International Journal of Computing and Digital Systems
2025, VOL. 18, NO. 1, 1-17

http://dx.doi.org/10.12785/ijcds/1571120148

MammoFusionNet: A Unified Approach for Breast Lesion Classification and Segmentation Using Modified Involutional Neural Networks

Md. Fahim-Ul-Islam[1] **and Jia Uddin**[2]

[1]*Department of Computer Science and Engineering, Brac University, Dhaka, Bangladesh*
[2]*Department of AI and Big Data, Endicott College, Woosong University, Daejeon 34606, South Korea*

Received 12 February 2025, Revised 30 April 2025, Accepted 1 May 2025

Abstract: Breast cancer impacts more than 1.5 million people worldwide annually, with early detection essential for differentiating between malignant and benign lesions. The precision of diagnostics frequently relies on the proficiency of healthcare practitioners, resulting in possible misdiagnoses. Consequently, the development of efficient computer-aided procedures is imperative for enhancing early detection and decreasing death rates. Although numerous research concentrate on lesion identification from ultrasound images, they often tackle classification or segmentation independently. To address this issue, we present MammoFusionNet, a multi-task learning architecture that simultaneously executes the classification and segmentation of breast lesions. MammoFusionNet integrates a modified Involution Neural Network (INN) into a UNet architecture, enhancing its flexibility to many imaging modalities and resolutions. To validate the model, we used the Breast Ultrasound Images Dataset (BUSI) which consists of two classes: Benign and Lesion. MammoFusionNet exceeds the performance of leading models such as InceptionResNetV2, EfficientNetB7, and ResNet50 in classification tasks, attaining 99.12% accuracy, 98% AUC, and a 99% F1-score. Additionally, it surpasses segmentation models like Swin-Unet and DeepLabV3, achieving an 82% IOU, 90% mean IOU, and a 75% Dice score. The integration of classification and segmentation tasks improves overall performance, establishing MammoFusionNet as an effective instrument for breast cancer diagnosis in clinical environments.

Keywords: Breast cancer, Ultrasonography, Tumor classification, Lesion Segmentation, Involutional Neural Network, Multi-Tasking

1. Introduction

Breast cancer is the second most lethal kind of cancer following lung cancer, with a comparatively high mortality rate. The World Health Organization (WHO) reports that breast cancer impacts about 1.5 million individuals globally each year [2], [1], [3]. It is oldest known kinds of cancer, was originally detected in Egypt in roughly 1600 BC [4]. It could come from a benign or malignant, which are two different types of lesions. To identify these tumors, medical practitioners need an active assessment method [13]. However, even for experts, it is sometimes fairly difficult to recognize tumors [5]. The majority of breast cancer patients do not identify they have the disease and eventually die before receiving the correct treatment. Therefore, early detection of breast cancer is vital to minimizing the fatality rate.

Visual assessment of radiographic (X-ray) pictures is one of the most successful early detection tools for breast cancer (known as mammograms) [14]. The extensive use of screening mammography is encouraged for the early identification of breast cancer. According to research, the incorporation of ultrasonography can double the overall rate of cancer diagnosis and minimize the number of needless biopsies [6], [7]. Tumors that are unseen on mammograms in breasts with dense tissue can be discovered utilizing ultrasonic imaging [10]. Moreover, Breast ultrasound (BUS) imaging is the most widely utilized technology for diagnosis and categorization of early breast lesions due to its cost, real-time analysis, patient safety, and non-invasive nature [11], [16]. However, the ultrasound pictures created by BUS must be analyzed by radiologists with skill and training. Consequently, it is imperative to implement computer systems that assist novice radiologists in image processing to detect breast cancer, as erroneous classification of breast lesions can pose significant risks to patients. Nonetheless, alterations in tissue characteristics, including shape, grey level, dimensions, intensities, and positioning, render the

E-mail address: fahim.islam@bracu.ac.bd, jia.uddin@wsu.ac.kr

classification of normal, benign, and cancerous tissues a formidable endeavor when utilizing this method. Additionally, eye inspection has a limited specificity and a high sensitivity. Ultrasound screening can be applied in conjunction to imaging technology to potentially minimize the quantity of needless biopsies.

The rapid breakthroughs in machine learning, particularly deep learning, have greatly heightened the medical community's interest in harnessing these approaches to boost the accuracy of cancer detection through imaging [8]. Machine learning holds the potential to assist healthcare practitioners in the early diagnosis of malignant tumors [17]. Despite its advantages, the use of these approaches in cancer screening is associated by high hazards of false positives and false negatives [22].

Deep learning networks utilize a multi-layered structure that facilitates hierarchical learning and the incremental extraction of features from input, beginning with simple patterns and progressing into more complicated abstractions [9]. By autonomously recording a complete range of information, deep learning algorithms may produce highly precise outcomes, making them exceptionally useful for classifying medical images and recognizing characteristics like lesions [23], [24]. Zakareya et al. combined GoogLeNet and residual blocks with granular computing and attention processes, obtaining 93% accuracy on ultrasound and 95% on histopathological images [20]. Martinez et al. discussed by combining heterogeneous data and feature selection, deep learning increases prescreening accuracy and minimizes false negatives compared to older methods [12]. Abunasser et al. proposed BCCNN model classifies breast cancer into eight kinds using advanced deep learning techniques to enhance diagnosis from X-ray, MRI, and CT images [15].

The performance of convolutional neural networks (CNNs) confronts various restrictions, and popular models generally aim to boost their accuracy by increasing the depth of convolutional layers [21]. However, the spatial diversity of relevant regions in an image offers issues for determining a suitable kernel size for convolution processes, as the information may be spread at varied scales. A smaller kernel is appropriate for more localized information, whereas a larger kernel is preferable for broader distributions. Deep networks are more sensitive to overfitting, and guaranteeing efficient gradient updates throughout the network becomes problematic, given the increasing computational expense of stacking lengthy convolution layers [18], [19].

To address these challenges, MammoFusionNet is introduced as a powerful multi-task learning network engineered to concurrently address classification and segmentation tasks in the analysis of breast cancer lesions. MammoFusionNet utilizes a modified Invariant Neural Network integrated into the UNet framework to improve adaptability and efficacy in intricate medical image processing. It tackles the difficulty of creating an operation that is both location-specific and channel-agnostic by integrating the INN layer, which generates kernels conditioned on particular spatial positions. This method facilitates the efficient processing of input tensors with changing resolutions, enhancing the model's adaptability to various imaging modalities and resolutions.

MammoFusionNet enhances accuracy and resilience by concurrently training classification and segmentation tasks inside a cohesive framework, using the synergy between them through shared feature maps. INN layers' dynamic adjustment of receptive fields offers enhanced flexibility in feature extraction relative to traditional convolutional layers. The collaborative refinement of classification and segmentation branches during training improves overall performance, rendering MammoFusionNet an effective instrument for clinical applications in breast cancer diagnosis.

Primary contributions of this article are stated as below:

- A modified INN model is proposed to identify if an input ultrasonography image has a lesion or not.
- An augmentation strategy has been proposed that proved to be effective and increased the performance of the model.
- A multitasking learning approach has been implemented for dynamically utilizing both classification and segmentation tasks.

Section II describes the existing works in our classification and segmentation tasks in the medical domain for breast cancer disease. Section II. Furthermore, Section III depicts the data collection and augmentation and extraction methods applied to our study before the training stage. Besides, Section IV describes the methodology of our study integrating classification and segmentation methods of our MammoFuisonNet model. Besides, in Section V, we delve into the experimental setup for the training SOTA models with our MammoFuisonNet model while integrating evaluation metrics for our study. Finally, in Section VI we arrive at our manuscript's conclusion.

2. Related Works

There have been multiple researches on breast cancer lesion detection from diverse angles. This section highlights a few recent works on breast cancer lesion detection.

Rouhi et al. introduced two automated algorithms for discriminating between benign and malignant cancers in mammography. The first approach utilizes an autonomous region-growing method, where the level of significance is selected by an effectively trained neural network model. The second method combines a cellular neural network (CNN) for segmentation, with genetic algorithms (GA) improving the network parameters [25].

TABLE I. Overview of Breast Cancer Classification Approaches.

Study	Dataset(s)	Classification Task	Performance Metrics	Multi-tasking
Rouhi et al. [25]	MIAS, DDSM	Breast Cancer Classification (Binary Classification)	Acc: 96.47%, Sensitivity: 96.87%, Specificity: 95.94%	No
Aljuaid et al. [26]	Breast-Cancer-Database	Benign vs Malignant Binary Classification and Multiclass Breast Cancer Classification	Accuracy: 99.7%, 97.66%, and 96.94% (Binary), 97.81%, 96.07%, and 95.79% (Multiclass)	No
Abdur et al. [27]	Wisconsin Breast Cancer	Breast Cancer Classification	Accuracy: 99.3%, 98.06%, 97.35%	No
Khan et al. [29]	MIAS (109 cases)	Normal vs Mass Classification in Mammograms	Accuracy: 68%	No
Abdel-Nasser et al. [30]	Mini-MIAS, INBreast	Breast Density and Mass vs Normal Classification	Accuracy: 99%, AUC: 0.9325	No
Shan et al. [31]	283 Private Cases	Benign vs Lesion Classification	SVM: Accuracy: 77.7%, AUC: 84; RF: Accuracy: 78.5%, AUC: 83	No
Pratiwi et al. [32]	MIAS	Normal vs Abnormal and Benign vs Malignant Classification	Normal/Abnormal: Accuracy: 93.98%, Sensitivity: 97.22%; Benign/Malignant: Accuracy: 94.29%, Sensitivity: 100%	No
Adebiyi et al. [33]	Wisconsin Breast Cancer Dataset	Benign vs Lesion Classification (Binary)	96.4% (SVM with LDA), 95.6% (RF with LDA)	No
Our Proposed Approach	BUSI dataset	Benign vs Lesion Classification (Binary)	99.12% accuracy, 98% AUC, 99% F1-score, 82% IOU, 90% mean IOU, 75% Dice.	Yes

Aljuaid et al. established a methodology for categorizing mammary tumor into several categories relying on a convolutional neural network (CNN) model. The approach that is suggested tries to categorize breast cancers into subclasses other than benign or malignant, such as lobular carcinoma, fibroadenoma, etc. [26].

