Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH
AND DEVELOPMENT

| RESEARCH ARTICLE

Transforming Breast Cancer Identification: An In-Depth Examination of Advanced Machine Learning Models Applied to Histopathological Images

Rejon Kumar Ray[1], Ahmed Ali Linkon[2], Mohammad Shafiquzzaman Bhuiyan[3], Rasel Mahmud Jewel[4], Nishat Anjum[5]✉, Bishnu Padh Ghosh[6], Md Tuhin Mia[7], Badruddowza[8], Md Shohail Uddin Sarker[9], Mujiba Shaima[10]

[1]Department of Business Analytics Business Analytics, Gannon University, USA
[2]Department of Computer Science, Westcliff University, Irvine, California
[3]Department of Business Administration, Westcliff University, Irvine, California, USA
[4]Department of Business Administration, Westcliff University, Irvine, California, USA
[5]Department of Computer Science, University of South Dakota, Vermillion, South Dakota, USA
[6]School of Business, International American University Angeles, CA, USA
[7]School of Business, International American University Angeles, CA, USA
[8]Department of Computer & Info Science, Gannon University, Erie, Pennsylvania, USA
[9]Department of Computer & Info Science Gannon University, Erie, Pennsylvania, USA
[10]Department of Computer Science, Monroe College, New Rochelle, New York, USA

Corresponding Author: Nishat Anjum, **E-mail**: nishat.anjum@coyotes.usd.edu

| ABSTRACT

Breast cancer stands as one of the most prevalent and perilous forms of cancer affecting both women and men. The detection and treatment of breast cancer benefit significantly from histopathological images, which carry crucial phenotypic information. To enhance accuracy in breast cancer detection, Deep Neural Networks (DNNs) are commonly utilized. Our research delves into the analysis of pre-trained deep transfer learning models, including ResNet50, ResNet101, VGG16, and VGG19, for identifying breast cancer using a dataset comprising 2453 histopathology images. The dataset categorizes images into two groups: those featuring invasive ductal carcinoma (IDC) and those without IDC. Through our analysis of transfer learning models, we observed that ResNet50 outperformed the other models, achieving impressive metrics such as accuracy rates of 92.2%, Area under Curve (AUC) rates of 91.0%, recall rates of 95.7%, and a minimal loss of 3.5%.

| KEYWORDS

Breast Cancer Identification; Machine Learning Models; Histopathological Images

| ARTICLE INFORMATION

ACCEPTED: 04 January 2024 **PUBLISHED:** 28 January 2024 **DOI:** 10.32996/jcsts.2024.6.1.16

1. Introduction

Breast cancer poses a significant health risk affecting both women and men, making it one of the most prevalent and life-threatening forms of cancer. The critical role of histopathological images in the detection and treatment of breast cancer, providing essential phenotypic information, has prompted the widespread utilization of Deep Neural Networks (DNNs) [1,2]. This study delves into a comprehensive analysis of pre-trained deep transfer learning models, specifically ResNet50, ResNet101, VGG16, and VGG19, to identify breast cancer using a dataset comprising 2453 histopathology images. The dataset categorizes images into two groups: those featuring invasive ductal carcinoma (IDC) and those without IDC.

The research methodology involves the systematic acquisition of histopathology images from accessible sources, followed by a meticulous pre-processing stage. Subsequently, the study employs the conventional hold-out validation approach, training, testing, and validating the pre-trained transfer learning models on the comprehensive histopathology images dataset. The findings of the evaluation, presented in Figure 3, showcase the superior performance of ResNet50, outperforming other models with remarkable accuracy rates of 92.2%, an Area under Curve (AUC) of 91.0%, recall rates of 95.7%, and a minimal loss of 3.5%.

The importance of early detection and effective treatment of breast cancer is emphasized through statistics provided by the American Cancer Society (ACS), estimating over 287,850 new cases and a projected death toll of 43,780 in the USA for the year 2022. Breast tissue is categorized into four distinct forms, each carrying different health implications. Various methods, including X-ray mammography, temperature measurement, and advanced imaging techniques, contribute to the detection of breast cancer. The study also explores the role of genomics and histopathology, with a focus on the valuable insights offered by histopathology images in the early stages of cancer treatment.

