IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Developing A Predictive Model For Breast Cancer Detection Using 2d Mammography And Deep Learning Techniques

[1]Dr.K. Srinivas, [2]Mohammad Samreen, [3]U. Lakshmi Prasanna , [4]K. Soundharya, [5]K. Hima Bindu

[1]Professor, [2-5]Student

[1-5]Department of Computer Science and Engineering,

Dhanekula Institute of Engineering and Technology, Vijayawada, India

Abstract: Breast Cancer continues to be a major health concern for women worldwide, causing significant illness and death. Early detection is crucial because it increases the chances of successful treatment, making accurate diagnosis vital. In this approach, a hybrid predictive model was developed that leverages deep learning and machine learning techniques. Convolutional Neural Network (CNN) extracts features and captures intricate patterns from mammographic images, while Support Vector Machine (SVM) serves as the classification model to differentiate between Benign and Malignant cases. If malignant, the hybrid model further analyses extracted features to predict the cancer stage, aiding in severity assessment and treatment planning. The model's performance is evaluated using standard metrics, demonstrating superior accuracy and reliability compared to traditional methods. This integrated approach assists radiologists in making more precise and timely diagnoses, ultimately leading to better patient outcomes.

***Keywords* - SVM, CNN, Mammography, Classifiers.**

I. INTRODUCTION

With countless women affected around the globe, Breast Cancer is a significant challenge to health care today and finding its presence early can be vital for diagnosis and treatment. The recent integration of imaging and machine earning has greatly helped improve the effectiveness of diagnostic tools.

In this approach a hybrid predictive model was developed that uses mammography images for analysis to enhance breast cancer detection. The combination of Convolutional Neural Network (CNN) with ResNet50, it is a deep residual network architecture with 50 layers for feature extraction and Support Vector Machine (SVM) for tumor classification will illustrate how complex structures are detected in mammograms that suggest breast cancer.

Beyond classification, the hybrid model is capable of predicting the stage of cancer for tumor with malignant cases, and thus, assists in determining grade of the disease. This approach improves not only the accuracy of the diagnosis but also enables radiologists to make faster and more informed decisions. This model surpasses the conventional technique and helps in the early detection of breast cancer, when treatment is most effective and the outcomes for patients can greatly improve.

II. LITERATURE REVIEW

1. "**Deep Learning for Breast Cancer Diagnosis from Mammograms Tsochatzidis, L., Costaridou, L., & Pratikakis, I. (2019).**" In this study, researchers explored convolutional neural networks (CNNs) for detecting abnormalities in mammograms. The study emphasizes the advantage of CNNs in extracting high-level features directly from image data without requiring manual feature engineering. The model achieved high accuracy, demonstrating its capability in identifying benign and malignant cases. The researchers also highlighted the role of data augmentation and preprocessing in enhancing the model's performance

2. **"A dataset for breast cancer histopathological image classification, A Spanhol, LS Oliveira, C petitjean L Heutte, 2015."** In this paper, researches introduce a dataset of 7909 breast cancer histopathology images acquired on 82 patients, which is now publicly available from http://web.inf.ufpr.br/vri/breast-cancer-database. The dataset includes both benign and malignant images. The task associated with this dataset is the automated classification of these images in two classes, which would be a valuable computer-aided diagnosis tool for the clinician. In order to assess the difficulty of this task, we show some preliminary results obtained with state-of-the-art image classification systems.

3. "**Intelligent Ultrasound Imaging for Enhanced Breast Cancer Diagnosis: EnsembleTransfer Learning Strategies ,2024",** This paper introduces an intelligent breast cancer ultrasound image diagnosis system based on transfer learning and ensemble stacking models. State-of-the-art TL models (VGG-16, VGG-19, Inception V3) are hybridized with Multi-Layer Perceptrons (MLP) and Support Vector Machines (SVM) to achieve improved diagnostic accuracy and stability. The Inception V3 + Stacking yields excellent AUC (0.947) and CA (0.858). Future research may concentrate on enhancing scalability, real-time processing, and generalization using varied datasets and sophisticated methods such as deep reinforcement learning.

4. "**Support-vector networks C Cortes, V Vapnik - Machine learning, 1995",** The support-vector machine (SVM) is a robust learning model used for classifying two groups. It utilizes optimal hyperplanes, non-linear solution surfaces via kernel functions, and soft margins to manage training errors. Although its decision surface may seem straightforward, SVM offers impressive performance and effective capacity control, making it versatile for a range of tasks.

