Research Article

A Selective Ensemble Classification Method Combining Mammography Images with Ultrasound Images for Breast Cancer Diagnosis

Jinyu Cong,1 Benzheng Wei,2 Yunlong He,1 Yilong Yin,3 and Yuanjie Zheng1

1School of Information Science and Engineering, Key Lab of Intelligent Computing & Information Security in Universities of Shandong, Institute of Life Sciences, Shandong Provincial Key Laboratory for Distributed Computer Software Novel Technology, and Key Lab of Intelligent Information Processing, Shandong Normal University, Jinan 250358, China
2College of Science and Technology, Shandong University of Traditional Chinese Medicine, Jinan 250004, China
3School of Computer Science and Technology, Shandong University, Jinan 250100, China

Correspondence should be addressed to Benzheng Wei; wbz99@sina.com and Yuanjie Zheng; zhengyuanjie@gmail.com

Received 21 March 2017; Accepted 28 May 2017; Published 27 June 2017

Academic Editor: Po-Hsiang Tsui

Copyright © 2017 Jinyu Cong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Breast cancer has been one of the main diseases that threatens women's life. Early detection and diagnosis of breast cancer play an important role in reducing mortality rates of breast cancer, which can improve the cure rate, relieve patients' sufferings, and guarantee the patients' life quality [3, 4].

1. Introduction

Breast cancer is one of the most common malignant tumors in female [1]; it has been the first leading of death among tumors in patients under the age of 45 years old in both Asian and European countries [2]. Early detection and diagnosis of breast cancer play an important role in reducing mortality rates of breast cancer, which can improve the cure rate, relieve patients' sufferings, and guarantee the patients' life quality [3, 4].

Mammography and ultrasound are two most commonly used screening modalities for early detection and treatment of breast cancer. Mammography uses X-ray to noninvasively image the internal structure of human breast, which can be used to show the masses, calcifications, and any other suspicious area that could be identified as signs of breast cancer [5]. However, screening with mammography alone is limited in its ability to detect tumors in dense breasts which is typically linked to a higher risk of cancer [6]. As a supplement modality to mammography, ultrasound has been proven to depict the shape, border, and internal echo features of small, node-negative breast tumors clearly [7]. Meanwhile, the low cost [8] and better performance in dense breast [9, 10] improve the application of ultrasound images in breast cancer diagnosis. In particular for the Asian women, who are with higher density breast tissue, the breast cancer diagnosis based on the mammography has certain limitations. It has been proven that using ultrasound in conjunction with mammography resulted in significantly increased cancer detection rates [11–14]. Moreover, the radiologist's clinical experience and subjective judgment directly affect the accuracy of the diagnosis; the radiologist who lacks clinical experience may make an inaccurate diagnosis or miss a diagnosis. At the same time, the inherent high frequency noise and the shadow of medical images may also affect the accuracy of diagnosis. Therefore, a computer-aided diagnosis (CAD) using ultrasound with mammography which can improve the accuracy and specificity is needed in clinical application [15]. The selective ensemble learning [16] is mentioned for the fusion...
of the ultrasound and mammography in an efficient way. Compared with traditional ensemble learning, the selective ensemble learning can generate the base classifiers with stronger generalization ability. And it is also have the smaller size and faster speed.

In this paper, we propose a selective ensemble classification method using mammography with ultrasound for breast cancer diagnosis. First, instead of the BI-RAD feature, we extract the texture and morphological features to obtain more information of the lesion's edge, shape, and texture on ultrasound and mammography. Then three classifiers, KNN, SVM, and Navies Bayes, which have been used to diagnose breast cancer by many researchers are employed for the ultrasound and mammography, respectively. To ensure the accuracy and the generalization of the base classifier, the new indicator \( R \) is proposed to choose the appropriate base classifier. Finally, the results of breast cancer diagnosis is obtained by the integrated classifier by majority vote. The extensive experimental results show that the ensemble classification method is efficient for breast cancer diagnosis in our database. The indicator \( R \) can choose the base classifiers with high accuracy and generalization.

The rest of the paper is organized as follows. In Section 2, we explain our method. The indicators are mentioned in Section 3. Experimental results of diagnosis are explained in Section 4, and, finally in Section 5, we conclude our work and give the prospect of the future work.