The dataset collection and pre-processing stages involve utilizing the Breast Histopathology Images dataset from Kaggle, comprising 162 whole-mount slide photos of breast cancer samples. The images are categorized into IDC positive and IDC negative, with ResNet50 identified as the most optimal transfer learning model for breast cancer detection.

The research methodology, results, and discussions underscore the significance of transfer learning models in the realm of medical imaging and diagnosis. The findings contribute valuable insights into the efficacy of different models, with ResNet50 standing out as a promising candidate for accurate and early breast cancer detection.

Several methods are employed for breast cancer detection, including X-ray mammography, breast temperature measurement, Portion Emission Tomography (PET), Computed Tomography (CT), ultrasound (US), and Magnetic Resonance Imaging (MRI). The diagnosis encompasses two approaches: genomics and histopathology. Histopathology images, microscopic representations of breast tissue, prove invaluable in the early stages of cancer treatment. A burgeoning field known as radio-genomics explores multi-scale correlations between gene expression data and medical imaging, providing new insights into understanding and managing breast cancer [4,5].

2. Related Work

Khan et al (2023) In this research, the focus is on developing a sustainable machine-learning model for the identification of benign or malignant breast cancer types. The goal is to propose an efficient model for outbreak detection. The study evaluates the performance of five machine learning algorithms, namely XGBoost, Naïve Bayes, Decision Tree, Random Forest, and Logistic Regression, in predicting human health behavior. XGBoost emerges with the highest accuracy at 95.42% and demonstrates strong performance in sensitivity (98.5%), specificity (97.5%), and F-1 score (99%). The results indicate that XGBoost holds significant potential in breast cancer prediction. However, further research is necessary to refine and apply it for potential commercial use in the healthcare industry. In our research, the focus is on developing a sustainable machine-learning model for the identification of benign or malignant breast cancer types. The goal is to propose an efficient model for outbreak detection. The study evaluates the performance of five machine learning algorithms, namely XGBoost, Naïve Bayes, Decision Tree, Random Forest, and Logistic Regression, in predicting human health behavior. XGBoost emerges with the highest accuracy at 95.42% and demonstrates strong performance in sensitivity (98.5%), specificity (97.5%), and F-1 score (99%). The results indicate that XGBoost holds significant potential in breast cancer prediction.

Wolberg et al (1990) This paper utilizes breast cytology diagnosis as a case study to showcase the application of the method in medical diagnosis and decision-making. During sample collection, 11 cytological characteristics of breast fine-needle aspirates, known to differ between benign and malignant samples, were graded on a scale of 1 to 10. Out of these, nine characteristics exhibited significant differences between benign and malignant samples. Mathematically, these characteristic values for each sample were represented as points in a nine-dimensional space of real variables. Linear programming was employed to establish planes that effectively separated benign and malignant points. Accurate separation was achieved in 369 out of 370 samples, with one misclassified malignant case likely due to the definite benign nature of the fine-needle aspirate cytology and the clear malignancy in the excised cancer, leading us to believe that the tumor was overlooked during aspiration. The mathematical approach presented here is transferable to other medical diagnostic and decision-making scenarios.

Medjahed et al (2013) The diagnosis of this cancer poses a significant challenge in cancer research. Within the field of artificial intelligence, machine learning serves as a discipline that enables machines to evolve through a learning process. In bioinformatics, and more specifically in breast cancer diagnosis, machine learning is extensively applied. One widely used supervised learning method in this context is K-nearest neighbors (K-NN). The application of K-NN in medical diagnosis is particularly intriguing. The effectiveness of the results is heavily influenced by the choice of distance metrics and the parameter "k," representing the number of nearest neighbors. This study delves into the assessment and evaluation of various distances applicable to the K-NN algorithm. Additionally, the analysis involves experimenting with different values of the "k" parameter and employing various classification rules to determine the classification of a given sample.

