Information Theory in Cancer Biology

Sultan Alam   
School of Computer Science and Engineering  
*VIT Bhopal University*Sehore, 466114, India  
mohammadsultanalam@vitbhopal.ac.in

Ayush Johari  
School of Computer Science and Engineering  
*VIT Bhopal University*Sehore, 466114, India  
ayush.johari2021@vitbhopal.ac.inAbhishek Rawat  
School of Computer Science and Engineering  
*VIT Bhopal University*Sehore, 466114, India  
abhishek.rawat2021@vitbhopal.ac.in

Dhananjay Sharma  
School of Computer Science and Engineering  
*VIT Bhopal University*Sehore, 466114, India  
dhananjay.sharma2021@vitbhopal.ac.inManan Gupta  
School of Computer Science and Engineering  
*VIT Bhopal University*Sehore, 466114, India  
manangupta2021@vitbhopal.ac.in

Yashveer Patle  
School of Computer Science and Engineering  
*VIT Bhopal University*Sehore, 466114, India  
yashveer.patle2021@vitbhopal.ac.in

Abstract

This paper investigates the application of independent component analysis (ICA) for dimensionality reduction in breast cancer decision support systems and its impact on the performance of various classifiers, including artificial neural networks (ANN), k-nearest neighbors (k-NN), radial basis function neural networks (RBFNN), and support vector machines (SVM). Utilizing the Wisconsin Diagnostic Breast Cancer (WDBC) dataset, which originally comprises 30 features across 569 instances, ICA effectively reduces the dimensionality to a single feature. Experimental results demonstrate that ICA enhances the sensitivity and computational efficiency of RBFNN and SVM classifiers, achieving higher diagnostic accuracy for malignant tumor identification while slightly reducing performance in k-NN and ANN. The study employs cross-validation and partitioning techniques to evaluate classifiers based on metrics such as accuracy, sensitivity, specificity, and Youden’s index. Results show that the reduced feature set improves RBFNN’s accuracy from 87.17% to 90.49% and increases SVM sensitivity from 96.07% to 97.47%, validating ICA as a powerful tool for feature reduction in cancer diagnosis. These findings highlight the potential of ICA in reducing computational complexity without compromising diagnostic accuracy, making it a viable solution for scalable and efficient breast cancer classification systems.

# Introduction

Breast cancer is one of the leading causes of– death among all cancers for women. Early detection and correct diagnosis of cancer are essential for the treatment of the disease. How- ever, the traditional approach to cancer diagnosis depends highly on the experience of doctors and their visual inspections. Naturally, human beings can make mistakes due to their limitations. Humans can recognize patterns easily. How- ever, they fail when probabilities have to be assigned to observations [2]. Although several tests are applied, exact diagnosis may be difficult even for an expert. That is why automatic diagnosis of breast cancer is investigated by many researchers. Computer aided diagnostic tools are intended to help physicians in order to improve the accuracy of the diagnosis [3]–[5]. A study was carried out to demonstrate that the machine learning may improve the accuracy of diagnosis. In Brause’s work, the result shows that the most experienced physician can diagnose with 79.97% accuracy while 91.1% correct diagnosis is achieved with the help of machine learning [6].

Tumors are classified as benign and malignant. Benign tumors are not cancerous or life threatening. However, these can increase the risk of getting breast cancer. Malignant tumors are cancerous and more alarming than benign tumors. Although significant studies are performed for early detection, about 20% of all women with malignant tumors die from this disease [7].

Support vector machine (SVM) is an effective statistical learning method for classification [14]. SVM is based on finding optimal hyperplane to separate different classes mapping input data into higher-dimensional feature space. SVM has advantage of fast training technique, even with large number of input data [15], [16]. Therefore, it has been used for many recognition problems such as object recognition and face detection [17]–[19].

