

Ensemble learning enabled classification of breast cancer with shear wave elastography images



Jiaming Huang, Kunhou Zhou, Yingdan Hu, Zhaochen Liu, Xiaochuan Tian, Lintao Zhang

Supervisor: Zhaoyang Liu Collaborator: Zhongshan Hospital, Fudan University (Shanghai, CN) SURF ID: SURF-2024-0396

Abstract

In this experiment, we explored four ML approaches, including XGB, SVM, RF, LR, and nine DL approaches, including AlexNet, GoogLeNet, ShuffleNet, ResNet (34, 50, 101), Vision Transformer (ViT), EfficientNetV2 (s, m, l), MobileNet (V2, V3_L, V3_S), DenseNet (121, 161, 169, 201) and MobileViT (s, x_s, xx_s), to classify breast cancer with SWE images. Each original SWE image is cropped into four types of images (RE-SWE, RE-US, SWE, US). Results demonstrate the model with the highest AUROC for each image type. To give a visual interpretability, four CAM-based approaches are employed. Moreover, a webserver is constructed to facilitate usage. Future research involves ensemble learning by averaging predictions from four models with the highest AUROC for each image type.

1. Introduction

A glimpse of ML and DL in advancing image classification

ML and DL have advanced medical image classification at an astonishing pace. Parallel ML methods of SVM, RF, and XGB accelerate hyperspectral skin cancer image classification [1]. ResNet-34, pre-trained on ImageNet, classifies dental radiographs with high accuracy [2]. ViT is used to classify thyroid nodules with contrastive learning in ultrasound images [3]. ShuffleNet improved image classification method for cervical precancerous lesions [4]. Moreover, based on hierarchical local information and GoogLeNet-based representations, NHL pathological image classification achieves high accuracy [5].

Shear wave elastography (SWE)

SWE is an ultrasound-based imaging technique that measures tissue stiffness. It works by generating shear waves through focused ultrasound pulses within the tissue [6]. The speed of these waves, which varies with tissue stiffness, is then tracked by the ultrasound system. Stiffer tissues cause faster wave propagation, while softer tissues slow the waves down. This data is used to create a color-coded elastogram, visually mapping the stiffness of the tissue in real time.

Motivations

SWE offers informative images for breast cancer classification. Though CNN is already used to capture these characteristics [7], its performance can be improved with more advanced models. In this study, four traditional ML methods and nine DL methods are evaluated and used for breast cancer classification with SWE images. To seek for the best training data, the original image set is cropped into four types of images, which are RE-SWE, SWE, RE-US, and US, respectively. Results demonstrated the ROC plots with AUROC labeled for each model; specifically, the model with the highest AUROC is further stated for each image type. Future research involves ensemble learning to elevate the overall performance by taking account of comprehensive prediction results from these well-performed models.

2. Method

Here demonstrated representative 3 ML models and 4 DL model are demonstrated below:

3 ML models

Logistic Regression (LR)

Logistic Regression uses a linear combination of input features, followed by the logistic function, to model the probability of a binary outcome.

Support Vector Machine (SVM)

SVM separates points with hyperplanes that maximize the margin between classes and can handle nonlinearity through kernel methods

eXtreme Gradient Boost (XGB)

XGB includes regularization to prevent overfitting, parallel tree boosting for faster training, and the ability to handle sparse data.

4 DL models

ResNet-50

ResNet-50 [8] is a deep CNN with 50 layers with residual, which helps mitigate the vanishing gradient problem by allowing gradients to flow through the network more effectively during training.

Vision Transformer (ViT)

ViT [9] applies the transformer architecture to image recognition tasks by dividing an image into patches and processing them as a sequence of tokens, eventually allowing the model to capture long-range dependencies.

ShuffleNet

ShuffleNet [10] applies grouped convolution first and then channel shuffle. After that, depthwise separable convolution is employed, which is comprised of depthwise convolution and pointwise convolution.

GoogLeNet

GoogLeNet [11], a 144-layer network, features Inception blocks composed of four branches, utilizing 1x1 convolutions for dimensionality reduction and mapping, with two auxiliary classifiers and average pooling instead of fully connected layers.

