SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I6P105

Volume 12 Issue 6, 54-64, June 2025

Original Article

Exponential Spider Wasp-Optimized Deep Learning Model for Type-2 Diabetes Detection Using Gene Expression Data

Chetan Nimba Aher[1], Madhavi Ranjeet Patil[2], Neha Ganvir[3], Sumati Manoj Jagdale[4]

[1]*Department of Computer Engineering, AISSMS Institute of Information Technology, Pune, India.*
[2]*Department of Electronics and Telecommunication, Shree Siddheshwar Women's College of Engineering, Solapur, India.*
[3]*Department of Electronics and Telecommunication, K J College of Engineering and Management Research, Pune, India.*
[4]*Department of Electronics and Telecommunication, Bharati Vidyapeeth's College of Engineering for Women, Pune, India.*

[1]*Corresponding Author : chetan.aher07@gmail.com*

Received: 03 April 2025 Revised: 04 May 2025 Accepted: 05 June 2025 Published: 27 June 2025

Abstract *- Type-2 Diabetes Mellitus (T2DM) occurs by insulin dysfunction, a chronic disease. Moreover, the human body cannot react with high sugar due to low secretion of insulin, which increases the blood sugar level. Due to the resistance of insulin or its low production, T2DM patients suffer a lot. The existing diagnosis process faces challenges such as low reliance on testing data, limited accessibility and chances for misdiagnosis. A new model, Exponential Spider Wasp Optimization, is introduced to address these issues, enabling the Quantum Dilated Convolutional Neural Network (ESWO_QDCNN) to detect T2DM. Initially, gene expression data is considered input from the gene expression dataset. Afterwards, the data transformation process is performed using Box-Cox transformation. Next, the feature selection process is performed by employing weighted Euclidean distance. Lastly, T2DM detection is attained by utilizing QDCNN, which is tuned using Exponential Spider Wasp optimization (ESWO). Here, the hybrid approach ESWO is developed by utilizing Exponential Weight Moving Average (EWMA) and Spider Wasp Optimizer (SWO). In addition to this, ESWO_QDCNN has achieved 91.524% accuracy, 90.854% sensitivity and 92.290% specificity.*

Keywords *- Gene Expression Data, Type2 Diabetes mellitus, Deep learning, Quantum Dilated Convolutional Neural Network, Spider Wasp Optimizer.*

1. Introduction

Gene expression data is the collection of information that helps to compare the various levels of messenger RNA (mRNA) in a cell structure. It also helps to study the protein formation in genes and the contribution of protein to cell function. Gene expression data is used in the biomedical field to diagnose diseases, especially diabetes. Diabetes mellitus is a metabolic disorder that occurs due to the rise in blood glucose and causes imperfection in the secretion of insulin [5]. Diabetes mellitus occurrence has increased as a result of changes in lifestyle, irregular food habits, overweight and obesity, and an ageing population [5, 7].

The various types of diabetes are Type-1 diabetes, Type-2 diabetes, and gestational diabetes. The increased level of blood sugar can cause T2DM in the human body. Generally, the blood glucose gets high due to low insulin production or insulin resistance. Due to this, the human body cannot react with insulin. This type of variation in insulin dysfunction does not cause any noticeable symptoms in humans, but it impacts the body immediately. Identifying insulin dysfunction is very challenging [4, 9, 10].

Detecting T2DM disease earlier gives a better solution [4]. Identifying diabetes-associated genes and their insights into disease diagnosis mechanisms involve different techniques [5]. Various Artificial Intelligence (AI) and Machine Learning (ML) techniques have been introduced by different researchers to automate the diagnosis of various diseases, including Diabetes [6]. In the field of AI, classification techniques analyse the data and distinguish whether the patient has the disease or not. Many supervised and unsupervised ML techniques identify important genes in gene databases. These methods are used to understand the structure of gene networks and help develop disease prediction models [5, 8]. To diagnose the disease, Deep learning (DL) methods are employed in the current scenario, and the most common DL techniques that are used in detecting T2DM are Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Deep Neural

Networks (DNN). CNN is used to analyze gene expression data to identify potential diabetes-related abnormalities. RNN can be used to analyze blood glucose levels to identify patterns and predict future trends. DNN can be applied to complex datasets, including clinical and lifestyle factors, to predict diabetes risk. DL-enabled techniques are considered to have higher accuracy than any other conventional ML approach.

The major aim is to design ESWO_QDCNN for the detection of Type-2 diabetes. Primarily, the input gene expression data obtained from the database undergoes data transformation. The data transformation process converts the input gene expression data from one format or structure into another format or structure by utilizing Box-Cox transformation. Subsequently, the transformed data is processed through the feature selection process, where suitable features are selected. This process is carried out by employing weighted Euclidean distance. Finally, the T2DM detection is performed using QDCNN, which is trained using ESWO. Here, the optimized hybrid approach ESWO is developed by combining EWMA and SWO.

An important contribution of the proposed model is discussed as follows.

- Proposed ESWO_QDCNN for T2DM detection: A potent model is designed to improve the detection of T2DM named ESWO_QDCNN. Here, ESWO-trained QDCNN undergoes the detection process.

The structure of the remaining sections is as follows: Section 1.2 reviews the literature on T2DM detection methods and their limitations, Section 2 explains the ESWO_QDCNN methodology, Section 3 presents the evaluation results of ESWO_QDCNN, and Section 4 provides the conclusion of ESWO_QDCNN for T2DM detection.

1.1. Motivation

In the biomedical field, gene expression data are preferred to diagnose T2DM. The traditional techniques available for disease detection are unreliable, and more skilled people are required to analyze data. So, an effective and automative technique is significant in detection. Motivated by this fact, a new T2DM detection method is introduced by analyzing classical approaches. Some shortcomings of the traditional detection methods are discussed in this section.