The objective of the proposed study is to analyse the effect of feature reduction using ICA on classifications of the tumors as benign or malignant. Thus, the dimension of WDBC dataset is reduced into only one feature using ICA. The reduced data is subdivided into test and training data using 5/10-fold cross-validation and 20% partitioning to evaluate the performance of 𝑘*-*NN, ANN, RBFNN, and SVM. Performance measures including accuracy, specificity, sensitivity, 𝐹-score, Youden’s index, and discriminant power are computed and the receiver operating characteristic (ROC) curve is plotted to compare the classifiers. Section 2 summarises background knowledge on dataset, ICA, 𝑘-NN, ANN, RBFNN, SVM, and performance measures. In Section 3, the methodology deployed in this study is described. In Sections 4 and 5 experimental results are presented and discussed.

Therefore, it is not surprising that information theory has found many theoretical advancements and witnessed myriad applications dealing with biological data [8]. This is particularly true in the umbrella field of computational biology and bioinformatics that deals with computational

applications of mathematical and statistical methods in the study of biological systems and processes [3]. In this domain, information theory is widely used for model development and data analysis for a variety of biologically derived data types ranging from molecular, sequence and phenotypic data in genomics and genetics to gene expression, protein and spectral data in transcriptomics, proteomics, and metabolomics, respectively [17], [24].

# Litreature Review

## Applications of Information Theory in Cancer Research

Several studies have applied information theory to analyse cellular communication in cancer. For example, entropy and mutual information have been used to assess how effectively cells transmit signals in key cancer-related pathways. These

metrics help quantify how genetic and molecular disruptions impact cancer development, providing insight into tumour behaviour and potential vulnerabilities [25].

## Role of GNNs in Cancer Biology

Graph neural networks (GNNs) have emerged as powerful tools for modelling complex biological interactions in cancer research. GNNs allow researchers to analyse molecular structures, signalling pathways, and genetic networks to predict outcomes such as drug responses or cancer progression. They can incorporate multi-omics data, offering a more holistic understanding of cancer biology.

GNNs are superficially similar to graphical models, in that both perform learning over graph structures. Bayesian networks (BNs), or probabilistic directed acyclic graphs (DAGs) learned from the data, are arguably the most popular graphical models in life science applications. BNs can incorporate both data-driven learning and existing knowledge, and allow for probabilistic reasoning and propagation over the DAGs. A major feature of BNs is that they filter out superficial (transitive, non-direct) dependencies, thus arriving at sparse DAGs suggesting directional casualties. A question is often asked: what are the principal differences between BNs and GNNs, especially from the life sciences application perspective? Here, we compare the underlying fundamentals of a GNN (specifically, a GCN) and a BN [29].

## DNA Methylation and Epigenetics in Cancer

Information theory has also been applied to understand DNA methylation, an epigenetic process known to play a critical role in cancer. For example, studies have shown that certain methylation patterns can serve as key drivers of cancer by altering gene expression. Understanding these epigenetic modifications through the lens of information theory provides valuable insights into how cancers develop and evolve. The combination of computational models, such as GNNs, with information theory has allowed for significant progress in cancer research. However, there remain challenges in scaling these models to handle large, high-dimensional datasets. Additionally, more research is needed to refine these approaches to achieve greater predictive accuracy in clinical applications [27].

## Artifical Neural Network

Feedforward neural network (FFNN) is most popular ANN structure due to its simplicity in mathematical analysis and good representational capabilities [27], [28]. FFNN has been used successfully to various applications such as control, signal processing, and pattern classification. FFNN architecture. 𝑁 states the number of input patterns and 𝑀 states the number of neurons in hidden layer. Neurons in the hidden layer receive weighted inputs from a previous layer and transfer output to the neurons in the next layer in FFNN, and these computations can be described as

where 𝑤0 is bias, 𝑤i is the weight of each input neuron, 𝑥i is input neuron, 𝑦net is composed of the summation of weighted inputs denotes the nonlinear activation function, 𝑦 is the observed output value of neural network, and 𝐸 is the error between output value and network result [29].

