

DMADL-CTO: Hybrid Distributed Mixed Attention-based deep learning model for Cerebellar Ataxia Detection

Dr. Edara Sreenivasa Reddy¹, Mr. Sunil Prem Kumar Prathipati^{2*}

¹Professor, Department of Computer Science and Engineering (SCOPE), VIT Amaravathi, India.

^{2*}Research Scholar, Department of Computer Science and Engineering, Acharya Nagarjuna University, India.

Email: sunilpremnelson@gmail.com

ARTICLE INFO	ABSTRACT
Received: 15 Nov 2024	<p>Cerebellar ataxia (CA) is a disability disease that originated in the cerebellum of the human brain and caused several awkward condition as in-organized body balances, eye movements, inability to gait, and extremities. Because of these symptoms, early detection of CA is mandatorily targeted by the researchers using different conventional methods. These existing detection methods faced various disadvantages, as poor performance effectiveness, complexity problems, increased error rate, and computational time consumption. To overcome these issues in the existing methods, a proposed model is designed as Hybrid Distributed Mixed Attention-based Deep Learning approach enabled Competitive Tuning Optimization (Hybrid DMADL-CTO) model that precisely detects the abnormalities along with its class labels. The developed model possessed a distributive learning approach that captures detective outcomes with minimal operational time. Meanwhile, the model's ability and effectiveness are effectually enhanced by the integration of Competitive tuning optimization (CTO) and mixed attention mechanisms. The developed algorithm eliminates the optimization problem and attained better durability, consistency, and convergence rate significantly. Additionally, the mixed attention mechanism minimized the complexity problem and achieved effective performance efficacy. Thus, the effectiveness of the research model is compared with other related techniques based on the performance metrics, such as accuracy, recall, and precision for which the proposed method acquired 97.67%, 97.99%, and 95.89% respectively.</p> <p>Keywords: Neurological diseases, Distributive learning, Gait analysis, Cerebellar Ataxia, and Frame selection.</p>
Revised: 20 Dec 2024	
Accepted: 08 Jan 2025	

INTRODUCTION

In human anatomy, the cerebellum plays a significant role in the brain's nervous system, which effectively controls voluntary movement coordination, eye movement, and maintains the balance of muscles. In this context, sudden damage causes control loss, cerebellar functional damage, and impairments of body parts, which is termed as cerebellar ataxia. The impairments of body parts include limb in-coordination, speech impairment, eye movement abnormalities, and gait instability [1] [2] [3] [4][5]. In addition to this, the CA caused abnormalities in precise movement that led to disability for doing simple tasks like walking, eating, and dressing. Moreover, the major clinical abnormalities are dysarthria, ophthalmoparesis, dysphagia, peripheral neuropathy, nystagmus, and hyposmia. At the initial stage of cerebellar ataxia, it impacts the daily activities of affected patients. Similarly, at the moderate stage that caused difficulty in walking with the support of some walking aid. Comparably, at the advanced stage that caused severe impact by carrying the affected patient wheelchair-bound and needing other support to do daily activities [6]. To avoid these obstacles, early recognition is important and provides proper treatment for patients, who suffer from CA [7].

In general, CA is diagnosed by numerous imaging techniques with the inclusion of infections, neoplastic, inflammatory, toxic, vascular, paraneoplastic, degenerative etiologies, and hereditary. Furthermore, the traditional imaging techniques for the detection of these infections might lead to huge challenges in detection management, increased computational cost, and high time consumption. Therefore, the obtained limitation in the traditional methods was effectively overcome by machine learning (ML) techniques, which provide a high-order evaluation with active mining [8]. The most frequent ML approaches used in the detection of CA are Naïve Bayes (NB), Logistic

Regression (LR), Adaptive boosting, Stochastic Gradient Descent (SGD), Random Forest (RF), and Decision Tree (DT) classifiers [9]. Based on these ML techniques, the model provides better detection accuracy by resolving prediction problems effectively [10][11]. However, the ML approaches had some interpretability issues during the detection of cerebellar ataxia, to tackle these challenges, Deep learning (DL) methods were introduced significantly. Some of the DL-based methods are Bidirectional Long Short-Term Memory (BiLSTM), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Neural Network (DNN), and Extreme gradient boosting (XGB) [12][13], which were implemented in the detection of Cerebellar ataxia.

Furthermore, the integration of traditional methods required extra manual support to attain better prediction results. Moreover, the ML-based detection methods also suffered from complexity issues. In NB evaluation, the model faced irrelevant and imbalanced feature issues that did not generate better detection results [14][15]. Similarly, the classification approaches of ML-based techniques have the challenges of imbalanced data collection, lack of training, generalization problems, and decreased accuracy [16] [17][18]. These challenges are overcome by DL-based approaches. Furthermore, the DL-based techniques also suffered from some disadvantages such as high computational complexities, increased memory resources, poor prediction results, and imbalance issues. In addition, the irrelevant feature issues are resolved by advanced strategies as gait analysis, spatiotemporal features, random sampling, and other methodologies [19]. Because of these integrations, the model provides better improvement in disease prediction and achieves better detection accuracy [20][21]. Therefore, the challenges of existing ML and DL-based techniques are effectively addressed by the newly developed model that provides better detection results with increased ataxia severity classification and feature importance score.

The Hybrid DMADL-CTO significantly detects and classifies the abnormalities of CA with the inclusion of a standard classifier approach along with the advances of developed optimization algorithms and attention mechanisms. In this context, the detection process is performed under optimal frames, which are selected by an optimized multi-metric frame selection method. Further, the pre-processing and ROI extraction methods are performed accordingly. Along with this, the main objective for the research is given as follows.

- **Optimized Multi-metric frame selection:** To perform an accurate detection process, optimal key frames are selected with the integration of the CTO algorithm. In this, the competitive strategy and learning factor capability are effectively combined to identify the optimal frames. Thus, the obtained keyframes are preceded into further detection process.
- **Mixed attention-based Hybrid Distributed Deep Learning approach enabled Competitive Tuning Optimization:** The insertion of distributed learning in the CNN-LSTM standard classifier enables accurate abnormalities detection along with its class labels. In this method, the operational time of the model is reduced effectively, and shows better performance in CA detection. Moreover, the ability and scalability of the model are improved by the integration of mixed attention mechanisms that also reduce the overlapping problems and complexity issues effectually.

The remainder part of the research article is scheduled as follows, Section 2 intercepts the prominent related works along with the approaches and research scope. Section 3 describes the developed methodology for detecting CA disorder along with the execution of mathematical modelling. Further, section 4 explores the experimental setup and performance effectiveness of the proposed model, additionally comparing the achieved results with other existing methods. Finally, section 5 summarizes the conclusion of the research and presents the future scope directions.

2. LITERATURE REVIEW

The section summarizes some of the prominent existing methods and provides a detailed explanation of their advances and limitations.

M. Shanmuga Sundari, and Vijaya Chandra Jadala, [22] introduced an ensemble ML technique to detect the abnormalities of CA. In the method, the severity of the disease was effectively classified and attained better accessibility and efficacy. However, the inherent variants of the ML approaches consumed poor sensitivity and stability that resulted in minimal detection accuracy. Pavithra Durganivas Seetharama, and Shrishali Math, [3] implemented an extreme gradient boosting method along with the integration of feature selection and ranking

approaches (FSR-XGB) for detecting the CA disorder. In this method, the misclassification issues and optimization error problems were eliminated precisely. However, the performance efficacy and detection accuracy were reduced, which was caused by the cross-validation feature mechanism.

Massimiliano Pau, et al. [23] developed a quantitative characteristic mechanism for detecting the disturbances of severe gait in CA neurological diseases. In this method, the progression and rehabilitation of the disease were monitored precisely and attained better performance efficacy. However, the quantitative mechanism suffered from generalizability, scalability, and computational complexity issues. Karan Bania, and Tanmay Tulsidas Verlekar, [4] presented a Graph Convolutional Network enabled with the utilization of an ataxia dataset. Thus, the developed model is termed as AtGCN model, which effectually strengthens the detection accuracy and provides better prediction results accurately. In this evaluation, the model faced high error loss that affected the training capability and validation time adequately.

Dante Trabassi, et al. [24] employed a conditional tabular Generative Adversarial Network (ctGAN) to capture the applicability of hereditary CA. The model provided better ability and consistent explainability in disease diagnostic performance. Meanwhile, the dependability and robustness of the model were decreased during the evaluation of the detection mechanism. Kyriakos Vattis, et al. [25] established a CNN model to garner the effectiveness of cerebellar ataxia. The model utilized frequency and time partial derivatives for evaluating the manifestation of ataxias. Based on these derivative techniques, the model significantly determined the multi-context with dynamic encapsulation. However, the CNN classifier suffered from scalability, lack of leveraged information, decreased quantitative assessment, and poor performance ability during early diagnosis.

Jennifer Faber, et al. [26] introduced a fastsurferCNN model combined with a U-Net segmentation network (CerebNet) for detecting and segmenting the CA. The CerebNet method was very compact and consumed minimal GPU validation time. Further, the data augmentation technique effectually addressed the replicability challenge in small-scale datasets. While performing the detection process under the complied cross dataset, the generalizability, reliability, accuracy, and sensitivity were decreased significantly. Robin Cabeza-Ruiz, *et al.* [8] developed a CNN classifier to detect the abnormalities of CA from magnetic resonance images of the brain. Additionally, the CNN method detected the abnormality classes with high-performance efficacy in a precise manner. During evaluation, the model affected the estimated volume loss and caused generalizability problems. Due to this, the operational time for the detection process was increased enormously.

