International Journal of Engineering and Techniques - Volume 12 Issue 4, July - August - 2026

Multi-Modal Cancer Detection Systems: Advances in Early Diagnosis

Nisha A.Wagh[1], Mr. S. G. Shah[2]

[1] *Department of Computer Science & Engineering, Deogiri Institute of Engineering and Management Studies, Chh Sambhajinagar, India*
[2] *Department of Computer Science & Engineering, Deogiri Institute of Engineering and Management Studies, Chh Sambhajinagar, India*
Corresponding author email: nishawagh8560@gmail.com

Abstract — Early and accurate cancer diagnosis is critical for improving treatment outcomes and reducing the mortality rate. This study presents a deep learning-based multimodal cancer detection framework designed for the analysis of ultrasound, MRI, and CT scan images of the breast, liver, and thyroid organs. The proposed system utilizes EfficientNet-based architectures for the binary classification of cancerous and noncancerous medical images. EfficientNetB0 was employed for smaller ultrasound datasets to achieve computational efficiency and reduce overfitting, whereas EfficientNetB3 was utilized for more complex MRI and CT scan datasets to enhance feature extraction capability and classification performance. To improve model generalization, EfficientNetB3 was pretrained on approximately 117,000 medical images across 135 classes before fine-tuning on the target datasets. The framework was implemented using Python and PyTorch, along with supporting libraries, including OpenCV, NumPy, Pillow, and Torchvision. The experimental evaluation demonstrated promising results, achieving accuracies of 94% for breast ultrasound, 90% for liver ultrasound, 85% for thyroid ultrasound, 85% for breast MRI, 93% for liver MRI, and 80% for liver CT scan images. The results indicate that the proposed multimodal approach combined with transfer learning can effectively improve medical image classification and support early cancer diagnosis.

Keywords — Multi-modal Cancer Detection, Deep Learning, EfficientNet, Transfer Learning, Medical Image Analysis, Ultrasound, Magnetic Resonance Imaging (MRI), Computed Tomography (CT).

I. INTRODUCTION

Cancer is one of the most life-threatening diseases worldwide and remains a major concern in modern healthcare systems due to its increasing incidence and mortality rates. According to global health studies, millions of new cancer cases are diagnosed every year, and delayed diagnosis often leads to poor treatment outcomes and reduced survival rates. Therefore, the early identification of cancerous abnormalities is essential for improving patient prognosis, enabling timely treatment, and reducing healthcare costs. Medical imaging technologies have become an important component in cancer diagnosis because they allow physicians and radiologists to visualize internal body structures and identify abnormal tissue growth non-invasively.

Several medical imaging modalities are widely used for cancer detection, including ultrasound imaging, Magnetic Resonance Imaging (MRI), and Computed Tomography (CT) scans. Each imaging technique provides distinct diagnostic information and contributes differently to the disease analysis. Ultrasound imaging is commonly used because of its affordability, portability, and non-ionizing nature, making it suitable for breast, liver, and thyroid examinations. MRI provides high-resolution soft-tissue visualization and is highly effective in identifying tumor boundaries and tissue abnormalities. CT scans generate detailed cross-sectional anatomical images and are extensively used to detect structural changes in organs, such as the liver. Although these imaging modalities are individually effective, relying on a single imaging technique may not always provide sufficient diagnostic accuracy because of variations in image quality, noise, tissue complexity, and disease characteristics.

Traditional cancer diagnosis primarily depends on the manual examination of medical images by radiologists and healthcare experts. This process requires significant clinical experience and may be affected by human fatigue, subjective interpretation, and interobserver variability. Moreover, the increasing volume of medical imaging data has created challenges in terms of processing time and diagnostic consistency. These limitations have encouraged researchers to explore artificial intelligence and deep learning approaches for automated medical image analysis and computer-aided diagnosis systems development.

Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable success in image classification, object detection, and pattern recognition tasks. In the field of medical imaging, deep learning models can automatically learn discriminative features from raw image data without manual feature engineering. This capability makes deep learning highly suitable for identifying complex cancer-related patterns in heterogeneous medical imaging datasets. Recent advancements in transfer learning have further improved the performance of deep learning systems by utilizing pretrained models that are trained on large-scale datasets. Transfer learning helps reduce training time, improve convergence, and enhance generalization capability, especially when medical datasets are limited in size.