5. "**Large scale deep learning for computer aided detection of mammographic lesions Thijs Kooi a, Geert Litjens a, Bram van Ginneken a, Albert Gubern-Mérida a, Clara I. Sánchez a, Ritse Mann a, Ard den Heeten b, Nico Karssemeijer a (2016)",** In this paper researchers have shown that a deep learning model in the form of a Convolutional Neural Network (CNN) trained on a large data set of mammographic lesions outperforms a state-of-the art system in Computer Aided Detection (CAD) and therefore has great potential to advance the field of research. A major advantage is that the CNN learns from data and does not rely on domain experts, making development easier and faster.

6. "**Improved Breast Cancer Detection using Modified ResNet50-Based on Gradient-Weighted Class Activation Mapping",** This article introduces a novel approach to identifying breast cancer that involves the utilization of a deep learning model utilizing the ResNet50 framework, coupled with heat mapping and gradient weighted class activation mapping (Grad-Cam). The suggested method was primarily assessed using the FDDM dataset of Subtracted Contrast Enhanced Spectral Mammography (CDD-CESM) images. The outcomes from this model were then contrasted with those of five other well-known models: VGG16, VGG19, MobileNetV2.

III. METHODOLOGIES

1. Mammographic Data Collection: The 2D mammography images dataset we use here is taken from the Kaggle, where we have publicly available datasets. This dataset is used to train and test the hybrid model.

2. Data Pre-processing: Pre-processing of mammographic images encompasses various steps towards quality and consistency improvement. These include noise reduction, contrast enhancement, scaling, normalization, and application of data augmentation techniques to enhance the system's stability and provide the best input for feature extraction.

3. Hybrid Model Design: Hybrid model combines feature extraction and classification methods in order to ensure high accuracy in detection as well as staging of breast cancer.

- **Convolutional Neural Network (CNN) with Resnet50 Based Feature Extraction:** A convolutional neural network with Resnet50 is used to learn automatically high-level features from pre-processed mammographic images. It extracts important fine spatial patterns and texture information to distinguish malignant and benign tissues.
- **SVM Classification:** The feature vectors so extracted by the CNN are input to a Support Vector Machine classifier. SVM is used because of its effectiveness in handling high-dimensional data and its ability to come up with robust decision boundaries for accurate classification between cancer and benign cases.

4. Prediction of Cancer Stage: For malignant cases, the hybrid model then proceeds to predict the stage of cancer based on the extracted features. A stage-labelled data is trained with a different classification module to assist in severity evaluation and personalized treatment planning.

5. Model Training: The hybrid model is trained from pre-processed mammographic images and their respective diagnostic labels. Optimization in model performance is done by utilizing techniques such as backpropagation, stochastic gradient descent, and cross-validation.

6. Performance Evaluation: The hybrid model performance is assessed with common metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve.

7. Registration and login:

- The registration module allows users to create an account by providing personal details such as name, email, Password.
- The login module allows user to login by using email and password to access dashboard.

This information is stored in a secure database.

8. Error Handling: In case of errors during image upload or processing, the system generates user friendly error messages

- Invalid File Format: If a non-image file is uploaded, an error message is displayed.
- Processing Errors: If the image cannot be processed (e.g., due to corruption), the system notifies the user and suggests re-uploading the image.

9. Practical Guidance for Medical Experts: The output here we get are used to assist the radiologists so that they can automate the classification process and also save time in manual evaluations.

IV. PROPOSED METHOD

Developing a Predictive Model for Breast Cancer Detection using 2D Mammography Images and Deep Learning techniques can be utilized to enhance the accuracy and reliability of breast cancer detection by integrating the power of Convolutional Neural Networks (CNNs) with the ResNet50 architecture and Support Vector Machines (SVM).