The most common empirical measure to assess effectiveness

is the accuracy for classifier and it is calculated by

Sensitivity measures the proportion of actual positives which are correctly identified and specificity measures the proportion of negatives which are correctly identified. These are formulated by:

𝐹-score is a measure of test accuracy. It considers both precision and the recall to compute. These are calculated by

where 𝛽 is the bias and 𝐹-Score is balanced when 𝛽 = 1. It

favors recall when 𝛽 < 1 and favors precision otherwise.

In this study, an attempt has been made to evaluate the performance of the classifiers computing the afore mentioned measures for 5/10-fold cross-validations (CV) and 20% data

partitioning. For 5-CV or 10-CV, the data are divided into 5 or 10 subsets, and each subset is sequentially deployed as test

data while others are deployed as training data.

# .METHODOLOGY

In this study, the original 30 features of WDBC data and reduced one feature using ICA are deployed to evaluate the classifiers performance on breast cancer decision. Thus, the proposed model shown in Figure 4 is applied to WDBC data that have 30 features and 569 instances (patients) were used to train and test the models. First, the dimensionality of the data is reduced using ICA and partitioned into subsamples using 5/10-CV and 20% partitioning to evaluate the classifiers [7]-[9]. The subsamples have been used sequentially to train and test ANN, RBFNN, SVM, and 𝑘*-*NN. The outputs of the classifiers have been evaluated to find out performance measures. First, ICA is used to compute ICs. Since the first IC has distinctly large eigenvalue given in Figure 5, it has been selected as a feature vector. In other words, one IC can successfully identify the thirty features with the retained 98.205% of nonzero eigenvalues. In addition, the distribution of the IC is given in Figure 6 to indicate its distinguishing capability. The data are divided into subsets using 5/10-CV and 20% partitioning to test and train classifiers. After training process, the test data are used to evaluate diagnostic performances of the classifiers in terms of sensitivity, specificity, accuracy, 𝐹-score, Youden’s index, DP, and ROC curve [11].

The discriminant function of the hyperplane can be

described by the following equation:

*𝑔 (𝑥) = wT x+ 𝑏*

where 𝑥 describes data points, 𝑤 is a coefficient vector,

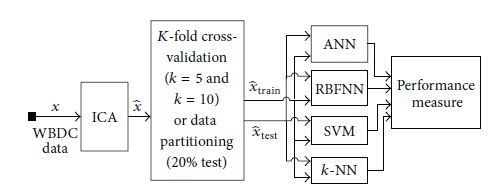
and 𝑏 shows offset from the origin. In case of linear SVM

𝑔(𝑥) ≥ 0 for the closest point on the one of the classes,

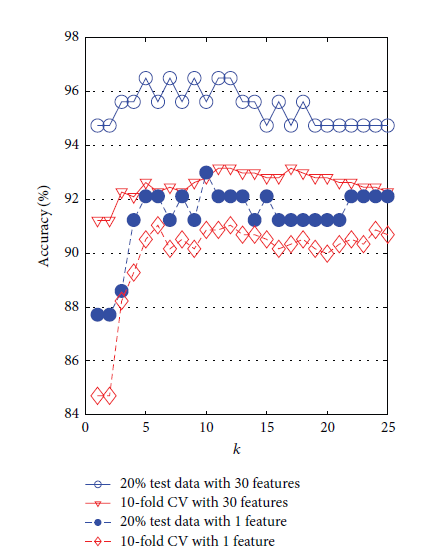
𝑔(𝑥) < 0 for the closest point belongs to another class.

## KNN classifiers

For training processes, 𝑘-NN classifier, one-dimensional Euclidean distance, between test and training samples [35]. The results of 𝑘-NN classifier are obtained for the 𝑘 values from 1 to 25, and then the performance measures at the best 𝑘 value are stored. The model of ANN is selected as feed forward neural network with one hidden layer Moreover, the activation function of the hidden layer of the network has been chosen as log-sigmoid transfer function [13]-[15]. In order to train the network, gradient descent with momentum and adaptive learning rate backpropagation algorithm is used. RBFNN is also evaluated varying the spread value (𝜎). For SVM, linear, quadratic, and RBF kernels are used to explore which type of separating hyperplane is more suitable for breast cancer classification [18].



