[1] **Abebe Alemu Balcha***
[2] **Anteneh Girma (PhD)**
[3] **Mesfin Abebe (PhD)**
[4] **Moges Zenebe (PhD)**

A Hybrid Genetic Algorithm with a K-Means++ Clustering Model to Accurately Determine BI-RAD Scoring for Breast Mammography Image Screening

Journal of Electrical Systems

Abstract: - Recent studies have revealed that breast cancer is responsible for 25.84% of all cancer-related deaths. In Africa and Ethiopia, new breast cancer cases account for 29.46% and 31.85% of the total new cancer cases, respectively. Researchers are highlighting the need to improve false positive (FP) and false negative (FN) values in the confusion matrix to develop an effective early detection model for breast cancer. Artificial intelligence (AI) plays a crucial role in improving breast cancer diagnosis, as late detection significantly reduces survival rates. In Ethiopia, 72.56% of breast cancer diagnoses are in an advanced stage (95% CI; 68.46-76.65%), highlighting the urgent need for early detection. Furthermore, the lack of radiologists contributes to delays in the annual reading and classification of BI-RAD. It is important to note that most women diagnosed with breast cancer receive negative or non-invasive results. Mammography image data is obtained from the local diagnosis center where a senior radiologist provides recommendations on the diagnosis results and the next follow-up steps for patients. A retrospective research project focuses on screening breast mammography images to identify the level of BI-RAD. A computational intelligence framework has been developed that uses OpenCV for image classification and hybridizing genetic algorithms with K-Means++ clustering. The model accurately screens BI-RAD levels and provides results for patients based on the recommendations of senior radiologists, yielding significant results in the research article.

Keywords: Breast Cancer, BI-RAD, Mammography, Computational intelligence, Genetic Algorithm, K-Means++.

I. INTRODUCTION

In a global landscape where the Cancer occurs when a mutation or genetic change interfaces with the orderly process or cells grow uncontrollably to form a mass tumor [1]. Initially, breast cancer is confined to 85% in the lining of the ducts or lobes and does not cause symptoms. A breast is made up of three main parts. Lobules: Gland that produces milk. Ducts: tubes that carry milk. Connector tissue: fibrous and fatty tissue surrounded and holds everything together [2]. Breast cancer spreads outside the breast through blood vessels and lymph vessels, [3].

Table 1: Acronyms

Short form	Description
AI	Artificial Intelligence
BC	Breast Cancer
BIRAD	Breast Imaging Reporting and Data System
CI	Computational Intelligence
CNN	Convolutional Neural Network
CSO	Cuckoo Search Optimization
DL	Deep Learning
DDSM	Digital Database for Screening Mammography
DICOM	Digital Imaging and Communications in Medicine
FP	False Positives
FN	False Negatives
GA	Genetic Algorithm
GA-KM++	Genetic Algorithm K-Means++
Local MBIDS	Local Mammography Breast Image Dataset MIAS
MIAS	Mammographic Image Analysis Society

[1] Adama Science and Technology University. Adama, Ethiopia abebe. alemu@research.astu.edu.et
[2] Computer Science /Cyber-Security University of the District of Columbia. Washington DC. anteneh.girma@udc.edu
[3] Adama Science and Technology University Adama, Ethiopia. mesfin.abebe@astu.edu.et
[4] Addis Ababa University moges. zenebe@aau.edu.et

Short form	Description
ML	Machine Learning
MSE	Minimum square Error
SVM	Support Vector Machine
TP	True Positives
TN	True Negatives

Breast Cancer Staging: The most common staging system is the TNM (Tumor, Node, and Metastasis). TNM helps medical professionals speak the same language. T- Size of the tumor measured in cm. N- number of positive lymph nodes. M- refers to whether the cancer has moved to the distant [4, 5].

Stage of breast cancer: Early breast cancer is categorized as zero to III whereas Late stage or Advanced Stage breast cancer is IV or Metastatic Breast Cancer (MBC). Stage-0 Carcinoma in situ and Stage-I: determined by tumor size and lymph nodes with several cancers. The benign type of breast cancer is not invasive cancer, whereas malignant tumors are cancerous, and the wavering of the diagnosis eventually spreads beyond the original tumor to other parts of the body. The cancer spreads to the nearby lymph node and goes to the regional metastasis or other called distant metastasis [5, 6].

Different research has been conducted using artificial intelligence to diagnose BC. Machine learning and Deep learning algorithms are implemented for the detection of breast cancer from medical images. [7] Deep learning algorithms are not yet superior or inferior in terms of performance when compared to clinicians. An acceptable diagnostic performance analogous to the algorithms was observed. Research papers show that diagnosis performance depends on data quality and good feature selection. Based on current data analysis and research, radiologists miss 15% to 35% breast cancer from mammography image data. Even though AI has a gap in the early detection of BC [8], AI has a good impact on improving performance in the diagnosis of breast cancer with data intensive.

These can improve sensitivity and ensure fewer false positives than radiologists. However, it may have the risk of overfitting the training data, resulting in a breakable, degraded performance in certain settings. Machine learning has a tradeoff between accuracy and intelligibility [9]. More accurate models are usually not intelligible; in contrast, more intelligible models offer significantly worse accuracy, [10]. To ensure AI technology applies to cancer diagnosis and prognosis successfully, first, resolve the challenges that AI faces. For example, input data cannot be used directly from the medical imaging technology; this is a crucial step to extract features from the imaging data and process. [10], give emphases on the development and popularization of technology, in addition, the weights coefficient in the neural network models is tested, and calculated, and the confidence interval is reasonable, so medical interpretation needs further research. Screening of another patient from the positive is a time-consuming task. Countries like Ethiopia have limited radiologists for screening the image, and to calcification process takes time. For this reason, such a study has a solution to easily screen the image by BI-RAD scoring [6, 11]. According to the researchers, the screening is good for age groups of 40 to 45 years. In our country, these age groups are around 23%.

Globally, mammographic screening is the gold standard, as it is effective in reducing breast cancer mortality in women over 40 years of age. However, it is a complex public health strategy that requires resources for infrastructure and coordination. World Health Organization (WHO) recommends that screening programs be carried out only when their effectiveness has been demonstrated in the local context, that is, where resources are sufficient to finance a population-based service and the prevalence of the disease is sufficiently high to justify this investment. A recent study [12] in four countries in Sub-Saharan Africa revealed that the overall prevalence of breast cancer screening was only 12.9% during the study period. According to a document by the EMoH [13], initiatives outlined in the national guide aim to enhance the survival rate of breast cancer by prioritizing early detection at stage I and stage II. The goal is to achieve a diagnosis of invasive cancer in over 60% of cases by the year 2030. Screening women within the defined eligible age group allows for early diagnosis and swift treatment, improving the prognosis and outcomes by addressing the disease early. The comprehensive evaluation, imaging, tissue sampling, and pathology assessments are all completed within 60 days stated as an objective. Given the variation in the progression of different cancers based on their underlying biology, health systems need to differentiate between malignant and benign breast findings effectively. The focus of early breast cancer diagnosis is on individuals showing signs and symptoms of the disease, in contrast to cancer screening, which aims to detect preclinical cancer or precancerous lesions in a healthy target population [13]. This research concentrates on

screening mammography images using BI-RAD scoring to determine the stage, and it will be one of the inputs for the success of the document.

Problem and hypothesis: There are thousands of diagnostic images coming to the center within the limited number of radiologists. As we have seen the images are classified as negative, needed follow-up visit, or positive. Those customers take their time for waiting the result to have a summary of the senior radiology report reading the image. This is time-consuming and prone to errors in reading the image and interpreting the result, resulting in false negative and false positive results.

By doing this research, customers can leave the diagnosis center just after the imaging by taking the result; in addition, the other groups also have supportive results for the radiologist to decide on a follow-up visit or additional imaging. This minimizes false positives and false negatives in addition to the cost-effectiveness of the number of radiologists who use their time effectively on most screening patients. This enhances the screening process of the mammography image which is taken from the breast.

Method: We have collected 4092 breast cancer images along with diagnosis recommendations from senior radiologists. The images are organized based on the BI-RAD scores provided in the radiologist summary report. Out of these, 2062 images do not have explicit scores, but they do contain information about density, mammography view, and sizes. We have used the remaining data for training. Our approach involves using OpenCV for data classification and training, K-Means++ clustering, and a Genetic algorithm for optimization. The image data goes through preprocessing, which includes sizing, scaling, grayscale conversion, and noise reduction, normalization, filtering, and selecting the region of interest.

Result: We have after the process and training of the image Accuracy clustering of the image according to the BI-RAD scoring is 96.02, precision 95.34, and F1-score of 92.80.

Contribution: We have developed a computational intelligence model to screen, analyze, and interpret mammography images to detect potential breast abnormalities. Our prepared BI-RAD dataset provides a comprehensive and standardized collection of mammography images for training and testing the model. Our research approach aims to enhance the accuracy and efficiency of breast cancer screening by utilizing advanced computational methods to aid radiologists in the early detection and diagnosis of breast abnormalities. The goal of this research is to assess the stage of the breast image based on the BI-RAD scoring, thus streamlining the screening process, and reducing the requirement for specialized resources and time. Additionally, it aims to provide patients with immediate information about their results.

