Densely Convolutional Networks for Breast Cancer Classification with Multi-Modal Image Fusion

Eman Hamdy
College of Computing and Information Technology
Arab Academy for Science, Technology and Maritime Transport Alex,
Egyp
eman.hamdy666@adj.aast.edu

Osama Badawy
College of Computing and Information Technology
Arab Academy for Science, Technology and Maritime Transport Alex,
Egypt
obadawy@aast.edu

Mohamed Zaghloul
College of Engineering and Technology
Arab Academy for Science, Technology and Maritime Transport Alex,
Egypt
Dr_mszaghloul@aast.edu

Abstract: *Breast cancer is the main health burden worldwide. Cancer is located in the breast, starts when the cell grows under control and begins as in-situ carcinoma and when spread into other parts known as invasive carcinoma. Breast cancer mass can early be found by image modality when discovering mass early can easily diagnose and treated. Multimodalities used for the classification of breast cancer Such as mammography, ultrasound, and Magnetic resonance imaging. Two types of fusion are used earlier fusion and later fusion. Early fusion it's a simple relation between modalities while later fusion gives more interest to fusion strategy to learn the complex relationship between various modalities as a result, can get highly accurate results when using the later fusion. When combining two image modalities (mammography, ultrasound) and using an excel sheet containing the age, view, side, and status attribute associated with each mammographic image using DenseNet 201 with Layer level fusion strategy as later fusion by making connections between the various paths and same path by using Concatenated layer. Fusing at the feature level achieves the best performance in terms of several evaluation metrics (accuracy, recall, precision area under the curve, and F1 score) and performance.*

Keywords: *Breast cancer, classification, deep learning, densenet, diagnostic imaging, multimodal imaging.*

Received April 3, 2022; accepted April 28, 2022
https://doi.org/10.34028/iajit/19/3A/6

1. Introduction

Breast cancer is one of the most extreme critical diseases and the second most common among all cancers. Incidence and mortality in Egypt. Normally, relying on one modality has the hazard of missing tumors or false diagnosis and low accuracy. To start a pathologist with one source of modality leads to uncertainty. The main problem with visually diagnosing medical images is that diagnosis is restricted to experts as it is based on their experiences. The late detection of tumors and abnormalities can delay serving severe cases. Having misclassified breast cancer as normal breast results in not detecting cancer in the early stages, which may lead to the development of cancer to higher cancer levels, and may also start to spread within the body parts. Can get multi-information about a goal by using multi-modal images that assist to extract features from various perspectives and make whole information. The fusion strategy in multi-modal fusion techniques is split into 3 categories like layer level fusion, input-level fusion network, and decision level fusion Figure 1 illustrates the category of level fusion networks [1]. The objectives of this research are to combine two image modalities (Mammography and Ultrasound) and use an excel sheet containing the age, view, side, and status attribute associated with each mammographic image for early diagnosis of breast cancer in the early stage to be easy to treat and improve the prognosis and chances of survival through timely clinical treatment to patients and to support pathologists for taking the right decision. The system will help to reduce the patient time and early identification of the diseases.

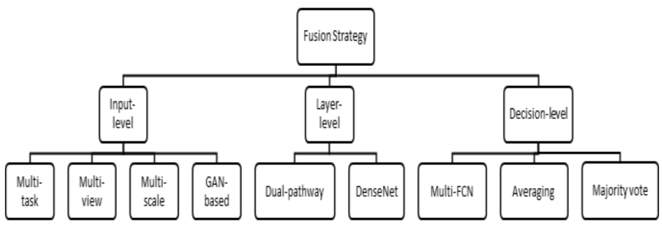


Figure 1. Three categories of fusion strategy [6].