*Figure 1: The basic model of the study*



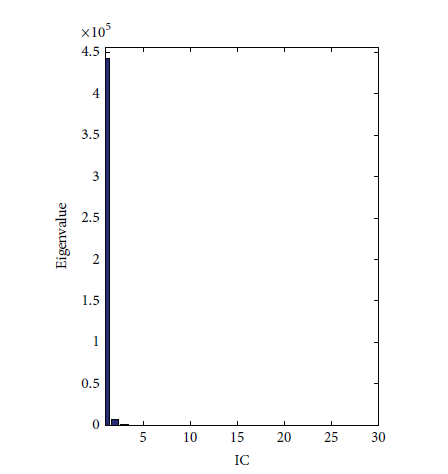
*Figure 2: The graph of accuracy of KNN classifier*

## Gene Data Matrix

* The gene expression data matrix (experimental or synthetic) typically provides the genes in the rows and samples and experimental conditions in the columns.
*  The goal of the network inference method is to use the expression matrix to infer the set of regulatory interactions (direct or indirect regulation) between any two genes in the GRN, thereby predicting the edges in the network.
* Because mutual information (MI) provides the ability to capture nonlinear dependencies between two variables, several methods for reconstructing GRNs use MI or associated information theoretic scores to infer the regulatory relationships.
* *J*th column is given by the MI between genes Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersted. This often leads to confusion because equations do not balance dimensionally.

## WBC DataSet

Sensitivity refers successfully identified malignant samples in cancer classification. Thus, higher sensitivity means higher diagnostic capability of malignant tumors and it can be used to help physicians to diagnose cancerous mass more correctly [33]. The accuracy and sensitivity measures of previous classification studies and this study on WDBC dataset are given to compare the effect of feature reduction using ICA. It should be noted that the studies on WDBC differ from studies on WBC dataset which consists of 699 instances with 10 attributes. Higher number of features used to classify breast cancer as benign and malignant results in slightly better accuracy.

*Figure 3: Corresponding eigrnvalues of the WBC data*

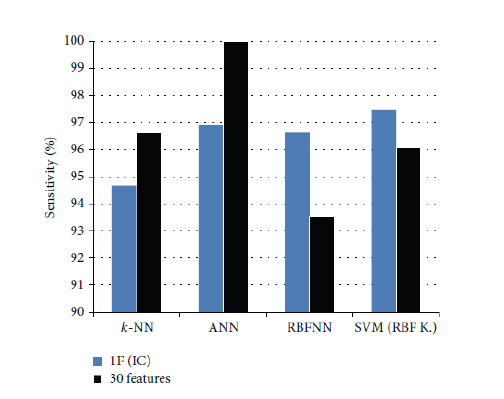
# Results and analysis

One-dimensional feature vector of WDBC data reduced using ICA issued for training and testing the classifier. The accuracy, sensitivity, and specificity of one dimensionality have been performed using 5/10 CV technique and 20% of data as test data. Also, the success of the breast cancer classification is generally evaluated on the basis of sensitivity value because the classifying of the malignant mass is more important than the benign mass [32].

## KNN Accuracy

The accuracy of the 𝑘-NN classifier has been computed for varying 𝑘 values between 1and 25. The comparison graph of the effect of ICA on accuracy of 𝑘-NN classifier is shown in Figure 7 The maximum accuracy results when 20% test data with 30 features is 96.49% where 𝑘 = 5. However, reduced one feature vector using ICA provides the accuracy of 92.98% where 𝑘 = 5 and 20% test data is selected [27]. Moreover, the accuracy of 𝑘-NN classifier is decreased from 93.15% (30 features) to 91.04% (1 feature by ICA) when 10-CV issued to test and train. Accuracy graph of ANN has been plotted varying neuron numbers in the hidden layer for

10/5-CV and 20% test data.

