A Novel Artificial Intelligence-Based Hybrid System to Improve Breast Cancer Detection Using DCE-MRI

Volkan KAYA[1,*], İsmail AKGÜL[1], Erdal KARAVAŞ[2], Sonay AYDIN[3], Ahmet BARAN[1]

[1] Department of Computer Engineering, Faculty of Engineering and Architecture, Erzincan Binali Yıldırım University, Türkiye
[2] Department of Radiology, Faculty of Medicine, Bandırma Onyedi Eylül University, Balıkesir, Türkiye
[3] Department of Radiology, Faculty of Medicine, Erzincan Binali Yıldırım University, Erzincan, Türkiye

Abstract. The interpretation of breast magnetic resonance imaging (MRI) in the healthcare field depends on the good knowledge and experience of radiologists. Recent developments in artificial intelligence (AI) have shown advances in the field of radiology. However, the desired levels have not been reached in the field of radiology yet. In this study, a novel model structure is proposed to characterize the diagnostic performance of AI technology for individual breast dynamic contrast material–enhanced (DCE) MRI sequences. In the proposed model structure, Inception-v3, EfficientNet-B3 and DenseNet-201 models were used as hybrids together with the Yolo-v3 algorithm to detect breast and cancer regions. In the proposed model, DCE-MRI sequences (T2, ADC, Diffusion, Non-Contrast Fat Non-Suppressed T1, Non-Contrast Fat Suppressed T1, Contrast Fat Suppressed T1, and Subtraction T1) were evaluated separately and validation was made, thus providing a unique perspective. According to the validation results, the model structure with the best performance was determined as Yolo-v3 + DenseNet-201. With this model structure, 92.41 accuracy, 0.5936 loss, 92.44% sensitivity, and 92.44% specificity rates were obtained. In addition, it was determined that the results obtained without using contrast material in the best model were 91.53% accuracy, 0.9646 loss, 92.19% sensitivity, and 92.19% specificity. Therefore, it is predicted that the need for contrast material use can be reduced with the help of this model structure.

Key words: artificial intelligence, artificial neural network, deep learning, DCE-MRI, breast cancer detection and classification

1. INTRODUCTION

The recent trend toward incorporating artificial intelligence (AI) into health care, particularly radiology, has increased medical physicians' expectations for the possible impact of AI on their everyday practice. However, AI technology is still evolving, and has not yet reached a stable and widespread acceptance level [1-3].

Breast imaging, particularly breast magnetic resonance imaging (MRI), has been critical in the rapid advancement of breast cancer management. Many studies showed that dynamic contrast material–enhanced (DCE) MRI achieved the highest sensitivity of any imaging modality in detecting breast cancer. Previous breast MRI studies have proved that a reliable identification of cancer (ductal carcinoma in situ or invasive carcinoma) was possible, regardless of radiographic breast density, stage, tumor type, or postsurgical changes. Breast MRI, by recognizing the prominent MRI features of benign and malignant disease, as well as the distinct morphologic and kinetic characteristics associated with various malignant tumor subtypes, enables radiologists to make more accurate diagnoses than other more conventional imaging modalities, so that beneficial treatment changes may occur [4, 5].

Despite the numerous advancements in breast MRI, several problems remain. The expensive cost, need for contrast medium, and lengthy exam duration are some of the primary barriers preventing breast MRI from being extensively associated. Additionally, background enhancement can impair lesion identification by masking or simulating lesions. Furthermore, false negative examinations are caused by errors in perception, interpretation, and management [6, 7]. AI is on the verge of overcoming some or all of these constraints. Previous studies emphasized that AI architectures can efficiently detect benign and malignant lesions on routine DCE-MRI examinations, ultrafast DCE-MRI, and maximum intensity projection (MIP) images of routine DCE-MRI [8, 9]. AI techniques was also found to be effective in benign-malign lesion differentiation by using peritumor tissue as input [10]. It also can decrease the number of benign/unnecessary biopsies [11]. While the

*e-mail: vkaya@erzincan.edu.tr

diagnostic accuracy of AI systems in differentiating benign from malignant lesions and their ability to improve radiologists' performance [5] have been evaluated previously, the diagnostic performance of AI systems for individual sequences and their ability to eliminate or reduce the need for contrast medium use have not been extensively studied.

The current work aims to contribute to the growing body of knowledge regarding the use of AI in breast MRI by characterizing the diagnostic performance of AI technology for individual breast DCE-MRI sequences. Additionally, we predicted that the need for contrast media use could be reduced thanks to the proposed new hybrid model structures. The contributions of this study to the literature can be summarized as follows:

1) Application of 3 new hybrid model structures for breast cancer detection.

2) High detection rate with the recommended Yolo-v3 + DenseNet-201 model structure.

3) Presenting a unique perspective by evaluating 7 different DCE-MRI sequences separately.

4) Reducing the amount of contrast material with the proposed model structures.

5) Evaluation and comparison of images taken from cancer and normal patients by both the specialist physician and the proposed new hybrid model.

2. MATERIALS AND METHODS
2.1. System Configuration
The performance of Yolo-v3, Inception-v3, EfficientNet-B3, and DenseNet-201 models used in the study to detect breast cancer was tested using Python programming language in Google Colaboratory [12] environment with an NVIDIA Tesla K80 graphics processor. In addition, using the C# programming language, indexes were created automatically, the filenames of the patient images were encoded according to a certain format for patient confidentiality, and the mixed indexes were placed in a certain order.

2.2. Image Data Acquisition and Preprocessing
The dataset was obtained using breast MRI images examined and labeled by Erzincan Binali Yıldırım University Mengücek Gazi Education and Research Hospital specialist physicians to detect breast cancer (Clinical Research Ethics Committee Decision: E-22746194-025.11-458920).

This dataset contains 73522 breast MRI images in DICOM format, belonging to 64 patients in total, 36853 of 32 Cancer patients, and 36669 of 32 Normal patients. Each patient has more than a thousand breast MRI images in DICOM format in 16 folders (Topogram, T2, ADC, Diffusion, Non-Contrast Fat Non-Suppressed T1, Non-Contrast Fat Suppressed T1, Contrast Fat Suppressed T1-1, Contrast Fat Suppressed T1-2, Contrast Fat Suppressed T1-3, Contrast Fat Suppressed T1-4, Contrast Fat Suppressed T1-5, Subtraction T1-1, Subtraction T1-2, Subtraction T1-3, Subtraction T1-4, and Subtraction T1-5). DICOM files consisting of 16 folders were collected under 8 folders (Topogram, T2, ADC, Diffusion, Non-Contrast Fat Non-Suppressed T1, Non-Contrast Fat Suppressed T1, Contrast Fat Suppressed T1, and Subtraction T1).

Images in DICOM format were coded and renamed according to the sequence number by taking the first two letters of the patient's first and last names with the help of C# programming language for patient confidentiality. Images in DICOM format, each in different sizes, were initially converted to JPG format at a gray level without changing their size. Using the C# programming language, the images in 8 different folders of each patient were gathered under the labels of cancer and normal.

