

Ribonucleic Acid (RNA) Cancer Clustering Using Directional Local-Binary-Pattern and Self-Organizing Map Network (SOM)

Shaima Miqdad Mohamed Najeeb^{1*}, Dr. Razan Abdulhammed², Dr. Raid Rafi Omar Al-Nima³, and Haider TH. Salim ALRikabi⁴

^{1*}Department of Cyber Security and Cloud Computing Techniques Engineering, Northern Technical University Mosul, Iraq. shaimamiqdad76@ntu.edu.iq, <https://orcid.org/0000-0001-9746-8117>

²Department of Cyber Security and Cloud Computing Techniques Engineering, Northern Technical University Mosul, Iraq. rabdulhammed@ntu.edu.iq, <https://orcid.org/0000-0002-5939-6009>

³Department of Artificial Intelligence Technology Engineering, Northern Technical University Mosul, Iraq. raidraf1@gmail.com, <http://orcid.org/0000-0002-9673-453X>

⁴Electrical Engineering Department, College of Engineering, Wasit University, Wasit, Iraq. hdhiyab@uowasit.edu.iq, <https://orcid.org/0000-0002-5201-2084>

Received: January 18, 2025; Revised: February 28, 2025; Accepted: April 10, 2025; Published: May 30, 2025

Abstract

This study helps medical professionals to find out the genetic information related to five types of cancer. The information stored in RNA is characterized by its abundance, which is responsible for forming proteins in the human body. If there is any harm in this acid, this leads to the formation of cancerous diseases. The RNA contains huge numbers of genes so It is difficult for researchers to isolate the affected genes from the healthy genes. This paper presents a method for clustering the mutated genes for five types of cancer by using a novel feature extraction method with one of the traditional unsupervised machine learning algorithms (Self-Organized Map SOM). The proposed feature extraction method is called the Bi- directional Local-binary-pattern Signal (BLS), it is inspired from three traditional feature extraction methods named the Local Binary Pattern (LBP), Local Line Binary Pattern (LLBP), and Enhanced One-dimensional Local Binary Patterns (EOLBP) to extract the important textures for five types of cancer genes. This method finds the average for both direction features of a specific window in the RNA gene sequence and slides this window along the RNA to extract the features that are the most indicative features of cancerous genes. This study successfully clustered five types of cancer genes using the SOM classifier based on the features extracted by the proposed method. The study used a benchmarked dataset to test the proposed method. The accuracy for distinguishing the muted genes using the proposed method is 99.5%.

Keywords: Ribonucleic Acid RNA, Self-Organized Map, Local-binary-pattern, Cancer Detection.

1 Introduction

The abnormal division behavior of cells is a disease called cancer that damages the surrounding human tissues (Clauwaert et al., 2025; Zhang et al., 2025). In many cases, this disease may cause the death of

Journal of Internet Services and Information Security (JISIS), volume: 15, number: 2 (May), pp. 125-135.

DOI: 10.58346/JISIS.2025.12.009

*Corresponding author: Department of Cyber Security and Cloud Computing Techniques Engineering, Northern Technical University Mosul, Iraq.

the patient if the disease is diagnosed in its advanced stages (Ali et al., 2021) The early discovery of cancer may increase the chance of recovery (Sha et al., 2025; Han et al., 2025) For this reason, researchers try to find new methods for detecting this disease, but it is a complicated issue to find a new detection method and an efficient way to diagnose the existence of this kind of disease (Assegid & Ketema, 2023; Xiao et al., 2018; Alaidi et al., 2025; Zhang et al., 2025). One of them is by diagnosing the noncoding RNA, which plays a crucial role in cancer progression (Zhang et al., 2024; Mou & Harries, 2025). All vital activities within the cell are the result of gene sequences in RNA (Zhou et al., 2024; Shaath et al., 2022; Chen et al., 2021). RNA is responsible for several activities within the cell, such as the production of proteins and transcription (Yang et al., 2024; Wang et al., 2018). However, RNA-seq data analysis is facing some challenges, such as a lack of gene labels, high dimensionality, noise, and sparsity (Khosroabadi et al., 2025). These issues degrade tasks such as biomarker vital signs and classification of cell-type (Panahi et al., 2025; Wang et al., 2025).

Machine learning (ML) is a technique that allows a machine to develop its ability to learn without the need to be reprogrammed (Sarker, 2021; Gharibian & Ghorbani, 2007). It has enormous applications, such as target tracking, image recognition, face detection, etc (Hamad et al., 2024; Al_Airaji et al., 2025; Akshyaa Raaja Shri et al., 2023). ML usually needs a preprocessing stage that deals with and excludes unnecessary data (Najeeb & Ali, 2021; Najeeb & Ali, 2021) This reduces the computation time of the classifier. Also, it requires a feature extraction stage, which makes an ML classifier produce an accurate performance (Al-Dabag & Ozkurt, 2019; Lokman et al., 2018). ML includes a variety of algorithms that can be learned under a supervisor (supervised algorithms) or by itself (unsupervised algorithms) (Sun et al., 2024; Han et al., 2025).