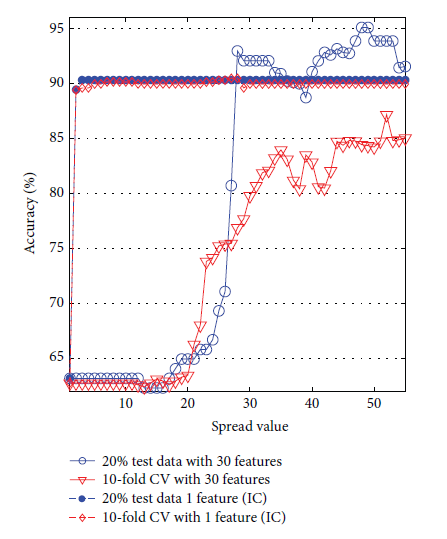


*Figure 4: Sensitivity measures of the classifiers*

## RBFN Accuracy

The accuracy graph of RBFN is shown in Figure 9. Referring to the accuracy graph of RBFNN, maximum accuracy, 95.12%, is obtained where spread value is 48 for 20% test data. This value is decreased to 90.35% when reduced one-dimensional feature vector by ICA is used.

However, when 10-CV is used, the effect of ICA increases the accuracy from 87.18% (with 30 features) to 90.49% (with 1 feature reduced by ICA).



*Figure 5: The accuracy of RBFNN*

## Support Vector Machine (SVM)

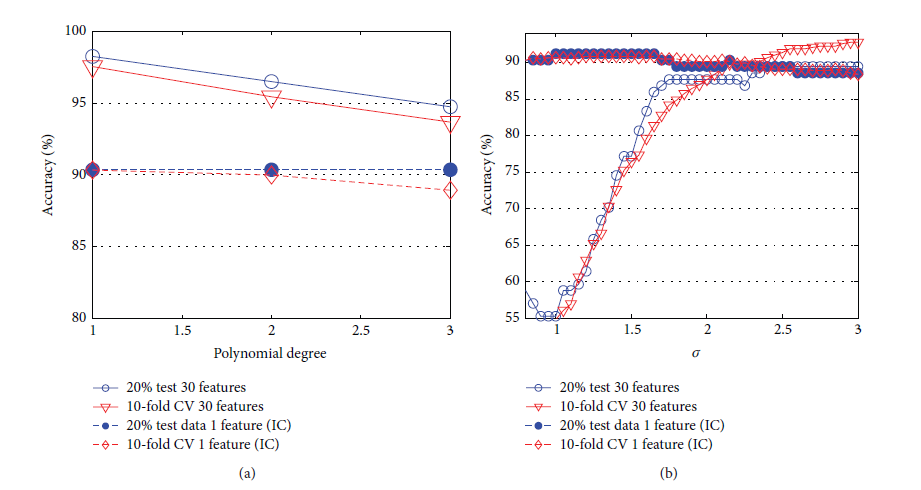
Accuracy evaluation of SVM has been computed for kernel functions including linear, polynomial, and RBF with kernel function parameters such as RBF sigma value for RBF kernel and polynomial degree for polynomial kernel. The accuracy graph of SVM classifier is presented in Figure 10 where the axes of polynomial degree indicate linear kernel when its value equals one.

Generally, SVM classifier with linear kernel provides more accurate result than polynomial and RBF kernel. Its accuracy is 98.25% for 30 features and 90.35% for reduced 1 feature when 20% of data is used as test data. In contrast to polynomial kernel, effect of ICA increases the accuracy of SVM with RBF kernel from 89.47% (30 features) to 91.23% (1 feature). When 10-CV is used, the accuracy is decreased from 97.54% (30 features, linear kernel) and 95.25% (30 features,(RBF kernel) to 90.33% and 90.86%(reduced 1 feature by ICA).