1.1 The paper is structured into the following sections:

Section 1 - Introduction: This section offers an overview of the challenges associated with breast cancer diagnosis and imaging screening, specifically focusing on mammography using the BI-RAD scoring system.

Section 2 - Literature Review This section provides a comprehensive analysis of the existing research literature on breast cancer, emphasizing the challenges encountered by radiologists in screening mammography images, key image features, and the current landscape of breast cancer diagnosis. It also discusses the influential benchmark research that informed our model.

Section 3 - Methods: In this section, we elaborate on our model for breast image screening based on the BI-RAD scoring system. This includes a discussion of the datasets used, data size, preprocessing methods, and tools, as well as the development and demonstration of our research models.

Section 4 - Results: This section presents a detailed description of the results obtained from research articles, our model analysis, training, and testing results on the dataset. We also compare our findings with existing research and employ plotting tools to visually represent the impact of our model in relation to current research.

Section 5 - Discussion and Conclusion This section delves into the strengths, weaknesses, and potential future research directions. We explore the potential impact of the research in the field, identify gaps, and provide guidance for researchers interested in working with the dataset and results. This final section concludes the research work by acknowledging the contributions of all participants and extends an invitation for further research in this area within a highly computational environment.

II. RELATED WORKS

Breast cancer is a disease in which abnormal breast cells grow out of control and form tumors. If left unchecked, the tumors can spread throughout the body and become fatal. Breast cancer cells begin inside the milk ducts the milk-producing lobules of the breast. The earliest form (insitu) is not life-threatening and can be detected in the early stages. Cancer cells can spread to nearby breast tissue (invasion). This creates tumors that cause lumps or

thickening. Invasive cancers can spread to nearby lymph nodes or other organs (metastasize). Metastasis can be life-threatening and fatal, [14, 15]. Cancer is the second leading cause of death in the United States, exceeded only by heart disease. One in every four deaths in the United States is due to cancer. Five years from 2012 to 2018 Breast cancer data classified by race, and age with data source.

In 2019 [16], in the United States, 264,121 new cases of Female Breast cancer were reported among women, and 42,280 women died from this cancer. For every 100,000 women, 130 new Female Breast cancer cases were reported, and nineteen women died of this cancer. According to WHO [2, 12], released on March 13, 2024, breast cancer caused 670,000 deaths globally in 2022. Half of all breast cancers occur in women with no specific risk factors other than sex and age. Breast cancer was the most common cancer among women in 157 countries out of 185 in 2022. 0.5–1% of breast cancers occur in men [11].

Global estimates reveal striking inequities in the breast cancer burden according to human development. For example, in countries with a very high Human Development Index (HDI), 1 in 12 women will be diagnosed with breast cancer in their lifetime and 1 in 71 women die of it [12]. In contrast, in countries with a low HDI; while only 1 in 27 women is diagnosed with breast cancer in their lifetime, 1 in 48 women will die from it.

Breast cancer can have combinations of symptoms, especially when it is more advanced. Most people will not experience symptoms when the cancer is still early, hence the importance of early detection. Symptoms of breast cancer can include a lump or thickening of the breast, often without pain change in the size, shape, or appearance of the breast, dimpling, redness, pitting, or other changes in the skin change in the appearance of the nipple, or the skin surrounding the nipple (areola) abnormal or bloody fluid from the nipple. Ductal carcinoma in situ (DCIS) is the presence of abnormal cells within a milk duct in the breast. DCIS is considered the earliest form of breast cancer. DCIS is noninvasive, meaning it has not spread out of the milk duct and has a low risk of becoming invasive. Breast cancer is a complex disease, and its stages help determine the extent of its spread. The breast cancer stages:

- Stage 0 (In Situ): Abnormal cells are confined to the milk ducts or lobules. It is non-invasive and highly treatable.
- Stage I: Small tumors (less than 2 cm) that have not spread beyond the breast.
- Stage II: Larger tumors (2-5 cm) or those with limited lymph node involvement.
- Stage III: Locally advanced cancer with significant lymph node involvement.
- Stage IV: Metastatic cancer that has spread to distant organs.

Calcifications are small deposits of calcium that show up on mammograms as bright white specks or dots on the soft tissue background of the breasts. The calcium readily absorbs the X-rays from mammograms. Calcifications usually do not show up on ultrasounds and never show up on breast magnetic resonance imaging. Calcifications are a common finding in mammograms and are especially common after menopause [17].

A fibroadenoma is a solid, painless, non-cancerous (benign) breast tumor. It is the most common type of non-cancerous breast tumor and usually does not increase the risk of breast cancer. Fibroadenomas can be classified according to their size and growth. There are four types: **Simple fibroadenomas** are the most common type. They tend to be smaller, and the cells look the same throughout the tissue. **Complex fibroadenomas** are more common in people older than 35 years. They tend to be larger and may have calcifications or cysts in them. Finding any kind of lump in your breast tends to set off alarm bells. But if you have a painful lump in your breast, you may have a breast cyst – which can result from normal hormonal changes. Breast cysts can occur at any age but are most common in women in their 30s and 40s.

Table 2: Stage Distribution (%) of New Cancer Cases, All Ages, All race, Female [18],

Measure from	Percent	Count
Localized	66.00%	832,142
Regional	25.80%	325,488
Distant	5.80%	73,054
Unstaged	2.40%	30,785

Table 3: Stage Distribution (%) of New Cancer Cases, All Ages, white, Female [18],

Measure from	Percent	Count
Localized	67.40%	697,141
Regional	24.90%	257,415
Distant	5.50%	56,678
Unstaged	2.20%	22,457

Table 4: Stage Distribution (%) of New Cancer Cases, All Ages, Black, Female [18],

Measure from	Percent	Count
Localized	57.60%	85,810
Regional	31.50%	46,946
Distant	8.30%	12,382
Unstaged	2.50%	3,774

Table 5: SSEER 5-Year Relative Survival Rates, by Stages all ages, 2012-2018 [18],

Measure from	5-Year Relative Survival (%)	Lower 95% C.I.	Upper 95% C.I.
Localized	99.1%	98.9	99.2
Regional	86.1%	85.8	86.4
Distant	30%	29.1	30.8
Unstaged	60%	58.3	61.7

From Table 23,4, the distribution of new breast cancer is classified by the distance as localized, regionally distant, and unstaged. From the data, distant distribution is on average 6.5% of the data were categorized by race.

In Table 5 the survival rate of distant distribution of breast cancer is around 30% whereas localized is 99.1% [18]. The prevalence of counts is 1,093,984 with a rate of 65.58% of 5-year USA on January 1, 2019. The 5-year relative survival rate is 90.6% for all races, while whites 91%, Black people 82.6% and other races 91.2% have a relative survival rate. U.S. 5-Year Age-Adjusted Mortality Rates, 2016-2020, 19.6% (210,505) all races, 26.6% (32,478) Black including Hispanic and 19.1% (170,082). [7] suggests that deep learning algorithms are not yet superior or inferior in terms of performance compared to clinicians. An acceptable diagnostic performance with analogous deep-learning algorithms was observed in the breast. It is important to note that no matter how well constructed the algorithm is, its diagnostic performance depends on the volume of raw data and quality training DL algorithms require reliable high-quality image inputs and AI-related method guides have been published, with many still under development [9]. AI in the early detection of BC was highlighted as a gap and methodological concern. The current mammography radiologist has 15% to 35 % missing breast cancer. AI-based algorithms are data-intensive, and the performance of the result is affected by the quality of data sets [8, 14]. There are published papers about breast cancer detection using AI methods, however, there is a gap in the early detection of breast cancer. The research paper [14], with opportunities to use various deep learning mechanisms to predict patients' status such as LSTM, GAN, and RNN as these types of research have not yet been conducted in the field. Focuses only on the accuracy metric to evaluate your performance and does not address the parameters of the confusion matrix and the AUC. This is insufficient since the accuracy metric does not distinguish between false positive and false negative classifications. Future studies should include at least AUC and F-scores to assess the performance of each model.

[19] The rise and dissemination of AI in breast cancer screening is poised to improve breast cancer risk assessment and enable personalized screening recommendations. However, many technical challenges related to the inherent properties of mammographic imaging have yet to be addressed, especially as AI developments transition to digital breast tomosynthesis. Furthermore, to accelerate the validation of AI breast cancer risk models and their transition to clinical implementation, it is paramount to enhance their reproducibility, interpretability, and robustness using large and heterogeneous datasets. With creative AI solutions to improve accuracy, validate performance, and cultivate trust in decision-making, AI will transform the way breast cancer screening is performed [19]. In the research paper [20], twelve studies on AI mammography systems for routine breast screening were reviewed. It was discovered that in six smaller studies, AI outperformed individual radiologists in accuracy. However, it remained uncertain whether the lower accuracy in other cases was attributed to variations in the case mix or the expertise of the radiologists. The two largest retrospective cohort studies in Europe revealed that all AI systems were less accurate than the consensus of two radiologists, and 34 out of 36 AI systems were less precise than a single reader. Furthermore, a separate large retrospective study indicated that while AI exhibited higher sensitivity compared to the original first reader decision, it showed lower specificity and was less accurate than a consensus reading.