Jagadeeswara Rao and Sivaprasad suggested a ML system for determining pancreatic cancer. They proposed two methods; the first used a bunch of classifiers for efficient gene classification and the second method called weighted gene co-expression network analysis (WGCNA) to classify RNA gene sequences that cause cancer (Maghsoodi, 2014; Jagadeeswara Rao & Sivaprasad, 2024) Laiqa Rukhsar et al. proposed a method with three stages; the first converts the RNA values into 2D images then features are extracted using Deep Learning (DL), and in the last stage eight DL algorithms are used for classification. Convolution neural network CNN is found to have the best performance among the eight algorithms (Rukhsar et al., 2022). Elbashir et al., suggested a method to detect tumors from RNA (Faris et al., 2024; Fairmont-State-University, 2020). Images are constructed from RNA and then Array-Array Intensity Correlation (AAIC) with filtering is used to remove the outliers and artifacts from the RNA images. The resultant data are applied to a CNN with certain parameters for detecting tumors (Ganesh & Siva Kumar, 2021). This method achieved optimal performance (98% accuracy) (Elbashir et al., 2019; Wang et al., 2018). Tehnan et al. proposed a combination of CNN with the Ebola Optimization Search Algorithm (EOSA) to detect breast cancer based on the cancer genome atlas dataset. This dataset was first preprocessed before being fed to the proposed method. This method provides good performance (Mohamed et al., 2023). Erkan Bostanci et al. used more than one ML, and DL algorithm to recognize colon cancer based on healthy RNA-seq and colon cancer patients (Kumar et al., 2024; Mehta & Reddy, 2024). All algorithms achieve good performance above 95% and the best algorithm (Bidirectional Long Short-Term Memory BiLSTM) achieves 98% (Bostanci et al., 2023) Comitani et al. utilized a CNN for clustering RNA childhood cancer. They applied their method to 13,313 transcriptomes and this method successfully diagnosed the disease in about 85% (Comitani et al., 2023). García-Díaz et al. proposed a Grouping Genetic Algorithm (GGA) with an ML algorithm as a fitness function to cluster five types of cancer based on an RNA-seq dataset. The proposed method shows optimal performance with 98% accuracy (García-Díaz et al., 2020). Kim et al. express more than ML algorithms to identify 21 types of

tumors using RNA-seq. They explore neural network (NN), support vector machine (SVM), k-nearest neighbors (kNN), and random forest (RF) algorithms. They found that NN provides the best performance with 94% accuracy (Kim et al., 2020). Nojima et al. used three ML algorithms (SVM, eXtreme gradient boosting and random forest (RF)) to predict drug-tolerant persister (DTP) in patient-derived organoids (PDOs) based on single-cell RNA sequencing (scRNA-Seq) data. The study shows an effective identification for DTP cell based on scRNA-Seq data for familial adenomatous polyposis (FAP) and provides candidate treatments (Nojima et al., 2025).

The organization of this paper is as follows; sec.II provides a brief description of the related theoretical subjects. Sec.III explains the prediction system used to classify cancer. It also includes a detailed explanation of the proposed method for extracting features from RNA. Sec. IV illustrates the experimental results of the proposed system. The last sec. provides the conclusions drawn from this research. Theoretical background

A. Local Binary Pattern

This method was originally used to analyze the texture of an image. It is a simple and effective technique for labeling image pixels by thresholding each pixel in a given window according to a focal pixel. First, the image is divided into 3×3 sub-windows. The center value of each sub-window is compared to the neighboring values starting from the upper left corner and in a clockwise manner, see Figure. 1 (Al-Nima, 2017; Arora et al., 2025; (Momani & Kóczy, 2025).

