An Improved Breast Cancer Classification Method Using an Enhanced AdaBoost Classifier

Yousef K. Qawqzeh[1]*, Abdullah Alourani[2], Sameh Ghwanmeh[3]
Information Technology College, University of Fujairah, Fujairah, 125212, United Arab Emirates[1, 3]
Department of Computer Science and Information, College of Science in Zulfi, Majmaah University
Al-Majmaah, 11952, Saudi Arabia[2]

***Abstract*—The goal of this research is to create a machine learning (ML) classifier that can improve breast cancer (BC) diagnosis and prediction. The principle components analysis (PCA) technique is used in this work to minimize the dimensions of the BC dataset and achieve better classification metrics. The developed classifier outperformed others in terms of F1 score and accuracy score. Using the original BC dataset, four different classifiers are applied to determine the best classifier in terms of performance metrics. The used classifiers were RandomForest, DecisionTree, AdaBoost, and GradientBoosting. The RandomForest classifier obtained (95.7%) f1 score and (94.5%) accuracy score, the DecisionTree classifier obtained (93%) f1 score and (91%) accuracy score, the GradientBoosting classifier obtained (95%) f1 score and (93.5%) accuracy score, and the AdaBoost classifier obtained (95.8%) f1 score and (94.5%). The AdaBoost classifier was utilized to create the final model using the reduced PCA dataset because it scored the highest performance metrics. The developed classifier is named as "pcaAdaBoost". The optimized pcaAdaBoost achieved higher performance metrics in terms of f1 score (99%) and accuracy score (98.8%). The results reveal that the optimized pcaAdaBoost scored highest performance measures in terms of cross-validation and testing outcomes, with an overall accuracy of (99%). The improved results justify the use of dimensionality reduction in high-dimension datasets to reduce complexity and improve performance measures.**

***Keywords*—Breast cancer; diagnosis; prediction; AdaBoost, RandomForest; PCA**

I. Introduction

Several studies and technologies have been conducted worldwide to screen for and investigate the risk of Breast Cancer (BC). Despite significant advances in screening and patient management, BC represents one of the common malignancies in women worldwide and it was ranked as the second most likely cause of cancer mortality. Statistically, there were 268,600 new cases of BC diagnosed in American women in 2019, with 41,760 deaths [1-3]. BC is a very diverse illness with numerous forms and subtypes. Approximately 95 % of BCs responded to endocrine and targeted therapy, and their prognosis and survival rates are generally favorable. However, the widely used screening instrument is a two-dimensional mammography, which can detect tumors that are too small to perceive. The breast is compressed between two rigid plates in a conventional mammogram, and X-rays have been used to capture images of the breast tissue. Such techniques are invasive, costly, and tedious to conduct. With the advent of new computational power in terms of big data, machine learning (ML), and data science (DS), scholars have attempted to apply such new computation techniques to the analysis of BC datasets, as well as to develop new promising low cost and fast BC classification techniques. However, to reduce processing time and to increase prediction performance, data reduction techniques are used. Removing irrelevant input data and get rid of redundant inputs would probably enhance classifier's capability in terms of performance measures. In this study, the PCA technique is used in this work to minimize the dimensions of the BC dataset and achieve better classification metrics. The generated classifier outperformed others in terms of F1 score and accuracy score. Using the original BC dataset, four different classifiers are applied to determine the best classifier in terms of performance metrics. Therefore, the following classifiers, RandomForest, DecisionTree, AdaBoost, and GradientBoosting, are used and evaluated. This work utilized the PCA method to reduce features from the original BC dataset. This method improved the performance of the ML model on hand and enabled better data visualization. In this way, the PCA is used to reduce the dimensions of the BC dataset, making it less sparse and more statistically significant.

*Corresponding Author.