Table 6: Breast Cancer diagnosis result using machine learning techniques [21],

Training Data	Technique	Challenge	Result
Mammogram	CRNN, FC-CSO	Bluer images not work,	Accuracy= 98.4%

Training Data	Technique	Challenge	Result
images		need filter	Specificity= 99.9%, F1-score=74.5%
Wisconsin BC Data sets	KNN	failed to work with large. data set	Accuracy= 97.51%
Mammogram images	CNN | Logistic Regression	To improve need large data-set	precision= 98.5%
DCE-MR Image	Multi-variant ML Model	need uniform scanning and contrast protocol	AUC= 77.1%
Mammogram images	Segmentation	need to scale to improve ac-curacy	Matching ratio 96.3%
mammogram images	KNN	need enhancement for classification	Accuracy= 94.44%
Digital mammogram	SVM	challenge to interpret for high dimensional matrix	Accuracy= 96.55, Sensitivity= 96.97%, Specificity= 96.29%
Mammography images	SVM	need to improve accuracy. by large dataset	Accuracy= 87.2%, AUC= 94%

According to Kumar's research, [21], mammogram image classification using CRNN techniques achieves an accuracy rate of 98.4%, a specificity of 99.9%, and F1 scores of 74.5%. However, CRNN techniques struggle to classify blurred images and require additional filtration. KNN techniques achieve an accuracy rate of 94.44%, but further improvements are needed for better classification. SVM techniques for digital mammogram images have an accuracy rate of 96.55%, a sensitivity of 96.97%, and a specificity of 96.20%. However, the challenge with digital mammograms lies in their high-dimensional matrix, which makes interpretation complex. Large data sets are required to improve the accuracy of SVM for mammography image datasets, as the current accuracy rate is 87.2% with an AUC of 94%. In summary, the highest precision in breast cancer diagnosis using mammography image data is approximately 98.4% with a specificity rate of 99.9% in CRNN and a sensitivity rate of 96.97% in SVM, Table 12.

From the research paper [22], the experimental investigation employed a total of 1646 mammography pictures from 4 subjects. The radiologist can follow this model when analyzing mammogram images. The prediction model provides results with a precision of 0.75 for BI-RAD 0,1 and 2 are 0.75, 0.96, and 0.6, respectively.

Based on the paper [23], the SVM-based model could classify BI-RADS categories and malignant-benign discrimination with an accuracy rate of 86.42% and 92.59%, respectively. However, the CNN-based model showed a precision of 79.01% and 83.95% for the BI-RADS categories and benign malignant discrimination, respectively. These results showed that a well-designed machine learning-based classification model can give better results than a deep learning model. The system [24] achieves classification accuracy, positive predictive value, negative predictive value, and Matthew's correlation coefficient of 84.5%, 84.4%, 94. 8% and 79. 3%, respectively.

According to the paper [25], the classifier model achieves an F1 score of 0.87 and a sensitivity of 0.82. With the addition of suspicious mass detection, sensitivity increases to 0.89, albeit at the expense of a slightly lower F1-score of 0.79.

At the end of the paper [26], the performance of this work is demonstrated by an overall accuracy of 94.22%, an average sensitivity of 95.31%, an average specificity of 99.15% and an area under curve (AUC) of 0.9723. When applied to the screening for breast cancer for Asian women who are more likely to have dense breasts, this model is expected to give a higher accuracy than others in the literature, since it was trained using mammograms taken from Taiwanese women.

Table 7: BI-RAD Classification result from other study with the [23]

ref	task	Dataset	methods	Accuracy %
[7]	Bi-RADs Classification	DDSM	MLP	88.02
	Mass malignancy Classification			83.85

ref	task	Dataset	methods	Accuracy %
[8]	BI-RAD Classification	DDSM	BPNN	84.50
[9]	BI-RAD Classification	a public dataset	DNN	94.22
[10]	BI-RAD Classification	In Breast	CNN	83.40
[11]	Mass malignancy Classification	DDSM	FFnn	98.10 (Sensitivity)
[12]	Mass malignancy Classification	a public dataset	CNN	90.50
[13]	Mass malignancy Classification	DDSM	SVM	80.00
[14]	Microcalsification Classification	a public dataset	SVM	80.00
[15]	Microcalsification Classification	a public dataset	MLP	82.00
[23]	BI-RAD Classification	a public dataset	SVM	86.42
[23]	Mass malignancy classification	a public dataset	SVM	92.59

In the research paper [27], a BI-RADS category was predicted as the output of block-based images segmented from mammogram datasets. The research achieved an accuracy of 94.22%, an average sensitivity of 95.31%, and an average specificity of 99.15%.

According to the paper [23] Classification has been done using a multi-layer perceptron with separate schemes; first, classify masses to distinguish the BI-RADs 2 to 5 and second classify the abnormalities as benign-malignant. The results were encouraging, evaluated on 480 mammographic masses extracted from the digital database for screening mammography. From this research paper, the confusion matrix has been classified as malignant which is positive as BC, and benign negative as BC. There are 121 TN and 13 FP, with 90.30% accuracy in classifying benign, whereas, from the classes of malignant, there are 10 FN and 48 TP, which has 82.76% accuracy from the actual malignant. The accuracy of well classified as 88.02%

According to the paper [28] BC detection methods with KM++SCO have an accuracy of 96.42% with the data sets Mini-MIAS, with DDSA 96.45% and 96.92% with the dataset of BCDR. Gener- ally KM++CS) has an average 96.27% accuracy in the detection of BC. Whereas the research paper, [29] uses k-means clustering for the determination of stages of BC based on the size of the cancerous cells in different steps of the process. In [30] also uses K-means for the diagnosis of BC and it is reliable to identify a malignant from a benign tumor.

III. METHODS

Data Collection: In 2021, Mammography breast image data was collected in DICOM format from the Pioneer Diagnostic Center in Addis Ababa. A total of 4,092 images were obtained from 376 patients. All data are female image and varies age from 30 to 75. The information on the density is also registered on the summary page in parallel with the BI-RAD and recommendations. Organize the BI-RAD with six classes from 0 to 5, and others. Regarding a mammography image view, there are 1249 CC views, and 1109 MLO the rest 1734 are categorized as other views.

Table 8: Local datasets classified by BI-RAD

No	BI-RAD	Number	Remark
1	BI-RAD-0	382	
2	BI-RAD-1	367	
3	BI-RAD-2	314	
4	BI-RAD-3	740	
5	BI-RAD-4	189	
6	BI-RAD-5	77	
7	Other	2023	

Mammography images typically include bilateral craniocaudal (CC) and mediolateral oblique (MLO) views, which are standard for routine screening. The CC view entails the X-ray beam travelling from the head toward the feet, whereas the MLO view involves the X-ray beam moving from the inside to the outside. We have received approval from the diagnostic center to use these data for research purposes. All patient identification names, IDs, and image numbers are removed and changed anonymously.

Data Preprocessing: The collected data was pre-processed to remove unnecessary information and noises from the image. All the information on the summary page given by the senior radiologist is properly organized into categories. Other means in this research do not clearly state the BI-RAD information on the summary page. We used well-categorized image data for training. We have two data sheets, one that has detailed information about the image and the other is the image with the BI-RAD category.

Tools: RadiAnt DICOM Viewer is an application for processing and displaying medical images in DICOM format. We used the RadiAnt DICOM Viewer to read the DICOM format breast images, annotate, archive, export to other formats, and synchronize [31]. In addition to this, we utilized Sante DICOM Editor Tools.

DICOM is the international standard for medical images and related information. It defines the formats for medical images that can be exchanged with the data and quality necessary for clinical use.

Data processing and Presentation: In the process of handling image data, Jupyter Notebook and Python libraries are utilized for tasks such as preprocessing, training, and testing. For image preprocessing, we implement OpenCV, optimization, and feature selection and comparison of the result implemented different algorithms GA, K-means with our GA-K-Means++ model. The test accuracy results are then showcased using Jupyter Notebook and a variety of libraries. This visual representation offers a clear and organized method for data analysis, making it easier to derive meaningful insights from the results.