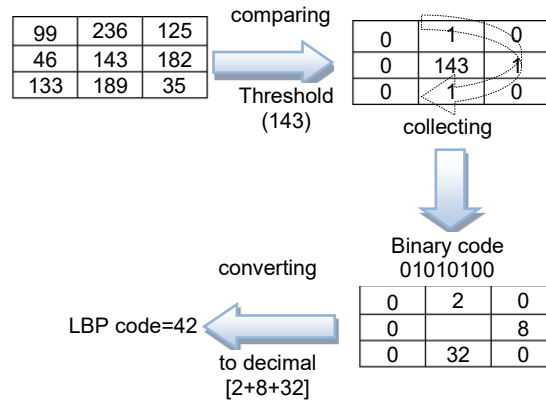


Figure 1: Evaluating the LBP Example (Al-Nima, 2017)

Local Line Binary Pattern (LLBP) is a modification of LBP. This modification extracts features in the same way as LBP but it takes two vector windows (vertical and horizontal windows) rather than 2D windows (Nanni & Lumini, 2009) Enhanced One-dimensional Local Binary Patterns (EOLBP) is another modification of LBP and this modification is used to cover 1D signals and analyze the signal texture for one direction (Hussein et al., 2023).

B. Self-Organizing Map SOM

SOM is one of the artificial neural algorithms that can be trained without any previous supervision. This mean there is no teacher for learning the algorithm. The main goal of this algorithm is to transform the high dimensional space into low dimensional space. SOMs utilize a competitive learning rule where its output node contend with themselves to represent a certain patterns within the input space. SOM

architecture is composed of a completely connected weighted link between the input layer nodes and the output layer nodes, see Figure 2 (Han & Kamber, 2006; Franklin et al., 2025)

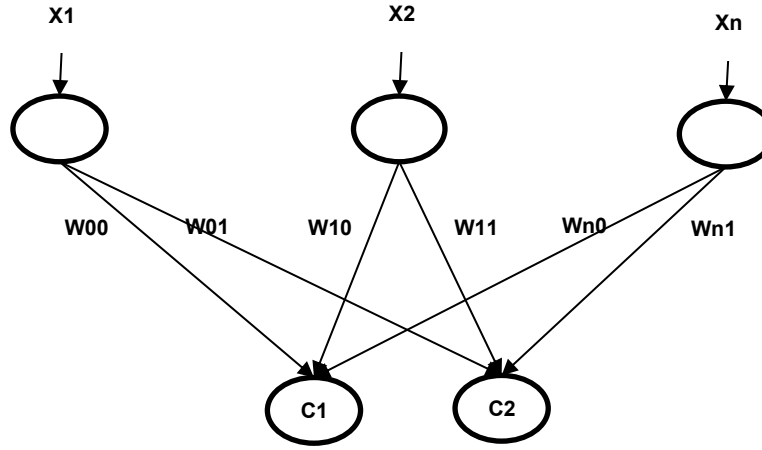


Figure 2: The Architecture of the SOM (Han & Kamber, 2001).

2 Methodology

This study presents a method for clustering five types of cancers. This method consisted of two stages; the proposed feature extraction method and a traditional SOM classifier. Figure shows the block diagram of the RNA prediction system.

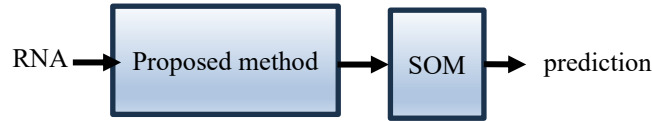


Figure 3: Block Diagram of the Prediction System

Proposed Method (Bi-directional LBP)

This method is designed to extract classifiable features from the RNA sequence. It is derived from the LBP method which finds the distinctive texture pattern for an image. It is inspired by LBP technique to identify texture patterns in an image. This method adapted the LBP concept to the 1D domain to extract the Bi-directional features from the one-dimensional RNA-seq. the key idea is to slide a fixed-size (nine samples) window along the RNA-seq samples with one sample stride and extract the context pattern bi-directionally relative to the center of the window. The proposed method is illustrated in the following steps:

1. Utilizing a window with 9 samples from the signal. The center of the window (the 5th sample) is denoted as X_c and the remaining samples are denoted as X_n for $n=0$ to 7
2. For each window, two-dimensional binary codes are evaluated using equations 1 and 2 with the thresholding function (equation 3).

$$V_{LR} = \sum_{n=0}^7 s(x_n - x_c) 2^n \quad (1)$$

$$V_{RL} = \sum_{n=0}^7 s(x_n - x_c) 2^{7-n} \quad (2)$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (3)$$

3. The resultant final feature referred to the $BLBP_C$ code is obtained by combining the two directional binary values (V_{RL} and V_{LR}) using the Euclidean norm:

$$BLBP_C = \sqrt{v_{LR}^2 + v_{RL}^2} \quad (4)$$

4. Striding the window one stride along the signal.

5. Repeat step 2 to 4 until the window scan the entire RNA-seq to obtain all $BLBP_C$.

Figure 3 illustrates the flowchart of the proposed method to extract features from the RNA.

