The Utility of Deep Learning in Breast Ultrasonic Imaging: A Review

Tomoyuki Fujioka [1], Mio Mori [1,*], Kazunori Kubota [1,2], Jun Oyama [1], Emi Yamaga [1], Yuka Yashima [1], Leona Katsuta [1], Kyoko Nomura [1], Miyako Nara [1,3], Goshi Oda [4], Tsuyoshi Nakagawa [4], Yoshio Kitazume [1] and Ukihide Tateishi [1]

[1] Department of Diagnostic Radiology, Tokyo Medical and Dental University, Tokyo 113-8510, Japan; fjokmrad@tmd.ac.jp (T.F.); kubotard@dokkyomed.ac.jp (K.K.); ooymmrad@tmd.ac.jp (J.O.); ymgdrnm@tmd.ac.jp (E.Y.); 11.ruby.89@gmail.com (Y.Y.); leonah@jcom.home.ne.jp (L.K.); nomura.kyoko@kameda.jp (K.N.); miyako641@gmail.com (M.N.); ktzmmrad@tmd.ac.jp (Y.K.); ttisdrnm@tmd.ac.jp (U.T.)

[2] Department of Radiology, Dokkyo Medical University, Tochigi 321-0293, Japan

[3] Department of Breast Surgery, Tokyo Metropolitan Cancer and Infectious Diseases Center Komagome Hospital, Tokyo 113-8677, Japan

[4] Department of Surgery, Breast Surgery, Tokyo Medical and Dental University, Tokyo 113-8510, Japan; oda.srg2@tmd.ac.jp (G.O.); nakagawa.srg2@tmd.ac.jp (T.N.)

* Correspondence: m_mori_116@yahoo.co.jp; Tel.: +81-3-5803-5311; Fax: +81-3-5803-0147

check for updates

Abstract: Breast cancer is the most frequently diagnosed cancer in women; it poses a serious threat to women's health. Thus, early detection and proper treatment can improve patient prognosis. Breast ultrasound is one of the most commonly used modalities for diagnosing and detecting breast cancer in clinical practice. Deep learning technology has made significant progress in data extraction and analysis for medical images in recent years. Therefore, the use of deep learning for breast ultrasonic imaging in clinical practice is extremely important, as it saves time, reduces radiologist fatigue, and compensates for a lack of experience and skills in some cases. This review article discusses the basic technical knowledge and algorithms of deep learning for breast ultrasound and the application of deep learning technology in image classification, object detection, segmentation, and image synthesis. Finally, we discuss the current issues and future perspectives of deep learning technology in breast ultrasound.

Keywords: breast; ultrasound; deep learning; machine learning; artificial intelligence; neural network

1. Introduction

Breast cancer is the most common cancer and the second leading cause of cancer-related death in women [1]. Owing to its advantages, such as safety, convenience, and low cost, ultrasound is used to detect and diagnose breast lesions when abnormalities are identified by other imaging modalities or on palpation [2]. In addition, breast ultrasound is expected to emerge as a complementary screening method for women with mammographically dense breasts, and the screening practice is expected to detect tumors at an early stage and reduce breast cancer mortality in women [3].

The use of breast ultrasound is increasing, and radiologists and clinicians spend significant time examining huge volumes of breast ultrasonic images. This has become a major problem in many countries because it leads to increased medical costs and worsens the patient case. Although breast ultrasonic imaging is performed using the Breast Imaging Reporting and Data System [4], it is difficult to maintain control over the accuracy of diagnostic imaging in areas in which experts are scarce.

In recent years, artificial intelligence (AI), especially deep learning methods, has accomplished outstanding performance in automatic speech recognition, image recognition, and natural language processing. The applications of deep learning in biomedical fields include all the medical levels, from genomic applications such as gene expression to public medical health management such as prediction of demographic information or infectious disease epidemics [5,6]. During the previous few decades, advances in high-throughput technology have significantly increased the amount of biomedical data. These biomedical data require effective and efficient computational tools for their storage, analyses, and interpretation. Deep learning-based algorithmic frameworks can resolve these challenges [7,8].

Furthermore, research on deep learning technology has been actively increasing in the field of medical imaging. It has been applied to radiologic images, for instance, to detect tuberculosis on chest radiographs, detect and diagnose lung nodules on chest CT, and segment brain tumor on MRI [9–11]. Furthermore, deep learning has proven useful in the field of breast imaging, and deep learning-based diagnostic support systems are becoming more commonly available [12–17]. The technology is also used in pathological imaging and is reportedly useful in counting the number of cells or mitosis, grading tumor tissue, segmenting the nucleus, and estimating the risk of recurrence with Oncotype DX [18–21].

Deep learning has also been used in breast ultrasonic imaging and will have a great influence in the future. Effective use of deep learning is expected to assist clinicians and help improve the standard of care. Health care workers involved in breast care must understand the current state of deep learning in breast ultrasound and consider how it will evolve.

This review discusses the basic technical knowledge and algorithms of deep learning for breast ultrasound and its application in image classification, segmentation, and image synthesis. Finally, we discuss current issues and future perspectives of deep learning in breast ultrasound.

## 2. What Are AI, Machine Learning, and Deep Learning?

AI is a broad area of computer science related to building smart machines that can perform tasks that normally require human intelligence [22]. Machine learning is a term introduced by Arthur Samuel in 1959 to describe one of the technologies of AI (Figure 1). Machine learning provides a system with the ability to automatically learn and improve from experience without explicit programming. In classic machine learning, expert humans discern and encode features that appear distinctive in the data, and statistical techniques are used to organize or segregate the data based on these features [22,23].
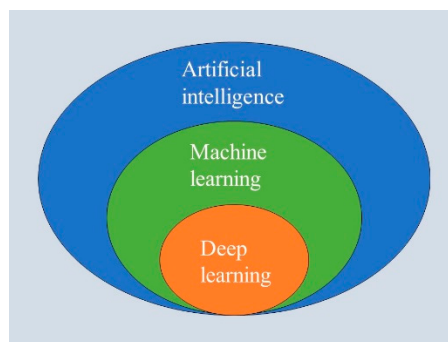
Figure 1. Schematic diagram of artificial intelligence, machine learning, and deep learning.

