Semi-supervised GAN-based Radiomics Model for Data Augmentation in Breast Ultrasound Mass Classification

Ting Pang a,b, Jeannie Hsiu Ding Wong c,e, Wei Lin Ng c, Chee Seng Chan a

a Center of Image and Signal Processing, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, 50603, Malaysia
b Center of Network and Information, Xinxing Medical University, Xinxian, 453000, PR China
c Department of Biomedical Imaging, Faculty of Medicine, University of Malaya, Kuala Lumpur, 50603, Malaysia

d A R T I C L E   I N F O

Article history:
Received 3 August 2020
Accepted 16 February 2021

Keywords:
Deep learning radiomics
Semi-supervised learning
Generative adversarial network
Data augmentation
Breast cancer classification
Ultrasound imaging

A B S T R A C T

Background and Objective: The capability of deep learning radiomics (DLR) to extract high-level medical imaging features has promoted the use of computer-aided diagnosis of breast mass detected on ultrasound. Recently, generative adversarial network (GAN) has aided in tackling a general issue in DLR, i.e., obtaining a sufficient number of medical images. However, GAN methods require a pair of input and labeled images, which require an exhaustive human annotation process that is very time-consuming. The aim of this paper is to develop a radiomics model based on a semi-supervised GAN method to perform data augmentation in breast ultrasound images.

Methods: A total of 1447 ultrasound images, including 767 benign masses and 680 malignant masses were acquired from a tertiary hospital. A semi-supervised GAN model was developed to augment the breast ultrasound images. The synthesized images were subsequently used to classify breast masses using a convolutional neural network (CNN). The model was validated using a 5-fold cross-validation method.

Results: The proposed GAN architecture generated high-quality breast ultrasound images, verified by two experienced radiologists. The improved performance of semi-supervised learning increased the quality of the synthetic data produced in comparison to the baseline method. We achieved more accurate breast mass classification results (accuracy 90.41%, sensitivity 87.94%, specificity 85.86%) with our synthetic data augmentation compared to other state-of-the-art methods.

Conclusion: The proposed radiomics model has demonstrated a promising potential to synthesize and classify breast masses on ultrasound in a semi-supervised manner.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

According to the World Health Organization (WHO), breast cancer is one of the most frequent cancer among women, with 15% of deaths yearly [1]. It is essential to have an early screening as it can provide timely treatment and reduce the death rate. In recent years, ultrasound imaging (Fig 1), which does not involve ionizing radiation, has become one of the most important screening tools for breast cancer [2]. However, manual examination of medical images is a tedious and exhausting process, exacerbated by the shortage of qualified radiologists, leading to a drop in diagnostic efficiencies [3]. Thus, an automatic diagnosis of ultrasound examinations using computer-aided diagnosis (CAD) is essential.

Recently, deep learning [4] has become a de facto solution for many research domains. For instance, in medical imaging, radiomics based on convolutional neural network (CNN) is an emerging method in CAD of breast cancer [5-6]. Generally, deep learning radiomics (DLR) requires a lot of training samples to produce an effective model, and this is a huge challenge in the medical domain [7]. On one hand, collection of a handful of medical images is difficult due to patient confidentiality issues, and thus most of the medical data are imbalanced (i.e., usually, images of malignant masses are much lesser than those of the benign or normal breast parenchyma). On the other hand, the annotation process of medical images by the radiologists is also very time-consuming [8]. To solve this problem, traditional affine transformation, such as flipping, rotation, brightness changing, cropping etc., have been used for augmenting the dataset. However, such augmentation methods with small modifications (e.g., move the image a few pixels to the left or rotate it 45 degrees) does not provide the deep model to learn additional information from the medical images [9].

https://doi.org/10.1016/j.cmpb.2021.106018
0169-2607 © 2021 Elsevier B.V. All rights reserved.
Generative adversarial network (GAN), inspired by the game theory, is a popular sophisticated model for dataset augmentation to learn the rich features and model high-dimensional distributions of data [10-11]. Basically, it includes a pair of networks in competition, i.e., a generator G explores to generate fake images and a discriminator D is regarded as learned prior for discrimination between the real and fake images repeatedly until the GAN model produces high-quality synthetic images. Therefore, GAN emerges as an effective method to address the challenges of limited medical images for DLR models [12-13]. However, most approach/implementation utilizes supervised GAN that requires a large number of labeled images, which is time-consuming and manually exhausting.