II. Literature Reviews

Scholars defined ML as a subset of artificial intelligence (AI). It denotes a mathematical model that is used to make decisions or predictions using a training dataset. It is frequently referred to as an evolving prediction model that will improve classification capabilities in a variety of fields including disease diagnosis and screening as in medical industry [4]. It is necessity to reduce the danger of diseases, infections, disorders, or pandemics using a proactive ML model [5-6]. To deal with the increasing complexity of the vast data and convert it into meaningful scientific knowledge for the benefit of humanity, new bioinformatics methods must be developed [7-8]. The use of ML classification techniques in medical diagnosis applications is highly valued [6]. However, traditional classification may not have performed as well as planned, raising the necessity of such investigations that could improve the current classification technologies in medical sectors. The goal of medical AI and ML research is to create applications that use AI technologies to aid practitioners in providing treatment based on better decision making [9-10]. Some research has been conducted on comprehensible AI in order to solve the downfalls of AI analysis tools being black boxes. In comparison to AI systems such as deep learning, XAI can present model's explanations and decision-making

capabilities [11]. Many traditional ML approaches for classification problems are used like logistic regression (LR), support vector machine (SVM), decision tree (DT), and RandomForest (RF). The RF is a ML classification method that comprised of several decision trees in an ensemble. The outcome of these DT elections symbolizes the RF decision. Regardless of the used ML algorithm, several evaluation methods such as F1_score, accuracy_score, precision, recall, AUC, and ROC have been commonly used to assess the effectiveness of each proposed method [12]. One of the most important areas of medical based ML applications is BC classification. BC has now surpassed lung cancer as being the most prevalent malignancy afflicted in women worldwide [13]. For the prognosis and diagnosis of BC disease, researchers developed a SVMs based classifier in contrast to Bayesian classifiers and ANN. However, they gave implementation summary for the findings of the evaluated classifiers [14]. Another study [15] used the BC dataset to classify BC disease using several ML methods: Knn, RF, SVM, DT, and LR. They evaluate the results of each method and concluded that the SVM algorithm outperformed the others with a performance accuracy of (97.2%). Another study [16] tried to examine the findings and to analyze several ML approaches for the detecting of BC using the same dataset. Another study [17] proposed an efficient recursive neural network (RNN) approach for BC classification using RNN and "Keras–Tuner" enhancement method in which they claimed that, the developed model achieved high performance accuracy. However, the study in [16] showed that Logistic regression classifier beats the other classifiers in predicting BC disease using BC Wisconsin (Diagnostic) data set (BCWD).

III. Research Methods

The current study attempts to minimize the dimensionality of the dataset before selecting the optimal features to be fed to the classifier. The proposed method employs the simplest model that meets the performance requirements of the complicated models. Accordingly, the dataset shall be dimensionally reduced in order to improve model performance and eliminate extraneous features. Therefore, the proposed model begins with data preprocessing, followed by feature selection, dimensionality reduction, and classification (Fig. 1). In this methodology, four different supervised classification algorithms are used, RandomForest, AdaBoost, GradientBoost, and DecisionTree respectively. For this study, the classifier that obtained the best performance measure (Accurecy_score, and F1_score) is selected to perform classification process. However, after selecting the best classifier, the principle components analysis (PCA) procedure is taken place to perform dimensionality reduction. The reduced dataset is then fed to the chosen classifier to implement BC classification. The developed model is then validated using k-fold cross validation, tested using a subset of the original dataset (30%), and lastly, performance measures are evaluated.

A. Dataset Description

In this study, BCWD data is being utilized for model development purposes. The data include 31 features along with the class feature (target). The cell nuclei detected in the breast image clip are represented by the independent indices. Moreover, the dependent index contains binary outcome: zero indicates benign, and one describes malignant. The output will be classified as being benign or malignant. However, the shape of the used dataset is (569,31), and its descriptive analysis is shown in Fig. 2. Additionally, the bar chart (Fig. 3) depicts the count of the target variable to be malignant (M) or benign (B).