Deep learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning (Figure 1). It can make flexible decisions according to the situation by learning a large amount of data and automatically extracting common feature quantities [23,24]. Deep learning can generally be viewed as applying single-layer traditional machine learning methods (i.e., projection, convolution, hyperplane separation, and dot product) in a sequential, layer-wise manner, thus resulting in computational architectures with a depth dimension. Fully automatic training via the back-propagation algorithm is based on gradient descent, where graphics processing units are generally necessary to manage computational complexity [24].

Deep learning has dramatically improved research in areas such as speech recognition, visual image recognition, and object detection [24]. In image processing, the deep learning architecture called a convolutional neural network (CNN) has become its mainstream feature. Several types of CNNs have been developed. CNN consists of input and output layers, along with the important components of the convolution, max pooling, and fully connected layers (Figure 2) [25]. CNN typically includes multiple convolutional and pooling layers that facilitate the learning of more and more abstract features. The convolution layer extracts a feature from the input image and passes these results to the next layer. Convolution uses small squares of the input data to learn the features of the image, thus maintaining the relationships between the pixels and resulting in activation. Repeated application of the same filter to the input produces a map of activation, a feature map, which reveals the location and intensity of the features detected in the input, such as the image [25,26]. The pooling layers section reduces the spatial size of the activation maps for minimizing the likelihood of overfitting. Spatial pooling is called subsampling or downsampling, which reduces the dimensionality of each map but retains important information. Although there are several types of spatial pooling, max pooling has been the most common type [25,26]. The purpose of a fully connected layer is to obtain the results of the convolutional/pooling process and use them to classify the images into labels. The output of the convolution/pooling is flattened into a single vector of values. Fully connected layers connect all neurons in one layer to all neurons in another layer. The inputs from the neurons in the previous layer to those in the next layer are combined into a signal by the connection, a linear transformation process, and the signal is then output to the next layer of neurons via an activation function, a non-linear transformation. The rectified linear unit (Relu) function is commonly used as the activation function [25,26]. The output layer is the last layer that produces the given outputs for the program. The neurons are designed to rationalize and improve the end result of the iterative process, and the softmax function is typically used for multiple classification tasks [22,25].
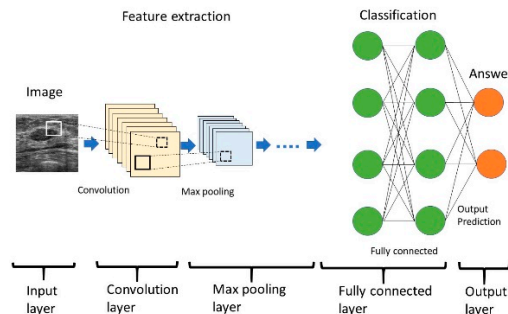
Figure 2. Structure of a convolutional neural network.

With increasing depth and complexity, CNN offers amazing performance evolution compared with traditional computer vision technology [25]. CNN has attracted significant attention since winning the international image classification competition ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012 [22,23].

AI is a broad area of computer science. Machine learning provides a system with the ability to automatically learn and improve from experience without explicit programming. Deep learning is part of a broader family of machine learning methods based on artificial neural networks.

CNN consists of multiple layers. Input data are first processed with convolutional and max pooling layers for feature extraction. These data are further processed in fully connected layers for classification, and the final prediction is output. Layers between the input and output layers are called hidden layers.

Learning methods include supervised learning, which learns a large amount of data and automatically acquires its characteristics; unsupervised learning, which classifies data in various dimensions; and reinforcement learning, which obtains the correct answer via a repeated trial-and-error method.

We must collect datasets (training, validation, and test data) to perform deep learning. The model first fits the training dataset. This is a set of examples used to match model parameters (such as the

weight of connections between neurons in an artificial neural network). The fitted model is then used to predict the response of the observations of the second dataset called the validation dataset. Finally, the test dataset is used to provide an unbiased evaluation of the final model that fits the training dataset [22].

Several techniques are available for efficient deep learning. Data augmentation is one solution when insufficient enough training data are available for constructing a deep learning model. Image datasets can be augmented via flipping, cropping, rotating, and adjusting the contrast, sharpness, and white balance. The diversification of training datasets helps prevent overfitting and enhances model generalization [25]. Transfer learning is an efficient learning method that makes one trained model useful for another area. Transfer learning has the advantages of being able to learn quickly and achieving high accuracy even with small amounts of data [25].

3. Development of AI Research on Breast Ultrasound

The development of AI research on breast ultrasound has led to an increase in publications in this area. PubMed was searched through 31 October 2019 to extract articles on imaging and AI used for breast ultrasound using the following search query: "artificial intelligence" or "machine learning" OR "deep learning" AND "breast ultrasound." Among these articles, only those that described studies on AI for breast imaging were selected.