Among the various deep learning architectures, EfficientNet has gained considerable attention because of its optimized network-scaling strategy and superior performance with lower computational complexity. The EfficientNet models utilize compound scaling of network depth, width, and image resolution to achieve improved accuracy while maintaining computational efficiency. Therefore, different variants of EfficientNet can be selected based on the dataset size, image complexity, and hardware requirements. Lightweight variants, such as EfficientNetB0, are suitable for smaller datasets and faster training, whereas advanced variants, such as EfficientNetB3, provide enhanced feature extraction capabilities for complex medical images.

This study presents a deep learning-based multimodal cancer detection framework designed for the analysis of ultrasound, MRI, and CT scan images of the breast, liver, and thyroid organs. The proposed system performs binary classification to identify cancerous and non-cancerous medical images using EfficientNetB0 and EfficientNetB3 architectures. EfficientNetB0 was employed for smaller datasets, such as breast ultrasound and liver ultrasound images, to minimize computational complexity and reduce overfitting. EfficientNetB3 was used for thyroid ultrasound, MRI, and CT scan datasets because of its improved representational capability and better performance on complex imaging modalities.

To further enhance model generalization and feature learning, EfficientNetB3 was pretrained on approximately 117,000 medical images belonging to 135 different classes before fine-tuning the target cancer datasets. The framework was implemented using Python and deep learning libraries, including PyTorch, NumPy, OpenCV, Pillow, and Torchvision. The experimental results demonstrated promising classification performance across multiple imaging modalities, highlighting the effectiveness of transfer learning and EfficientNet architectures in medical image analysis.

The primary objective of this study was to develop an efficient and reliable multimodal cancer detection system capable of supporting early diagnosis and assisting healthcare professionals in clinical decision-making. By integrating multiple imaging modalities and optimized deep learning architectures, the proposed framework aims to improve diagnostic accuracy, reduce dependency on manual interpretation, and contribute to the advancement of intelligent healthcare systems.

II. RELATED WORK

New developments in lightweight deep learning have greatly improved how we diagnose breast cancer, especially through analyzing tissue images. A study by Kaushik and Sharma suggests a new MobileNetV3 framework that uses transfer learning to classify tissue samples as benign or malignant. By using pre-trained ImageNet weights and specific data enhancement methods, the model improves feature extraction even with limited data. The results show a decent diagnostic ability, with a ROC-AUC of 0.78 and an accuracy of 0.76, meaning it can fairly well tell apart cancerous from non-cancerous patterns. However, there are some issues like false positives and changes in validation loss that need fixing. Overall, this research shows that compact models like MobileNetV3 can be useful for efficient breast cancer detection, supporting the move towards better AI tools for early cancer screening[1].

Ultrasound imaging is very important for checking for breast cancer because it is safe, easy to access, and good at finding problems in dense breast tissue. A study looked at a new deep learning model called MobileNetV3 for classifying ultrasound images. This model is lightweight and works well for real-time diagnosis. By using transfer learning with MobileNetV3, the researchers improved how features are taken from low-contrast ultrasound images, which can be hard to classify in traditional ways. The model includes ways to adjust and enhance data to work better while using less computing power, making it good for healthcare settings. The method shows strong results in telling apart benign and malignant lesions, showing how effective it is at dealing with ultrasound noise and differences. This study highlights the growing use of smaller CNN models in clinical settings, especially where quick results are important for early breast cancer diagnosis[2].

CNN-based systems are being used more to help find breast cancer, especially in places with limited access to advanced imaging tools. This study shows how CNN models can work with ultrasound images and IoT technology to help diagnose cancer early in areas with fewer resources. By using the BUSI ultrasound dataset, the system can reduce noise, improve image quality, and identify important features to tell apart benign, malignant, and normal breast tissue. The model is very accurate, scoring 99.31%, which is better than older machine learning methods like quadratic SVM. This shows that CNNs are good at managing variations and issues in ultrasound images. An important part of this work is its focus on integrating healthcare with IoT, allowing ultrasound images to be sent remotely and providing real-time support for decisions. This is especially helpful for rural or underserved areas. Overall, this approach shows a move towards smart cancer detection systems that offer quick, automated second opinions and help improve early screening efforts worldwide[3].