Abdur et al. did a study to investigate which classifier worked best in breast cancer diagnosis, using data visualization and generating four forecasting models: quadratic SVM, logistic regression (LR), k-nearest neighbors (KNN), and ensemble classification (EC). These strategies were created to boost the accuracy of breast cancer diagnosis. The suggested models revealed an accuracy of up to 99.3% on the benchmark dataset, surpassing the results of past studies and delivering substantial insights for possible clinical applications [27].

Houssein et al. offered an analysis that emphasizes current advancements in the identification and categorization of breast cancer using machine learning and deep learning technology. This study revealed how breast cancer was classified. Following a review of alternative neural network algorithms and specific structures for the identification and categorization of cancer of the breast, it next offers an overview of several machine learning processes. We also present a quick description of the main picture modalities to provide a comprehensive grasp of the topic [28].

Khan et al. assessed the efficacy of six alternative approaches of guided feature extraction in the categorization of tumors in digital scans. These methods employed Gabor filter banks to extract directional textural information, revealing the geometrical features of both big quantities and normal cells at various directions and intervals. However, these ROIs sometimes included normal tissues along with masses, leading to false positives throughout the classification process [29].

Shan et al. largely introduced important computational components from the Breast Imaging Reporting and Data System (BI-RADS) to construct a computer-aided diagnostic (CAD) system for breast ultrasonography. A collection of 283 pathology-verified benign and malignant tumors was utilized to create and examine computational characteristics that match ultrasonography BI-RADs classifications [31] .

Abdel-Nasser et al. advised the deployment of a high-resolution method which takes advantage of the complementing data offered by several pictures of the identical item. The recommended CAD method has four phases: high-resolution calculation, region of interest extraction, extraction of attributes, and categorization. Researchers tested the capabilities of five substance techniques with the suggested CAD structure which is the histogram of directed gradients, the local binary patterns, the phase congruency-based localized digital pattern, the levels of

gray relationship framework features, and the shape of the lacunarity spectrum. They demonstrate that the performance of the textural methodologies under consideration is increased by our super-resolution-based methodology, which outperforms the state-of-the-art for identifying benign from malignant tumors [30].

The implementation of RBFNN for mammography classification with Gray-level Co-occurrence Matrix (GLCM) texture-based characteristics is detailed in Pratiwi et al.'s work. According to the computational testing, RBFNN performs breast cancer classification. Their investigation findings showed that the accuracy of RBFNN is 93.98% for normal and abnormal diagnosis, 14% more than BPNN, and 94.29% for benign and malignant classification, 2% better than BPNN [32].

Adebiyi et al. employed supervised learning methods, random forest (RF,) and support vector machine (SVM), paired with linear discriminant analysis (LDA) for feature extraction, to boost breast cancer diagnostic accuracy [33].

Table I depicts the existing works and their method analyzed to further work on our study.

3. Data Acquisition and Preparation

To construct a deep learning healthcare system, it is necessary to have access to a well-structured dataset. In this context, we exploited the Breast Ultrasound Images Dataset (BUSI) [35], an openly accessible resource. The BUSI dataset features ultrasound pictures of breast tissue obtained from women aged 25 to 75. The images were obtained in 2018 utilizing the LOGIQ E9 ultrasound system and the LOGIQ E9 Agile ultrasound equipment at Baheya Hospital for Early Detection and Treatment of Women's Cancer in Cairo, Egypt. The dataset includes 600 female patients and is divided into three groups: benign, malignant, and normal. It contains a total of 1,578 ultrasound images, with an average resolution of 500×500 pixels. Specifically, the collection comprises 266 photos categorized as normal, 891 images as benign, and 421 images as malignant. Figure 1 depicts the sample of photos from dataset BUSI. Table II represents the minimum and maximum height and width of each class.

TABLE II. DATASET BUSI: DATA DESCRIPTION

Class	Qty	W-Min	W-Max	H-Min	H-Max
Normal	266	391	709	391	709
Benign	891	273	719	324	719
Malignant	421	287	719	351	719

1) Data Augmentation

The augmentation techniques discussed in the article have been employed for classifying lesions. We evaluated at numerous data augmentation techniques that have been applied subsequently in various research. Understanding the effects of data augmentation and determining if one type of augmentation is more beneficial than another are our primary priorities. After an extensive studies of previously employed augmentation technique and few trail runs we finalized on our data augmentation technique, that proved to be very efficient.

The techniques we utilized are presented below in Table III:

Data augmentation plays a crucial role in enhancing the robustness and generalization ability of deep learning models, particularly in medical image analysis, where annotated data is often limited. By artificially increasing the size and diversity of the dataset, augmentation techniques help the model learn more generalized features, reducing overfitting and improving its ability to handle unseen variations in new data. This is especially important in medical imaging, where variations in lesion shape, size, orientation, and lighting conditions are common.

TABLE III. Data Augmentation Parameters

Parameter	Value
RandRotation	[-15, 15]
RandScale	[0.80, 1.20]
RandXShear	[-25, 25]
RandYShear	[-25, 25]
RandXTranslation	[-12, 12]
RandYTranslation	[-12, 12]
Zoom Range	0.20
Fill Mode	"reflect"

4. Methodology

In this study, we present a novel approach to breast cancer lesion segmentation and classification tasks in our multi-tasking MammoFusionNet model. Our input images are fed into the MammoFusionNet model that is intended for cancer segmentation and classification where we used INN as the backbone of our model. The architecture of the MammoFusionNet builds on the U-Net framework and is structured into distinct segments: the encoding path, the classification head, the decoding path, and the segmentation head. Figure 2 depicts our proposed MammoFuisonNet model used in our study.

A. Involutional Neural Network

Involution presents a new operation that is spatially-specific and channel-agnostic, unlike conventional convolution, which is spatially-agnostic and channel-specific [36]. Involution dynamically produces kernels for each spatial position according to the input feature map, enabling the process to adjust to diverse visual patterns. The kernel generation process entails a bottleneck architecture consisting of two linear transformations, wherein W_0 and W_1 are

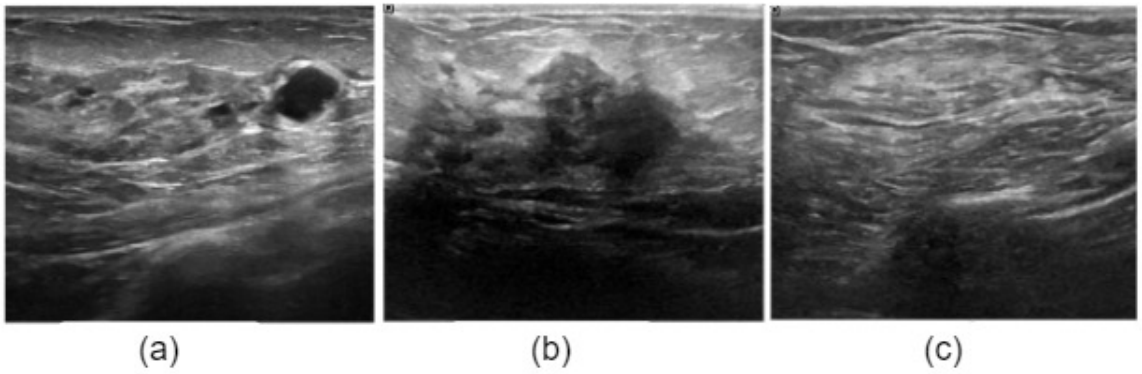

Figure 1. The BUSI dataset contains three classes: (a) for Benign, (b) for Malignant, and (c) for Normal.

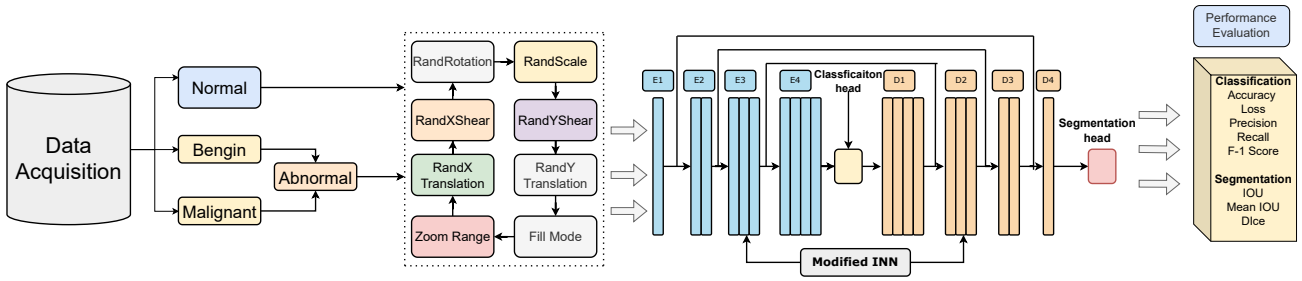

Figure 2. The Overall Workflow

utilized on the local feature vector at each spatial location, followed by the application of an activation function such as Batch Normalization or a non-linear activation σ. These dynamically generated kernels are uniformly distributed across all channels but differ across spatial positions. By consolidating channels to utilize a common involution kernel, the procedure minimizes redundancy and enhances efficiency. Upon generation, the kernels are applied to the input feature maps by a multiply-add operation, thereby convolving the kernels with the input feature map at each spatial location. Figure 4 depicts architecture of our utilized INN for encoding and decoding block our MammmoFuison-Net architecture.