Mirsadeghi et al (2021) This research aims to highlight key findings related to the prognosis and diagnosis of MBCA in various aspects. Initially, SVM, ANN, RF, and EARN models are employed to predict drivers and passengers. Subsequently, biological insights into these predictions are discussed, leveraging gene set enrichment analysis. The study further involves statistical validation and comparison of all learning methods through evaluation metrics. The application of the ReactomeFIVIz tool in pathway enrichment analysis (FDR < 0.03) for the top 100 genes predicted by EARN leads to the proposal of a novel gene set panel for MBCA. This panel includes genes such as HDAC3, ABAT, GRIN1, PLCB1, KPNA2, NCOR1, TBL1XR1, SIRT4, KRAS, CACNA1E, PRKCG, GPS2, SIN3A, ACTB, KDM6B, and PRMT1. Additionally, a comparative analysis is conducted with results from 983 primary tumor samples of breast invasive carcinoma (BRCA) sourced from the Cancer Genome Atlas (TCGA). The comparison reveals that using EARN, the ROC-AUC reaches 99.24% for MBCA and 99.79% for BRCA, outperforming three individual classifiers in each case.

Naji er al (2021) This research employs five distinct machine learning algorithms, namely Support Vector Machine (SVM), Random Forest, Logistic Regression, Decision tree (C4.5), and K-Nearest Neighbours (KNN), on the Breast Cancer Wisconsin Diagnostic dataset. Following the generation of results, the study conducts a performance evaluation and comparison among these diverse classifiers. The primary aim is to predict and diagnose breast cancer utilizing machine learning algorithms, with a focus on assessing effectiveness through measures such as the confusion matrix, accuracy, and precision. The findings indicate that Support Vector Machine surpasses all other classifiers, achieving the highest accuracy at 97.2%. The entire process is conducted within the Anaconda environment, utilizing the Python programming language and the Scikit-learn library.

3. Methodology

The methodology is initiated by acquiring histopathology images from readily available sources and progressing to the subsequent pre-processing stage. Subsequently, employing the conventional hold-out validation approach, we engage in the training, testing, and validation of pre-trained transfer learning models, including ResNet-50, ResNet-101, VGG-16, and VGG-19 [9,10,11] utilizing the comprehensive histopathology images dataset. The systematic analysis and processing of the obtained results enable the discernment of the most optimal transfer learning model for precisely detecting breast cancer. This structured approach ensures a thorough evaluation of various models, leading to the identification of the model that exhibits superior performance in the context of breast cancer detection, thus contributing valuable insights to the field of medical imaging and diagnosis.

3.1 Dataset Collection and Preprocessing

The breast cancer dataset utilized in this research, known as Breast Histopathology Images, is sourced from the publicly accessible platform Kaggle [16]. The initial dataset comprises 162 whole-mount slide photos of breast cancer samples scanned at 40x, as reported by the dataset source [17]. These images were subsequently extracted and categorized into two groups: Invasive Ductal Carcinoma (IDC) positive and IDC negative. Notably, IDC represents the most prevalent subtype of breast cancer. For our study, a total of 2453 histopathology images were employed. The differentiation between non-cancerous and cancerous breast histopathology images is visually represented in Figure 1 and Figure 2, respectively. This comprehensive dataset forms the foundation for our investigation into the performance of various transfer learning models in the detection of breast cancer, offering valuable insights into the realm of medical imaging research.

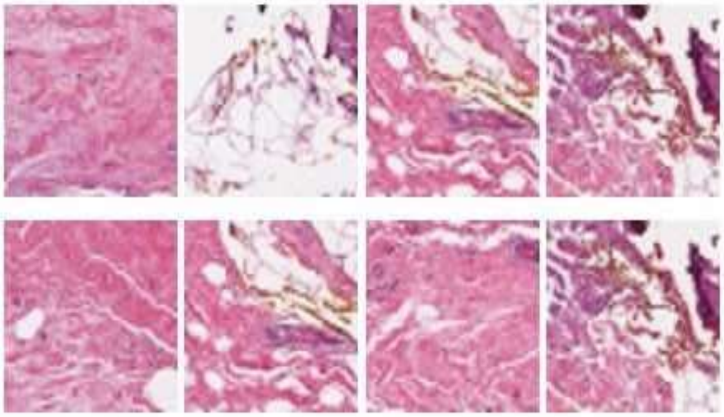

Figure 3: displays histopathology images of breasts without cancer.

The pre-processing stage is integral to preparing the data for effective training purposes. In the context of histopathology images, this phase primarily focuses on feature extraction through procedures such as image reading, resizing, de-noising, segmentation, and smoothing edges. A common pre-processing step for the Breast Cancer (BC) histopathology image dataset involves accurately delineating Invasive Ductal Carcinoma (IDC) areas within whole mount slides [18,19]. This meticulous pre-processing is crucial as it significantly enhances the accuracy achieved by a transfer learning model employed for image classification or detection. By optimizing the data through these pre-processing techniques, we ensure that the subsequent application of transfer learning models yields superior results, contributing to the overall efficacy of the breast cancer detection process.

