Title:

Circulating tumor cells (CTCs) enumeration and machine-learning based diagnostic biomarkers for breast cancer detection

Authors

Chun-Yu Liu1,2,3, Yu-Hsiang Lin1, Yi-Fang Tsai1,2,4, Po-Yen Lu5, Yu-Hsuan Li5, Ji-Lin Chen2, Chi-Cheng Huang2,4,6, Yen-Shu Lin1,2,4, Ta-Chung Chao1,2,3, Chin-Jung Feng1,2,4, Chih-Yi Hsu1,7, Jen-Hwey Chiu2,4,8, Chyong-Mei Chen5,\*, Ling-Ming Tseng1,2,4,\*

Affiliations

1School of Medicine, College of Medicine, National Yang Ming Chiao Tung University, Taipei 11217, Taiwan.

2Comprehensive Breast Health Center, Department of Surgery, Taipei Veterans General Hospital, Taipei 11217, Taiwan.

3Division of Medical Oncology, Department of Oncology, Taipei Veterans General Hospital, Taipei 11217, Taiwan.

4Division of Breast Surgery, Department of Surgery, Taipei Veterans General Hospital, Taipei 11217, Taiwan.

5Institute of Public Health, College of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan.

6Department of Public Health, College of Public Health, National Taiwan University, Taipei, Taiwan.

7Department of Pathology and Laboratory Medicine, Taipei Veterans General Hospital, Taipei 11217, Taiwan.

8Institute of Traditional Medicine, School of Medicine, National Yang Ming Chiao Tung University, Taipei 11217, Taiwan.

|  |
| --- |
| This folder contains the R code for implementing the proposed method.  To use the new code, please conduct the following two steps.  Step 1: Implement the code “code\_SVM\_linear\_Model.r” to construct the svm-forest machine. In this file, readers need to determine the features by setting biomarker.name=c("Age", "CK18", "MGB", "WBC","Platelet") or biomarker.name=c("Age", "CK18", "MGB").  Step 2: Implement “code\_prediction for test data.r” to predict the test data, test\_data.csv. |