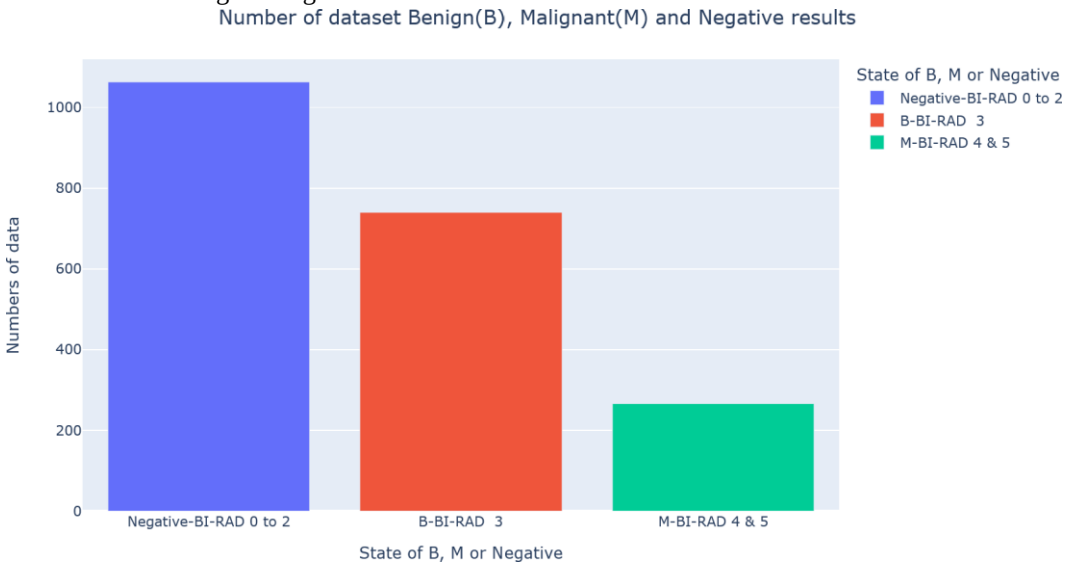

Fig. 1. Mammography Breast image categorized by BI-RAD

In Figure 1, most of the sample data lies in the categories from 0 to 2 which does not need a further diagnosis, whereas 3, 4, and 5 are very small to the total dataset as shown in Figure 2.

```
print(" Total number of CC    " + str(VCC))
print(" Total number of MLO   " + str(VMLO))
print(" Total number of Other View   " + str(otherV))
print(" Total number of High density   " + str(DH))
print(" Total number of Hetrogenous Density " + str(DHT))
print(" Total number of Scattered density   " + str(DSC))
print(" Total number of Extrem " + str(DEX))
print(" Total number of Fatty   " + str(DFT))
print(" Total number of other " + str(otherD))
total_Mammo_View_Data = VCC + VMLO + otherV
total_MassDensity_data = DH + DHT + DSC + DEX + DFT + otherD
print(" Total number of mammography view " + str(total_Mammo_View_Data))
print(" Total number of density data " + str(total_MassDensity_data))
```

```
Total number of CC   1249
Total number of MLO  1109
Total number of Other View  1734
Total number of High density  85
Total number of Hetrogenous Density 4
Total number of Scattered density  499
Total number of Extrem 8
Total number of Fatty  68
Total number of other 3428
Total number of mammography view 4092
Total number of density data 4092
```

Fig. 2. Image data in pre-processing

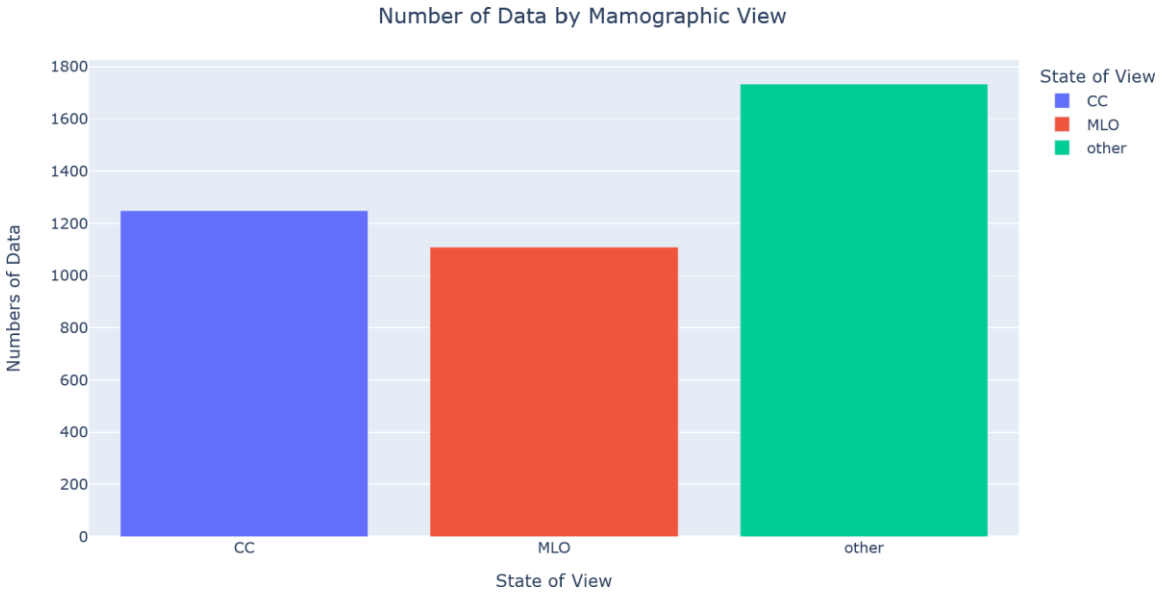

Fig. 3. Mammography data categorized by the image view.

Figure 3 processed data by the views of the image as CC MLO and others. In Figure 4 the image is compressed and processed into grayscale, this helps to minimize the complexity of the image, and optimize the performance and computational resources.