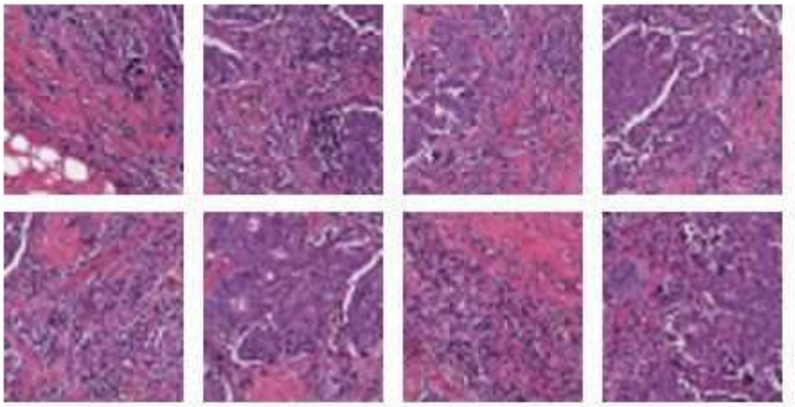

Figure 4: illustrates histopathology images depicting breasts affected by cancer.

3.2 Data Validation

The validation process holds substantial importance in the overall study, as the selection of an appropriate validation procedure tailored to a specific dataset is crucial. Opting for hold-out validation techniques, which yielded optimal outcomes [20], we partitioned our data into three segments: training (80%), testing (10%), and validation (10%). This division was achieved through the hold-out process, ensuring a robust evaluation methodology for the effectiveness of our chosen models in the context of breast cancer detection.

3.3 Transfer Learning Approach

In the realm of image classification or recognition within machine learning, transfer learning stands out as a widely employed technique. This study utilizes four distinct pre-trained transfer learning models, namely ResNet50, ResNet-101, VGG16, and VGG19 [21,22], for breast cancer detection. Below is a concise overview of these transfer learning models:

ResNet50 and ResNet101: ResNet, derived from residual networks [19], is designed to address the challenge of training deeper neural networks. It achieves this by incorporating shortcut connections, allowing the skipping of entire blocks of convolutional layers. This innovative approach mitigates performance deterioration and reduces errors associated with adding more layers. In the creation of a residual block, a shortcut link is introduced for each pair of three-by-three filters. The introduction of residual learning has significantly advanced image recognition and classification by enhancing accuracy and tackling complex tasks. ResNet50 represents a 50-layered residual network, while ResNet101 and ResNet152 present further variations in the ResNet framework [23,24]. These models contribute to overcoming challenges in training deeper networks and addressing issues of deterioration and disappearing gradients.

VGG16 and VGG19, integral components of the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), originated as a proposal from the Oxford Visual Geometry Group. These models have demonstrated their prowess by achieving top rankings in the ImageNet classification challenge, an extensive dataset comprising over 14 million images categorized into 1000 classes. The VGG network, serving as a feature extractor pre-trained on the ImageNet dataset, exhibits remarkable proficiency in extrapolating images, making it a versatile tool applicable to various tasks such as image classification and detection [28]. The VGG architecture encompasses two variants: VGG16 and VGG19. VGG16 comprises 16 layers, with the initial 13 being convolutional, followed by 3 fully connected layers, 4 max-pooling layers for volume size reduction, and a final fully connected layer with a softmax activation function. On the other hand, VGG19 is an extended version with 19 layers, including 16 convolutional neural networks, 3 fully connected layers, 5 Maxol layers, and 1 SoftMax layer . The distinct characteristics of VGG16 and VGG19, coupled with their exceptional performance in image recognition challenges, underscore their significance in the domain of deep learning applications.

The utilization of these pre-trained transfer learning models exemplifies the study's commitment to leveraging state-of-the-art techniques in machine learning for the specific task of breast cancer detection.

4. Result and Discussion

The evaluation of ResNet50, ResNet101, VGG16, and VGG19 models aimed to determine the most effective model for breast cancer detection based on histopathology images. The performance metrics, including Accuracy, Area Under Curve (AUC), Recall, and Loss, are presented in chart 1.