Hybrid deep learning models that mix feature extraction from pre-trained convolutional neural networks (CNNs) with traditional machine-learning classifiers are becoming popular for diagnosing breast cancer using ultrasound images. The study compares seven well-known CNNs—VGG-16, VGG-19, ResNet-18, ResNet-50, EfficientNet-B0, MobileNet-V3, and DenseNet—along with a support vector machine (SVM) classifier. This combination helps classify ultrasound images into benign, malignant, and normal categories more accurately. The authors used methods like data augmentation, normalization, and transfer learning to create a system that can handle differences in ultrasound images and the skills of different operators. Their results show that these hybrid models work better than standard CNN classifiers, with EfficientNet B0+SVM achieving the highest accuracy of 98.84% and MobileNet-V3+SVM getting 98.41% on tests. These results show the benefits of using smaller CNNs with SVMs for better performance on limited medical datasets. The study emphasizes that hybrid deep learning systems can improve early breast cancer detection in real clinical settings, especially where resources are limited[4].

Deep learning models are important for improving the accuracy and efficiency of detecting breast cancer using ultrasound. This study looks at several advanced CNN models—MobileNet, MobileNetV2, Xception, VGG16, and VGG19—to see how well they can automatically classify breast ultrasound images as healthy or cancerous. The authors used a clear process that included preparing data, enhancing it, normalizing it, and extracting important features to evaluate how well each model performed based on accuracy, precision, recall, F1-score, and AUC. The Xception model showed the best results with 94% accuracy and an AUC of 0.98, performing better than lighter models like MobileNet and deeper ones like VGG19.The study explains how certain types of convolutions used in Xception and MobileNet help in effectively learning from noisy ultrasound images while being efficient. By comparing different CNN models, the study shows that deep learning methods are becoming more reliable in helping radiologists, reducing mistakes in interpretation, and allowing for earlier diagnosis in places where ultrasound is the main imaging tool[5].

Machine learning tools for predicting liver cancer are becoming popular because they can look at medical data to find patterns linked to cancer. This study uses different machine learning methods to assess signs of liver disease, focusing on techniques like decision trees and support vector machines for early detection. By looking at patient information and lab test results, the model identifies risk levels based on important features. The study points out that simple machine learning methods are still valuable in real healthcare situations where advanced imaging might not be available. While these models can help screen for liver disease quickly and cheaply, their effectiveness can be affected by uneven data, non-imaging factors, and challenges in complex cases. Despite this, the research shows that machine learning is still useful for early liver disease prediction and could be used in systems that support clinical decisions[6].

Deep learning helps to accurately find liver tumors in CT images, especially when the tumors are hard to see. This study introduces a new method that combines a specific image processing technique with two types of neural networks called U-Net and GoogleNet to better identify and classify liver tumors. The researchers used a dataset called 3D-IRCADb01 and applied techniques to improve image quality and increase the amount of data available. The results show that U-Net performs much better than GoogleNet, with a high accuracy rate of 98.65% and excellent results in identifying tumor boundaries. This study shows that using deep learning networks like U-Net is very effective for finding liver tumors in CT scans. These results suggest that advanced models can help doctors diagnose tumors more consistently, lessen their workload, and aid in early detection during real-time use[7].

Multiparametric ultrasound (mpUS) imaging is a promising non-invasive way to detect early liver fat, which can lead to liver cancer. The study uses different ultrasound techniques, like contrast-enhanced ultrasound (CEUS), shear-wave elastography (SWE), and H-scan imaging, to measure changes in liver stiffness and tissue structure. The research was done on rats and showed important changes in a specific measurement that indicates early fat buildup in the liver. Unlike regular ultrasound, mpUS provides a more detailed view of liver health, showing small changes that happen before tumors appear. The study also checks the ultrasound results against tissue samples, proving that mpUS is a good early diagnostic tool. This approach could replace the need for painful liver biopsies and help find liver disease earlier, leading to better screening for people at risk of liver cancer[8].