The annual number of publications did not change from 2010 to 2017, but it increased significantly from 2018 to 2020, with more than 20 publications in 2018 and more than 40 publications in 2019. Meanwhile, 50 papers were published in the first 10 months of 2020 (Figure 3). Many of these studies described imaging classification, object detection, segmentation, and synthetic imaging of breast lesions.
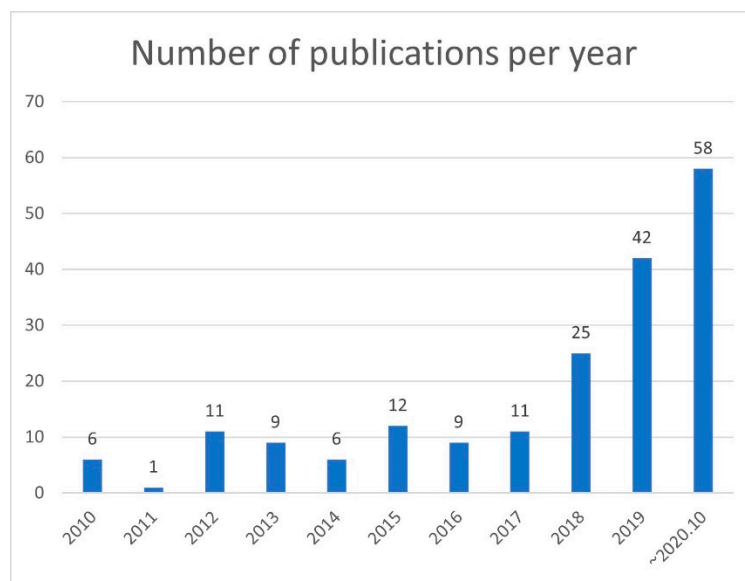
Figure 3. Number of publications per year.

The number of publications indexed on PubMed was obtained using the following search query: "artificial intelligence" or "machine learning" OR "deep learning" AND "breast ultrasound." PubMed was accessed on 31 October 2020. Since 2018, the number of papers on AI of breast ultrasound has increased rapidly.

4. Image classification

Image classification is a method for identifying and predicting what an image represents (Figure 4). The development of machine learning, especially deep learning with CNN, has facilitated the creation of highly accurate models. Since AlexNet [27] won the ILSVRC with overwhelming results, deep learning techniques have acquired a leading role in image classification. Subsequently, capable networks such

as VGGNet [28], GoogLeNet [29], ResNet [30], and DenseNet [31] have been developed successively with deeper layers of CNN.
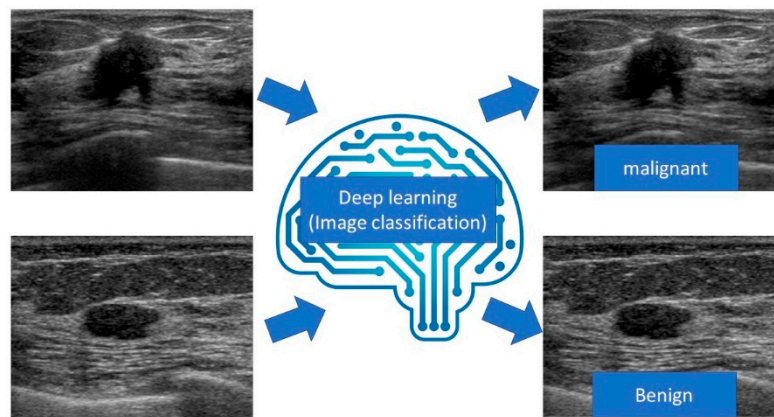
Figure 4. Diagram of image classification for breast ultrasound.

In the image classification of breast ultrasound, many reports discussed the distinction between benign and malignant lesions on B-mode images (Table 1). Han et al. collected a large number of B-mode images to train a deep neural network with 3154 malignant and 4254 benign samples using GoogLeNet to distinguish the malignancy of breast masses on ultrasound. They reported that the deep learning model had an accuracy of 91%, a sensitivity of 86%, a specificity of 93%, and an area under the curve (AUC) of >0.9 [32].

Table 1. Deep learning models for image classification.

Purpose (Type of Image)	Model	Number of Training Set Images	Number of Test Set Images	Result	Study
Breast lesions (B-mode images)	GoogLeNet	6579	829	Sensitivity: 86% Specificity: 96% Accuracy: 90% AUC: 0.9	[32]
Breast lesions (B-mode image)	GoogLeNet Inception v2	937	120	Sensitivity: 95.8% Specificity: 87.5% Accuracy: 92.5% AUC: 0.913	[33]
Breast lesions via CAD (B-mode image)	Koios DS	Over 400,000	900	AUC without CAD: 0.83 AUC with CAD: 0.87	[34]
Breast lesions (SWE image)	PGBM and RBM	227	Five-fold cross-validation	Sensitivity: 88.6% Specificity: 97.1% Accuracy: 93.4% AUC: 0.947	[35]
Breast lesions (SWE image)	DenseNet 169	304	73	Sensitivity: 85.7% Specificity: 78.9% AUC: 0.898	[36]
Axillary lymph nodes (B-mode image)	VGG-M model	118	Five-fold cross-validation	Sensitivity: 84.9% Specificity: 87.7% Accuracy: 86.4% AUC: 0.937	[37]

CAD, computer-assisted diagnosis; PGBM, point-wise gated Boltzmann machine; RBM, restricted Boltzmann machine; AUC, area under the curve.

We also collected B-mode images of 480 benign and 467 malignant masses as training data. A deep learning model was constructed using the CNN architecture GoogLeNet Inception v2, and 48 images of benign masses and 72 images of malignant masses were analyzed as the test data. The deep learning model had an accuracy of 95.8%, a sensitivity of 87.5%, a specificity of 92.5%, and an AUC of 0.913, and its diagnostic performance was equal or superior to that of radiologists [33].

Mango et al. evaluated the utility of Koios DS Breast, a machine learning-based diagnostic support system (Koios, https://koiosmedical.com/), by performing reading tests using the ultrasound images of 900 breast lesions. Among 15 physicians, the mean reader AUC for cases reviewed using ultrasound only was 0.83 versus 0.87 for ultrasound plus Koios DS Breast. Thus, Koios DS Breast improved the accuracy of sonographic breast lesion assessment while reducing inter- and intra-observer variability [34]. This diagnostic support system has been approved by the US Food and Drug Administration (FDA) and commercially launched.