Table 1: Model Performance among Different Deep Learning Models

Model	Accuracy %	AUC %	Recall %	Loss %
ResNet50	92.20	91.00	95.70	3.5
ResNet101	90.01	88.20	90.20	5.3
VGG16	89.10	90.10	88.40	7.2
VGG19	84.3	85.0	92.0	7.3

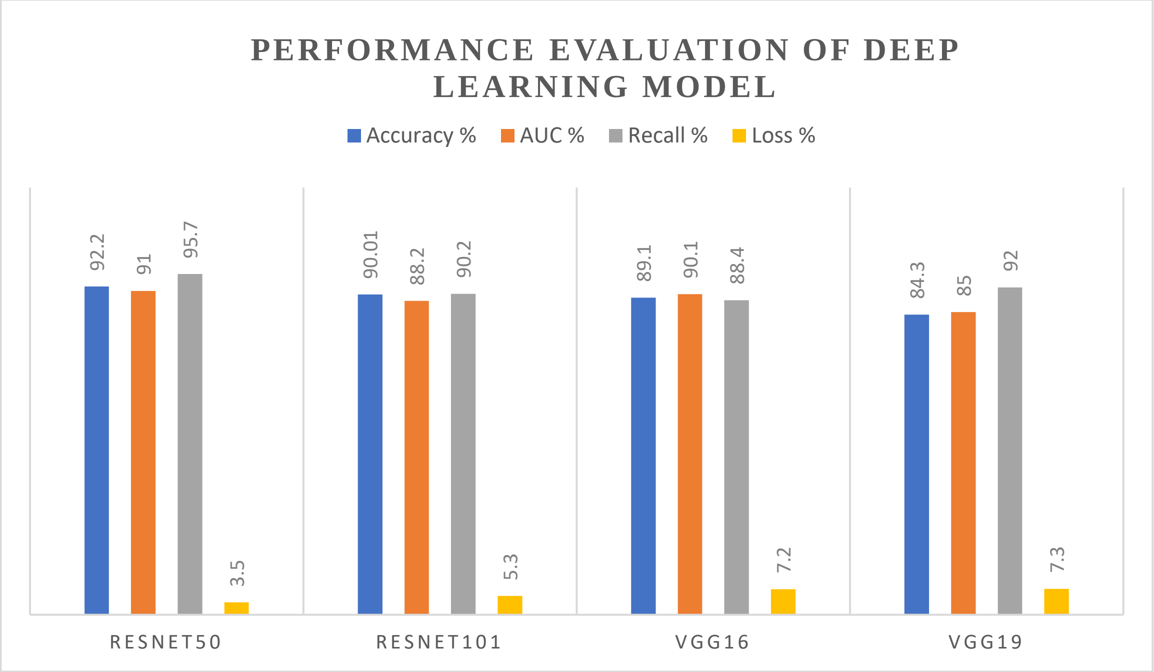

Chart 1: Model Performance among Different Deep Learning Models

In considering four key performance metrics—accuracy, AUC, recall, and loss—we sought to identify the optimal pre-trained transfer learning model among the four. While accuracy is a significant metric, it alone may not provide a comprehensive assessment of model performance. Therefore, the inclusion of AUC, recall, and loss values becomes crucial in understanding the overall efficacy of the models. A higher AUC value signifies superior model performance in distinguishing between positive and negative categories.

ResNet50 emerged as the top-performing transfer learning model across all evaluated metrics, showcasing not only high accuracy but also exceptional AUC, recall, and a minimal loss of 3.5%. The loss function serves as a vital indicator, offering insights into variance from training and validation values. Lower loss values indicate improved model performance. ResNet50 outperformed competitors like ResNet101, VGG16, and VGG19, with lower loss percentages (5.3% compared to 7.2%, and 7.3%, respectively).

Although ResNet101 and VGG16 demonstrated similar performance in terms of accuracy, AUC, and recall, ResNet101 had a slightly higher loss (5.3% compared to 7.2%). VGG19 exhibited the weakest overall performance among the four pre-trained transfer learning models, with 84.3% accuracy, 85.0% AUC, 92.0% recall, and a 7.3% loss. These findings underscore the superiority of ResNet50 in effectively identifying breast cancer based on comprehensive performance metrics.