The performance measures of 𝑘-NN, ANN, RBFNN, and SVM classifiers such as sensitivity, specificity, accuracy, 𝐹-score, discriminant power (DP), and Youden’s index are given in Table 4 to compare the effect of ICA on the classification [32]. Discriminant power evaluates how well a classifier distinguishes between positive and negative samples. DP of ANN and SVM with 30 original features differs from 3 which means good discriminant. When ICA is used to reduce to one dimensionality, DP falls to 2.769 (SVM) and 2.655 (ANN). In other words, discriminants turn to fair.

A higher value of Youden’s index shows better ability to avoid failure. 𝑘-NN results in the highest value of Youden’s index; refer to Table 4. Youden’s index is used to plot the ROC curve of a classifier. The true positive rate (sensitivity)

is plotted in function of the false positive rate (1−Specificity) can be used to compute area under the ROC curve (AUC) and 95% confidence interval (CI). AUC equals 1 when all test data is assigned to true class labels [26]. Higher AUC indicates that higher accuracy 95% CI is another indicator of the ROC curve which can be used to test whether a classifier can distinguish the classes. If its value is not 0.5, ANN, RBFNN, and SVM classifiers using one-dimensional feature vector reduced by ICA and 30 features are presented in Figure 11. for cut-off points in a ROC curve. The ROC curve. The criterion values of the ROC curves of classifiers are given in Table 5. AUC of the ANN (0.966) and SVM (0.949) results in higher value when 30 original features are used. However, when classification with 1 feature reduced by ICA decreases the accuracy of 𝑘-NN, ANN, and SVM. However, it increases the accuracy of RBFNN. The afore mentioned classification methods are analysed in terms of computing time given in Table 6 to compare the computational complexities to the classifications with the original 30 features.



*Figure 6: The accuracy of SVM classifiers*

##### Conclusion

In this paper, the effect of dimensionality reduction using independent component analysis (ICA) on breast cancer decision support systems with several classifiers such as artificial neural network (ANN), 𝑘-nearest neighbour (𝑘-NN), radial basis function neural network (RBFNN), and support vector machine (SVM) is investigated. The results of the applied original thirty features of Wisconsin diagnostic breast cancer (WDBC) is compared with the reduced one dimension by ICA. The accuracy rates of the classifications with thirty original features except RBFNN have slightly decreased from 97.53%, 91.03%, and 95.25% to 90.5%, 91.03%, and 90.86%, respectively. However, the one-dimensional feature vector causes RBFNN classifier to be more distinguishing with the increased accuracy from 87.17% to 90.49%.

The study of WDBC data creators set has the highest accuracy (97.50%) using multi surface method tree (MSMT) with 3 selected features. Similarly, hybrid methods are more successful than the others. Breast cancer classifications using probabilistic neural network (PNN) with hybrid feature reduction using discrete wavelet transform (DWT) and ICA or classification using SVM with 6-dimensional feature space obtained by 𝐾-means algorithm [41] have accuracy rates of 96.31% and 97.38%for 10-CV. Particularly, SVM based studies [36, 38] with 30 features have near scores to our one-dimensional results.

Furthermore, the sensitivity rates which define the successfully recognized malignant samples are increased from 93.5% to 96.63% for RBFNN and from 96.07% to 97.47% for SVM, while the others have slight decrease at the rate between 0.96% and 3.09%. If the objective is to increase the rate of the successfully identified malignant breast cancer using RBFNN or decrease computational complexity without loss of the high accuracy rate, feature reduction applying ICA can be a high performance solution.

##### References

1. [1] I. Christoyianni, E. Dermatas, and G. Kokkinakis,

“Fast detection of masses in computer-aided mammography,” IEEE Signal Processing Magazine, vol. 17, no. 1, pp. 54–64, 2000.