2. Background

2.1 Breast Cancer

Breast cancer is carcinoma usually feels like a solid or firm lump that is usually irregular in type and may feel like it's connected to skin or tissue deep in the breast. [10] Medical imaging reveals inside structures covert by the skin and bones, it is used as a diagnostic tool for several diseases and has a vital role in monitoring treatment effects and predicting outcomes. Breast cancer in its early stages is most facilely and effectively treated. Survival rates decrease when women present with advanced cases; thus, a significant strategy for decreasing the carcinoma death rate is growing the speed of cases that are discovered through the first stages of the malady. unluckily, ladies discover cancer at a later stage in resource-poor countries than ladies elsewhere, partially because of the dearth of mass screening programs in various such countries. Can raise the average of cancer cases that are diagnosed in their earliest stages when making regular screening of all ladies aged fifty and over. Can evaluate a breast lump by Physical examination and breast cancer imaging modalities. The digital Digital Mammography (DM), ultrasound (US), and Resonance Imaging (MRI) are multimodalities used for the classification of carcinoma to early discovery of mass may be facilely diagnosed and treated.

1. DM:
Using mammography is the most effective to early diagnosing breast cancer in the early stage to detect and diagnose abnormalities of the breast. It consists of two images of each breast: Craniocaudal (CC) and Medial-Lateral-Oblique (MLO) [5, 9]. The advantages of mammograms are safe radiation, improved treatment of early disease, and death-rate reduction.[2]

2. Ultrasound (US):
Ultrasound images also known as sonograms; it is used as complementary to mammography. Ultrasound confirmed a high sensitivity for determining abnormalities in dense breasts and for women younger than 35 years of age. Ultrasound is advocated recommended to be used as a complement to DM due to the fact of its availability, inexpensiveness in contrast to different modalities, and well tolerated by patients.[1]

3. MRI:
Magnetic fields and radio waves MRI are used in diagnostic technology that is used to take a nitty-gritty picture of the body's soft tissue, such as the breast. Breast MRI photos can display spotless views of breast soft tissues than MGs, US, or Computerized Tomography (CT) pictures, MRI is mostly requested once the lesion has been diagnosed and the physician needs to get certain information around the quantity and extent of the disease.

2.2. Convolutional Network (DenseNet)

Convolutional network (DenseNet) is a structure in which every layer is hooked up to each different layer in a feed-forward style. Dense blocks contain n dense layer with dense connections. The layers in a dense block are connected using a dense connection, this means each layer receives feature maps from all previous layers and passes its feature maps to all following layers [2, 7]. Figure 2 illustrates layer dense block. The different architectures of DenseNets have been brief in Table 1. each architecture consists of four Dense Blocks with a different number of layers and three transition layers. The DenseNet 201 has (6, 12, 48, 32) layers in the four dense blocks. The fourth dense block is followed by a classification layer that accepts the feature maps of all layers of the network to perform the classification. Each dense block contains 1x1 convolution and 3x3 convolution to reduce the number of channels rather than the input. Each Convolution (Conv) layer shown in the table corresponds to the series Batch Normalization-Rectified Linear Units - Convolution (BN-ReLU-Conv). Advantages of DenseNet are robust gradient flow, more diversified features, and maintaining low complexity features.[2, 7]

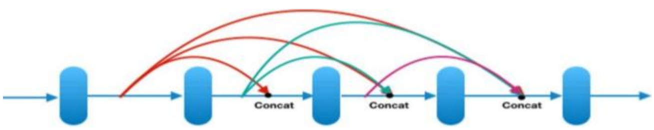


Figure 2. Concatenation between the layers.

Table 1. DenseNet architectures [2, 7].

