the PDF file uses mammograms from the CBIS-DDSM dataset . This dataset is a widely used public dataset in numerous research studies and comprises 2,620 mammograms, along with their corresponding masks and metadata.

Advantages:

- The proposed architecture is an end-to-end solution that can automatically detect breast cancer in mammograms and provide an explanation for the diagnosis.
- The use of CNNs for feature extraction can improve the accuracy of the system by automatically learning relevant features from the mammogram images.
- The use of CBR for classification and explanation generation can provide a transparent and interpretable decision-making process, which is important in the healthcare field.

Disadvantages:

- The proposed architecture may require a large amount of training data to achieve high accuracy, which can be a challenge in the medical domain due to the limited availability of annotated data.
- The use of CNNs for feature extraction can be computationally expensive and may require powerful hardware to achieve real-time performance.
- The proposed architecture may not be able to handle all types of breast cancer cases, and its performance may vary depending on the characteristics of the dataset and the quality of the mammogram images.

Accuracy:

Model	Accuracy (%)	Precision (%)	Recall (%)
KNN (K=7)	84.25	86.60	80.49
GAUSSIAN PROCESS	84.05	86.22	80.50
SVM (Kernel RBF)	83.23	84.42	80.91
MLP (4 Layers)	82.62	**91.05**	71.78
KNN (K=3)	80.98	83.04	77.17
Ada Boost	80.57	89.67	68.46
Gaussian NB	78.32	76.06	81.74
Quadratic Discriminant Analysis	77.30	73.05	85.48
SVM (Kernel = Linear)	75.66	81.77	65.15
Decision Tree	75.66	82.45	64.32
Our approach	**86.71**	82.42	**91.34**

2. **Breast cancer anomaly detection based on the possibility theory with a clustering paradigm**

Dataset Used:

CBIS-DDSM dataset

Advantages:

- Improved accuracy in mass and microcalcification detection, which is crucial for early breast cancer diagnosis.
- Better handling of the high uncertainty level in breast tissue anomaly classification, which can lead to more reliable and consistent results.

- Possibility-necessity based decision making can provide a more efficient and effective way of assigning samples to different classes.
- Clustering can create a more relevant partitioning of data, yielding a new redefinition of classes, based on the different clusters, which can be more efficient in case of sparse classes.

Disadvantage:

- The proposed approach may require a large amount of data to train the models and achieve high accuracy.
- The possibility-based modeling formalism may be more complex and difficult to implement than other traditional methods.
- The proposed approach may not be suitable for all types of breast tissue anomalies or may not generalize well to other datasets.

3. **Breast tumor segmentation in digital mammograms using spiculated regions**

Dataset Used:

Advantages:

-The technique accurately extracts the mass core (central region) and spiculated regions which contain the tumor border details.

-The technique significantly discriminated between the shape of malignant and benign masses.

-The technique resulted in higher accuracy and fewer calculations required to extract the considerable and spiculated regions of the masses effectively.

-The different thresholds did not change the final segmentation perceptibly.

Disadvantages:

-The α, β and γ thresholds affect the extraction of the spiculated regions and mass core, so future research could focus on optimizing these thresholds.

-More effective pre-processing methods could be applied to help extract the details of the mass core and the spiculated parts more precisely.

Tumor classification methods could be used to identify the strengths and weaknesses of the proposed method

Methodology:

-The present research introduces a new method for segmentation of tumor mammograms that extracts the spiculated regions and the mass core. Generally, the pixels of a spiculated region are located along a line and the pixels of the mass core regions are similar. The proposed method extracts these regions using the differences between a pixel and its adjacent pixels.

-The proposed method uses three thresholds to delete redundant pixels from the spiculated regions and the mass core. These regions then are merged to form the segmented tumor

4. YOLO Based Breast Masses Detection and Classification in Full-Field Digital Mammograms

Advantage:

-The use of anchor boxes concept in YOLO-V3 that are generated by applying k-means clustering on the dataset allows for the detection of most of the challenging cases of masses and their correct classification.

-Augmenting the dataset using different approaches, and comparing with other recent YOLO based studies, found that augmenting the training set only is the fairest and most accurate to be applied in realistic scenarios

Disadvantage:

-There is only one limitation for our study coming from the small size of the annotated mammograms dataset, while this is overcome by augmenting the original INbreast mammograms.

-YOLO-V3 is a good model for detecting small objects, but it has some limitations. For example, it can't detect small masses that are close together very well.

-Future work in breast cancer detection and classification should focus on improving the accuracy of YOLO-V3 for detecting small masses.

-This could be done by using more detailed datasets and by collecting more FFDM mammograms.

-The proposed method should also be tested on new mammograms and its robustness should be evaluated.

-In the future, the authors also plan to study the features of other breast abnormalities, such as micro-calcifications, and find a good approach to detect them accurately.

Methodology:

-We propose an end-to-end computer-aided diagnosis system based on You Only Look Once (YOLO). The proposed system first preprocesses the mammograms from their DICOM format to images without losing data. Then, it detects masses in full-field digital mammograms and distinguishes between the malignant and benign lesions without any human intervention. YOLO has three different architectures, and, in this paper, the three versions are used for mass detection and classification in the mammograms to compare their performance. The use of anchors in YOLO-V3 on the original form of data and its augmented version is proved to improve the detection accuracy especially when the k-means clustering is applied to generate anchors corresponding to the used dataset. Finally, ResNet and Inception are used as feature extractors to compare their classification performance against YOLO.