2.1 Challenges

The drawbacks obtained in the related methods are furnished below,

- While detecting the abnormalities of CA, the CerebNet model consumed minimal efficacy of detection accuracy, operating speed, reliability, and sensitivity. Due to this, the generalizability and interpretability of the model were also reduced gradually [27].
- During the evaluation of early diagnosis under the utilization of the CNN classifier, the leveraged information of the disease was not acquired effectively. Because of this, the model did not provide precise detection outcomes and consumed a lot of time for evaluation.
- The CNN model suffered from certain disadvantages, as high-volume loss, decreased performance efficacy, large time consumption, and complexity issues.
- The major limitation of AtGCN was high error loss, due to this, the ability of the model was reduced appropriately and caused an impact on interpretability and generalizability.
- The FSR-XGB method obtained high false positive rate with minimal performance efficacy because of the integration of the cross-validation process.

2.2 Problem statement

CA is a neurological disorder that occurs in the cerebellum of brain activity, which affects the movement of muscles and other body parts effectively. Mainly, CA causes walking disabilities and getting paralyzed more easily, and these issues are significantly overcome by early diagnosis methods. Numerous conventional methods were built to detect

the abnormalities of CA, but these methods faced certain challenges, as lack of training, imbalance data collection, high-cost consumption, complexity issues, overfitting risk, lower model performance, and generalizability problems [28]. To address these limitations, a Hybrid DMADL-CTO is designed in this research that provides desired detection results. Meanwhile, the model is integrated with advanced feature strategies as gait pose kinematic structural flowmap features, and mixed attention mechanism. According to this context, the developed Hybrid DMADL-CTO reduced the operational time and training execution by the distributed learning method. Let the input for the research be garnered from the READISCA dataset [29], which is mathematically notated as,

$$G = \sum_{b=1}^r k_b \quad (1)$$

where, G indicates the video dataset, k_b denotes the input video, and r mentions the total video. From this video sequence, instructive frames are selected precisely based on the evaluation of Structural Similarity Index Measure (SSIM), Peak Signal-to-Noise Ratio (PSNR), and Signal-to-Interference-Noise Ratio (SINR) similarity measures. These similarity measures are integrated with the CTO algorithm for selecting the frames to perform further processing. Thus, the selected optimal frames are represented as,

$$D(G) = \min \left(\text{Threshold} - \left(\frac{SSIM + PSNR + SINR}{3} \right) \right) \quad (2)$$

where, $D(G)$ refers the optimal frames from the given input video. In this research, the threshold limit is given as 0.8 for performing the detection process. Therefore, the entire selected frames are illustrated as,

$$D(k_b) = \{D_1, D_2, \dots, D_i, \dots, D_v\} \quad (3)$$

where, k_b denotes the input video, D_i represents the i^{th} frames, and D_v indicates the total frames. Further, the selected frames tend towards pre-processing and ROI extraction methods to eliminate the redundancy and noise attributes effectively. Meanwhile, a mixed attention mechanism is combined with a distributive learning approach to achieve desired detection outcomes with minimal occurrence of overfitting issues and time complexity problems. Based on these integrated mechanisms, the Hybrid DMADL-CTO effectively detects the behavior as normal or abnormal. Thus, the performance of the Hybrid DMADL-CTO is strengthened by reducing the categorical cross-entropy loss function, which is expressed as,

$$Ca.loss = - \sum_{b=1}^r t_b \log(\hat{t}_b) \quad (4)$$

where, t_b defines the true probability distribution, \hat{t}_b represents the predicted probability distribution. By reducing the attained loss function, the Hybrid DMADL-CTO effectively predicts the class as,

$$t_b = \begin{cases} 0, & \text{for } k_b \text{ is normal} \\ 1, & \text{for } k_b \text{ is abnormal with slight cerebellar ataxia} \\ 2, & \text{for } k_b \text{ is abnormal with moderate cerebellar ataxia} \\ 3, & \text{for } k_b \text{ is abnormal with advanced cerebellar ataxia} \end{cases} \quad (5)$$

where, t_b mentions the actual class types that range as $(0-4)$.

3. SYSTEM MODEL FOR CEREBELLAR ATAXIA DETECTION

Figure 1 illustrates the system model for CA disorder detection, in which the detection process is achieved with the determination of input stage, frame selection stage, pre-processing stage, feature extraction stage, model training stage, and outcome stage. In this context, the optimal frames are selected from the obtained input video sequences to reduce unnecessary repetition. The selected frames are further furnished into the pre-processing stage to eliminate the noise distortions, and the instructive features are extracted from the feature extraction stage. After getting the informative features, which are transferred into the model training phase where effective performance is achieved and showed precise detection results at the outcome stage.

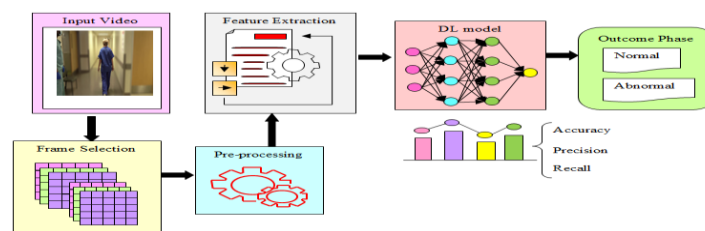


Figure 1: System model for cerebellar ataxia

3.1 Hybrid Distributed Mixed Attention-based Deep Learning approach enabled Competitive Tuning Optimization for cerebellar ataxia detection

The main objective of the research is to develop a Hybrid DMADL-CTO to identify the abnormalities of CA neurological disorder. Initially, the input for the research is captured from the READISCA dataset. The obtained input video sequence is further subjected to an optimized Multi-metric frame selection phase, in which the optimal frames are selected by the symmetrical measures of SSIM, PSNR, and SINR with the integration of the developed CTO algorithm [30]. In this context, the CTO algorithm is derived by the hybridizing character of Imperialist Competitive Algorithm (ICA) and Teaching Learning Based Optimization (TLO) [31], which effectively degrades the local optima issues and achieves a better convergence rate while selecting the optimal frames. Further, the selected frames are allowed into the pre-processing stage, where the quality of the frames is enhanced by reducing the contrast, noise distortions, redundancy, and inaccurate recordings effectively. Meanwhile, the quality-enhanced frames are fed into ROI extraction, in which the texture and structure of the particular region get enhanced and proceed for further processing. Then, the extracted region is transmitted into the feature extraction phase that performs the extraction process by gait pose kinematic structural flowmap features, which indulges skeleton-based gait features, pose estimated features, kinematic spatiotemporal features, and deep flowmap-based structural features. Furthermore, the extracted features are allowed into the Hybrid DMADL-CTO to detect the abnormalities of CA. Moreover, the performance efficacy of the Hybrid DMADL-CTO is enhanced by tuning the hyperparameters using the developed CTO algorithm methods. In addition to this, the mixed attention mechanism extracts the spatial and channel information of the features to minimize the overfitting and computational complexity problems. Therefore, the overall schematic representation of Hybrid DMADL-CTO is mentioned in Figure 2.

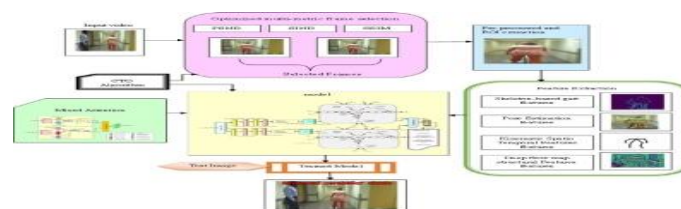


Figure 2: Schematic representation of Hybrid DMADL-CTO model

3.2 Input video dataset

The research garnered the input video from the READISKA dataset [27] respectively, which has 155 video sequences of 6 seconds time duration with gait performance. By considering these video sequences, the model detects the CA disorder precisely. Thus, the video sequence input of the Hybrid DMADL-CTO is depicted as,

$$G = \{k_1, k_2, \dots, k_b, \dots, k_r\} \quad (6)$$

The obtained long-range video sequences are segmented into frames for further evaluation, which is mathematically illustrated as,

$$D(k_b) = \sum_{i=1}^v D_i \quad (7)$$

3.3 Optimized multi-metric Frame selection

The input video sequences are allowed into the frame selection stage, in which the optimal frames are selected by multi-metric parameter evaluation with the integration of optimization. The multi-metric parameter indulges SINR, SSIM, and PSNR, in this the pixels and distributed features of every frame are compared with other frames. Conceptually, the CTO algorithm takes the input sequence and assigns the lower bound as 0, the upper bound value is the mean value that varies based on the evaluation of multi-metric parameters. Thus, the multi-metric parameter representation is depicted as,

$$Mul_{met} = \left(Threshold - \left(\frac{SINR + SSIM + PSNR}{3} \right) \right) \quad (8)$$

In the research, every frame has its own multimetric value (Mul_{met}), based on this the CTO algorithm selects the frame index, which has minimal Mul_{met} value. From the evaluation of the above equation, the attained value is minimal to the threshold (0.8) value, then the frames are assigned as optimal frames. Thus, the selected optimal frames are in the dimensions as $(1 \times 1920 \times 1080 \times 3)$, and the mathematical notation of the selected frames is expressed in equation (3).

3.3.1 Competitive Tuning Optimization Algorithm for optimal frame selection

The optimal video frames are selected based on the utilization of the Competitive Tuning Optimization (CTO) Algorithm, which inhibits the characters of both of ICA [32], and TLO [33] methods. In this context, the imperialist competitive algorithm has strong competitive strategy behavior that provides feasible and ideal results to increase the power of empires. Based on these strategies, the strong empires captured more power and got stronger enough whereas, the weak empires reduced their power and the converge rate collapsed accordingly. Moreover, in Teaching Learning optimization extends the capability to extract the information based on learning factors. Thus, the competitive strategy behavior and extraction capability capture the information with better convergence speed, high reliability, and lower computational cost, which also overcomes the disadvantages of poor sensitivity, high complexity error, and local optima issues effectively.