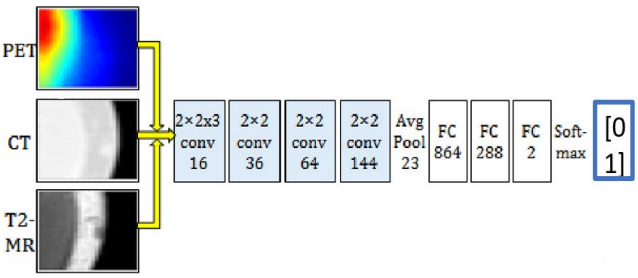
Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-121
Convolution	112 x 112	7 X 7 conv, stride 2			
Pooling	56 x 56	3 x 3 max pool, stride 2			
Dense Block (1)	56 X 56	1 x 1 conv 3 X 3 conv x 6	1 x 1 conv 3 X 3 conv x 6	1 x 1 conv 3 X 3 conv x 6	1 x 1 conv 3 X 3 conv x 6
Transition Layer (1)	56 X 56	1 x1 conv			
	28 x 28	2 x 2 average pool, stride 2			
Dense Block (2)	28 x 28	1 x 1 conv 3 X 3 conv x 12	1 x 1 conv 3 X 3 conv x 12	1 x 1 conv 3 X 3 conv x 12	1 x 1 conv 3 X 3 conv x 12
Transition Layer (2)	28 x 28	1 x1 conv			
	14 x 14	2 x 2 average pool, stride 2			
Dense Block (3)	14 X 14	1 x 1 conv 3 X 3 conv x 24	1 x 1 conv 3 X 3 conv x 32	1 x 1 conv 3 X 3 conv x 48	1 x 1 conv 3 X 3 conv x 64
Transition Layer (3)	14 X 14	1 x1 conv			
	7 x 7	2 x 2 average pool, stride 2			
Dense Block (4)	7 x 7	1 x 1 conv 3 X 3 conv x 16	1 x 1 conv 3 X 3 conv x 32	1 x 1 conv 3 X 3 conv x 32	1 x 1 conv 3 X 3 conv x 48
Classification Layer	1 x 1	7 x 7 global average pool			
		1000D fully-connected, softmax			

3. Related Work

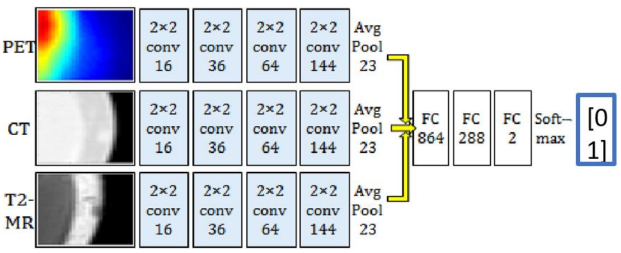
More recent attention has focused on medical image analysis using deep learning. In [3] convert the image to 256x256 pixels using 22 layers (27 layers including pooling layers), and part of these layers are a total of 9 inception modules is GoogLeNet architecture as a result Convolution Neural Network (CNN) model achieves a sensitivity of 0.958, the accuracy of 0.925 and specificity of 0.925 and radiologists achieve a sensitivity of 0.583, the accuracy of 0.658and specificity of 0.604. The Strength make the comparison between the CNN model and three radiologists in Time for reading the images. The Weakness When converting an image to 256×256 pixels might loss of information and, thus, affect the performance of the models [3].

Figure 3 performs the three image fusion schemes based on the CNN with different structures and combined into a single framework. The proposed picture segmentation framework can study the multi-modality images using diverse fusing schemes at the identical time; the results that at the same time as all of the fusion schemes output perform the single-modality schemes, fusing at the feature stage can usually acquire excellent performance for each accuracy and computational cost. It uses three image fusion schemes using CNN with a different network [5].

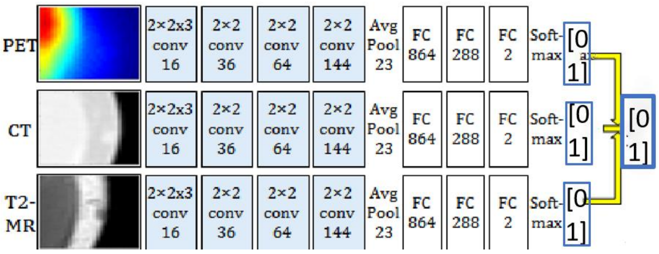
In [11] Category the network architectures into input-level fusion network illustrate in Figure 4, layer-level fusion network illustrates in Figure 5, and decision-level fusion network illustrates in Figure 6. multi-information about a target (tumor, tissue) is provided by Multi-modality, So can detect the tumor. As a result, can detect mass or areas of interest. Multi-modal fusion achieves higher accuracy than single-modal networks [11].



a) Type-I fusion network implementing the feature-level fusing.



b) Type-II fusion network implementing the classifier-level fusing.



c) Type-III fusion network implementing the decision-level fusing.