B. Encoding Path

The architecture's encoding element utilizes downsampling blocks that include involution layers, batch normalization, and Leaky ReLU activation functions. These blocks gather high-level characteristics by gradually reducing the spatial dimensions of the input feature maps while simultaneously increasing their depth. INN layers are utilized for their capacity to effectively collect comprehensive context on a global scale, using fewer parameters in comparison to traditional convolutional layers. Each downsampling block comprises an INN layer, which is subsequently followed by batch normalization and Leaky ReLU activation. This arrangement facilitates the extraction of discriminative features. The output feature maps generated during the encoding path are subsequently utilized by the classification and segmentation heads for further processing. This hierarchical feature extraction technique allows the model to effectively collect and depict essential information required for cancer lesion categorization. The initial number of feature maps is 3, and it doubles with each downsampling step. However, to efficiently utilize computational resources, the maximum number of feature maps is limited to 64.

C. Classification head

The classification head, located at the bottom layer of the architectural framework, consists of a convolutional block followed by global average pooling, effectively accommodating inputs of different sizes. Subsequent layers comprise fully connected layers, reinforced with Leaky Rectified Linear Unit (ReLU) activation, culminating in a softmax layer, thereby producing assessment probabilities relevant to fire categorization. In this classification job, the INN Layer efficiently incorporates global context, consequently boosting knowledge of the encompassing setting, a critical component for distinguishing varied levels of input photos. Moreover, its parameter efficiency stands out, since it conducts global context aggregation with less parameters compared to normal convolutional layers [36]. This not only increases the model's performance but also mitigates the risk of overfitting, assuring solid predictions during both the training and testing phases.

D. Decoding Path

The decoding pipeline of the Inv-Net is characterized by the employment of upsampling blocks integrating two INN layers, which are vital in recovering feature maps to

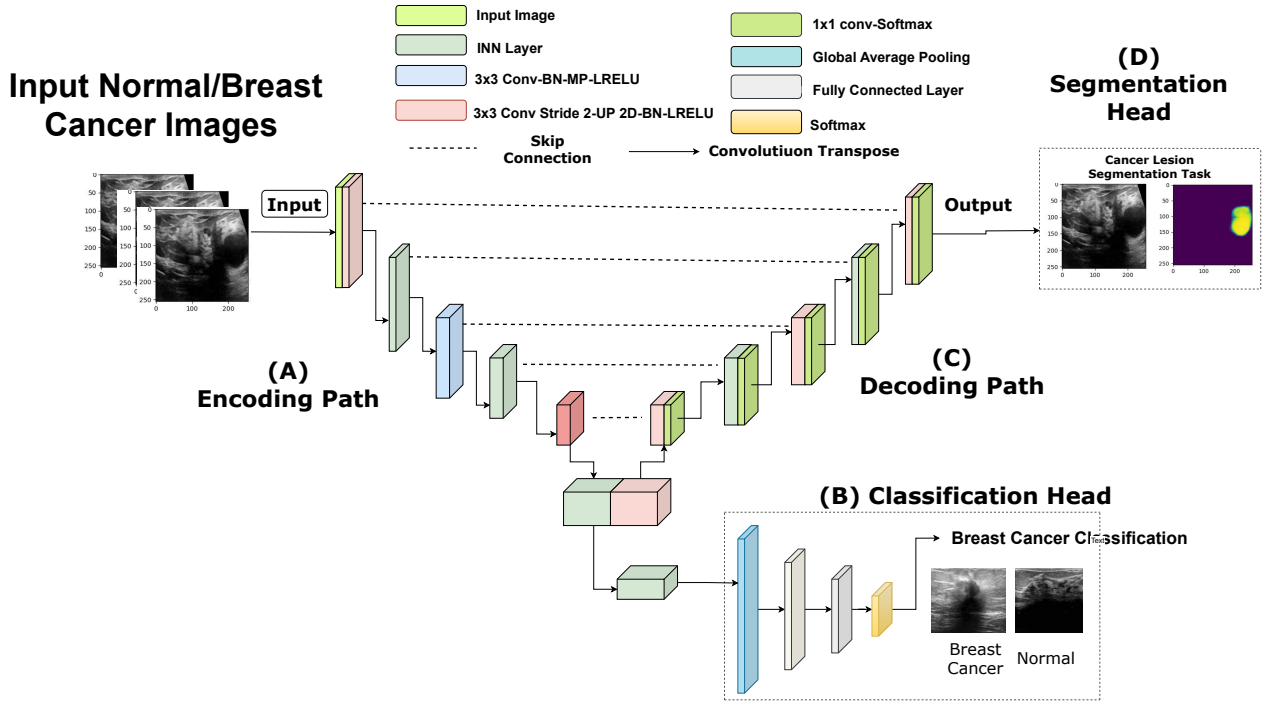

Figure 3. Our Proposed MammoFuisonNet Model

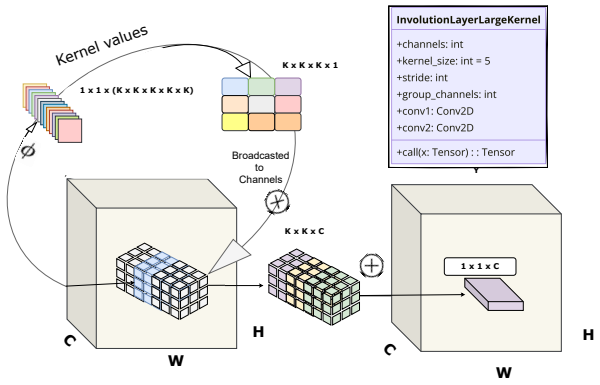

Figure 4. Architecture of Involutional Neural Networks

their original input dimensions. Each block is defined by the insertion of two INN layers alongside batch normalization and Leaky Rectified Linear Unit (ReLU) activation. Intervening between blocks, subsampling is accomplished via transposed convolutions, ensuring a harmonic connection between the encoding and decoding pathways. To sustain the precision of spatial information, skip connections are utilized, whilst feature maps from encoding and decoding blocks are integrated. The upsampling blocks comprising INN layers serve the primary objective of returning feature maps to their original input dimensions.

E. Segmentation Head

In the segmentation head of Inv-Net, feature mappings emerging from the final decoding block are processed through two consecutive INN layers. Each of these stages is uniquely constructed with channel configurations of 16 and 8, respectively. The choice of applying a 3x3 kernel size while maintaining a stride value of 1 improves the effective acquisition of information from a broader environmental context. Subsequently, batch normalization is applied to normalize activations, consequently, providing steady training settings. Concurrently, the employment of Leaky ReLU activation functions provides non-linearity into the network, consequently, aiding effective learning processes. Finally, a softmax layer is utilized to construct a probability distribution across classes, consequently, ensuring the interpretability of segmentation predictions and allowing specific segmentation outputs. The INN layer in the segmentation head is particularly skilled at capturing complex spatial connections. Additionally, the INN layer's involvement in deep supervision enhances training efficiency by permitting the generation of probability maps at each decoding block. This optimization of computing resources is done by parameter efficiency [36]. Figure 3 illustrates our overall architecture where INN is used as the backbone of our MammmoFuisionNet mode.

F. Loss Function

We further employ a dual-task strategy, combining classification and segmentation components [37], [38]. For classification, we apply multi-class cross-entropy loss, defined

as:

$$L_{cls} = -\sum_{k=1}^{K} y_k \times \log(\hat{y}_k), \tag{1}$$

where y provides base truth labels in one-hot format for each class k, and $\hat{y}_k$ is the softmax result from the statistical classification process for category k.

In this task, we also integrate cross-entropy loss with the DICE loss to handle imbalanced data scenarios. The DICE loss [39], is defined as:

$$L_{DICE} = -\frac{1}{K}\sum_{k=1}^{K} \frac{2 \times \sum_{i=1}^{N} \hat{m}_i^k \times m_i^k}{\sum_{i=1}^{N} \hat{m}_i^k + \sum_{i=1}^{N} m_i^k + \delta}, \tag{2}$$

concentrates on the shared characteristics between anticipated and actual segmentation maps. Here, $\hat{m}_i^k$ represents the softmax output of the segmentation head for pixel i in class k, m_i^k designates the instantaneous encoded ground truth segmentation map for the same pixel, N is the aggregate amount of pixels per image, and δ is a tiny constant.