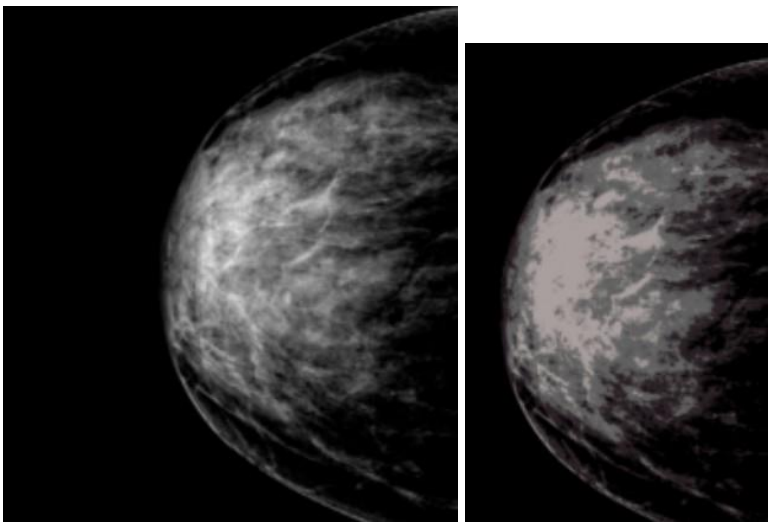

Fig. 4. Converted image into gray scale.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4092 entries, 0 to 4091
Data columns (total 9 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   No                4092 non-null   int64
 1   PID               4092 non-null   object
 2   ImageID           3260 non-null   object
 3   Mammographic_view 2389 non-null   object
 4   Mass_Density      1975 non-null   object
 5   Cal_MB            370 non-null    float64
 6   BI_RAD            2069 non-null   object
 7   Final-Assesment   371 non-null    object
 8   Recommondation    2187 non-null   object
dtypes: float64(1), int64(1), object(7)
memory usage: 287.8+ KB
```

Fig. 5. Processing data information

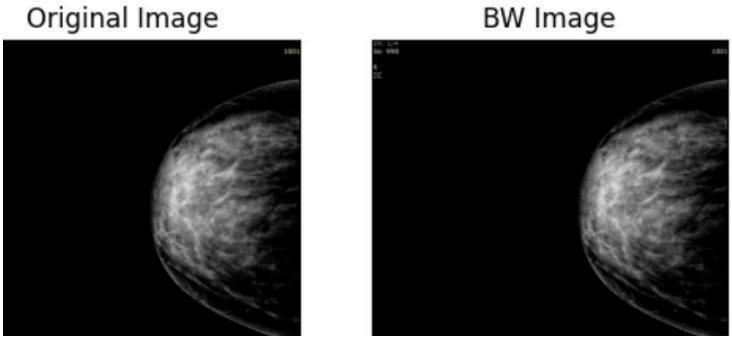

Fig. 6. Result of Original and BW image

Figure 4 prepares the image by converting it into grayscale during the pre-processing. Figure 5 the processed image data information, and Figure 6 comparing the original image and after the gray-scale process. Next to this, Figure 7, we did the process to identify the region of interest in the image data, by removing the noise. This helps to minimize the complexity and enhance the performance during the processing of the model.

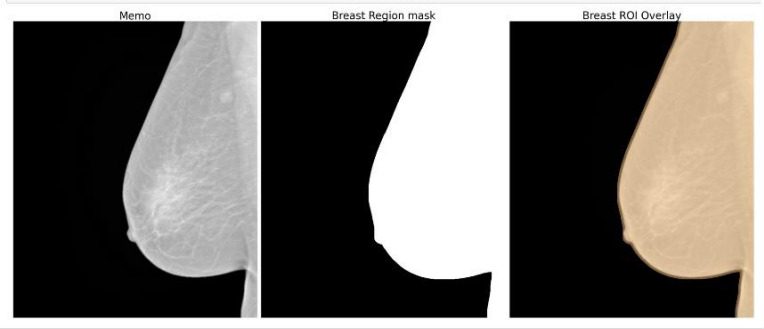

Fig. 7. Result of the image after ROI

Figure 8 shows the number of the BI-RAD data ready for the training by the proposed algorithm. Figure 9 illustrates the image which is ready for training after reprocessing.

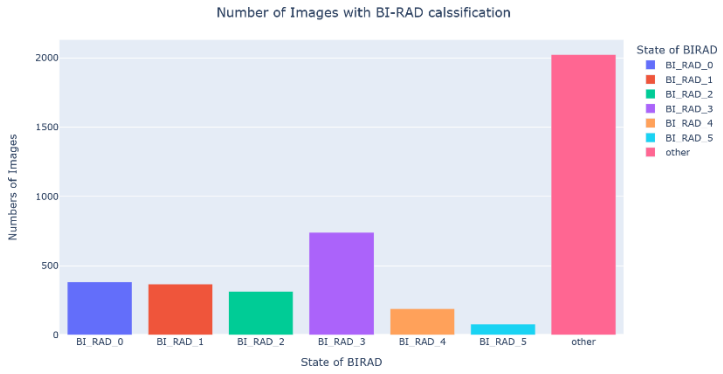

Fig. 8. Displaying the state of BI-RAD Score

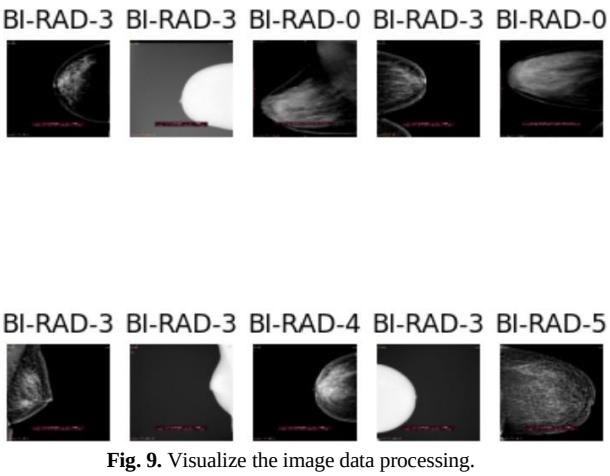

Fig. 9. Visualize the image data processing.

Figure 8 lists the number of image data by BI-RAD classification. Figure 9 describes the mammography breast image data after classifying by the BI-RAD. These data were used for data training to screen the patient using BI-RAD scoring.

Inclusion and Exclusion Criteria: We include all data that contain suggestions and recommendations from senior radiologists as a summary. Of 4948 image data, 95 patient data lacked a summary and have been classified as 'other'. Now we have left 281 patient data, consisting of 3890 images, for the study. Based on the summary note, we have classified the mammography image data using the BI-RAD scale (0-5), as provided by the senior radiologists. BI-RADS are a scoring system used by radiologists to describe mammogram findings.

- BI-RAD-0: Incomplete. If the findings are unclear, more tests are needed. Additional mammograms or ultrasounds may be recommended to get a clearer picture of breast tissue.
- BI-RAD-1: Negative. The breast tissue appears normal, with no masses, calcifications, or abnormalities. Regular monitoring with mammograms is recommended.
- BI-RAD-2: Benign. Although an abnormality (such as scar tissue) can be detected, it is non- cancerous. Continued monitoring with regular mammograms is recommended.
- BI-RAD-3, benign. An abnormality is noncancerous, but follow-up is needed. Another mammogram in six months checks for changes. Fewer than 2% of category three findings develop into cancer.
- BI-RAD-4, Suspicious for malignancy, (2% - 95%), with 4a being low level (2% -9%), 4b inter- mediate (10% - 49%), and 4c highly suspicious,
- BI-RAD-5, Highly suggestive of malignancy
- BI-RAD-6, Known biopsy-proven malignancy.

Data processing Design Model: OpenCV is used for tasks such as image detection, preprocessing, compression, selecting regions of interest (RoI), and data training. The GA algorithm is utilized for feature extraction, and K-Means is employed for training, as shown in Figure 10.

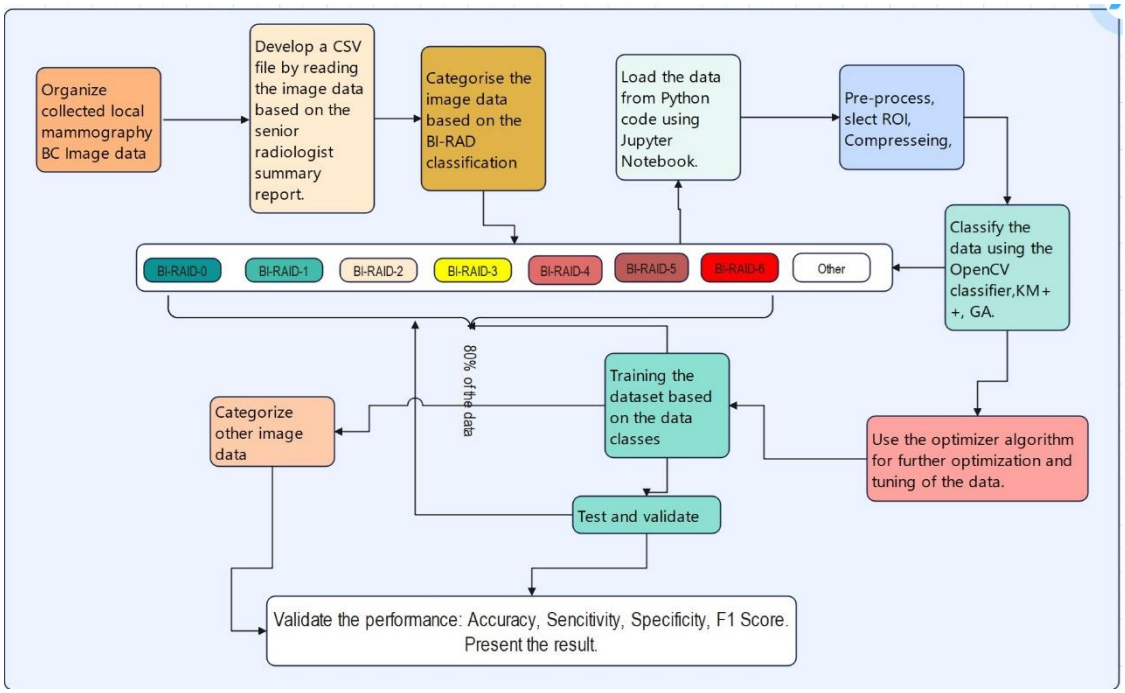

Fig. 10. Data processing

In the Pseudo code, one emphasizes for K-Means clustering, whereas two for the algorithm of the Genetic algorithm employed in the training process of the data. We proposed.

Algorithm 1 Algorithm for K-Means++ clustering