Zhang et al. built a deep learning architecture and evaluated its performance in the differentiation of benign and malignant breast tumors on a set of shear wave elastography (SWE) images using 135 benign and 92 malignant tumors. The deep learning architecture displayed good classification performance, with an accuracy of 93.4%, a sensitivity of 88.6%, a sensitivity of 97.1%, and an AUC of 0.947 [35].

We also gathered 158 images of benign masses and 146 images of malignant masses as training data for SWE. Deep learning models were constructed using several state-of-the-art architectures. We analyzed the SWE images of 38 benign and 35 malignant masses as test data. The best model (DenseNet 169) had a sensitivity of 85.7%, a specificity of 78.9%, and an AUC of 0.898. The deep learning models had equal or better diagnostic performance than radiologist readings [36].

Coronado-Gutiérrez et al. developed quantitative ultrasound image analysis techniques using deep learning to noninvasively diagnose axillary lymph node involvement in breast cancer using 118 lymph node ultrasound images. The achieved accuracy of this method was 86.4%, and its sensitivity and specificity were 84.9 and 87.7%, respectively [37].

After inputting an image, the deep learning algorithm identifies and predicts what the image represents.

5. Object Detection

Object detection refers to detecting the location and category (class) of a defined object in an image (Figure 5). Similar to image classification, with the development of deep learning with CNN, faster and more accurate models have been created. In 2014, Regions with Convolutional Neural Networks (R-CNN) was developed to successfully apply the CNN algorithm to the task of object detection [38]. Subsequently, high-speed, high-precision object detection models such as Spatial Pyramid Pooling (SPP)-net [39], Fast R-CNN [40], Faster R-CNN [41], You Only Look Once (YOLO) [42], Single Shot MultiBox Detector (SSD) [43], feature pyramid networks [44], and RetinaNet [45] were developed.
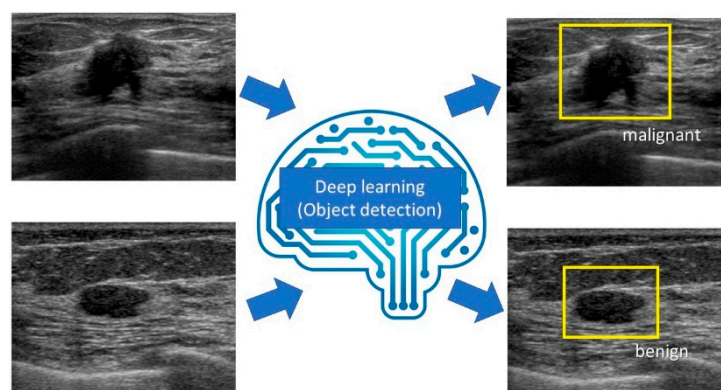
Figure 5. The diagram of object detection for breast ultrasound.

Cao et al. studied the existing state-of-the-art CNN methods (Fast R-CNN, Faster R-CNN, YOLO, and SSD) for breast lesion detection using breast ultrasound images. They collected a dataset consisting of 579 benign and 464 malignant lesions and submitted the ultrasound images to manual annotation by experienced clinicians. They found that YOLO and SSD perform significantly better than the other methods in detecting breast lesions [46].

Jiang et al. compared the diagnostic accuracy and interpretation time of screening automated breast ultrasound (ABUS) for women with dense breast tissue with and without the use of the deep learning-based computer-aided detection (CAD) system QVCAD (QViewMedical, https://www.qviewmedical.com/); 18 radiologists interpreted a test set of 185 screening ABUS images (52 with and 133 without breast cancer). The AUC was 0.848 with the CAD system, compared with 0.828 without the CAD system, which was statistically noninferior. The mean interpretation time was 3.55 min per case without the CAD system, compared with 2.4 min with the CAD system. This diagnostic support system has been approved by the FDA and reported to be useful in the US [47]. Yang et al. [48] and Xu et al. [49] also demonstrated that CAD systems improve the diagnostic performance for ABUS interpretation (Table 2).

**Table 2.** Deep learning models for object detection.

Purpose (Type of Image)	Model	Number of Training Set Images	Number of Test Set Images	Result	Study
Object detection of breast lesions (B-mode image)	SSD300	860	183	Precision rate: 96.89% Recall rate: 67.23% F1 score: 79.38%	[46]
Object detection of breast lesions by CAD (ABUS image)	QVCAD	Over 20,000	185	AUC without CAD: 0.828 AUC with CAD: 0.848	[47]
Object detection of breast lesions by CAD (ABUS image)	QVCAD	Over 20,000	1485	AUC without CAD: 0.88 AUC with CAD: 0.91 Sensitivity without CAD: 67% Sensitivity with CAD: 88%	[48]
Object detection of breast lesions by CAD (ABUS image)	QVCAD	Over 20,000	1000	AUC without CAD: 0.747 AUC with CAD: 0.784	[49]

CAD, computer-assisted diagnosis; AUC, area under the curve; ABUS, automated breast ultrasound.

After inputting an image, deep learning systems detect the location and category (class) of a defined object in the image.

6. Segmentation

Semantic segmentation is a method that can associate labels and categories with all pixels in an image and divide the object into multiple regions at the pixel level (Figure 6).
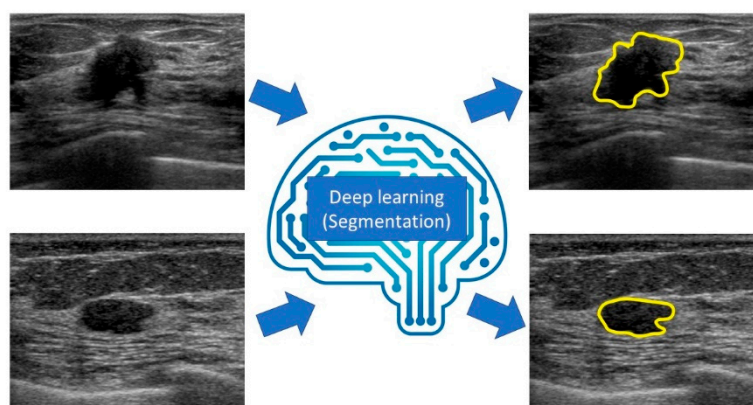
Figure 6. Diagram of segmentation for breast ultrasound.