1. [2] N. Salim, Medical Diagnosis Using Neural Network, Faculty of Information Technology University, 2013, <http://www.generation5>. org/content/2004/MedicalDiagnosis.asp.
2. A. Tartar, N. Kilic, and A. Akan, “Classification of pulmonary nodules by using hybrid features,” Computational and Mathematical Methods in Medicine, vol. 2013, Article ID 148363
3. Kilic, O. N. Ucan, andO. Osman, “Colonicpolypdetectionin CT colonography with fuzzy rule based 3D templatematching,” Journal of Medical Systems, vol. 33, no. 1, pp. 9–18, 2009.
4. [5] A.Mert,N.Kilic¸, andA.Akan, “Evaluation of bagging ensemble method with time-domain feature extraction for diagnosing of arrhythmia beats,” Neural Computing and Applications, vol. 24, no. 2, pp. 317–326, 2014.
5. R. W. Brause, “Medical analysis and diagnosis by neural networks,” in Proceedings of the 2nd International Symposium on Medical Data Analysis (ISMDA ’01), pp. 1–13, Madrid, Spain, October 2001.
6. T. S. Subashini, V. Ramalingam, and S. Palanivel, “Breast mass classification based on cytological patterns using RBFNN and SVM,” Expert Systems with Applications, vol. 36,no. 3, pp. 5284– 5290, 2009.
7. M. N. Gurcan, H.-P. Chan, B. Sahiner, L. Hadjiiski, N. Petrick,and M. A. Helvie, “Optimal neural networkarchitecture selection:improvement in computerized detection of microcalcifications,” Academic Radiology, vol. 9, no. 4, pp. 420–429, 2002.
8. A. P. Dhawan, Y. Chitre, C. Bonasso, and K.Wheeler, “Radialbasisfunction based classification of mammographic microcalcifications using texture features,” in Proceedings of the 17th IEEE Engineering in Medicine and Biology Annual Conference, pp. 535–536, September 1995.
9. A. T. Azar and S. A. El-Said, “Superior neuro-fuzzy classification systems,” Neural Computing and Applications, vol. 23, no. 1, supplement, pp. 55–72, 2012.
10. M. Jia, C. Zhao, F. Wang, and D. Niu, “A new method for decision on the structure of RBF neural network,” in Proceedings of the 2006 International Conference on Computational Intelligence and Security, pp. 147–150,November 2006.
11. J. K. Sing, S. Thakur, D. K. Basu,M. Nasipuri, and M. Kundu, “High-speed face recognition using self-adaptive radial basis function neural networks,” Neural Computing & Applications, vol. 18, no. 8, pp. 979–990, 2009.
12. R. Huang, L. Law, and Y. Cheung, “An experimental study: on reducing RBF input dimension by ICAand PCA,” in Proceedings of the 2002 International Conference on Machine Learning and Cybernetics, vol. 4, pp. 1941–1945, November 2002.
13. V. N. Vapnik, Statistical Learning Theory, John Wiley & Sons, New York, NY, USA, 1989.
14. M. F. Akay, “Support vector machines combined with feature selection for breast cancer diagnosis,” Expert Systems with Applications, vol. 36, no. 2, pp. 3240–3247, 2009.
15. B.Wang, H. Huang, and X.Wang, “A support vector machine based MSM model for financial short-term volatility forecasting,” Neural Computing and Applications, vol. 22, no. 1, pp. 21– 28, 2013.
16. M. Pontil and A. Verri, “Support vector machines for 3D object recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 6, pp. 637–646, 1998.
17. J. Zhou, G. Su, C. Jiang, Y. Deng, and C. Li, “A face and fingerprint identity authentication systembased onmulti-route detection,” Neurocomputing, vol. 70, no. 4–6, pp. 922–931, 2007.
18. E. Gumus, N. Kilic, A. Sertbas, and O. N. Ucan, “Evaluation of face recognition techniques using PCA, wavelets and SVM,” Expert Systems with Applications, vol. 37, no. 9, pp. 6404–6408, 2010.
19. ]S. Kara, A. G¨uven, and S. Ic¸er, “Classification of macular and optic nerve disease by principal component analysis,” Computers in Biology and Medicine, vol. 37, no. 6, pp. 836–841, 2007.