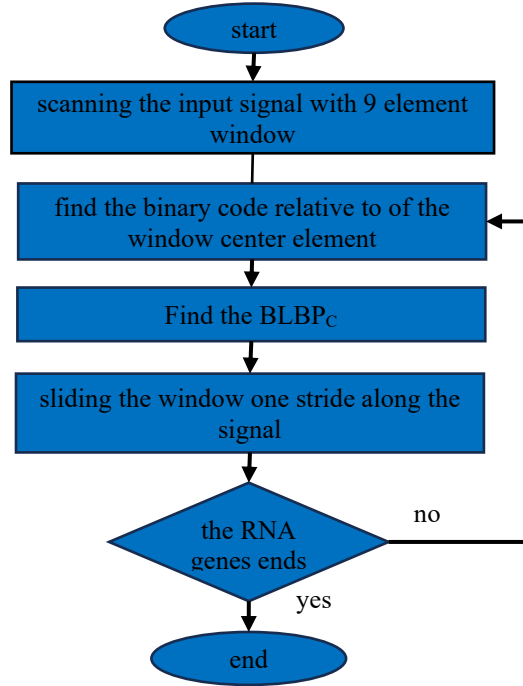


Figure 3: Flowchart of The Feature Extraction Proposed Method

3 Experimental Results

Dataset

The dataset was provided by the SYNAPSE website which provides different datasets and projects shared by researchers. This data set, called TCGA-PANCAN-HiSeq, is one of the datasets provided by this website. It has profiles for 801 subjects which contain the gene expression with a cancer type. Each RNA sample contains 20531 genes and these genes represent the feature of the subject. “BRCA”, “KIRC”, “COAD”, “LUAD” and “PRAD” are the cancer types that are embedded in the subject RNA (Fairmont-State-University, 2020).

Training the System

Since there are 801 subject RNA samples. 401 samples are used to train the proposed system while 400 RNA samples are used to test the system. SOM was utilized for clustering the five genes that cause cancer disease so the topology of the SOM has five output topologies, see Figure 4. After training the system, MSE was used to evaluate the SOM clustering result during testing the system, because the testing cancer genes were known in advance.

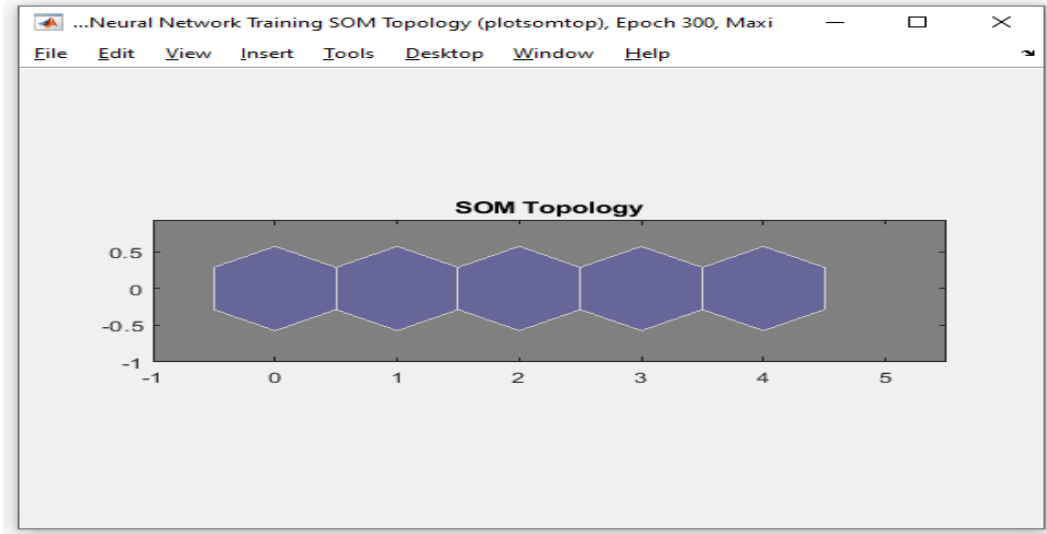


Figure 4: The SOM Topology

The proposed method shows a significant performance of 99.5% accuracy during testing the system for all five cancers. Table 1 shows the actual and predicted sample counts for the five types of cancer. To summarize the system performance, Table 2 shows the confusion matrix for testing 400 RNA samples to cluster five types of cancer. This matrix illustrates the robustness of the proposed method since its accuracy for predicting the five types of cancer falls between 99%-100% and the system's overall performance is 99.5%.