Models based on the CNN architecture that enable high-precision and high-speed semantic segmentation, such as Fully Convolutional Network (FCN) [50], Segnet [51], and U-net [52], have been developed. They are widely used in various industries that require high-precision image mapping, such as medical imaging, autonomous driving, industrial inspection, and satellite imagery. Zhang et al. developed Residual-Dilated-Attention-Gate-U-net, an improved version of U-net, and performed segmentation of lesions on breast ultrasound images. Their model displayed better breast lesion

segmentation (precision rate: 88.58%, recall rate: 83.19%, F1 score: 84.78) than the traditional CNN model [53].

Hu et al. proposed combining a dilated fully convolutional network with a phase-based active contour model for automatic tumor segmentation. Their models exhibited high robustness, accuracy, and efficiency (Dice similarity coefficient: 88.97%, Hausdorff distance: 35.54 pixels, mean absolute deviation: 7.67 pixels), similar to the manual segmentation results [54].

Kumar et al. developed a Multi U-net algorithm based on CNNs that automatically segments breast masses. The model effectively segmented the breast masses, achieving a mean Dice coefficient of 0.82, a true positive fraction of 0.84, and a false positive fraction of 0.01 [55] (Table 3).

Table 3. Deep learning models for segmentation.

Purpose (Type of image)	Model	Number of Training Set Images	Number of Test Set Images	Result	Study
Segmentation of breast lesions (B-mode image)	RDAU-NET	857	205	Precision rate: 88.58% Recall rate: 83.19% F1 score: 84.78	[53]
Segmentation of breast lesions (B-mode image)	Combining DFCN with a PBAC model	400	170	Dice similarity coefficient: 88.97% Hausdorff distance: 35.54 pixels Mean absolute deviation: 7.67 pixels	[54]
Segmentation of breast lesions (B-mode image)	Multi U-net algorithm	372	61	Mean Dice coefficient: 0.82 True positive fraction: 0.84 False positive fraction: 0.01	[55]

CAD, computer-assisted diagnosis; RDAU-NET, Residual-Dilated-Attention-Gate-U-net; DFCN, dilated fully convolutional network; PBAC, phase-based active contour; AUC, area under the curve.

After inputting an image, the deep learning system associates labels and categories with every pixel in the image and divides the object into multiple areas at the pixel level.

7. Image synthesis

Image synthesis involves the generation of realistic images using a computer algorithm. This method is useful for improving image quality by reconstructing and generating training data by generating virtual images (Figure 7). One of the most interesting breakthroughs in the field was the advent of generative adversarial networks (GANs), which consist of effective machine learning frameworks used to train unsupervised generative models [56]. GAN is a special type of neural network model in which two networks are trained simultaneously, with one focusing on image generation and the other focusing on discrimination. Variational autoEncoder (VAE) is another effective generative model. Autoencoder (AE) is a type of neural network that learns the data encodings from the dataset in an unsupervised manner. Basically, VAE contains two parts; one is an encoder that learns efficient data encoding from the dataset and passes it into a bottleneck architecture and the other is a decoder that uses latent space in the bottleneck layer to regenerate images similar to that in the dataset. VAE differs from AE in that it provides a statistical manner to describe the samples of the dataset in latent space. Therefore, in VAE, the encoder outputs the probability distribution to the bottleneck layer instead of a single output value [57].

In recent years, a competent generative model based on deep learning was developed, and its effectiveness was confirmed. Image reconstruction using its technology makes it possible to generate images with few artifacts. In addition, a high-frame-rate image can be generated by performing image interpolation. Because the virtual image generated by the generative model does not contain personal information, it is considered useful for research and education that is open to the public [58]. We succeeded in generating realistic virtual images of breast tumors and virtual interpolation images of tumor development using deep convolutional GANs [59,60]. In addition, we used a GAN-based anomaly detection model to distinguish normal tissues from benign and malignant masses. The model had a sensitivity of 89.2%, a specificity of 90.2%, and an AUC of 0.936 [61]. Han et al. proposed a semi-supervised segmentation network based on GANs. They found that the model achieved higher segmentation accuracy than state-of-the-art semi-supervised segmentation methods [62].

After inputting a real image, the deep learning machine synthesizes realistic virtual images.
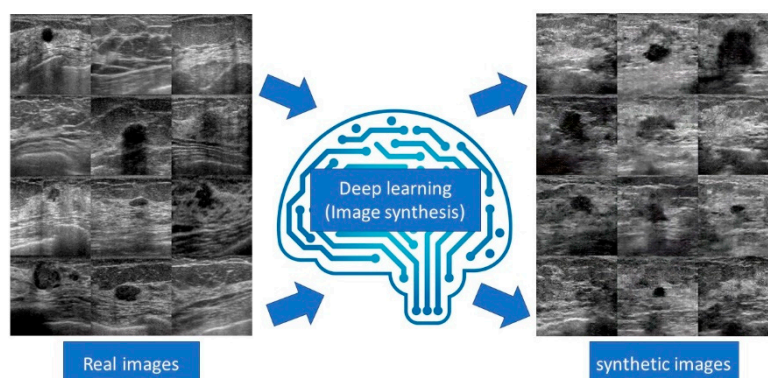

Figure 7. Diagram of image synthesis for breast ultrasound.