The images in the cancer and normal labels were examined in the presence of specialist physicians and the final version of the dataset was created by using the images deemed appropriate. Since there are more than one thousand breast MRI images of each patient in this dataset, 2545 images from the images of cancer patients, in which the cancer region is prominent, are labeled under the cancer class. In addition, 29757 images suitable for use from breast MRI images of normal patients are labeled under the normal class. The sample images included in this dataset are shown in Fig. 1.
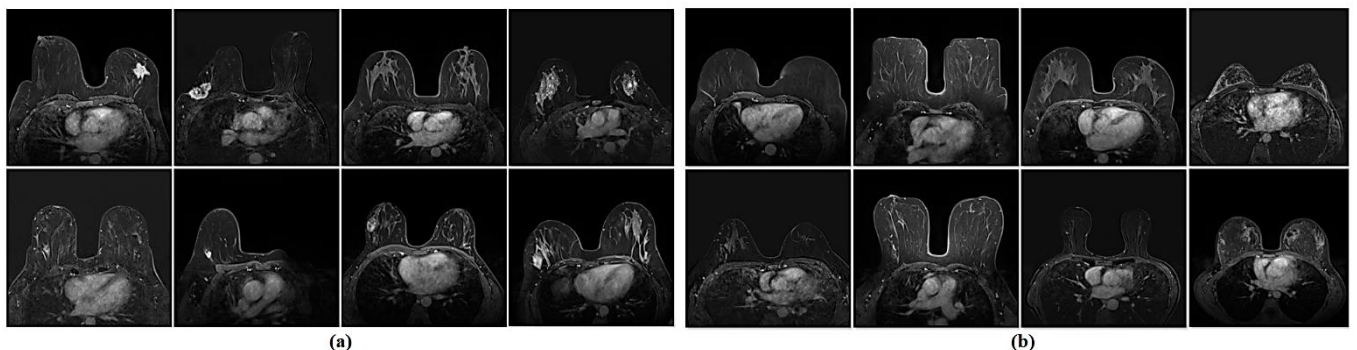

(a)   (b)

Fig. 1. Dataset sample images (a) Cancer, (b) Normal

2.3. Proposed AI Model Structure

In this study, CNN model structures of Yolo-v3 [13], Inception-v3 [14], EfficientNet-B3 [15], and DenseNet-201 [16] in the literature were used to create a new model structure. Using the breast MRI image dataset, the Yolo-v3 + Inception-v3, Yolo-v3 + EfficientNet-B3 and Yolo-v3 + DenseNet-201 model structures were applied and the cancer region was determined and classified from the breast MRI images. The model structure that detects and classifies the patient's cancerous breast area from breast MRI images is shown in Fig. 2.
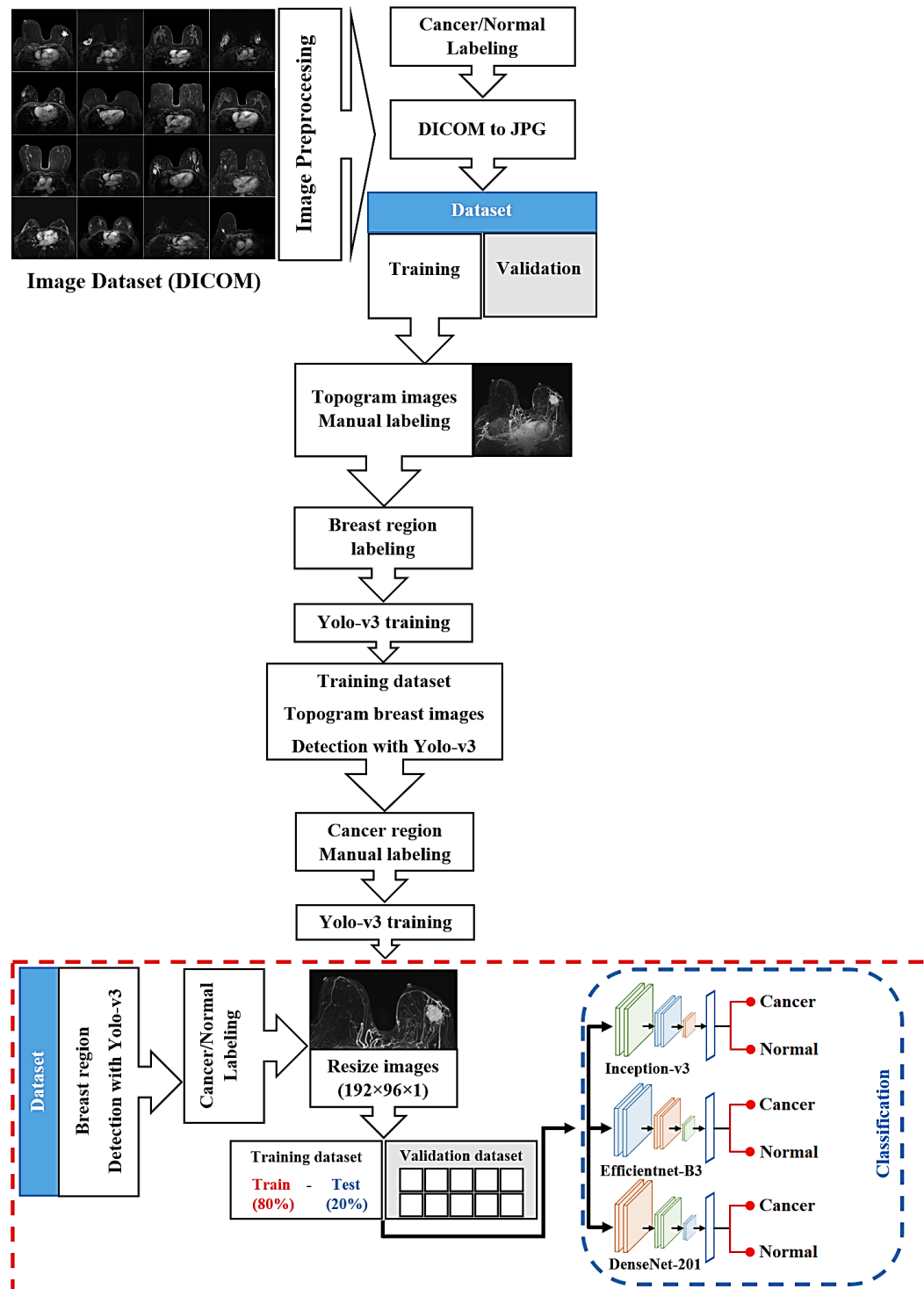

Fig. 2. Proposed AI Model Structure

3

After pre-processing the patient data taken in DICOM format in the proposed model structure shown in Fig. 2, the images in the dataset were labeled in JPG format under the Cancer and Normal class and turned into a dataset. Then, images of 50 patients out of 64 patients in this dataset were divided into the training dataset, and images of 14 patients were divided into the validation dataset. For breast region cross-section training with Yolo-v3, breast regions in all topogram images (304) of 50 patients in the training dataset were manually selected and their coordinates were labeled. These labeled coordinates were trained at 2000 epochs using the Yolo-v3 algorithm and "breast area cross-section weights" were obtained.