1.2. Literature Survey

Hu, Y. et al. [1] designed a Weighted Gene Co-Expression Network Analysis (WGCNA) for detecting shared genes of Pancreatic cancer and T2DM. It has the potential to identify the disease and provide a pathway to treat patients with T2DM. However, it had only limited public database availability, and the prolonged sample in the laboratory caused the results to be inaccurate. Yang, Y. et al. [2] presented Degree Matrix Network Entropy (DMNE) for diagnosing the various levels of T2DM. It was used to detect various levels of T2DM development and identify the important genes involved in T2DM disease occurrences. However, it could not generate gene comparison data. Li, J. et al. [3] introduced a Support Vector Machine (SVM) based model for identifying T2DM. It was a successful prediction model for conventional diagnosis markers, allowing clinicians to treat patients individually and effectively. However, it was devoid of thorough analysis and validation using a bigger sample size. Middha, K. et al. [4] developed a Competitive Multi-Verse Rider Optimizer (CMVRO)-based hybrid deep learning scheme for detecting T2DM. It had the potential to predict the disease using tuned classifiers and was efficient in detecting the disease. Even though it has some advantages, it lacks an optimization algorithm and advanced classifiers for the accurate prediction of T2DM disease.

1.3. Challenges

The following discusses the various challenges encountered by existing approaches.

- The method presented in [2] facilitated the extraction of more computable features from gene expression data. However, it was incapable of performing well in large datasets with diverse gene expression data.
- In [3], the designed approach achieved the identification of T2DM. However, it failed in detailed investigation and validation of clinical data with increased sample size to verify the prognosis of the gene signature involved in type 2 diabetes.
- Major challenges in Type 2 Diabetes Mellitus (T2DM) detection include the struggle of early detection due to its asymptomatic nature and time-consuming testing procedures. In addition, detecting T2DM with vast datasets has become very challenging for identifying the disease.

2. Methodology

This study proposes an Exponential Spider Wasp Optimization (ESWO) enabled Quantum Dilated Convolutional Neural Network (QDCNN) for the detection of Type-2 Diabetes Mellitus (T2DM) using gene expression data. The methodology consists of multiple stages: data acquisition, transformation, feature selection, and classification.

2.1. Proposed Exponential Spider Wasp Optimization enabled Quantum Dilated Convolutional Neural Network for T2DM detection

T2DM cause serious health issues, and if it is untreated over a long time, it may cause heart disease, stroke, kidney failure, and loss of vision. Early detection of the disease

gives better hope for the patients. Nowadays, the current diagnosis techniques are time-consuming and slow in detecting the disease. Hence, improved detection techniques are significant in assisting doctors in the diagnosis phase. Here, ESWO_QDCNN is presented for T2DM detection. Initially, the input gene expression data is from the database, and it is subjected to the data transformation process. The data transformation process will convert the input data from one format or structure into another format or structure by utilizing Box-Cox transformation. Next, the transformed data will be allowed in the feature selection process, where suitable features are selected. This process is carried out by employing weighted Euclidean distance. Lastly, the T2DM detection is accomplished by ESWO-trained QDCNN. Here, the optimized hybrid approach ESWO is developed by integrating EWMA and SWO. Figure 1 reveals the pictorial presentation of ESWO_QDCNN for T2DM detection.

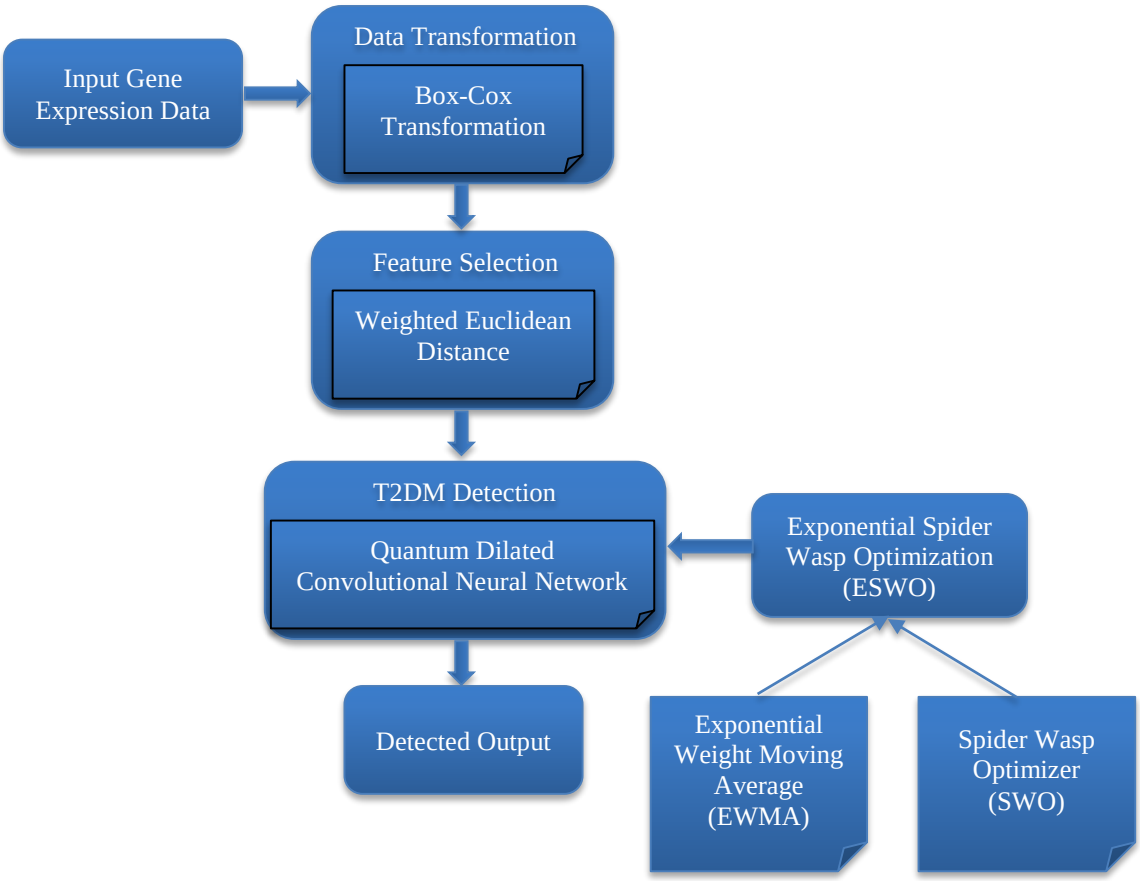


Fig. 1 Pictorial presentation of ESWO_QDCNN for T2DM detection

2.1.1. Data Acquisition

An input Gene expression data is acquired from the gene expression dataset [18], which can be formulated by,

$$G = \{G_1, G_2, \ldots G_a, .. G_b\} \quad (1)$$

Where, G_a represent a^{th} input gene expression data, whereas G_b Denotes the total gene expression data from the dataset G.