In another aspect, GAN also has been effectively applied to breast imaging augmentation in mass classification. Wu et al. showed that when trained on the same CNN classifier, the conditional GAN-augmented artificial mammography images produced higher AUC than the data augmentation using the affine transformation [14]. Guan and Loew used GAN to generate samples for data augmentation (to reduce overfitting), combined with the original data fed into a CNN for breast mass classification [15]. Muramatsu et al. performed experiments on breast masses detected on mammogram using shared data with domain transformation and GAN [16]. However, most of these studies focused on mammographic imaging, and to the best of our knowledge, there is no research utilizing GAN for augmenting datasets in breast ultrasound images for breast mass classification. Besides, these aforementioned methods employed supervised GAN that require complicated labeling processes.

To alleviate the above constraints, in this paper, we explored a GAN-based radiomics model in a semi-supervised manner to augment imaging datasets for breast mass classification on ultrasound.

First, we exploited a semi-supervised GAN architecture, particularly the TripleGAN (TGAN) [17] to synthesize breast masses, which was then verified visually via a professional assessment. Secondly, we demonstrated the effectiveness of semi-supervised learning by comparing the classification accuracy using the generated images trained on datasets with distinct amounts of unlabeled data between our GAN architecture and baseline method. Finally, we demonstrated that the proposed method can achieve a better classification performance for breast mass than other state-of-the-art methods.

The contributions of this work are mainly in the three following aspects: 1) We showed that it is possible to generate high-quality breast ultrasound masses using the GAN method for data augmentation in deep learning classification, and this is verified by professional assessments. 2) Our TGAN-based data augmentation architecture improves the quality of the synthetic data produced. 3) The proposed radiomics model has improved the classification results for a better breast cancer diagnosis.

2. Related Works

2.1. Deep Learning Radiomics for Breast Cancer Classification

Handcrafted radiomics have been used for breast cancer classification [18-19]. Although it is useful for clinical diagnosis, these conventional methods are very costly as they need to be carefully designed to extract useful features and then segment the regions of interest (ROI) from the medical images. Hence, recently researchers have deployed deep learning in radiomics to enhance the screening benefits [6-7,20-21]. Due to the huge storage of digital breast images and the more aggressive forms of breast cancer, many researchers in this area are devoted to use deep learning to alleviate the burden of designing the specific features for breast cancer diagnosis.

The first implementation of DLR in breast cancer was done by Arvalo et al. in 2015 [22]. They applied CNN architecture to discover features for breast mass classification. CNN architecture showed an increased area under the curve (AUC) of receiver operating characteristics (ROCs) with the handcrafted radiomics method. With the widespread use of deep learning, more scholars designed different DLR architectures for breast cancer diagnosis. Kooi et al. showed that the location information and context features could improve the CNN network [23]. Zhou et al. first integrated CNN and shear-wave elastography (SWE) to classify breast ultrasound tumors [24]. Shin et al. proposed a hybrid semi-supervised deep learning network based on Fast R-CNN (used a strongly annotated dataset) and multiple-instance learning (used a weekly annotated dataset) for breast mass localization and classification [25]. Wang et al. proposed to use CNN to reduce false-positives and avoid complicated feature extraction after a local phase-based candidates’ selection in breast tumors detection [26]. Nevertheless, these studies focus on receiving better diagnostic results, not considering the problem of limited breast datasets.

2.2. Generative Adversarial Network Applied in Medical Imaging

Due to data protection issues and complex annotated labels by the radiologists, collecting a large number of medical images for the purpose of training an effective deep learning model is very challenging. In medical images, a lot of GAN variants (DCGAN [27], InfoGAN [28], LAPGAN [29], CycleGAN [30], et al.) have been developed for reconstruction [31], synthesis [32], segmentation [33] and detection [34]. Alternatively, in the current medical domain, GAN has been applied into medical image classification task and achieved better performance. For instance, Adar et al. combined DCGAN and ACGAN to generate samples which is beneficial for the live mass classification task [9]. Salehinejad et al. used five different GANs to generate and classify five different classes on chest X-rays [35]. Although GAN has made great achievements in aiding medical image augmentation for classification, the drawback is obvious. That is, most of them focus on applying supervised GAN, which heavily relies on the labeling of a large number of images for training in medical imaging classification.