By using these weights, the breast regions in the topogram images were determined with Yolo-v3, and the breast regions were cut. Then, for cancer region cross-section training from these cut breast regions, 155 images containing the cancer region were selected manually and their coordinates were labeled. These labeled coordinates containing the cancer region were trained at 1500 epochs using the Yolo-v3 algorithm and "cancer region cross-section weights" were obtained.

Using the "breast region cross-section weights" obtained with Yolo-v3, the breast regions in 73522 (36853 cancer, 36669 normal) breast MRI images in the whole dataset were detected and cut, and 58946 (36853 cancer, 36669 normal) breast region images were has been obtained. Of the 29189 images in the cancer class, 2545 breast images containing the cancer region were determined in the presence of a specialist physician, and the final form of the cancer class was obtained. All of the breast region images of 29757 normal patients were used under the normal labeled class. Thus, a total of 32302 image data was obtained and the final version of the dataset was created. Then, images under cancer/normal classes were resized to $192 \times 96 \times 1$ for model training.

After the other images (T2, ADC, Diffusion, Non-Contrast Fat Non-Suppressed T1, Non-Contrast Fat Suppressed T1, Contrast Fat Suppressed T1, and Subtraction T1) except the topogram of 50 patients in the training dataset were gathered together, they were divided into 80% training and 20% test dataset. In addition, 10 validation datasets were created by using images other than the topogram of 14 patients included in the validation dataset. Validation-1 for T2, Validation-2 for ADC, Validation-3 for Diffusion, Validation-4 for Non-Contrast Fat Non-Suppressed T1, Validation-5 for Non-Contrast Fat Suppressed T1, Validation-6 for Contrast Fat Suppressed T1, Validation-7 for Subtraction T1, Validation-8 for T2 + ADC + Diffusion + Non-Contrast Fat Non-Suppressed T1 + Non-Contrast Fat Suppressed T1 combination, Validation-9 for Contrast Fat Suppressed T1 + Subtraction T1 combination, and Validation-10 datasets for all images except topogram were created. Inception-v3, EfficientNet-B3, and DenseNet-201 model training were carried out according to the training-test dataset, and with the validation datasets that were not used in the model training, the model was final tested and its cancer/normal classification was verified.

The number of images belonging to the training-test datasets is given in Table 1, and the image numbers of the validation datasets are given in detail in Table 2.

Table 1. Number of images in the training-test dataset belonging to 50 patients

	Training (%80)	Test (%20)	Total (%100)
Number of Images	20373	5094	25467

Table 2. Number of images in the validation dataset belonging to 14 patients

Validation Dataset	Explanation	Number of Images
Validation-1	T2	218
Validation-2	ADC	23
Validation-3	Diffusion	23
Validation-4	Non-Contrast Fat Non-Suppressed T1	236
Validation-5	Non-Contrast Fat Suppressed T1	575
Validation-6	Contrast Fat Suppressed T1	2954
Validation-7	Subtraction T1	2806
Validation-8	Validation-1 – Validation-5	1075
Validation-9	Validation-6 – Validation-7	5760
Validation-10	Validation-1 – Validation-7	6835
Total		**6835**

Using the training-test datasets given in Table 1, Inception-v3, EfficientNet-B3, and DenseNet-201 models were trained with 30 epochs, 64 batch size, Adamax optimization algorithm, and Sigmoid activation function parameters.

3. RESULTS AND DISCUSSION

In the study, to detect the cancer region from breast MRI images, firstly the breast region was cross-sectioned with the Yolo-v3 algorithm and then the cancer region was detected. In the study, the Yolo-v3 algorithm was run for 2000 epochs to get the breast region cross-section, and a loss value of 0.020 and an average loss value of 0.033 were obtained. Similarly, to detect the cancer region, the Yolo-v3 algorithm was run at 1500 epochs, and a loss value of 0.220 and an average loss value of 0.180 were obtained. These low loss values show that the Yolo-v3 model trainings are successful.

Thanks to the training-test dataset (50 patient images) obtained as a result of the Yolo-v3 model training, Inception-v3, EfficientNet-B3, and DenseNet-201 model structures were trained to classify cancer/normal. The accuracy, loss, sensitivity, and specificity performance metric values obtained as a result of the training are given in Table 3, Table 4, Table 5, and Table 6, respectively.

4

Table 3. Accuracy values obtained as a result of the training-test of the models

Epoch	Inception-v3		EfficientNet-B3		DenseNet-201	
	Train	Test	Train	Test	Train	Test
1	0.9626	0.9246	0.9116	0.9246	0.9624	0.9246
2	0.9849	0.9580	0.9219	0.9246	0.9844	0.9548
3	0.9898	0.9804	0.9267	0.9256	0.9900	0.9831
4	0.9924	0.9847	0.9388	0.9433	0.9919	0.9831
5	0.9924	0.9925	0.9551	0.9564	0.9932	0.9912
6	0.9939	0.9921	0.9654	0.9556	0.9937	0.9815
7	0.9958	0.9894	0.9739	0.9641	0.9937	0.9943
8	0.9965	0.9878	0.9803	0.9704	0.9954	0.9943
9	0.9966	0.9935	0.9848	0.9729	0.9969	0.9947
10	0.9979	0.9953	0.9857	0.9753	0.9966	0.9959
11	0.9991	0.9953	0.9881	0.9688	0.9975	0.9963
12	0.9971	0.9963	0.9888	0.9808	0.9977	0.9798
13	0.9985	0.9963	0.9910	0.9819	0.9979	0.9923
14	0.9991	0.9961	0.9930	0.9853	0.9986	0.9955
15	0.9989	0.9976	0.9937	0.9839	0.9989	0.9974
16	0.9998	0.9971	0.9934	0.9845	0.9985	0.9753
17	0.9988	0.9973	0.9948	0.9845	0.9990	0.9973
18	0.9992	0.9982	0.9963	0.9863	0.9994	0.9949
19	0.9989	0.9941	0.9965	0.9870	0.9997	0.9976
20	0.9989	0.9971	0.9963	0.9872	0.9999	0.9976
21	0.9998	0.9980	0.9978	0.9814	0.9998	0.9927
22	1.0000	0.9980	0.9971	0.9825	0.9991	0.9967
23	1.0000	0.9980	0.9973	0.9886	0.9997	0.9971
24	1.0000	0.9980	0.9974	0.9829	0.9993	0.9973
25	1.0000	0.9980	0.9979	0.9892	0.9993	0.9982
26	1.0000	0.9980	0.9984	0.9853	0.9998	0.9986
27	1	0,9982	0,9975	0,9892	1.0000	0,9982
28	1	0,9982	0,9985	0,9804	1.0000	0,9982
29	1	0,9980	0,9979	0,9876	1.0000	0,998
30	1	0,9980	0,9983	0,9906	1.0000	0,9982

Table 5. Sensitivity values obtained as a result of the training-test of the models