2.2. Box-Cox Transformation for Data Transformation

Data transformation is the transformation of the structure of the data. It is useful in identifying significant variations throughout the procedure; otherwise, it will go unnoticed [11]. Here, Box-Cox transformation is employed to perform data transformation. This approach is much preferable due to its flexibility, ability to handle non-normality and effectiveness in improving model performance. It is an organized approach to make data more suitable for statistical modelling, ensuring better results and more reliable inferences. Hence, this method is used for data transformation. The expression mentioned below is used for processing the data to perform data transformation and G_a is given as an input.

$$B_a^{(\lambda)} = \begin{cases} \lambda^{-1(G_a^{\lambda}-1)} & if \lambda = 0, \lambda \neq 0 \\ log(G_a) & \end{cases} \quad (2)$$

$$G^{(\lambda)} = X\beta + \in \quad (3)$$

Where $G_a$is the input data, $G^{(\lambda)}$ is the λ- transformed data, X is the design matrix, $\beta$is the set of parameters associated with λ-transformed data, ∈is the error term, and B_aIs the output obtained from Box-Cox transformation with the dimension of $[m \times n]$.

2.3. Feature Selection using Weighted Euclidean Distance

Feature selection is essential for reducing the size of datasets as they have varied sizes and will change drastically. Here, feature selection is employed by weighted Euclidean distance [12].

Weighted Euclidean distance improves accuracy by integrating objective and subjective features with adjustable weights, allowing for better customization of significant features. It has improved flexibility in handling diverse data types, making it more effective.

The resulting data from the data transformation B_a with the dimension of $[m \times n]$is provided as the input and S_a is the selected feature having a dimension of $[m \times l]$ where $n > l$.

$$S_{a[m\times l]} = \sqrt{\sum_i w_i (x_i - y_i)}^2 \quad (4)$$

Where, $w_i$represent weight, $x_i$represents feature to be selected and $y_i$represents the target.

2.4. T2DM Detection using QDCNN

T2DM detection is a time-consuming and unreliable process. Due to human intervention, an effective model is necessary. Here, QDCNN, which is trained using ESWO, performs the detection. The selected features with dimensions involved in the feature selection phase are subjected to detection.

2.4.1. Architecture of QDCNN

QDCNN [13] adopts the structure of a Convolutional Neural Network by integrating quantum layers with classical layers, and the quantum circuit can be placed anywhere in the mode. The quantum dilated convolutional layer is the dilation convolution that is performed in the convolution layer. Figure 2 illustrates the architecture of ESWO_QDCNN.

Convolution Operation

In CNN, the convolutional layer performs the convolution operation; hence, it has an important role. The convolution operation involves multiplying a set of weights with the input in convolutional networks, and it is a linear process.

Dilated Convolution

In dilated convolution, the convolution process is performed on the selected feature. S_a. In addition to the convolution layer, the dilated convolution layer has an extra hyperparameter called the dilation rate.

$$Q_a[c,d] = \sum_q \sum_l k[q.l] . S_a [c + q.r, d + l.r] \quad (5)$$

Where, Q_a is the output, S_a is the input, c and d are the location indices of Q_a, k is the filter and r is the dilation rate. The spatial resolution O_w and O_h of the resulting feature map, extracted from an $C_w \times C_h$ input image by a $t \times u$ kernel can be calculated as:

$$O_w = \left(\frac{C_w - t + 2p}{s}\right) + 1 \quad (6)$$

$$O_h = \left(\frac{C_h - u + 2p}{s}\right) + 1 \quad (7)$$

Where, O_w and O_h are the spatial resolution, s is the stride, $C_w \times C_h$ represents the input image by a $t \times u$ kernel, and p is the padding.

Quantum Convolution

Convolution is performed based on quantum circuits in this phase; hence, it is called quantum convolution. Quantum convolution consists of three modules to perform convolution, such as

Encoding Module: The existing data is encoded into a quantum state within the convolutional circuit. Therefore, the classical information is encoded in the initial state of a qubit. This type of encoding is referred to as single-variable or qubit encoding.

Entangled Module: The encoding module is applied to the cluster of single and multi-qubit gates obtained from the encoded quantum state. The single and multi-quit gates are associated with quantum convolution for extracting task-specific features. From the entanglement module, the final quantum states are measured.

Decoding Module: In a decoding module, output classical vectors are extracted by mapping the quantum state to the classical vector. The main task of the convolution layer is to extract a classical output vector $f(j,\theta)$ By using the mapping from the quantum state:

$$M: |j,\theta\rangle \rightarrow f(j,\theta) \quad (8)$$

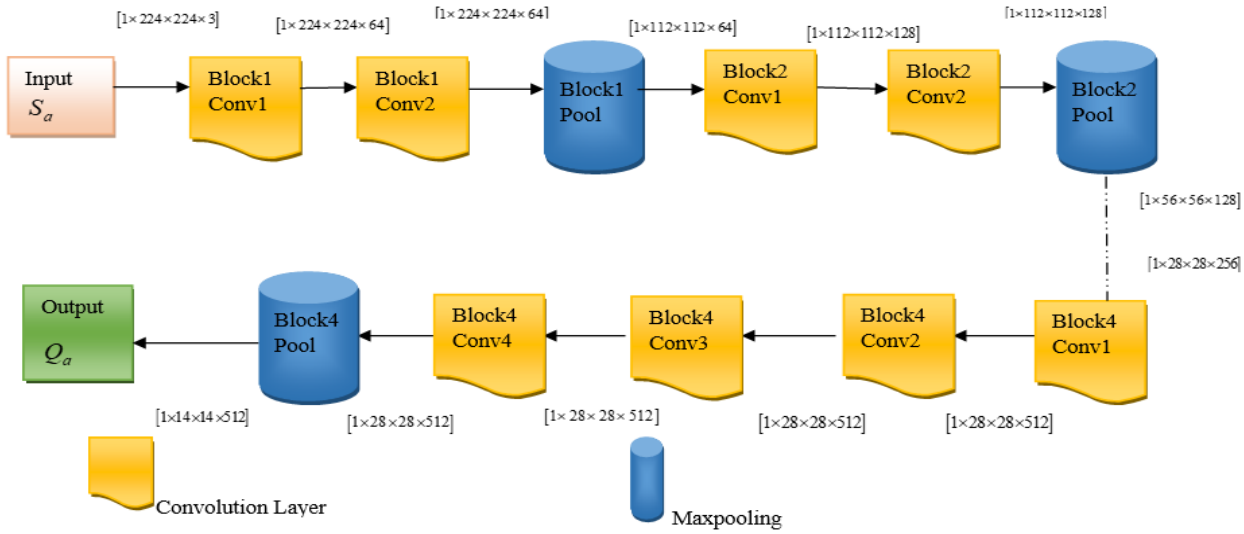


Fig. 2 Architecture of ESWO_QDCNN

2.4.2. Tuning of QDCNN Using ESWO

The parameters of QDCNN are optimized using the ESWO algorithm. Here, the ESWO algorithm is employed using the Exponentially Weighted Moving Average and Spider Wasp Optimizer algorithm. SWO algorithm inherited the characteristics of spider wasps, such as addressing challenges, unique hunting behaviour, nesting features, and mating. Solution Encoding: The solution encoding is used to obtain optimal solution in a D dimensional search space $D = [1 \times n]$, where n represents the learning parameter of QDCNN. Fitness function: The fitness function obtains optimal solutions from QDCNN. It is represented as,

$$F = \frac{1}{b}\sum_{a=1}^{b}[T - Q_a]^2 \tag{9}$$

Where F is the fitness function? Is the output obtained from QDCNN the targeted output?

Algorithm Steps:

The algorithm steps of ESWO_QDCNN for T2DM detection are as follows.

Step 1: Initialization of Parameters

The population of spiders and wasps was initialized in the search space [16]. Here R represents spider and W represents Wasp, N are the number of features e and g are the number of spiders and wasps.

$$R = \begin{bmatrix} R_{1,1} & \cdots & R_{1,N/2} & \cdots & R_{1,N} \\ \vdots & \ddots & \vdots & / & \vdots \\ R_{e/2,1} & \cdots & R_{e/2,N/2} & \cdots & R_{e/2,N} \\ \vdots & / & \vdots & \ddots & \vdots \\ R_{e,1} & \cdots & R_{e,N/2} & \cdots & R_{e,N} \end{bmatrix}_{e\chi N} \tag{10}$$

$$W = \begin{bmatrix} W_{1,1} & \cdots & W_{1,N/2} & \cdots & W_{1,N} \\ \vdots & \ddots & \vdots & / & \vdots \\ W_{g/2,1} & \cdots & W_{g/2,N/2} & \cdots & W_{g/2,N} \\ \vdots & / & \vdots & \ddots & \vdots \\ W_{g,1} & \cdots & W_{g,N/2} & \cdots & W_{g,N} \end{bmatrix}_{g\chi N} \tag{11}$$

Step 2: Estimate Fitness Function

The fitness function is used to determine the optimal solution for T2DM detection. The fitness function is already computed in Equation (9).

Step 3: Crossover based on Spider Movement

To perform crossover, the original number of features is extracted from the features of spiders and wasps, and the newly formed subset of each iteration proceeds to a new population [16]. Here are the features selected from spider wasps, the reminder, the number of spiders and wasps, and the number of features selected, representing the top population.

$$Fs_g{}^R = R_v\left[H\left(g, \frac{N_R}{2}\right)\right]\left[:\frac{N_{Fs}}{2}\right] \tag{12}$$

$$Fs_g{}^W = W_v\left[H\left(g, \frac{N_W}{2}\right)\right]\left[\frac{N_{Fs}}{2}:\right] \tag{13}$$

Step 4: Mutation based on Wasp Movement

Changing certain features in the wasp population gives certain feature changes to perform mutation operations [16].

$$L_{ge}{}^W = W_v\left[H\left(g, \frac{N_W}{2}\right)\right][e] \times [J(0,1)] \tag{14}$$

Here, L_{ge}^{W} the mutated value of Wasp from the subset of a number of spiders and Wasp returns an integer, which g is the number of spiders and wasps.

Step 5: Updated Solution

The updated solution is the integration of EWMA [14] with the SWO algorithm [15]. The standard equation from the SWO algorithm is given as,

$$\overrightarrow{RW_e}^{z+1} = \overrightarrow{RW_e}^{z} + \mu_1 * \left(\overrightarrow{RW_U}^{z} - \overrightarrow{RW_V}^{z}\right) \quad (15)$$

Let us assume,

$$\overrightarrow{RW_e}^{z+1} = I_e(z+1) \quad (16)$$

$$\overrightarrow{RW_e}^{z} = I_e(z) \quad (17)$$

$$\overrightarrow{RW_U}^{z} = I_{eU}(z) \quad (18)$$

$$\overrightarrow{RW_v}^{z} = I_{eV}(z) \quad (19)$$

Then Equation (15) becomes,

$$I_e(z+1) = I_e(z) + \mu_1 * \left(I_{eU}(z) - I_{eV}(z)\right) \quad (20)$$

The general equation from EWMA is given as,

$$I_e{}^{E}(z) = \lambda I_e(z) + (1-\lambda) I_e{}^{E}(z-1) \quad (21)$$

$$I_e(z) = \frac{1}{\lambda}\left[I_e{}^{E}(z) - (1-\lambda) I_e{}^{E}(z-1)\right] \quad (22)$$

Substituting Equation (17) in Equation (16),

$$I_e(z+1) = \frac{\left[I_e{}^{E}(z) - (1-\lambda) I_e{}^{E}(z-1)\right] + \left[\mu_1 * (I_{eU}(z) - I_{eV}(z))\right] * \lambda}{\lambda} \quad (23)$$

Where Z represents the iteration number, $I_e{}^{E}(z-1)$ is the individual's position using EWMA of z^{th} item. $I_e{}^{E}(z+1)$ Is the updated ESWO solution for training QDCNN?

Step 6: Re-Evaluation of Fitness Function

The fitness function will iterate continually till it attains the optimal solution.

Step 7: Termination

In the termination phase, the optimization will be achieved using the repeated iteration of ESWO by training QDCNN using the algorithm.

3. Results and Discussion

This section discusses the result obtained from the T2DM detection process of ESWO_QDCNN, along with the metrics and dataset.