```
Require: Initialize:
 1: Choose the number of clusters ← k. 2:
Randomly initialize ← k centroids.
 3: while convergence or n-number of iterations do
 4:     Assignment step:
 5:     for data point do
 6:         calculate the distance: ← each centroid.
 7:         Assign the data point: ← the nearest centroid; 8:
        end for
 9:     Update step:
10:     for cluster do
11:         calculate new centroid: ← assigned to that cluster;
            {NOTE: calculate the new centroid by taking the mean of all data points;}  12:
            Check for convergence:
13:         if centroids not change significantly or converged then
14:             stop the algorithm.
15:         else
16:             repeat steps ← 4 and 7.
17:         end if
18:     end for
19: end while=0
```

Distance to Cluster Centers: In the genetic algorithm for clustering, the fitness function is calculated according to equation 1.

$$\text{fitness} = \frac{1}{\sum_{k=1}^{N_c} \sum_{j=1}^{N_k} \sqrt{\sum_{i=1}^{F} \left(C_{ki} - P_{jj}\right)^2}}$$

Where:

- N_c: is the number of clusters.
- Nk: is the number of samples within the cluster k.
- F: is the number of features representing the sample.
- Ck: is the center of the cluster k.
- C: the number of clusters

The following steps are followed to calculate the fitness,

- Loop through the cluster centers to calculate the Euclidean distance between all samples and all cluster centers.
- Assign each sample to the cluster with the least Euclidean distance.
- Another loop goes through the clusters to calculate the sum of all the distances in each cluster. If a cluster has zero samples, then its total distance is zero.
- Sum the distances in all clusters.
- Calculate the inverse of the sum of distances. This is the fitness value.

Algorithm 2 A generalized pseudo-code for a genetic algorithm.

Require: t
:= 0;
Require: T
:=N; 1:
start ←
GA() ;
2: *initpopulation* ← *P* (*t*);
initialize random population of individuals;
3: *evaluate* ← *P* (*t*); evaluate fitness of all initial
population; 4: **while** $t \leq T$ **do**
5: $t \leftarrow t + 1$;
6: $P \leftarrow selectparentsP(t)$; {select a sub-population for offspring production.}
7: *recombine* ← *P* (*t*); {recombine the "genes" of selected parents, cross over with
probability.} 8: *mutate* ← *P* (*t*); {perturb the mated population stochastically, mutation with
probability.} 9: $P \leftarrow survive_p(t)$; {select the survivors from actual fitness.}
10: **end while**
11: *return* ← *bestFit*
12: end GA. =0

Algorithm 3 Pseudo code for K-means++ with Genetic Algorithm: (GA-KMeans++)

Require: Initialize Population: Randomly generate an initial population of solutions; 1:
Evaluate Fitness:
2: **for** solution in the population: **do**
3: performs clustering.
4: Calculate the fitness of each solution ← clustering quality
5: sum of squared distances: ← from points to their cluster centers;
6: **end for**
7: **while** best **do**
8: Selection: ← Select a subset of the population based on fitness;
9: Crossover: ← Perform crossover on the selected solutions to create new offspring.
10: Mutation: ← Apply mutation to the offspring to introduce variability.
11: Evaluate Offspring: ← Perform clustering on the offspring ← (Calculate the fitness of each offspring)
12: Replacement: ← Replace the worst-performing solutions in the population with the new offspring. 13:
Termination: ←
14: **if** check termination condition is met **then**
15: terminate the
algorithm 16: **else**
17: repeat steps 8-14
18: **end if**
19: **end while**
20: Output: ← Return the best solution found. =0

Performance measurement to evaluate the performance of a multi-class classification problem, we use the area under the curve (AUC) and the receiver operating characteristics (ROC) curve. The ROC curve shows the trade-off between true positive rate (TPR) and false positive rate (FPR). An AUC value close to one indicates an excellent model with good separability. We aim for high precision and accuracy to ensure better model performance. Low sensitivity and precision make analysis challenging, so we can use the F-1 score, which combines recall and precision into a single metric. Specificity and sensitivity are inversely related. In the breast cancer protection classification process, the quality of the result is measured using performance measurement tools. In this research, we work on:

1. Minimum Square: The minimum square error is calculated as equation 2, for validating the performance of the process of GA, we need to show the graph of the cost training. To be successful or to have the best performance the error should be minimal.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{y_i} - y_i^{*})\ \text{^}2$$

Where y is the true output (actual) and y^* is the predicted result of the model output. MSE can be particularly useful when we have continuous data or regression problems.

2. Performance measures have been performed with the confusion matrix, the F1 score, precision, sensitivity, and specificity.

Based on the confusion matrix for binary classifier different rates are computed. Accuracy defines the Overall classification, it responds, how often is the classifier correct?

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{TPR(Recall)} = \text{Sensitivity} == \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{F1} - \text{Score} = \frac{2TP}{2TP + FP + FN}$$

$$\text{FPR} = 1 - \text{Specificity} = \frac{FP}{TN + FP}$$

IV. RESULTS

In our research, we introduced a novel approach that utilizes k-means optimized with a genetic algorithm (GA) to analyze locally arranged data sets. The aim was to tackle screening challenges in breast cancer diagnosis, by comparing outstanding research results.

Performance of existing research works: The study [23], uses two machine learning models, the first SVM-based model classifies BI-RAD categories and malignant-benign discrimination, with an overall accuracy rate of 86.42% and 92.59% respectively. Whereas using the CNN-based model gets an accuracy of 79.01 and 83.95% respectively. The research was done from 264 mammogram images of 139 patients after working on the data augmentation. The results of shown in Table 9, based on the research work [23], the malignant-benign discrimination SVM is outperformed in the listed performance measurements.

Table 9: Performance metrics of the SVM and CNN algorithm for the discrimination of benign and malignant [23],

Algorithm	**class**	**Accuracy**	**Precision**	**Recall**	**F1-score**
SVM	Benign	92.59	90.60	90.60	90.60
	malignant	92.59	93.90	93.90	93.90
CNN	Benign	83.95	78.80	81.30	80.00
	malignant	83.95	87.50	85.70	86.60

From our locally organized mammography breast image data set, we run in the Genetic algorithm model, K-means++, and a hybrid model of GA with KM++. The hybrid algorithm outperforms the other algorithm. Besides this it has good performance compared to SVM and CNN described in the [23] and shown in Table 10.

Table 10: Performance metrics of the GA, KM++ and GA-KM++ algorithm for the discrimination of benign and malignant

