



An Optimal Stacked ResNet-BiLSTM-Based Accurate Detection and Classification of Genetic Disorders

K. Nandhini¹ · G. Tamilpavai²

Accepted: 13 February 2023 / Published online: 19 May 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Gene is located inside the nuclease and the genetic data is contained in deoxyribonucleic acid (DNA). A person's gene count ranges from 20,000 to 30,000. Even a minor alteration to the DNA sequence can be harmful if it affects the cell's fundamental functions. As a result, the gene begins to act abnormally. The sorts of genetic abnormalities brought on by mutation include chromosomal disorders, complex disorders, and single-gene disorders. Therefore, a detailed diagnosis method is required. Thus, we proposed an Elephant Herd Optimization-Whale Optimization Algorithm (EHO-WOA) optimized Stacked ResNet-Bidirectional Long Term Short Memory (ResNet-BiLSTM) model for detecting genetic disorders. Here, a hybrid EHO-WOA algorithm is presented to assess the Stacked ResNet-BiLSTM architecture's fitness. The ResNet-BiLSTM design uses the genotype and gene expression phenotype as input data. Furthermore, the proposed method identifies rare genetic disorders such as Angelman Syndrome, Rett Syndrome, and Prader-Willi Syndrome. It demonstrates the effectiveness of the developed model with greater accuracy, recall, specificity, precision, and f1-score. Thus, a wide range of DNA deficiencies including Prader-Willi syndrome, Marfan syndrome, Early Onset Morbid Obesity, Rett syndrome, and Angelman syndrome are predicted accurately.

Keywords Genetic disorder · Stacked ResNet-bidirectional long short term memory model · Elephant herd optimization · Whale optimization algorithm · Angelman syndrome · Prader-Willi syndrome

1 Introduction

DNA or deoxyribonucleic acid is a blueprint for whole lives and it is a basic determinant of the functioning of the body. DNA is the hereditary material of all living organisms which includes the genetic information of the whole cell [1]. The information is stored in the form of codes which are built using Guanine, Adenine, Thymine, and Cytosine chemical bases.

✉ K. Nandhini
nandhukk28@gmail.com

¹ Department of Computer Science and Engineering, Anna University, Chennai, India

² Department of Computer Science and Engineering, Government College of Engineering, Tirunelveli, India

Human DNA comprises nearly three billion bases and most of them are similar for all people. These base sequences define the information presented for building an organism. The DNA that carries genetic information is called the gene and is placed inside the nucleus. The gene is organized as lengthy structures called chromosomes and the genome includes 23 chromosomes [2]. The human chromosome comprises 10% Ribonucleic acid (RNA), 15% DNA, and 75% proteins. Every human has a gene count of approximately 20,000 to 30,000. Even a small DNA sequence modification is pathogenic if it has negative impacts on the biological process of the cell. This leads to abnormal functioning of the gene [3]. Genetic disorders induced by a mutation include chromosomal disorders (abnormalities in chromosomes), complex disorders (related to numerous gene mutations), and single-gene disorders (caused by a mutation in a single gene). Thus it is necessary for a precise diagnosis process.

Genome sequencing helps to acquire precise data on nucleic acid [4]. The primary human gene sequencing approach utilizes a three-phase divide and conquers method which is done by randomly cutting the DNA into numerous segments. It leads to higher costs and did not produce complete maps and they were complicated. In adult and pediatric medicine, genome sequencing techniques quickly analyze the genome at increasing resolution for the identification of single-gene or oligogenic genetic disorders [5]. Even though the various advantages posed by the whole-genome shotgun (WGS) sequencing and clone-by-clone method, it has several drawbacks like producing less accuracy. WGS is a short gun sequence reader. The increased possibility to discover the genetic component of health issues, at a lower cost than that of the old procedures, is a clear benefit of next-generation sequencing techniques. The high-throughput—next-generation sequencing (HT-NGS) technologies minimized the cost of genome sequencing [6] Also, it develops new diagnostic and therapeutic ways of treating diseases like cancer. Thus even the thousands of single personal gene with the suspicion of complex disease is sequenced. These have several challenges like adapting to newer platforms.

Disorders like Pelizaeus-Merzbacher disease and Potocki-Lupski syndrome are brought on by genomic rearrangements. The discovery of uncommon CNVs as a potential contributor to complex diseases including, Parkinson's, Alzheimer's, HIV-susceptibility psoriasis, and autism. This copy number variant (CNVs) and rare variants, of genomic variation, influence the disease phenotypes heavily. Machine learning developed many algorithms for enabling the computer to analyze the mass complicated datasets related to genomes [7]. That is a promising platform for developing sequence-based diagnostics. AI technology can help greatly with COVID-19 genome sequencing [8]. It further assists for develop vaccines and medicines for getting therapeutic and preventive agents for Covid19 [9] AI methods are affordable and agile. Computer vision can thus identify phenotypic features from medical images and provide recommendations for genetic testing similar to those made by an expert pathologist or dysmorphologist. The issues related to these are ethical considerations, the time required, and the complexity of data collection. Interpretability lacking is a prevalent issue in deep learning. AI algorithms should be flexible for all diseases [10].

Recently, the field of image recognition is paying greater attention to ReNets, which have various benefits. ResNets has the advantages of being able to train networks with a large number of depth layers with simplicity, which has not been affected by vanishing gradient issues, and for optimization learning from the residual mapping ie too difficult, readily convergent, and improved generalization. Thus, for image recognition modeling, BiLSTM is yet another reliable and promising Deep Learning (DL) model. The important aspects of this paper are.

- The genomic disorder detection model named EHO-WOA optimized Stacked ResNet-BiLSTM is proposed in this paper. It predicts various DNA deficiencies like Prader-Willi syndrome, Marfan syndrome, Early onset Morbid Obesity, Rett syndrome, and Angelman syndrome with high accuracy.
- Introducing a hybrid EHO-WOA algorithm for evaluating the fitness of the Stacked ResNet-BiLSTM architecture.
- The Gene expression phenotype and genotype are taken as the input data to the ResNet-BiLSTM design.
- The proposed model offers better accuracy, recall, specificity, precision, and f1-score to classify Angelman Syndrome, Rett Syndrome, and Prader-Willi Syndrome when evaluated with different datasets.

The rest of the paper is organized as below: Sect. 2 describes the related works to genomic disorder prediction. Section 3 describes the proposed EHO-WOA optimized Stacked ResNet-BiLSTM model method utilized for genome disorder prediction. Section 4 describes four datasets selected for justifying the proposed method and outcomes of these examinations conducted using the datasets. Section 5 concludes the paper.

2 Related Works

Machine Learning optimized polygenic scores for blood cell traits to identify sex-specific trajectories and genetic correlations with the disease were presented by Xu et al. [11]. A typical approach of thresholding ($P + T$) and pruning and learning methods of five are elastic nets (EN), multilayer perceptron (MLP), convolutional neural network (CNN), Bayesian ridge (BR), and LDpred2 were all tested for their relative performance in order to build Polygenic scores (PGS) through 26 blood cell features. Using INTERVAL and UK Biobank blood cell trait data this method improved PGS. The full allelic spectrum and genotype difficulty were not utilized in this method. Nasir et al. [12] introduced an advanced genome disorder prediction model empowered with deep learning. The development of an Advanced Genome Disorder Prediction Model (AGDPM) for the genome multiclass prediction disorder utilizing a lot of data was using the AlexNet as an efficient convolution neural network architecture. Testing and comparison of AGDPM with the pre-trained AlexNet neural network model show that AGDPM performs work efficiently, with testing and training accuracy of 89.89% and 81.25%. However, this method was not implemented in many genetic disorders.