3.1. Experiment Setup

ESWO_QDCNN, designed for T2DM detection, is executed using the PYTHON tool.

3.2. Dataset Description

In the Gene Expression Dataset [18], human islets were extracted from organ donors' pancreas using collagenase digestion, density gradient purification, hand selection, and two days of culture in M199 culture media. It contains selected columns of [0 - 147] and a dataset length of (14, 22284). Thirteen samples are used in the Affymetrix Human Genome U133A Array platform. It contains an additional file, GSE25724_RAW.tar, which is 26.0 MB in size and is of the TAR (of CEL) file type.

3.3. Evaluation Metrics

Accuracy, sensitivity and specificity metrics are used to assess the ESWO_QDCNN approach.

3.3.1. Accuracy

Accuracy [17] is the model's proportion of correct predictions made while evaluating the samples. It is formulated as,

$$Accuracy = \frac{Z_{pos}+Z_{neg}}{Z_{pos}+Y_{pos}+Y_{neg}+Z_{neg}} \quad (24)$$

Here, Z_{pos} it denotes true positives and negatives, whereas it specifies false positives and false negatives.

3.3.2. Sensitivity

Sensitivity [17] is calculated as the percentage of cases that are predicted correctly, which is modelled as,

$$Sensitivity = \frac{Z_{pos}}{Z_{pos}+Y_{neg}} \quad (25)$$

3.3.3. Specificity

Specificity [17] is evaluated by the percentage of non-diabetic individuals correctly identified by the specific test. It is represented as,

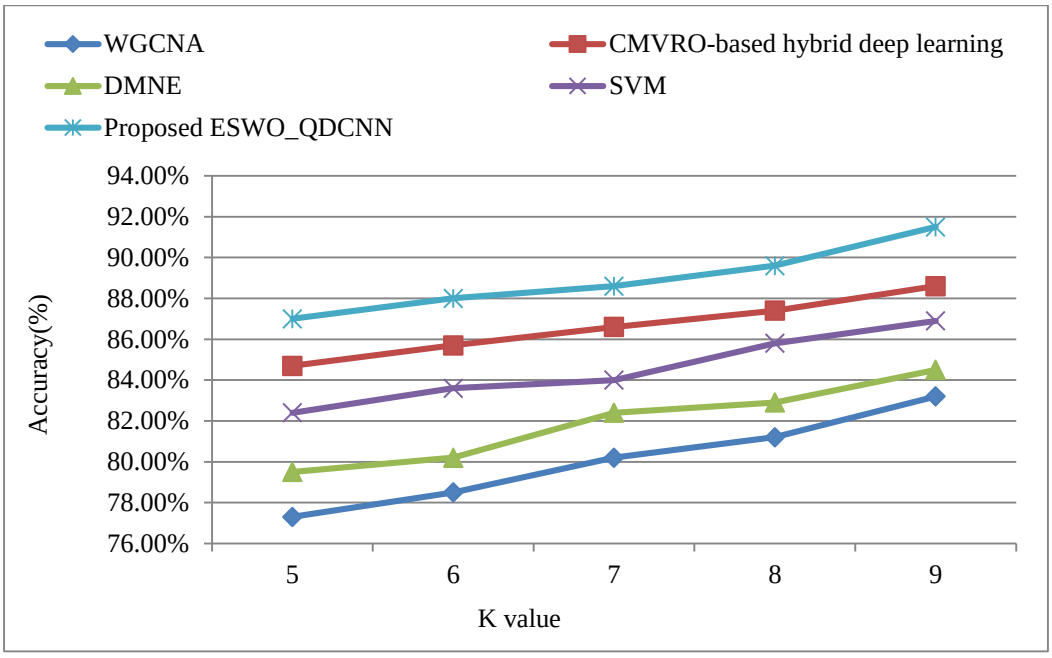
$$Specificity = \frac{Z_{neg}}{Z_{neg}+Y_{pos}} \quad (26)$$

3.4. Comparative Analysis

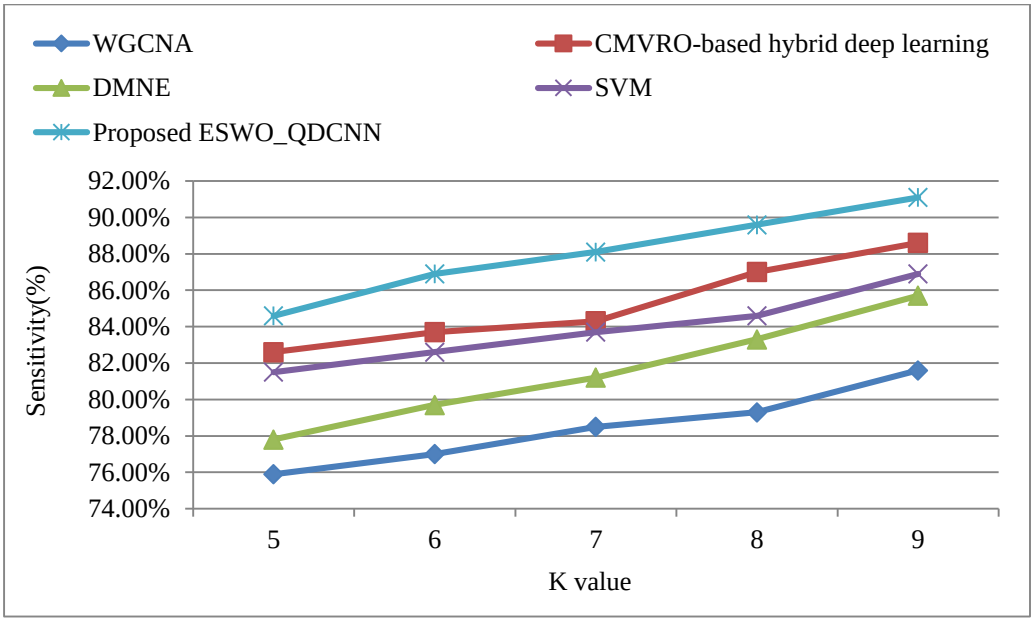
The methods WGCNA [1], DMNE [2], SVM [3] and CMVRO [4] are used to compare the performance of the proposed ESWO_QDCNN and show its efficiency.