Table 1: The Actual and Predicted Sample Numbers of the Five Cancers After Testing the System

TP samples	39	72	145	73	71
Predicted samples	37	71	146	75	71

4 Conclusion

The experimental results show the proposed feature extraction method efficiently captures features generated from RNA-seq. These features significantly improve the traditional SOM classifier to identify the cancerous genes which achieving an interesting classification accuracy of 99.5%. The bidirectional RAN features obtained from the proposed BLBPc method enable the SOM classifier to categorize between five distinct cancer types. The overall system performance improvement is illustrated in the confusion matrix presented in the Table 2, which summarizes the classification of 400 RNA test samples. Figure 5 shows a visual comparison between the original RNA-seq and the extracted features.

Table 2: Confusion Matrix for Testing 400 Samples Using the Proposed Method

predicted	Expected				
	398	0	0	2	0
	0	399	1	0	0
	0	0	397	1	0
	0	0	1	396	0
	0	0	0	0	400

This leads to the conclusion that when the correct features are extracted, this helps the classifier to predict the disease correctly. All the evaluation methods (recall, accuracy, ... etc.) are driven from Table 2 proves this conclusion.

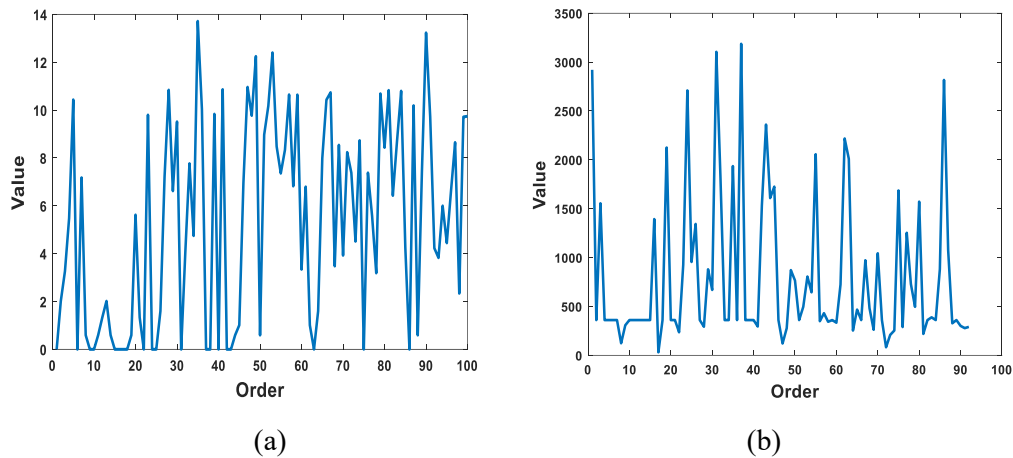


Figure 5: The Proposed Feature Extraction Method Rna-Seq Outcome, (A) The Original Rna-Seq Sample, (B) The Corresponding Blbp Code For The Rna-Seq

References

- [1] Akshyaa Raaja Shri, I. T., Dheenadhayalan, S., Rahul, B., Vishnu, B., & Suganya, S. (2023). Face recognition using LSTM in image processing. *International Journal of Advances in Engineering and Emerging Technology*, 14(1), 124–129.
- [2] Al_Airaji, R. a. M., Aljazeera, I. A., & ALRikabi, H. T. (2025). Challenges and approaches of face recognition: A review. In *AIP Conference Proceedings*. AIP Publishing. 3169 (1)
- [3] Alaidi, A. H. M., Ramadhan, Z. A., Alrubaye, J. S., ALRikabi, H. T. S., Mutar, H. A., & Svyd, I. (2025). AI-based monkeypox detection model using Raspberry Pi 5 AI Kit. *Sustainable Engineering and Innovation*, 7(1), 1-14.
- [4] Al-Dabag, M. L., & Ozkurt, N. (2019). EEG motor movement classification based on cross-correlation with effective channel. *Signal, Image and Video Processing*, 13, 567–573.
- [5] Ali, M. S., Miah, M. S., Haque, J., Rahman, M. M., & Islam, M. K. (2021). An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. *Machine Learning with Applications*, 5, 100036. <https://doi.org/10.1016/j.mlwa.2021.100036>
- [6] Al-Nima, R. R. O. (2017). *Signal processing and machine learning techniques for human verification based on finger textures* (Master's thesis). Newcastle University. <https://theses.ncl.ac.uk>