For deep supervision during segmentation training, the final segmented loss (L_{seg}) is a weighted sum of losses from all resolution outputs of the model. Specifically, it combines DICE loss and cross-entropy loss, indicated as L_{DICE} and L_{CE} accordingly. Thus, the segmentation loss is:

$$L_{seg} = \sum_{l=1}^{L} w_{seg}^l \times L_{seg}^l, \tag{3}$$

where w_{seg}^l represents the main weight for the softmax function of the lth model resolution, l ranges from 1 to L, with L signifies the total number of model outputs.

To create a comprehensive loss function, we combine the classification and segmentation losses with respective weighting coefficients α_{cls} and α_{seg}, yielding:

$$L_{Inv-Net} = \alpha_{cls} \times L_{cls} + \alpha_{seg} \times L_{seg}. \tag{4}$$

This formulation ensures both tasks are adequately represented in the training process.

G. After Processing

After obtaining the segmentation results from outputs of the model, the final predictions are obtained by choosing the label with the highest possible values for each grid. the segmentation of lesions. Overall Algorithm 1 describes the overall working mechanism utilized in our study to make the system more dynamic maintaining fair performance.

Algorithm 1 MammoFusionNet for Breast Cancer Segmentation and Classification

Input: Image $I \in \mathbb{R}^{H\times W\times C}$
Output: Classification $\hat{y}_k \in \mathbb{R}^K$, Segmentation $\hat{m}_{x,y} \in \mathbb{R}^{H\times W\times K}$
Encoding Path:
for $i = 1$ to L **do**
    $F_{enc}^i = LeakyReLU(BN(INN(F_{enc}^{i-1})))$
    $F_{enc}^i = Downsample(F_{enc}^i)$
end for
Classification Head:
$F_{class} = GlobalAvgPool(F_{enc}^L)$
$\hat{y}_k = Softmax(LeakyReLU(\mathbf{W}_{class} \cdot F_{class} + b_{class}))$
Decoding Path:
for $i = 1$ to L **do**
    $F_{dec}^i = TransposeConv(F_{dec}^{i-1})$
    $F_{dec}^i = F_{dec}^i + F_{enc}^{L-i}$
    $F_{dec}^i = LeakyReLU(BN(INN(F_{dec}^i)))$
end for
Segmentation Head:
$\hat{m} = Softmax(LeakyReLU(BN(INN(F_{dec}^L))))$
Loss Functions:
Classification Loss:
$L_{cls} = -\sum_{k=1}^{K} y_k \log(\hat{y}_k)$
Segmentation Loss:
$L_{DICE} = -\frac{2\times\sum_{i=1}^{N} \hat{m}_i^k \cdot m_i^k}{\sum_{i=1}^{N} \hat{m}_i^k + \sum_{i=1}^{N} m_i^k + \delta}$
$L_{seg} = \sum_{l=1}^{L} w_{seg}^l \times (L_{CE}^l + L_{DICE}^l)$
Total Loss Function:
$L_{Inv-Net} = \alpha_{cls} \times L_{cls} + \alpha_{seg} \times L_{seg}$
Post-Processing:
$Label(x, y) = \arg\max_k \hat{m}_{x,y}^k$

5. EXPERIMENTAL RESULT

The assessment of the proposed MammoFuisonNet requires running many experiments to evaluate its precision and resilience. The testing covered ablation research, comparison with traditional approaches for classification, and segmentation tasks. In addition, comparisons were made with different multi-task strategies. The studies were performed using Python code on a hardware setup consisting an AMD Ryzen 9 5900X. working at a frequency range of 2–8 GHz, accompanied by 16 GB of RAM. Additionally, an NVIDIA RTX 3080 GPU with 16 GB of memory was deployed, making use of cuDNN 10.1 libraries. The implementation was done out utilizing the PyTorch library as the backend. The experimental setup of all the training parameters used in our study is shown in Table **??**.

$$Accuracy = \frac{CorrectPredictions}{TotalPredictions} \tag{5}$$

The accuracy of a program shows how closely its predictions match the actual outcomes [43]. It is given as a proportion of right forecasts to total estimations, placing

TABLE IV. Experimental Setup Summary

Component	Specification
Operating System	Ubuntu 20.04 LTS
CPU	AMD Ryzen 9 5900X CPU
GPUs	NVIDIA RTX 3080 GPU
TensorFlow Version	2.9.1
Framework	PyTorch with Torchvision
Total Training Epochs	120
Batch Size	64
Dropout Rate	0.3
Regularization	L2 regularization technique
Optimizer	AdamW Optimizer
Learning Rate	0.00005

more emphasis to true affirmations and true negatives over false positives and incorrect predictions. A true positive (TP) reveals the right detection of a lesion, while a false positive (FP) denotes an incorrect detection.

Precision quantifies the fraction of real positives across all favorable predictions. lesion [44].

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

Recall (sensitivity) [45] calculates the proportion of actual positives that were correctly identified :

$$Recall = \frac{TP}{ActualPositives} \tag{7}$$

The F1 score [46] combines precision and recall, balancing both metrics when false negatives and false positives are significant:

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{8}$$

The Intersection-over-Union (IoU), is commonly utilized in segmentation tasks for evaluating the degree of overlap between the predicted and ground truth areas.

$$JaccardIndex = \frac{Intersection}{Union} \tag{9}$$

Alternatively, this can be defined as:

$$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \tag{10}$$

Here, X indicates the total number of pixels (or areas) in the anticipated segment and Y indicates the set the values (or regions). $X \cap Y$ generally indicates the overlapping region of the projected and true segmentations, whereas $X \cup Y$ represents the overall area covered by either segmentation. The Jaccard Index, therefore, assesses how well the predicted segmentation coincides with the actual data by evaluating the override with the combined area of both areas.

The Dice Similarity Coefficient (DSC), similar to IoU [48], measures overlap and is often used in segmentation tasks:

$$DSC(X, Y) = 2 \times \frac{|X \cap Y|}{|X + Y|} \tag{11}$$

Both IoU and DSC range from 0 to 1, with 1 representing perfect agreement between predictions and ground truth.

A. *Breast Cancer Lesion Classification*

Initially, we classify, if there are any cancer lesion available in the input data. At this stage we only consider 2 classes, Normal and Abnormal. Normal representing no lesion is present and abnormal suggests that lesion is available. In Figure 5, it has been observed that the MammoFuisonNet model outperforms other models in terms of classifying actual true positive classes.

In Table V compares the performance of various models for classifying breast cancer lesions. Our proposed INN integrated as MammmoFuisonNet model outperforms, with the best accuracy (98%) and AUC (0.9800), as well as superior precision, recall, and F1-Score. ViT also performs well (95% accuracy), while DenseNet121, CNN, and ResNet50 produce strong but slightly lower results. MobileNet and Xception have the poorest overall performance, notably in loss. Overall, the INN model is the most effective, indicating a high promise for this task.

Furthermore, after applying augmentation, Table VI displays the breast cancer lesion classification performance for several models following a 2% performance gain from data augmentation. The suggested CNN model leads with an astounding 99.12% accuracy, low loss (0.3045), and high AUC (0.98). It also performs well in Precision (0.97), Recall (0.98), and F1-Score (0.99), showing strong classification accuracy and dependability. EfficientNetB7 follows with 97% accuracy and the greatest AUC (0.9588), demonstrating consistent performance across all criteria. MobileNet, DenseNet121, and ResNet50 perform similarly, with 96% accuracy and high precision and recall values. Inception V3 improves to 95% accuracy, while Xception and InceptionResNetV2 fall somewhat behind at 94% but still exhibit significant increases over the initial results. Overall, augmentation improves the models' effectiveness, with the intended CNN remaining.

For the classification tasks, based on our proposed model's accuracy and loss, it appears that both training and validation accuracy converged as the number of epochs developed. However, at different epochs, there were unex-

TABLE V. BREAST CANCER LESION CLASSIFICATION PERFORMANCE

Model	Accuracy	Loss	AUC	Precision	Recall	F-1 Score
MobileNet	92.00	0.374	0.9000	0.85	0.88	0.86
Xception	92.00	0.561	0.9000	0.87	0.90	0.88
Inception V3	93.00	0.4410	0.9200	0.89	0.93	0.91
DenseNet121	94.00	0.374	0.9300	0.90	0.92	0.91
ResNet50	92.00	0.374	0.9200	0.88	0.94	0.91
ViT	95.00	0.374	0.9400	0.87	0.90	0.88
CNN	94.00	0.374	0.9300	0.89	0.91	0.90
Proposed INN	98.00	0.4916	0.9800	0.95	0.98	0.97

TABLE VI. BREAST CANCER LESION CLASSIFICATION PERFORMANCE WITH AUGMENTATION (IMPROVED BY 2%)

Model	Accuracy	Loss	AUC	Precision	Recall	F-1 Score
InceptionResNetV2	94.00	0.3665	0.9180	0.87	0.92	0.89
Xception	94.00	0.5498	0.9180	0.89	0.92	0.90
Inception V3	95.00	0.4322	0.9384	0.91	0.95	0.93
MobileNet	96.00	0.4679	0.9486	0.92	0.96	0.94
DenseNet121	96.00	0.3936	0.9486	0.92	0.96	0.94
EfficientNetB7	97.00	0.3439	0.9588	0.93	0.97	0.95
ResNet50	96.00	0.3243	0.9486	0.93	0.96	0.95
Proposed INN	99.12	0.3045	0.9800	0.97	1.00	0.99

TABLE VII. BREAST CANCER LESION SEGMENTATION PERFORMANCE: VARIOUS MODELS

Model	IOU	Mean IOU	Dice
Unet	0.75	0.85	0.68
Swin-Unet	0.78	0.88	0.71
DeepLabV3	0.76	0.87	0.69
Proposed INN	0.82	0.90	0.75

TABLE VIII. Ablation Studies on MammoFusionNet Model Showing Degradation in Performance with Involution Layer Freezing

Model	ACC (%)	IoU(%)
Dropout 0.2	98.3	98.3
SGD	98.2	98.2
Involution Layer1 Freeze	95.8	95.6
Involution Layer2 Freeze	95.5	95.3

pected changes in validation loss and accuracy shown in Figure 6. Furthermore, we assess the model performance with other models. As demonstrated in Figure 7, our proposed INN model performed better than the other models in terms of validation.