This paper talks about the important need for good computer tools to help find and classify liver cancer early. It suggests a new method using deep learning to improve this process. The study shows how using different types of information can help in medical analysis, especially since liver tumors can be serious and come in different forms. The method starts by using the Haar wavelet transform to pick out important features while keeping the data simple and small, which helps with speed. Then, it uses the Extreme Learning Machine (ELM) for classification because it trains quickly and is accurate. The GWO algorithm, which is inspired by nature, helps adjust the ELM settings to make it work better. When tested on a diverse dataset, the GWO-ELM method achieved a classification accuracy of 99.41%.This shows that the system is strong and effective, proving that GWO-ELM and Haar wavelet transform work well together for detecting liver tumors accurately[9].

This study aims to improve the accuracy of regular B-mode Ultrasound (BUS) for detecting liver cancer by using information from Contrast-Enhanced Ultrasound (CEUS).While BUS is commonly used, CEUS gives important extra details about blood flow in tissues. The researchers use a method called supervised transfer learning (TL) to share knowledge between these two types of ultrasound. They apply a special tool called Support Vector Machine Plus (SVM+) that works well with paired data. The authors introduce a new version called Nonparallel Hyperplane based SVM+ (NHSVM+) to better connect the information from CEUS to BUS. To handle the different types of CEUS data, they also create a Multi-Kernel Learning (MKL)-based NHSVM+ (MKL NHSVM+) algorithm. This MKL method combines the different types of information more effectively than traditional methods. The final model showed strong results, with an average accuracy of 88.18%, proving that using multi-source transfer learning can improve BUS-based systems for liver cancer detection[10].

This paper talks about a new computer system that helps doctors find thyroid nodules in ultrasound images. It uses a better version of a model called YOLOv5, which is made to be both accurate and fast. The research comes from the growing number of thyroid cancer cases and the challenges doctors face when diagnosing using ultrasound images. The main improvement is adding two new parts to the YOLOv5 model: the Coordinate Attention (CA) module, which helps the model understand where things are in the images, and the Label Smoothing Regularization (LSR) module, which makes the model more reliable even if some data is incorrect. The new system was tested on 191 ultrasound images and showed a high accuracy score of 95.3%.An additional study showed that the new parts improved accuracy by 4.4%.This work shows that the system is very accurate and fast, detecting images in just 8.4 milliseconds, making it a good option for doctors to use[11].

This study looks at how machine learning (ML) can help improve the diagnosis of thyroid cancer, which is currently difficult because it is complex and depends on doctor's experience. The main goal was to create an ML model using a wide range of clinical data. The researchers tested different ML methods, such as Logistic Regression and Decision Trees, and used techniques like Stacking and Voting. They also made sure to clean the data by fixing missing values, changing categories into numbers, and removing incorrect data points. The results showed that their method achieved about 82% accuracy in detecting thyroid cancer. Logistic Regression performed the best among the models tested. The authors believe their findings show that ML can lead to better and more efficient systems for diagnosing thyroid cancer, which can improve patient care[12].

This paper looks at how well a computer can find thyroid nodules using ultrasound images, suggesting that it could be as accurate as a procedure called fine-needle aspiration biopsy (FNAB).The study focuses on using Convolutional Neural Networks (CNNs), which can learn important features from images on their own. It checks how well the system can identify nodules by looking at three key parts: the way predictions are made, different ways to improve results, and how to measure errors. The training of the model includes important steps to prepare the ultrasound images, like making them standard, reducing noise, improving contrast, and adding extra data to ensure consistency and accuracy. The study concludes that the CNN method, tested with real clinical data, could significantly improve the finding and understanding of thyroid nodules in medical settings[13].

This study looks at how to tell apart non-cancerous and cancerous thyroid nodules in ultrasound images, trying to make the diagnosis faster using deep learning. This is important because ultrasound can be tricky to read, and mistakes can happen due to noise and low image quality. The researchers created deep learning models using a type of network called Convolutional Neural Networks (CNNs), specifically testing an improved version called Xception. They added a tool called the Convolutional Block Attention Module (CBAM), which helps the model focus on important parts of the images. They trained the models on a set of 110 non-cancerous and 132 cancerous thyroid cases. The results showed that the original Xception network worked very well, with a high accuracy of 95.058%.However, adding the CBAM did not improve the model's performance. Overall, the study shows that deep learning models can effectively and accurately classify thyroid nodules automatically[14].