- [7] Arora, N., Budhiraja, I., Garg, D., Garg, S., Choi, B. J., & Hossain, M. S. (2025). Revolutionizing facial image retrieval: Multi-block and mean based local binary patterns with sign and magnitude analysis. *Alexandria Engineering Journal*, 116, 601-608.
- [8] Assegid, W., & Ketema, G. (2023). Harnessing AI for early cancer detection through imaging and genetics. *Clinical Journal for Medicine, Health and Pharmacy*, 1(1), 1–15.
- [9] Bostanci, E., Kocak, E., Unal, M., Guzel, M. S., Acici, K., & Asuroglu, T. (2023). Machine learning analysis of RNA-seq data for diagnostic and prognostic prediction of colon cancer. *Sensors*, 23(6), 3080. <https://doi.org/10.3390/s23063080>
- [10] Chen, S., Ben, S., Xin, J., Li, S., Zheng, R., Wang, H., ... & Wang, M. (2021). The biogenesis and biological function of PIWI-interacting RNA in cancer. *Journal of hematology & oncology*, 14(1), 93.
- [11] Clauwaert, J., McVey, Z., Gupta, R., Yannuzzi, I., Basrur, V., Nesvizhskii, A. I., ... & Prensner, J. R. (2025). Deep learning to decode sites of RNA translation in normal and cancerous tissues. *Nature Communications*, 16(1), 1275.
- [12] Comitani, F., Nash, J. O., Cohen-Gogo, S., Chang, A. I., Wen, T. T., Maheshwari, A., ... & Shlien, A. (2023). Diagnostic classification of childhood cancer using multiscale transcriptomics. *Nature medicine*, 29(3), 656-666.
- [13] Dawood, M. S., & Aldabag, M. L. (2023). Design of a 1D-CNN to classify surface EMG (SEMG) signals. *Nexo Revista Científica*, 36(6), 1020–1037. <https://doi.org/10.5377/nexo.v36i06.17458>
- [14] Elbashir, M. K., Ezz, M., Mohammed, M., & Saloum, S. S. (2019). Lightweight convolutional neural network for breast cancer classification using RNA-seq gene expression data. *IEEE Access*, 7, 185338–185348.
- [15] Fairmont-State-University. (2020). *DATASET*. <https://www.coursehero.com/file/69705853/stat437-Global-Proj2docx/>
- [16] Faris, R. A., Mosa, Q., & Albdairi, M. (2024). Robust classification for sub brain tumors by using an ant colony algorithm with a neural network. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 15(2), 270–285. <https://doi.org/10.58346/JOWUA.2024.12.018>
- [17] Franklin, G. L., Allende-Arandía, M. E., Paladio-Hernández, A., & Kurczyn, J. A. (2025). Spatio-temporal variability of coastal currents along the Yucatán Peninsula using Self-Organising Maps. *Regional Studies in Marine Science*, 83, 104098. <https://doi.org/10.1016/j.rsma.2025.104098>
- [18] Ganesh, R., & Siva Kumar, R. (2021). Diagnosis of brain tumor using artificial neural network. *International Academic Journal of Innovative Research*, 8(1), 6–10. <https://doi.org/10.9756/IAJIR/V8I1/IAJIR0802>
- [19] García-Díaz, P., Sánchez-Berriel, I., Martínez-Rojas, J. A., & Diez-Pascual, A. M. (2020). Unsupervised feature selection algorithm for multiclass cancer classification of gene expression RNA-Seq data. *Genomics*, 112(2), 1916-1925.
- [20] Gharibian, F., & Ghorbani, A. A. (2007). Comparative study of supervised machine learning techniques for intrusion detection. In *Fifth Annual Conference on Communication Networks and Services Research (CNSR'07)* (pp. 350-358). IEEE.
- [21] Hamad, A. A., Salih, T. A., & Mahmoud, A. F. (2024). Integration of DeepSORT and RFID technology for enhanced human tracking. *NTU Journal of Engineering and Technology*, 3(4), 17–25.
- [22] Han, J., & Kamber, M. (2001). *Data mining: Concepts and techniques*. Morgan Kaufmann Publishers.
- [23] Han, J., & Kamber, M. (2006). *Data mining: Concepts and techniques* (2nd ed.). Morgan Kaufmann Publishers.