Algorithm	**class**	**Accuracy**	**Precision**	**Recall**	**F1-score**
GA	Benign	93.92	92.55	92.55	93.00
	malignant	93.90	94.50	94.50	94.00
KM++	Benign	93.00	87.60	89.94	91.00
	malignant	93.00	93.50	93.50	93.50
GA-KM++	Benign	95.85	94.80	94.30	95.00
	malignant	95.95	95.50	94.70	95.30

When comparing the benign with malignant discrimination with the algorithm of GA-KM++ more of the malignant outperforms shown in Figure 11.

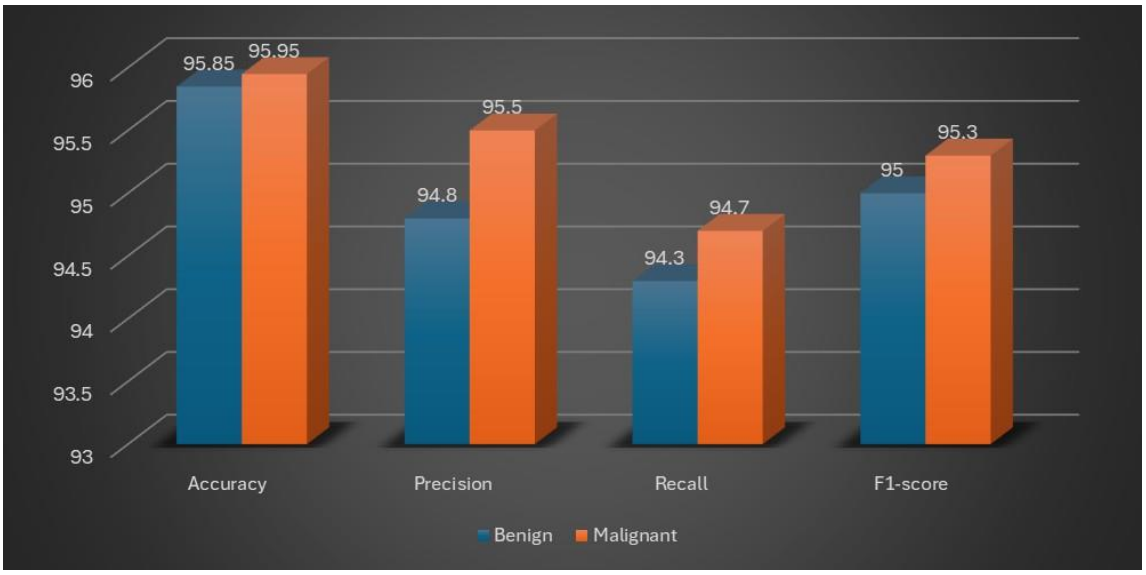

Fig. 11. Classification performance as Benign or Malignant

According to [23], the highest accuracy is registered in the BI-RAD-2 and 5, 95.06 with the SVM algorithm, and the highest F1-score is registered in the BI-RAD-4, as described in Table 11. In the case of CNN, the highest accuracy is registered in BI-RAD 3, 91.36 and F1- Score in BI-RAD 4, as described in Table 12.

Table 11: Performance metrics of the SVM algorithm for classifying as BI-RAD 2,3,4 and 5. [23],

class	Accuracy	Precision	Recall	F1-score
BI-RAD-2	95.06	83.30	83.30	83.30
BI-RAD-3	91.36	70.00	63.60	66.70
BI-RAD-4	91.36	85.30	93.50	89.20
BI-RAD-5	95.06	96.00	88.90	92.30

Table 12: Performance metrics of the CNN algorithm for classifying as BI-RAD 2,3,4 and 5. [23],

class	Accuracy	Precision	Recall	F1-score
BI-RAD-2	88.90	66.70	61.50	64.00
BI-RAD-3	91.36	70.00	63.60	66.70
BI-RAD-4	84.00	79.40	81.80	80.60
BI-RAD-5	93.80	88.00	91.70	89.80

From the algorithms, SVM and CNN the overall accuracy has been registered as 86.42% and 79.01% respectively.

Table 13: performance metrics of the GA algorithm for classifying as BI-RAD from our local data set (LMBCDS)

class	Accuracy	Precision	Recall	F1-score
BI-RAD-0	92.70	86.65	85.60	89.60
BI-RAD-1	92.24	86.45	85.40	89.10
BI-RAD-2	94.28	90.40	84.65	88.45
BI-RAD-3	92.30	90.38	84.49	88.10
BI-RAD-4	90.50	83.54	85.90	85.60
BI-RAD-5	89.39	68.92	75.70	75.30

Table 14: performance metrics of the KM++ algorithm for classifying as BI-RAD from our local data set (LMBCDS)

class	Accuracy	Precision	Recall	F1-score
BI-RAD-0	90.25	85.70	84.90	88.10
BI-RAD-1	89.90	85.70	84.90	88.70
BI-RAD-2	89.25	85.40	83.65	87.80

BI-RAD-3	89.25	85.40	83.60	87.50
BI-RAD-4	86.80	81.24	83.15	84.60
BI-RAD-5	77.10	70.12	78.30	77.00

Table 15: performance metrics of the GA-KM++ algorithm for classifying as BI-RAD from our local data set (LMBCDS)

class	**Accuracy**	**Precision**	**Recall**	**F1-score**
BI-RAD-0	95.67	89.95	87.40	92.80
BI-RAD-1	95.67	89.92	87.36	92.65
BI-RAD-2	96.02	95.34	86.45	92.10
BI-RAD-3	93.86	95.31	86.45	92.00
BI-RAD-4	90.80	86.29	87.00	87.90
BI-RAD-5	89.39	72.43	75.70	77.50

Table 13,14 and 15 show the performance of the dataset with the BI-RAD screening with GA, KM++, and the hybrid model, and the hybrid model has a significant high performance comparing to the individual algorithms. From Figure 12 the highest accuracy is registered in the BI-RAD-2.

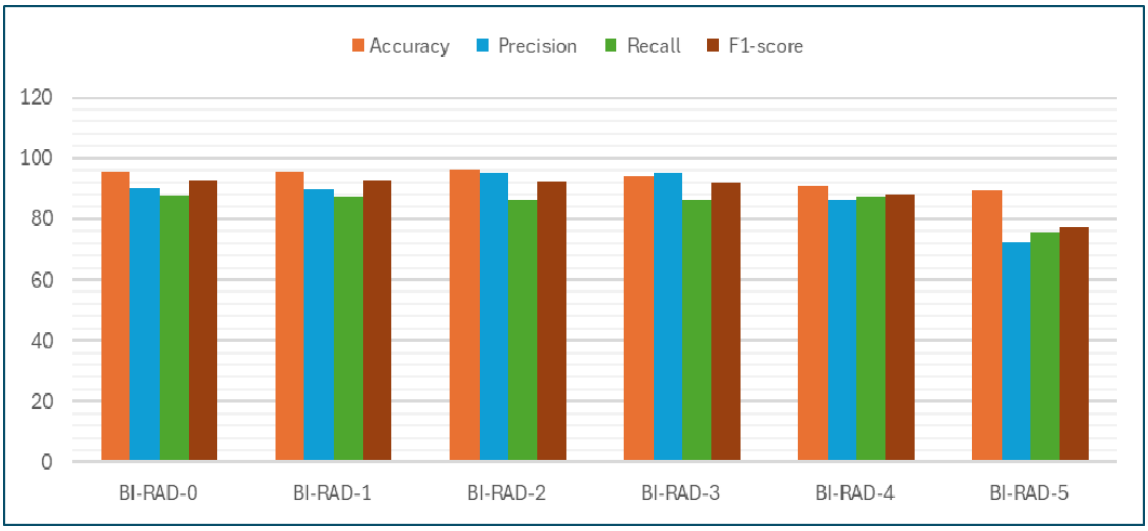

Fig. 12. Screening performance using BI-RAD

In Table 16 SVM and CNN algorithms in the research paper [23] have registered the accuracy of BI-RAD-2 to 5, similarly, we took the BI-RAD2 to 5 with the algorithms GA and the hybrid, the hybrid has the best performance for BI-RAD 2 and 3, whereas SVM has the best performance for BI-RAD 4 and 5. In our case, the data for BI-RAD 4 and 5 are not more than 100, and because of that, the performance becomes less, see Figure 13

Table 16: Accuracy comparison of the algorithms for classifying as BI-RAD 2,3,4 and 5. [23],

class	**SVM**	**CNN**	**GA**	**GA-KM++**
BI-RAD-2	95.06	88.90	94.28	96.02
BI-RAD-3	91.36	91.36	92.30	93.86
BI-RAD-4	91.36	84.00	90.50	90.80
BI-RAD-5	95.06	93.80	89.39	89.39

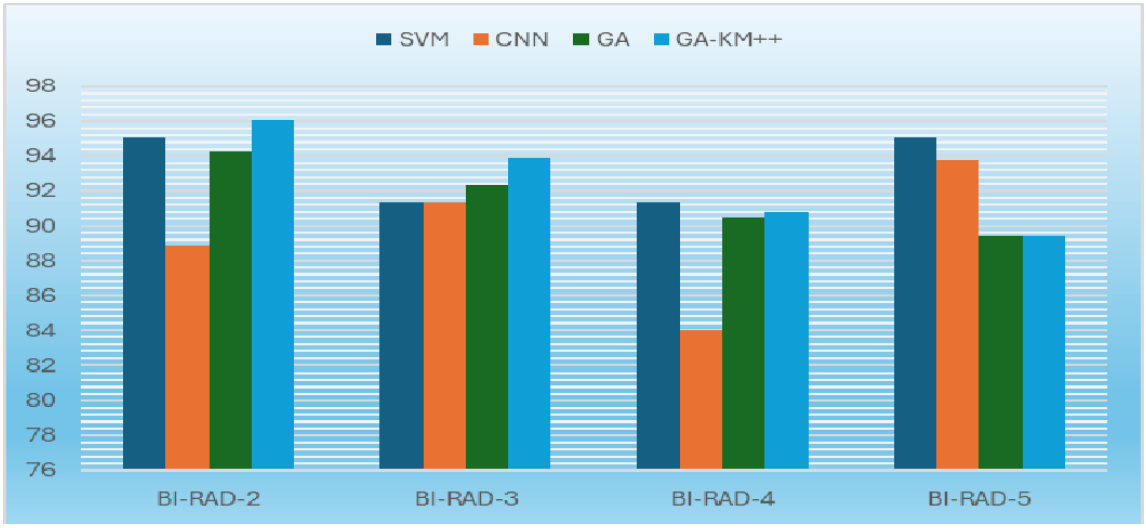

Fig. 13. Screening performance using BI-RAD

In Figure 14 comparison in the algorithms SVM according to [23] and our dataset with the algorithm GA-KM++ has better accuracy in the BI-RAD2 and 3, whereas SVM has better in the BI-RAD 4 and 5.

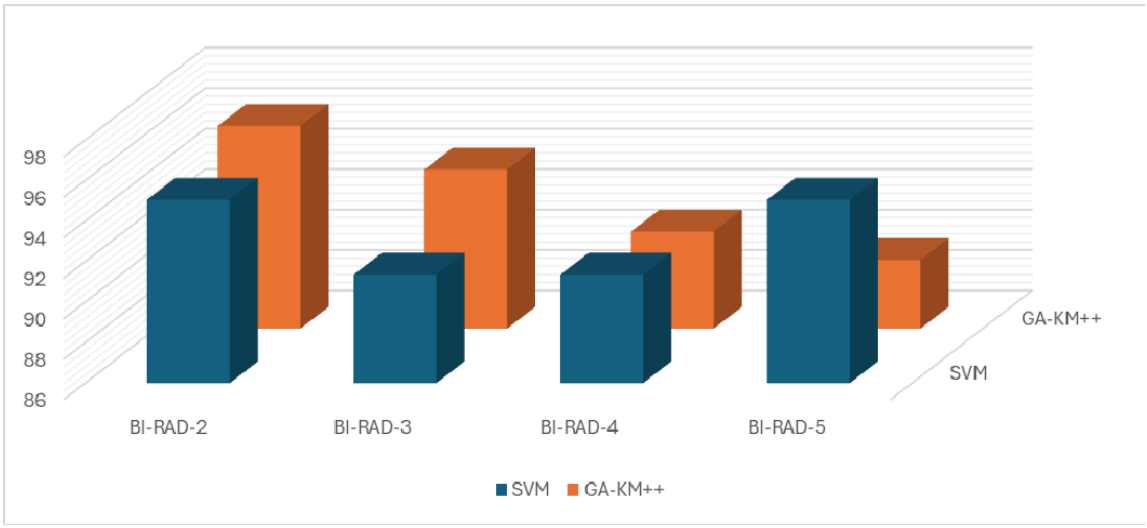

Fig. 14. Screening performance using BI-RAD

By applying datasets MINI-MIAS, BCDR, and DDSM for the algorithms KM++CSO [24] and for the algorithms GA, KM++, and GA-KM++ with our data set and DDSM, the result is shown in Table 17. Using the DDSM data set, our hybrid model has outperformed the KM++ CSO algorithm. However using our dataset, the result is less accurate than the other dataset. This shows that the dataset may need further pre-processing or the quality of the data or the size of the data affects the outcome of the result.

Table 17: Performance Comparison of our model with other work [24]

Method	**Dataset**	**ACC %**	**SEN %**	**SPEC %**
KM++ CSO	MINI-MIAS	96.92	96.78	97.10
	BCDR	96.42	96.90	96.16
	DDSM	95.49	95.68	95.10
GA	DDSM	94.59	95.44	95.20
	Local MBIDS	93.90	94.60	94.90
KM++	DDSM	93.72	95.59	96.26
	Local MBIDS	93.05	94.95	95.02
KM++ GA	DDSM	96.05	96.13	95.94
	Local MBIDS	95.88	95.92	95.67

V. Discussion and Conclusion

Discussion: In our research paper, we have meticulously organized a locally collected mammography image dataset, which comprises a total of 4092 images obtained from 376 patients. Our dataset includes a summary of approximately 2960 data sets categorized as BI-RAD 0 to 5, as per the evaluation by a senior radiologist, with the remaining data sets classified as "other". Furthermore, we have differentiated the data based on image view types, namely CC and MLO, as well as considering the density of the images. Our study involved a comprehensive comparison of results obtained from various methods and models applied to the mammography image dataset. Through our analysis, we observed that the hybrid model we developed exhibited superior performance in screening patients based on their BI-RAD images. However, we believe that further optimization and data cleaning are necessary to enhance the efficacy of our approach. This is due to the varying performance levels observed in the accuracy of screening when employing algorithms such as SVM, GA, and K-Means++. Notably, the GA-K-Means++ model did not outperform the others for BI-RAD 4 and 5, due to the smaller number of image data, which impacted the efficiency of the training process.

Conclusion: The proposed GA-K-Means++ model plays a crucial role in analyzing patients' breast mammography image data. Most of the data observed at the diagnosis center are classified as BI-RAD 1 and 2, indicating no need for further diagnosis and allowing patients to leave the center directly. This model assists radiologists in making informed decisions based on the image data. Additionally, the research can determine whether the image is malignant or benign, providing valuable support to medical doctors in guiding patient consultation. While we currently employ preprocessing tasks such as resizing, grayscale conversion, normalization, and noise reduction, the results suggest that further optimization processes and data cleaning could further enhance our work.