To detect mitochondrial and multifactorial genetic disorders Rahman et al. [13] presented an IoMT-based Machine learning Model. Machine learning techniques Support Vector Machine (SVM) and K-Nearest Neighbors (K-NN) were utilized to train the model for the testing phase. By utilizing Performance matrices such as Sensitivity, Precision, F1-score, and Accuracy the results were evaluated. SVM and K-NN were compared. When compared to the KNN model, SVM was obtained with the highest training accuracy of 86.6% but SVM required more training time. DeepGestalt is a facial analysis framework that Gurovich et al. [14] created for the classification of genetic syndromes. Using deep learning techniques and a face recognition dataset, the system learns facial representation. The Casia-WebFace dataset was used to train the model initially, and it was then adapted to the syndromes domain with the use of verified images of patients. However, this model required more datasets to produce effective results.

French et al. [15] designed Whole Genome Sequencing (WGS) reports that genetic conditions are continuously in ill children. This approach was performed for the analysis of

the future group of families who were appointed PICU and NICU in the specific site and the whole genes were assigned uniformly as the phenotype of the children. The WGS was used to diagnose the most effective of detecting the particular gene condition through data interpretation. The experimental result outperformed the diagnosis of the genetic disease early, more accurately than the existing method. On the other hand, acceptance of these approaches could be difficult. Aromolaran et al. [16] designed an important gene detection in *Drosophila melanogaster* based on the sequence and functional features using the Machine Learning (ML) method. This model was used to predict the *Drosophila melanogaster* gene based on a variety of features such as gene networks, protein–protein interactions, functional annotations, evolutionary conservation, and nucleotide and protein sequences. Here, the performance matrices Receiver Operating Characteristic-Area under the ROC Curve (ROC-AUC), Precision Recall-Area under the ROC Curve (PR-AUC), and F1 score were treated in this method. However, this method does not contain more organisms.

Arlothet al. [17] utilized a Multivariate genotype–Phenotype association for directly integrating regularity information by “Deep WAS”. Which utilized, directly integrates predictions of a single variant into a multi-variant GWAS setting. The GWAS settings identify genetic variants associated with the quantitative disease. This technique utilized a chromatin feature in a cell type which leads to identifying individual SNPs in the non-coding region with gene regularity capacity with a joint contribution to disease risk. Meanwhile, large performance losses occurred, so more research on communication was still needed to avoid simple integration. Guo et al. [18] utilized a Deep Learning (DL) System to predict phenotype from Genome sequencing. The Deep metabolism to predict cell phenotypes from transcriptomics data. The settings improved to predict phenotypes with high accuracy and high speed. This model offers accuracy and speed value of ($pcc > 0.92$) and (< 30 min for > 100 GB data using a single GPU), and high robustness. This model Deep Metabolism to bridge the gap between genotype and phenotype and serve as a springboard. Meanwhile, the discussed method was not applied in field application.

3 Proposed Genetic Disorder Prediction Framework

Based on different genetic profiles, this study presents an EHO-WOA optimized Stacked ResNet-BiLSTM model to identify different rare genetic disorders such as Angelman Syndrome, Rett Syndrome, and Prader-Willi Syndrome. The proposed architecture integrates the BiLSTM with ResNet to extract the local abstract and temporal features. This section provides a brief description of the different optimization algorithms, stacked ResNet-BiLSTM architecture formation, and the Hybrid EHO-WOA framework utilized in this study.

3.1 Bi-directional Long Short-Term Memory (Bi-LSTM)

Long Short Term Memory (LSTM) that can store information in both the backward and forward directions is known as BiLSTM [25]. The BiLSTM can take into account both the future and past data at once by utilizing the connection after and before when understanding the context. Also, two LSTM layers are present in every memory block. Two hidden layer states are obtained by the backward and forward LSTM layers which are represented by $R'_t, t \in [L, 1]$ and $R_t, t \in [1, L]$. The output is then obtained by connecting the two states of the hidden layer. The input sequence’s future and past data are obtained via the backward and forward LSTM layers. The state of the hidden layer D_t contains backward \overleftarrow{d}_t and forward

\vec{d}_t at a time which is denoted as t . LSTM's input is $y = (y_1, y_2, \dots, y_L)$

$$\vec{d}_t = \overrightarrow{LSTM}(d_{t-1}, y_t, s_{t-1}), t \in [1, L] \tag{1}$$

$$\overleftarrow{d}_t = \overleftarrow{LSTM}(d_{t+1}, y_t, s_{t+1}), t \in [L, 1] \tag{2}$$

The prior and current states of a cell are represented by s_{t-1} and s_t .

$$D_t = \left[\vec{d}_t, \overleftarrow{d}_t \right] \tag{3}$$

The time series' length is represented by L .

3.2 Proposed Stacked ResNet-BiLSTM model

The fundamental ResNet-BiLSTM model must be designed prior to combining multiple fundamental models utilizing ensemble learning [26]. Figure 1 depicts the fundamental design of the ResNet-BiLSTM model. The Gene expression phenotype and genotype are the input data to the ResNet-BiLSTM design. Local feature extraction in Genome data using ResNet, BiLSTM extracted features examination for temporal dimension, and resultant regression forecast using a fully connected layer are the main elements in the Stacked ResNet-BiLSTM technique. ResNet, Dense, and BiLSTM are the 3 components in the fundamental ResNet-BiLSTM prediction technique which is also referred to as the 37-layer deep network.

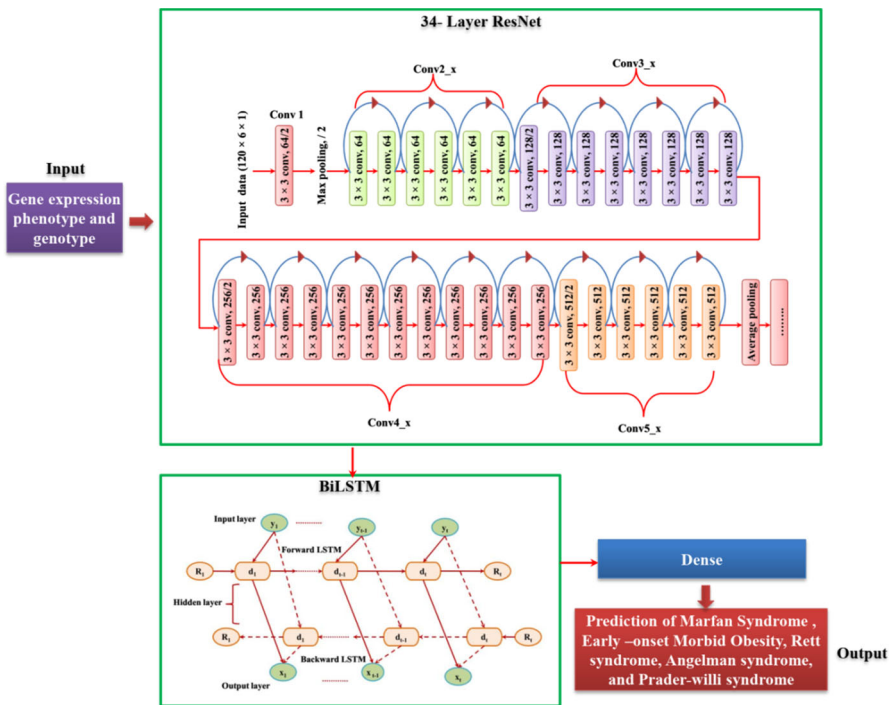


Fig. 1 Fundamental design of the ResNet-BiLSTM model

3.2.1 ResNet

To perform as a traditional CNN, the residual blocks (sequence of fundamental blocks) are used in the design of ResNet. A unique shortcut link construction is employed by the residual blocks for learning preferred mapping. The 2 stacked non-linear layers set the preferred basic mapping $I(a)$. Here, to generate residual learning, applying the identity mapping or skip link within the fundamental block consecutively avoids $H(a)$ transformation. Instead of learning $I(a)$ to map, residual function $H(a)$ based residual mapping is learned by the network. Here, the ResNet-34 model is utilized. This model is also referred to as the CNN with 33-layer as there is an elimination of a fully connected layer.