Figure 3. Three categories of fusion strategy using CNN, the yellow arrows indicate the fusion location [5].

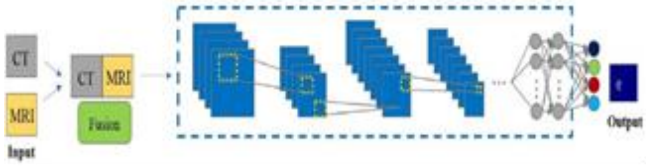


Figure 4. Input-level fusion network [11].

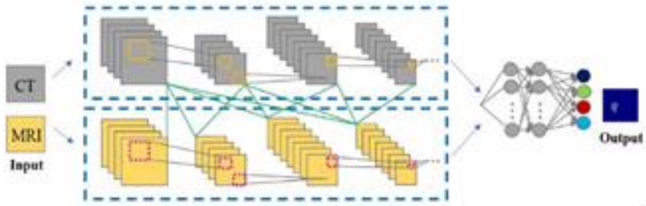


Figure 5. Layer-level fusion network [11].

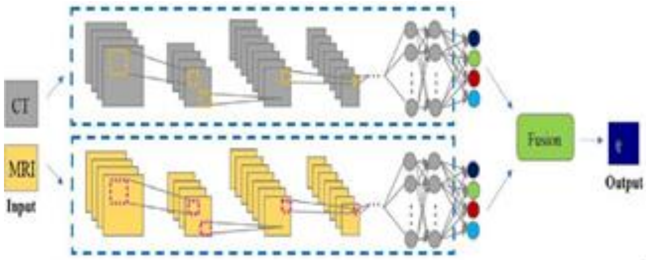


Figure 6. Decision-level fusion network [11].

4. The Proposed System

The proposed deep learning structure goal at enhance the overall effectiveness (Accuracy) and efficiency (Performance). The proposed model provides use combines two image modalities (Mammography and Ultrasound) and uses an excel sheet containing the age, view, side, and status attribute associated with each mammographic image with a layer-level fusion strategy. Figure 7 illustrates the pipeline of using deep learning in a multi-modal medical image. Starting from data preparation using (data augmentation and reshaping) and network architecture (DenseNet) the fusion strategy (layer-level).

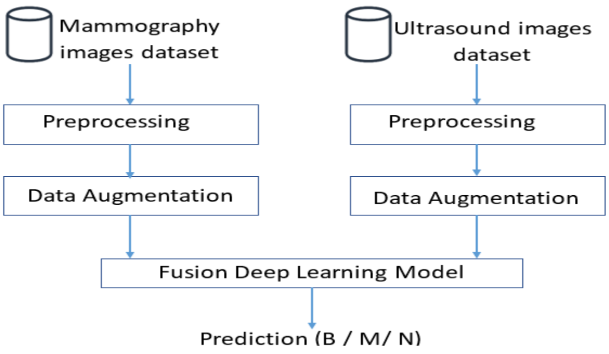


Figure 7. Pipeline of multi-modal.

4.1. Datasets

The datasets chosen for this research are Mini-Digital Database for Screening Mammography (Mini-DDSM) is a dataset used in breast cancer classification categorized into normal, benign, and malignant, its a free online dataset with a size of 45 GB, this data set comes with an excel sheet that gives direct access to all image attributes (age and views) and format (.png). Figure 8 illustrates Mini-DDSM [8].