Inspiration: The CTO algorithm is obtained through the inspiration of imperialistic competition and teaching learning procedure, which is implemented in the research model to detect the optimal frames for pursuing the detection process and tunes the hyperparameters of the model for providing effective outcome results. The competitive strong strategy obtained from the ICA and the learning factor extracting capability character of TLO are combined effectively to attain relevant results. These merged character algorithms provide better selection results and effective detection validation precisely. Thus, the CTO algorithm in the research model provides accurate results during the identification of optimal frames, and the hyperparameters are tuned effectively to achieve better detection results precisely.

Initialization: In the random solution positions, the initial position of the solution is determined as P under the utilization of the CTO algorithm, which is mentioned as,

$$J = \{J_1, \dots, J_p, \dots, J_q\} \quad (9)$$

where, J_p indicates the competitive group of the solution (video frames), and J_q denotes the total number of solutions.

Objective function: In Hybrid DMADL-CTO, the CTO algorithm is widely used to select the optimal frames for making the further detection process simpler with minimum computation cost and speed. While selecting an optimal frame, the fitness measure of the CTO algorithm attains minimum value for the multi-metric measures, which is represented as,

$$L = fit_Fe_{sel} = (\min(\text{multi-metric})) \quad (10)$$

The efficient frames are selected by the minimization of multi-metric parameters, in which the multi-metric parameters indulge SINR, SSIM, and PSNR. Based on these parameters, the optimal frames are effectively selected and proceed for further evaluation.

Solution Update: The solution of the CTO algorithm is updated based on the learning factor of every group. If the learning factor attains high value then which has the ability to observe the information of real environment experience. Meanwhile, the learning factor attains minimum value, which inhibits poor learning capacity and also needs a constant teaching environment. Due to this minimal learning factor, the poor performance solution in the teaching environment is termed as the worst solution. Conceptually, the CTO algorithm captures the information of the worst solution for exploring better learning experiences.

Learning Factor Evaluation

The learning factor rate is measured for every q solution, and is computed as,

$$F = \left(\frac{L(J_p^m) + L(J_p^{m-1})}{2} \right) \quad (11)$$

where, m denotes the current iteration, J_p^m indicates the current competitive group solution at the current iteration, J_p^{m-1} defines the current competitive group solution at the previous iteration. Based on these learning factors, the obtained learning and teaching phase solutions are updated until maximum iteration.

Phase 1: Exploitation phase- $[L \leq 1]$

If the attained fitness measure is greater than or equal to 1, then it handles the exploitation phase. In this research, the exploitation phases include the Global learner, weaker learner, and personal learner phases. Due to these phases, the learner improves the knowledge about the interactions

a) Global Learner Phase $\left[F(J_p^{m+1}) \geq F_{glo} \leq F_{thres} \right]$

This phase is attained when the learning factor of the updated current competitive group solution $F(J_p^{m+1})$, is greater than the global learning factor F_{glo} , and less than the threshold learning factor F_{thres} . In this phase, the learner attained better learning performance with the inclusion of a linear learning experience that enhances the detection accuracy of the model. Thus, the mathematical representation of the updated Global learner phase is given as,

$$J_{p1}^{m+1} = J_p^m + g_1(J_a^m - J_p^m) + g_2(J_{glo}^m - J_p^m) + g_3(J_{best}^m - J_p^m) + g_4(J_p^m - J_p^{m-1}) + rnd[1 + ran(0,1)] + E \times g_5 \quad (12)$$

where, J_a^m indicates the previous competitive group solution at the current iteration, J_{glo}^m represents the global competitive group solution at the current iteration, J_{best}^m mentions the best competitive group solution at the current iteration, E defines the interactive experience, ran refers the random variable that ranges from 0 to 1, and g_1, g_2, g_3, g_4 denotes the random values, which are mathematically illustrated as,

$$g_1 = l_{\max} - \frac{m_{\max} \times (l_{\max} - l_{\min})}{m_{\max}} \quad (13)$$

$$g_2 = l_{\max} + (l_{\max} - l_{\min}) \times \left(\frac{m_{\max} - m}{m_{\max} - 1} \right) \quad (14)$$

where, l_{\max} denotes the maximum inertia factors, l_{\min} refers the minimum inertia factors, m_{\max} defines the maximum iteration. Similarly, the random value of g_3, g_4 is expressed as,

$$g_3 = \sum_q M(J_p^m) \times \|J_{best}^m - J_p^m\| \quad (15)$$

$$g_4 = \sum_q Q(J_p^m) \times \|J_p^m - J_p^{m-1}\| \quad (16)$$

where, $M(J_p^m)$ mentions the global factor of p^{th} learner, and $Q(J_p^m)$ expresses the local factor p^{th} learner, which is illustrated by,

$$M(J_p^m) = \begin{cases} 0, & \|J_{best}^m - J_p^m\| \leq F_{thres} \\ 1, & otherwise \end{cases} \quad (17)$$

$$Q(J_p^m) = \begin{cases} 1, & \|J_p^m - J_p^{m-1}\| \leq F_{thres} \\ 0, & otherwise \end{cases} \quad (18)$$

In equation (12), the term g_5 represents the interactive degree, which is depicted as,

$$g_5 = d_0 - (d_0 - d_{\min}) \cdot \cos(\exp^{2-l} - 1) \quad (19)$$

where, d_0 indicates the initial learning factor value, d_{\min} defines the minimum value of learning factor, and $\cos(\exp^{2-l} - 1)$ represents the self-learning nature of learners.

b) Personal Learner Phase $\left[F(J_p^{m+1}) \geq F_{glo} \geq F_{thres} \right]$

This phase begins when the learning factor of updated competitive group solution $F(J_p^{m+1})$, is greater than the global learner factor F_{glo} , and also greater than the threshold learning factor F_{thres} . In this phase, the obtained group learning and group performance information from the global learner phase is shared with every individual solution

to achieve better individual supporting knowledge, which also enhances the overall competitive group performance. Based on this condition, the updated personal learner phase is given by,

$$J_{p2}^{m+1} = J_p^m + g_6 \left(J_{p1}^{m+1} - J_{glo}^m \right) + g_7 \left(J_{p1}^{m+1} - J_a^m \right) \quad (20)$$

where, J_{p1}^{m+1} defines the updated solution of the global learner phase, and g_6, g_7 illustrates the random value that ranges from 0 to 1. The threshold value for the learner phase is mentioned as,

$$F_{thres} = \frac{F(J^m) + F(J^{m-1}) + F(J^{m-2}) + \dots + F(J^{m_{max}})}{m_{max}} \quad (21)$$

where, $F(J^m)$ indicates the learning factor of the competitive group solution at the current iteration, $F(J^{m-1}), F(J^{m-2})$ denotes the learning factor of the competitive group solution of previous iterations, and $F(J^{m_{max}})$ defines the learning factor of the competitive group solution of maximum iterations. By considering the group and individual learning performance ability under successive iterations, the random factor is determined as,

$$ran_{fac} = \frac{F(J_{p1}^m) + F(J_{p2}^m)}{2} \quad (22)$$

where, $F(J_{p1}^m)$ indicates the global learning phase with current iteration, and $F(J_{p2}^m)$ represents the personal learning phase with current iteration.

c) Weak Learner Phase $\left[F(J_p^{m+1}) \leq F_{glo} \right]$

This phase activates when the learning factor updated competitive group solution $F(J_p^{m+1})$, is greater than the global learner factor F_{glo} , which also observes the knowledge of the worst solution. In this phase, the worst solutions are taken into account to provide a better learner value score, which enhances the overall mean of the solution group. Thus, the updated weak learner solution is illustrated as,

$$J_{p3}^{m+1} = J_p^m + g_8 \left(J_{best} - A \times J_{mean} \right) + g_9 \left(J_{p2}^{m+1} - J_{weak}^m \right) + g_{10} \left(J_{p1}^{m+1} - J_p^m \right) \quad (23)$$

where, J_{weak}^m denotes the weak learner solution at the current iteration, J_{mean}^m represents the average mean knowledge, J_{best} indicates the best solution, A indicates the teaching factor, g_8, g_9, g_{10} mentions the random value, which is mathematically depicted as,

$$g_8 = m_{max} \left(1 - \frac{\max \left(F(J_{best}), F(J_{mean}) \right)}{\max \left(F(J_{best}) \right)} \right) \quad (24)$$

$$g_9 = m_{max} \left(1 - \frac{\max \left(F(J_{glo}^m), F(J_{p2}^{m+1}) \right)}{\max \left(F(J_{p2}^{m+1}) \right)} \right) \quad (25)$$

$$g_{10} = m_{\max} \left(1 - \frac{\max(F(J_{p1}^{m+1}), F(J_p^m))}{\max(F(J_{p1}^{m+1}))} \right) \quad (26)$$

From the above equation, the average mean knowledge value is illustrated as,

$$J_{\text{mean}} = \left(\frac{J_{p1}^{m+1} + J_{p2}^{m+1}}{2} \right) - g_{11} (J_{p1}^{m+1} - J_{\text{best}}^m) - g_{12} (J_{p1}^{m+1} - J_{p2}^{m+1}) \quad (27)$$

where, g_{11}, g_{12} denotes the random values, which is mathematically depicted as,

$$g_{11} = d_0 - (d_0 - d_{\max}) \cdot \cos(\exp^{2-l} - 1) \quad (28)$$

$$g_{12} = h_{\max} - \frac{(h_{\max} - h_{\min}) \times (F(J_{\text{best}}) - F(J_{\text{weak}}))}{F(J_{\text{best}}) - F(J_{\text{weak}})} \quad (29)$$

where, d_{\max} mentions the maximum value of the learning factor, h_{\max}, h_{\min} indicates the maximum and minimum weight value.