20. A. Hyv¨arinen and E. Oja, “Independent component analysis: algorithms and applications,” Neural Networks, vol. 13, no. 4-5,pp. 411–430, 2000.
21. M. P. S. Chawla, “A comparative analysis of principal component and independent component techniques for electrocardiograms,” Neural Computing and Applications, vol. 18, no. 6, pp. 539–556, 2009.
22. S. D. Villalba and P. Cunningham, “An evaluation of dimension reduction techniques for one-class classification,” Artificial Intelligence Review, vol. 27, no. 4, pp. 273–294, 2007.
23. W. H.Wolberg,W. N. Street, and O. L. Mangasarian, “Machine learning techniques to diagnose breast cancer from imageprocessed nuclear features of fine needle aspirates,” Cancer Letters, vol. 77, no. 2-3, pp. 163–171, 1994
24. K. H. Liu, B. Li, Q. Q. Wu, J. Zhang, J. X. Du, and G. Y. Liu, “Microarray data classification based on ensemble independent component selection,” Computers in Biology and Medicine, vol. 39, no. 11, pp. 953–960, 2009.
25. 2013, <http://research.ics.tkk.fi/ica/fastica/>.
26. N. Sivri, N. Kilic, and O. N. Ucan, “Estimation of stream temperaturein Firtina Creek (Rize-Turkiye) using artificial neural network model,” Journal of Environmental Biology, vol. 28, no. 1, pp. 67–72, 2007.
27. O. A. Abdalla, M. H. Zakaria, S. Sulaiman, and W. F.W. Ahmad, “A comparison of feed-forward back-propagation and radial basis artificial neural networks: A Monte Carlo study,” in Proceedings of the International Symposium in Information Technology (ITSim ’10), vol. 2, pp. 994–998, Kuala Lumpur, Malaysia, June 2010.
28. B. E. Boser, I. M.Guyon, and V.N.Vapnik, “Atraining algorithmfor optimal margin classifiers,” in Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, pp. 144– 152, July 1992.
29. V.N.Vapnik, The Nature of Statistical Learning Theory, Springer, New York, NY, USA, 1995.
30. R. Courant and D. Hilbert, Methods of Mathematical Physics, Wiley, New York, NY, USA, 1953.
31. W. J. Youden, “Index for rating diagnostic tests,” Cancer, vol. 3, no. 1, pp. 32–35, 1950.
32. L. L. Pesce and C. E. Metz, “Reliable and computationally efficient maximum-likelihood estimation of proper binormal ROC curves,” Academic Radiology, vol. 14, no. 7, pp. 814–829, 2007.
33. J. Hamidzadeh, R. Monsefi, and H. S. Yazdi, “DDC: distance based decision classifier,” Neural Computing & Applications, vol. 21, no. 7, pp. 1697–1707, 2012.
34. A. M. Krishnan, S. Banerjee, C. Chakraborty, and A. K. Ray, “Statistical analysis of mammographic features and its classification using support vector machine,” Expert Systems with Applications, vol. 37, no. 1, pp. 470–478, 2010.
35. S. C. Bagui, S. Bagui, K. Pal, and N. R. Pal, “Breast cancer detection using rank nearest neighbor classification rules,” Pattern Recognition, vol. 36, no. 1, pp. 25–34, 2003.
36. N. H. Sweilam, A. A. Tharwat, and N. K. Abdel Moniem, “Support vector machine for diagnosis cancer disease: a comparative study,” Egyptian Informatics Journal, vol. 11, no. 2, pp. 81–92, 2010.
37. O. L. Mangasarian, W. N. Street, and W. H. Wolberg, “Breast cancer diagnosis and prognosis via linear programming,” Operations Research, vol. 43, no. 4, pp. 570–577, 1995.
38. A. Mert, N. Kılıc¸, and A. Akan, “An improved hybrid featurereduction for increased breast cancer diagnostic performance,” Biomedical Engineering Letters, vol. 4, no. 3, pp. 285–291, 2014.
39. B. Zheng, S. W. Yoon, and S. S. Lam, “Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms,” Expert Systems with Applications, vol. 41, no. 4, pp. 1476–1482, 2014.