**Journal of Electronics, Electromedical Engineering, and Medical Informatics**
**Multidisciplinary: Rapid Review: Open Access Journal** **Vol. 7, No. 2, April 2025, pp: 283-294; eISSN: 2656-8632**

**RESEARCH ARTICLE** **OPEN ACCESS**

Manuscript received December 7, 2024; Revised January 26, 2025; Accepted January 28, 2024; date of publication February 22, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i2.650

How to cite: Saravanakumar C, Marirajan S, Pandian A, and Durgadevi K., "Advancing Genomic Diagnostics: Fast Fourier Transform Optimization and Machine Learning in Huntington's Disease Detection", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 7, no. 2, pp. 283-294, April 2025.

# Advancing Genomic Diagnostics: Fast Fourier Transform Optimization and Machine Learning in Huntington's Disease Detection

**Saravanakumar C, Marirajan S, Pandian A, and Durgadevi K**

Department of Electronics and Communication Engineering, SRM Valliammai Engineering College, Tamilnadu, India

Corresponding author: Saravanakumar C (e-mail: kanchi.saravana@gmail.com)

**ABSTRACT** Optimizing the Fast Fourier Transform (FFT) for genomic data analysis offers a significant advancement in addressing challenges related to sequential input processing and computational efficiency. By integrating advanced signal processing techniques such as Infinite Impulse Response (IIR) filtering, the proposed approach effectively identifies spectral characteristics and dominant frequencies in DNA sequences. This framework demonstrates improved accuracy and reduced computational overhead, making it highly suitable for large-scale and real-time genomic applications. Machine learning models were employed to classify Huntington's Disease (HD)-associated and normal DNA sequences, using spectral features as predictive markers. Among the models evaluated, K-Nearest Neighbors (KNN) achieved perfect scores across all performance metrics, including Classification Accuracy (CA), Area Under the Curve (AUC), Precision, Recall, Matthews Correlation Coefficient (MCC) and F1 Score. Support Vector Machine (SVM) and Neural Networks also delivered competitive results, emphasizing the effectiveness of combining signal processing with machine learning for medical diagnostics and genomic studies. The computational efficiency of the proposed FFT algorithm was validated using 2,300 genomic sequences, with 90% demonstrating enhanced processing speeds compared to traditional methods. These improvements were particularly notable for longer sequences, showcasing the algorithm's capability in high-throughput genomic analysis. This approach is particularly impactful for investigating complex conditions like Huntington's disease, where rapid and accurate identification of genetic markers is essential. This work underscores the potential of integrating FFT optimization with machine learning to revolutionize genomic data processing and disease detection. Beyond advancing computational genomics, the proposed methodology offers a foundation for broader bioinformatics applications, including the analysis of other genetic disorders and real-time clinical diagnostics, contributing to the evolution of precision medicine.

**INDEX TERMS** Classification Accuracy, Fast Fourier Transform, Genomic Data Analysis, Huntington's Disease, Machine Learning, Signal Processing.

## I. INTRODUCTION

Since the introduction of the Fast Fourier Transform (FFT) [1], significant advancements have been made in developing FFT algorithms tailored for various applications. These advancements include methods focused on reducing computational complexity, such as the split-radix algorithm, as well as techniques optimized for specific use cases, including decimation in time and frequency, handling subsets of input or output points, and processing input samples of arbitrary lengths. However, most conventional FFT algorithms assume that all input samples are readily available prior to computation [2]. As a result, their primary design goal is to minimize arithmetic operations, often overlooking scenarios where input samples arrive sequentially.

In many real-world applications, particularly in streaming systems or real-time signal processing, input data is received sequentially, making it imperative to optimize the FFT process for such conditions. Reducing the time required to complete FFT operations in these scenarios is crucial for minimizing buffer memory requirements and adhering to strict timing

constraints, such as the "gap time," which represents the allowable interval between receiving the final input sample and transmitting the first output sample [3]. Addressing these challenges requires FFT algorithms designed to operate efficiently with sequentially received data.