5. **Early detection and classification of abnormality in prior mammograms using image-to-image translation and YOLO techniques**

Advantages:

-The study used the YOLO architecture model to detect and classify suspicious lesions in mammograms. This model has been shown to be effective in other object detection tasks, and it was able to correctly localize and identify three different types of lesions: Mass, Calcification, and Architectural Distortion.

-The study also integrated the Prior mammograms from all used follow-up screenings to provide an early detection and classification on initial screened mammograms. This is important because it can help to identify abnormalities that may have been missed in earlier screenings.

-The study also used image-to-image translation techniques to create new translated Prior mammograms that can overcome the misalignment between the two screenings due to temporal and texture changes. This helps to improve the accuracy of the detection and classification results.

Disadvantages:

-One limitation of the study is that it was only conducted on a small dataset of mammograms. This means that the results may not be generalizable to a larger population.

-Another limitation is that the study used a two-stage approach to detection and classification. This means that the model first detected the lesions, and then it classified the lesions. This can lead to errors if the model incorrectly detects a lesion.

Methodology:

The study used the following methodology:

-The researchers collected a dataset of mammograms.

-They used the YOLO architecture model to train a model to detect and classify suspicious lesions in mammograms.

-They used image-to-image translation techniques to create new translated Prior mammograms that can overcome the misalignment between the two screenings due to temporal and texture changes.

-They tested the model on the new translated Prior mammograms.

-They evaluated the results of the model.

6. A deep learning architecture with an object-detection algorithm and a convolutional neural network for breast mass detection and visualization

Methodology:

The methodology used in this study involves an integrated deep learning architecture with an object-detection algorithm and a convolutional neural network

Algorithm:

The object-detection algorithm used in this study is YOLO v5

Merits:

The proposed system can reduce the time and effort required for radiologists to review mammograms and improve the accuracy of breast mass detection.

Demerits:

False positives, which can lead to reviewer frustration and diminish confidence in a tool's effectiveness, are a potential shortcoming of object-detection systems.

7. Automated breast cancer detection in mammography using ensemble classifier and feature weighting algorithms

Methodology:

The classification performance of the proposed breast cancer detection strategy, based on ensemble classifier and feature weighting algorithms, was evaluated experimentally to determine the resulting diagnostic accuracy

Algorithm 1: Artifact removal in the breast region.

Algorithm 2: Abnormal area removal in ROIs.

Merits:

The proposed scheme could be highly practical and effective for helping radiologists to detect breast cancer, greatly reducing the incidence of false positives and false negatives.

Demerits:

A label or artifact is attached to the breast region, which cannot be removed using the proposed region extraction method. Therefore, we hope region extraction can be improved to separate the label or artifact from the breast region.

Large breast cancer datasets cannot be employed in the proposed scheme because they are very difficult to obtain.

8. Breast cancer: Classification of suspicious regions in digital mammograms based on capsule network

Methodology:

The Methodology used in this study is Capsule based Deep learning model to classify suspicious masses in the breast into normal, benign, and malignant.

Algorithm:

Dynamic Routing algorithm

Merits:

The proposed model shows good performance for binary classification and multiclassification of suspicious breast masses, particularly for extremely dense mammograms.

Demerits:

The binary classification of masses into normal and abnormal achieves an accuracy of 96.03% whereas the multiclassification of breast masses into normal, benign, and malignant scores an accuracy of 77.78% which is less than binary classification.

9. Early detection and classification of abnormality in prior mammograms using image-to-image translation and YOLO techniques

Methodology:

The Methodology used in this study is the YOLO architecture model for simultaneous detection and classification of breast lesions was proposed.

Merits:

Our methodology could early detect locations of Mass lesions much earlier than the expert diagnoses using follow-up screenings within 3 to 6 years, which is relatively late for breast cancer diagnosis.

Demerits:

Preparing the right format of the training configuration for the YOLO-based model.

10. Mass segmentation and classification from film mammograms using cascaded deep transfer learning

Methodology:

Breast cancer diagnosis can be done from film mammograms by using the proposed cascaded U-net++Xception deep learning pipeline without clinical data.

Merits:

The proposed model may be used to reduce the workload of radiologists for mass detection, segmentation, and classification which are all crucial mammogram interpretation steps.

The proposed model is useful for breast cancer mass segmentation and classification on BCDR dataset and may be useful on additional mammogram datasets.

Demerits:

The limitations are:

1) A cascaded approach increases the requirement for both computational resources and the amount of training data;

2) BCDR and other datasets lack performance comparison between different radiologists so there is an uncertainty for setting an accomplishment level for the deep learning model;

3) The obtained performance results are valid for BCDR dataset. Application of the model on other datasets may require additional training.

11. Evaluation of deep learning models for detecting breast cancer using histopathological mammograms Images

Methodology:

The Methodology used in this study is Convolution Neural Network(CNN) is a deep learning method which is very effective in classifying images, object detection and image recognition.

Demerits:

Due to the limited volume of imaging data, the network models suffer from low accuracy.