5. Conclusion and Future Work

In conclusion, this study has delved into the realm of breast cancer detection utilizing advanced deep-transfer learning models applied to histopathology images. Breast cancer, a prevalent and life-threatening form of cancer, affects both men and women, underscoring the need for effective detection methods. Histopathological images have proven crucial in this regard, providing phenotypic information that aids in accurate diagnosis and treatment.

The investigation incorporated pre-trained transfer learning models, including ResNet50, ResNet101, VGG16, and VGG19, analyzing their performance on a dataset comprising 2453 histopathology images categorized into invasive ductal carcinoma (IDC) positive and IDC negative. The findings revealed that ResNet50 outperformed other models, demonstrating superior accuracy (92.2%), Area under Curve (AUC) (91.0%), recall (95.7%), and minimal loss (3.5%). This highlights the efficacy of ResNet50 in enhancing breast cancer detection accuracy.

The discussion emphasizes the significance of early detection, as evidenced by alarming statistics from the American Cancer Society (ACS), estimating over 287,850 new cases of breast cancer in 2022 with a projected death toll of 43,780 in the USA. The World Health Organization's (WHO) 2020 report on cancer further provides insights into the global impact of various cancer types. The study employed a comprehensive dataset sourced from Kaggle, comprising 162 whole-mount slide photos initially, and later expanded to 2453 histopathology images. The dataset's categorization into IDC positive and IDC negative facilitated a detailed analysis of transfer learning models' performance.

The methodology involved meticulous data preprocessing, including image reading, resizing, de-noising, segmentation, and edge smoothing. The validation process utilized hold-out validation techniques, crucial for robust evaluation. The adoption of state-of-the-art transfer learning models demonstrates the study's commitment to leveraging advanced techniques for breast cancer detection.

Model performance evaluation showcased ResNet50 as the optimal choice, surpassing ResNet101, VGG16, and VGG19 in accuracy, AUC, recall, and loss. The results emphasize the importance of considering multiple metrics for a comprehensive assessment of model performance.

In charting the future direction, the study advocates for continued exploration of advanced technologies, such as deep transfer learning models, in the field of medical imaging for enhanced breast cancer detection. The findings contribute valuable insights to the ongoing efforts to improve diagnostic capabilities, ultimately impacting patient outcomes and mortality rates in the battle against breast cancer.

References

[1] J. Miah, R. H. Khan, S. Ahmed and M. I. Mahmud, "A comparative study of Detecting Covid 19 by Using Chest X-ray Images– A Deep Learning Approach," 2023 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2023, pp. 0311-0316, Doi: 10.1109/AIIoT58121.2023.10174382.

[2] Haque, M. S. (2023). Retail Demand Forecasting Using Neural Networks and Macroeconomic Variables. Journal of Mathematics and Statistics Studies, 4(3), 01–06. https://doi.org/10.32996/jmss.2023.4.3.1

[3] R. H. Khan, J. Miah, M. M. Rahman and M. Tayaba, "A Comparative Study of Machine Learning Algorithms for Detecting Breast Cancer," 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2023, pp. 647-652, doi: 10.1109/CCWC57344.2023.10099106.

[4] Khan, R. H., Miah, J., Arafat, S. M. Y., Syeed, M. M. M., & Ca, D. M. (2023). Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks. arXiv preprint, arXiv:2310.17729.

[5] R. H. Khan, J. Miah, M. M. Rahman, and M. Tayaba, "A Comparative Study of Machine Learning Algorithms for Detecting Breast Cancer," 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2023, pp. 647-652, doi: 10.1109/CCWC57344.2023.10099106.

[6] Wolberg, W. H., & Mangasarian, O. L. (1990). Multisurface method of pattern separation for medical diagnosis applied to breast cytology. *Proceedings of the national academy of sciences*, *87*(23), 9193-9196.

[7] Medjahed, S. A., Saadi, T. A., & Benyettou, A. (2013). Breast cancer diagnosis by using k-nearest neighbor with different distances and classification rules. *International Journal of Computer Applications*, *62*(1).

[8] Mirsadeghi, L., Haji Hosseini, R., Banaei-Moghaddam, A. M., & Kavousi, K. (2021). EARN is an ensemble machine learning algorithm to predict driver genes in metastatic breast cancer. *BMC Medical Genomics*, *14*(1), 122.