This paper talks about using a Convolutional Neural Network (CNN), a type of deep learning model, to make thyroid cancer diagnosis faster and more accurate. The problem it tries to solve is that diagnosing biopsy images takes a lot of time and needs expert analysis. A strong point of this research is that it uses high-quality images of thyroid biopsies. These high-resolution images are better than usual low-resolution ultrasound images because they allow a clearer look at the cells in the thyroid. The CNN model was trained to sort these images into two groups: benign (non-cancerous) and malignant (cancerous).It was tested and got an accuracy of 84%.The authors point out that this level of performance, along with good sensitivity and specificity, shows that the model could speed up the diagnosis process, leading to quicker and better treatment for cancer patients[15].

This paper looks at using Support Vector Machine (SVM) algorithms to classify three types of cancer: lung, breast, and liver, using MRI scans. The study focuses on improving the accuracy of cancer diagnosis by using machine learning to analyze medical images. The main method is SVM, a strong tool that works well with complex data, to tell apart cancerous and non-cancerous tissues. The research follows a careful process of taking images, preparing them, extracting important features, and training the SVM model. The main aim is to accurately classify different cancer types using the same imaging method. The results show that the SVM approach can be effective for cancer detection, and the authors believe this method could help bring advanced machine learning into cancer diagnosis for better and earlier results[16].

This research aims to create effective prediction models using machine learning to detect three common types of cancer: prostate, lung, and breast cancer, early on. Finding cancer early is important for better patient care and lower death rates. The study uses different machine learning methods, like Logistic Regression, Random Forest, and Gradient Boosting, to examine data that includes clinical and lifestyle factors. A key part of the research is handling the data well, including using a method called SMOTE to fix class imbalance, which is a frequent problem in medical data. Each model's performance is carefully tested using measures like accuracy, precision, and recall. The results show that these machine learning models are good at predicting these three cancers, which could help doctors with early screening and treatment[17].

III. METHODOLOGY

A. Proposed Framework

The proposed framework is a deep learning-based multi-modal cancer detection system developed to classify cancerous and non-cancerous medical images obtained from different imaging modalities. The framework analyzes ultrasound, MRI, and CT scan images of breast, liver, and thyroid organs using transfer learning with EfficientNet architectures.

The proposed methodology consists of the following stages:

1. Medical image acquisition
2. Image preprocessing
3. Dataset preparation
4. Transfer learning using EfficientNet architectures
5. Model training and validation
6. Binary classification
7. Performance evaluation

B. Dataset Collection

The proposed framework utilizes medical images collected from publicly available datasets. Three imaging modalities were considered to increase the robustness and applicability of the proposed system.

The datasets include:

- Ultrasound Images
  - Breast
  - Liver
  - Thyroid
- MRI Images
  - Breast
  - Liver
- CT Scan Images
  - Liver

Each dataset consists of two classes:

- Cancerous
- Non-Cancerous

The datasets were divided into training, validation, and testing subsets to evaluate the model under unseen conditions.

C. Image Preprocessing

Medical images obtained from different datasets often vary in resolution, intensity distribution, and imaging quality. Therefore, preprocessing was performed before model training to improve image consistency and feature extraction.

The preprocessing pipeline includes:

- Image resizing
- Pixel normalization
- Conversion to RGB format (where required)
- Data augmentation
- Tensor conversion

Image augmentation was applied only to the training dataset to increase data diversity and reduce overfitting. The augmentation techniques include:

- Random horizontal flip
- Random rotation
- Random cropping
- Brightness adjustment
- Random affine transformation

All preprocessing operations were implemented using the Torchvision library.

D. Transfer Learning

Instead of training a deep neural network from scratch, transfer learning was adopted to reduce training time and improve model generalization.

EfficientNetB3 was initially pretrained on approximately 117,000 medical images belonging to 135 different classes. The pretrained weights enabled the network to learn generalized medical image features before fine-tuning on the target cancer datasets.

The pretrained model was subsequently fine-tuned separately for each imaging modality to improve classification accuracy.

E. EfficientNet-Based Feature Extraction

The proposed framework employs two variants of the EfficientNet architecture according to dataset size and image complexity.

EfficientNetB0 :-

EfficientNetB0 was selected for:

- Breast ultrasound images
- Liver ultrasound images

These datasets contain relatively fewer training samples. Therefore, EfficientNetB0 provides an effective balance between computational efficiency and classification accuracy while minimizing the risk of overfitting.