Phase 2: Exploration Phase- [$L > 1$]

If the attained fitness measure is greater than 1, then it remains in the exploration phase. In this research, the exploration phase inhibits the Adaptive learning phase and the Random Learning phase. Because of these phases, the prolonged and observational experience makes the model detection accuracy precise.

a) Adaptive Learning Phase [$F(J_p^{m+1}) \geq \text{ran}_{\text{fac}}$]

When the updated learning factor of the current competitive group $F(J_p^{m+1})$, is greater than or equal to the random factor ran_{fac} , then the adaptive learning phase occurs significantly. In this phase, the prolonged experience of the solution is observed by scheduled learning. Thus, the mathematical representation for adaptive learning is given by,

$$J_{p4}^{m+1} = J_{\text{ran}}^{\text{out}} + 0.5 \left[(A \times J_3^m + A \times J_4^m + J_p^m + (\alpha J_p^{m-1} + \alpha(1-\alpha)J_p^{m-2} + \alpha(2-\alpha)(1-\alpha)J_p^{m-3})) \right] \quad (30)$$

where, $J_{\text{ran}}^{\text{out}}$ defines the solution of the random learner in the outside search boundary, J_3^m, J_4^m denotes the solution experiment at current iteration, α indicates the adaptive fusion degree, which is depicted as,

$$\alpha = m_{\max} \times \left(1 - \frac{m}{m_{\max}} \right) \times (J_{\text{glo}}^m) \quad (31)$$

b) Random Learning Phase [$F(J_p^{m+1}) < \text{ran}_{\text{fac}}$]

When the updated learning factor of the current competitive group $F(J_p^{m+1})$, is less than the random factor ran_{fac} , then a random learning phase has occurred respectively. In this phase, the position of the solution is explored randomly to attain acquired knowledge with a teaching mechanism. Thus, the updated random learning phase is determined as,

$$J_{p5}^{m+1} = J_p^m + ran \times (J_{ran} - J_p^m) \quad (32)$$

where, J_{ran} indicates the random learner solution, and ran indicates the random numerical. Thus, the CTO algorithm effectually selects the optimal frames and the hyperparameters of the Hybrid DMADL are tuned significantly to detect the abnormalities of cerebellar ataxia.

Termination condition: When $m < m_{max}$ the global position of the competitive group solution is evaluated by the learning factor under the exploitation and exploration phase. In these phases, the learning factor of the solution is analyzed by global, personal, and weak learners, additionally in adaptive and random learning phases, which effectually reduces the imbalance issue and complexity problems. Therefore, the CTO algorithm execution in the frame selection phase effectually selects the optimal frames for pursuing further evaluation. The entire flow map representation for the CTO algorithm is illustrated in Figure 3.

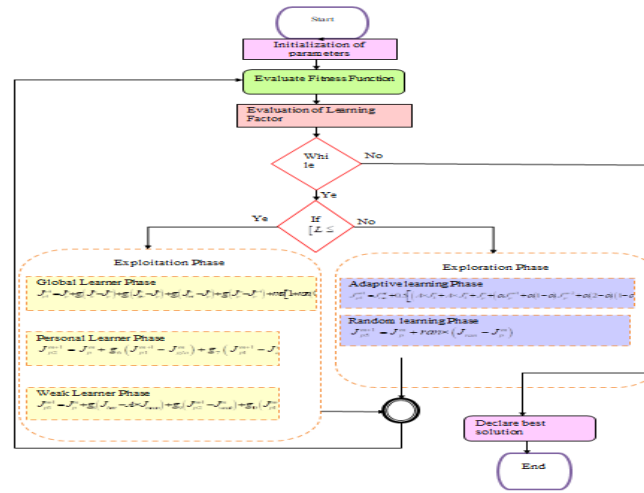


Figure 3: Flow map for CTO algorithm

3.4 Pre-processing and ROI extraction of Selected Video Frames:

The selected optimal frames are subjected to the pre-processing stage to enhance the quality of every frame by reducing the undesired phenomena, noise distortions, redundant data, and contrast influences. When eliminating these distortions, the quality of the frames is strengthened, and the weighted method enhances the contrast and brightness of the frame and proceeds for further evaluation. The obtained dimensions of the pre-processed frames are $(1 \times 1920 \times 1080 \times 3)$, and that is expressed as,

$$D^*(k_b) = \{D_1^*, D_2^*, \dots, D_i^*, \dots, D_v^*\} \quad (33)$$

Thus, the quality-enhanced frames are fed into ROI extraction, here the structure and texture of the image region are effectually extracted. Conceptually, the exact regions of images are extracted using morphological operations, which analyse the gradient and contour texture of the region. From this perspective, the shape and similar front objects with identical intensity and colour are effectively discovered in a particular curve, which connects all the continuous points in the boundary significantly. This contour analysis is widely used in image acquisition and region extraction. Further, the size of the image is rescaled as $(n \times 224 \times 224 \times 3)$ the attained ROI extracted image is depicted as H , and is given as,

$$H = RoI(D^*(k_b)) \quad (34)$$

3.5 Feature extraction using gait pose kinematic structural flowmap features

The extracted ROI image is further allowed into the feature extraction stage, where instructive features are extracted significantly. In the research, the gait pose kinematic structural flowmap features are utilized to extract the informative features of the RoI image. The developed feature extraction techniques indulge Skeleton-based gait features, pose estimation, Kinematic Spatiotemporal features, and deep flow map-based structural features, which are briefly elaborated as follows,

The skeleton-based gait features extract the walking pattern of humans with the utilization of skeletal mechanisms. In this, the extraction process is done based on the integration of a deep learning neural network with an open pose system, which effectually estimates the pose of the human body. Additionally, the orientation and position of limbs and joints are determined precisely with the representation of skeletal structure. Moreover, the neural network indulges convolutional and pooling layers, which extract high-level feature information and emphasize better real-time performance, accuracy scrutiny, and improved robustness conditions [34]. Thus, the outcome of skeleton-based gait features is depicted as B , which is mathematically denoted as follows, and the achieved dimension is $(1 \times 224 \times 224)$.

$$B = H \otimes T + e \quad (35)$$

where, T defines the convolution filter, e indicates the learnable parameter, and \otimes denotes the 2-dimensional convolutional operation.

The pose estimation feature method extracts the gait features of humans, along with the specifications of the human face, legs, body, and hand joints effectually. Conceptually, the model tracks the human body movements for managing diverse capabilities, colour space conversion, frame resizing, and video stream manipulation in real-time scenarios. In this research, the estimation analysis is done by a bottom-up technique that assesses the actual pose of the person based on the linkable pairs of the body parts [35]. Thus, the obtained dimension of pose estimation is $(1 \times 224 \times 224 \times 3)$, and the outcome is expressed as N .

The Kinematic spatiotemporal feature extraction process extracts the informative features based on the technique of Multi-Layer Perceptron. In this, the spatial and temporal features are applied in the axes to perform the image classification with minimal processing time. Conceptually, the model contains multiple blocks that are widely used to extract the spatiotemporal features at united insertion length [36] with dimensions $(1 \times 224 \times 224)$, and the outcome is defined as,

$$C = H + f_2 \eta(f_1 \text{norm}(H)) \quad (36)$$

$$S = C + f_4 \eta(f_3 \text{norm}(C)) \quad (37)$$

where, C represents the insertion data length, f_1, f_2, f_3, f_4 denotes the MLP parameter, η refers the activation function, S indicates the outcome of kinematic spatiotemporal features, and norm indicates the normalization function.

The pre-trained residual network is influenced by the architecture of the VGG-19 model possesses a deep network structure with numerous convolutional layers of 33 filters. Conceptually, the output feature map contains the same number of filters, which are then allowed into the Local Ternary pattern for extracting the structural flow-map features. The process effectively evaluates the correlation of pixel intensity along with the neighborhood image. Based on this correlation information, the pixel representation is converted into ternary code to capture the local texture accurately [37]. Thus, the obtained deep flow map-based structural features are mathematically mentioned as,

$$K = \sum_{c=0} R(O_{nei} - O_{cen}).3^c \quad (38)$$

$$R = \begin{cases} 1, & O \geq thr \\ 0, & O = thr \\ -1, & O \leq thr \end{cases} \quad (39)$$

where, R represents the gradient size, O_{nei} indicates the neighbor pixel intensity, O_{cen} denotes the current pixel intensity, and thr refers the threshold value. Thus, the achieved dimensions of Deep flowmap-based structural features are $(1 \times 224 \times 224)$.

Therefore, the extracted features under gait pose kinematic structural flowmap features are expressed as W , which is obtained by the concatenation of Skeleton-based gait features, pose estimation, Kinematic Spatio temporal features, and deep flow map-based structural features. Thus, the entire concatenate feature is represented in the form of a vector with dimensions $(1 \times 224 \times 224 \times 6)$, and the achieved feature vector is represented as,

$$W = \{B \parallel N \parallel S \parallel K\} \quad (40)$$

where, B denotes the outcome of skeleton-based gait features, N represents the outcome of pose estimation, S represents the outcome of kinematic spatio temporal features, and K mentions the outcome of deep flow map-based structural features.