Ensemble learning

Ensemble learning involves combining the predictions from diverse models by averaging the outputs of these models. The ensemble can leverage the strengths of each model, reducing the risk of overfitting and improving overall prediction accuracy.

3. Workflow

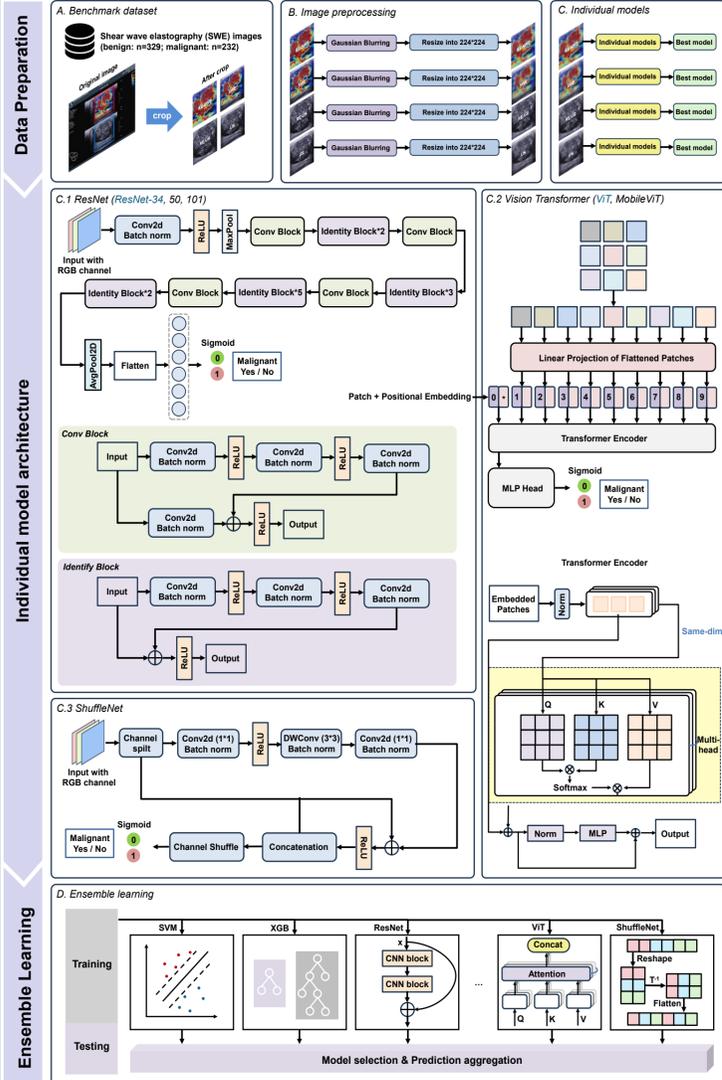


Figure 1. The overall workflow of model construction. Detail procedures include benchmark datasets, image preprocessing, individual models, and ensemble learning.

4. Results and analysis

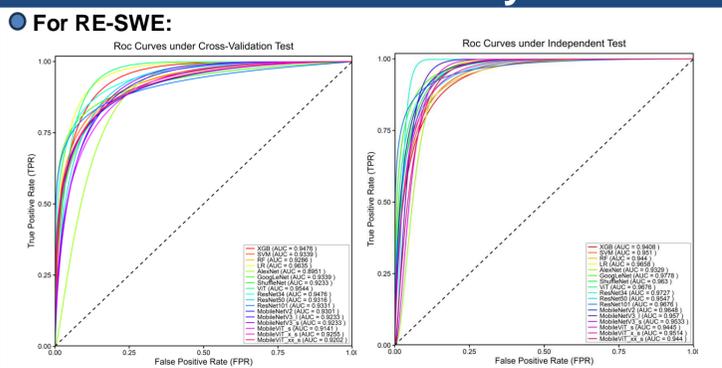


Figure 2. ROC curves under cross-validation test and independent test set for RE-SWE images

(Result 1) ViT outperforms others under the cross-validation test, while GoogLeNet outperforms others under the independent test.