2. Method

In this paper, we propose a selective ensemble classification method combining mammography with ultrasound images to diagnose breast cancer. Extracting the feature on ultrasound images and mammography images, respectively, we sampled these images with 10-fold cross-validation to train the base classifiers (SVM, KNN and Naive Bayes for mammography and ultrasound features). Considering that integrating many of the classifiers may be better than integrating all of them, we obtain different integrated classifiers generated by three of all base classifiers and rank them by the indicator \( R \). We choose the best integrated classifier and obtain the final result by majority vote to diagnose the breast cancer. The flow chart of the selective ensemble method has been shown in Figure 1.

2.1. Feature Extraction. In this paper, we extract the minimum bounding rectangle around the lesion on ultrasound images and mammography images to calculate the gray-level cooccurrence matrix (GLCM) which can provide a method for generating texture features. GLCM are computed in four directions (0°, 45°, 90°, and 135°) with 8-pixel distance. The feature used for classification are: correlation, contrast, energy and entropy. Morphological features are beneficial to distinguishing between benign and malignant tumors on ultrasound images, so we extract three morphological features: depth-to-width (\( D : W \)) ratio, elliptic-normalized circumference ratio of the lesion and its equivalent ellipse, and it represents the irregular boundary of the lesions. The benign tumors will stop growing, while the malignant will not stop growing. The larger size is always with the malignant tumor.

2.2. The Selective Ensemble Classification Method Combining Mammography Images with Ultrasound Images. As is well known, the integrated classifier may be better than the single classifier. So we propose a selective ensemble classification method combining mammography images with ultrasound images to diagnose breast cancer. We choose the KNN, SVM, and Naive Bayes as the base classifier because they are implemented simply and diagnose breast cancer effectively. The key in ensemble learning is the accuracy and diversity of base classifiers. We use the indicator \( R \) to rank the single classifier which is based on the accuracy and its diversity. \( R \) can be described as follows:

\[
R_i = \mu_1 * ACC_i - \mu_2 * DF_i. \tag{1}
\]

\( ACC_i \) is the accuracy of the single classifier; double fault (DF) can reflect the diversity of the classifier \([18]\). As is show in Table 1, \( N^{\text{mou}} \) means the number of different classifier's results with \( m \) and \( n \) (\( m \) and \( n \) are 0 or 1). The smaller \( R \) is, the better the diversity the classifier will be. The different \( \mu \) can impact the choice of the classifier (\( \mu_1 = 0.4, \mu_2 = 0.6 \)):

\[
DF_{i,k} = \frac{N_{00}^0 + N_{10}^0 + N_{01}^0 + N_{00}^{11}}{N_{00}^{11} + N_{10}^0 + N_{01}^0 + N_{00}^{11}} \tag{2}
\]

\[
DF_i = \frac{\sum_{k=1}^{L} DF_{i,k}}{L-1}.
\]

3. Indicators

We use six indicators to evaluate the performance of the system, accuracy, sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), and the area under receiver operating characteristic curve (AUC).

The accuracy is described as the ratio of the correct samples distinguished by the classifier to the total samples. The accuracy can be described as follows:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} * 100\%, \tag{3}
\]

where the mean of TP, TN, FP, and FN can be seen in the Table 2.

The sensitivity is described as the ratio of the malignancy distinguished by the integrated classifier (TP) to the real malignancy (TP + FN). It can measure that how many times
malignancy can be diagnosed. The specificity is described as the ratio of the benign tumor distinguished by the integrated classifier (TN) to the real benign tumor (TN + FP). It measures that how many times benign tumors was misdiagnosed as malignancy. The NPV can measure the diagnosis accuracy of benign tumor. The PPV can measure the diagnosis accuracy of malignancy tumor. These four indicators can be described as follows:

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\%,
\]
\[
\text{Specificity} = \frac{TN}{TN + FP} \times 100\%,
\]
\[
\text{PPV} = \frac{TP}{TP + FP} \times 100\%,
\]
\[
\text{NPV} = \frac{TN}{TN + FN} \times 100\%.
\]

In addition to the above, we also use the AUC which means the area of the ROC (receiver operating characteristic) curve to appraise the performance.