3.2.2 BiLSTMs

One of the optimized Recurrent Neural Network versions is the BiLSTM. It can forecast and process essential results over extensive delays and time series intervals. For the construction of a stacked BiLSTM structure, the bidirectional LSTM layers are incorporated. *size_of_the_input*, *batch_size*, and *time_step* are the 3 parameters in the design of BiLSTM. In progress, BiLSTM's batch size is determined as 50 and its size is the same as the whole model size. The value of the time step is determined as 1 for the BiLSTM within a stacked BiLSTM section. It represents the feature series length operated at once is fixed as 1. The value of 64 is determined as the *size_of_the_input*, which represents the feature sequence length as 64. Each and every LSTM layer in the BiLSTM produces a vector series as an outcome and the succeeding LSTM layer gets this output as the input. The most complicated time-series data representation is facilitated by those hidden layer hierarchies. Therefore, there is a need to obtain data at various levels.

3.2.3 Dense

The 2 fully connected layers (i.e., dense layers) are present in the 3rd section of the ResNet-BiLSTM design. The fundamental model predicts the output of Marfan syndrome, Early-onset Morbid Obesity, Rett Syndrome, Angelman Syndrome, and Prader-Willi Syndrome on the basis of the linear activation function. The loss function is considered as the Mean Square Error (MSE) in which the Elephant Herd optimization (EHO) and Whale Optimization Algorithm (WOA) optimizer are employed for optimizing the model. Up to 300 epochs, the fundamental model has been trained, in which the pre-processed input data is taken in various monitoring sites. Dropout, Regularization, and Early stopping are the optimization techniques that are followed to minimize overfitting and speed up model convergence. With respect to a specified ratio, the dropout function determines a few neurons to 0 in a random manner within a hidden layer. A number of essential features are learned by the model through adopting a dropout strategy, accordingly, the noise can be minimized. In the stacked BiLSTMs component, the second and first layer is used by the dropout, thus allowing its generalized potential.

In addition, based on the present validation output, the judgment has to be taken to terminate training. This is done prior to tests. Overfitting starts when the validation loss raises and the training loss reduces. In this situation, automatically the training is ended. Consider the Early stopping parameter *optimal_weights_restore = True*, and *patience = 30*. Thus, without the development in the model, the number of epochs to be accepted is determined and then, the model's weight values are stored in its optimum condition. For training each and

every model related to their dataset, the five-fold validation technique is employed. Subsequently, the stacked generalization approach is adopted to build the stacked ResNet-BiLSTM (Metamodel). For attaining regression and classification process with the best efficiency, the high-level method is adopted by the stacked generalization (which is the ensemble technique) for combining a number of fundamental models. For evading overfitting and generalizing the model's capability, every fundamental model uses k -fold cross-validation. Evaluating the fundamental models' predictable performance and balancing every sub-model involvement with the integrated forecasting considerably increases the model's standard performance. Figure 2 indicates the stacking process of the ResNet-BiLSTM. For making the stacked ResNet-BiLSTM architecture, the meta-learning device is the linear regression, and the sub-model is every individual ResNet-BiLSTM design. The below-explained steps are the exact executions:

1. Partitioning the original data as test dataset $M1$ and training dataset $N1$.
2. Using the training dataset, every fundamental model of stage 0 is trained, and using the test dataset those models are tested.
3. For the training meta-model of stage 1, the newest $N2$ training dataset is created. This new dataset is created by employing the related objective values and input as the validation dataset's forecasted values.

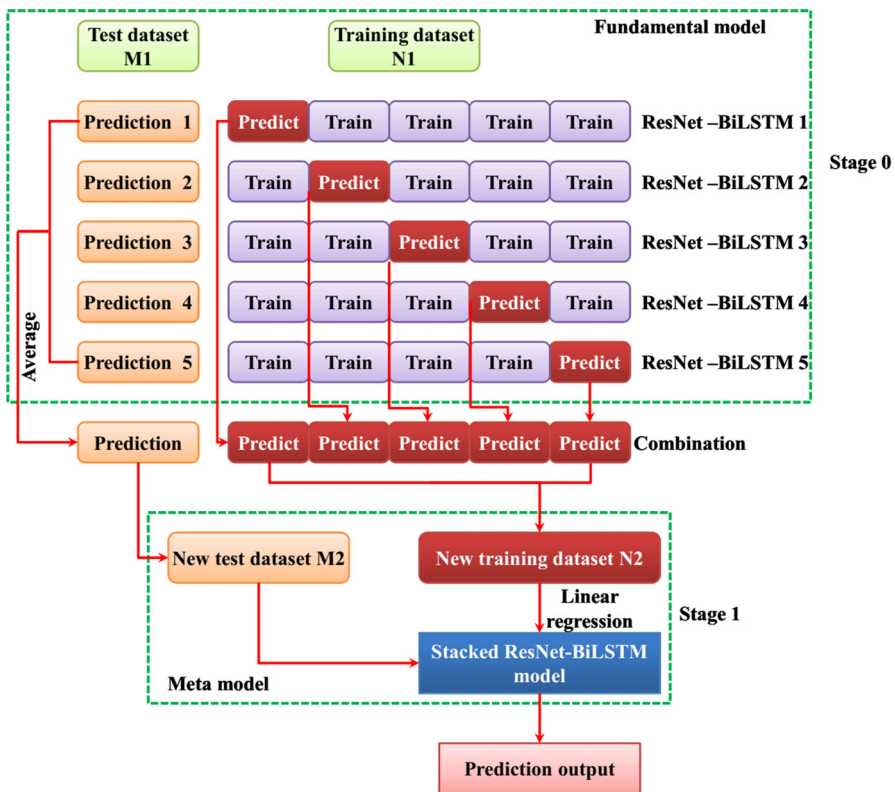


Fig. 2 Stacking process of the ResNet-BiLSTM model

4. For achieving the output prediction at a meta-model, the newest *M2* test dataset is created by averaging the *M1* test dataset's predictions.

The output of the ResNet-BiLSTM model is presented as follows:

- *Rett syndrome*: The development of the brain is impacted by the developmental condition and rare inherited neurological known as Rett syndrome. Language and motor skills gradually deteriorate as a result of this disorder. Women are more commonly impacted by Rett syndrome. The first six months of a baby's life look to be healthy, but as time goes on, speech, coordination, and hand use quickly deteriorate. Then, symptoms might remain stable for years. Despite the lack of a cure, physiotherapy, speech therapy, medication, and nutritional support can help control symptoms, avoid problems, and enhance the quality of life.
- *Prader-Willi Syndrome (PWS)*: A portion of chromosome 15 that was passed down from the father frequently gets deleted, leading to Prader-Willi syndrome. Prader-Willi syndrome is most frequently characterized by behavioral issues, short height, and intellectual disability. Delays in puberty and obesity-causing persistent appetite are indications of hormones. Although there is no treatment for Prader-Willi syndrome patients will gain from a restricted diet. Thus, hormone therapy is a treatment option.
- *Marfan syndrome (MFS)*: MFS is a heritable disorder, which is caused by a mutation in the fibrillin-1 (FBN1) gene. In this gene, more than 500 unique mutations were identified. Among the identified mutation, the environmental factors, and a high degree of penetration, intra, and inter-familial in this phenotype.
- *Angelman syndrome (AS)*: AS and interstitial duplication 15q autism were the genomic disorder that is caused by maternal deletion of the 15q11.2-q13 region. The UBE3A gene involves in the maternal phenotype which is caused by autism. The changes in chromatin and gene expression can be evaluated by int dup(15) and three reciprocal (AS) deletions which may be caused by both the AS and autism phenotypes. The changes in chromatin could influence genes by formaldehyde-assisted isolation of regularity elements (FAIRE) which leads to identifying 1104 regions of open chromatin in AS deletion. The Microarray research revealed 976 genes that were activated in int dup (15) versus AS deletion PBMC (p-value 0.05) and 1225 genes that were increased in int dup(15). Genes including UBE3A, ATP10A, and HERC2 at the 15q locus showed a significant outcome. As a result of their increased transcription in int dup(15) samples, many genes involved in chromatin remodeling, neurogenesis, and DNA repair were found at FAIR peaks in AS deletion samples. With the use of this investigation, a novel collection of genes and pathways related to the pathogenesis of both Angelman syndrome and autism have been discovered. As per the Affymetrix procedures, 100 ng of total RNA from each individual was used as the starting material for RNA synthesis and amplification. Internal chip control was used to normalize the signal data in the hybridized Affy Human Gene st v1 chip. The normalized data were then saved to a text file for further examination of expression using the EXPANDER program.
- *Early onset Morbid Obesity*: The signs of metabolic syndrome (MetS) and Obesity are raised among pediatric patients (teenagers and children) worldwide [24]. The danger of metabolic disorders and Cardiovascular disease (CVD) is increased at a later age because of the signs of MetS at a young age. Early onset morbid obesity is associated with the risk score of MetS risk in youth. An elevated MetS risk arises in early-onset obesity before 5 years of age. One of the public health issues worldwide is Obesity. High-density lipoprotein (HDL) cholesterol levels become low, largest waist circumference, elevated blood pressure, fasting hyperglycemia, and hypertriglyceridemia are the 5 risk factors of