ResNet50, it's one of the highly efficient deep learning models, eliminates and solves issues of vanishing gradient problem in deep networks by using residual connections, enabling the extraction of intricate and relevant features from breast cancer mammographic images. Here the pre-trained ResNet50 model serves as the backbone for feature extraction, which delivers high-dimensional and strong representation to the input images. After feature extraction with ResNet50, the model then applies SVM as the classifier for differentiating malignant and benign cases. If the tumor is tagged as malignant, the model then predicts the stage of the tumor, which will help in assessment of severity as well as in planning treatment. SVM works magnificently on high-dimensional data and is most appropriately suited to implement in medical image processing when high accuracy is crucial and required. The combination of CNN and SVM brings the power of deep feature learning and classification accuracy based on machine learning. This hybrid method not only gives greater accuracy in classification but also allows system to handle then skewed or imbalanced data more effectively, which results in a greater sensitivity and specificity.

V. MODEL DESCRIPTION

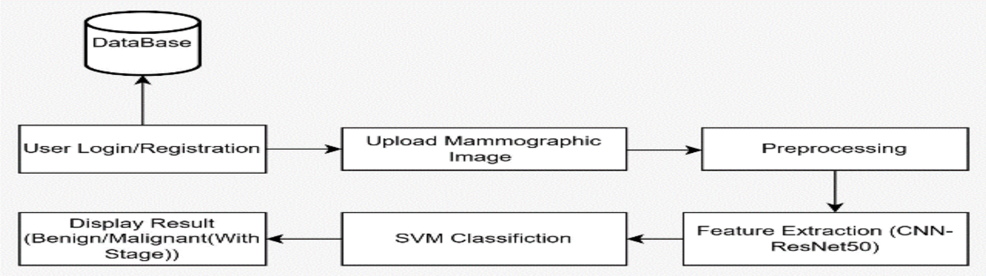

Figure: System Architecture

The system is designed for Breast Cancer detection using mammography images. It begins with user registration and login. Here if user is new, user need to register using email, name, password otherwise he needs to login to access dashboard for analysis. The process starts with uploading mammography image then the image undergoes pre-processing stage of normalization, noise reduction, re-sizing and other enhancement techniques. Next the uploaded images undergo feature extraction through CNN with ResNet50 to capture detailed spatial and texture patterns critical for proper classification. The feature extracted are then goes to SVM classification to differentiate between benign and malignant cases. If the tumor is detected as malignant the system further analysis the features to predict the cancer stage frim 0 to 4, facilitating severity rating and treatment planning. Finally, the outputs are displayed to the users along with graphical representation. Hence, the user data is been monitored by the admin

VI. IMPLEMENTATION

The implementation process for Breast cancer Detection using 2D mammography images integrates CNN with ResNet50 and SVM for image classification, by combining deep learning techniques for feature extraction and classification with a web-based interface using flask, to allow users to upload images and view the results. Here's a step-by-step explanation of the implementation:

1. User Login / Registration.
2. Upload Mammography Image
3. CNN Feature Extraction
4. SVM Classification
5. Result with confidence graph:
   - If Benign
     Display Benign results
   - Else
     Malignant Result with Stage

SUMMARY OF ALGORITHMS USED:

1. **Convolutional Neural Network (CNN):** ResNet50, which is a 50-layer deep residual network, is employed to extract features. It is a pre-trained model utilized to train large image databases. It is capable of extracting useful and high-level pattern or feature from the image.

2. **Support Vector Machine (SVM):** SVM is the supervised machine learning algorithm applied in classification of extracted features by ResNet50. SVM is utilized to identify a best hyperplane to classify varying classes (malignant and benign cases).

DATASET DESCRPTION:

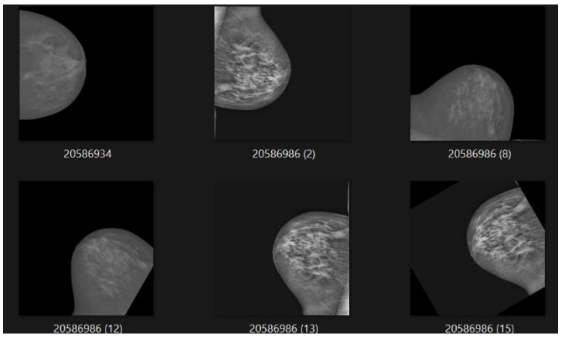

Figure: Sample random images

The dataset used in this study consists of 7,632 files/mammographic images collected from the publicly available dataset in Kaggle. It contains the folders namely test, train, validate with the mammography images in it. During pre-processing each image is pre-processed and resized to 224x224 pixels for compatibility with the ResNet50 model. Here we divided the dataset as 80% training, 10% validation, and 10% testing to ensure robust model evaluation.