3.6 Distributed Learning based Convolutional Neural Network- Long short Term Memory Model for detecting cerebellar Ataxia

In this neurological disorder research, the extracted features are fed into Hybrid DMADL-CTO to detect the abnormalities of cerebellar ataxia. In this context, the Hybrid DMADL-CTO is obtained by the combination of LSTM, CNN, and distributive learning techniques. In general, the LSTM classifier effectually performs complex structure evaluation and retains long data dependencies with high robustness but, it suffers from computational cost, interpretability challenges, and complexity issues. To overcome these issues, a Hybrid DMADL-CTO model is designed in this research, which evaluates the detection operation by distributed learning method. Because of this, the computation cost is decreased gradually and attains better performance efficacy with high detection accuracy. Meanwhile, the contextual information of long-range dependencies and model ability is enhanced by the integration of a mixed attention mechanism. Additionally, the convergence rate of the model is improved through the usage of CTO algorithm. By considering these combinations, the Hybrid DMADL-CTO achieved better interpretability and flexibility measures with effective performance memory. The schematic representation of the Hybrid DMADL-CTO architecture is mentioned in Figure 4

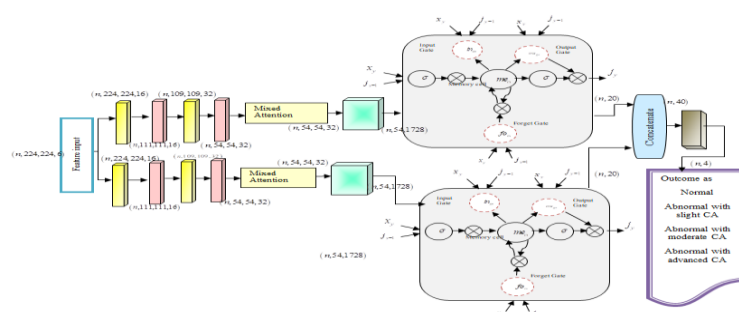


Figure 4: Schematic representation of Hybrid DMADL-CTO Architecture

The above figure 4 represents the architecture of the Hybrid DMADL-CTO, in which the input is garnered from the extracted features with dimensions $(n, 224, 224, 6)$, which is fed into the convolutional layer that is positioned in a distributed form. These distributed evaluations reduced the time consumption and provided better performance results. In this context, the convolutional layer captures the temporal and spatial dependencies during the training process and the acquired dimension is $(n, 222, 222, 16)$. The attained dimension is then allowed into the max pooling layer, which selects the maximum value by reducing the dimensions as $(n, 111, 111, 16)$. These convolutional and max-pooling evaluations are performed in an iterative manner with attained dimension $(n, 54, 54, 32)$. Further, the achieved dimension is subjected to the mixed attention layer that enhances the efficiency of complex data structure and achieves better interpretability, flexibility and decreases the gradient vanishing issue. Conceptually, the mixed attention mechanism effectually extracts the contextual information of long-range dependencies with the integration of global averagepooling at the output layer that minimizes the overfitting issues and computation complexity [35]. The obtained dimension for the mixed attention mechanism is $(n, 54, 54, 32)$. Furthermore, the outcome of the mixed attention module is subjected to the time-distributed layer or flattened layer that switches the dimensions by converting the sequential input into the historical period of time sequence data with more long-range features. Thus, the achieved dimension of the flattening layer is $(n, 54, 1728)$. Moreover, the flattened outcome is allowed into the LSTM network layer, where the gradient explodes and vanishing problems are reduced gradually. On this the updated memory information is stored in three diverse gates they are input, forget, and output. In this, the input cell contains the information of the memory cell, whereas the forget gate clears and handles the cell memory, and the output gate manages the exposed memory content. Additionally, the LSTM constraints contain hyperbolic tangent and sigmoid activation functions to perform element-wise multiplication, which effectively overcomes the long dependency issues [36]. Thus, the outcome of the memory constraints is denoted as V with dimension $(n, 20)$, which is mathematically illustrated as,

$$V = \sigma \left(me_{ce} \otimes \tanh \left(s_{x_y} + n_{j_{y-1}} \right) \right) \quad (41)$$

where, me_{ce} represents the memory cell, \otimes mentions the element-wise multiplication, \tanh indicates the hyperbolic tangent activation function, σ refers the sigmoid activation function, s, n defines the parameters of memory constraints, x_y represents the input sequence, and j_{y-1} mentions the previous hidden state sequence. Thus, the attained outcome from the distributed LSTM constraints is concatenated $(n, 40)$, which is further fed into a dense layer, where, the actual prediction results are obtained based on the previously extracted feature information, and the dimension is attained as $(n, 4)$ class labels of normal, abnormal with slight ataxia, abnormal with moderate ataxia, and abnormal with advanced ataxia. Further, the performance efficacy and ability of the model are improved by the integration of the CTO algorithm, which is implemented in the layer of the model that reduced the vanishing explode and local optimization problems by tuning the weights and bias of Hybrid DMADL-CTO significantly, which is briefly explained in section 3.3. During training, the ability of the model is enhanced by CTO algorithm with an maximal fitness measure of accuracy, and is mathematically depicted as,

$$L' = fit_mo_{tr} = \left(\max (Accuracy) \right) \quad (42)$$

Based on this fitness measure, the CTO algorithm effectively tunes the bias and weights of Hybrid DMADL to achieve accurate detection of CA along the class labels. Thus, the predicted outcome of the model is represented as \hat{t}_b .

3.6.1 Mixed Attention Mechanism

In the research, the mixed attention comprises both spatial and self-attention modules to enhance the efficacy and performance of the Hybrid DMADL-CTO, which is implemented in the layer of the model to evaluate the detection tasks. Conceptually, the mixed attention captures the correlation of inter-spatial features with the insertion of both max pooling and average pooling layers. These pooling operations are done in the channel axis that highlights the instructive regions and concatenates the attained features to generate an effectual feature descriptor. Further, the feature descriptor is allowed into a convolutional layer to provide an attention map of spatial representatives [37]. The computed spatial attention map with the insertion of the sigmoid activation function is described as,

$$Y_{spa} = \sigma \left(co^{7 \times 7} \left(AvgPool_{spa}; \max pool_{spa} \right) \right) \quad (43)$$

where, $co^{7 \times 7}$ represents the convolution operation with filter size 7×7 , Y_{spa} mentions the spatial attention map, and σ indicates the sigmoid activation function. Moreover, the computational complexities and memory resource consumption are effectively reduced by the utilization of mixed attention modules, which garnered the semantic information of long-range dependencies accurately. In this context, the input feature vector is mapped by sequence-to-sequence transformation for capturing the contextual information. Additionally, the sum weights are computed by the matrix multiplication method, in which the input feature vectors are further classified into small-size vectors in the role of query, key, and value in the parallel form [38]. From this perspective, the overall average weights of the vectors are determined, and the information of diverse vector representations is effectual. Thus, the computed weighted average is given as,

$$Y_{se} = \sum we_{ig} \otimes (qu.ke.va) \quad (44)$$

where, we_{ig} indicates the weighted parameter, qu refers the parameter of the query vector, ke represents the parameter of the key vector, and va mentions the parameter of the value vector. Therefore, the overall architecture of the mixed attention module is expressed in Figure 5.

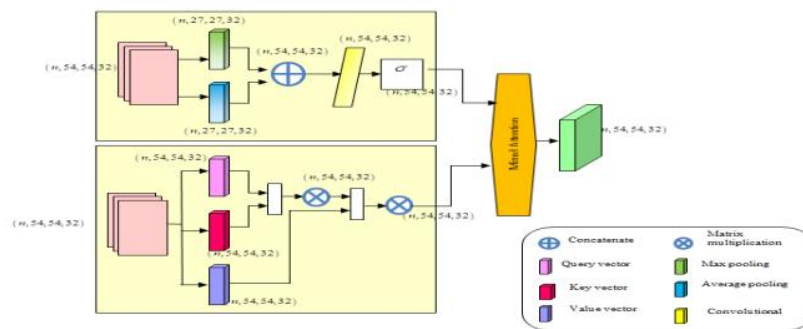


Figure 5: Architecture of Mixed attention mechanism

4. RESULTS

The CA detection performed using Hybrid DMADL-CTO provides effectual results with the utilization of numerous evaluation measures of Accuracy, Recall, and precision. The Hybrid DMADL-CTO is significantly analyzed by the READISCA dataset, thus the achieved performance effectiveness is briefly determined in the section.

4.1 System Implementation

The Hybrid DMADL-CTO research of CA is executed under the software of Pycharm of version 2024.2.1 in the application of Python version 3.7 in the operating system of Windows 11 configuration, which has the memory of 128 GB and 16 GB of ROM and RAM respectively. Furthermore, the operating system has a clock speed that ranges from 3GHz.[38][39][40]

4.2 Dataset Specification

The input for the CA research is captured from the READISCA dataset, the detailed explanation of this dataset is given as follows,

READISCA dataset [41] contains 155 long video sequences with the inclusion of gait performance, in which the video subject is 6 seconds long and that is in an anonymized version. From this perspective, the dataset has four class labels normal, abnormal with slight CA, abnormal with moderate CA, and abnormal with advanced CA. These ataxia disorder classes are effectively detected by the Hybrid DMADL-CTO model.

4.3 Experimental Outcomes

In this section, the outcome of every individual stage of the model under the detection of CA is illustrated in Figure 6.



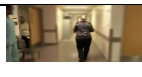











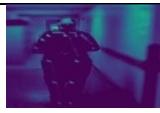

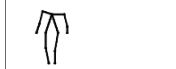


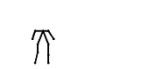




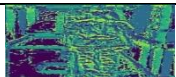
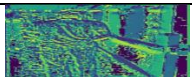

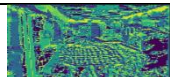

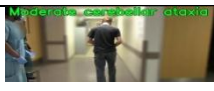
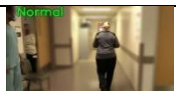

Input Video					
Selected Frame					
Pre-processed Image					
Feature Extraction	Skeleton-based Gait features				
	Kinematic Spatio Temporal features				
	Pose Estimation				
	Deep flow-based Structural pattern				
Detected Result					

Figure 6: Experimental outcomes of Hybrid DMADL-CTO

4.4 Performance Metrics

To evaluate the effectiveness of the Hybrid DMADL-CTO, numerous evaluation measures are utilized in the research as accuracy, recall, and precision under training Percentage (TP), and K-Fold validation (KF), which are briefly described in table 1.