4. Experiments

To evaluate the effectiveness of our method, we do some experiments using the indicators mentioned above. The database used in this paper contains 142 medical images, including 71 ultrasound images and 71 mammography images. These images are acquired by 71 patients, 40 benign tumors and 31 malignant tumors. The programs are implemented with Matlab R2010b and WEKA, and data was entered into the computer on Intel 2.93 G dual-core processor with 3.29 G RAM.

4.1. The Performance of Single Classifier Compared with the Integrated Classifiers in Ultrasound Images and Mammography Images, Respectively. First, we compared the performance of single classifier with the integrated classifiers based on ultrasound images or mammography images, respectively. As is shown in Table 3, the performance of integrated classifiers is better than the Naive Bayes and KNN on ultrasound images. Compared with SVM, it is better at accuracy, specificity, and PPV, while it is little worse at sensitivity, NPV, and AUC. The integrated classifier and the SVM tend to have different strengths on ultrasound images.

In Table 4, the single classifier has poor performance on all indicators compared with integrated classifiers on mammography images. (Classifier-M means the results of the mammography images using this classifier. And Classifier-U means the results of the ultrasound images using this classifier. For example, SVM-M means the results of the mammography images using SVM.) The ROC is shown in the Figures 2 and 3. Obviously, the integrated classifier is better than the single classifier.

4.2. The Effectiveness and Necessity of Combining the Ultrasound Images with Mammography Images. We propose the ensemble method based on multimodal images (mammography and ultrasound images) according to its complementary feature. In Table 5 and Figure 4, we compared the performance on ultrasound images and mammography images with multimodal images to prove the effectiveness and necessity. The performance on multimodal images is superior to that of mammography images. It is worse than the performance in ultrasound images only in specificity and PPV, but it is superior in accuracy, sensitivity, NPV, and AUC. Typically, the ensemble method can obtain the sensitivity of 96.77%. It is much higher than the integrated classifier on
Table 3: The performance of single classifier compared with the integrated classifiers on ultrasound images.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes-U</td>
<td>71.83%</td>
<td>54.84%</td>
<td>85%</td>
<td>70.83%</td>
<td>73.91%</td>
<td>0.6831</td>
</tr>
<tr>
<td>SVM-U</td>
<td>84.51%</td>
<td>87.10%</td>
<td>82.50%</td>
<td>89.19%</td>
<td>79.41%</td>
<td>0.8427</td>
</tr>
<tr>
<td>KNN-U</td>
<td>73.24%</td>
<td>77.42%</td>
<td>70%</td>
<td>80%</td>
<td>66.67%</td>
<td>0.7250</td>
</tr>
<tr>
<td>The integrated classifier (ultrasounds)</td>
<td><strong>85.92%</strong></td>
<td>83.87%</td>
<td><strong>87.50%</strong></td>
<td>87.50%</td>
<td><strong>83.87%</strong></td>
<td>0.8363</td>
</tr>
</tbody>
</table>

Table 4: The performance of single classifier compared with the integrated classifiers on mammography images.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes-M</td>
<td>78.87%</td>
<td>77.42%</td>
<td>80%</td>
<td>82.05%</td>
<td>75.00%</td>
<td>0.7653</td>
</tr>
<tr>
<td>SVM-M</td>
<td>54.93%</td>
<td>45.16%</td>
<td>62.5%</td>
<td>59.52%</td>
<td>48.28%</td>
<td>0.5202</td>
</tr>
<tr>
<td>KNN-M</td>
<td>67.61%</td>
<td>58.06%</td>
<td>75%</td>
<td>69.77%</td>
<td>64.29%</td>
<td>0.6290</td>
</tr>
<tr>
<td>The integrated classifier (mammography)</td>
<td><strong>83.10%</strong></td>
<td><strong>80.05%</strong></td>
<td><strong>85.00%</strong></td>
<td><strong>85.00%</strong></td>
<td><strong>80.65%</strong></td>
<td><strong>0.8089</strong></td>
</tr>
</tbody>
</table>

**Figure 2:** The ROC of the KNN, SVM, and Naive Bayes compared with the integrated classifiers on ultrasound images.

**Figure 3:** The ROC of the KNN, SVM, and Naive Bayes compared with the integrated classifiers on mammography images.

**Figure 4:** The ROC of the integrated classifier on ultrasound images or mammography images compared with the integrated classifiers on multimodal images.

ultrasound images. The higher the sensitivity is, the more the malignancy can be diagnosed. It benefits to early detect breast cancer, so the system with higher sensitivity is valuable in clinical application. So the ensemble method combining the mammography images with ultrasound images has better performance. It is necessary and effective to diagnose breast cancer on multimodal images.