3.4.1. Analysis based on K-Value

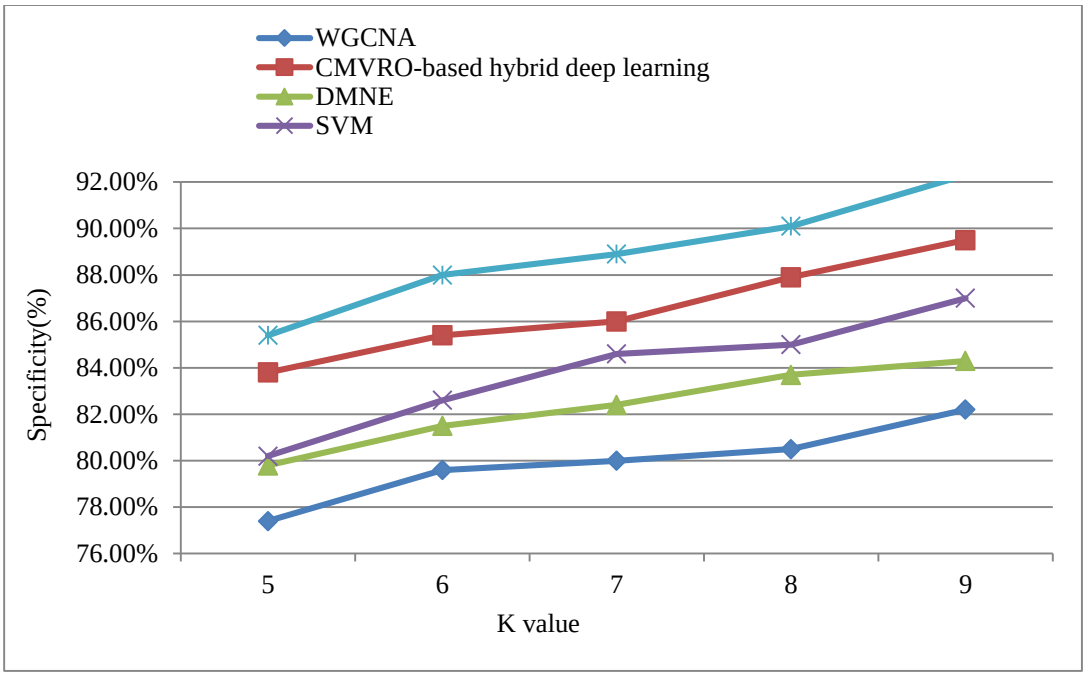
Figure 3 represents the study of ESWO_QDCNN with other classical methods based on metrics such as accuracy, Sensitivity and specificity by changing the K-value. Figure 3 (a) illustrates the accuracy of ESWO_QDCNN and other methods. For K-Value=9, ESWO_QDCNN reaches an accuracy of 91.524%, whereas other methods reached 83.257%, 84.524%, 86.954% and 88.541%. It shows that the performance of ESWO_QDCNN is improved by 9.032%, 7.648%, 4.993% and 3.259%. Figure 3 (b) shows the Sensitivity of ESWO_QDCNN and other classical methods. When K-Value=9, ESWO_QDCNN achieved a Sensitivity of 90.854%, whereas other methods achieved 81.564%, 85.632%, 86.933% and 88.521%. It demonstrates that the Sensitivity of ESWO_QDCNN is improved by 10.226%, 5.748%, 4.316% and 2.567%. Figure 3 (c) represents the specificity of ESWO_QDCNN. When considering a K-value of 9, the traditional methods achieved the specificity of 82.365%, 84.256%, 86.924% and 89.521%, whereas ESWO_QDCNN attained 92.290%. It shows the performance improvement of ESWO_QDCNN by 10.753%, 8.704%, 5.814% and 2.999%.



(a)



(b)

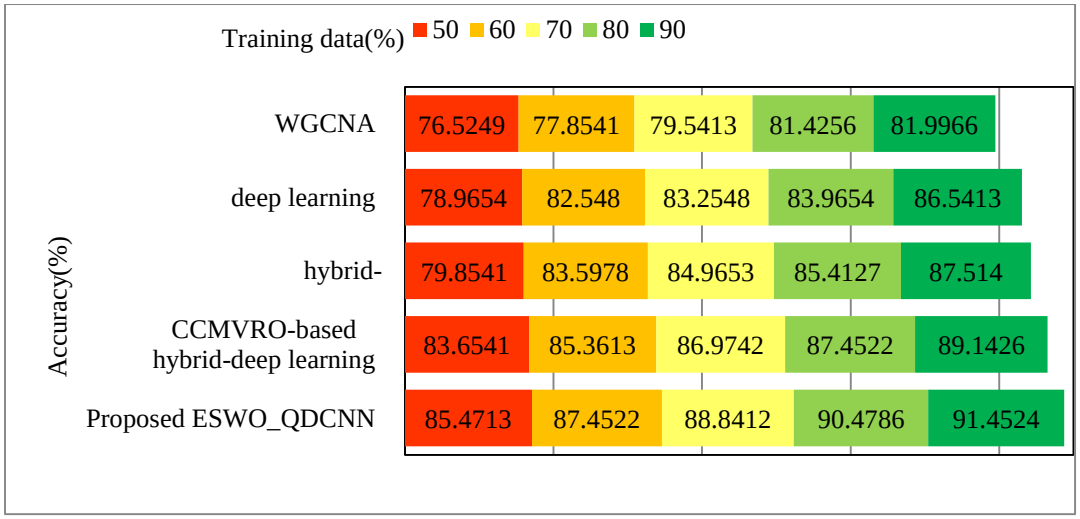


(c)