To evaluate the performance of such algorithms, this work introduces the concept of "completion delay," defined as the additional time needed to complete FFT operations after all N input samples have been received. This delay is measured in clock cycles, with one cycle corresponding to the sampling duration of the input sequence. At lower sampling rates, FFT computations may be completed within the same clock cycle as the final input sample, resulting in minimal delay [4]. Conversely, at higher sampling rates or with slower FFT operations, the completion delay may extend over several clock cycles. Thus, algorithms that achieve shorter completion delays are highly desirable for improving the efficiency of FFT operations in systems with sequential data acquisition as the one in Gene data sets [5].

The field of bioinformatics has emerged as one of the most dynamic and innovative areas within modern scientific research. Despite the human genome comprising approximately three billion genetic elements, the pace at which genomic data is being generated continues to accelerate exponentially. One of the key challenges faced by scientists today is the effective interpretation of these vast genomic sequences [6]. Researchers worldwide are dedicated to decoding the information embedded within DNA sequences, exploring various aspects such as metadata peculiarities, nucleotide sequence characteristics, classification techniques, and genetic disorder analysis [7].

Huntington's disease (HD) is a genetic neurodegenerative complaint that culminates in multiple aspects of a person's physical, cognitive, and emotional well-being. Caused by a genetic mutation in the huntingtin (HTT) gene, HD is characterized by the progressive degeneration of brain cells, leading to a broad spectrum of symptoms. The HTT gene mutation leads to an unusual expansion of the CAG nucleotide distribution, with individuals carrying 40 or more CAG repeats being guaranteed to develop the disease. This mutation disrupts the production of the huntingtin protein, which plays essential but not fully understood roles in brain development and function [8]. Over time, the mutant protein accumulates in brain cells, particularly in the striatum, resulting in significant neuronal damage and the onset of HD symptoms.

Huntington's disease generally presents among the ages of thirty and fifty, however juvenile-onset instances may arise in younger individuals. The disease is often referred to as a "family disorder" due to its fifty percent possibility of acquiring the mutated gene [9]. Globally, HD affects approximately 41,000 symptomatic individuals in the United States, with over 200,000 people at risk.

The clinical presentation of HD is multifaceted. Early symptoms often include personality changes, mood swings, depression, and cognitive impairments such as memory loss and difficulty making decisions. Physical manifestations include involuntary movements (chorea), slurred speech, swallowing difficulties, and unintentional weight loss [10]. As the disease progresses, motor functions deteriorate further, with individuals experiencing stiffness, reduced mobility, and challenges in communication. Advanced stages necessitate full-time care due to the cumulative impact on reasoning, movement, and daily functioning, with complications such as pneumonia or heart failure often leading to death [11].

Despite its devastating impact, there is currently no cure for HD, nor treatments capable of halting or reversing its progression [12]. Interventions focus on alleviating mood disturbances, managing involuntary movements, maintaining physical strength, and addressing communication and nutritional challenges. Social and community support also play a vital role in improving the wellbeing for people with HD and their caregivers, who often face significant emotional and physical challenges in providing care.

The complexity of HD, with its overlapping motor, cognitive, and psychiatric symptoms, has drawn comparisons to conditions such as ALS, Parkinson's disease, and Alzheimer's disease [13]. These parallels underscore the urgency of advancing research into effective treatments and the importance of exploring novel therapeutic strategies to demand the unmet needs of HD patients.

Currently, substantial research is focused on understanding and addressing Huntington's disease (HD). Researchers are investigating both conventional diagnostic methods and advanced machine learning (ML) approaches to differentiate HD from other neurodegenerative disorders and enhance diagnostic precision [14].

The proposed work employs advanced signal processing techniques to explore specific genetic configurations, demonstrating their effectiveness in feature extraction for classification systems. The primary goal is to establish a framework for categorizing HD samples by applying signal processing methods to nucleotide sequences relevant to Huntington's disease. A classification model is developed to detect the presence of HD, highlighting the potential of integrating bioinformatics and signal processing to tackle key challenges in genomic analysis and medical diagnostics.

## II. RELATED WORKS

The study in [15] examines the Fast Fourier Transform (FFT), a crucial method in digital signal processing, and emphasizes its limitations, including discrete frequency intervals and spectrum leakage. This work presents Prism Signal Processing (PSP) as a supplementary approach to tackle these difficulties. PSP employs linear phase FIR filters for real-time analysis, improving frequency resolution and facilitating the detection of low-amplitude tones. The suggested method integrates PSP with FFT using low-pass filters and heterodyning to reduce spectral leakage and clarify multiple peaks. This method, while computationally costly, provides substantial improvements in spectral analysis for applications requiring high precision.