4.3. The Different Selection of Base Classifier. We have proved that the integrated classifiers are better than single classifier. However, there are so many ways to integrate. We use the indicator $R$ to select the base classifier in our work. We ranked the different integrated classifiers with $R$ and chose the top 10 integrated classifiers ($C_1, C_2, \ldots, C_{10}$). Then we tested their performance to demonstrate the effectiveness of indicator $R$. The performance of the different integrated
Table 5: The effectiveness and necessary of multimodal images.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>The integrated classifier based on multimodal images (C1)</td>
<td>88.73%</td>
<td>96.77%</td>
<td>82.50%</td>
<td>97.06%</td>
<td>81.08%</td>
<td>0.8968</td>
</tr>
<tr>
<td>The integrated classifier on ultrasound images</td>
<td>85.92%</td>
<td>83.87%</td>
<td>87.50%</td>
<td>83.87%</td>
<td>83.63%</td>
<td></td>
</tr>
<tr>
<td>The integrated classifier on mammography images</td>
<td>83.10%</td>
<td>80.05%</td>
<td>85%</td>
<td>85%</td>
<td>80.65%</td>
<td>0.8089</td>
</tr>
</tbody>
</table>

The classifier is shown in Table 6 and Figure 5. We can see that the best value of accuracy, sensitivity, specificity, NPV, PPV, and AUC appeared primarily in the integrated classifiers C1 and C2. So indicator $R$ can effectively choose base classifiers. Considering that the sensitivity is more important in the clinical application and the integrated classifier C1 performs much better than the integrated classifier C2 in accuracy and sensitivity, we choose integrated classifier C1 which is ranked first by indicator $R$ to diagnose breast cancer.

4.4. The Performance of the Classifier-Fusion Method Compared with the Feature-Fusion Method. In this paper, we integrated the classifier instead of integrating the feature to prevent the characteristic redundancy. To test the efficiency, we compared it with the different fusion methods. As is shown in Table 7 and Figure 6, the classifier-fusion method is better than the feature-fusion method in all indicators.

4.5. The Performance of Our Method Compared with GASEN. GASEN (Genetic Algorithm based Selected Ensemble) [16] is the first selective ensemble learning method proposed by Zhou et al. To evaluate the effectiveness of our method, we compared the performance of our method with GASEN. As is shown in Table 8, our method is better than GASEN in all indicators. The proposed selective ensemble method in this paper is more suitable for breast cancer diagnosis rather than GASEN.

5. Conclusion

In this paper, we propose a new selective ensemble method to diagnose breast cancer by combining ultrasound images with mammography images. When generating the integrated classifier, we choose the suitable base classifier using new indicator $R$. Then we test our method in a multimodal database containing 71 breast ultrasound images and 71 mammography images. The selective ensemble method is efficient in diagnosing the breast cancer; we can obtain an accuracy of 88.73% and sensitivity of 97.06%. We also prove that the classifier-fusion method is better than the feature-fusion method in all indicators in our database. In our work, we just choose the simple feature. We will concentrate on the relationship of features ultrasound and mammography images in the future.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by Natural Science Foundation of China (NSFC) (no. 61572300), Natural Science Foundation of Shandong Province in China (no. ZR2015FM010 and no. ZR2014FM001), Taishan Scholar Program of Shandong Province in China (no. TSHW201502038), a Project of Shandong Province Higher Educational Science and Technology
Table 6: The performance of the different selection of classifiers.