MetS. Even though all these 5 risks are met by some children, one of these risks is met by at least 30% of children with obesity. Moreover, other than non-Hispanic white adolescents of 9%, the MetS occurrence is more in Hispanic adolescents (11%).

3.3 Elephant Herding Optimization

The Elephant herding behavior is parroted in the EHO [27]. There are some favored rules present in this algorithm that is listed below,

- The elephant population comprises a group and each family contains a certain number of elephants.
- The leader of the group is represented as a matriarch (which means a strong elephant). Each group contains a Matriarch.
- According to the guidance of the Matriarch, each elephant in the group were live together.
- At each generation, a certain number of elephants have left the group (worst candidate)

The work in the EHO algorithm takes place in two sections Group updating operator and separating operator.

3.3.1 Group Updating Operator

As stated before, there are several elephant groups in the population, and a fixed number of elephants are present in each group d .

Utilizing the below Equation, the elephants present in the group change their position in accordance with their matriarch $A_{ebest,c}^p$

$$A_{n,c}^{p+1} = A_{n,c}^p + \beta \times (A_{ebest,c}^p - A_{n,c}^p) \times r \tag{4}$$

From the above equation, $A_{ebest,c}^p$ is an updated value on the elephant n^{th} (where, $n = 1, 2, 3, 4, \dots, F$ depicted the total elephant present in the group) of the group c^{th} ($c = 1, 2, 3, 4, \dots, E$ represented total groups) for the following iteration, and the scaling factor depicted as $\beta \in [0, 1]$ that defines the matriarch’s influence $A_{ebest,c}^p$ on the present elephant $A_{n,c}$. The uniformly distributed random number is represented as r and it lies between the range of $[0,1]$. The below equation update the matriarch value ($A_{ebest,c}^p$)

$$A_{ebest,c}^{p+1} = \alpha \times A_{cen,c}^p \tag{5}$$

The scaling factor as mentioned above defines the group center effect $A_{cen,c}^p$ of a group leader’s updated position $A_{ebest,c}^{p+1}$ through the following iteration. By the below equation, the group center value is calculated

$$A_{cen,c} = \frac{1}{F} \times \sum_{n=1}^F A_{n,c} \tag{6}$$

Here, the c^{th} group elephant is depicted as $A_{n,c}$.

3.3.2 Separating Operator

The worst solution individuals are replaced by randomly initialized individuals using this operator. It expands the EHO algorithm’s population’s capacity for exploration and diversification. The below equation alters the worst individuals.

$$A_{eworst,c} = low_b + (up_b - low_b + 1) \times r \tag{7}$$

The individual upper and lower bound is represented as up_b and low_b . where $A_{eworst,c}$ is the group’s worst member of the c^{th} clan, who is to be altered by the new member who was chosen at random. As they upgraded their location based on the group center, the EHOs group updating operator prevents the group head from searching outside the group. The separation operator substitutes the weakest candidates with persons with random initialization. As a result, EHO has performance issues related to trapping and premature convergence into local optima. A good balance between exploitation and exploration is needed to avoid these issues. It is a challenging procedure, though, because population-based algorithms are stochastic. As a result, we proposed an enhanced EHO method in this study, which extended the capabilities of the original EHO algorithm.

3.4 Whale Optimization Algorithm

The WOA imitates humpback whales’ haunting demeanor [28, 29]. To reflect the demeanor of the whale the algorithm consists of two phases including the exploration phase and the exploitation phase. The exploration phase is defined as the process of seeking prey. Encircling the target and utilizing the bubble-net attack technique are the exploitation phases.

3.4.1 Exploitation Phases

The below equation depicts encircling prey humpback whale behavior. The algorithm’s leading solution is assumed to be the target prey. The remaining solutions will attempt to terminate the target prey.

$$\vec{V} = \left| \vec{\sigma} \cdot \vec{Y}_a(s) - \vec{Y}(s) \right| \tag{8}$$

$$\vec{Y}(s + 1) = \left| \vec{Y}_a(s) - \vec{\lambda} \cdot \vec{V} \right| \tag{9}$$

where \vec{Y}_a represented as a best-so-far solution, v indicates the current iteration, the coefficients vectors are the $\vec{\sigma}$ and $\vec{\lambda}$

$$\vec{\lambda} = 2\vec{b} \cdot \vec{q} - \vec{a} \tag{10}$$

$$\vec{\sigma} = 2 \cdot \vec{q} \tag{11}$$

where \vec{q} in $(0, 1)$ represented as a random vector, \vec{b} indicates a linear reduce coefficient, in the iteration from 2 to 0 which will reduce. By adopting $\vec{\lambda}$ and $\vec{\sigma}$ vectors the current position various positions compared to the best-so-far solution are controlled. The algorithm supposes that the best-so-far solution is prey based on Eq. (9), changes in humpback whales’ current position close to the prey, and simulates a situation in which the prey is surrounded. Two mathematical models are described as follows that simulate the humpback whale bubble-net attack.

1. Shrinking encircling mechanism: The value of the vector \vec{b} is linearly decreased to apply this model. The coefficient vector a fluctuation range, based on the random vector \vec{q} the vector b , is between $(-\vec{\lambda}, \vec{\lambda})$, \vec{b} decreases from 2 to 0 during iteration.
2. Spiral updating position: The humpback whale surrounds the prey in a logarithmic spiral motion after the model first determines the distance between itself and the prey. The below equation shows the mathematical model

$$\vec{Y}(s + 1) = \vec{V}' \times f^{\mu\delta} \times \cos(2\omega\delta) + \vec{Y}_a(s) \tag{12}$$

The distance between a humpback whale and its prey is represented by $\vec{V}' = |\vec{Y}_a(s) - \vec{Y}(s)|$. In $(-1, 1)$ δ is a random value. A constant called μ determines how the logarithmic spiral will look. The humpback whale will dive deep when the prey’s position is determined in the exploitation phase starting to produce spiral bubbles around the prey, and then moving upstream toward the surface. In order to update the humpback whale position during the iterative process, the hunting behavior supposes that both the shrinking circle and the spiral-shaped path have a similar implementation probability.

3.4.2 Exploration Phase

From the current best solution the algorithm pressures the solution to be far during the exploration phase and explores the search space at random. Because of this, the Whale Optimization Algorithm chooses the reference solution at random using a $\vec{\lambda}$ random value vector that is higher than +1 or less than -1. As a result of this technique and > 1 , the algorithm can carry out worldwide exploration. The below equation describes the mathematical model.

$$\vec{V} = \left| \vec{\sigma} \cdot \vec{Y}_{rand} - \vec{Y} \right| \tag{13}$$

$$\vec{Y}(s + 1) = \vec{Y}_{rand} - \vec{\lambda} \cdot \vec{V} \tag{14}$$

where \vec{Y}_{rand} is the solution’s position vector, which was selected at random from the current population.