The Periodic Interpolation Method introduced in [16], an adaptable and fruitful approach for calculating scalar potentials within a unit cell of an endlessly periodic array. FFT-PIM addresses one-dimensional, two-dimensional, and three-dimensional periodicities for both dynamic and static potential issues, supporting situations with or without periodic phase shifts. FFT-PIM utilizes fast converging Green's function series to partition the potential into near-zone and far-zone components. Applications encompass wave propagation, micromagnetic solvers, rendering it optimal for periodic integral equation challenges.

The investigation in [17] explores the respiratory rate (RR) as a critical marker of heat stress and respiratory ailments in animals, emphasizing a technique for forecasting RR in recumbent Holstein cows by RGB and infrared (IR) imaging. A dataset including 95 videos was annotated to delineate the flank area of interest (ROI), utilizing YOLOv8 for automatic ROI recognition. Respiratory frequencies were filtered and breaths per minute were computed using a pipeline that incorporated fast Fourier transform (FFT) and inverse FFT. Evaluation metrics indicated RMSEP values of 8.3 breaths/min for cows and 13.0 for calves, underscoring the method's efficacy in precise respiratory rate estimation.

In order to meet the demand for energy-efficient, real-time signal processing, the research technique discussed in [18] introduces a unique FPGA-based Fast Fourier Transform (FFT) processor. Conventional FFT processor architectures encounter constraints, such as elevated power consumption, which limits their applicability in low-power scenarios. A hybrid radix encoder featuring a 2 fold multiplier with an improved truncated Kogge-Stone adder is presented to address this issue. The developed FFT processor attains elevated throughput (78.036 Gbps), lowered consumption of power, and negligible delay. This processor supports FFT widths upto 4096 points, designed for biomedical and 5G applications.

The work in [19] examines the necessity for effective health monitoring applications in a context of growing patient-to-physician disparity. It presents innovative algorithms for assessing blood pressure and heart rate through photoplethysmography signals obtained from an Android phone's flashlight camera. The proposed Android application provides an economical, rapid, and energy-efficient solution, compatible with Android 5.0 and subsequent versions. Experimental results indicate an accuracy reaching 98% in comparison to digital medical instruments, surpassing prior methodologies. Future endeavors seek to amalgamate data collecting and machine learning methodologies to facilitate emotion analysis for the identification of depression, hence augmenting the application's functionality.

Due to its reliance on adders and multipliers as essential components, the butterfly operation is the Fast Fourier Transform (FFT) system's most computationally demanding stage. The work in [20] presents a low-power butterfly unit for FFT architectures by consolidating addition and subtraction operations into one, hence decreasing common expressions and cutting transistor count. Simulations indicate substantial enhancements in area, power, and delay, with diminished power consumption realized through the optimization of supply voltage and load capacitance. The proposed unit accommodates 8, 12, and 16-bit data widths, resulting in a reduction of 16 clock cycles for 8-bit operations during single-clock execution. Monte Carlo simulations validate negligible fluctuations in process parameters, while complexity is diminished by 43% relative to conventional designs by the integration of the Gate Diffused Inputs approach.

Recent improvements in Human-Computer Interaction (HCI) have profoundly influenced signal processing within the healthcare sector, especially with the interpretation of ECG, EMG, and EEG signals. The ECG, which provides significant insights into an individual's emotional state, has demonstrated potential for application in biometric identification, including gender classification. The research in [21] investigates the utilization of the RF algorithm for classification employing ECG data from the ECG ID Database. The findings indicate that raw data attains an accuracy of 55.000%, but filtered data enhances this to 65.806%, highlighting the significance of data preparation in improving classification efficacy.

The review in [22] critically assesses contemporary pharmacological strategies for Huntington's disease, emphasizing the significance of mutant huntingtin protein, mitochondrial failure, excitotoxicity, and neuroinflammation. Emerging therapeutic targets, including protein homeostasis, neurotransmitter systems, and mitochondrial function, are examined. The study emphasizes clinical trial results, the repurposing of current pharmaceuticals, and advocates for the incorporation of personalized medicine and combination therapy as prospective avenues in Huntington's disease treatment.