Epoch	Inception-v3		EfficientNet-B3		DenseNet-201	
	Train	Test	Train	Test	Train	Test
1	0.9603	0.9250	0.9035	0.9250	0.9619	0.9250
2	0.9851	0.9584	0.9188	0.9250	0.9843	0.9559
3	0.9898	0.9809	0.9251	0.9273	0.9900	0.9832
4	0.9923	0.9848	0.9354	0.9441	0.9920	0.9832
5	0.9925	0.9924	0.9534	0.9557	0.9932	0.9912
6	0.9939	0.9922	0.9648	0.9563	0.9936	0.9815
7	0.9958	0.9895	0.9736	0.9641	0.9938	0.9939
8	0.9964	0.9877	0.9805	0.9709	0.9954	0.9942
9	0.9966	0.9936	0.9844	0.9730	0.9969	0.9947
10	0.9979	0.9953	0.9858	0.9771	0.9966	0.9959
11	0.9991	0.9953	0.9880	0.9680	0.9976	0.9961
12	0.9971	0.9963	0.9886	0.9805	0.9977	0.9797
13	0.9985	0.9963	0.9908	0.9814	0.9979	0.9924
14	0.9991	0.9961	0.9928	0.9854	0.9986	0.9955
15	0.9989	0.9977	0.9939	0.9840	0.9989	0.9975
16	0.9998	0.9971	0.9937	0.9846	0.9984	0.9754
17	0.9987	0.9973	0.9946	0.9848	0.9990	0.9973
18	0.9992	0.9982	0.9963	0.9865	0.9993	0.9949
19	0.9990	0.9941	0.9965	0.9873	0.9997	0.9977
20	0.9989	0.9971	0.9963	0.9871	0.9999	0.9977
21	0.9998	0.9980	0.9977	0.9811	0.9998	0.9928
22	1.0000	0.9980	0.9970	0.9826	0.9991	0.9967
23	1.0000	0.9980	0.9972	0.9887	0.9997	0.9971
24	1.0000	0.9980	0.9974	0.9832	0.9993	0.9973
25	1.0000	0.9980	0.9979	0.9893	0.9993	0.9982
26	1.0000	0.9980	0.9985	0.9854	0.9998	0.9986
27	1.0000	0.9982	0.9975	0.9893	1.0000	0.9982
28	1.0000	0.9982	0.9984	0.9807	1.0000	0.9982
29	1.0000	0.9980	0.9978	0.9875	1.0000	0.9980
30	1.0000	0.9980	0.9983	0.9906	1.0000	0.9982

Table 4. Loss values obtained as a result of the training-test of the models

| Epoch	Inception-v3		EfficientNet-B3		DenseNet-201	
	Train	Test	Train	Test	Train	Test
1	0.1290	0.3641	0.5234	0.2883	0.1296	0.3729
2	0.0484	0.1884	0.3789	0.2780	0.0505	0.1804
3	0.0338	0.0636	0.3124	0.2228	0.0343	0.0577
4	0.0238	0.0483	0.2555	0.2173	0.0260	0.0572
5	0.0212	0.0313	0.1653	0.4024	0.0212	0.0335
6	0.0189	0.0244	0.1225	0.3183	0.0198	0.0655
7	0.0137	0.0483	0.0909	0.1490	0.0197	0.0207
8	0.0103	0.0459	0.0656	0.1149	0.0132	0.0214
9	0.0111	0.0239	0.0497	0.0919	0.0106	0.0197
10	0.0064	0.0182	0.0446	0.1177	0.0100	0.0146
11	0.0035	0.0287	0.0365	0.1337	0.0071	0.0201
12	0.0084	0.0237	0.0390	0.0715	0.0080	0.0638
13	0.0055	0.0215	0.0267	0.0708	0.0055	0.0344
14	0.0029	0.0297	0.0222	0.0598	0.0040	0.0218
15	0.0039	0.0164	0.0201	0.0608	0.0046	0.0124
16	0.0014	0.0175	0.0190	0.0762	0.0041	0.0748
17	0.0046	0.0188	0.0164	0.0676	0.0031	0.0156
18	0.0021	0.0118	0.0127	0.0569	0.0028	0.0228
19	0.0042	0.0234	0.0120	0.0541	0.0014	0.0168
20	0.0040	0.0159	0.0116	0.0564	0.0006	0.0184
21	0.0003	0.0157	0.0085	0.0853	0.0006	0.0334
22	0.0000	0.0164	0.0097	0.0832	0.0028	0.0164
23	0.0000	0.0170	0.0098	0.0581	0.0011	0.0176
24	0.0000	0.0176	0.0079	0.0931	0.0025	0.0136
25	0.0000	0.0179	0.0074	0.0513	0.0024	0.0114
26	0.0000	0.0184	0.0047	0.0720	0.0009	0.0110
27	0.0000	0.0166	0.0076	0.0466	0.0000	0.0142
28	0.0000	0.0168	0.0055	0.0972	0.0000	0.0146
29	0.0000	0.0179	0.0062	0.0526	0.0000	0.0149
30	0.0000	0.0181	0.0052	0.0515	0.0000	0.0153

Table 6. Specificity values obtained as a result of the training-test of the models

Epoch	Inception-v3		EfficientNet-B3		DenseNet-201	
	Train	Test	Train	Test	Train	Test
1	0.9627	0.9250	0.9031	0.9250	0.9593	0.9250
2	0.9851	0.9582	0.9173	0.9250	0.9844	0.9549
3	0.9899	0.9807	0.9243	0.9246	0.9902	0.9832
4	0.9925	0.9848	0.9373	0.9434	0.9919	0.9832
5	0.9927	0.9926	0.9545	0.9567	0.9932	0.9910
6	0.9941	0.9922	0.9642	0.9553	0.9937	0.9813
7	0.9958	0.9893	0.9737	0.9650	0.9937	0.9945
8	0.9963	0.9883	0.9800	0.9707	0.9956	0.9942
9	0.9966	0.9936	0.9844	0.9729	0.9969	0.9947
10	0.9979	0.9953	0.9857	0.9748	0.9967	0.9959
11	0.9991	0.9953	0.9880	0.9703	0.9976	0.9963
12	0.9971	0.9963	0.9891	0.9811	0.9977	0.9799
13	0.9985	0.9963	0.9905	0.9822	0.9979	0.9924
14	0.9991	0.9961	0.9928	0.9854	0.9986	0.9955
15	0.9989	0.9977	0.9935	0.9840	0.9989	0.9975
16	0.9998	0.9971	0.9934	0.9844	0.9984	0.9756
17	0.9988	0.9973	0.9946	0.9838	0.9990	0.9973
18	0.9992	0.9982	0.9963	0.9855	0.9993	0.9949
19	0.9989	0.9941	0.9964	0.9869	0.9997	0.9977
20	0.9989	0.9971	0.9963	0.9877	0.9999	0.9977
21	0.9998	0.9980	0.9976	0.9816	0.9998	0.9928
22	1.0000	0.9980	0.9971	0.9824	0.9991	0.9967
23	1.0000	0.9980	0.9971	0.9887	0.9997	0.9971
24	1.0000	0.9980	0.9973	0.9830	0.9993	0.9973
25	1.0000	0.9980	0.9979	0.9895	0.9993	0.9982
26	1.0000	0.9980	0.9983	0.9854	0.9998	0.9986
27	1.0000	0.9982	0.9975	0.9895	1.0000	0.9982
28	1.0000	0.9982	0.9985	0.9801	1.0000	0.9982
29	1.0000	0.9980	0.9978	0.9883	1.0000	0.9980
30	1.0000	0.9980	0.9982	0.9904	1.0000	0.9982