[9] Naji, M. A., El Filali, S., Aarika, K., Benlahmar, E. H., Abdelouhahid, R. A., & Debauche, O. (2021). Machine learning algorithms for breast cancer prediction and diagnosis. *Procedia Computer Science*, *191*, 487-492.

[10] Miah, J., Cao, D. M., Sayed, M. A., & Haque, M. S. (2023). Generative AI Model for Artistic Style Transfer Using Convolutional Neural Networks. arXiv preprint arXiv:2310.18237

[11] Khan, R. H., Miah, J., Arafat, S. M., Syeed, M. M., & Ca, D. M. (2023). Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks. arXiv preprint arXiv:2310.17729.

[12] Haque, M. S., Amin, M. S., & Miah, J. (2023). Retail Demand Forecasting: A Comparative Study for Multivariate Time Series. *arXiv e-prints*. https://doi.org/10.48550/arXiv.2308.11939

[13] Miah, J., Cao, D. M., Sayed, M. A., Taluckder, M. S., Haque, M. S., & Mahmud, F. (2023). Advancing Brain Tumor Detection: A Thorough Investigation of CNNs, Clustering, and SoftMax Classification in the Analysis of MRI Images. *arXiv preprint arXiv:2310.17720*.

[14] Syeed, M. M., Khan, R. H., & Miah, J. (2021). Agile Fitness of Software Companies in Bangladesh: An Empirical Investigation. *International Journal of Advanced Computer Science and Applications*, *12*(2).

[15] Miah, J., & Khan, R. H. (2019, November). Service Development of Smart Home Automation System: A Formal Method Approach. In *Proceedings of the 2019 2nd International Conference on Computational Intelligence and Intelligent Systems* (pp. 161-167).

[16] R. H. Khan, J. Miah, M. A. R. Rahat, A. H. Ahmed, M. A. Shahriyar and E. R. Lipu, "A Comparative Analysis of Machine Learning Approaches for Chronic Kidney Disease Detection," 2023 8th International Conference on Electrical, Electronics and Information Engineering (ICEEIE), Malang City, Indonesia, 2023, pp. 1-6, doi: 10.1109/ICEEIE59078.2023.10334765.

[17] R. H. Khan and J. Miah, "Performance Evaluation of a new one-time password (OTP) scheme using stochastic Petri net (SPN)," 2022 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2022, pp. 407-412, doi: 10.1109/AIIoT54504.2022.9817203.

[18] Khan, R. H., Miah, J., Arafat, S. M., Syeed, M. M., & Ca, D. M. (2023). Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks. *arXiv preprint arXiv:2310.17729*.

[19] R. H. Khan, J. Miah, S. M. Y. Arafat, M. M. M. Syeed and D. M. Ca, "Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks," 2023 15th International Conference on Innovations in Information Technology (IIT), Al Ain, United Arab Emirates, 2023, pp. 104-109, doi: 10.1109/IIT59782.2023.10366426.

[20] Md Abdur Rakib Rahat, MD Tanvir Islam, Duc M Cao, Maliha Tayaba, Bishnu Padh Ghosh, Eftekhar Hossain Ayon, Nur Nob, Taslima Akter, Mamunur Rahman, & Mohammad Shafiquzzaman Bhuiyan. (2024). Comparing Machine Learning Techniques for Detecting Chronic Kidney Disease in Early Stage. *Journal of Computer Science and Technology Studies*, *6*(1), 20–32. https://doi.org/10.32996/jcsts.2024.6.1.3

[21] MD Tanvir Islam, Eftekhar Hossain Ayon, Bishnu Padh Ghosh, MD, Salim Chowdhury, Rumana Shahid, Aisharyja Roy puja, Sanjida Rahman, Taslima Akter, Mamunur Rahman, Mohammad Shafiquzzaman Bhuiyan, & Tuan Ngoc Nguyen. (2024). Revolutionizing Retail: A Hybrid Machine Learning Approach for Precision Demand Forecasting and Strategic Decision-making in Global Commerce. *Journal of Computer Science and Technology Studies*, *6*(1), 33–39. https://doi.org/10.32996/jcsts.2024.6.1.4