The HD study in [23] uses genetic, neurobiological, and clinical markers to understand Huntington's disease (HD) before clinical diagnosis in people with a genetic mutation. The study estimates the beginning and first development of HD-related variables relative to diagnosis. This study investigated 438 persons who had the HD gene mutation but did not meet HD diagnostic criteria or show functional impairment. Basic cognitive, motor, psychiatric, and neuroimaging assessments were tested for predictive significance using nonlinear models, with time to diagnosis as the predictor. The expected time to diagnosis was associated with most clinical and neuroimaging markers, with alterations appearing one to two decades before the predicted diagnosis.

In addition to other molecular processes such as the depletion of neurotrophic factor, mitochondrial malfunction, and abnormal synaptic plasticity in central spiny neurons, the research displayed in [24] demonstrates that one of the most important characteristics of HD is the growth of polyglutamine duplications in proteins. Animal models have been extremely helpful in expanding our understanding of the pathophysiology of HD as well as the therapeutic approaches that have been developed. Animal models that mimic the pathophysiology of the illness are vital for investigating the

underlying mechanisms of HD and evaluating prospective therapies. Although there are limited treatments currently available for HD, these models continue to be essential. The importance of in-vitro and in-vivo models for HD research is discussed in this review, along with the role that these models play in the process of drug screening for this debilitating disorder.

The work in [25] underscores the important fact that neuro degenerative alterations that occur as a result of striatal dysfunction in Huntington's disease (HD) lower sensitivity to linguistic principles, notably in the handling of referential enslavements that are constrained by grammatical constraints. Additionally, HD makes it more difficult to piece together events by use of verb phrases and the thematic arguments they contain. The significance of cortical-subcortical loops in reference processing is highlighted by these findings, which also indicate that there is a possibility that these impairments would occur simultaneously with other cognitive deficiencies, such as disorder of theory of mind [26]. The study shows disparities in language sensitivity between spontaneous speech and controlled studies, with spontaneous speech revealing early indicators of linguistic degeneration. In terms of clinical applications, the study calls for a focus on language understanding at the grammatical level.

In [27], the brain samples of one hundred and fifty seven HD patients and its controls were examined using gene profiling and dimensionality reduction. This was followed by the application of ML procedures such as decision tree, rule induction, random forest, and generalized linear model. A high level of cross-validation accuracy was achieved by these models, which revealed 66 candidate genes associated to HD. The accuracy ranged from 89.49% to 97.46%. The genes that were discovered were linked to important biological processes such as the regulation of transcription and the cytoskeleton. The results of this study indicate that the HTT gene mutation may be responsible for the disruption of these genes, which in turn contributes to the pathogenesis of HD [28].

In DSP-based exon prediction, the critical initial step involves converting nucleotide bases into numerical values, as highlighted in [29]. The choice of numerical mapping scheme plays a significant role in shaping the characteristics of the DNA sequence, enabling the accurate identification of exon regions. Over the past two decades, several nucleotide mapping techniques have been effectively used as preprocessing steps for exon prediction [30]. The proposed sequence mapping method demonstrates superior performance in predicting exon regions compared to existing approaches [31].

## III. MATERIALS AND METHODS

Deoxyribonucleic acid (DNA) is a complex molecule composed of two long polynucleotide chains arranged in a double helix structure. This structure is responsible for storing the biological information necessary for the growth, development, and reproduction of all living organisms, including humans, animals, plants, and viruses. DNA is made up of four nucleotides: adenine (A), guanine (G), cytosine (C), and thymine (T). The adenine always pairs with thymine, and cytosine always pairs with guanine. This pairing guarantees the accurate transmission of genetic information, which is crucial for DNA replication and cell function. The sequence of these nucleotides forms the genetic code, providing instructions for the production of proteins that are essential for life processes.