5

In addition, accuracy, loss, sensitivity, and specificity graphs for each model are shown in Fig. 3. When the results are given in Table 3 and Fig. 3a are examined, it is seen that, after 30 epochs, the highest test success accuracy in classifying breast cancer was achieved with DenseNet-201 (99.81%) and the lowest test success accuracy with EfficientNet-B3 (99.06%). It is seen that the 99.80% test success accuracy achieved with the Inception-v3 model is very close to the result obtained in the DenseNet-201 model.
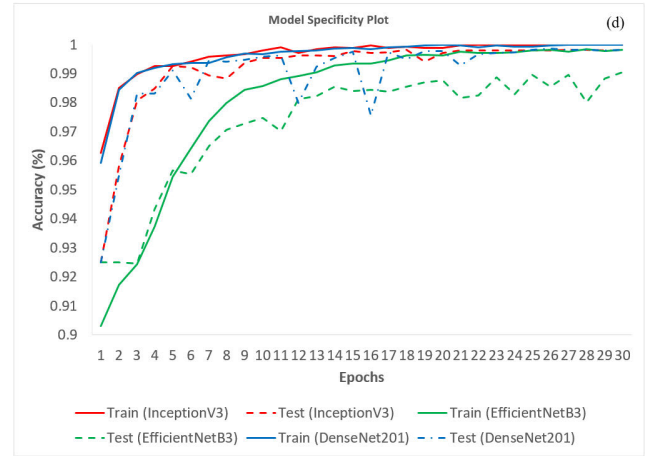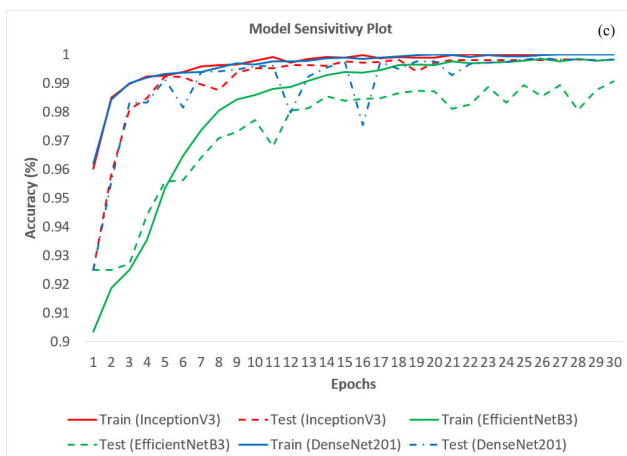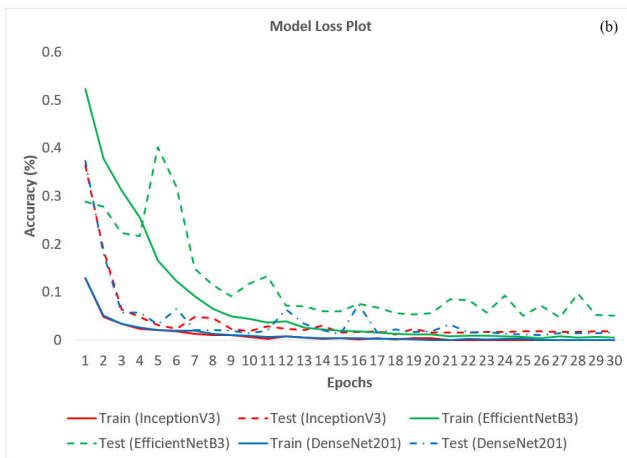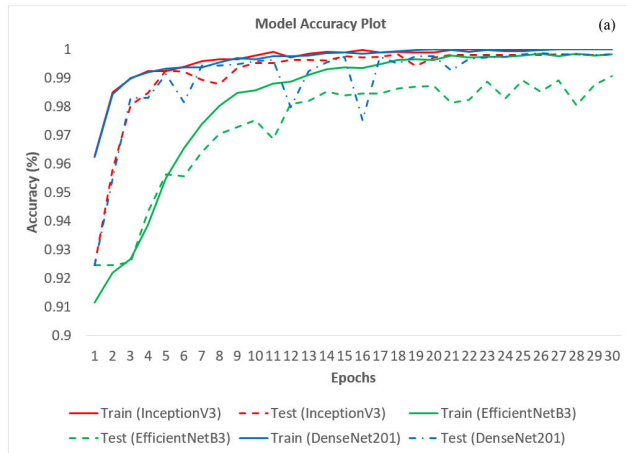
Fig. 3. Graphs obtained as a result of the training-test of the models (a) accuracy, (b) loss, (c) sensitivity, (d) specificity.

Similarly, when the results are given in Table 4 and Fig. 3b are examined, it is seen that the lowest test loss value was reached with DenseNet-201 (0.0153) and the highest test loss value was reached with EfficientNet-B3 (0.0515). It is seen that the test loss value of 0.0181 reached with the Inception-v3 model is close to the result obtained in the DenseNet-201 model. When the sensitivity results given in Table 5 and Fig. 3c and the specificity results given in Table 6 and Fig. 3d are examined, it is seen that the results obtained are very close to the success accuracy rates. This shows that model trainings have been carried out successfully. In addition, confusion matrix was obtained according to the test dataset used in the training of Inception-v3, EfficientNet-B3, and DenseNet-201 models, and the results are given in Table 7.

Table 7. Confusion matrix values of the models

Models	Cancer		Normal	
	True Positive	False Negative	True Negative	False Positive
Inception-v3	375	9	4709	1
EfficientNet-B3	339	45	4707	3
DenseNet-201	375	9	4710	0

When Table 7 is examined; It was observed that DenseNet-201 model achieved the most successful classification rate by classifying 375 as true, 9 as false of 384 cancerous breast images, and all 4710 normal breast images as true. Unlike the DenseNet-201 model, the Inception-v3 model was found to be the second-best model by misclassifying 1 of the normal breast images. It was observed that the most misclassification of cancerous and normal breast images was made in the EfficientNet-B3 model. Therefore, it has been observed that the DenseNet-201 model has superior classification performance in cancer/normal classification compared to other models. In addition, to test the model performance, the validation success rates of the Inception-v3, EfficientNet-B3, and DenseNet-201 model structures were tested by using the images of 14 patients in the validation dataset (10 subsets). The obtained validation test results are given in Table 8.