3.4.3 Overview of Whale Optimization Algorithm

The WOA adaptive alteration of the coefficient vector $\vec{\lambda}$ enables it to transition between exploitation and exploration with ease. When the $|\vec{\lambda}| > 1$ exploration is carried out when $|\vec{\lambda}| < 1$ the exploration is carried out. The Whale Optimization Algorithm is capable of switching between a spiral-shaped path and a diminishing circle while the exploitation phase. $\vec{\lambda}, \vec{\sigma}$ the Whale Optimization Algorithm’s two main internal parameters, required to be changed.

3.5 Hybrid Elephant Herding Optimization- Whale optimization Algorithm

The aim of the hybrid EHO-WOA is to integrate the EHO’s fast convergence rate and WOA’s good searchability. Therefore, the challenge is how to create the best benefits of WOA and EHO. To address this issue, based on the value of fitness the population is split into two parts. The EHO carried out the population’s best half and the WOA carried out the population’s

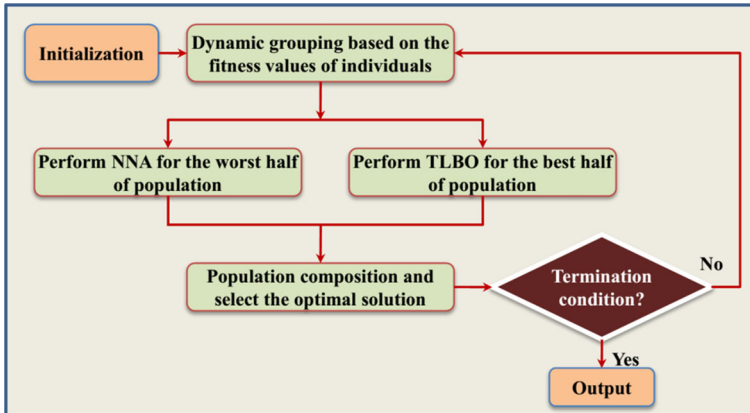


Fig. 3 Hybrid EHO-WOA framework

worst half. Figure 1 describes the framework of hybrid EHO-WOA. This algorithm has the three main steps of optimization listed in Fig. 3.

3.5.1 Dynamic Grouping Stage

According to the individual's fitness value after each looping, the population is split into two groups which mean a dynamic population in a hybrid EHO-WOA strategy. Dynamic grouping in hybrid EHO-WOA refers to the division of the population into two portions depending on individual fitness levels at the end of each loop. Using EHO, the population's best half is carried out and using WOA, the population's worst half is carried out. According to the further two reasons this stage is designed. Initially, similar behavior present in search operators may cause the search space to become less diverse. It is difficult to escape the solutions once they become stuck in a local minimum. The various search operators are present in EHO and WOA. Therefore, the hybrid EHO-WOA can improve the ability to escape the local optimal through a dynamic grouping step. In the Next step, EHO concentrates on local searches while WOA enhances global searches during the search process, which can aid hybrid EHO-WOA in striking an effective balance between exploitation and exploration.

3.5.2 EOA Stage

Hybrid EHO-WOA employs EHO to optimize the population's best half, which is mostly taken into account below. By optimizing the population's best half, this step takes full advantage of the WOA's rapid convergence speed, which can significantly speed up the convergence of the hybrid EHO-WOA.

3.5.3 WHO Stage

Hybrid EOA-WHO uses WHO to optimize the worst half of the population which depends on the below concern. In the account of WHO's superior global exploration, this phase not only somewhat prevents the population's worst half from entering local optimum solutions but also enables the discovery of better global optimal solutions.

3.5.4 Proposed EHO-WOA Implementation

The planned hybrid EHO-WOA implementation is explained in detail in this section. Following is a step-by-step description of how to execute hybrid EHO-WOA.

Step 1: Initialization:

This step is to initialize the parameters. The value M_{max} describes the function evaluation's maximum number, the size of the population is depicted as M , The designed variable's lower bound and upper bound is represented as lb , and ub , the problem dimension is represented as U the $f(.)$ is the fitness function. the number of iterations at a time p is currently 1, α the modification factor is 1, set the value to 0 for the function evaluation number $M_{current}$ Population initialization. Based on the parameter initialization a random population A^p is created.

Weight Matrix initialization. F represents the number $0.5 * N$. When using hybrid EHO-WOA, WOA is used to optimize F population members in each loop. Therefore the matrix weight V^p is a $F \times F$ square matrix.

Step 2: population evaluation

$$M_{current} = M_{current} + M \tag{15}$$

The objective function is used to calculate each individual's fitness values, and the best solution H^p is chosen. Therefore, the function evaluation's current number $M_{current}$ is upgraded by the above equation.

Step 3: Termination standard:

The method terminates if the function evaluation current number exceeds the maximum number of function evaluations. Else it goes to step 4.

Step 4: Dynamic grouping mechanism.

According to the fitness value, the individual populations are sorted. After sorting A^p population is described as \hat{A}^p . \hat{A}^p the best half of the population T^p and the \hat{A}^p worst of the population is O^p . Additionally, the optimal solution to speed up convergence is shared by T^p and O^p is explained as

$$A_{opt}^p = H^p \tag{16}$$

$$A_{best}^p = H^p \tag{17}$$

Step 5: Population optimization.

The WOA has two phases namely the exploration phase and the exploitation phase. The exploration phase is carried out from Eq. 8 through Eq. 12. Firstly encircling prey humpback whale behavior is represented by Eq. 8 and Eq. 9. Coefficient vectors $\vec{\lambda}$ $\vec{\sigma}$ are performed in Eqs. 10 and 11. The algorithm supposes that the best-so-far solution is prey based on Eq. 9. Equation 12 performed the bubble-net attack. The Exploration Phase is carried out from EqS. 13, 14. When the $|\vec{\lambda}| > 1$ exploration is carried out when $|\vec{\lambda}| < 1$ the exploration is carried out. Therefore the function evaluation current number can be upgraded by the below equation

$$M_{current} = M_{current} + 0.5 * M \tag{18}$$

EHO optimization contains two stages a Group updating operator and the Separating operator. The group updating operator is carried out from Eq. (4) to (6) and the separating operator is implemented in Eq. 7. Therefore the function evaluation current number can be upgraded by the below equation

$$M_{current} = M_{current} + M \quad (19)$$

Step 6: Population compositions:

O^{p+1} Population and T^{p+1} population are integrated into the A^{p+1}
 α^{p+1} is a modification factor and the iteration number p is upgraded by $p = p + 1$

Step 7: Continue to Step 3.

3.6 Parameter optimization of ResNet-BiLSTM model

On the basis of the ResNet-BiLSTM model classifier, EHO-WOA is taken as the integrated method. For achieving the most possible outcomes, experiment utilizing tenfold cross-validation, and $n = 1$, the number of nearest neighbors in the ResNet-BiLSTM model classifier. The WOA is integrated with the EHO algorithm to generate the proposed EHO-WOA. Here, every solution is computed using the below-expressed fitness function.

$$Fit = \alpha \rho_T(C) + \mu \frac{|T|}{|M|} \quad (20)$$

For balancing the number of chosen features and the accuracy of classification, the 2 parameters μ and α are introduced. The total number of features is denoted as $|M|$, and the chosen number of features for every solution is indicated as $|T|$. Then, for the employed classifier, the error rate of classification is represented as $\alpha \rho_T(C)$.