To analyze DNA sequences more effectively, they are often encoded into numerical data that can be processed using computational methods, particularly signal processing techniques. This conversion allows researchers to perform various operations on the data, such as sequence alignment, pattern recognition, and prediction of protein structures. Each nucleotide is assigned a specific numerical value, and different encoding methods are used, such as binary encoding or more complex schemes based on the properties of the nucleotides. Once encoded, the sequences can be analyzed using mathematical models to detect significant patterns and features within the DNA, which tends to better impression of biological processes and disease mechanisms.

A Frequency-Based Weighted Encoding Scheme is proposed to convert the text into numbers. The idea is to quantify the appearance of each nucleotide in a given DNA sequence, turning the sequence into a numerical vector that retains essential information about the sequence's composition.

In this scheme, each character of the DNA sequence is encoded based on its frequency relative to the total number of characters in the sequence. For instance, consider the sequence in Eq. (1) [3].

$$\text{originalSeq(n)} = \text{" AGATCGATGA"} \quad (1)$$

In this case, there are ten nucleotides in total, and we calculate the frequency of each base. In this specific sequence, adenine (A) appears four times, guanine (G) appears three times, thymine (T) appears two times, and cytosine (C) appears once. The next step involves calculating the frequency of each nucleotide by dividing its occurrence by the total length of the sequence. Therefore, the encoded values for each nucleotide are as follows: A = 0.4, G = 0.3, T = 0.2, and C = 0.1. This frequency ratio is used as the numeric encoding for the respective nucleotides. After applying this method across all positions in the sequence, we obtain the encoded sequence. For example, the original DNA sequence in Eq. (1) would be represented numerically by following the encoding scheme as the sequence in Eq. (2) [3]