6

**Table 8.** Validation dataset final test results of 14 patients

Validation Dataset	Explanation	Number of Images	Accuracy (%)	Loss	Sensitivity (%)	Specificity (%)
Validation-1	T2	218	86.70	1.3121	87.05	87.05
Validation-2	ADC	23	72.29	2.0149	78.05	78.05
Validation-3	Diffusion	23	75.32	1.4136	80.98	80.98
Validation-4	Non-Contrast Fat Non-Suppressed T1	236	91.53	0.9646	92.19	92.19
Validation-5	Non-Contrast Fat Suppressed T1	575	85.57	1.3518	85.58	85.58
Validation-6	Contrast Fat Suppressed T1	2954	89.98	0.8567	90.05	90.05
Validation-7	Subtraction T1	2806	92.41	0.5936	92.44	92.44
Validation-8	Validation-1 – Validation-5	1075	86.23	1.3299	86.40	86.40
Validation-9	Validation-6 – Validation-7	5760	91.16	0.7285	91.16	91.16
Validation-10	Validation-1 – Validation-7	6835	89.36	0.9892	89.55	89.55

The validation dataset was compared by dividing it into 10 subsets as shown in Table 8. When Table 8 is examined, it is seen that the highest accuracy, sensitivity, and specificity rates and the lowest loss value were obtained from Subtraction T1 images in the Validation-7 subset, while the lowest accuracy, sensitivity, and specificity rates and the highest loss value were obtained from ADC images in the Validation-2 subset. When Table 8 is viewed from a wide perspective, an accuracy of 86.23% was achieved in the Validation-8 subset containing breast MRI images taken without contrast, and 91.16% in the Validation-9 subset containing breast MRI images taken with contrast. This remarkable situation shows that the results obtained from the images taken without contrast given in Validation-8 are close to the results obtained from images taken with contrast in Validation-9. Also noteworthy is that the accuracy rate (91.53%) obtained from Non-Contrast Fat Non-

Suppressed T1 images in the Validation-4 cluster is close to the Subtraction T1 accuracy rate (92.41%) in the Validation-7 cluster, which has the highest rate. This shows that with the proposed model structure used in the study, successful results can be obtained in non-contrast-applied breast MRI images. It is thought that obtaining successful results using non-contrast breast MRI images with the proposed model structure will bring a great innovation to the literature and the field of health.

3.2. Performance Test Results

With the proposed model structure, the performance test was applied to detect the cancer region from the breast MRI images taken from 14 patients who were not used in the study. The performance test applied is shown in the model diagram in Fig. 4.
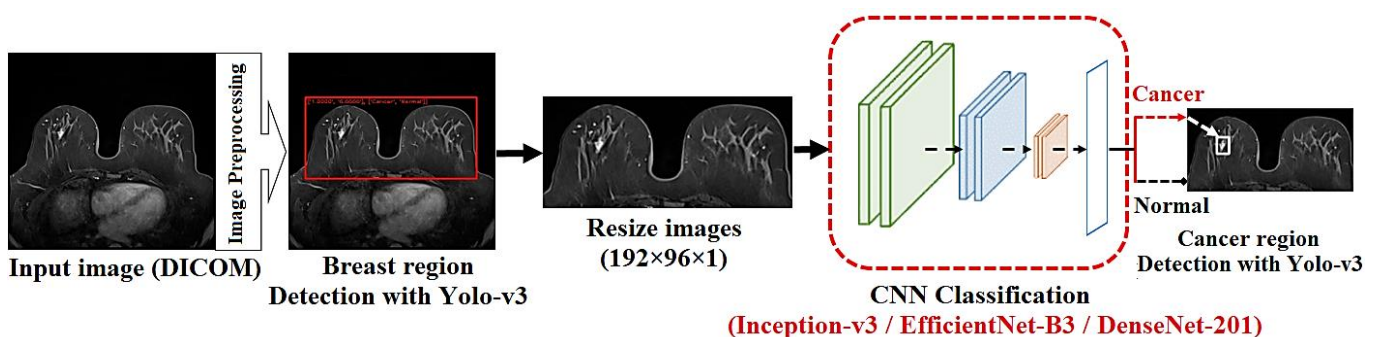
Fig. 4. Performance test model diagram.

When Fig. 4 is examined, breast MRI images taken in DICOM format of each patient were converted to JPG format for use in the model network. Then, the breast region was determined by using the breast region cross-section weights obtained with the Yolo-v3 algorithm and converted to 192×96×1 pixel dimensions. The resulting breast region was classified as cancer/normal using Inception-v3, EfficientNet-B3, and

DenseNet-201 models. As a result of the classification, the cancerous region on the breast image was determined using the Yolo-v3 cancer region cross-section weights and framed. The performance test results performed on 14 patients using the Yolo-v3 + DenseNet-201 model, where the best results were obtained, are given in Table 9.

7

Table 9. Performance test results of 14 patients

Patient No.	Cancer/ Normal	Total Number of Images	Number of Images Determined as Cancer/Normal by the Specialist Physician	Number of Images Detected Cancer/Normal with Yolo-v3 + DenseNet-201	Success Rate (%)
1		1169	35	33	94.29
2		1169	83	81	97.59
3		1169	14	10	71.43
4	Cancer	1164	47	38	80.85
5		933	12	10	83.33
6		1164	62	50	80.65
7		1169	116	82	70.69
Total		**7939**	**369**	**304**	**82.38**
8		1081	1081	1081	100
9		1169	1169	1169	100
10		1169	1169	1169	100
11	Normal	1169	1169	1155	98.80
12		1169	1169	1169	100
13		981	981	981	100
14		1152	1152	1152	100
Total		**7890**	**7890**	**7876**	**99.82**