Strengths: Our research presents a robust concept aimed at providing radiologists with the ability to effectively screen and categorize patients into BI-RAD groups 1 to 2, which represent the majority within this classification. It is important to highlight that BI-RAD-0, according to the senior radiologist's comments, was indicated to be incomplete or not easily interpretable. Consequently, our work encourages further exploration of the developed model and dataset to facilitate breast image screening, enabling the identification of image stages and the provision of information to patients before further diagnostic prescriptions are made.

Limitations: One of the limitations of our research lies in the fact that the pre-processing focused primarily on the BI-RAD scope, neglecting other important aspects of the dataset such as density, image view, size, and other relevant features. Additionally, the optimization process for the data was not exhaustive due to limitations in computational resources. As a result, our research is open to further exploration, allowing for potential modifications to the results based on the given dataset and model. Furthermore, we suggest that researchers consider our results as a baseline for future studies.

References

[1] D. P. Acharjya and C. L. Chowdhary, "Breast Cancer Detection Using Hybrid Computational Intelligence Techniques", IGI Global, 2018, pp. 251-280.

[2] WHO,"International Agency for Research of Cancer", WHO, URL:https://gco.iarc.fr/today/home.2022.

[3] J.,Seladi-Schulman, "Breast Cancer Prognosis: Survival Rates by Stage, Age, and Race." Health Line, 2022, https://www.healthline.com/health/breast-cancer/survival-facts-statistics.

[4] https://www.uptodate.com/contents/tumor-node-metastasis-tnm-staging-classification-for-breast- cancer.2024.

[5] https://www.cancer.org/cancer/types/breast-cancer/understanding-a-breast-cancer- diagnosis/stages-of-breast-cancer.html.2024.

[6] https://cancer.ca/en/cancer-information/cancer-types/breast/staging.2024.

[7] Peng Xue, Jiaxu Wang, Dongxu Qin, Huijiao Yan, Yimin Qu, Samuel Seery, Yu Jiang and Youlin Qiao. Deep learning in image-based breast and cervical cancer detection: a systematic review and meta-analysis. nPJ Digital Medicine,2022 www.nature.com/npjdigitalmed.

[8] Shah SM, Khan RA, Arif S, Sajid U. Artificial intelligence for breast cancer analysis: Trends & directions. Comput Biol Med. 2022 Mar;142:105221. doi: 10.1016/j.compbiomed.2022.105221. Epub 2022 Jan 6. PMID: 35016100.

[9] Freeman, K., Geppert, J., Stinton, C., Todkill, D., Johnson, S., Clarke, A., & Taylor-Phillips, S. (2021). Use of artificial intelligence for image analysis in breast cancer screening programs: systematic review of test accuracy. BMJ (Clinical research ed.), 374, n1872. https://doi.org/10.1136/bmj.n1872.

[10] Shigao Huang, J. Y. (2019). Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. Elsevier: https://www.elsevier.com/open-access/userlicense/1.0/.

[11] National Cancer Institute: Surveillance, Epidemiology, and End Results Program. https://seer.cancer.gov/statistics-network/explorer/overview.html

[12] WHO. https://www.who.int/news-room/fact-sheets/detail/breast-cancer. Updated date March 13, 2024. access date Jun 27,2024.
[13] NATIONAL GUIDELINE FOR BREAST HEALTH, EARLY DIAGNOSIS AND TIMELY BREAST CANCER MANAGEMENT IN ETHIOPIA, (2024-2028). APRIL,2024. Addis Ababa, Ethiopia.
[14] Nassif, A. B., Talib, M. A., Nasir, Q., Afadar, Y., & Elgendy, O. (2022). Breast cancer detection using artificial intelligence techniques: A systematic literature review. Artificial intelligence in medicine, Elsevier, Vol 127, 102276. https://doi.org/10.1016/j.artmed.2022.102276.
[15] Maha M. Alshammari, Afnan Almuhanna, and Jamal Alhiyafi. Mammography Image-Based Diagnosis of Breast Cancer Using Machine Learning: A Pilot Study. Sensors 2022, 22, 203. https://doi.org/10.3390/s22010203.
[16] https://www.verywellhealth.com/how-common-breast-cancer-6748850.September,2024.
[17] https://www.breastcancer.org/screening-testing/mammograms/calcifications. July, 2024.
[18] NIH National Cancer Institute. Surveillance, Epidemiology, and End Results program. URL: https://seer.cancer.gov/ Dated: August 2022.
[19] Gastounioti, A., Desai, S., Ahluwalia, V. S., Conant, E. F., & Kontos, D. (2022). Artificial intelligence in mammographic phenotyping of breast cancer risk: a narrative review. Breast cancer research: BCR, 24(1), 14. https://doi.org/10.1186/s13058-022-01509-z.
[20] Karoline Freeman, Julia Geppert, Chris Stinton, Daniel Todkill, Samantha Johnson, Aileen Clarke, Sian Taylor-Phillips, 2021. Use of artificial intelligence for image analysis in breast cancer screening programs: systematic review of test accuracy. BMJ 2021;374:n1872. http://dx.doi.org/10.1136/bmj.n1872.
[21] Kumar, Y., Gupta, S., Singla, R. et al. A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis. Arch Compute Methods Eng 29, 2043–2070 (2022). https://doi.org/10.1007/s11831-021-09648-w
[22] Ruban, S., Jabeer, M., Basti, R.S. (2024). BI-RADS Score Prediction Using AI for Breast Cancer Screening. In: Shukla, S., Sayama, H., Kureethara, J.V., Mishra, D.K. (eds) Data Science and Security. IDSCS 2023. Lecture Notes in Networks and Systems, vol 922. Springer, Singapore. https://doi.org/10.1007/978-981-97-0975-5_18
[23] Fuat Türk, Erkan Akkur and O. Erogul. (2024). BI-RADS CATEGORIES AND BREAST LESIONS CLASSIFICATION OF MAMMOGRAPHIC IMAGES USING ARTIFICIAL INTELLIGENCE DI- AGNOSTIC MODELS. DOI: 10.14311/NNW.2023.33.023
[24] Boumaraf, S., Liu, X., Ferkous, C., & Ma, X. (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, 7695207. https://doi.org/10.1155/2020/7695207
[25] Kebede, S.R., Waldamichael, F.G., Debelee, T.G. et al. Dual view deep learning for enhanced breast cancer screening using mammography. Sci Rep 14, 3839 (2024). https://doi.org/10.1038/s41598-023- 50797-8
[26] Tsai, K. J., Chou, M. C., Li, H. M., Liu, S. T., Hsu, J. H., Yeh, W. C., Hung, C. M., Yeh, C. Y., & Hwang, S. H. (2022). A High-Performance Deep Neural Network Model for BI-RADS Classification of Screening Mammography. Sensors (Basel, Switzerland), 22(3), 1160. https://doi.org/10.3390/s22031160
[27] Chokri, F. and Hayet Farida, M. (2017), Mammographic mass classification according to Bi-RADS lexicon. IET Comput. Vis., 11: 189-198. https://doi.org/10.1049/iet-cvi.2016.0244
[28] Wisaeng K. (2022). Breast Cancer Detection in Mammogram Images Using K-Means++ Clustering Based on Cuckoo Search Optimization. Diagnostics (Basel, Switzerland), 12(12), 3088. https://doi.org/10.3390/diagnostics12123088.
[29] Karmilasari , Suryarini Widodo, Matrissya Hermita, Nur Putri Agustiyani, Yuhilza Hanum and Lussiana ETP, “Sample K-Means Clustering Method for Determining the Stage of Breast Cancer Malignancy Based on Cancer Size on Mammogram Image Basis” Inter- national Journal of Advanced Computer Science and Applications(IJACSA), 5(3), 2014. http://dx.doi.org/10.14569/IJACSA.2014.050312
[30] Noor Salah Hassan, Adnan Mohsin Abdulazeez, Diyar Qader Zeebaree and Dathar A. Hasan. Asian Journal of Research in Computer Science 9(1): 23-38, 2021; Article no.AJRCOS.68284 ISSN: 2581-8260
[31] Medixant. RadiAnt DICOM Viewer [Software]. Version 2024.1. May 4, 2024. (9600) URL: https://www.radiantviewer.com