- [24] Han, J., Park, S., Kim, L. A., Chung, S. H., Kim, T. I., Lee, J. M., ... & Lee, H. (2025). Machine Learning-Enabled Non-Invasive Screening of Tumor-Associated Circulating Transcripts for Early Detection of Colorectal Cancer. *International journal of molecular sciences*, 26(4), 1477. <https://doi.org/10.3390/ijms26041477>
- [25] Hussein, A. B., Al-Nima, R. R. O., & Han, T. (2023). Supplying choral with advanced techniques for stuttered people. *European Journal of Interdisciplinary Research and Development*, 21, 1-25.
- [26] JagadeeswaraRao, G., & Sivaprasad, A. (2024). An integrated ensemble learning technique for gene expression classification and biomarker identification from RNA-seq data for pancreatic cancer prognosis. *International Journal of Information Technology*, 16(3), 1505–1516.
- [27] Khosroabadi, Z., Azaryar, S., Dianat-Moghadam, H., Amoozgar, Z., & Sharifi, M. (2025). Single cell RNA sequencing improves the next generation of approaches to AML treatment: Challenges and perspectives. *Molecular Medicine*, 31(1), 33.
- [28] Kim, B.-H., Yu, K., & Lee, P. C. (2020). Cancer classification of single-cell gene expression data by neural network. *Bioinformatics*, 36(5), 1360-1366.
- [29] Kumar, B. S., Karpagavalli, S., Keerthana, K., & Krishnaja, A. (2024). Automatic segmentation of colon cancer using sam AI. *Archives for Technical Sciences*, 2(31), 296–304. <https://doi.org/10.70102/afts.2024.1631.296>
- [30] Lokman, M., Dabag, A., Ozkurt, N., Miqdad, S., & Najeeb, M. (2018). Feature selection and classification of EEG finger movement based on genetic algorithm. In *2018 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1–5). IEEE.
- [31] Maghsoodi, M. (2014). A new method to determine cluster heads based on principal component analysis in gene regulation process. *International Academic Journal of Science and Engineering*, 1(1), 104–111.
- [32] Mehta, V., & Reddy, P. (2024). Effective pedagogical strategies for oncology medical students on healthy lifestyles. *Global Journal of Medical Terminology Research and Informatics*, 1(1), 9–15.
- [33] Mohamed, T. I., Ezugwu, A. E., Fonou-Dombeu, J. V., Ikotun, A. M., & Mohammed, M. (2023). A bio-inspired convolution neural network architecture for automatic breast cancer detection and classification using RNA-Seq gene expression data. *Scientific Reports*, 13(1), 14644.
- [34] Momani, A. A., & Kóczy, L. T. (2025). A robust fingerprint identification approach using a fuzzy system and novel rotation method. *Pattern Recognition*, 159, 111134.
- [35] Mou, Z., & Harries, L. W. (2025). Integration of single-cell and bulk RNA-sequencing data reveals the prognostic potential of epithelial gene markers for prostate cancer. *Molecular Oncology*.
- [36] Najeeb, S. M. M., & Ali, S. M. (2021). Finding the discriminative frequencies of motor electroencephalography signal using genetic algorithm. *TELKOMNIKA*, 19(1), 285–291.
- [37] Nanni, L., & Lumini, A. (2009). A multi-matcher system based on knuckle-based features. *Neural Computing and Applications*, 18, 87-91.
- [38] Nojima, Y., Yao, R., & Suzuki, T. (2025). Single-cell RNA sequencing and machine learning provide candidate drugs against drug-tolerant persister cells in colorectal cancer. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 167693. <https://doi.org/10.1016/j.bbadis.2025.167693>
- [39] Panahi, B., Hassani, M., & Gharajeh, N. H. (2025). Integrative analysis of RNA-seq data and machine learning approaches to identify biomarkers for Rhizoctonia solani resistance in sugar beet. *Biochemistry and Biophysics Reports*, 41, 101920.
- [40] Rukhsar, L., Bangyal, W. H., Ali Khan, M. S., Ag Ibrahim, A. A., Nisar, K., & Rawat, D. B. (2022). Analyzing RNA-seq gene expression data using deep learning approaches for cancer classification. *Applied Sciences*, 12(4), 1850. <https://doi.org/10.3390/app12041850>