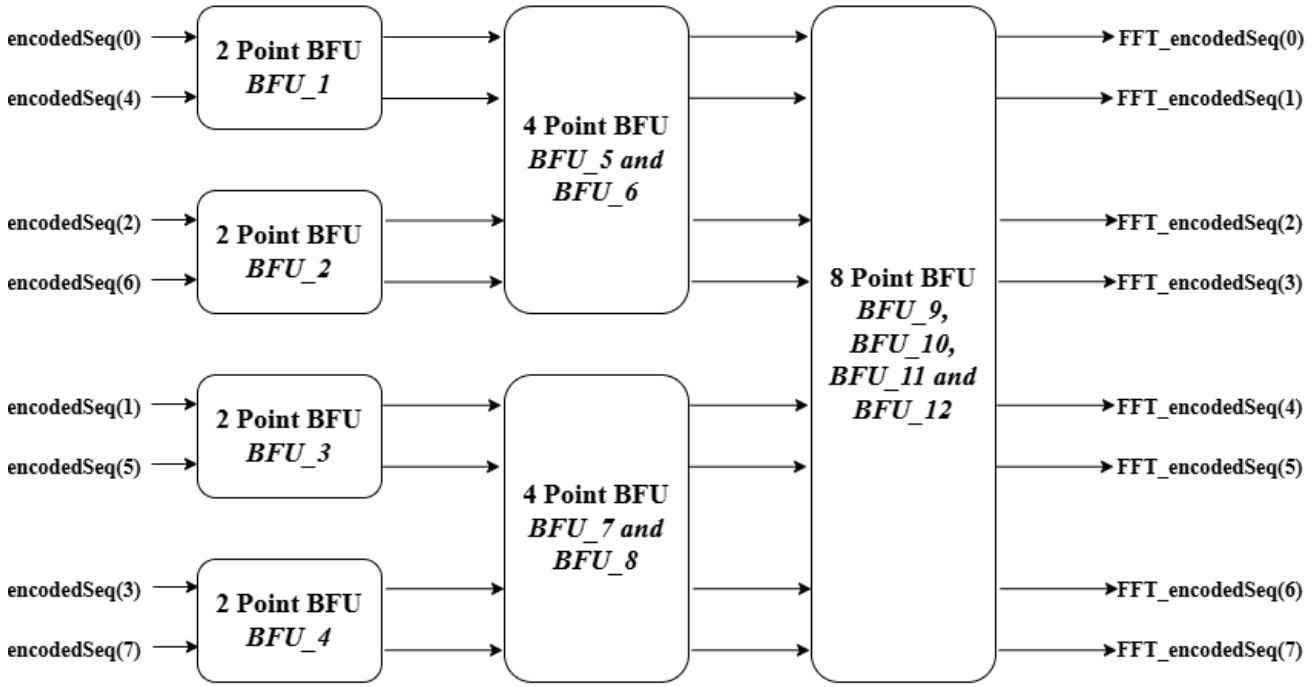



**FIGURE 1.** SFG of Fast Fourier Transform

$$\text{encodedSeq(n)} = [\ 0.4,0.3,0.4,0.1,0.1,0.3,0.4,0.2,0.3,0.4] \quad (2)$$

This numerical representation preserves the nucleotide composition of the original DNA sequence while transforming it into a format suitable for computational analysis. This encoding method is highly useful when DNA sequences need to be processed using signal processing techniques, such as in bioinformatics applications, where large volumes of DNA sequence data need to be analyzed efficiently.

When a DNA sequence is encoded into a numerical format, it becomes crucial to implement a technique that can minimize the data size while preserving the integrity and accuracy of the encoded information. To achieve this, the Discrete Fourier Transform (DFT) can be applied to the encoded numerical sequence. This process generates X_encodedSeq(n), which highlights the periodicity of three inherent in the DNA chain. However, the DFT technique is computationally intensive, as it demands a significant number of multipliers and adders. The computational effort required grows with the size of the DNA sequence, making it less

$$\begin{aligned}
X[0] &= [(x[0]+x[4])+(x[2]+x[6])W_8^0]+[(x[1]+x[5])+(x[3]+x[7])W_8^0]W_8^0\\
X[1] &= [(x[0]-x[4])+(x[2]-x[6])W_8^2]+[(x[1]-x[5])+(x[3]-x[7])W_8^2]W_8^1\\
X[2] &= [(x[0]+x[4])-(x[2]+x[6])W_8^0]+[(x[1]+x[5])-(x[3]+x[7])W_8^0]W_8^2\\
X[3] &= [(x[0]-x[4])-(x[2]-x[6])W_8^2]+[(x[1]-x[5])-(x[3]-x[7])W_8^2]W_8^3\\
X[4] &= [(x[0]+x[4])+(x[2]+x[6])W_8^0]-[(x[1]+x[5])+(x[3]+x[7])W_8^0]W_8^0\\
X[5] &= [(x[0]-x[4])+(x[2]-x[6])W_8^2]-[(x[1]-x[5])+(x[3]-x[7])W_8^2]W_8^1\\
X[6] &= [(x[0]+x[4])-(x[2]+x[6])W_8^0]-[(x[1]+x[5])-(x[3]+x[7])W_8^0]W_8^2\\
X[7] &= [(x[0]-x[4])-(x[2]-x[6])W_8^2]-[(x[1]-x[5])-(x[3]-x[7])W_8^2]W_8^3
\end{aligned} \quad (3)$$

$$X[0] = [(x[0] + x[4]) + (x[2] + x[6])W_8^0] + [(x[1] + x[5])W_8^0 + (x[3])W_8^0] + (x[7])W_8^0$$
$$X[1] = [(x[0] - x[4]) + (x[2] - x[6])W_8^2] + [(x[1] - x[5])W_8^1 + (x[3])W_8^3] - (x[7])W_8^3$$
$$X[2] = [(x[0] + x[4]) - (x[2] + x[6])W_8^0] + [(x[1] + x[5])W_8^2 - (x[3])W_8^2] - (x[7])W_8^2$$
$$X[3] = [(x[0] - x[4]) - (x[2] - x[6])W_8^2] + [(x[1] - x[5])W_8^3 + (x[3])W_8^1] - (x[7])W_8^1$$
$$X[4] = [(x[0] + x[4]) + (x[2] + x[6])W_8^0] - [(x[1] + x[5])W_8^0 + (x[3])W_8^0] - (x[7])W_8^0$$
$$X[5] = [(x[0] - x[4]) + (x[2] - x[6])W_8^2] - [(x[1] - x[5])W_8^1 + (x[3])W_8^3] + (x[7])W_8^3$$
$$X[6] = [(x[0] + x[4]) - (x[2] + x[6])W_8^0] - [(x[1] + x[5])W_8^2 - (x[3])W_8^2] + (x[7])W_8^2$$
$$X[7] = [(x[0] - x[4]) - (x[2] - x[6])W_8^2] - [(x[1] - x[5])W_8^3 + (x[3])W_8^1] + (x[7])W_8^1 \quad (4)$$

efficient for handling large datasets. This limitation can be effectively addressed by leveraging the Fast Fourier Transform, a algorithm for evaluating the DFT in a faster way. The FFT significantly reduces the number of required operations, especially for large