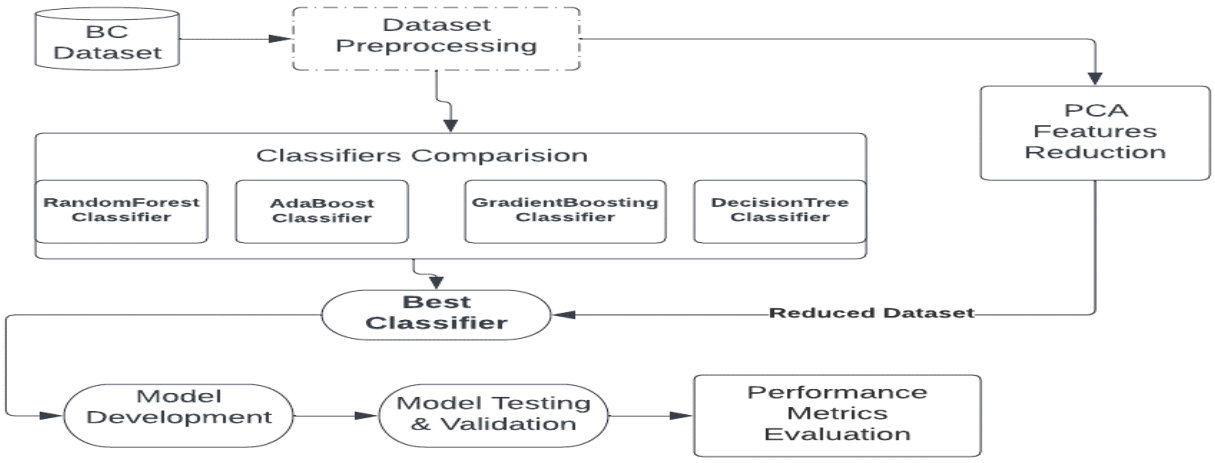

Fig. 1. An enhanced methodology for BC classification using PCA.

```
1 data.shape
```

```
(569, 31)
```

```
1 data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 569 entries, 0 to 568
Data columns (total 31 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   id                      569 non-null    int64
 1   radius_mean             569 non-null    float64
 2   texture_mean            569 non-null    float64
 3   perimeter_mean          569 non-null    float64
 4   area_mean               569 non-null    float64
 5   smoothness_mean         569 non-null    float64
 6   compactness_mean        569 non-null    float64
 7   concavity_mean          569 non-null    float64
```

```
1 table = tabulate.tabulate(data, tablefmt='html')
2 display(HTML(table))
```

0	842302	17.99	10.38	122.8	1001	0.1184	0.2776	0.3001
1	842517	20.57	17.77	132.9	1326	0.08474	0.07864	0.0869
2	8.43009e+07	19.69	21.25	130	1203	0.1096	0.1599	0.1974
3	8.43483e+07	11.42	20.38	77.58	386.1	0.1425	0.2839	0.2414
4	8.43584e+07	20.29	14.34	135.1	1297	0.1003	0.1328	0.198
5	843786	12.45	15.7	82.57	477.1	0.1278	0.17	0.1578
6	844359	18.25	19.98	119.6	1040	0.09463	0.109	0.1127
7	8.44582e+07	13.71	20.83	90.2	577.9	0.1189	0.1645	0.09366
8	844981	13	21.82	87.5	519.8	0.1273	0.1932	0.1859
9	8.4501e+07	12.46	24.04	83.97	475.9	0.1186	0.2396	0.2273
10	845636	16.02	23.24	102.7	797.8	0.08206	0.06669	0.03299

Fig. 2. Descriptive analysis of BC dataset.

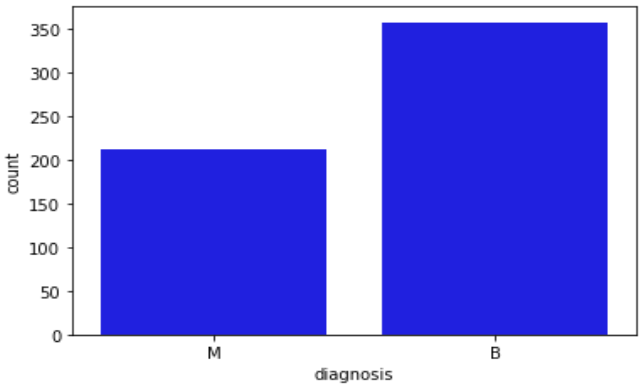

Fig. 3. Count for class variable (diagnosis).

B. Dimensionality Reduction

Features reduction represents a very important preprocessing stage that eliminates redundancy, inconsistent, and unimportant features to optimize learning, classification accuracy, and minimize training cost [18]. One of the its approaches that diminishes computation cost for the learning process is called "Principal Component Analysis (PCA)". Moreover, features reduction is useful in several realms because it reduces the computational burden as well as other unfavorable characteristics of high-dimensional areas. Many scholars recommended the utilization of features reduction techniques to improve computational power and to enhance performance accuracy [5]. Therefore, the literature has several uses of dimensionality reduction techniques such as, Zhao and Du [19] in which they advocated for the use of the "feature_based" spectral-spatial" classification (SSFC) structure. Another study by Xu. Y et. al [20] proposed spontaneous removal of piece picture from side to side deep learning. However, the PCA is commonly used method for features reduction.