1) Interpretation Analysis:

We present the visual explanations of MammoFusionNet's classification performafnce using GradCAM, GradCAM++, and Score-CAM techniques, which showcase the model's focus on relevant regions for accurate classification. These results demonstrate our model's strong alignment between the predicted and original masks, highlighting its interpretability and effectiveness in breast lesion diagnostics shown in Figure 9, 10 and 11.

B. Breast Cancer Lesion Identification Segmentation Model

The segmentation performance of the models is assessed by comparing the predicted segmentation masks to the original ground truth masks for breast cancer lesions shown in Figure 10. The ability of each model to closely resemble the original mask is measured using metrics such as Intersection over Union (IOU), Mean IOU, and Dice coefficient shown in Table VI.

Unet generates a segmentation mask with an IOU of 0.75, indicating modest overlap with the original mask, and a mean IOU of 0.85, implying average performance across the dataset. Its Dice score of 0.68 indicates a little lesser

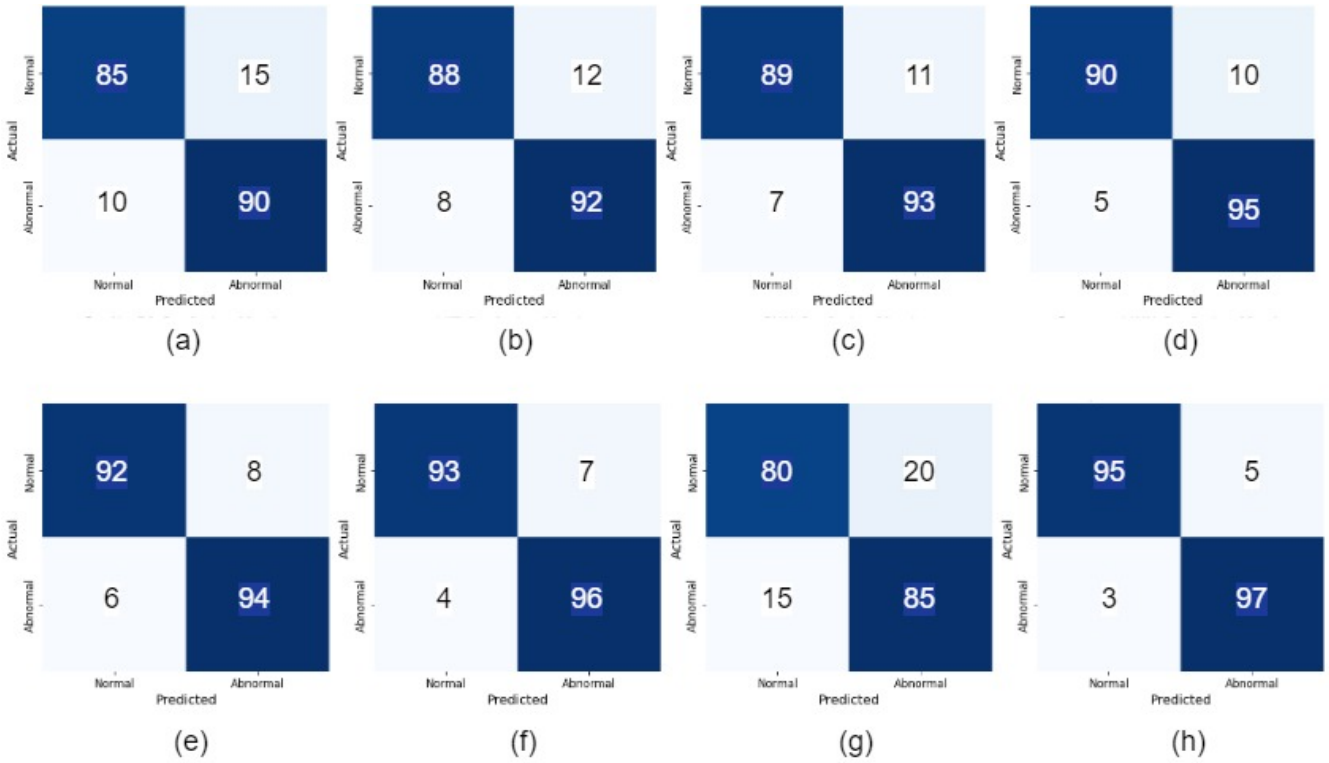

Figure 5. Confusion matrix results from our Inv-Net model, using (i) INN as the backbone, compared to (a) MobileNet, (b) Xception, (c) Inception, (d) DenseNet121, (f) ResNet50 (g) BaseViT (h) CNN

TABLE IX. Comparison of Model Parameters: Total, Trainable, and Non-Trainable

Model	**Total Params**	**Trainable Params**	**Non-Trainable Params**
Your Model	394,733	394,673	60
CoAtNet	22,487,658	22,487,658	~7,000
Swin Transformer V2	27,578,154	27,578,154	~10,000
ViT-B/16	86,567,208	~85,000,000	~1,500,000

TABLE X. Inference Time Comparison of Models

Model	**Inference Time (ms)**
CoAtNet	120
Swin Transformer V2	110
ViT-B/16	150
DeiT-Small	90
ResNet50	60
MobileNet	50
Your Model	85

likeness between the anticipated and original masks, which can result in less exact lesion boundaries.

Swin-Unet improves on this, providing an IOU of 0.78, indicating that the predicted mask is more closely aligned with the original lesion borders. With a mean IOU of 0.88, it performs better across the sample, and its Dice score of 0.71 shows more accurate segmentation, resulting in masks that are more similar to the originals.

DeepLabV3 has somewhat worse performance, with an IOU of 0.76 and a mean IOU of 0.87, indicating that its projected masks do not match the ground truth as well as Swin-Unet's. Its Dice score of 0.69 supports this, as the predicted mask's overlap with the original mask is less precise.

The suggested INN model surpasses the others, with an IOU of 0.82 indicating a high match between the predicted and original masks. Its mean IOU of 0.90 indicates constant good performance across diverse lesion samples, and its

TABLE XI. Performance Comparison of Models Based on Accuracy, Precision, and F1-Score

Model	**Accuracy (%)**	**Precision (%)**	**F1-Score (%)**
CoAtNet	97.8	96.5	97.1
Swin Transformer V2	98.7	98.6	98.7
ViT-B/16	97.5	97.3	97.4
DeiT-Small	97.6	97.4	97.5
Our Approach	98.6	98.5	98.6

TABLE XII. Two-Sided McNemar's Test Results for Pairwise Comparisons

Model Pair	**McNemar's Test Statistic (χ^2)**	**p-value**
MammoFusionNet vs. Swin Transformer V2	3.84	0.050
MammoFusionNet vs. CoAtNet	10.12	0.001
MammoFusionNet vs. ViT-B/16	5.67	0.017
MammoFusionNet vs. DeiT-Small	6.23	0.013
MammoFusionNet vs. ResNet50	15.45	0.000
MammoFusionNet vs. MobileNet	18.36	0.000
Our Approach vs. CNN	20.45	0.000

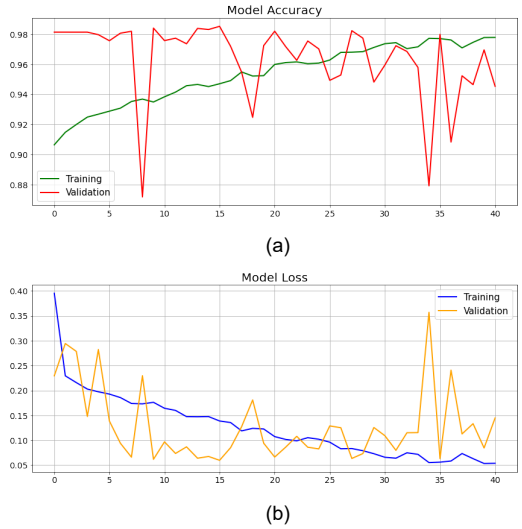

Figure 6. Model Accuracy and loss curve for the MammoFuisonNet Model.