EfficientNetB3 :-

EfficientNetB3 was selected for:

- Thyroid ultrasound images
- Breast MRI images
- Liver MRI images
- Liver CT scan images

These datasets contain more complex anatomical structures requiring richer feature representations. EfficientNetB3 offers improved feature extraction capability through increased network depth, width, and image resolution.

F. Model Training

The extracted deep features were learned through supervised training using binary labels representing cancerous and non-cancerous classes.

The models were implemented using Python and PyTorch. Binary Cross Entropy with Logits Loss (BCEWithLogitsLoss) was employed as the loss function for binary classification.

The network parameters were optimized using the Adam optimizer, while transfer learning significantly reduced convergence time.

G. Classification

After feature extraction, the fully connected classification layer predicts whether the input image belongs to the cancerous or non-cancerous category.

The classification process was independently performed for all imaging modalities:

- Breast Ultrasound
- Liver Ultrasound
- Thyroid Ultrasound
- Breast MRI
- Liver MRI
- Liver CT Scan

This modality-specific learning strategy enables the framework to effectively capture unique imaging characteristics associated with each organ.

H. Performance Evaluation

The effectiveness of the proposed framework was evaluated using classification accuracy on the test datasets.

The obtained accuracies are summarized below:

Imaging Modality	Organ	Model	Accuracy
Ultrasound	Breast	EfficientNetB0	94%
Ultrasound	Liver	EfficientNetB0	90%
Ultrasound	Thyroid	EfficientNetB3	85%
MRI	Breast	EfficientNetB3	85%
MRI	Liver	EfficientNetB3	93%
CT Scan	Liver	EfficientNetB3	80%

The experimental results demonstrate that selecting EfficientNet architectures according to dataset characteristics improves classification performance while maintaining computational efficiency across multiple medical imaging modalities.

IV. RESULTS AND DISCUSSION

The proposed multi-modal cancer detection framework was evaluated using ultrasound, MRI, and CT scan datasets of breast, liver, and thyroid organs. The experimental analysis was conducted to measure the effectiveness of EfficientNetB0 and EfficientNetB3 architectures in binary classification of cancerous and non-cancerous medical images. The models were implemented using PyTorch and trained using transfer learning techniques to improve feature extraction and classification performance across multiple imaging modalities.

EfficientNetB0 was utilized for smaller ultrasound datasets including breast and liver ultrasound images due to its lightweight architecture and reduced computational complexity. EfficientNetB3 was employed for thyroid ultrasound, MRI, and CT scan datasets because of its enhanced representational capability and superior feature extraction performance for complex medical images. To improve generalization capability, EfficientNetB3 was pretrained on approximately 117,000 medical images belonging to 135 different classes before fine-tuning on the target datasets.

The experimental results demonstrated promising classification performance across all imaging modalities. The breast ultrasound dataset achieved the highest accuracy of 94%, indicating the effectiveness of EfficientNetB0 in extracting discriminative features from ultrasound images. Liver MRI classification achieved 93% accuracy using EfficientNetB3, demonstrating the capability of deeper EfficientNet architectures in handling high-resolution and complex imaging data. Liver ultrasound images achieved 90% accuracy, while thyroid ultrasound and breast MRI datasets both achieved 85% classification accuracy. The liver CT scan dataset obtained an accuracy of 80%, which may be attributed to higher image variability, complex lesion patterns, and lower soft tissue contrast in CT imaging.

The obtained results indicate that transfer learning significantly improved model convergence and feature generalization across different datasets. The adaptive selection of EfficientNet architectures based on dataset complexity also contributed to improved computational efficiency and reduced overfitting. The proposed framework demonstrated the ability to effectively classify cancerous and non-cancerous images from heterogeneous medical imaging modalities, highlighting the potential of deep learning-based systems in supporting early cancer diagnosis.

Table 1 presents the classification accuracy achieved across different imaging modalities and organs.