8. Discussion

In this study, we explained the basic technical knowledge and algorithms of deep learning in breast ultrasound and discussed current research trends. Deep learning techniques in breast ultrasound have evolved in recent years, and the number of related publications has increased. Research on deep learning models for breast ultrasound is actively being conducted, but a few computer-aided diagnosis support systems based on deep learning have been approved by public institutions. To our knowledge, three machine learning- or deep learning-based diagnostic imaging support systems for mammograms have been approved by the FDA: HealthMammo (Zebra Medical Vision, https://www.zebra-med.com/), ProFoundAI (iCAD, https://screenpoint-medical.com/), and Transpara (ScreenPoint Medical, https://screenpoint-medical.com/) [63]. These systems detect lesions such as masses and calcifications and score the malignant probability of the lesion. In breast ultrasound, as previously mentioned, two breast ultrasound systems, namely Smart Ultrasound (Koios) for 2D-B-mode imaging and QVCAD (QViewMedical) for ABUS, have been authorized by the FDA.

Contrarily, in Japan, no deep learning-based diagnostic imaging support systems for breast imaging have gained approval from the Pharmaceuticals and Medical Devices Agency. However, we believe that it will be possible to perform breast imaging in Japan using deep learning-based diagnostic imaging support systems by importing foreign-developed systems or developing a system unique to the country in the near future.

However, there are some obstacles to the introduction of deep learning-based diagnostic support systems for breast imaging. First, it is necessary to clarify who is responsible for the diagnosis. Previous studies revealed that even efficient deep learning models can misdiagnose lesions similarly to humans. At present, deep learning systems remain a supplementary modality for doctors, but there will be debates concerning whether these systems can replace doctors if their accuracy is improved in the future. If a deep learning system makes a serious mistake, it will be necessary to conclude whether the developer, seller, or doctor is responsible for the misdiagnosis.

Diagnostic systems based on deep learning have a complex structure. Even if the results are correct, doctors do not know the basis on which the deep learning system reached the diagnosis. This is called the black box problem. It will be necessary to develop and research deep learning systems that can both provide a diagnosis and clarify the reason for the same [64,65].

The diagnostic performance of deep learning models can be significantly reduced if there is a significant difference between the deep learning system-trained and deep learning system-adapted populations. To use deep learning diagnostics properly, it will be necessary to understand in what population (e.g., prevalence, severity, and type of disease) and for what purpose (e.g., screening and scrutiny) the deep learning system was trained.

It is expected that deep learning support for image classification, object detection, segmentation, and image synthesis will improve diagnostic performance, streamline physician work, and reduce

health care costs in the screening, differential diagnosis, efficacy assessment, and surveillance of breast cancer. Physicians will be able to spend more time communicating with patients and doctors in other departments, which will improve the quality of hospital care. Although some competent deep learning systems have been developed, the systems are only applicable for a limited purpose or modality. In the future, it will be necessary to develop a highly versatile deep learning system that combines multiple datasets such as images from several modalities, blood test data, and clinical information such as patients' symptoms to comprehensively diagnose diseases and manage treatment. A high-performance system may be better than clinicians and even replace them. Clinicians with a better understanding of deep learning and willingness to use such systems will be able to play a leading role in this imminent change.

9. Conclusions

Deep learning represents a powerful, efficient, highly accurate tool that reduces physicians' workload. Deep learning of breast ultrasound images is expected to be further developed and clinically applied in various situations such as image classification, object detection, and segmentation. We must have basic knowledge of deep learning, anticipate the problems that will occur when it is introduced, and prepare to address those problems.

References

1. Siegel, R.L.; Miller, K.D.M.; Jemal, A. Cancer statistics, 2018. *CA: A Cancer J. Clin.* **2018**, *68*, 7–30. [CrossRef] [PubMed]
2. Kornecki, A. Current Status of Breast Ultrasound. *Can. Assoc. Radiol. J.* **2011**, *62*, 31–40. [CrossRef] [PubMed]
3. Ohuchi, N.; Suzuki, A.; Sobue, T.; Kawai, M.; Yamamoto, S.; Zheng, Y.-F.; Shiono, Y.N.; Saito, H.; Kuriyama, S.; Tohno, E.; et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): A randomised controlled trial. *Lancet* **2016**, *387*, 341–348. [CrossRef]
4. D'Orsi, C.; Sickles, E.; Mendelson, E.; Morris, E. *ACR BIRADS®Atlas, Breast Imaging Reporting and Data System*; American College of Radiology: Reston, VA, USA, 2013.
5. Mahmud, M.; Kaiser, M.; Hussain, A.; Vassanelli, S. Applications of deep learning and reinforcement learning to biological data. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2063–2079. [CrossRef] [PubMed]
6. Cao, C.; Liu, F.; Tan, H.; Song, D.; Shu, W.; Li, W.; Zhou, Y.; Bo, X.; Xie, Z. Deep Learning and Its Applications in Biomedicine. *Genom. Proteom. Bioinform.* **2018**, *16*, 17–32. [CrossRef] [PubMed]
7. Ravi, D.; Wong, C.; Deligianni, F.; Berthelot, M.; Andreu-Perez, J.; Lo, B.; Yang, G.-Z. Deep Learning for Health Informatics. *IEEE J. Biomed. Heal. Inform.* **2017**, *21*, 4–21. [CrossRef]
8. Zemouri, R.; Zerhouni, N.; Racoceanu, D. Deep learning in the biomedical applications: Recent and suture status. *Appl. Sci.* **2019**, *9*, 1526. [CrossRef]
9. Lakhani, P.; Sundaram, B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* **2017**, *284*, 574–582. [CrossRef]
10. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [CrossRef]
11. Akkus, Z.; Galimzianova, A.; Hoogi, A.; Rubin, D.L.; Erickson, B.J. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *J. Digit. Imaging* **2017**, *30*, 449–459. [CrossRef]
12. Geras, K.J.; Mann, R.M.; Moy, L. Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives. *Radiology* **2019**, *293*, 246–259. [CrossRef] [PubMed]