3. Materials and Methods

In this section, we first depict the dataset and its preprocessing. Then we discuss the semi-supervised GAN-based method for generating images and build the DLR architecture for the task of breast mass classification based on CNN.

3.1. Materials and Preprocessing

The dataset of breast ultrasound images obtained in this research was from the University of Malaya Medical Centre (UMMC).
Table 1
The labeled and unlabeled training and testing data.

<table>
<thead>
<tr>
<th></th>
<th>labeled</th>
<th></th>
<th>unlabeled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>benign</td>
<td>malignant</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>492</td>
<td>435</td>
<td>231</td>
</tr>
<tr>
<td>Testing</td>
<td>122</td>
<td>109</td>
<td>58</td>
</tr>
</tbody>
</table>

The study was approved by the institutional review board (MREC ID: 2019822-7771). We adopted 1447 images (each image had one tumor and the original image format was DICOM) scanned from 357 female patients (mean age = 51.6 y) between January 2018 and January 2019. These included 767 benign and 680 malignant masses. All data were checked by experienced radiologists according to the Breast Imaging Reporting and Data System (BI-RADS) [36], and the malignancies were proven via histopathological biopsies. Among all, 1158 (80%) were used as training data, and 289 (20%) were used as testing data. The distribution of labeled data and unlabeled data were 80% and 20%, respectively. Table 1 presents the number of labeled images and unlabeled images checked by the radiologists.

For the pre-processing step, feature-wise processing (FWP) [37] was applied to reduce the deep learning model processing time on raw images in advanced. First, in order to have a balance between image resolution and computational complexity of the deep learning model [35], the images were cropped to 128x128 pixels as the ROIs (each image has one ROI). Fig 2 shows some samples of benign and malignant breast mass ROIs in the ultrasound images. Then, to initially resolve the prior problem of limited images, the conventional data augmentation method (where each ROI was flipped horizontally and vertically) was adopted for training data, seen in Fig 3. Hence, the final number of breast dataset was 4341.

3.2. Methods

3.2.1. Semi-supervised GAN for breast masses synthesis

This method was inspired by the TripleGAN (TGAN) proposed by Li et al. [17]. Conventionally, the existing GAN for semi-supervised learning has a single discriminator, which shares two incompatible learning roles; that (i) identify the fake samples and (ii) predict labels, leading to the G and D having different optimal [38]. Besides, G ignores the labeled features resulting in its inability to control the semantics of the generative samples. These problems were addressed in the TGAN architecture, where it has three players: the generator and the classifier characterize the conditional distributions between images and labels, respectively; and the discriminator focuses on identifying fake image-label pairs. First, generator G synthesizes fake labelled data under the condition that it not only focuses on the fake images, but is given conditional real labels and noise. Then the classifier C achieves real unlabeled data (Here, just to be consistent with G, unlabeled means artificial labels) under the condition that it is given real images and fake labels. Finally, the real labeled data and the synthetic data are fed into the discriminator D to learn the features of the image-label pair, to discriminate the images and classify the synthetic images, as shown in Fig 4.

The adversarial losses for the three parts are defined in the formulas below:

\[
L_G = \min \left\{(1 - \alpha)E_{(x, y)} \sim p(x, y)[\log(1 - D(G(y, x), y))] \right\} \quad (1)
\]

\[
L_C = \min \left\{\alpha E_{(x, y)} \sim p(x, y)[\log(1 - D(x, y))] \right\} + \mathcal{R}_C \quad (2)
\]

\[
L_D = \max \left\{E_{(x, y)} \sim p(x, y)[\log D(x, y)] + \alpha E_{(x, y)} \sim p(x, y)[\log(1 - D(x, y))] \right\} \quad (3)
\]

Where \(\alpha \in (0, 1)\) is a constant which commands the relative importance between G and C. x is the input breast tumor image, and y is the label (benign or malignant). The input-label pair dataset \((x, y)\) convergence in a joint distribution \(p(x, y)\). For the generator G, the image x is generated by a noise \(z \sim \mathcal{N}(0, 1)\) from the latent space labeled as y, which is \(y = G(z, x)\). For classifier C, a pseudo label \(y\) is assigned to image x by the model. For the discriminator D, it compares the input-label pair data obtained from G and C with the ground truth \((x, y)\) for image and label judgment, respectively. The global equilibrium will be achieved if \(p(x, y) = p_r(x, y) = p_g(x, y)\) (See the theoretical analysis in [17]). To reach the global optimum, it introduces equivalent Kullback–Leibler divergence (relative entropy loss) \([39-40]\) \(\mathcal{R}_C\) to C.