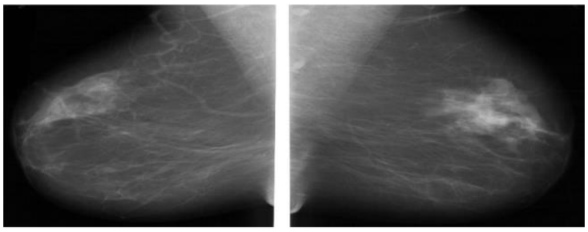
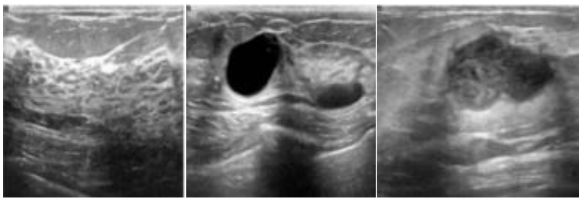


Figure 8. Illustrates mini-ddsm.

The dataset chosen for breast ultrasound is Breast Ultrasound Images (BUSI) is a dataset used in breast cancer classification categorized into normal, benign, and malignant with an average image size of 500×500 pixels with images format (.png), its free online dataset, breast ultrasound images contain 780 images collected by 600 female patients in the ages of 25 to 75 years old and collected in 2018 from Baheya hospital for early detection & treatment of women's cancer, Cairo, Egypt, Figure 9 illustrates (BUSI dataset). [1]



a) Normanl b) Benign c) Malignant

Figure 9. Illustrates breast US [1].

4.2. Preprocessing and Data Augmentation

A variety of preprocessing steps is utilized in the dataset to minimize the processing system necessities and to extend the robustness of the models. Use data Augmentation because the number of available medical photos is always constrained. One technique to overcome the scarcity of data is data augmentation [4]. Need to execute a geometric transformation on the main dataset, Modify the pixel positions of the photo whilst retaining the authentic features.Data augmentation is very probably to enhance accuracy considering the model can see more samples. Data augmentation becomes a usually used approach to grow the size by using zoom-range, rescale, rotation, vertical-flip, and Horizontal –flip. The datasets have different images size so data processing is necessary for data reshaping. the data is reshaped image to (64,64) pixels. The train/test split was implemented with 70% to the training set and 30% to the test set. Train models with memory limitations using the size of Random Access Memory (RAM) smaller than the giant training data used to train the model, to overcome this problem use a batches approach to divide the data into 32 images for each patch, and the data was transferred batch by batch to RAM when each batch finish the training, it is ignored and the next batch of images is created and stocked in RAM.

4.3. Network Architecture Fusion of Multi-Modal

The purpose of this study is to combine two image modalities (Mammography and Ultrasound) and use an excel sheet containing the age, view, side, and status attribute associated with each mammographic image using DenseNet201 based on a layer-level fusion strategy. Medical image analysis using DenseNet 201 has a high improvement due to the flow of information within the same path and various paths which facilitate the combination between the mammograms features and ultrasound features this leads to highly accurate results. this is because various features are combined from various views and get all information about the images and attributes to get from an excel sheet that contains different information about the image decrease the effect of the over-fitting problem. When using DenseNet 201 has a significant to improving the breast cancer classification in terms of benefit from a connection between layers. There are two types of fusion early and late fusion, early fusion make simple relation between two modalities while when using late fusion achieve high results due to getting complex features from different views and excel sheet So late fusion has a high result than early fusion. The layer level fusion used different images are used as single input then each path is fused in the decision layer to achieve the last end outcome. The network is densely connected both inside the same path and throughout various paths. Mammography doesn't work well in dense breasts, so mammography alone misses a lot of cancerous cases, as a result, using ultrasound with mammography will improve the accuracy of breast cancer classification. Table 2. Shows the parameters of the DenseNet and the proposed HyperDens. Figure 10 illustrates layer-level fusion strategy using DenseNet

201 giving at that point the liberty of the model to learn where and the way the uncommon modalities must be processed and blended.

Table 2. The layers used in the DenseNet201 and the output size according to input size 64x64 pixel.