[22] Duc M Cao, Md Abu Sayed, Md Abu Sayed, Md Tuhin Mia, Eftekhar Hossain Ayon, Bishnu Padh Ghosh, Rejon Kumar Ray, Aqib Raihan, Taslima Akter, & Mamunur Rahman. (2024). Advanced Cybercrime Detection: A Comprehensive Study on Supervised and Unsupervised Machine Learning Approaches Using Real-world Datasets. *Journal of Computer Science and Technology Studies*, *6*(1), 40–48. https://doi.org/10.32996/jcsts.2024.6.1.5

[23] Ghosh, B. P., Bhuiyan, M. S., Das, D., Nguyen, T. N., Jewel, R. M., Mia, M. T., & Cao, D. M. (2024). Deep Learning in Stock Market Forecasting: Comparative Analysis of Neural Network Architectures Across NSE and NYSE. *Journal of Computer Science and Technology Studies*, *6*(1), 68-75.

[24] Amin, M. S., Ayon, E. H., Ghosh, B. P., MD, M. S. C., Bhuiyan, M. S., Jewel, R. M., & Linkon, A. A. (2024). Harmonizing Macro-Financial Factors and Twitter Sentiment Analysis in Forecasting Stock Market Trends. *Journal of Computer Science and Technology Studies*, *6*(1), 58-67.

[25] Mia, M. T., Ray, R. K., Ghosh, B. P., Chowdhury, M. S., Al-Imran, M., Das, R., ... & Puja, A. R. (2023). Dominance of External Features in Stock Price Prediction in a Predictable Macroeconomic Environment. *Journal of Business and Management Studies*, *5*(6), 128-133.

[26] Shoukat, S., & Rabby, S. (2017). A Total Journey of Footwear With Material Analysis. Journal of Scientificand Engineering Research, 4(6), 187–191

[27] Shoukat, S. (2016). Now and then: The neckline history of women. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 26(2), 33-52.

[28] Shoukat, S., & Islam, M. A. (2016). Development of Wedding Costume from Bengal to Bangladesh.

[29] S. Islam, M. N. Islam Khan, M. Zannatun Ferdus, S. J. Islam and M. Abul Kashem, "Improving Throughput using Cooperating TDMA Scheduling of Wireless Sensor Networks," 2020 International Conference on Computing and Information Technology (ICCIT-1441), Tabuk, Saudi Arabia, 2020, pp. 1-4, doi: 10.1109/ICCIT-144147971.2020.9213713.

[30] S. J. Islam, S. Islam, M. Z. Ferdus, M. N. Islam Khan, M. A. Kashem and M. S. Islam, "Load Compactness and Recognizing Area Aware Cluster Head Selection of Wireless Sensor Networks," 2020 International Conference on Computing and Information Technology (ICCIT-1441), Tabuk, Saudi Arabia, 2020, pp. 1-4, doi: 10.1109/ICCIT-144147971.2020.9213750.

[31] M. Z. Ferdus, S. Islam and M. A. Kashem, "An Innovative Load Balancing Cluster Composition of Wireless Sensor Networks," 2019 Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2019, pp. 1-4, doi: 10.1109/GCAT47503.2019.8978385.

[32] M. Z. Ferdus, M. Nurul Islam Khan, S. Islam and M. A. Kashem, "VFLT: SQA Model for Cyber Physical System," 2019 Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2019, pp. 1-4, doi: 10.1109/GCAT47503.2019.8978289.

[33] Miah, Jonayet and Ayon, Eftekhar Hossain Ayon and Ghosh, Bishnu Padh Ghosh and Mia, Md Tuhin and ., Badruddowza and Sarker, Md Shohail Uddin Sarker and Islam, MD Tanvir, Enhancing Viral DNA Sequence Classification Using Hybrid Deep Learning Models and Genetic Algorithm Optimization (January 11, 2024). Available at SSRN: https://ssrn.com/abstract=4692259

[34] Islam, M.M. *et al.* (2020). An Empirical Study to Predict Myocardial Infarction Using K-Means and Hierarchical Clustering. In: Bhattacharjee, A., Borgohain, S., Soni, B., Verma, G., Gao, XZ. (eds) Machine Learning, Image Processing, Network Security and Data Sciences. MIND 2020. Communications in Computer and Information Science, vol 1241. Springer, Singapore. https://doi.org/10.1007/978-981-15-6318-8_11