VII. RESULTS

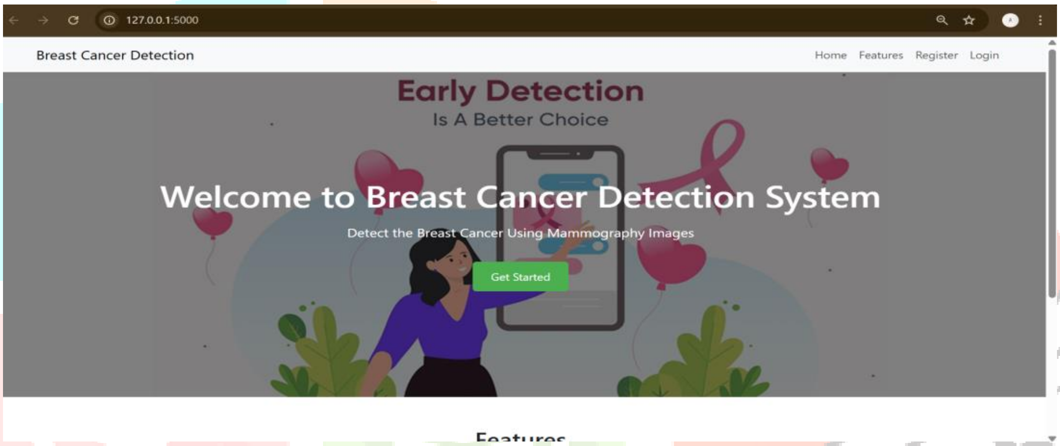

Fig 1: Landing Screen for breast cancer detection. It's a user interface where user can login, register.

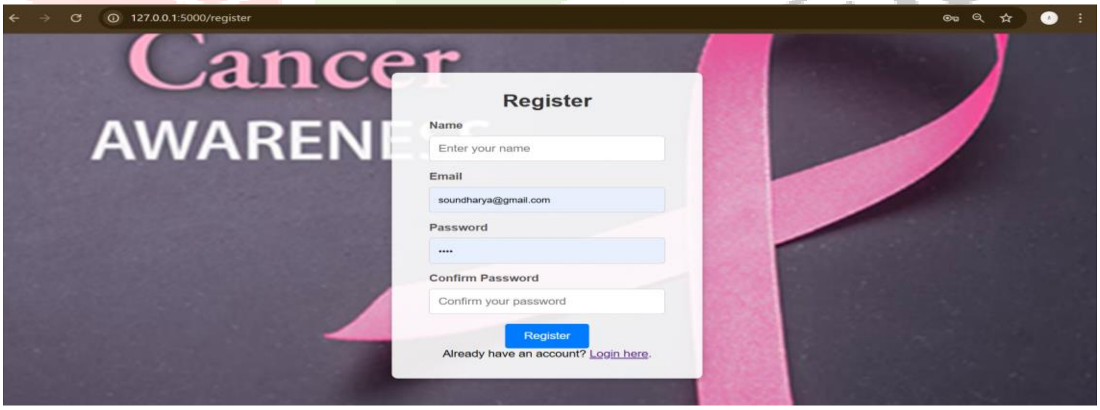

Fig 2: User Registration Interface, where users can register using their personal information and this interface effectively guides users through the account creation process.

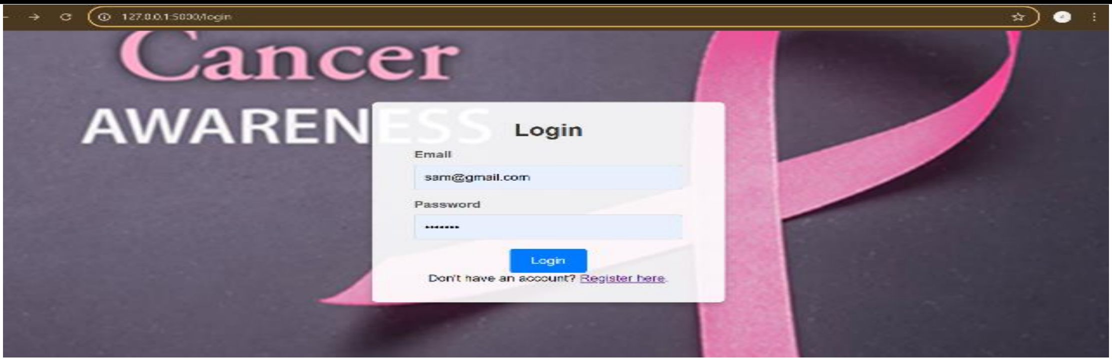

Fig 3: Login Screen, where users can login using their information to access dashboard and this login screen effectively guides users through the account access process.