When Table 9 is analyzed on a patient basis, the cancer region was correctly identified with different success rates in all 7 cancer patients. Incorrect cancer detection was made in only 1 of 7 normal patients. Therefore, 13 out of 14 patients were correctly identified and a success rate of 92.86% was achieved. When Table 9 is analyzed in terms of the number of images, 304 images of 369 cancer regions belonging to 7 patients with cancer were correctly detected and a success rate of 82.38% was achieved. Of the 7890 images of 7 normal patients, only 14 were incorrectly detected as cancer. The success rate of normal patients was found to be 99.82%. According to the performance test results made with the proposed model structure, high accuracy determinations were made and the sample images obtained from the test results are shown in Fig. 5. In Fig. 5b, a sample image of the patient (Patient No: 11) who was wrongly diagnosed with cancer is also included.
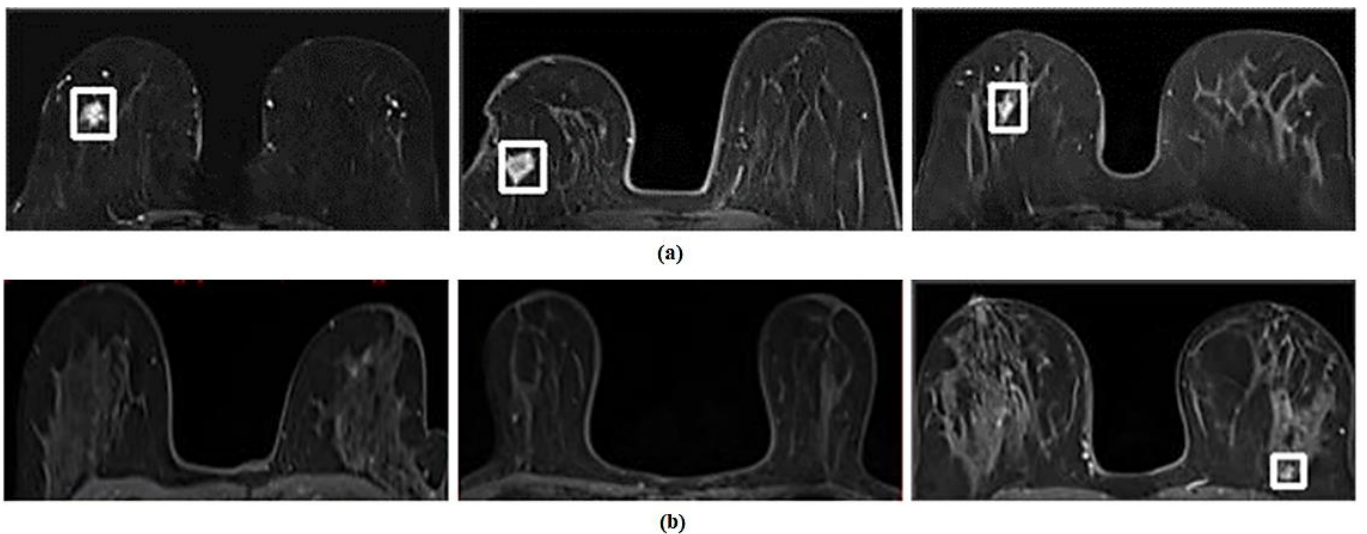
Fig. 5. Sample images of the performance test (a) Cancer, (b) Normal.

4. Discussion

In recent years, the results obtained on medical images using deep learning methods have shown more successful performances than traditional methods [8, 17]. In this study, a model structure based on artificial intelligence technology was developed to characterize diagnostic performance in individual breast DCE-MRI sequences. In addition, it is predicted that the need for contrast material can be reduced in breast MRI with the help of artificial intelligence technology.

Our proposed approach has been compared with many studies based on deep learning using MR images for breast cancer

8

classification. Our suggested approach is with 92.41% accuracy rate outperformed many studies [9, 10, 18-24]. Although there are studies showing a higher success rate than our model [25-27], the sensitivity and specificity of 92.44% in our model were found to be higher than these studies (Table 10).

Table 10. Comparison of the proposed model with existing methods in breast cancer

References	Model	Number of patients	Accuracy	Sensitivity	Specificity
[9]	3D RetinaNet	462	90.00%	95.00%	-
[10]	Random Forest + ResNet-50	133	89.00%	95.00%	74.00%
[18]	3D ResNet-50	536	85.00%	66.70%	93.20%
[19]	VGG-19 + SVM	1979	91.00%	90.00%	76.00%
[20]	DenseNet-121 DenseNet-169 InceptionResNet-V2 Inception-V3 NasNetMobile Xception	286	89.50%	74.50%	96.00%
[21]	ResNet-50	903	85.90%	76.90%	81.40%
[22]	CNN	130	87.70%	86.10%	91.20%
[23]	Transfer Learning	60	85.00%	89.00%	-
[24]	DenseNet ResNet	1794	85.80%	90.00%	82.60%
[25]	VGG-19	273	92.80%	89.50%	94.30%
[26]	Resnet-101	438	94.20%	74.40%	95.30%
[27]	Proposed approach + SVM	448	93.70%	95.60%	87.20%
Proposed approach	**Yolo-v3 + Inception-v3 Yolo-v3 + EfficientNet-B3 Yolo-v3 + DenseNet-201**	**64**	**92.41%**	**92.44%**	**92.44%**

In our study, unlike other studies, DCE-MRI sequences (T2, ADC, Diffusion, Non-Contrast Fat Non-Suppressed T1, Non-Contrast Fat Suppressed T1, Contrast Fat Suppressed T1, and Subtraction T1) were evaluated separately and a unique perspective was presented. Therefore, in our study, characterization of diagnostic performance in individual breast DCE-MRI sequences was emphasized rather than the accuracy rate directly obtained in other studies. Although the highest success rate was 92.41% with the Subtraction T1 sequence in our study, a high accuracy of 91.53% was also achieved in the Non-Contrast Fat Non-Suppressed T1 sequence. In this way, it was concluded that the correct diagnosis can be made in the images obtained without contrast application.

5. CONCLUSIONS

In this study, a total of 32302 breast MRI images taken from 64 (32 normal and 32 cancer) patients were used to detect breast cancer regions. A novel model structure has been created to detect breast area and cancer on the breast. In the model structure, Yolo-v3 algorithm was used for region detection, and Inception-v3, EfficientNet-B3, and DenseNet-201 model structures were used for classification. To validate the models, DCE-MRI sequences (T2, ADC, Diffusion, Non-Contrast Fat Non-Suppressed T1, Non-Contrast Fat Suppressed T1, Contrast Fat Suppressed T1, and Subtraction T1) were evaluated separately.

The best detection and classification according to these model structures were performed with the Yolo-v3 + DenseNet-201 model, and 92.41 accuracy, 0.5936 loss, 92.44% sensitivity, and 92.44% specificity rates were obtained, respectively. In addition, the proposed best model structure was subjected to performance testing with 14 patients who were never used. The correct diagnosis was made in 13 of 14 (7 cancer, 7 normal) patients, and a success rate of 92.86% was achieved.

As a result, an artificial intelligence model structure based on 7 different DCE-MRI sequences has been developed for breast cancer detection and classification. The proposed artificial intelligence model may be useful in increasing the diagnostic confidence and efficiency of radiologists in breast cancer detection and classification. In future studies, different CNN models can be used to determine the cancer region in real-time in breast cancer detection and classification.

REFERENCES
[1] F. Pesapane, C. Volonté, M. Codari, and F. Sardanelli, "Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States," *Insights into imaging*, vol. 9, no. 5, pp. 745-753, 2018, doi: 10.1007/s13244-018-0645-y.
[2] M. Codari, S. Schiaffino, F. Sardanelli, and R. M. Trimboli, "Artificial intelligence for breast MRI in 2008–2018: a systematic mapping review," *American Journal of Roentgenology*, vol. 212, no. 2, pp. 280-292, 2019, doi: 10.2214/AJR.18.20389.

[3] D. Özdemir, and N. N. Arslan, "Analysis of Deep Transfer Learning Methods for Early Diagnosis of the Covid-19 Disease with Chest X-ray Images," *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, vol. 10, no. 2, 628-640, 2022, doi: 10.29130/dubited.976118.

[4] R. M. Warren, L. Pointon, D. Thompson, R. Hoff, F. J. Gilbert, A. Padhani, D. Easton, S. R. Lakhani, and M. O. Leach, "Reading protocol for dynamic contrast-enhanced MR images of the breast: sensitivity and specificity analysis," *Radiology*, vol. 236, no. 3, pp. 779-788, 2005, doi: 10.1148/radiol.2363040735.