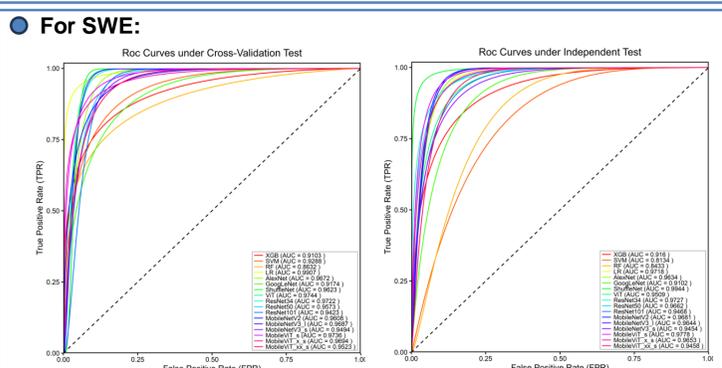


Figure 3. ROC curves under cross-validation set and independent test set for SWE images

(Result 2) LR outperforms others under the cross-validation test, while ShuffleNet outperforms others under the independent test.

4.1 ROC Plots (Continued)

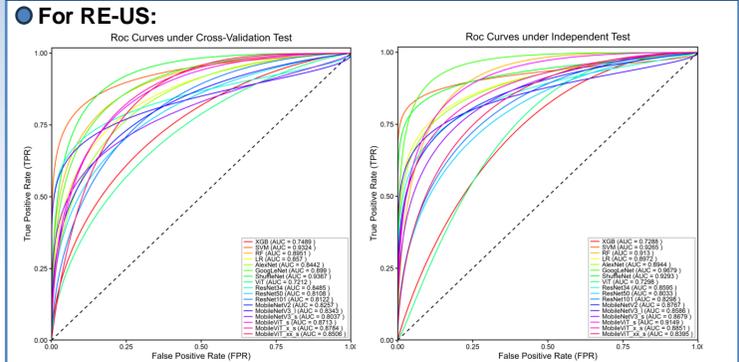


Figure 4. ROC curves under cross-validation test and independent test set for RE-US images

(Result 3) ShuffleNet outperforms others under the cross-validation test GoogLeNet, while outperforms others under the independent test.

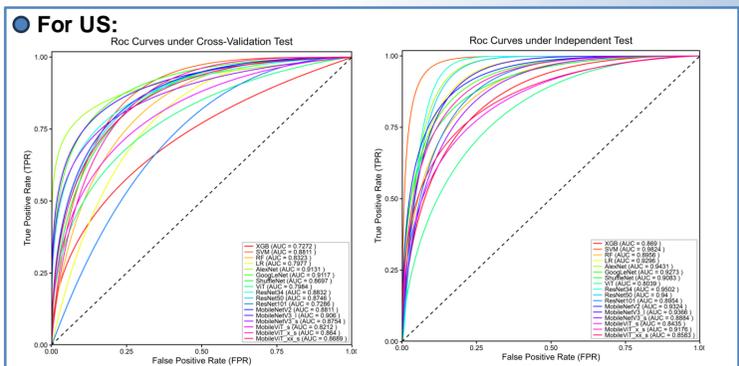


Figure 5. ROC curves under cross-validation test and independent test set for US images

(Result 4) AlexNet outperforms others under the cross-validation test ResNet-34, while outperforms others under the independent test.

4.2 Visualization Interpretability

To give a visual interpretability, Grad-CAM, Grad-CAM++, Layer-CAM, and Score-CAM are applied with MobileViT (specifically, _xx_s) on a typical malignant RE-SWE image (Figure 5).

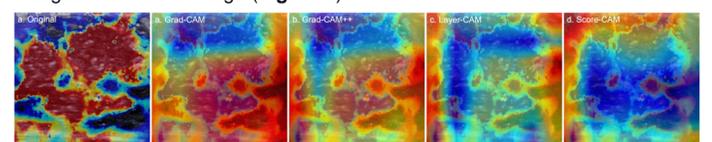


Figure 5. Visual interpretability toward a typical malignant RE-SWE image with Grad-CAM, Grad-CAM++, Layer-CAM, and Score-CAM.

(Result 5) Grad-CAM pays more attention to tumor areas than others, thus functioning better for visualization interpretability.