4 Results and Discussion

This paper developed a genomic disorder prediction method called EHO-WOA optimized Stacked ResNet-BiLSTM. In this work utilizing various comparative analyses and experimental analyses, the performance of our proposed method is evaluated. This is done in this section by selecting four datasets comprising diseases like Angel syndrome, Rett syndrome, Prader-Willi Syndrome, and Marfan syndrome. Following this, the prediction accuracy, sensitivity, specificity, and f-score outputs of the method are compared with those of the existing Support Vector Machine, AlexNet, K-Nearest Neighbour, and Convolutional Neural Network methods. Execution of this work is performed in software namely MATLAB. The parameter description is given in Table 1.

4.1 Dataset Description

This section mainly presents a brief description of the different datasets utilized in this study. Each dataset contains the gene expression profiles associated with different genetic disorders such as Rett syndrome, Angelman syndrome, Marfan syndrome, and Prader-Willi Syndrome.

Table 1 Description of parameters

Parameter	Learning rate
Learning rate	0.01
Dropout rate	0.5
No. of neurons in the LSTM hidden layer	64
Batch size	64
Word vector dimension	250

4.1.1 Length-Dependent Gene Misregulation in Rett Syndrome (MeCP2)

Any abnormalities in the MeCP2 lead to an autism-featured neurological disorder called Rett syndrome (RTT)[19]. This gene encodes the methyl DNA binding protein which acts as a transcriptional repressor. This misregulation of genes that happened in human RTT brains coincides with the start and the seriousness of symptoms in MeCP2 mutant mice and recommends the alteration of long gene expression leads to RTT pathology. Encoding proteins of long genes with neuronal functions and overlay with genes have connected with fragile X syndrome and autism. MeCP2 suppresses lengthy genes by binding methylated CA dinucleotides in the genes. Thus the MeCP2 gene mutation causes neurological dysfunction.

4.1.2 Angel Man Syndrome

It is otherwise called happy puppet syndrome or puppet-like syndrome [20]. The characteristics of this angel man syndrome are delayed development, severe speech impairments, problems in balancing and movement, and intellectual disability. RNF4 and UBE3A are the genes of humans associated with this disease. It also affected flies and rats.

4.1.3 A Marfan Syndrome Gene Expression Phenotype in Cultured Skin Fibroblasts

The mutation of the fibrillin-1 gene is a heritable connective tissue disorder [21] caused by MFS. It comprises identifiable aortic aneurismal subtypes and it accounts for five percent of thoracic aortic aneurysms. This analysis identified 4132 and ten genes selected for quantitative RT-PCR validation.

4.1.4 Sleep Abnormalities in Rare Genetic Disorders: Angel man Syndrome, Rett Syndrome, and Prader-Willi Syndrome—ARP 5207 (Human)

Sleep is crucial for a healthy life. Any disruption in sleep causes negative impacts on the health and brain development, and brain function of children [22]. People with Rett Syndrome (RTT), Angelman Syndrome (AS), Early-onset Morbid Obesity (EMO), and Prader-Willi Syndrome (PWS) mostly have sleep disorders. Thus it causes health disorders related to sleep and will cause further problems with memory and learning. This study has the following key objectives: Analyze the sleeping habits of people with early-onset morbid obesity, Angelman syndrome, Rett syndrome, and Prader-Willi syndrome. and comparing normal people's sleep behavior with the above disease-affected people.

4.2 Performance Metrics

Various parameters utilized for accessing the genome sequencing capability of the method are given below from Eqs. 21–25. These metrics are accuracy(P_A), specificity(P_{SP}), sensitivity(P_{SE}), precision(P_{PR}), and -score(P_F). The four classes utilized in the below equation are true positive(P_{TP}), true negative(P_{TN}), false positive(P_{FP}), and false negative(P_{FN}).

$$P_A = \frac{P_{TP} + P_{TN}}{P_{TP} + P_{TN} + P_{FP} + P_{FN}} \tag{21}$$

$$P_{SP} = \frac{P_{TN}}{P_{TN} + P_{FP}} \tag{22}$$

$$P_{SE} = \frac{P_{TP}}{P_{TP} + P_{FN}} \tag{23}$$

$$P_{PR} = \frac{P_{TP}}{P_{TP} + P_{FP}} \tag{24}$$

$$P_F = 2 \times \frac{P_R * P_{PR}}{P_R + P_{PR}} \tag{25}$$

4.3 Comparative Analysis

The cross-validation technique is utilized for assessing the proposed EHO-WOA optimized Stacked ResNet-BiLSTM method which holds the parameter K. IT represents the number of segmented data sample groups. Otherwise, this process is recognized as K-fold cross-validation. During training, this method split the dataset into K folds within a particular time. In this approach, ten-fold and five-fold cross-validation is selected to evaluate the precision-recall (precision value plotted on Y-axis and recall value plotted on X-axis) curve and ROC curve (Plotted for True positive values versus false positive values). Figures 4 and 5 represents the evaluation done based on the precision-recall and receiver operating curve. The ROC value achieved is 0.9958 and the precision-recall curve achieved the value of 0.9852.

Figure 6 clearly explains the training accuracy of the proposed EHO-WOA optimized Stacked ResNet-BiLSTM method. The training accuracy is analyzed for 20, 30, 50, 60, and 80 epochs until the 100 epochs. For an effective system, the training accuracy must possess

Fig. 4 Cross-validation results for ROC curve

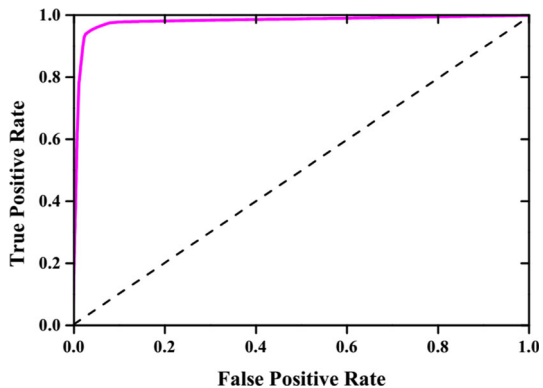


Fig. 5 Cross-validation results for precision-recall curve

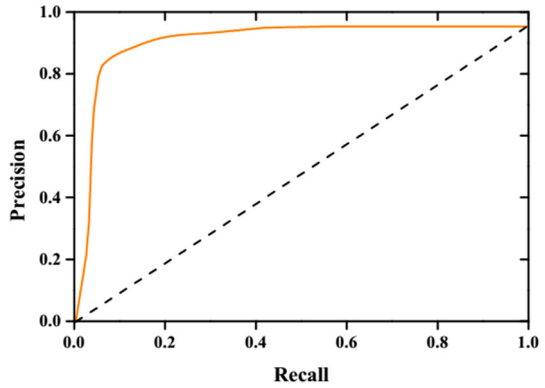
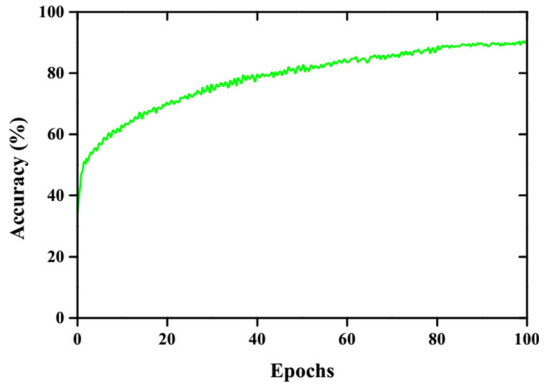
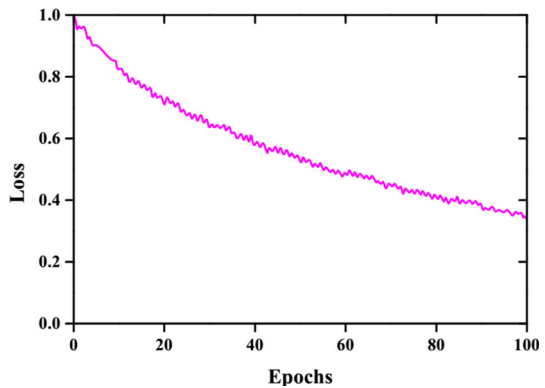


Fig. 6 Training accuracy of the proposed method



higher values and the loss should be the lowest. The observation results showed that the method has the best training accuracy with 98.524%. The training loss that occurred during the training of the EHO-WOA optimized Stacked ResNet-BiLSTM method is shown in Fig. 7. Training loss refers to the errors that occurred during the training of a particular method. Loss is also calculated for various epochs till the hundred epochs. It has less training loss with a

Fig. 7 Training loss of the proposed method



value of 9.52%. With the highest training accuracy and the lowest training loss, the proposed genome sequencing method has shown better performance.