C. Classification

Classification represents the most important task in supervised learning techniques [21-23]. It normally utilized to separate the dataset into a unique class as per the values in the dependent variable [24-25]. To select the best classifier that provides the optimal performance metrics based on BC dataset, four different classifiers are used, namely, RandomForest, Adaboost, Gradientboost, and DecisionTree are used. A brief detail-on each classifier is shown below:

1) Random forest classifier: Because of its ease of implementation and high versatility, it is one of the most often used supervised learning algorithms. It is a collection of prediction trees capable of handling large datasets.

2) Adaboost classifier: AdaBoost was among the first applications to employ the boosting technology. It accomplishes this by integrating numerous weak classifiers into a single strong classification method.

3) Gradientboost classifier: It is an ensemble, functional gradient iterative approach that reduces a "loss function" by repeatedly selecting a function who points towards the negative slope.

4) Decision tree classifier: The decision tree can be defined as a supervised learning technique in which it is commonly implemented for solving binary classification problems. However, such a technique bases its decision on some rules.

IV. Results

In this study, initially four different classifiers were employed to get performance metrics using BC dataset. The established approaches were assessed using accuracy score and F1 score metrics. However, the model with the best metrics was used to develop the enhanced model. After data preprocessing and visualization, the dataset is reduced into two main components namely, first principal component and second principal component. Fig. 4 illustrates the reduced PCA components.

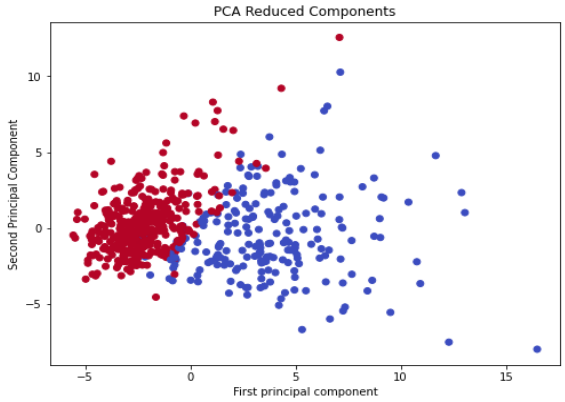

Fig. 4. The reduced PCA components (first & second principal components).

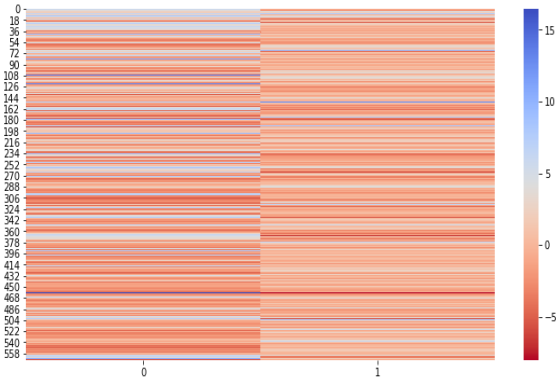

Fig. 5. Heatmap of the reduced PCA features.

As it can be seen in Fig. 4, the reduced dataset consists of two principal components in which they can represent the original dataset with no loss of information. It can be noticed that the reduced dataset can be used to clearly discriminate between the two classes of the target variable ('diagnosis'). The reduced dataset represents a dumpy matrix array in its rows represents the principal components and each column in it relates back to the original indices. Such relationship can be visualized as heatmap (Fig. 5).

The classifier that obtained the higher performance metrics is selected to implement the final model using the reduced PCA dataset. The RandomForest classifier obtained (95.7%) f1_score and (94.5%) accuracy_score in which it occupied position number two among the selected classifiers. The DecisionTree classifier obtained (93%) f1_score and (91%) accuracy_score. The GradientBoosting classifier obtained (95%) f1_score and (93.5%) accuracy_score. Eventually, the AdaBoost classifier obtained (95.8%) f1_score and (94.5%) accuracy_score (Table I).