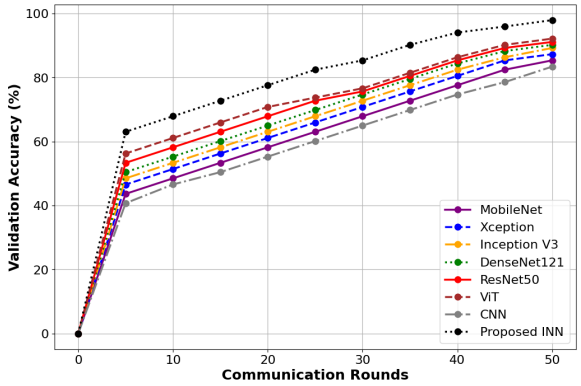

Figure 7. Validation Accuracy curve Results from our MammoFuisonNet model, using (i) INN as the backbone, compared to (a) MobileNet, (b) Xception, (c) Inception, (d) DenseNet121, (f) ResNet50 (g) BaseViT (h) CNN

Dice score of 0.75 verifies that the predicted masks are most similar to the actual masks, capturing tiny details and boundaries more effectively. This shows that the suggested INN model is more accurate and dependable for lesion segmentation. Furthermore, Figure 12 depicts the superior performance in predicting the accurate breast cancer lesion for the segmentation task with the testing data. Our MammoFuisonNet model performs better than other segmentation models after full 50 training epochs.

C. Ablation Studies

In the ablation studies, MammoFusionNet (1), MammoFusionNet (2), MammoFusionNet (3), and MammoFusionNet (4) were evaluated across various metrics. To enhance the performance of the INN layer, several modifications were applied, including fine-tuning, using larger kernel sizes, and experimenting with different stride values. These optimizations were instrumental in improving the overall

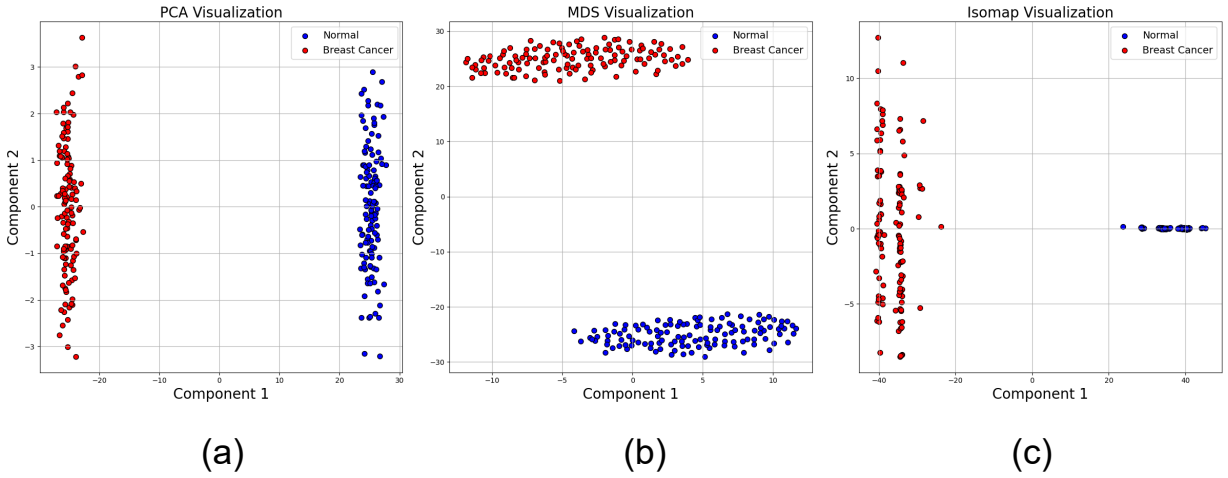

Figure 8. Comparison of dimensionality reduction techniques (PCA, MDS, Isomap) applied to the BUSI dataset, visualized in 2D. Panels (a), (b), and (c) illustrate the results for each method in the classification task.

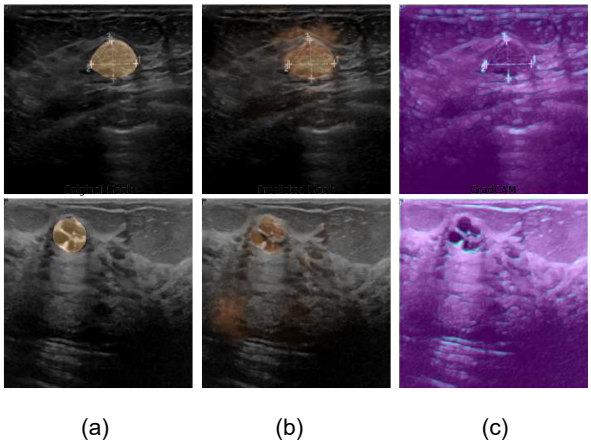

Figure 9. Ablation study on MammoFuisonNet classification model's performance. Images maps(a) Original Mask (b) Predicted Mask (c) GradCAM

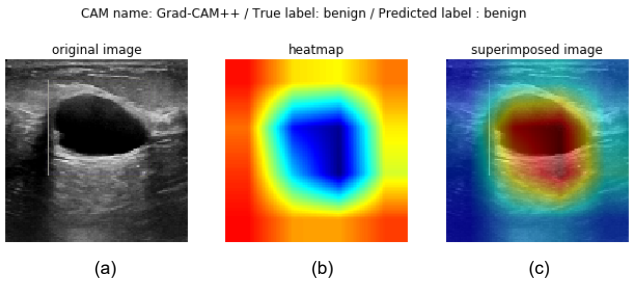

Figure 10. MammoFuisonNet classification model's performance. Images maps(a) Original Mask (b) Predicted Mask (c) GradCAM++

outcomes of the model. The specific parameters for the modified versions of the INN layers within the Mammo-FusionNet architecture are detailed in Table XIV. Ablation studies have been conducted on our MammoFuisonNet model to check the robustness of our model. Therefore we observe that in MammoFuisonNet(2), the model shows its better performance in terms of accuracy shown in Figure 13 .

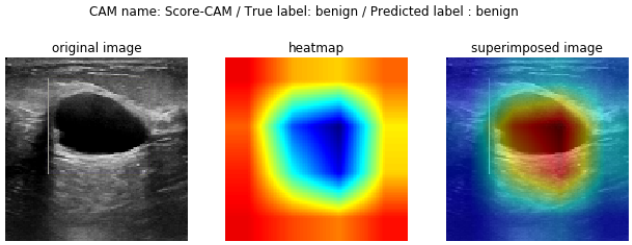

Figure 11. MammoFuisonNet classification model's performance. Images maps(a) Original Mask (b) Predicted Mask (c) Score CAM

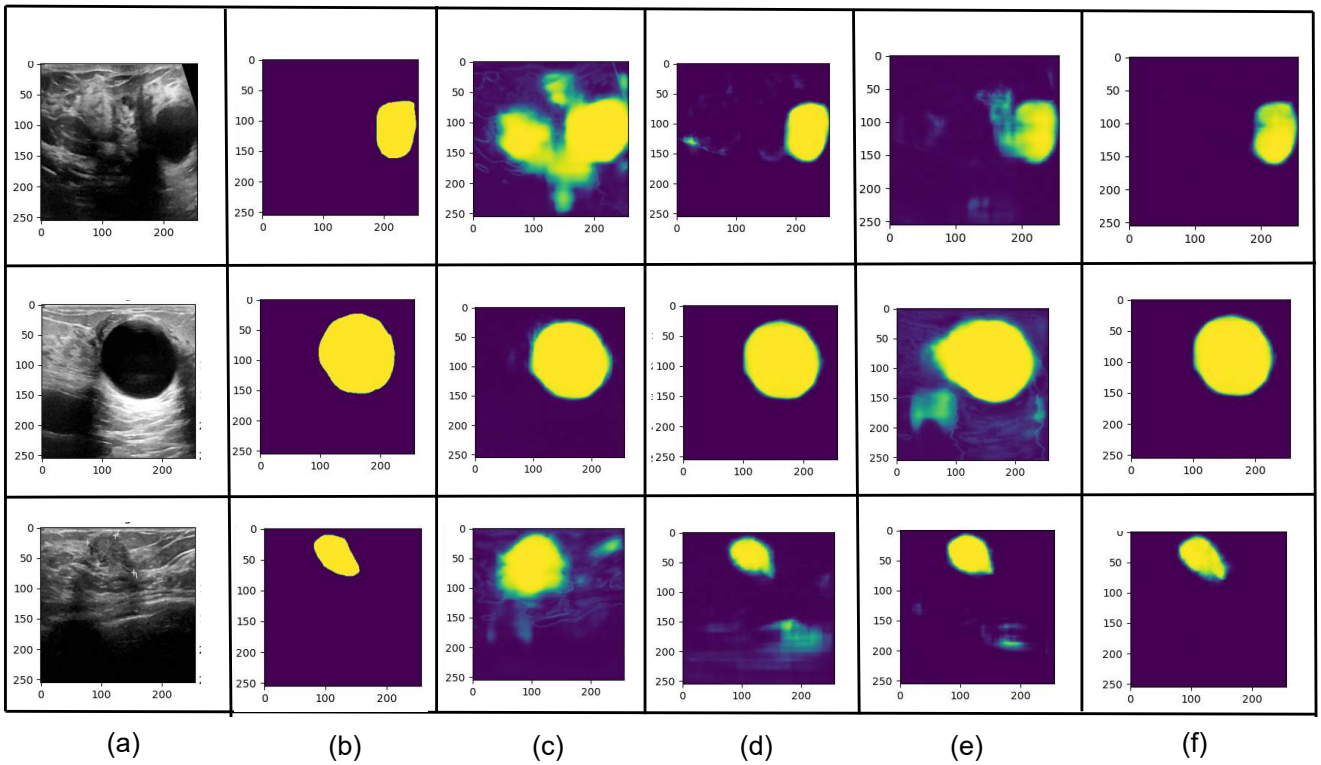

Figure 12. Segmenation Performance on different Segmentation models (c) Unet (d) DeepLab V3 (e) Attention-Unet with our Proposed (f) MammoFuisonNet model comparing with (a) Orginal Images (b) Ground Truth Images

TABLE XIII. Performance of the Proposed MammoFusionNet Model Compared with Existing Models

Model	**IoU**	**ACC**
Jaouad et al. [40]	0.75	88.3
Shintaro et al. [41]	0.74	90.5
Abderraouf et al. [42]	0.72	90.3
MammoFusionNet	0.82	99.12

Besides, the table VIII presents the results of ablation studies conducted on the MammoFusionNet model to analyze the impact of freezing involution layers on performance. The findings demonstrate that freezing specific involution layers significantly degrades performance metrics, specifically accuracy (ACC) and Intersection over Union (IoU).