Sr. No.	Mode	Type	Accuracy
1	Ultrasound	Breast	94%
2	Ultrasound	Liver	90%
3	Ultrasound	Thyroid	85%
4	MRI	Breast	85%
5	MRI	Liver	93%
6	CT Scan	Liver	80%

Overall, the experimental findings demonstrate that the proposed multi-modal framework provides reliable classification performance across multiple cancer imaging modalities. The combination of transfer learning, EfficientNet architectures, and multi-modal medical imaging contributes to improved diagnostic accuracy and supports the development of intelligent computer-aided cancer diagnosis systems for early detection.

V. CONCLUSION AND FUTURE WORK

This research presented a deep learning-based multi-modal cancer detection framework using ultrasound, MRI, and CT scan images for the early diagnosis of breast, liver, and thyroid cancers. The proposed system utilized EfficientNetB0 and EfficientNetB3 architectures to perform binary classification of cancerous and non-cancerous medical images across multiple imaging modalities. By integrating transfer learning with EfficientNet models, the framework achieved effective feature extraction and improved classification performance on heterogeneous medical datasets.

Experimental results demonstrated promising accuracy across different modalities, with breast ultrasound and liver MRI datasets achieving the highest classification performance. The adaptive selection of EfficientNet architectures based on dataset size and image complexity contributed to improved computational efficiency and reduced overfitting. Furthermore, pretraining EfficientNetB3 on a large-scale medical image dataset enhanced model generalization capability and supported reliable performance across diverse imaging conditions.

The findings of this study indicate that multi-modal deep learning systems can significantly support early cancer diagnosis and assist healthcare professionals in medical image interpretation. The proposed framework highlights the potential of artificial intelligence in developing accurate, scalable, and clinically supportive computer-aided diagnostic systems. Future work can focus on integrating additional imaging modalities, explainable AI techniques, and larger standardized datasets to further improve the robustness and real-world applicability of intelligent cancer detection systems.

VI. DECLARATIONS

A. Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

B. Conflict of Interest

The authors declare no conflict of interest.

C. Data Availability

The dataset used in this study is publicly available through the Kaggle Cancer Detection Dataset. Additional materials and implementation details are available from the corresponding author upon reasonable request.

D. Ethics Statement

This study did not involve human participants, animals, or any identifiable personal data. Therefore, ethical approval was not required.

E. Author Contributions

Conceptualization, Nisha Wagh; Methodology, Nisha Wagh; Software, Nisha Wagh; Validation, Nisha Wagh; Investigation, Nisha Wagh; Writing — original draft, Nisha Wagh; Writing — review and editing, Nisha Wagh; Supervision, Mr. S. G. Shah . All authors have read and agreed to the published version of the manuscript.

VII. ACKNOWLEDGMENTS

The authors sincerely thank the Department of Computer Science & Engineering, Deogiri Institute of Engineering & Management Studies, Chh. Sambhajinagar, India, for providing the infrastructure, resources, and academic support required for this research. The authors are especially grateful to Prof. S. G. Shah for his valuable guidance, constructive suggestions, and continuous encouragement throughout the study. The authors also acknowledge the contributors of the publicly available Cancer Detection Dataset and the open-source software community whose resources supported the successful completion of this work.

VIII. REFERENCES

[1] P. Kaushik and P. Sharma, "A MobileNetV3-Based Approach to Breast Cancer Detection Through Transfer Learning and Histopathological Image Analysis," 2024 International Conference on Advances in Computing,

Communication and Materials (ICACCM), Dehradun, India, 2024, pp. 1-6, doi: 10.1109/ICACCM61117.2024.11059095.

[2] P. Kaushik and P. Sharma, "Breast Cancer Detection using MobileNetV3: A Deep Learning Approach for Ultrasound Image Classification," 2025 Fourth International Conference on Power, Control and Computing Technologies (ICPC2T), Raipur, India, 2025, pp. 1-5, doi: 10.1109/ICPC2T63847.2025.10958661.

[3] A. K. Dwivedi, A. Yadav and R. S. Gamad, "Breast Cancer Disease Prediction using Convolutional Neural Networks with Ultrasound Image Dataset for Healthcare-IOT Application," 2025 4th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0, Raigarh, India, 2025, pp. 1 6, doi: 10.1109/OTCON65728.2025.11070903.

[4] M. I. Rosadi, C. Fatichah and A. Yuniarti, "Breast Cancer Ultrasound Images Classification Using Hybrid Pre-Trained CNN and SVM," 2025 17th International Conference on Knowledge and Smart Technology (KST), Bangkok, Thailand, 2025, pp. 231-236, doi: 10.1109/KST65016.2025.11003377.