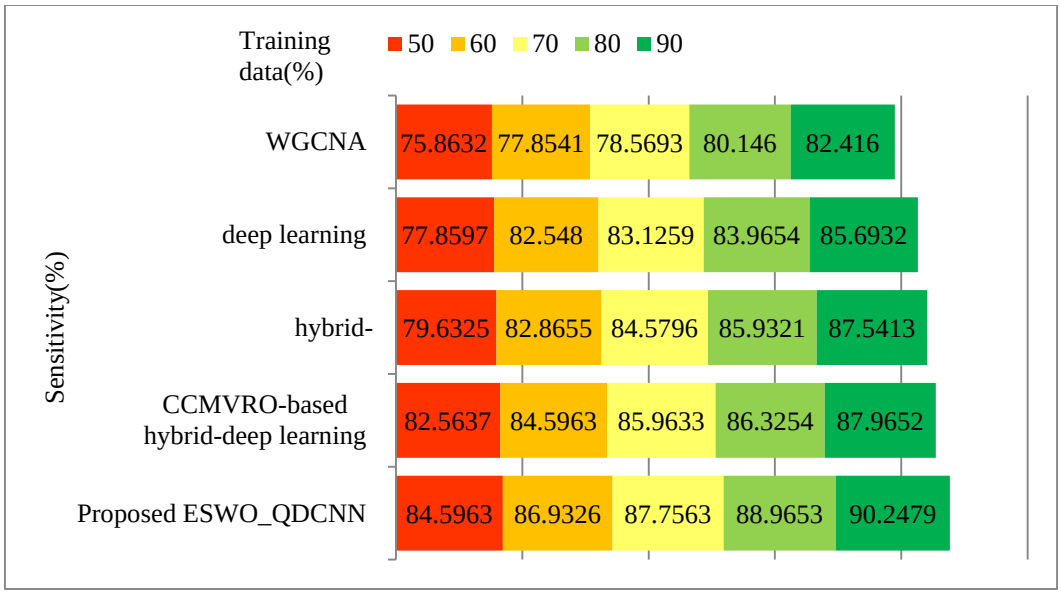
Fig. 3 Analysis based on K-value (a) Accuracy, (b) Sensitivity, and (c) Specificity.

3.4.2. Analysis based on Training Data

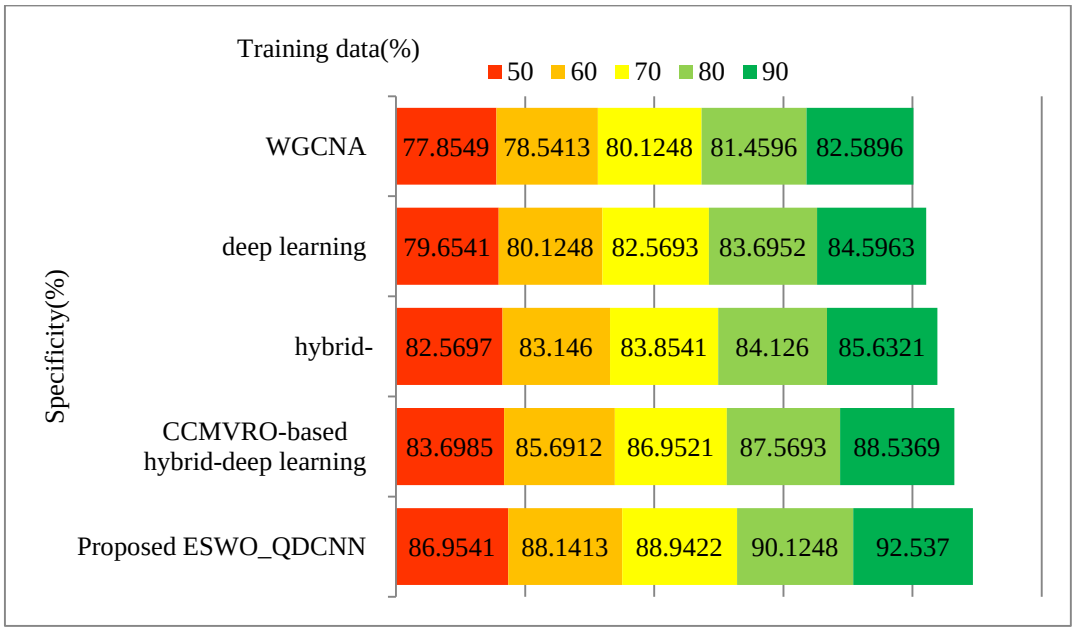
Figure 4 demonstrates the evaluation of ESWO_QDCNN with other comparative methods by varying training data depending on various metrics such as Accuracy, Sensitivity and Specificity. Figure 4 (a) displays the accuracy of ESWO_QDCNN and other methods with Training data=90%. Here, ESWO_QDCNN has an accuracy of 91.452%, and the other methods show an accuracy of 81.997%, 86.541%, 87.514% and 89.143%. It depicts the performance improvement is enhanced by 10.340%, 5.370%, 4.307% and %2.526 while comparing it with other methods. Figure 4 (b) shows the Sensitivity of ESWO_QDCNN, which was analyzed with other methods while considering training data = 90%. Here, the Sensitivity attained by ESWO_QDCNN is 90.248%, and the other traditional methods achieved 82.416%, 85.693%, 87.541% and 87.965%. It shows the performance improvement of ESWO_QDCNN by 8.678%, 5.047%, 2.999% and 2.529%. Figure 4 (c) portrays the specificity of ESWO_QDCNN compared with other methods while considering training data=90%. Here, ESWO_QDCNN achieved a specificity of 92.537%, and the other methods attained 82.590%, 84.596%, 85.632% and 88.537%. It shows performance improvement of ESWO_QDCNN by 10.750%, 8.581%, 7.462% and 4.323%.



(a)



(b)



(c)

Fig. 4 Analysis based on Training data, (a) Accuracy, (b) Sensitivity, and (c) Specificity.

3.5. Comparative Discussion

Table 1 shows the assessment of ESWO_QDCNN and other traditional approaches. When K-Value is 9, ESWO_QDCNN attained 91.524% accuracy, and the other classical methods attained 83.257%, 84.524%, 86.954% and 88.541%. The high accuracy rate shows that ESWO_QDCNN can capture better complex patterns in the gene-expression data. The Sensitivity achieved by ESWO_QDCNN while considering K-Value=9 is 90.854%, while other methods achieved 81.564%, 85.632%, 86.933% and 88.521%. The enhanced Sensitivity of ESWO_QDCNN shows robustness in handling imbalanced gene-expression data and can obtain optimized parameters. The classical methods obtained the specificity of 82.365%, 84.256%, 86.924% and 89.521%, whereas ESWO_QDCNN obtained the enhanced specificity of 92.290%. It denotes that ESWO_QDCNN can effectively identify negative instances, reduce false positives, and perform well in identifying T2DM. From the comparative analysis, it is obvious that ESWO_QDCNN is an effective approach for T2DM identification. Furthermore, ESWO_QDCNN attained an accuracy of 91.524%, Sensitivity of 90.854% and specificity of 92.290% for K-Value=9.