Layers	Output Size	DenseNet-201
Convolution	32 x 32	7 X 7 conv, stride 2
Pooling	16 x 16	3 x 3 max pool, stride 2
Dense Block (1)	16 X 16	1 x 1 conv 3 X 3 conv x 6
Transition Layer (1)	16 X 16	1 x1 conv
	8 x 8	2 x 2 average pool, stride 2
Dense Block (2)	8 x 8	1 x 1 conv 3 X 3 conv x 12
Transition Layer (2)	8 x 8	1 x1 conv
	4 x 4	2 x 2 average pool, stride 2
Dense Block (3)	4 X 4	1 x 1 conv 3 X 3 conv x 48
Transition Layer (3)	4 X 4	1 x1 conv
	2 x 2	2 x 2 average pool, stride 2
Dense Block (4)	2 x 2	1 x 1 conv 3 X 3 conv x 32
Classification Layer	1 x 1	global average pool
		fully-connected, softmax

4.4. Experimental Setup

Use Microsoft Azure cloud machines. The Azure cloud platform is more than 200 products and cloud services. Use Microsoft Azure instead of Google collab because of the long run time (collab runs for 24 hours) and Microsoft Azure is faster than Google collab in run time. Azure for Students (100$ free credit) with windows 10 Pro, 8GiB memory and use python language with Keras v2.3.1.

4.5. Model Validation

The data set has three different classes of mammography and ultrasound images, which are normal, and abnormal classes. If the image was classified as abnormal, further classification is needed to classify the abnormality into a benign or malignant mass as shown in Figure 11. The training and testing datasets were divided into subsets to calculate the performance of the model.

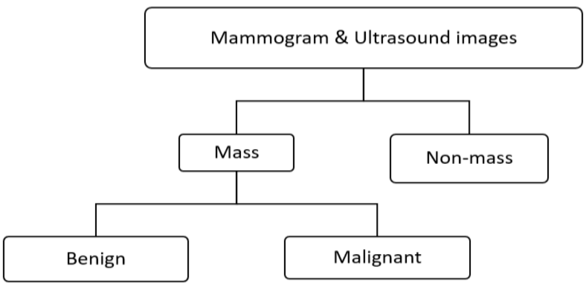


Figure 11. Dataset classification.

4.6. Implementation

All deep learning architectures are using batch size 32, the advantage of using a batch size requires less memory when training the model that's more serious if not able to fit all images in memory at the same time. Use Adam optimizer with Learning Rate (LR) 0.001 and use different evaluation metrics to monitor each epoch. Early Stopping to stop training model if the model stop for improvement, it is a form of regularization used to avoid overfitting. Use concatenate layer that concatenates a list of inputs by taking inputs and concatenating them along a specified dimension but the inputs must have the same size to exceed this problem use up-sampling to increase/up-samples the spatial dimensions of the feature maps by using Bicubic Interpolation.

4.7. Evaluation Criteria

To evaluate the proposed method, five appraisal metrics were used to appraise the accuracy of the competing methods: Accuracy, Precision, Recall, Area Under The Curve (AUC), F-measure.

Accuracy (Acc) = (TP plus TN) times (TP plus TN plus FPplus FN)-1 (1)

Precision= TP times (TP plus FP) -1 (2)

Recall= TP times (TP plus FN) -1 (3)

F Measure = 2 multiped by [(precision times recall) divided by (precision plus recall)] (4)

5. Result

Using the late fusion with the dense connection between layers has high accuracy than the use of a single path without lack in performance when using multi-modal medical images with an excel sheet for mammogram images for classification of breast cancer achieves high results.

Table 3. illustrates classification results of experiments shown use multi-modal denseNet fusion achieve high results. To ensure the effectiveness of using data augmentation when increasing the dataset make experiment 1 uses an ultrasound image dataset (BUSI) without data augmentation and experiment 2 uses an ultrasound image dataset(BUSI) with data augmentation. Experiment 3 uses a Mini-DDSM image dataset and excel sheet without data augmentation and experiment 4 uses a Mini-DDSM image dataset and excel sheet with data augmentation. Experiment 5 uses a Layer level fusion with dense connection to combine (Mini-DDSM image dataset with excel sheet and BUSI) with data augmentation due to the result of experiments 2 and 4 having high accuracy than experiments without data augmentation.