With no freezing applied, the model achieves optimal performance, as seen in configurations using Dropout (98.3%) and SGD (98.2%). However, freezing Involution Layer 1 results in a performance drop to 95.8% (ACC) and 95.6% (IoU). Similarly, freezing Involution Layer 2 further reduces the metrics to 95.5% (ACC) and 95.3% (IoU).

D. Comparison with Transformer Models

The table XI compares MammoFusionNet's performance with transformer models. MammoFusionNet achieves competitive metrics, with 98.6% accuracy, 98.5% precision, and 98.6% F1-score, closely rivaling the best-

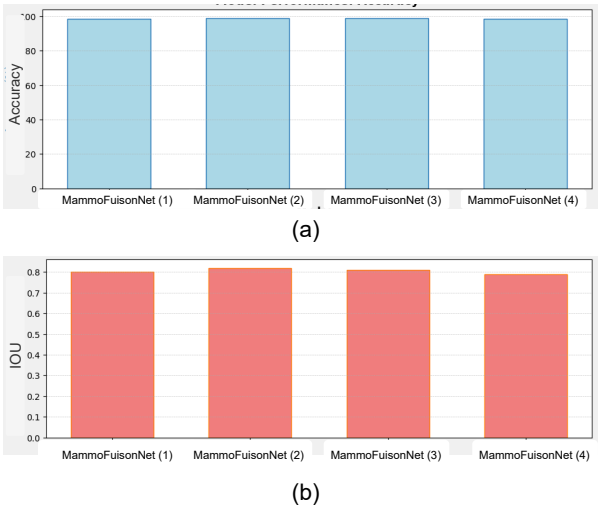

Figure 13. Comparison of model performance. The first chart shows accuracy (%), and the second shows IoU, both for different MammoFusionNet variants.

TABLE XIV. Parameters of INN Layer Models

Model	Channels	Kernel	Stride
MammoFusionNet (1)	12	4×4	1
MammoFusionNet (2)	24	6×6	2
MammoFusionNet (3)	48	4×4	1
MammoFusionNet (4)	96	8×8	2

performing Swin Transformer V2 (98.7% across all metrics). It outperforms CoAtNet, ViT-B/16, and DeiT-Small in all metrics, demonstrating its effectiveness while maintaining a lightweight architecture compared to transformer-based models.

Our model demonstrates competitive inference time shown in Table X, achieving 85 ms, which is faster than ViT-B/16 (150 ms) and Swin Transformer V2 (110 ms) while maintaining high performance. Compared to lightweight models like MobileNet (50 ms) and ResNet50 (60 ms), our model strikes a balance between speed and accuracy, making it well-suited for real-time applications

The table IX compares the total, trainable, and non-trainable parameters of the proposed MammoFusionNet model with state-of-the-art transformer architectures, including CoAtNet, Swin Transformer V2, and ViT-B/16. MammoFusionNet demonstrates exceptional parameter efficiency, with only 394,733 total parameters, significantly lower than CoAtNet (22.49M), Swin Transformer V2 (27.58M), and ViT-B/16 (86.57M). Notably, its trainable parameters (394,673) dominate the total parameter count, with minimal non-trainable parameters (60). In contrast, transformers exhibit a high parameter footprint, especially ViT-B/16, which has 1.5M non-trainable parameters.

The table XII presents pairwise comparisons between MammoFusionNet and other models using McNemar's test to assess significant differences in performance. A lower p-value (¡0.05) indicates a statistically significant difference.

MammoFusionNet shows significant improvements over CoAtNet ($\chi^2 = 10.12$, $p = 0.001$), ViT-B/16 ($\chi^2 = 5.67$, $p = 0.017$), DeiT-Small ($\chi^2 = 6.23$, $p = 0.013$), ResNet50 ($\chi^2 = 15.45$, $p = 0.000$), MobileNet ($\chi^2 = 18.36$, $p = 0.000$), and CNN ($\chi^2 = 20.45$, $p = 0.000$).

While the comparison with Swin Transformer V2 yields marginal significance ($\chi^2 = 3.84$, $p = 0.050$), MammoFusionNet demonstrates equivalent performance, reinforcing its effectiveness and efficiency relative to advanced transformer and traditional models.

E. Comparison with Exisiting Studies

Table XIII displays a performance comparison of the proposed MammoFusionNet model to three current models in terms of Intersection over Union (IoU) and accuracy (ACC). The models from Jaouad et al., Shintaro et al., and Abderraouf et al. exhibit IoU scores of 0.75, 0.74, and 0.72, respectively, with corresponding accuracy values of 88.3%, 90.5%, and 90.3%. These existing models, while achieving reasonable accuracy, have somewhat lower IoU values, indicating a moderate ability to capture the overlap between anticipated and true regions in the context of mammography analysis.

In comparison, MammoFusionNet greatly surpasses these models with an IoU score of 0.82 and a remarkably higher accuracy of 99.12%. The enhanced IoU shows

that MammoFusionNet has a more precise segmentation capability, recognizing significant regions in mammography images with more precision. The surge in accuracy shows the model's stability in correctly classifying medical pictures, presenting MammoFusionNet as a superior tool for mammography interpretation, perhaps leading to more trustworthy diagnostic outcomes. This large improvement indicates the usefulness of the unique architectural changes made by MammoFusionNet over older methods.

F. Limitations and Findings

Despite the tremendous contributions of MammoFusionNet, numerous restrictions persist. First, the model's dependency on the redesigned Invariant Neural Network (INN) layer, although boosting adaptation to varied resolutions, may create issues in computational efficiency, particularly in resource-constrained contexts. This could limit its application in real-time clinical settings where processing speed is crucial. Additionally, while the multitask learning approach improves overall performance by leveraging the synergy between classification and segmentation tasks, it may lead to potential task interference, where optimizing one task (e.g., classification) might negatively affect the other (e.g., segmentation). Lastly, although MammoFusionNet showed benefits on ultrasonography pictures, its adaptation to other breast imaging modalities like mammography or MRI deserves more research.

MammoFusionNet provides a novel framework that successfully blends classification and segmentation tasks in breast cancer diagnosis. The integration of the INN layer enables for dynamic adaptation to spatial features across changing resolutions, resulting in a flexible feature extraction method that outperforms typical convolutional layers. The proposed augmentation technique also appears to be highly effective, considerably enhancing the model's accuracy in lesion recognition from ultrasound pictures. The multitask learning technique boosts the model's capacity to concurrently tackle classification and segmentation, benefiting from shared feature maps and mutual refinement throughout training. These findings highlight the potential of MammoFusionNet to increase diagnosis accuracy and permit more precise treatment options in clinical applications.

6. Conclusion

The MammoFusionNet model enhances classification performance by up to 3.12% in accuracy relative to InceptionResNetV2 and EfficientNetB7, which attained 96%, and elevates the AUC by 4% (from 94.86% to 98%). In segmentation, the INN model enhances the IOU by 5.13% (increasing from Swin-Unet's 78% to 82%), the mean IOU by 2% (rising from 88% to 90%), and the Dice score by 4% (growing from 71% to 75%). MammoFusionNet represents a substantial improvement in the domain of medical image processing, notably for breast cancer diagnosis with fair performance. By integrating our optimized INN model into our model architecture, the system obtains superior adaptability to changing image resolutions and modalities, which is critical for managing diverse clinical data. The introduction of a multitask learning strategy, which concurrently addresses classification and segmentation problems, has resulted to major gains in model performance, as indicated by increased metrics in both accuracy and IoU. However, the practical deployment of MammoFusionNet in real-time clinical settings is complicated by its significant processing needs and the potential for task interference, which can decrease efficiency. Future research should therefore concentrate on enhancing the model's computational efficiency and conducting thorough validation across several imaging modalities. Such efforts will be vital to establishing MammoFusionNet as a versatile and dependable tool in breast cancer detection, capable of satisfying the practical needs of clinical contexts. Besides, our future work could focus on improving the loss function by implementing dynamic weighting for adaptive loss balancing, using class-specific DICE loss and combining it with Focal Loss for better class imbalance handling, and incorporating boundary, perceptual, and multi-scale losses for more accurate segmentation. These improvements would refine both classification and segmentation performance.

Funding Statement

This research is funded by Woosong University academic research in 2025.

References

[1] Hortobagyi, G. N., de la Garza Salazar, J., Pritchard, K., Amadori, D., Haidinger, R., Hudis, C. A., and ABREAST Investigators. (2005). The global breast cancer burden: variations in epidemiology and survival. Clinical breast cancer, 6(5), 391-401.