[5] R. Sivakumar, "Deep Learning-Based Automated Breast Cancer Ultrasound Image Classification: A Study," 2025 Eleventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2025, pp. 1-5, doi: 10.1109/ICBSII65145.2025.11013993.

[6] F. I. Mir, G. Raj, K. Lata and A. S. Verma, "Deep Data Analysis for Liver Cancer Prediction Using Machine Learning Approaches," 2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N), Greater Noida, India, 2024, pp. 864-870, doi: 10.1109/ICAC2N63387.2024.10895272.

[7] Z. Naaqvi, S. Akbar, S. A. Hassan and Q. Ul Ain, "Detection of Liver Cancer through Computed Tomography Images using Deep Convolutional Neural Networks," 2022 2nd International Conference on Digital Futures and Transformative Technologies (ICoDT2), Rawalpindi, Pakistan, 10.1109/ICoDT255437.2022.9787429. 2022, pp. 1-6, doi:

[8] L. Basavarajappa, J. Li, H. Tai, J. Song, K. J. Parker and K. Hoyt, "Early Detection of Liver Steatosis using Multiparametric Ultrasound Imaging," 2021 IEEE International Ultrasonics Symposium (IUS), 10.1109/IUS52206.2021.9593500. Xi'an, China, 2021, pp. 1-4, doi:

[9] W. G. Negassa, S. Mishra and H. D. Bizuneh, "Liver Tumor Detection and Classification Using GWO-ELM Model," 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT), Bhubaneswar, India, 2023, pp. 526 531, doi: 10.1109/APSIT58554.2023.10201687.

[10] H. Zhang et al., "Multi-Source Transfer Learning Via Multi-Kernel Support Vector Machine Plus for B-Mode Ultrasound-Based Computer-Aided Diagnosis of Liver Cancers," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 10, pp. 3874-3885, Oct. 2021, doi: 10.1109/JBHI.2021.3073812.

[11] D. Yang et al., "Automatic Thyroid Nodule Detection in Ultrasound Imaging With Improved YOLOv5 Neural Network," in IEEE Access, vol. 12, pp. 22662-22670, 2024, doi: 10.1109/ACCESS.2024.3359367.

[12] T. A. Vu, N. A. Huyen, H. Q. Huy and P. T. V. Huong, "Enhancing Thyroid Cancer Detection Through Machine Learning Approach," 2023 12th International Conference on Control, Automation and Information Sciences (ICCAIS), Hanoi, Vietnam, 2023, pp. 188 193, doi: 10.1109/ICCAIS59597.2023.10382297.

[13] C. Kaushik Viknesh, S. Kanimozhi and R. Thirumalai Selvi, "Investigation of Thyroid Nodule Detection Using Ultrasound Images with Deep Learning," 2024 Tenth International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2024, pp. 1-7, doi: 10.1109/ICBSII61384.2024.10564034.

[14] Z. İLKILIÇ AYTAÇ, İ. İŞERİ and B. DANDIL, "Thyroid Cancer Detection Using Convolutional Neural Networks," 2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI), Zarqa, Jordan, 2023, pp. 1-4, doi: 10.1109/EICEEAI60672.2023.10590377.

[15] X. Yang and G. Zhu, "Ultrasound-based discrimination of thyroid cancer based on deep learning," 2024 5th International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), Shenzhen, China, 2024, pp. 939-943, doi: 10.1109/AIEA62095.2024.10692904.

[16] S. R, S. R. Navaneethakrishnan, R. B, K. S. Balamurugan and M. Sudhagar, "Machine Learning in Oncology: SVM-Based Classification of Lung, Breast and Liver Cancer from MRI Scans," 2024 IEEE 9th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Bahrain, Bahrain, 2024, pp. 1-6, doi: 10.1109/ICETAS62372.2024.11119794.

[17] C. Rokde, P. Kharwandikar, A. Kungwani and A. R. Dudhe, "Predictive Modeling for Early Cancer Detection: A Machine Learning Approach to Prostate, Lung, and Breast Cancer," 2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI), Wardha, India, 2024, pp. 1-6, doi: 10.1109/IDICAIEI61867.2024.10842741.