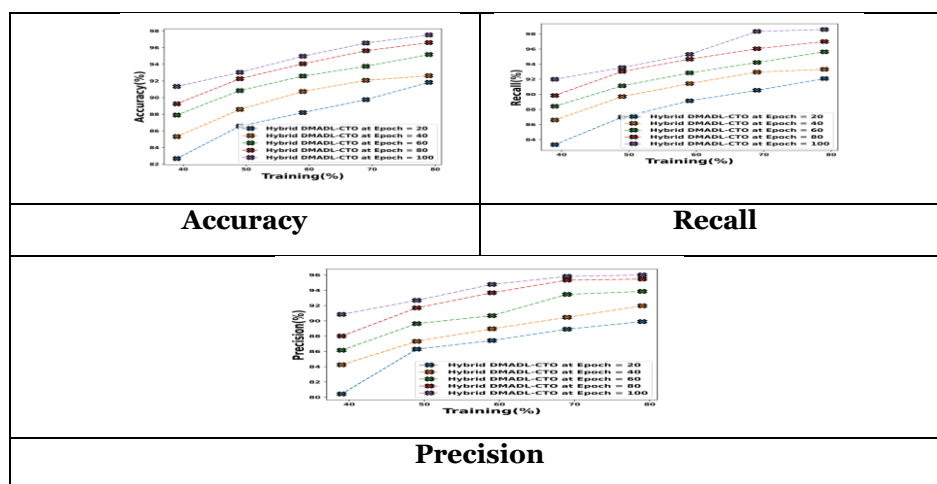
Table 1: Explanation of Performance metrics

Performance Metrics	Expression	Explanation
Accuracy	$\frac{Tr_u + Tr_w}{Tr_u + Tr_w + Fa_u + Fa_w}$	Measures the corrected classified predicted value to the overall number of predicted instances.
Recall	$\frac{Tr_u}{Tr_u + Fa_w}$	Computes the proportion of overall positive values obtained from the actual predicted value.
Precision	$\frac{Tr_u}{Tr_u + Fa_u}$	Determines the measure of positive instances from the overall actual positive predicted value.

where, Tr_u mentions the true positive value, Fa_u represents the false positive value, Tr_w indicates the true negative values, and Fa_w refers the false negative value respectively.

4.4.1 Performance Assessment of READISCA dataset under Training Percentage

The performance assessment of the model under the READISCA dataset for numerous evaluation metrics at a maximum TP of 80% with various epochs of 20, 40, 60, 80, and 100. The accuracy obtained by the proposed Hybrid DMADL-CTO under epoch 20 is 91.82%, further for epoch 40, the value is increased to 92.62%, for epoch 60 the value is 95.16%, for epoch 80 the value is 96.60%, and for epoch 100 the attained accuracy value is 97.52%, which shows an increased value while compared with other epochs. Further, the recall measured at 80% TP for epoch 20 is 92.08%, for epoch 40 is 93.32%, for epoch 60 is 95.63%, for epoch 80 is 96.98%, and for epoch 100, the achieved recall value is 98.59%. Similarly, for precision metric under epoch 20, the attained value is 89.89%, for epoch 40 is 91.96%, for epoch 60 is 93.85%, for epoch 80 is 95.47%, and for epoch 100 is 96.00% under 80% TP. By considering this effectiveness, the Hybrid DMADL-CTO increased the ability of detection framework under CA. Figure 7 shows the performance analysis based on TP.

**Figure 7:** Performance Assessment of READISCA dataset under Training Percentage

4.4.2 Performance Assessment of READISCA dataset under K-Fold validation

Figure 8, illustrates the performance efficacy of the model under the KF analysis of the READISCA dataset. In this, the performance assessment of the model is analyzed for a maximum KF of 10 under various epochs of 20, 40, 60, 80, and 100. The accuracy obtained by the proposed Hybrid DMADL-CTO under epoch 20 is 88.59%, further the

value attained for epoch 40, is 90.08%, the value achieved for epoch 60 is 91.89%, the value attained for epoch 80 is 95.37%, and the accuracy value attained for epoch 100 is 97.66%, which shows an increased value. Further, the value for recall metric under epoch 20 is 89.52%, the value attained for epoch 40 is 91.24%, the value achieved for epoch 60 is 93.38%, the value obtained for epoch 80 is 96.35%, and the recall metric value for epoch 100 is 97.99%. Similarly, the value for precision metric under epoch 20 is 87.53%, the value for epoch 40 is 88.84%, the value attained for epoch 60 is 90.55%, the value achieved for epoch 80 is 94.54%, and the value for precision metric at epoch 100 is 95.88% under maximum KF of 10.

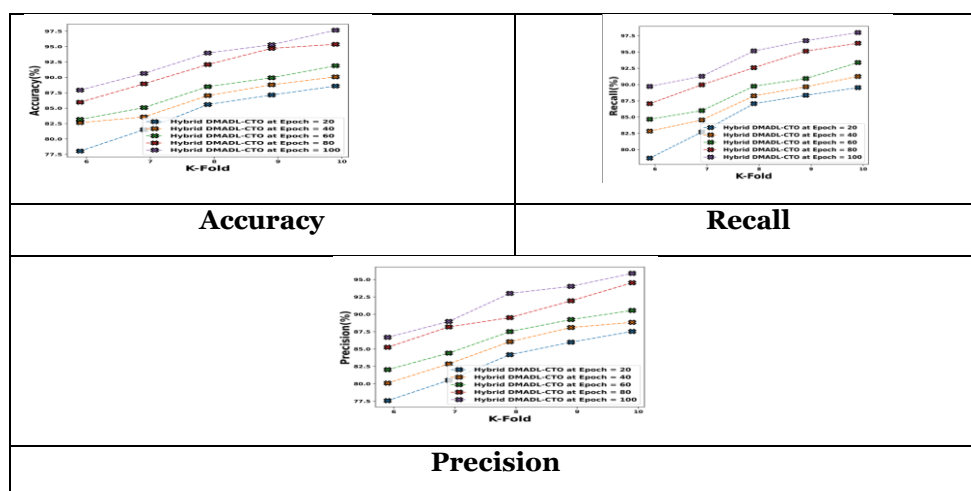


Figure 8: Performance Assessment of READISCA dataset under K-Fold validation

4.5 Comparative methods

The performance effectiveness of the Hybrid DMADL-CTO is compared with other related approaches as FWO-XGB [42], ODLCNN [43], DL-MCD-DT [44], AtGCN [45], ctGAN [46], FSR-XGB [47], Hybrid DRL [48], ICA-hybrid DRL [49], TLO-hybrid DRL [50], Hybrid DMADL [51], and Hybrid DRL-CTO techniques.

4.5.1 Comparative Assessment with Training Percentage

The comparative assessment of Hybrid DMADL-CTO under the READISCA dataset obtained effective results, which are illustrated in Figure 9. For 80% of TP, the accuracy value for Hybrid DMADL-CTO is 97.52%, which shows an enhancement of 16.60% for FWO-XGB, 12.61% for ODLCNN, 19.17% for DL-MCD-DT, 11.78% for AtGCN, 17.22% for ctGAN, 15.27% for FSR-XGB, 8.29% for Hybrid DRL, 5.63% for ICA-Hybrid DRL, 4.43% for TLO-hybrid DRL, 0.84% for hybrid DMADL, 2.63% for hybrid DRL-CTO. For recall metric, the value obtained by Hybrid DMADL-CTO is 98.59%, which is 18.03% higher than FWO-XGB, 13.05% greater than ODLCNN, 20.02% enhanced than DL-MCD-DT, 12.80% developed than AtGCN, 18.19% improved than ctGAN, 16.36% greater than FSR-XGB, 9.24% developed than Hybrid DRL, 6.10% enhanced than ICA-hybrid DRL, 4.50% increased than TLO-hybrid DRL, 0.88% improved than hybrid DMADL, and 2.69% higher than Hybrid DRL-CTO. Similarly, for precision metric, the value obtained by Hybrid DMADL-CTO is 96.00%, which shows an enhancement of 14.98%, 11.78%, 19.14%, 11.12%, 16.44%, 13.73%, 8.52%, 5.81%, 3.94%, 0.94%, and 2.445% over FWO-XGB, ODLCNN, DL-MCD-DT, AtGCN, ctGAN, FSR-XGB, Hybrid DRL, ICA-hybrid DRL, TLO-hybrid DRL, Hybrid DMADL, and Hybrid DRL-CTO method appropriately.