4.3 Webserver

To facilitate broader usage, a user-friendly webserver is developed, which can be freely visited at <http://www.biotools.tech/SWEBreCA-Pred/index.html>.

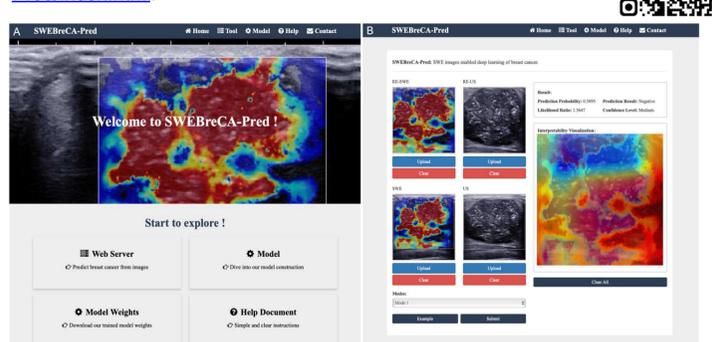


Figure 6. Screenshot of the webserver. A is for the main page and B is for the tool page.

4.4 Code Availability

For complete codes and ensemble learning results in this project, please visit my Github repository at: <https://github.com/Jiaming21/SWEBreCA-Pred.git>

5. Conclusion

In conclusion, GoogLeNet, ShuffleNet, GoogLeNet, and ResNet-34 are the four models with the highest AUROC under the independent test respectively. After the ensemble learning by averaging predictions from four models with the highest AUROC for each image type, the AUROC is further elevated.

5. References

- [1] B. Petracchi, E. Torti, E. Marenzi, and F. Loporati, 'Acceleration of Hyperspectral Skin Cancer Image Classification through Parallel Machine-Learning Methods', Sensors, vol. 24, no. 5, Art. no. 5, Jan. 2024, doi: 10.3390/s24051399.
- [2] J. E. Cejudo, A. Chaurasia, B. Feldberg, J. Krois, and F. Schwendicke, 'Classification of Dental Radiographs Using Deep Learning', Journal of Clinical Medicine, vol. 10, no. 7, Art. no. 7, Jan. 2021, doi: 10.3390/jcm10071496.
- [3] J. Sun et al., 'Classification for thyroid nodule using ViT with contrastive learning in ultrasound images', Computers in Biology and Medicine, vol. 152, p. 106444, Jan. 2023, doi: 10.1016/j.cbiomed.2022.106444.
- [4] S. Fang, J. Yang, M. Wang, S. Liu, and C. Liu, 'An Improved Image Classification Method for Cervical Precancerous Lesions Based on ShuffleNet', Computational Intelligence and Neuroscience, vol. 2022, Jan. 2022, doi: 10.1155/2022/9675628.
- [5] J. Bai, H. Jiang, S. Li, and X. Ma, 'NHL Pathological Image Classification Based on Hierarchical Local Information and GoogLeNet-Based Representations', BioMed Research International, vol. 2019, no. 1, p. 1065652, 2019, doi: 10.1155/2019/1065652.
- [6] J. M. Chang et al., 'Clinical application of shear wave elastography (SWE) in the diagnosis of benign and malignant breast diseases', Breast Cancer Res Treat, vol. 129, no. 1, pp. 89-97, Aug. 2011, doi: 10.1007/s10549-011-1627-7.
- [7] R. Hoffmann, C. Reich, and K. Skerl, 'Evaluating different combination methods to analyse ultrasound and shear wave elastography images automatically through discriminative convolutional neural network in breast cancer imaging', Int J CARS, vol. 17, no. 12, pp. 2231-2237, Dec. 2022, doi: 10.1007/s11548-022-02737-6.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, 'Deep Residual Learning for Image Recognition', in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
- [9] A. Dosovitskiy et al., 'An image is worth 16x16 words: Transformers for image recognition at scale', in Proc. Int. Conf. Learn. Represent., 2020, pp. 1-12.
- [10] X. Zhang, X. Zhou, M. Lin, and J. Sun, 'ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices', in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Jun. 2018, pp. 6848-6856, doi: 10.1109/CVPR.2018.00716.
- [11] C. Szegedy et al., 'Going deeper with convolutions', in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.