[5] Y. Jiang, A. V. Edwards, and G. M. Newstead, "Artificial intelligence applied to breast MRI for improved diagnosis," *Radiology*, vol. 298, no. 1, pp. 38-46, 2021, doi: 10.1148/radiol.2020200292.

[6] R. M. Mann, C. K. Kuhl, and L. Moy, "Contrast-enhanced MRI for breast cancer screening," *Journal of Magnetic Resonance Imaging*, vol. 50, no. 2, pp. 377-390, 2019, doi: 10.1002/jmri.26654.

[7] B. Reig, L. Heacock, K. J. Geras, and L. Moy, "Machine learning in breast MRI," *Journal of magnetic resonance imaging*, vol. 52, no. 4, pp. 998-1018, 2020, doi: 10.1002/jmri.26852.

[8] M. Adachi, T. Fujioka, M. Mori, K. Kubota, Y. Kikuchi, W. Xiaotong, J. Oyama, K. Kimura, G. Oda, T. Nakagawa, H. Uetake, and U. Tateishi, "Detection and diagnosis of breast cancer using artificial intelligence based assessment of maximum intensity projection dynamic contrast-enhanced magnetic resonance images," *Diagnostics*, vol. 10, no. 5, pp. 330, 2020, doi: 10.3390/diagnostics10050330.

[9] F. Ayatollahi, S. B. Shokouhi, R. M. Mann, and J. Teuwen, "Automatic breast lesion detection in ultrafast DCE-MRI using deep learning," *Medical Physics*, vol. 48, no. 10, pp. 5897-5907, 2021, doi: 10.1002/mp.15156.

[10] J. Zhou, Y. Zhang, K. T. Chang, K. E. Lee, O. Wang, J. Li, Y. Lin, Z. Pan, P. Chang D. Chow, M. Wang, and M. Y. Su, "Diagnosis of benign and malignant breast lesions on DCE-MRI by using radiomics and deep learning with consideration of peritumor tissue," *Journal of Magnetic Resonance Imaging*, vol. 51, no. 3, pp. 798-809, 2020, doi: 10.1002/jmri.26981.

[11] B. Reig, "Radiomics and deep learning methods in expanding the use of screening breast MRI," *European Radiology*, vol. 31, no. 8, pp. 5863-5865, 2021, doi: 10.1007/s00330-021-08056-9.

[12] Colab. Google Colaboratory [online]. Available: https://colab.research.google.com. [Accessed: 10 Sep. 2022]

[13] J. Redmon, and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint*, pp. 1-6, 2018, doi: 10.48550/arXiv.1804.02767.

[14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.

[15] M. Tan, and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *In International conference on machine learning, Proceedings of Machine Learning Research (PMLR)*, 2019, pp. 6105-6114.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700-4708.

[17] S. Mirbagheri, and M. Momeni, "A Hybrid Deep Learning Methodology for Breast Cancer Diagnosis using Magnetic Resonance Images," *Research Square*, pp. 1-16, 2022 doi: 10.21203/rs.3.rs-1604535/v1.

[18] S. Joo, E. S. Ko, S. Kwon, E. Jeon, H. Jung, J. Y Kim,. M. J. Chung and Y. H. Im, "Multimodal deep learning models for the prediction of pathologic response to neoadjuvant chemotherapy in breast cancer," *Scientific reports*, vol. 11, no. 1, pp. 1-8, 2021, doi: 10.1038/s41598-021-98408-8.

[19] Q. Hu, H. M. Whitney, H. Li, Y. Ji, P. Liu, and M. L. Giger, "Improved classification of benign and malignant breast lesions using deep feature maximum intensity projection MRI in breast cancer diagnosis using dynamic contrast-enhanced MRI," *Radiology: Artificial Intelligence*, vol. 3, no. 3, pp. 1-9, 2021, doi: 10.1148/ryai.2021200159.

[20] T. Fujioka, Y. Yashima, J. Oyama, M. Mori, K. Kubota, L. Katsuta, K. Kimura, E. Yamaga, G. Oda, T. Nakagawa, Y. Kitazume, and U. Tateishi, "Deep-learning approach with convolutional neural network for classification of maximum intensity projections of dynamic contrast-enhanced breast magnetic resonance imaging," *Magnetic Resonance Imaging*, vol. 75, pp. 1-8, 2021, doi: 10.1016/j.mri.2020.10.003.

[21] L. Wang, L. Chang, R. Luo, X. Cui, H. Liu, H. Wu, Y. Chen, Y. Zhang, C. Wu, F. Li, H. Liu, W. Guan, and D. Wang, "An artificial intelligence system using maximum intensity projection MR images facilitates classification of non-mass enhancement breast lesions," *European Radiology*, vol. 32, no. 7, pp. 4857-4867, 2022, doi: 10.1007/s00330-022-08553-5.

[22] Y. Wu, J. Wu, J. Dou, N. Rubert, Y. Wang, and J. Deng, "A deep learning fusion model with evidence-based confidence level analysis for differentiation of malignant and benign breast tumors using dynamic contrast enhanced MRI," *Biomedical Signal Processing and Control*, vol. 72, pp. 1-11, 2022, doi: 10.1016/j.bspc.2021.103319.

[23] G. Ye, S. He, R. Pan, L. Zhu, D. Zhou, and R. Lu, "Research on DCE-MRI Images Based on Deep Transfer Learning in Breast Cancer Adjuvant Curative Effect Prediction," *Journal of Healthcare Engineering*, vol. 2022, pp. 1-10, 2022, doi: 10.1155/2022/4477099.

[24] L. A. Kapsner, S. Ohlmeyer, L. Folle, F. B. Laun, A. M. Nagel, A. Liebert, H. Schreiter, M. W. Beckmann, M. Uder, E. Wenkel, S. Bickelhaupt, "Automated artifact detection in abbreviated dynamic contrast-enhanced (DCE) MRI-derived maximum intensity projections (MIPs) of the breast," *European Radiology*, vol. 32, no. 9, pp. 5997-6007, 2022, doi: 10.1007/s00330-022-08626-5.

[25] S. Eskreis-Winkler, N. Onishi, K. Pinker, J. S. Reiner, J. Kaplan, E. A. Morris, and E. J. Sutton, "Using deep learning to improve nonsystematic viewing of breast cancer on MRI," *Journal of Breast Imaging*, vol. 3, no. 2, pp. 201-207, 2021, doi: 10.1093/jbi/wbaa102.

[26] M. Z. Liu, C. Swintelski, S. Sun, M. Siddique, E. Desperito, S. Jambawalikar, and R. Ha, "Weakly supervised deep learning approach to breast MRI assessment," *Academic Radiology*, vol. 29, pp. S166-S172, 2022, doi: 10.1016/j.acra.2021.03.032.

[27] P. Jaglan, R. Dass, and M. Duhan, "An automatic and efficient technique for tumor location identification and classification through breast MR images," *Expert Systems with Applications*, vol. 185, pp. 1-11, 2021, doi:10.1016/j.eswa.2021.115580.