13. Le, E.; Wang, Y.; Huang, Y.; Hickman, S.; Gilbert, F. Artificial intelligence in breast imaging. *Clin. Radiol.* **2019**, *74*, 357–366. [CrossRef] [PubMed]

14. Wu, G.-G.; Zhou, L.-Q.; Xu, J.W.; Wang, J.-Y.; Wei, Q.; Deng, Y.-B.; Cui, X.-W.; Dietrich, C.F. Artificial intelligence in breast ultrasound. *World J. Radiol.* **2019**, *11*, 19–26. [CrossRef] [PubMed]

15. Sheth, D.; Giger, M.L. Artificial intelligence in the interpretation of breast cancer on MRI. *J. Magn. Reson. Imaging* **2020**, *51*, 1310–1324. [CrossRef] [PubMed]

16. Adachi, M.; Fujioka, T.; Mori, M.; Kubota, K.; Kikuchi, Y.; Wu, X.T.; Oyama, J.; Kimura, K.; Oda, G.; Nakagawa, T.; et al. Detection and Diagnosis of Breast Cancer Using Artificial Intelligence Based Assessment of Maximum Intensity Projection Dynamic Contrast-Enhanced Magnetic Resonance Images. *Diagnostics* **2020**, *10*, 330. [CrossRef]

17. Mori, M.; Fujioka, T.; Katsuta, L.; Kikuchi, Y.; Oda, G.; Nakagawa, T.; Kitazume, Y.; Kubota, K.; Tateishi, U. Feasibility of new fat suppression for breast MRI using pix2pix. *Jpn. J. Radiol* **2020**, *10*, 1075–1081. [CrossRef]

18. Pouliakis, A.; Karakitsou, E.; Margari, N.; Bountris, P.; Haritou, M.; Panayiotides, J.; Koutsouris, D.; Karakitsos, P. Artificial Neural Networks as Decision Support Tools in Cytopathology: Past, Present, and Future. *Biomed. Eng. Comput. Biol.* **2016**, *7*, 1–18. [CrossRef]

19. Janowczyk, A.; Madabhushi, A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J. Pathol. Inform.* **2016**, *7*, 29. [CrossRef]

20. Baltres, A.; Al-Masry, Z.; R., R.M.Z.; Valmary-Degano, S.; Arnould, L.; Zerhouni, N.; Devalland, C. Prediction of Oncotype DX recurrence score using deep multi-layer perceptrons in estrogen receptor-positive, HER2-negative breast cancer. *Breast Cancer* **2020**, *27*, 1007–1016. [CrossRef]

21. Zemouri, R.; Omri, N.; Morello, B.C.; Devalland, C.; Arnould, L.; Zerhouni, N.; Fnaiech, F. Constructive Deep Neural Network for Breast Cancer Diagnosis. *IFAC-PapersOnLine* **2018**, *51*, 98–103. [CrossRef]

22. Chartrand, G.; Cheng, P.M.; Vorontsov, E.; Drozdzal, M.; Turcotte, S.; Pal, C.J.; Kadoury, S.; Tang, A. Deep Learning: A Primer for Radiologists. *Radiographics* **2017**, *37*, 2113–2131. [CrossRef] [PubMed]

23. Pesapane, F.; Codari, M.; Sardanelli, F. Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Eur. Radiol. Exp.* **2018**, *2*, 35. [CrossRef] [PubMed]

24. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]

25. Yasaka, K.; Akai, H.; Kunimatsu, A.; Kiryu, S.; Abe, O. Deep learning with convolutional neural network in radiology. *Jpn. J. Radiol.* **2018**, *36*, 257–272. [CrossRef]

26. Angermueller, C.; Pärnamaa, T.; Parts, L.; Stegle, O. Deep learning for computational biology. *Mol. Syst. Biol.* **2016**, *12*, 878. [CrossRef] [PubMed]

27. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

28. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.

29. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

30. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

31. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; Volume 2017, pp. 2261–2269.

32. Han, S.; Kang, H.-K.; Jeong, J.-Y.; Park, M.-H.; Kim, W.; Bang, W.-C.; Seong, Y.-K. A deep learning framework for supporting the classification of breast lesions in ultrasound images. *Phys. Med. Biol.* **2017**, *62*, 7714–7728. [CrossRef]

33. Fujioka, T.; Kubota, K.; Mori, M.; Kikuchi, Y.; Katsuta, L.; Kasahara, M.; Oda, G.; Ishiba, T.; Nakagawa, T.; Tateishi, U. Distinction between benign and malignant breast masses at breast ultrasound using deep learning method with convolutional neural network. *Jpn. J. Radiol.* **2019**, *37*, 466–472. [CrossRef]

34. Mango, V.L.; Sun, M.; Wynn, R.T.; Ha, R. Should We Ignore, Follow, or Biopsy? Impact of Artificial Intelligence Decision Support on Breast Ultrasound Lesion Assessment. *Am. J. Roentgenol.* **2020**, *214*, 1445–1452. [CrossRef]

35. Zhang, Q.; Xiao, Y.; Dai, W.; Suo, J.; Wang, C.; Shi, J.; Zheng, H. Deep learning based classification of breast tumors with shear-wave elastography. *Ultrasonics* **2016**, *72*, 150–157. [CrossRef] [PubMed]

36. Fujioka, T.; Katsuta, L.; Kubota, K.; Mori, M.; Kikuchi, Y.; Kato, A.; Oda, G.; Nakagawa, T.; Kitazume, Y.; Tateishi, U. Classification of Breast Masses on Ultrasound Shear Wave Elastography using Convolutional Neural Networks. *Ultrason Imaging* **2020**, *42*, 213–220. [CrossRef] [PubMed]

37. Coronado-Gutiérrez, D.; Santamaría, G.; Ganau, S.; Bargalló, X.; Orlando, S.; Oliva-Brañas, M.E.; Perez-Moreno, A.; Burgos-Artizzu, X.P. Quantitative Ultrasound Image Analysis of Axillary Lymph Nodes to Diagnose Metastatic Involvement in Breast Cancer. *Ultrasound Med. Biol.* **2019**, *45*, 2932–2941. [CrossRef] [PubMed]

38. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.

39. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916. [CrossRef] [PubMed]

40. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448.

41. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation Inc.: San Diego, CA, USA, 2015; pp. 91–99.

42. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

43. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2016; pp. 21–37.

44. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. *arXiv* **2016**, arXiv:1612.03144.

45. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.

46. Cao, Z.; Duan, L.; Yang, G.; Yue, T.; Chen, Q. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Med. Imaging* **2019**, *19*, 1–9. [CrossRef] [PubMed]

47. Jiang, Y.; Inciardi, M.F.; Edwards, A.V.; Papaioannou, J. Interpretation Time Using a Concurrent-Read Computer-Aided Detection System for Automated Breast Ultrasound in Breast Cancer Screening of Women with Dense Breast Tissue. *Am. J. Roentgenol.* **2018**, *211*, 452–461. [CrossRef] [PubMed]

48. Yang, S.; Gao, X.; Liu, L.; Shu, R.; Yan, J.; Zhang, G.; Xiao, Y.; Ju, Y.; Zhao, N.; Song, H. Performance and Reading Time of Automated Breast US with or without Computer-aided Detection. *Radiology* **2019**, *292*, 540–549. [CrossRef] [PubMed]

49. Xu, X.; Bao, L.; Tan, Y.; Zhu, L.; Kong, F.; Wang, W. 1000-Case Reader Study of Radiologists' Performance in Interpretation of Automated Breast Volume Scanner Images with a Computer-Aided Detection System. *Ultrasound Med. Biol.* **2018**, *44*, 1694–1702. [CrossRef]

50. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.

51. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]

52. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.

53. Zhuang, Z.; Li, N.; Raj, A.N.J.; Mahesh, V.G.V.; Qiu, S. An RDAU-NET model for lesion segmentation in breast ultrasound images. *PLoS ONE* **2019**, *14*, e0221535. [CrossRef] [PubMed]

54. Hu, Y.; Guo, Y.; Wang, Y.; Yu, J.; Li, J.; Zhou, S.; Chang, C. Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model. *Med. Phys.* **2019**, *46*, 215–228. [CrossRef] [PubMed]

55. Kumar, V.; Webb, J.M.; Gregory, A.; Denis, M.; Meixner, D.D.; Bayat, M.; Whaley, D.H.; Fatemi, M.; Alizad, A. Automated and real-time segmentation of suspicious breast masses using convolutional neural network. *PLoS ONE* **2018**, *13*, e0195816. [CrossRef] [PubMed]

56. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *arXiv* **2014**, arXiv:1406.2661. [CrossRef]

57. Zemouri, R. Semi-Supervised Adversarial Variational Autoencoder. *Mach. Learn. Knowl. Extr.* **2020**, *2*, 361–378. [CrossRef]

58. Yi, X.; Walia, E.; Babyn, P. Generative adversarial network in medical imaging: A review. *Med. Image Anal.* **2019**, *58*, 101552. [CrossRef]

59. Fujioka, T.; Mori, M.; Kubota, K.; Kikuchi, Y.; Katsuta, L.; Adachi, M.; Oda, G.; Nakagawa, T.; Kitazume, Y.; Tateishi, U. Breast Ultrasound Image Synthesis using Deep Convolutional Generative Adversarial Networks. *Diagnostics* **2019**, *9*, 176. [CrossRef]

60. Fujioka, T.; Kubota, K.; Mori, M.; Katsuta, L.; Kikuchi, Y.; Kimura, K.; Kimura, M.; Adachi, M.; Oda, G.; Nakagawa, T.; et al. Virtual Interpolation Images of Tumor Development and Growth on Breast Ultrasound Image Synthesis with Deep Convolutional Generative Adversarial Networks. *J. Ultrasound Med.* **2020**. [CrossRef]

61. Fujioka, T.; Kubota, K.; Mori, M.; Kikuchi, Y.; Katsuta, L.; Kimura, M.; Yamaga, E.; Adachi, M.; Oda, G.; Nakagawa, T.; et al. Efficient Anomaly Detection with Generative Adversarial Network for Breast Ultrasound Imaging. *Diagnostics* **2020**, *10*, 456. [CrossRef]

62. Han, L.; Huang, Y.; Dou, H.; Wang, S.; Ahamad, S.; Luo, H.; Liu, Q.; Fan, J.; Zhang, J. Semi-supervised segmentation of lesion from breast ultrasound images with attentional generative adversarial network. *Comput. Methods Programs Biomed.* **2020**, *189*, 105275. [CrossRef]

63. Benjamens, S.; Dhunnoo, P.; Meskó, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. *NPJ Digit. Med.* **2020**, *3*. [CrossRef] [PubMed]

64. Fujioka, T.; Yashima, Y.; Oyama, J.; Mori, M.; Kubota, K.; Katsuta, L.; Kimura, K.; Yamaga, E.; Oda, G.; Nakagawa, T.; et al. Deep-learning approach with convolutional neural network for classification of maximum intensity projections of dynamic contrast-enhanced breast magnetic resonance imaging. *Magn. Reson. Imaging* **2021**, *75*. [CrossRef] [PubMed]

65. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.