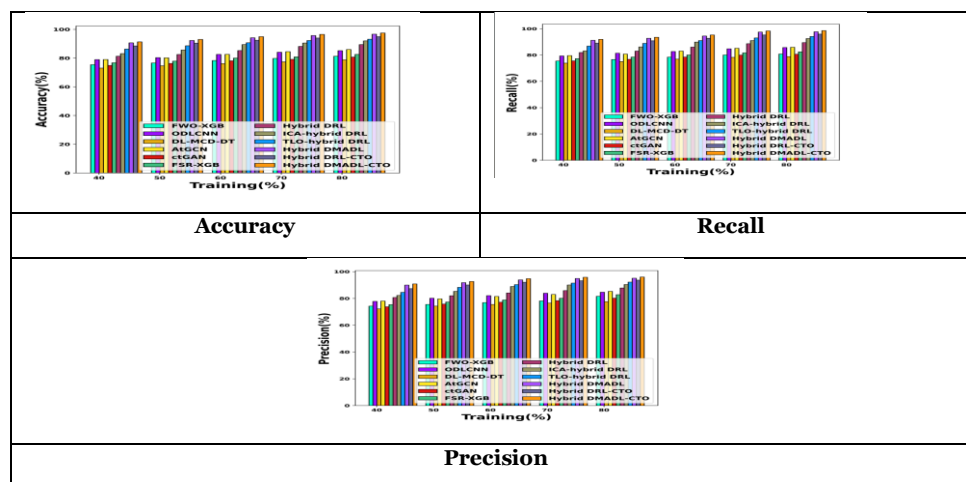


Figure 9: Comparative Assessment of READISCA dataset under Training Percentage

4.5.2 Comparative assessment with K-fold validation

Figure 10, depicts the comparative assessment of the Hybrid DMADL-CTO under the READISCA dataset of various related methods with a maximum KF of 10. In this comparative evaluation, the accuracy metric of Hybrid DMADL-CTO is 97.66%, which is 14.86% higher than FWO-XGB, 11.58% greater than ODLCNN, 16.47% enhanced than DL-MSC-DT, 11.09% developed than AtGCN, 14.75% improved than ctGAN, 13.02% greater than FSR-XGB, 9.15% developed than Hybrid DRL, 8.21% enhanced than ICA-hybrid DRL, 6.65% increased than TLO-hybrid DRL, 0.87% improved than hybrid DMADL, and 3.76% higher than Hybrid DRL-CTO. Further, the achieved value for recall metric under Hybrid DMADL-CTO is 97.99%, which shows an enhancement of 15.07%, 11.15%, 16.52%, 11.04%, 14.60%, 12.68%, 9.41%, 7.86%, 6.01%, 0.90%, and 3.46% over FWO-XGB, ODLCNN, DL-MCD-DT, AtGCN, ctGAN, FSR-XGB, Hybrid DRL, ICA-hybrid DRL, TLO-hybrid DRL, Hybrid DMADL, and Hybrid DRL-CTO method respectively. For Precision metric, the obtained value for Hybrid DMADL-CTO is 95.88%, which shows an enhancement of 13.37% for FWO-XGB, 10.70% for ODLCNN, 15.16% for DL-MCD-DT, 10.20% for AtGCN, 13.62% for ctGAN, 12.08% for FSR-XGB, 8.32% for Hybrid DRL, 7.20% for ICA-Hybrid DRL, 5.90% for TLO-hybrid DRL, 0.90% for hybrid DMADL, 3.41% for hybrid DRL-CTO.

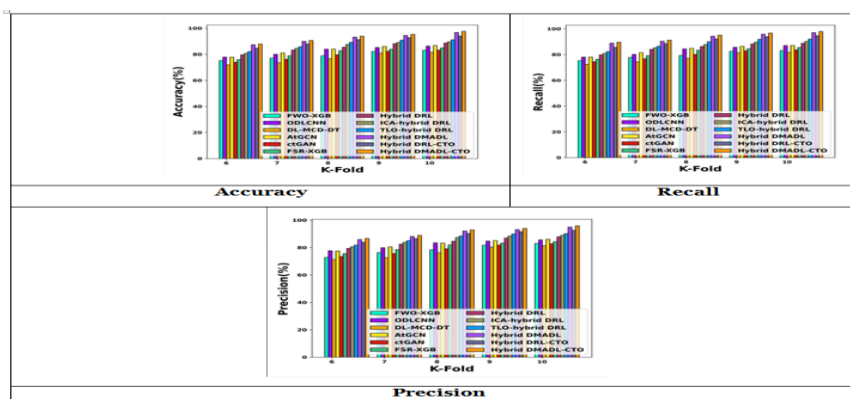


Figure 10: Comparative Assessment of READISCA dataset under K-fold validation

4.6 ROC curve

This section determines the results of the true positive rate (TPR) with respect to the false positive rate (FPR). These evaluation metrics provide accurate detection of CA with its class label classification. Because of these enhancements, the entire performance efficacy and ability of the model is improved effectually. The obtained ROC value for the Hybrid DMADL-CTO is 0.97, which shows an improved value when compared with other related methods as FWO-XGB, ODLCNN, DL-MCD-DT, AtGCN, ctGAN, FSR-XGB, Hybrid DRL, ICA-hybrid DRL, TLO-hybrid DRL, Hybrid

DMADL, and Hybrid DRL-CTO. These conventional methods attain the ROC values of 0.83, 0.85, 0.86, 0.90, 0.88, 0.89, 0.92, 0.93, 0.95, 0.960, and 0.965. Therefore, the ROC curve for the model is depicted in Figure 11.

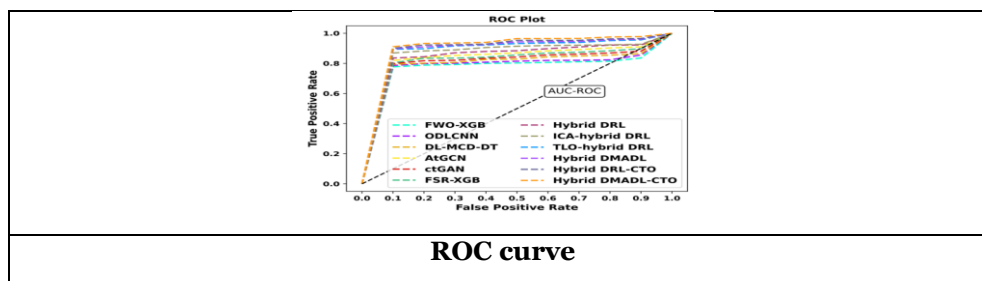
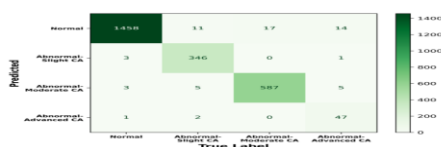


Figure 11: ROC evaluation of the proposed model

4.7 Confusion matrix

In the research, the Hybrid DMADL-CTO utilized the READISCA dataset, which effectively classifies the class labels as normal, abnormal slight CA, abnormal moderate CA, and abnormal advanced CA. In this context, the confusion matrix evaluation determines the predicted label with respect to the true label that provides precise detection with accurate class classification labels. In this evaluation, 1458 data are precisely classified as normal labels, 346 data are categorized as abnormal slight CA labels, 587 data are classified as abnormal with moderate CA labels, and 47 data are categorized as abnormal with advanced CA labels. Thus, the Hybrid DMADL-CTO attains a high detection rate with [precise classification of class labels when compared to other related methods. The attained confusion matrix is represented in Figure 12.



Confusion Matrix

Figure 12: Confusion matrix of proposed model

4.8 Comparative Discussion

The Hybrid DMADL-CTO achieved effective detection results when compared with other conventional approaches. In the FWO-XGB model, the operation time was increased, which caused overlapping and scalability issues. Similarly, the ODLCNN model requires manual imputation of features to achieve accurate detection results. Comparatively, the DL-MCD-DT model faced high computation complexity while validating the detection and classification process. Furthermore, the FSR-XGB model performed a cross-validation feature selection method that reduced the performance effectiveness of the model. Additionally, the AtGCN model faced high error loss that affected the validation time and training capability. The ctGAN model suffered from minimal dependability and robustness issues during the evaluation of the detection process. Moreover, the Hybrid DRL-CTO model consumed a huge operational time to provide accurate detection results, which caused a minimal decrease in its performance effectiveness gradually. To overcome these disadvantages, a Hybrid DMADL-CTO provides precise detection results with minimal time consumption, which are schematically tabulated in Table 2.

Table 2: Comparative Discussion of the Hybrid DMADL-CTO

Methods/Metrics	READISCA dataset					
	TP (90%)			KF (10)		
	Accuracy (%)	Recall (%)	Precision (%)	Accuracy (%)	Recall (%)	Precision (%)
FWO-XGB	83.15	83.22	83.06	81.33	80.81	81.62
ODLCNN	86.35	87.06	85.62	85.22	85.73	84.69
DL-MCD-DT	81.58	81.80	81.34	78.83	78.85	77.62
AtGCN	86.83	87.17	86.10	86.03	85.97	85.32
ctGAN	83.26	83.68	82.82	80.73	80.65	80.22
FSR-XGB	84.94	85.56	84.30	82.63	82.45	82.82
Hybrid DRL	88.73	88.77	87.90	89.43	89.48	87.82
ICA-Hybrid DRL	89.65	90.29	88.98	92.03	92.58	90.42
TLO-Hybrid DRL	91.17	92.10	90.22	93.20	94.15	92.22
Hybrid DMADL	96.82	97.10	95.02	96.70	97.72	95.10
Hybrid DRL-CTO	94.00	94.60	92.62	94.95	95.93	93.66
Hybrid DMADL-CTO	97.67	97.99	95.89	97.52	98.59	96.01

5. CONCLUSION

The Hybrid DMADL-CTO is significantly designed for detecting the abnormalities of CA along with its class labels. In this research, the CTO algorithm and mixed attention module are effectually integrated to elevate the performance efficacy and convergence rate by reducing the computational cost, complexities, and memory constraints problems. Additionally, the instructive optimal frames are selected by an optimized multi-metric frame selection method to perform further detection operations. Furthermore, the increased operational time of the previous model challenge is effectively addressed by the implementation of a distributed learning approach. Meanwhile, the mixed attention module boosts the ability and effectiveness precisely. Due to these methods, the Hybrid DMADL-CTO effectually classified the class labels of CA with enhanced performance efficiency, which is then compared with other related existing approaches with numerous evaluation measures under maximal TP and KF 3scores. Thus, the higher performance efficacy of the Hybrid DMADL-CTO is obtained under maximal TP validation of accuracy, recall, and precision measures that achieve 97.67%, 97.99%, and 95.89% respectively. Furthermore, future research for CA detection explores diverse learning techniques with the insertion of distinct optimization algorithms to achieve precise performance with better classification effectiveness.