3.2.2. Radiomics architecture for breast cancer classification

Our goal is to learn rich radiomics features from data distributions seen as the most important representatives of the input images through our model. Here, we built the semi-supervised GAN-based radiomics architecture for classification, as shown in Fig 5. First, we trained the TGAN architecture to obtain the synthetic images using the dataset combining with conventional data augmentation. Then we completed the breast cancer classification through the CNN model using the dataset combining with synthetic data generated by TGAN and conventional data augmentation. Here we employed Inception-V3 [41] as a CNN classification network. To easily perform the DLR architecture, all the three players in TGAN are typically implemented by multilayer neural networks.

4. Experiments and Discussion

In the following, we first describe the experimental implementation details. Then we present and discuss three experimented results: (1) high-quality synthetic ultrasound images; (2) better
Fig. 4: TGAN-based model. "Unlabeled" denotes artificial label.

Fig. 5: Overview of semi-supervised GAN-based radiomics architecture in breast mass classification. The dataset, combined with conventional data augmentation were taken as the input of the TGAN. Classifier C achieves artificial labels were given unlabeled real images, and Generator G synthesizes labeled images were given real labels and noise. Real labeled data and synthetic data were fed into the discriminator D for discrimination of the images and classification of synthetic images. Then the dataset combining with synthetic data and conventional data augmentation were taken as the input of CNN. The breast cancer classification was completed through the Inception-V3 network.

We used 5-fold cross-validation with case separation at the patient level. First, we selected the proper CNN by using Alexnet [42], VGG [43], GoogleNet (Inception-V3) and Resnet [44], respectively. After training on the initial dataset (without augmentation), we employed Inception-V3 as the classification network (it acquired the highest final testing accuracy). Then for TGAN, the batch size was 20, epoch was 1000, learning rate was \(2 \times 10^{-4}\), optimizer used was Adam (Adaptive Moment Estimation), and \(\alpha\) was 0.05. For CNN with data augmentation, the batch size was 20, epoch was 200, learning rate was \(1 \times 10^{-4}\) and optimizer was Adam. The detailed implementation networks for training on breast ultrasound tumor datasets are listed in Table 2. We utilized Tensorflow framework on a single GPU NVIDIA RTX 2070. We validated the classification per-
formance of the model on the testing dataset using the three metrics: accuracy (AC), sensitivity (SE), and specificity (SP) [21]. They are defined as follows:

\[
AC = \frac{TP + TN}{TP + FP + TN + FN}
\]

\[
SE = \frac{TP}{TP + FN}
\]

\[
SP = \frac{TN}{FP + TN}
\]

where TP denotes true positive (the number of correct malignancy), FP denotes false positive (the number of incorrect malignancy), FN denotes false negative (the number of incorrect benign), TN denotes true negative (the number of correct benign).

### 4.2. Results and Discussion

#### 4.2.1. Evaluation of synthetic data

The examples of generated breast masses during the TGAN training process are exhibited in Fig 6, including benign and malignant cases. In addition, we randomly selected 40 generated images (20 benign, 20 malignant), and they were evaluated by two experienced radiologists. Each expert examined whether the images were benign or malignant, and the evaluated results are summarized in Table 3. Then, we specially analyzed the causes of radiologists’ misjudgment in two generated samples (one benign mass and one malignant mass), shown in Fig 7.