6. Conclusions

This study investigated using multimodal fusion improved evaluation metrics for breast cancer classification. Using Layer level fusion with DensNet201 connection between the same path and different paths for multi-modal images fusion. When using the fusion strategy get features from various views of images and excel sheets to achieve better data analysis. Our proposed work enables the network whole liberty to look to various combinations that represent modalities features. Hard to find a prior study that has investigated two images modalities (Mammography and Ultrasound) used to classify breast cancer. In conclusion, the work shows the effect of using a combination of multi-modal image fusion to remedy the classification of breast cancer. From previous, the gained outcome of breast cancer.

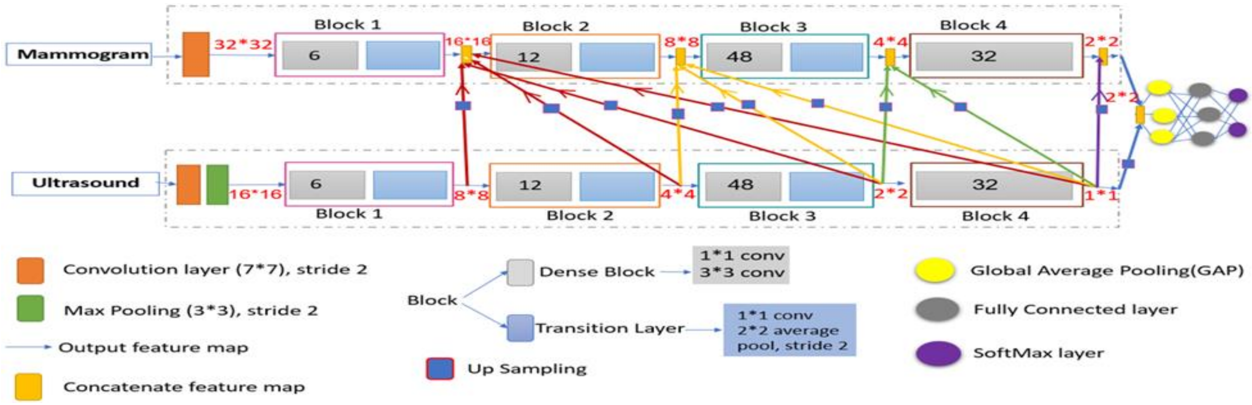


Figure 10. layer-level fusion strategy (Mammogram and Ultrasound Inputs).

Table 3. Classification results of experiments shown use multi-modal denseNet fusion achieve high results.

Path connectivity	Architecture	Dataset	Accuracy	Precision	Recall	AUC	F-measure
No connectivity between paths	Single path for Ultrasound input without data augmentation	BUSI	85.13%	84.9%	84.73%	89.3%	84.8%
	Single path for Ultrasound input using data augmentation	BUSI	86.89%	86.89%	86.89%	91.22%	86.05%
	Single path for mammogram with excel input without data augmentation	Mini-DDSM	87.89%	88.45%	88.47%	90.42%	88.46%
	Single path for mammogram with excel sheet input using data augmentation	Mini-DDSM	90.4%	91.8%	91.8%	92.6%	91.8%
Connectivity between paths	Multi-Modal DenseNet Fusion (Combine Breast ultrasound images and Mini-DDSM images With an excel sheet containing the information associated with each mammographic image)	Mini-DDSM And BUSI	95.2%	94.8%	94.5%	96.7%	94.65%

References

[1] Al-Dhabyani W., Gomaa M., Khaled H., and Fahmya A., "Deep Learning Approaches for Data Augmentation and Classification of Breast Masses Using Ultrasound Images," *Journal of Advanced Computer Science and Applications*, vol. 10, no. 5, pp. 1-11, 2019.

[2] Dolz J., Gopinath K., Yuan J., Lombaert H., Desrosiers C., and Ben Ayed I., "HyperDense-Net: A Hyper-Densely Connected CNN for Multi-Modal Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 38, no. 5, pp. 1116-1126, 2019.