Table 1. Comparative discussion of ESWO_QDCNN

Setups	Metrics/ Methods	WGCNA	DMNE	SVM	CMVRO	Proposed ESWO_QDCNN
K-Value=9	Accuracy (%)	83.257	84.524	86.954	88.541	**91.524**
	Sensitivity (%)	81.564	85.632	86.933	88.521	**90.854**
	Specificity (%)	82.365	84.256	86.924	89.521	**92.290**
Training Data=90%	Accuracy (%)	81.997	86.541	87.514	89.143	**91.452**
	Sensitivity (%)	82.416	85.693	87.541	87.965	**90.248**
	Specificity (%)	82.590	84.596	85.632	88.537	**92.537**

4. Conclusion

The early detection of T2DM is very crucial for eliminating certain complications in heart, kidney, vision and hearing problems in T2DM patients. Here, this condition is non-curable but can be prevented if it is diagnosed early. The main problems faced by T2DM detection methods are low accuracy, limited datasets, low processing, and inability to explore various datasets. As AI develops, various detection methods are being developed for effective disease detection and application in the real world. In this study, ESWO_QDCNN is considered for T2DM detection. Primarily, the input data is obtained from the Gene Expression dataset. Then, data transformation is achieved using Box-Cox transformation. Then, weighted Euclidean Distance is opted to perform feature selection. Lastly, the T2DM detection is achieved by employing ESWO_QDCNN. Here, the optimized hybrid approach ESWO is developed by utilizing the EWMA and SWO algorithms. Additionally, ESWO_QDCNN achieved 91.524% accuracy, 90.854% sensitivity and 92.290% specificity. In future, by enhancing model interpretability and decision-making integration, this system can enable more robust and reliable detection of T2DM in clinical environments.

References

[1] Yifang Hu et al., "Identification of the Shared Gene Signatures and Biological Mechanism in Type 2 Diabetes and Pancreatic Cancer," *Frontiers in Endocrinology*, vol. 13, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Yingke Yang et al., "Detecting the Critical States of Type 2 Diabetes Mellitus Based on Degree Matrix Network Entropy by Cross-Tissue Analysis," *Entropy*, vol. 24, no. 9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Jiabin Li et al., "Identification of Type 2 Diabetes Based on a Ten-Gene Biomarker Prediction Model Constructed Using a Support Vector Machine Algorithm," *BioMed Research International*, vol. 2022, no. 1, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Karuna Middha, and Apeksha Mittal, "An Effective Feature Selection Method for Type 2 Diabetes Mellitus Detection Using Gene Expression Data," *Intelligent Decision Technologies*, vol. 17, no. 3, pp. 595-606, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Amira M. Elsherbini et al., "Decoding Diabetes Biomarkers and Related Molecular Mechanisms by Using Machine Learning, Text Mining, and Gene Expression Analysis," *International Journal of Environmental Research and Public Health*, vol. 19, no. 21, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Md. Kowsher et al., "Prognosis and Treatment Prediction of Type-2 Diabetes Using Deep Neural Network and Machine Learning Classifiers," *2019 22nd International Conference on Computer and Information Technology*, Dhaka, Bangladesh, pp. 1-6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[7] Paul Zimmet, K.G.M.M. Alberti, and Jonathan Shaw, "Global and Societal Implications of the Diabetes Epidemic," *Nature*, vol. 414, no. 6865, pp. 782-787, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[8] Md. Maniruzzaman et al., "Statistical Characterization and Classification of Colon Microarray Gene Expression Data Using Multiple Machine Learning Paradigms," *Computer Methods and Programs in Biomedicine*, vol. 176, pp. 173-193, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Muhammad Fazal Ijaz et al., "Hybrid Prediction Model for Type 2 Diabetes and Hypertension Using DBSCAN-Based Outlier Detection, Synthetic Minority Over Sampling Technique (SMOTE), and Random Forest," *Applied Sciences*, vol. 8, no. 8, pp. 1-22, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Mehdi Boroumand, Mo Chen, and Jessica Fridrich, "Deep Residual Network for Steganalysis of Digital Images," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1181-1193, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jorge I. Vélez, Juan C. Correa, and Fernando Marmolejo-Ramos, "A New Approach to the Box–Cox Transformation," *Frontiers in Applied Mathematics and Statistics*, vol. 1, pp. 1-10, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[12] Jan de Leeuw, and Sandra Pruzansky, "A New Computational Method to Fit the Weighted Euclidean Distance Model," *Psychometrika*, vol. 43, no. 4, pp. 479-490, 1978. [CrossRef] [Google Scholar] [Publisher Link]

[13] Yixiong Chen, "Quantum Dilated Convolutional Neural Networks," *IEEE Access*, vol. 10, pp. 20240-20246, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Petar Čisar, and Sanja Maravić Čisar, "Optimization Methods of EWMA Statistics," *Acta Polytechnica Hungarica*, vol. 8, no. 5, pp. 73-87, 2011. [Google Scholar] [Publisher Link]

[15] Jinxue Sui, Zifan Tian, and Zuoxun Wang, "Multiple Strategies Improved Spider Wasp Optimization for Engineering Optimization Problem Solving," *Scientific Reports*, vol. 14, no. 1, pp. 1-29, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Himansu Das et al., "Enhancing Software Fault Prediction through Feature Selection With Spider Wasp Optimization Algorithm," *IEEE Access*, vol. 12, pp. 105309-105325, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] J. Pradeep Kandhasamy, and S. Balamurali, "Performance Analysis of Classifier Models to Predict Diabetes Mellitus," *Procedia Computer Science*, vol. 47, pp. 45-51, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[18] Chris Crawford, Gene Expression Dataset, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/crawford/gene-expression

[19] Madhukar G. Andhale, Renu Pathak, and Shirish Kulkarni, "Design of an Efficient Bio Inspired Optimization Model to Determine Optimal Power between Electric Motor and Internal Combustion Engine in Hybrid Electric Vehicles under Various Driving Scenarios," *International Journal of Technology Engineering Arts Mathematics Science*, vol. 3, no. 2, pp. 1-11, 2023. [Google Scholar] [Publisher Link]