From Fig 6, we can see that the quality of generated ultrasound images became better as the epoch increases. Though the final images have comparatively low resolution, they are visually similar to real ultrasound breast mass. The generated mass captured most of the characteristics and elements, such as the shape is oval and the margin is circumscribed for benign breast mass, while the malignant breast mass is irregular and ill-defined. Even annotations in the generated images can be seen, and this demonstrates the efficiency of TGAN in generating images. Besides, as shown in Table 3, we note that the experts respectively gave the correct ratio of judgement for classification (i.e., benign or malignant) as 82.5% and 80.0% (i.e., ([11+19]/40) and (15+17)/40). Therefore, the visual samples and expert evaluations proved that the generated breast masses have similar features to the real images and can be fed to the network to classify new breast masses. Moreover, it illustrates that the TGAN model learned most of the characteristics of breast cancer images. Fig 7(a) shows that the benign mass was wrongly determined as malignant owing to the spiculated margin with radiating lines. Fig 7(b) shows that the malignant mass was wrongly determined as benign owing to the well-circumscribed margin. Hence, though margins and shape of mass were both representative characteristics for diagnosis, the margin feature was more complicated and challenging to learn for synthetic data augmentation. Similarly, we infer that more attention should be paid to the margin when using deep learning radiomics for diagnosing breast masses.

#### 4.2.2. Effectiveness of semi-supervised learning

To assess the effectiveness of semi-supervised learning, we compared the classification accuracy for generated images respectively using our GAN architecture (TGAN) and the baseline InfoGAN [28] (most commonly used semi-supervised GAN method) trained on datasets with 289 (20%), 578 (40%), 1156 (80%) unlabeled breast

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Detailed implementation networks. MLP: Multilayer Perceptron, ReLU: Rectified Linear Unit, BN: Batch Normalization, De-conv: De-convolution, Conv: Convolution, NIN: Network-in-Network.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGAN</td>
<td>C</td>
</tr>
<tr>
<td>Input: Class 2 y, Noise: z ~ (−1, 1) 100 MLP 32768 units RelU, BN, Reshape 2048 × 4 × 4, 5 × 5 de-conv1024 RelU stride 2, BN</td>
<td>Input: 128×128 Ultrasound Images Gaussian Noise, 3 × 3 conv 512 RelU, 3 × 3 conv 512 RelU, 3 × 3 conv 512 RelU, 2 × 2 max-pooling stride 2, 0.5 dropout</td>
</tr>
<tr>
<td>5 × 5 de-conv 512 stride 2, BN</td>
<td>3 × 3 conv1024 RelU, 3 × 3 conv1024 RelU, 3 × 3 conv1024 RelU, 2 × 2 max-pooling stride 2, 0.5 dropout</td>
</tr>
<tr>
<td>5 × 5 deconv 3 stride 2, tanh</td>
<td>3 × 3 conv2048 RelU, NIN, 1024 RelU, NIN, 512 RelU, Global pooling, 2-class softmax</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>The evaluated results of generated breast masses. T: true, F: false.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert1</td>
<td>14 1 19 6</td>
</tr>
<tr>
<td>Expert2</td>
<td>15 3 17 5</td>
</tr>
</tbody>
</table>
Furthermore, further reduction was observed in the SNE visualization of the images generated by our method using t-SNE (t-distributed stochastic neighbor embedding) algorithm. Fig 8 shows that the benign and malignant masses have some overlaps, due to different shapes and confusing features (e.g. margin). The breast masses generated by the interpretable TGAN (Fig 8(b)) exhibit better separating localization for benign and malignant breast masses than the masses generated by the InfoGAN (Fig 8(a)). Further, the malignant masses of TGAN have grouped distribution while the malignant masses of InfoGAN have scattered distribution. This indicates that our method is able to increase the quality of the synthetic data produced with better performance of semi-supervised learning.