[3] Fujioka T., Kubota K., Mori M., Kikuchi Y., Katsuta L., Kasahara M., Oda G., Ishiba T., Nakagawa T., and Tateishi U., "Distinction Between Benign and Malignant Breast Masses at Breast Ultrasound Using Deep Learning Method With A Convolutional Neural Network," *Japanese journal of radiology*, vol. 37, no. 6, pp. 466-472 2019.

[4] Garg S. and Jindal B., "Skin Lesion Segmentation in Dermoscopy Imagery," *The International Arab Journal of Information Technology*, vol. 19, no. 1, pp. 29-37, 2022.

[5] Guo Z., Li X, Huang H., Guo N., and Li Q., "Medical Image Segmentation Based on Multi-Modal Convolutional Neural Network: Study on Image Fusion Schemes," *in Proceeding of the IEEE 15th International Symposium on Biomedical Imaging*, USA, pp. 903-907, 2018.

[6] Hamdy E., Zaghloul M., and Badawy O., "Deep Learning Supported Breast Cancer Classification with Multi-Modal Image Fusion," *in Proceeding of the 2021 22nd International Arab Conference on Information Technology*, Oman, pp. 1-7,

2021.

[7] Huang G., Liu Z., Maaten L., and Weinberger K., "Densely Connected Convolutional Networks," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, USA, pp. 4700-4708 2017.

[8] Kaggle.https://www.kaggle.com/cheddad/minidds2, Last Visited, 2021.

[9] Shen L., Margolies L., Rothstein J. Fludere., McBride R., and Sieh W., Reports S., "Deep Learning to Improve Breast Cancer Detection on Screening Mammography," *Scientific Reports*, vol. 9, no. 1, pp. 1-12, 2019.

[10] Singh A., *Demedicalizing Women's Health*, New Delhi, Gyan Publishing House, 2010.

[11] Zhou T., Ruan S. and Canu S., "A review: Deep Learning for Medical Image Segmentation Using Multi-Modality Fusion," *Array* vol. 3, pp. 100004, 2019.

Eman Hamdy was born in 1997 in Alexandria, Egypt. She received her B.Sc. in College of Computing and Information Technology and M.Sc. in Deep learning supported breast cancer classification with multi-modal image fusion from the Faculty of Computing and Information Technology, Alex., Egypt, Arab Academy for Science Technology and Maritime Transport (AASTMT) at Alexandria, Egypt. She is currently an Assistant Lecturer College of Computing and Information Technology, Arab Academy for Science and Technology and Maritime Transport, Alexandria, Egypt.

Osama Badawy was born in 1949 in Al Minya, Egypt. He received his B.Sc. in Electrical Engineering and M.Sc. in the Optimum Smoothing of Data Using Digital Filter from the Faculty of Military-Technical Engineering, Cairo, Egypt, and a Ph.D. in Expert Systems for Image Processing from the Faculty of Engineering, Ein Shams University at Cairo, Egypt. He is currently a Full Professor at the College of Computing and Information Technology, Arab Academy for Science and Technology and Maritime Transport, Alexandria, Egypt.

Mohamed Zaghloul Asst. prof. 2009 Ph.D. degree in Electrical Engineering 2002 from Menofia Faculty of Engineering. M.SC degree in Electrical Engineering from Faculty of Engineering, Alexandria University, 1991. Diploma in Electrical Engineering from Faculty of Engineering, Alexandria University, 1984. B.SC in Electrical Engineering from M T C, July 1977, Cairo. Expert for supervision on research projects with EU and master's thesis for students. Lecturer in artificial intelligent Faculty of Engineering/ Arab Academy for Science and Technology (AAST) & Maritime Transport. Alex. Egypt. Member of Egyptian industry comity and scientific research. Lecturer, in Electronic and Communications Department Tanta University and Kafr El Shek University. Lecturer, in Electronic and Communications Department Naval College.