<table>
<thead>
<tr>
<th>Classifier Selection</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 = Naive Bayes-M + KNN-U + SVM-U</td>
<td>88.73%</td>
<td>96.77%</td>
<td>82.50%</td>
<td>97.06%</td>
<td>81.08%</td>
<td>0.8968</td>
</tr>
<tr>
<td>C2 = Naive Bayes-U + KNN-U + SVM-U</td>
<td>85.92%</td>
<td>83.87%</td>
<td>87.50%</td>
<td>87.50%</td>
<td>83.87%</td>
<td>0.8363</td>
</tr>
<tr>
<td>C3 = KNN-U + Naive Bayes-U + Naive Bayes-M</td>
<td>84.51%</td>
<td>83.87%</td>
<td>85%</td>
<td>87.18%</td>
<td>75.68%</td>
<td>0.8363</td>
</tr>
<tr>
<td>C4 = KNN-U + SVM-U + KNN-M</td>
<td>83.10%</td>
<td>90.32%</td>
<td>77.50%</td>
<td>91.18%</td>
<td>75.68%</td>
<td>0.8363</td>
</tr>
<tr>
<td>C5 = Naive Bayes-U + SVM-U + Naive Bayes-M</td>
<td>84.51%</td>
<td>80.65%</td>
<td>87.50%</td>
<td>85.37%</td>
<td>83.33%</td>
<td>0.8097</td>
</tr>
<tr>
<td>C6 = KNN-U + SVM-U + SVM-M</td>
<td>77.46%</td>
<td>83.87%</td>
<td>72.50%</td>
<td>85.29%</td>
<td>70.27%</td>
<td>0.7782</td>
</tr>
<tr>
<td>C7 = KNN-U + Naive Bayes-M + KNN-M</td>
<td>77.46%</td>
<td>77.42%</td>
<td>77.50%</td>
<td>81.58%</td>
<td>72.73%</td>
<td>0.7500</td>
</tr>
<tr>
<td>C8 = KNN-U + Naive Bayes-U + KNN-M</td>
<td>77.46%</td>
<td>74.19%</td>
<td>80%</td>
<td>80%</td>
<td>74.19%</td>
<td>0.7468</td>
</tr>
<tr>
<td>C9 = KNN-U + Naive Bayes-M + SVM-M</td>
<td>74.65%</td>
<td>77.42%</td>
<td>72.50%</td>
<td>80.56%</td>
<td>68.57%</td>
<td>0.7306</td>
</tr>
<tr>
<td>C10 = SVM-U + Naive Bayes-M + KNN-M</td>
<td>74.65%</td>
<td>70.97%</td>
<td>77.50%</td>
<td>77.50%</td>
<td>70.97%</td>
<td>0.7234</td>
</tr>
</tbody>
</table>

Table 7: The performance of the classifier-fusion method compared with the feature-fusion method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes-feature-fusion</td>
<td>69.01%</td>
<td>80.65%</td>
<td>60%</td>
<td>80%</td>
<td>60.98%</td>
<td>0.6919</td>
</tr>
<tr>
<td>SVM-feature-fusion</td>
<td>76.06%</td>
<td>64.52%</td>
<td>85%</td>
<td>75.56%</td>
<td>76.92%</td>
<td>0.7290</td>
</tr>
<tr>
<td>KNN-feature-fusion</td>
<td>67.61%</td>
<td>61.29%</td>
<td>72.50%</td>
<td>70.73%</td>
<td>63.33%</td>
<td>0.6879</td>
</tr>
<tr>
<td>The classifier-fusion method (C1)</td>
<td>88.73%</td>
<td>96.77%</td>
<td>82.50%</td>
<td>97.06%</td>
<td>81.08%</td>
<td>0.8968</td>
</tr>
</tbody>
</table>

Table 8: The performance of our method compared with GASEN.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>GASEN</td>
<td>69.01%</td>
<td>80.65%</td>
<td>60%</td>
<td>80%</td>
<td>60.98%</td>
<td>0.6919</td>
</tr>
<tr>
<td>Our method</td>
<td>88.73%</td>
<td>96.77%</td>
<td>82.50%</td>
<td>97.06%</td>
<td>81.08%</td>
<td>0.8968</td>
</tr>
</tbody>
</table>

Program (no. J15LN20), and a Project of Shandong Province Medical and Health Technology Development Program (no. 2016WS0577).

References