REFERENCES

- [1] Pau, M., Porta, M., Pau, C., Tacconi, P. and Sanna, A., "Quantitative characterization of gait patterns in individuals with spinocerebellar ataxi", *Bioengineering*, vol.10, no.7, pp.788,2023.
- [2] Shanmuga Sundari, M. and Jadala, V.C., "Neurological disease prediction using impaired gait analysis for foot position in cerebellar ataxia by ensemble approach", *Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije*, 64(3), pp.540-549, 2023.
- [3] Seetharama, P.D. and Math, S., "Ataxia severity classification using enhanced feature selection and ranking optimization through a machine learning model", *Indonesian Journal of Electrical Engineering and Computer Science*, 32(3), pp.1605-1613, 2023.
- [4] Bania, K. and Verlekar, T., "AtGCN: A Graph Convolutional Network For Ataxic Gait Detection", *arXiv preprint arXiv:2410.22862*, 2024.
- [5] Trabassi, D., Castiglia, S.F., Bini, F., Marinozzi, F., Ajoudani, A., Lorenzini, M., Chini, G., Varrecchia, T., Ranavolo, A., De Icco, R. and Casali, C., "Optimizing Rare Disease Gait Classification through Data Balancing and Generative AI: Insights from Hereditary Cerebellar Ataxia", *Sensors*, 24(11), p.3613, 2024.

- [6] Vattis, K., Oubre, B., Luddy, A.C., Ouillon, J.S., Eklund, N.M., Stephen, C.D., Schmahmann, J.D., Nunes, A.S. and Gupta, A.S., "Sensitive quantification of cerebellar speech abnormalities using deep learning models" IEEE Access, 2024.
- [7] Faber, J., Kügler, D., Bahrami, E., Heinz, L.S., Timmann, D., Ernst, T.M., Deike-Hofmann, K., Klockgether, T., van de Warrenburg, B., van Gaalen, J. and Reetz, K., "CerebNet: A fast and reliable deep-learning pipeline for detailed cerebellum sub-segmentation", Neuroimage, 264, p.119703, 2022.
- [8] Cabeza-Ruiz, R., Velázquez-Pérez, L., Linares-Barranco, A. and Pérez-Rodríguez, R., "Convolutional neural networks for segmenting cerebellar fissures from magnetic resonance imaging", Sensors, 22(4), p.1345, 2022.
- [9] Radmard, S., Zesiewicz, T.A. and Kuo, S.H., "Evaluation of cerebellar ataxic patients", Neurologic clinics, 41(1), pp.21-44, 2023.
- [10] Suzuki, M., Hirano, S., Otte, K., Schmitz-Hübsch, T., Izumi, M., Tamura, M., Kuroiwa, R., Sugiyama, A., Mori, M., Röhling, H.M. and Brandt, A.U., "Digital motor biomarkers of cerebellar ataxia using an RGB-Depth camera-based motion analysis system", The Cerebellum, 23(3), pp.1031-1041, 2024.
- [11] Seemann, J., Daghsen, L., Cazier, M., Lamy, J.C., Welter, M.L., Giese, M.A., Synofzik, M., Durr, A., Ilg, W. and Coarelli, G., "Digital gait measures capture 1-year progression in early-stage spinocerebellar ataxia type 2", Movement Disorders, 39(5), pp.788-797, 2024.
- [12] [Timm, E.C., Purcell, N.L., Ouyang, B., Berry-Kravis, E., Hall, D.A. and O'Keefe, J.A., "Potential Prodromal Digital Postural Sway Markers for Fragile X-Associated Tremor/Ataxia Syndrome (FXTAS) Detected via Dual-Tasking and Sensory Manipulation", Sensors, 24(8), p.2586, 2024..
- [13] Koziol, L.F., Budding, D., Andreasen, N., D'Arrigo, S., Bulgheroni, S., Imamizu, H., Ito, M., Manto, M., Marvel, C., Parker, K. and Pezzulo, G., "Consensus paper: the cerebellum's role in movement and cognition", The Cerebellum, 13, pp.151-177, 2014.
- [14] Guell, X., Gabrieli, J.D. and Schmahmann, J.D., "Embodied cognition and the cerebellum: perspectives from the dysmetria of thought and the universal cerebellar transform theories", Cortex, 100, pp.140-148, 2018.
- [15] Schmahmann, J.D., "The cerebellum and cognition", Neuroscience letters, 688, pp.62-75, 2019.
- [16] Stoodley, C.J. and Schmahmann, J.D., "Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies", Neuroimage, 44(2), pp.489-501, 2009.
- [17] Galna, B., Barry, G., Jackson, D., Mhiripiri, D., Olivier, P. and Rochester, L., "Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease", Gait & posture, 39(4), pp.1062-1068, 2014.
- [18] Schmitz-Hübsch, T., Du Montcel, S.T., Baliko, L., Berciano, J., Boesch, S., Depondt, C., Giunti, P., Globas, C., Infante, J., Kang, J.S. and Kremer, B., "Scale for the assessment and rating of ataxia: development of a new clinical scale", Neurology, 66(11), pp.1717-1720, 2006.
- [19] Rusz, J., Benova, B., Ruzickova, H., Novotny, M., Tykalova, T., Hlavnicka, J., Uher, T., Vaneckova, M., Andelova, M., Novotna, K. and Kadrnozkova, L., "Characteristics of motor speech phenotypes in multiple sclerosis", Multiple sclerosis and related disorders, 19, pp.62-69, 2018.
- [20] Kashyap, B., Pathirana, P.N., Horne, M., Power, L. and Szmulewicz, D., "Automated tongue-twister phrase-based screening for Cerebellar Ataxia using Vocal tract Biomarkers", In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 7173-7176). IEEE, July 2019.
- [21] Kashyap, B., Horne, M., Pathirana, P.N., Power, L. and Szmulewicz, D., "Automated topographic prominence-based quantitative assessment of speech timing in cerebellar ataxia", Biomedical Signal Processing and Control, 57, p.101759, 2020.
- [22] Mancini, M., King, L., Salarian, A., Holmstrom, L., McNamara, J. and Horak, F.B., "Mobility lab to assess balance and gait with synchronized body-worn sensors", Journal of bioengineering & biomedical science, p.007, 2011.
- [23] Shumway-Cook, A. and Horak, F.B., "Assessing the influence of sensory interaction on balance: a suggestion from the field", Physical therapy, 66(10), pp.1548-1550, 1986.
- [24] Hasegawa, N., Shah, V.V., Carlson-Kuhta, P., Nutt, J.G., Horak, F.B. and Mancini, M., "How to select balance measures sensitive to Parkinson's disease from body-worn inertial sensors—separating the trees from the forest", Sensors, 19(15), p.3320, 2019.

- [25] Mancini, M., Priest, K.C., Nutt, J.G. and Horak, F.B., "Quantifying freezing of gait in Parkinson's disease during the instrumented timed up-and-go test", In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 1198-1201). IEEE, August 2012.
- [26] Powell, L.E. and Myers, A.M., "The activities-specific balance confidence (ABC) scale", The Journal of Gerontology Series A: Biological sciences and Medical sciences, 50(1), pp.M28-M34, 1995.
- [27] The READISCA dataset," <https://rochester.app.box.com/v/AtaxiaDataset>" accessed on November 2024.
- [28] Buckley, E., Mazzà, C. and McNeill, A., "A systematic review of the gait characteristics associated with Cerebellar Ataxia", Gait & posture, 60, pp.154-163, 2018.
- [29] Atashpaz-Gargari, E. and Lucas, C., "Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition", In 2007 IEEE Congress on evolutionary computation (pp. 4661-4667). Ieee, September 2007.
- [30] Rao, R.V., Savsani, V.J. and Vakharia, D.P., "Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems", Computer-aided design, 43(3), pp.303-315, 2011.
- [31] Gao, S., Yun, J., Zhao, Y. and Liu, L., "Gait-D: skeleton-based gait feature decomposition for gait recognition", IET Computer Vision, 16(2), pp.111-125, 2022.
- [32] Singh, A.K., Kumbhare, V.A. and Arthi, K., "Real-time human pose detection and recognition using media pipe", In International conference on soft computing and signal processing (pp. 145-154). Singapore: Springer Nature Singapore, June 2021.
- [33] Oda, M., Furukawa, K., Navab, N. and Mori, K., "KST-Mixer: kinematic spatiotemporal data mixer for colon shape estimation", Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 11(4), pp.1050-1056, 2023.
- [34] Ghosal, P., Nandanwar, L., Kanchan, S., Bhadra, A., Chakraborty, J. and Nandi, D., "Brain tumor classification using ResNet-101 based squeeze and excitation deep neural network", In 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP) (pp. 1-6). IEEE, February 2019.
- [35] Zhou, L., Wang, M. and Zhou, N., "Distributed federated learning-based deep learning model for privacy MRI brain tumor detection", arXiv preprint arXiv:2404.10026, 2024.
- [36] Khan, M.A. and Kim, Y., "Cardiac Arrhythmia Disease Classification Using LSTM Deep Learning Approach. Computers", Materials & Continua, 67(1), 2021.
- [37] Woo, S., Park, J., Lee, J.Y. and Kweon, I.S., "Cbam: Convolutional block attention module", In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 3-19), 2018.
- [38] Al-Asadi, M.A. and Tasdemir, S., "Using artificial intelligence against the phenomenon of fake news: a systematic literature review", Combating fake news with computational intelligence techniques, pp.39-54, 2022.
- [39] Stoean, C., Stoean, R., Atencia, M., Abdar, M., Velázquez-Pérez, L., Khosravi, A., Nahavandi, S., Acharya, U.R. and Joya, G., "Automated detection of presymptomatic conditions in Spinocerebellar Ataxia type 2 using Monte Carlo dropout and deep neural network techniques with electrooculogram signals", Sensors, vol.20, no.11, pp.3032, 2020.
- [40] Procházka, A., Dostál, O., Cejnar, P., Mohamed, H.I., Pavelek, Z., Vališ, M. and Vyšata, O., "Deep learning for accelerometric data assessment and ataxic gait monitoring", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol.29, pp.360-367, 2021.
- [41] Seetharama, P.D. and Math, S., "Ataxic person prediction using feature optimized based on machine learning model", International Journal of Electrical and Computer Engineering (IJECE), vol.14, no.2, pp.2100-2109, 2024.