- [41] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>
- [42] Sha, Y., Ortiz, J. B., Bristow, S. L., Loranger, K., Meng, L., Zhao, X., ... & Xu, W. (2025). Refining the interpretation of variants of uncertain significance in hereditary cancer screening through integrated RNA sequencing. *Genetics in Medicine Open*, 3, 101914. <https://doi.org/10.1016/j.gimod.2025.101914>
- [43] Shaath, H., Vishnubalaji, R., Elango, R., Kardousha, A., Islam, Z., Qureshi, R., ... & Alajez, N. M. (2022, November). Long non-coding RNA and RNA-binding protein interactions in cancer: experimental and machine learning approaches. In *Seminars in cancer biology* (Vol. 86, pp. 325-345). Academic Press.
- [44] Sun, Q., Liu, Y., & Li, S. (2024). Weighted directed graph-based automatic seizure detection with effective brain connectivity for EEG signals. *Signal, Image and Video Processing*, 18(1), 899–909.
- [45] Wang, H., Fu, X., Liu, L., Wang, Y., Hong, J., Pan, B., ... & Yan, S. (2025). PhytoCluster: a generative deep learning model for clustering plant single-cell RNA-seq data. *aBIOTECH*, 1-13.
- [46] Wang, L., Xi, Y., Sung, S., & Qiao, H. (2018). RNA-seq assistant: Machine learning based methods to identify more transcriptional regulated genes. *BMC Genomics*, 19, 1–13.
- [47] Wu, Q., Fu, X., Liu, G., He, X., Li, Y., & Ou, C. (2025). N7-methylguanosine modification in cancers: From mechanisms to therapeutic potential. *Journal of Hematology & Oncology*, 18(1), 12.
- [48] Xiao, Y., Wu, J., Lin, Z., & Zhao, X. (2018). A semi-supervised deep learning method based on stacked sparse auto-encoder for cancer prediction using RNA-seq data. *Computer Methods and Programs in Biomedicine*, 166, 99-105.
- [49] Yang, S., Kim, S.-H., Yang, E., Kang, M., & Joo, J.-Y. (2024). Molecular insights into regulatory RNAs in the cellular machinery. *Experimental & Molecular Medicine*, 56(6), 1235-1249.
- [50] Zhang, Y., Fang, C., Zhang, L., Ma, F., Sun, M., Zhang, N., ... & Wu, J. (2025). Identification and Validation of Immune-Related Biomarkers and polarization types of macrophages in keloid Based on bulk RNA-seq and single-cell RNA-seq analysis. *Burns*, 107413.
- [51] Zhang, Y., Gu, X., Li, Y., Huang, Y., & Ju, S. (2024). Multiple regulatory roles of the transfer RNA-derived small RNAs in cancers. *Genes & Diseases*, 11(2), 597–613.
- [52] Zhou, Z., Zhang, J., Zheng, X., Pan, Z., Zhao, F., & Gao, Y. (2024). CIRI-deep enables single-cell and spatial transcriptomic analysis of circular RNAs with deep learning. *Advanced Science*, 11(14), 2308115. <https://doi.org/10.1002/advs.202308115>

Authors Biography



Shaima Miqdad Mohamed Najeeb is presently one of the lecturers of the Technical Engineering College of Computer and Artificial Intelligence, Northern Technical University, Mosul, Iraq. She received her B.E. degree in Computer Engineering in 1998 from Northern Technical University in Mosul, Iraq. Her M.Tech. A degree in computer Engineering focusing on "independent component analysis of EEG signal during various mental and movement tasks" in Iraq in 2013. Total Research Papers 10: Local 3 and International 7. Research interests include: - Artificial Intelligence - Information Technology- Digital Signal Processing.



Dr. Razan Abdulhammed is an esteemed Assistant Professor at Northern Technical University, Mosul, Iraq, specializing in Computer Engineering Technology. Her academic journey includes a Ph.D. in Computer Science and Engineering from the University of Bridgeport, USA, a Master's in Computer Engineering from the University of Mosul, and a Bachelor's in Computer Engineering Technology from the Technical College of Mosul. Dr. Abdulhammed has made significant contributions to the fields of intrusion detection systems, machine learning, IoT, and cybersecurity, with over 35 published research papers and participation in 25 conferences and seminars. Her work includes advancements in anomaly-based intrusion detection and dimensionality reduction techniques for network security. Among her accolades, she has received the Academic Achievement Award (2019) from Upsilon Pi Epsilon, the Engineering Achievement Award (2019) from the AAAEA-USA, and multiple honors from the University of Bridgeport, including First Place and Honorable Mention Awards. Dr. Abdulhammed's dedication to research and education continues to inspire innovation in her field.



Dr. Raid Rafi Omar Al-Nima received the BSc and MSc degrees in Technical Computer Engineering in 2000 and 2006, respectively. During 2006, he worked as an Assistant Lecturer at the Technical College of Mosul, Iraq. In 2011, he obtained the Lecturer scientific title in the same college. In 2017, he accomplished his PhD in the School of Electrical and Electronic Engineering at Newcastle University, UK. In 2020, he achieved the title of Assistant Professor in the Northern Technical University. His research interests are in the fields of pattern recognition, security, artificial intelligence, and image processing.



Haider TH. Salim ALRikabi is an Assistant Professor at the Faculty of Electrical Engineering Department, College of Engineering, Wasit University in Al Kut, Wasit, Iraq. He received his B.Sc. degree in Electrical Engineering in 2006 from the Al Mustansiriya University in Baghdad, Iraq. His M.Sc. degree in Electrical Engineering, focusing on Communications Systems from California State University/Fullerton, USA, in 2014. He is an author, coauthor, and Editor of some national and international journals and conference papers. His current research interests include communications systems with the mobile generation, control systems, intelligent systems, smart cities, renewable energies, signal processing, as well as image and speech processing, and the Internet of Things (IoT). To his credit, he has 20 articles in national databases and 95 in international databases.