Accuracy measures are the most direct method for predicting the efficiency of any method. Here the method is assessed accuracy metric for evaluating the diagnosing efficiency of the proposed EHO-WOA optimized Stacked ResNet-BiLSTM method while predicting genomic diseases. Accuracy is calculated by matching the predicted diseases with actual classes. The accuracy of the proposed method is higher with a value of 99.85% while equating with the proposed methods while utilizing all diseases in all datasets and is depicted in Fig. 8. The accuracy of the existing methods using the same datasets is given as follows: The AlexNet has the second-highest accuracy of 92.52%, and the KNN and CNN have the third and fourth-highest accuracy performance with outputs of 85.52% and 83.36%. The SVM has the lowest accuracy of 80.52. Thus the EHO-WOA optimized Stacked ResNet-BiLSTM method has the highest prediction accuracy.

The graphical analysis of the precision measurement of the proposed EHO-WOA optimized Stacked ResNet-BiLSTM model and various existing models is illustrated in Fig. 9. This performance metric refers to the correctly diagnosed diseases to the actual positive labels. The precision value of the proposed method is 98.52. The precision of the existing methods like SVM, CNN AlexNet, and KNN is 89.52%, 86.96%, 79.95%, and 84.041% respectively. Comparing the precision values of all the methods the EHO-WOA optimized

Fig. 8 Comparative accuracy of the accuracy

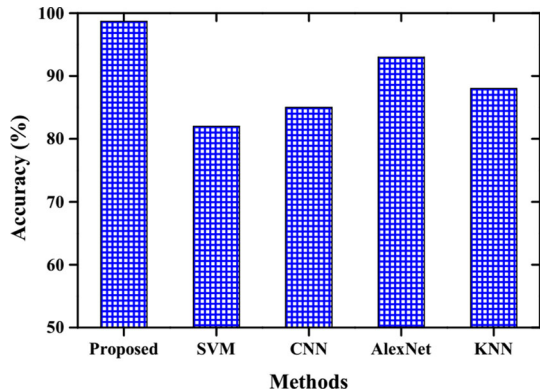


Fig. 9 Comparative accuracy of precision

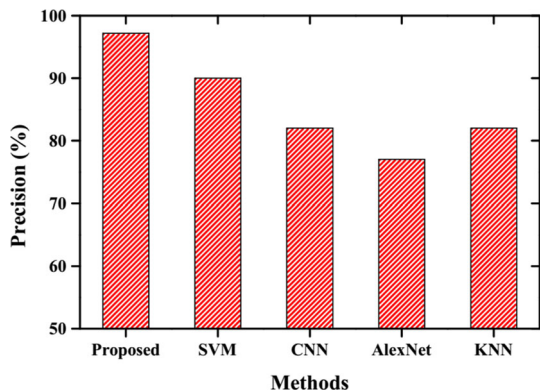
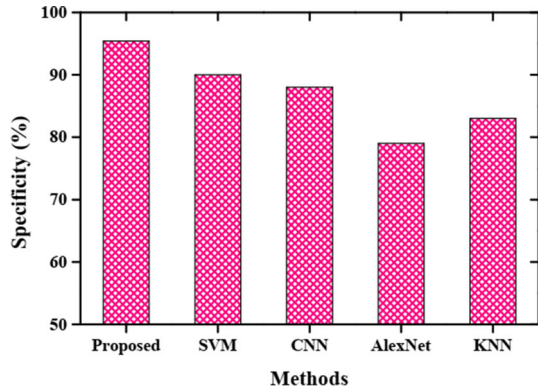


Fig. 10 Specificity analysis



Stacked ResNet-BiLSTM method has a higher precision value. Thus it predicts diseases more precisely than the other current methods.

Figure 10 shows the graphical representation of specificity analysis. In Fig. 10 the proposed EHO-WOA optimized Stacked ResNet-BiLSTM method is compared with the existing methods such as CNN, AlexNet, SVM, and KNN techniques. The EHO-WOA optimized Stacked ResNet-BiLSTM method has high specificity compared to existing methods. The proposed method achieved a specificity of 96%. Figure 11 indicates the Sensitivity analysis. In Fig. 11 we compare existing methods like CNN, AlexNet, SVM, and KNN techniques with the proposed method. The EHO-WOA optimized Stacked ResNet-BiLSTM method has high sensitivity when compared to existing methods. Our proposed method achieved 95% of sensitivity.

Figure 12 represented the Receiver Operating Characteristic (ROC) curve. We compare the existing methods such as CNN, AlexNet, SVM, and KNN techniques to the proposed EHO-WOA optimized Stacked ResNet-BiLSTM methods. The proposed approach has a high value when compared to the existing methodology. The ROC score for the proposed methodology is 0.98. The f-score is shown in Fig. 13. In Fig. 13 the proposed method is compared with the existing methods namely CNN, AlexNet, SVM, and KNN techniques. Comparing existing methods the proposed EHO-WOA optimized Stacked ResNet-BiLSTM method has a high f-score value of 98%.

Fig. 11 Sensitivity analysis

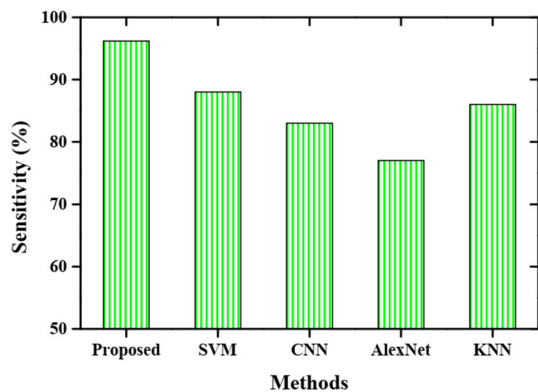


Fig. 12 ROC curve

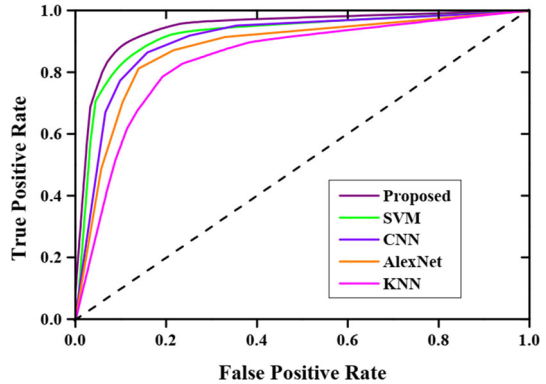
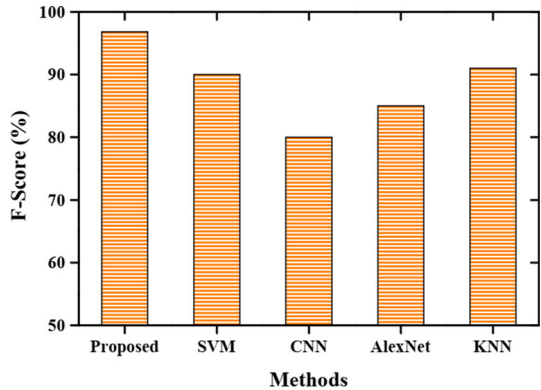


Fig. 13 F-score



5 Conclusion

To predict the genomic disorder we propose a new method called EHO-WOA optimized Stacked ResNet-BiLSTM. We integrate the EHO with the WOA for effective analysis and prediction of genomic disorder. Utilizing the datasets, MeCP2, Angelman syndrome, A Marfan syndrome gene expression phenotype in cultured skin fibroblasts, and Sleep Abnormalities in Rare Genetic Disorders dataset we forecast the diseases such as Rett syndrome, PWS, the MFS, Early onset Morbid Obesity and AS. The performance metrics such as Accuracy, Recall, specificity, precision, and f1-score are used for validating the result. The experimental findings of the aforementioned measures are compared with those of the various baseline approaches used in genomic prediction techniques, such as SVM, CNN AlexNet, and KNN. The value of the proposed method Area under the curve (AUC) is 0.98. The training accuracy and the training loss of EHO-WOA optimized Stacked ResNet-BiLSTM is 98.5 and 9.5. We planned to implement genomic disorder prediction in animals as a part of our future work.