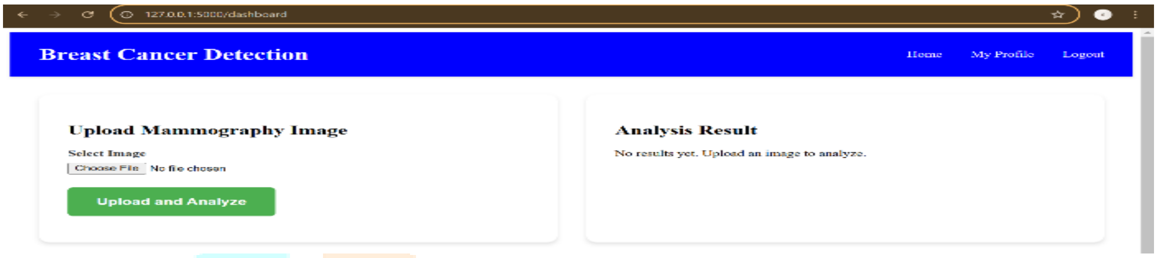

Fig 4: Input Screen where user Upload Mammography Image. This section allows users to upload mammogram images for analysis.

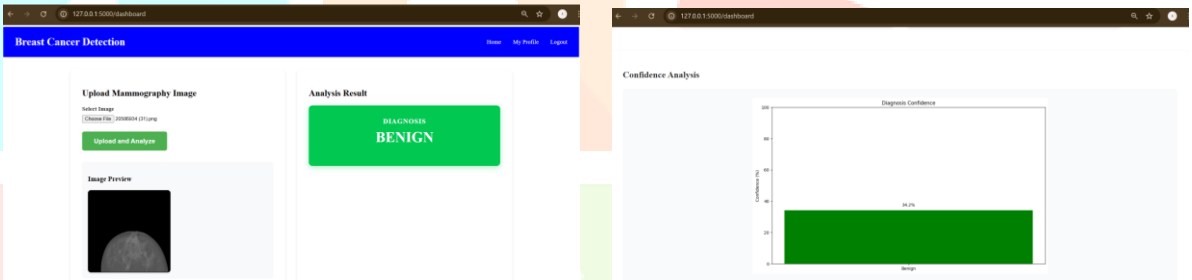

Fig 5: Benign Detection Result. When user upload image, if it is classified as benign this is how the analysis result looks with its diagnosis confidence graph.

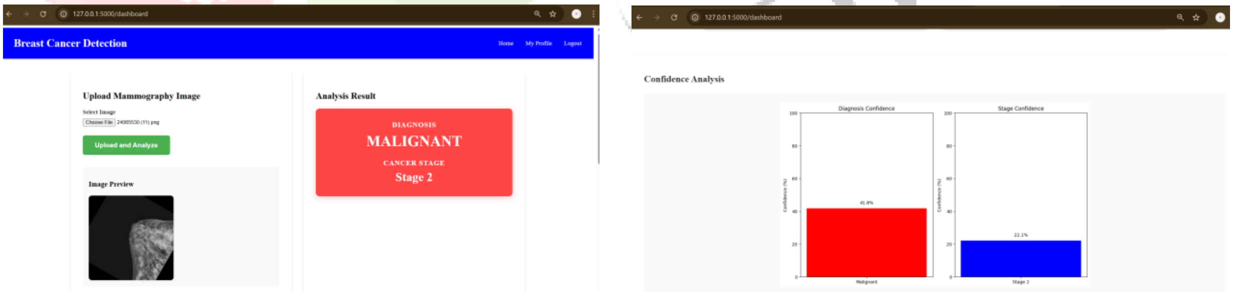

Fig 6: Malignant Detection Result. When user upload image, if it is classified as Malignant this is how the analysis result looks with its confidence analysis graph.