It can be seen from Table 4 that although all the classification accuracy reduces with the increasing number of unlabeled images, our TGAN-based radiomics model was 5% higher than the baseline method. Statistically significant differences (P < .001 for all comparisons) were observed for the AC of the two different approaches. It can be inferred that the unreal-labeled classifier of the TGAN model can better extract the radiomics features of unlabeled data in a semi-supervised manner. Fig 8 shows that the benign and malignant masses have some overlaps, due to different shapes and confusing features (e.g. margin). The breast masses generated by the interpretable TGAN (Fig 8(b)) exhibit better separating localization for benign and malignant breast masses than the masses generated by the InfoGAN (Fig 8(a)). Further, the malignant masses of TGAN have grouped distribution while the malignant masses of InfoGAN have scattered distribution. This indicates that our method is able to increase the quality of the synthetic data produced with better performance of semi-supervised learning.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>The classification accuracy for generated breast masses with different number of labeled inputs. Data in brackets represents the 95% CI.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>289 (20%)</td>
</tr>
<tr>
<td>infoGAN</td>
<td>77.28% [0.76, 0.81]</td>
</tr>
<tr>
<td>TGAN</td>
<td><strong>82.39% [0.78, 0.83]</strong></td>
</tr>
</tbody>
</table>
model in data augmentation for classification of breast masses in ultrasound. From the experiments, we have demonstrated that the GAN method can generate high-quality breast ultrasound mass images which can augment the training dataset. Additionally, the model was able to better extract the radiomics features of unla- beled data leading to an improvement in the quality of the synthetic data produced. Finally, our synthetic data augmentation can improve breast cancer classification performance compared with other state-of-the-art approaches. The limitation of the study is that this is a single institutional study. The application of the DLR model developed is yet to be tested on medical images from other centres or databases. We believe that this method can be easily applied to solve the problem of limited medical images for other diseases. However, these applications are beyond the scope of the current study.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

Acknowledgments

The study was supported by the Malaysian Ministry of Higher Education (MOHE) Fundamental Research Grant Scheme (FRGS) (FP017-2019A). [FRGS/1/2019/UMK03/UM/01/1]). We express our gratitude to the University of Malaya Medical Centre (UMMC) Medical Ethics Committee (No. 2019822-7771).

References


Table 5

<table>
<thead>
<tr>
<th>Input</th>
<th>AC</th>
<th>SE</th>
<th>SP</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN-AUG-GAN [10]</td>
<td>86.95%</td>
<td>84.36%</td>
<td>82.35%</td>
<td>21.8</td>
</tr>
<tr>
<td>CNN-AUG-DCGAN [27]</td>
<td>87.38%</td>
<td>84.95%</td>
<td>84.03%</td>
<td>22.6</td>
</tr>
<tr>
<td>CNN-AUG-InfoGAN [28]</td>
<td>88.72%</td>
<td>86.60%</td>
<td>85.18%</td>
<td>20.9</td>
</tr>
<tr>
<td>CNN-AUG-TGAN</td>
<td>89.41%</td>
<td>87.94%</td>
<td>85.86%</td>
<td>20.3</td>
</tr>
</tbody>
</table>

4.2.3. Assessment of classification performance

In this section, we explored the effectiveness of the classification accuracy of our method. First, we trained the same CNN (i.e., Inception-V3) on different inputs: 1) Raw data; 2) Raw data + AUG with only conventional data augmentation; 3) Raw data + AUG + Synthetic data with synthetic data augmentation (final 4326 available images selected from all generated data, 1835 benign and 2491 malignant). Fig 9 reports the training loss and training accuracy in each epoch (Here, the synthetic images are generated by TGAN). Second, we compared our classification results (CNN-AUG-TGAN) with AC, SE, SP, Training Time (T, one epoch) to three state-of-the-art GAN methods: CNN-AUG-GAN (traditional GAN), CNN-AUG-DCGAN (supervised learning), CNN-AUG-InfoGAN (semi-supervised learning). Table 5 shows the compared performance. Besides, Fig 10 shows one sample of predicting breast masses classification results using our method.

From Fig 9, we can see that the plots of all training loss and training accuracy converge to a point as epoch increases during training. Our method is robust, significantly improving performance without an early local optimal. Moreover, it can be seen from Table 5 that our method obtains the highest accuracy (90.41%), sensitivity (87.94%), specificity (85.86%) and shortest training time (20.3s, i.e. one epoch) on testing data. As shown in Fig 10, the final classification result can be accurately acquired using the model trained by our proposed network. Hence, from the final results demonstrating high accuracy in predicting the classification of breast mass, we are able to conclude that our method of using combinations of synthetic data augmentation is better than state-of-the-art methods in distinguishing breast masses as benign or malignant masses.

5. Conclusions

Semi-supervised GAN method can alleviate the labeling pressure of medical image augmentation for DLR in breast cancer. In this paper, we proposed a semi-supervised GAN-based radiomics