TABLE I. CLASSIFIERS PERFORMANCE METRICS

Classifier	F1-Score	Accuracy-Score
Random Forest	95.7%	94.5%
Decision Tree	93%	91%
Ada-Boost	95.8%	94.5%
Gradient-Boosting	95%	93.5%

V. DISCUSSIONS

The obtained results showed that AdaBoost classifier outperforms other used classifiers in terms of f1_score and accuracy_score. Thereby, it was used to implement the final enhanced model. On other hand, the RandomForest classifier achieved a very similar performance metrics to the AdaBoost classifier in terms of accuracy_score, but the AdaBoost classifiers achieved higher F1_score. Hence, the reduced PCA components were fed again to AdaBoost classifier to validate the enhancement made by the reduced dataset features using PCA components. Fig. 6 showed the enhanced performance metrics in which it achieved an overall accuracy of (99%). The new enhanced model is used to make predictions on a new dataset to validate its performance. The developed classifier was able to correctly classify the new data into its correct classes where they were 'Malignant' or 'Benign'. The final accuracy_score and F1_score is depicted in Fig. 7. The model obtained a noticeable higher f1_score (99%) and noticeable higher accuracy_score (98.8%).

```
flower_type = {0:'Malignant', 1:'Benign'}

flower_index = 7

y_test_np = np.array(y_test)

print(f'Actual --> {flower_type[y_test_np[flower_index]]}  --  Prediction --> {flower_type[pred[flower_index]]}')
```

```
Actual --> Malignant  --  Prediction --> Malignant
```

```
print('ENhanced classification metrics using PCA and AdaBoost Classifier')
print('Accuracy_Score is: ')
print(accuracy_score(y_test,pred))
print('F1_Score is: ')
print(f1_score(y_test,pred))
```

```
ENhanced classification metrics using PCA and AdaBoost Classifier
Accuracy_Score is:
0.9883040935672515
F1_Score is:
0.9908256880733944
```

Fig. 6. The enhanced BC classification performance metrics.

```
print(confusion_matrix(y_test,pred))
```

```
[[ 61   0]
 [  2 108]]
```

```
print('AdaBoost Classification Using PCA')
print(classification_report(y_test,pred))
```

```
AdaBoost Classification Using PCA
              precision    recall  f1-score   support

           0       0.97      1.00      0.98        61
           1       1.00      0.98      0.99       110

    accuracy                           0.99       171
   macro avg       0.98      0.99      0.99       171
weighted avg       0.99      0.99      0.99       171
```

Fig. 7. The improved performance metrics using AdaBoost and PCA.

As shown in Fig. 7, the developed model obtained higher accuracy score and higher F1 score. Therefore, it can be can concluded that, the reduced dataset using PCA components analysis can enhance classification performance in high-dimensions datasets. Furthermore, dimensionality reduction simplifies the classification process in ML, resulting in a better fit to the constructed classifier.

VI. Conclusions

This research utilized PCA technique to minimize the input features in the BC dataset seeking better enhancement of BC classification in terms of F1_score and accuracy_score. The developed model started with a performance metrics comparison between four supervised classification techniques namely, RandomForest, DecisionTree, AdaBoost, and GradientBoosting. The RandomForest classifier showed (95.7%) f1_score and (94.5%) accuracy_score, DecisionTree classifier obtained (93%) f1_score and (91%) accuracy_score, GradientBoosting classifier obtained (95%) f1_score and (93.5%) accuracy_score, and finally, AdaBoost classifier obtained (95.8%) f1_score and (94.5%) accuracy_score. Since the AdaBoost classifier scored the highest performance metrics, it used to implement the final model using the reduced PCA dataset. The developed classifier is named "pcaAdaBoost". The optimized pcaAdaBoost achieved higher performance metrics in terms of F1_Score (99%) and accuracy_score (98.8%). The results show that the optimized pcaAdaBoost has delivered the best results in terms of cross-validation and testing. with an overall accuracy of (99%). However, as per future works, the developed classifier should be trained and tested using different datasets to validate its ability to enhance performance metrics. Finally, the developed model is hoped to introduce a predictive tool for early diagnosis and classification of BC in our large society.