Acknowledgements The authors are grateful to all respondents who participated in this study and to the data collectors for their contribution.

Author Contributions All authors who participated in data analysis, drafting or revising the manuscript gave approval of the final version to be published.

Funding This study has received no external funding.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

Informed Consent Consent was secured from all of the respondents who participated in the study.

References

1. Howles CM (1996) Genetic engineering of human FSH (Gonal-F®). *Hum Reprod Update* 2(2):172–191
2. Zhang F, Rao S, Cao H, Zhang X, Wang Q, Xu Y, Sun J, Wang C, Chen J, Xu X, Zhang N (2021) Genetic evidence suggests posttraumatic stress disorder as a subtype of major depressive disorder. *J Clin Investig*. <https://doi.org/10.1172/JCI145942>
3. Van El CG, Cornel MC, Borry P, Hastings RJ, Fellmann F, Hodgson SV, Howard HC, Cambon-Thomsen A, Knoppers BM, Meijers-Heijboer H, Scheffer H (2013) Whole-genome sequencing in health care. *Eur J Hum Genet* 21(6):580–584
4. Hicks AL, Wheeler N, Sánchez-Busó L, Rakeman JL, Harris SR, Grad YH (2019) Evaluation of parameters affecting performance and reliability of machine learning-based antibiotic susceptibility testing from whole genome sequencing data. *PLoS Comput Biol* 15(9):e1007349
5. Ren Y, Chakraborty T, Doijad S, Falgenhauer L, Falgenhauer J, Goesmann A, Hauschild AC, Schwengers O, Heider D (2022) Prediction of antimicrobial resistance based on whole-genome sequencing and machine learning. *Bioinformatics* 38(2):325–334
6. Ng PC, Kirkness EF (2010) Whole genome sequencing. *Genet Var*. pp 215–226
7. Dias R, Torkamani A (2019) Artificial intelligence in clinical and genomic diagnostics. *Genome Med* 11(1):1–12
8. Ahmed I, Jeon G (2022) Enabling artificial intelligence for genome sequence analysis of COVID-19 and alike viruses. *Interdiscip Sci Comput Life Sci* 14(2):504–519
9. Bagabir S, Ibrahim NK, Bagabir H, Ateeq R (2022) Covid-19 and Artificial Intelligence: Genome sequencing, drug development, and vaccine discovery. *J Infect Public Health*
10. Libbrecht MW, Noble WS (2015) Machine learning applications in genetics and genomics. *Nat Rev Genet* 16(6):321–332
11. Xu Y, Vuckovic D, Ritchie SC, Akbari P, Jiang T, Grealey J, Butterworth AS, Ouwehand WH, Roberts DJ, Di Angelantonio E, Danesh J (2022) Machine learning optimized polygenic scores for blood cell traits identify sex-specific trajectories and genetic correlations with disease. *Cell Genomics* 2(1):100086
12. Nasir MU, Gollapalli M, Zubair M, Saleem MA, Mehmood S, Khan MA, Mosavi A (2022) Advance genome disorder prediction model empowered with deep learning. *IEEE Access* 10:70317–70328
13. Rahman AU, Nasir MU, Gollapalli M, Alsaif SA, Almadhor AS, Mehmood S, Khan MA, Mosavi A (2022) IoMT-based mitochondrial and multifactorial genetic inheritance disorder prediction using machine learning. *Comput Intell Neurosci*. <https://doi.org/10.1155/2022/2650742>
14. Gurovich Y, Hanani Y, Bar O, Nadav G, Fleischer N, Gelbman D, Basel-Salmon L, Krawitz PM, Kamphausen SB, Zenker M, Bird LM (2019) Identifying facial phenotypes of genetic disorders using deep learning. *Nat Med* 25(1):60–64
15. French CE, Delon I, Dolling H, Sanchis-Juan A, Shamardina O, Mégy K, Abbs S, Austin T, Bowdin S, Branco RG, Firth H (2019) Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. *Intensive Care Med* 45(5):627–636
16. Aromolaran O, Beder T, Oswald M, Oyelade J, Adebisi E, Koenig R (2020) Essential gene prediction in *Drosophila melanogaster* using machine learning approaches based on sequence and functional features. *Comput Struct Biotechnol J* 18:612–621
17. Arloth J, Eraslan G, Andlauer TF, Martins J, Iurato S, Kühnel B, Waldenberger M, Frank J, Gold R, Hemmer B, Luessi F (2020) DeepWAS: Multivariate genotype-phenotype associations by directly integrating regulatory information using deep learning. *PLoS Comput Biol* 16(2):e1007616
18. Guo W, Xu Y, Feng X (2017) DeepMetabolism: a deep learning system to predict phenotype from genome sequencing. *arXiv preprint arXiv:1705.03094*
19. Gabel HW (2022) GSE60074 - length-dependent gene misregulation in Rett syndrome (mecp2). *OmicsDI*. Retrieved November 23, 2022, from <https://www.omicsdi.org/dataset/geo/GSE60074>

20. Alliance of Genome Resources (2022) Retrieved November 24, 2022, from <https://www.alliancegenome.org/disease/DOID:1932>
21. Ruzzo WL (2022) GSE8759 - A Marfan syndrome gene expression phenotype in cultured skin fibroblasts. OmicsDI. Retrieved November 24, 2022, from <https://www.omicsdi.org/dataset/geo/GSE8759>
22. U.S. National Library of Medicine (2022) Sleep abnormalities in rare genetic disorders: Ang (ID 371673). National Center for Biotechnology Information. Retrieved November 24, 2022, from <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA371673>
23. Yao Z, Jaeger JC, Ruzzo WL, Morale CZ, Emond M, Francke U, Milewicz DM, Schwartz SM, Mulvihill ER (2007) A Marfan syndrome gene expression phenotype in cultured skin fibroblasts. *BMC Genomics* 8(1):1–13
24. Pacheco LS, Blanco E, Burrows R, Reyes M, Lozoff B, Gahagan S (2017) Peer reviewed: early onset obesity and risk of metabolic syndrome among chilean adolescents. *Prevent Chronic Disease*, 14
25. Peng T, Zhang C, Zhou J, Nazir MS (2021) An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* 221:119887
26. Cheng X, Zhang W, Wenzel A, Chen J (2022) Stacked ResNet-LSTM and CORAL model for multi-site air quality prediction. *Neural Comput Appl*, pp 1–18
27. Singh H, Singh B, Kaur M (2021) An improved elephant herding optimization for global optimization problems. *En Comput*, pp 1–33
28. Lee CY, Zhuo GL (2021) A hybrid whale optimization algorithm for global optimization. *Mathematics* 9(13):1477
29. Zhang Y, Jin Z, Chen Y (2020) Hybrid teaching–learning-based optimization and neural network algorithm for engineering design optimization problems. *Knowl-Based Syst* 187:104836

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.