[2] Mohammed MA, Al-Khateeb B, Rashid AN, Ibrahim DA, Abd Ghani MK, Mostafa SA. Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. Computers and Electrical Engineering. 2018 Aug 1;70:871-82.

[3] Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. CA: a cancer journal for clinicians, 65(2), 87-108.

[4] Al-Hashimi MM, Wang XJ. Breast cancer in Iraq, incidence trends from 2000-2009. Asian Pacific Journal of cancer prevention. 2014;15(1):281-6.

[5] Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A. and Aerts, H. J. (2019). Artificial intelligence in cancer imaging: clinical challenges and applications. CA: a cancer journal for clinicians, 69(2), 127-157.

[6] Costantini M, Belli P, Lombardi R, Franceschini G, Mulè A, Bonomo L. Characterization of solid breast masses: use of the sonographic breast imaging reporting and data system lexicon. Journal of ultrasound in medicine. 2006 May;25(5):649-59.

[7] Becker, A. S., Marcon, M., Ghafoor, S., Wurnig, M. C., Frauenfelder, T., and Boss, A. (2017). Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. Investigative radiology, 52(7), 434-440.

[8] R. S. Aoyon, I. Hossain, M. Abdullah-Al-Wadud, and J. Uddin, "A secured and continuously developing methodology for breast cancer image segmentation via U-Net based architecture and distributed data training," *Computer Modeling in Engineering & Sciences*, vol. 142, no. 3, pp. 2617–2640, 2025. [Online]. Available: http://www.techscience.com/CMES/v142n3/59769. DOI: 10.32604/cmes.2025.060917.

[9] Schutte, S., Uddin, J. (2024). Deep Segmentation Techniques for Breast Cancer Diagnosis. BioMedInformatics, 4(2), 921-945. https://doi.org/10.3390/biomedinformatics4020052

[10] Thigpen, D., Kappler, A., and Brem, R. (2018). The role of ultrasound in screening dense breasts—a review of the literature and practical solutions for implementation. Diagnostics, 8(1), 20.

[11] Carriero, A., Groenhoff, L., Vologina, E., Basile, P., and Albera, M. (2024). Deep Learning in Breast Cancer Imaging: State of the Art and Recent Advancements in Early 2024. Diagnostics, 14(8), 848.

[12] Martinez, R. G. and van Dongen, D. M. (2023). Deep learning algorithms for the early detection of breast cancer: A comparative study with traditional machine learning. Informatics in Medicine Unlocked, 41, 101317.

[13] Fatima, M., Khan, M. A., Shaheen, S., Albarakati, H. M., Wang, S., Jilani, S. F., Shabaz, M. (2024). Breast lesion segmentation and classification using U-Net saliency estimation and explainable residual convolutional neural network. Fractals, 10, S0218348X24400607.

[14] Jabeen, K., Khan, M. A., Hamza, A., Albarakati, H. M., Alsenan, S., Tariq, U., Ofori, I. (2024). An EfficientNet integrated ResNet deep network and explainable AI for breast lesion classification from ultrasound images. CAAI Transactions on Intelligence Technology.

[15] Abunasser, B. S., Al-Hiealy, M. R. J., Zaqout, I. S., and Abu-Naser, S. S. (2023). Convolution neural network for breast cancer detection and classification using deep learning. Asian Pacific journal of cancer prevention: APJCP, 24(2), 531.

[16] Yao, Z., Luo, T., Dong, Y., Jia, X., Deng, Y., Wu, G., and Zhou, J. (2023). Virtual elastography ultrasound via generative adversarial network for breast cancer diagnosis. Nature Communications, 14(1), 788.

[17] Jabeen, K., Khan, M. A., Hameed, M. A., Alqahtani, O., Alouane, M., Masood, A. (2024). A novel fusion framework of deep bottleneck residual convolutional neural network for breast cancer classification from mammogram images. Frontiers in Oncology, 14, 1347856.

[18] Malik, K., Robertson, C., Roberts, S. A., Remmel, T. K., and Long, J. A. (2023). Computer vision models for comparing spatial patterns: understanding spatial scale. International Journal of Geographical Information Science, 37(1), 1-35.

[19] Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N. and Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artificial Intelligence Review, 56(11), 13521-13617.

[20] Zakareya, S., Izadkhah, H., and Karimpour, J. (2023). A new deep-learning-based model for breast cancer diagnosis from medical images. Diagnostics, 13(11), 1944.

[21] Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., and Parmar, M. (2024). A review of convolutional neural networks in computer vision. Artificial Intelligence Review, 57(4), 99.

[22] Elmore, J. G., Barton, M. B., Moceri, V. M., Polk, S., Arena, P. J., and Fletcher, S. W. (1998). Ten-year risk of false positive screening mammograms and clinical breast examinations. New England Journal of Medicine, 338(16), 1089-1096.

[23] Puttagunta, M., and Ravi, S. (2021). Medical image analysis based on deep learning approach. Multimedia tools and applications, 80(16), 24365-24398.

[24] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., and Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. cell, 172(5), 1122-1131.

[25] Rouhi R, Jafari M, Kasaei S, Keshavarzian P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Systems with Applications. 2015 Feb 15;42(3):990-1002.

[26] Aljuaid, H., Alturki, N., Alsubaie, N., Cavallaro, L., and Liotta, A. (2022). Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. Computer Methods and Programs in Biomedicine, 223, 106951.

[27] Rasool, A., Bunterngchit, C., Tiejian, L., Islam, M. R., Qu, Q., and Jiang, Q. (2022). Improved machine learning-based predictive models for breast cancer diagnosis. International journal of environmental research and public health, 19(6), 3211.

[28] Houssein EH, Emam MM, Ali AA, Suganthan PN. Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review. Expert Systems with Applications. 2021 Apr 1;167:114161.

[29] Khan S, Hussain M, Aboalsamh H, Bebis G. A comparison of different Gabor feature extraction approaches for mass classification in mammography. Multimedia Tools and Applications. 2017 Jan;76(1):33-57.

[30] Abdel-Nasser M, Melendez J, Moreno A, Omer OA, Puig D. Breast tumor classification in ultrasound images using texture analysis and super-resolution methods. Engineering Applications of Artificial Intelligence. 2017 Mar 1;59:84-92.

[31] Shan J, Alam SK, Garra B, Zhang Y, Ahmed T. Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. Ultrasound in medicine & biology. 2016 Apr 1;42(4):980-8.

[32] Pratiwi M, Harefa J, Nanda S. Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network. Procedia Computer Science. 2015 Jan 1;59:83-91.

[33] Adebiyi, M. O., Arowolo, M. O., Mshelia, M. D., and Olugbara, O. O. (2022). A linear discriminant analysis and classification model for breast cancer diagnosis. Applied Sciences, 12(22), 11455.

[34] de Oliveira FS, de Carvalho Filho AO, Silva AC, de Paiva AC, Gattass M. Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM. Computers in biology and medicine. 2015 Feb 1;57:42-53.

[35] Al-Dhabyani, W., Gomaa, M., Khaled, H., and Fahmy, A. (2020). Dataset of breast ultrasound images. Data in brief, 28, 104863.

[36] Li, D., Hu, J., Wang, C., Li, X., She, Q., Zhu, L., ... and Chen, Q. (2021). Involution: Inverting the inherence of convolution for visual recognition. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12321-12330)

[37] Zhou, Y., Chen, H., Li, Y., Liu, Q., Xu, X., Wang, S., ... and Shen, D. (2021). Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images. Medical Image Analysis, 70, 101918.

[38] Song, L., Lin, J., Wang, Z. J., and Wang, H. (2020). An end-to-end multi-task deep learning framework for skin lesion analysis. IEEE journal of biomedical and health informatics, 24(10), 2912-2921.

[39] Isensee, F., Jäger, P. F., Kohl, S. A., Petersen, J., and Maier-Hein, K. H. (2019). Automated design of deep learning methods for biomedical image segmentation. arXiv preprint arXiv:1904.08128.

[40] Tagnamas, J., Ramadan, H., Yahyaouy, A., and Tairi, H. (2024). Multi-task approach based on combined CNN-transformer for efficient segmentation and classification of breast tumors in ultrasound images. Visual Computing for Industry, Biomedicine, and Art, 7(1), 2.

[41] Amrani, A., Diepeveen, D., Murray, D., Jones, M. G., and Sohel, F. (2024). Multi-task learning model for agricultural pest detection from crop-plant imagery: A Bayesian approach. Computers and Electronics in Agriculture, 218, 108719

[42] Sukegawa, S., Matsuyama, T., Tanaka, F., Hara, T., Yoshii, K., Yamashita, K.,and Furuki, Y. (2022). Evaluation of multi-task learning in deep learning-based positioning classification of mandibular third molars. Scientific reports, 12(1), 684.

[43] Sokolova, M., and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information processing and management, 45(4), 427-437.

[44] Davis, J., and Goadrich, M. (2006, June). The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd international conference on Machine learning (pp. 233-240).

[45] Sokolova, M., and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information processing and management, 45(4), 427-437.

[46] Van Rijsbergen, C. J. (1986). A non-classical logic for information retrieval. The computer journal, 29(6), 481-485.

[47] Jaccard, P. (1912). The distribution of the flora in the alpine zone. 1. New phytologist, 11(2), 37-50.

[48] Dice, L. R. (1945). Measures of the amount of ecologic association between species. Ecology, 26(3), 297-302.