VII. Data Availability

The used data in the development of this model and that is used to support the findings of this research can be accessed online at:

UCI Machine Learning Repository: Breast Cancer Wisconsin (Diagnostic) Data Set.

VIII. Conflicts of Interest

The authors stated that they do have no potential conflicts to disclose in relation to this study.

References

[1] R.L Siegel, K.D Miller, and A. Jemal, "Cancer Statistics, 2019," CA Cancer J. Clin, vol. 69, pp 7-34, 2019. doi: 10.3322/caac.21551.

[2] Y. Qawqzeh and K. Abdus Sattar, "Online Diagnostic Expert System for Detection of Breast Cancer in Saudi Arabia," International Journal of Computer Applications. 113. pp40-47, 2015. 10.5120/19833-1686.

[3] J.A Basurto-Hurtado, I.A Cruz-Albarran, M. Toledano-Ayala, M.A Ibarra-Manzano, L.A Morales-Hernandez, C.A Perez-Ramirez, "Diagnostic Strategies for Breast Cancer Detection: From Image Generation to Classification Strategies Using Artificial Intelligence Algorithms," Cancers 2022, 14, 3442. https://doi.org/10.3390/cancers14143442.

[4] X-D. Zhang, "Evolutionary computation. In: A matrix algebra approach to artificial intelligence," Springer, Singapore, pp 681–803, 2020.

[5] A.K.S. Ong et al." Utilization of random forest and deep learning neural network for predicting factors affecting perceived usability of a covid-19 contact tracing mobile application in Thailand," Int. J. Environ. Res. Public Health, 19, 6111, 2022, doi.org/10.3390/ijerph19106111.

[6] Y. Qawqzeh, M. T. Alharbi, A. Jaradat and K. N. Abdul Sattar, "A review of swarm intelligence algorithms deployment for scheduling and optimization in cloud computing environments," PeerJ Computer Science, vol. 7, pp. e696, 2021.

[7] D. Rajkumar, S. P. Raja, and A. Suruliandi, "Users' click and book-mark based personalization using modified agglomerative clustering for web search engine," International Journal on Artificial Intelligence Tools, 26, no. 06, 2017, doi: 10.1142/S0218213017300022.

[8] P. Divya, M. Pavithra, S. Jayalakshmi, P. Praveen kumar, "Application of Random Forest Algorithm in Bio Informatics," International Journal of Information Technology Insights & Transformations, Vol. 5, I 1, 2021, pp16-24.

[9] M. Alloghani, D. Al-Jumeily, A.J. Aljaaf, M. Khalaf, J. Mustafina, and S.Y. Tan, "The Application of Artificial Intelligence Technology in Healthcare: A Systematic Review," Commun. Comput. Inf. Sci. 2020, 1174, 248–261.

[10] M. Strzelecki and P. Badura, "Machine Learning for Biomedical Application," Applied Sciences, 2022; 12(4):2022. doi.org/10.3390/app12042022.

[11] Y. Zhang, Y. Weng, and J. Lund, "Applications of Explainable Artificial Intelligence in Diagnosis and Surgery," Diagnostics 2022, 12, 237. doi.org/10.3390/ diagnostics12020237.

[12] Y.K. Qawqzeh, M.M. Otoom, F. Al-Fayez, I. Almarashdeh, M. Alsmadi and G. Jaradat, "A Proposed Decision Tree Classifier for Atherosclerosis Prediction and Classification," IJCSNS, v 19, issue 12, 2019, pp197-202.

[13] WHO, "Preventing cancer," Accessed on: Oct. 18, 2022. [online] Available: https://www.who.int/activities/preventing-cancer.

[14] I. Maglogiannis, E. Zafiropoulos, and I. Anagnostopoulos, "An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers," Applied Intelligence, vol. 30, pp. 24–36, 2009.

[15] M. Naji and S. El Filali, "Machine learning algorithms for breast cancer prediction and diagnosis," Procedia computer science, 191(2021), pp487-492.

[16] M. Khan et al., "Machine Learning Based Comparative Analysis for Breast Cancer Prediction," Journal of Healthcare Engineering, Hindawi. V2022, pp1-14. doi.org/10.1155/2022/4365855.