VIII. CONCLUSION

Breast Cancer Detection using 2D mammography images leverages advancements in imaging technologies and machine learning to enhance diagnostic precision. Here we integrate CNN with ResNet50, for complex, intricate features extraction and SVM Classification to distinguish between benign and malignant cases (with stages). By automating the analysis process, this hybrid model serves as a assistive tool for radiologists, reducing workload and more consistent diagnosis.

IX. FUTURE SCOPE

The potential of this project in the future is wide open, with developments in technology and medicine consistently changing. One such direction is the use of state-of-the-art deep learning architectures such as

Vision Transformers (ViT) or EfficientNet that could enhance the robustness and accuracy of breast cancer detection further. Extending the system to evaluate various imaging modalities, e.g., MRI, ultrasound, or 3D mammography, would broaden its diagnostic function and offer richer information. Real-time detection with the power of cloud integration is yet another possibility that would allow medical professionals, especially those located far from city centers, to be able to utilize this technology hassle-free.
The model can be configured to facilitate patient-specific diagnosis using patient-specific information such as genomic information, history, and demographic information. Explainable AI (XAI) techniques can be employed to make the system predictions more transparent, further creating confidence among clinicians through the provision of understandable insight into its classification. Including a big diverse set of data to introduce variability to patient populations and between image quality differences also increases generalizability in the system, resulting in being able to efficiently function on different populations.
Lastly, the future should concentrate on clinical trials and regulatory compliance to prove the efficacy of the system in real-world applications and hasten its implementation in clinical practice. All these developments combined place the project as a revolutionary instrument in the early detection and treatment of breast cancer, leading to improved health outcomes and the potential to save lives.

REFERENCES

[1] Tsochatzidis, L., Costaridou, L., & Pratikakis, I. (2019). Deep learning for breast cancer diagnosis from mammograms—a comparative study. Journal of Imaging, 5(3), 37.

[2] Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2015). A dataset for breast cancer histopathological image classification. Ieee transactions on biomedical engineering, 63(7), 1455-1462.

[3] Rao, K. S., Terlapu, P. V., Jayaram, D., Raju, K. K., Kumar, G. K., Pemula, R., ... & Rakesh, S. (2024). Intelligent ultrasound imaging for enhanced breast cancer diagnosis: Ensemble transfer learning strategies. IEEE Access, 12, 22243-22263.

[4] Kooi, T., Litjens, G., Van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., ... & Karssemeijer, N. (2017). Large scale deep learning for computer aided detection of mammographic lesions. Medical image analysis, 35, 303-312.

[5] Gharaibeh, H., Nahar, K. M., Atitallah, N., Nasayreh, A., & Al, R. E. (2024). Improved Breast Cancer Detection using Modified ResNet50-Based on Gradient-Weighted Class Activation Mapping. Journal of Advanced Research in Applied Sciences and Engineering Technology, 190-208.

[6] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20, 273-297.

[7] Ragab, D. A., Sharkas, M., Marshall, S., & Ren, J. (2019). Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ, 7, e6201.

[8] Sharma, J., Rai, J. K., & Tewari, R. P. (2014, November). Identification of pre-processing technique for enhancement of mammogram images. In 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom) (pp. 115-119). IEEE.

[9] Zhang, Y., Weng, Y., & Lund, J. (2022). Applications of explainable artificial intelligence in diagnosis and surgery. Diagnostics, 12(2), 237.

[10] Nakach, F. Z., Idri, A., & Goceri, E. (2024). A comprehensive investigation of multimodal deep learning fusion strategies for breast cancer classification. Artificial Intelligence Review, 57(12), 327.

[11] Nasser, M., & Yusof, U. K. (2023). Deep learning based methods for breast cancer diagnosis: a systematic review and future direction. Diagnostics, 13(1), 161.

[12] Tan, Y. J., Sim, K. S., & Ting, F. F. (2017, November). Breast cancer detection using convolutional neural networks for mammogram imaging system. In 2017 International Conference on Robotics, Automation and Sciences (ICORAS) (pp. 1-5). IEEE.

[13] Priya, T. S., & Ramaprabha, T. (2021). Resnet based feature extraction with decision tree classifier for classificaton of mammogram images. Turkish Journal of Computer and Mathematics Education, 12(2), 1147-1153.