[17] H. Saleh, S.F. Abd-el ghany, H. Alyami, and W. Alosaimi, "Predicting Breast Cancer Based on Optimized Deep Learning Approach," Hindawi, Computational Intelligence and Neuroscience, Vol 2022, pp1-11. https://doi.org/10.1155/2022/1820777.

[18] S. Velliangiri, S. Alagumuthukrishnan, S.I.T. Joseph, " A Review of Dimensionality Reduction Techniques for Efficient Computation," Procedia Computer Science, V165, pp104-111. ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.01.079.

[19] W. Zhao and S. Du, "Spectral-Spatial Feature Extraction for Hyperspectral Image Classification: A Dimension Reduction and Deep Learning Approach," IEEE Trans. Geosci. Remote Sens. 54(8): 2016, 4544–4554.

[20] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I. Chang, "Deep Learning of Feature Representation with Multiple Instance Learning for Medical Image Analysis,". State Key Laboratory of Software Development Environment, Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education, Beihang University M," Icassp 1: 2014, 1645–1649.

[21] I.H. Sarker, "Machine Learning: Algorithms,". Real-World Applications and Research Directions. SN COMPUT. SCI. 2, 160, 2021. doi.org/10.1007/s42979-021-00592-x.

[22] N. Binsaif, "Application of Machine Learning Models to the Detection of Breast Cancer," Mobile Information Systems, Hindawi. V 2022, https://doi.org/10.1155/2022/7340689.

[23] M. Shanbehzadeh, H. Kazemi-Arpanahi, M.B Ghalibaf, A. Orooji, "Performance evaluation of machine learning for breast cancer diagnosis: A case study," Informatics in Medicine Unlocked, V 31, 2022, 101009, https://doi.org/10.1016/j.imu.2022.101009.

[24] Butt, Umair Muneer et al. "Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications," Journal of healthcare engineering vol. 2021 9930985. 29 Sep. 2021, doi:10.1155/2021/9930985.

[25] M.M Khan, T. Tazin, M.Z Hussain, M. Mostakim, T. Rehman, S. Singh, V. Gupta, O. Alomeir, "Breast Tumor Detection Using Robust and Efficient Machine Learning and Convolutional Neural Network Approaches," Exploration of Human Cognition using Artificial Intelligence in Healthcare. V 2022, https://doi.org/10.1155/2022/6333573.

AUTHORS' PROFILE

YOUSEF QAWQZEH received the PhD. degree in systems engineering from UKM University, Bangi, Kuala Lumpur, Malaysia, in 2011, where he is currently working as an associate professor in the college of information technology, Fujairah University. He is currently working on several projects in the fields of machine learning, data science, and bioinformatics. He has several publications in international journal and conferences. His research interest includes the early prediction of cardiovascular diseases using the photoplethysmography technique, the development of computer-aided diagnosis systems for early diagnosis of breast cancer using artificial intelligence and machine learning techniques, and the detection and prediction of high-risk diabetics using machine learning and artificial intelligence techniques.

ABDULLAH ALOURANI is an assistant professor at the Department of Computer Science and Information, Majmaah University, Saudi Arabia. He received his Ph.D. in computer science from the University of Illinois at Chicago, United States, his Master's degree in computer science from DePaul University in Chicago, United States, and his Bachelor's degree in computer science from Qassim University, Saudi Arabia. His current research interests are in the areas of software engineering, security, and artificial intelligence. He is a member of ACM and IEEE.

SAMEH GHWANMEH a full Professor of Computer Science and Engineering, and ICT. Obtained his PhD and MS in Computer Engineering from UK, in 1996 and 1993 respectively, and BS degree in Computer Engineering from Jordan, in 1985. More than 24 years of leadership experience in HE institutions. Was and still involved in many administrative and academic positions: University Chancellor (current), Registrar Director, Dean of Faculty of Information Technology, Director of Accreditation, Associate Dean, Program Chair, Academic Adviser to the Minister of Education and Higher Education in Jordan. The published work exceeds 50 international journal papers, and was an invited speaker in different international conferences and workshops. Received many national and international awards. Was involved in curriculum development for a number of undergraduate and graduate programs. Managed and coordinated several research projects that are funded by regional and international agencies such as European Commission.