DNA sequences, thereby enhancing overall efficiency. The relationship between the size of the DNA sequence and the reduction in computational steps achieved through FFT underscores its advantage. Given this evidence, the FFT is employed as a practical method to compute the DFT of encodedSeq(n), thereby optimizing the computational process. The FFT not only simplifies the calculations but also accelerates the processing of encoded DNA sequences, making it particularly suitable for applications requiring real-time or large-scale data analysis. The mathematical evaluation of the FFT is carried out using the Eq. (5)[3]:

$$FFT\_encodedSeq(k) = \sum_{n=0}^{N-1} encodedSeq(n).W_N^{kn} \quad (5)$$

The term $W_N^{kn}$ is called as twiddle factor.The computational process of the fast Fourier transform is represented through a signal flow diagram (SFG), as shown in FIGURE 1. For simplicity, let us consider the FFT_encodedSeq is denoted as X and encodedSeq is considered as x. Based on the visual representation in Eq. (3) [3], the butterfly computational segments follow a sequential input processing pattern. When input samples x[0] through x[4] are received, the initial segment (BFU_1) becomes eligible for calculation. Upon receiving x[5], the segment BFU_3 becomes potentially evaluable. As data continues to arrive, specifically x[6], subsequent segments like BFU_2, BFU_5, and BFU_7 can be computed, contingent upon the results from preceding segments BFU_1 and BFU_3 being successfully transmitted.

The remaining computational segments follow a similar computational progression. Upon careful analysis of the mathematical expressions, the terms can be strategically divided into two distinct groups for simplified comprehension. Group1 requires inputs x[1], x[2], x[4], and x[6] to complete its computations. Conversely, group2 necessitates waiting until x[7] arrives for full evaluation. Interestingly, if certain expressions involving x[1] and x[5], such as their summation or difference, can be pre-calculated before x[7] arrives, it becomes possible to extract x[7] from the group2 calculations, as demonstrated in the subsequent illustration as Eq. (4) [3].

Prior to receiving x[7], all butterfly modules are initialized by setting x[7] to zero, with the preliminary results labeled as x'old. Once x[7] is obtained, the sequence is updated using a generalized computational expression that incorporates the previous stage's values and specific weighted coefficients as in Eq. (6) [3]

$$X = [(x'_{old}) + (x)W_N^{(N-1)k}] \quad (6)$$

The computational strategy reveals an optimization where only half the number of distinct weight values (N/2) are necessary. This reduction translates to N/2 multiplications and N additions required to determine the Fast Fourier Transform (FFT) sequence after acquiring the (N-1)th sample. For small sample sizes like 8 points, the performance enhancement might appear negligible. However, when applied to extensive datasets such as DNA sequences, the computational efficiency becomes significantly pronounced. The method substantially reduces processing time by minimizing computational redundancy, with the speed efficiency of the FFT directly correlating to the increased sample complexity. The key innovation lies in the algorithmic approach that progressively refines the transformation process, enabling faster and more streamlined spectral analysis across larger, more complex datasets. The algorithm behind the process is given as ALGORITHM 1.

Spectral characteristics of DNA sequences were evaluated using Fourier Transform techniques, with power spectral density estimated through periodogram analysis. Infinite impulse response (IIR) filtering was employed to optimize signal processing and smoothing of DNA chain spectral data.

**ALGORITHM 1.** Optimized Fast Fourier Transform

**Step 1: Initialize Variables**

1.1 Set x[7]=0 (Pre-calculation for missing input)

1.2 Define twiddle factors W[n][k] for FFT computation

**Step 2: Pre-Calculation Phase (Before x[7] is Received)**

2.1 Compute initial butterfly module outputs using x[0]-x[6]:

2.2 b[1]: Pass x[2] directly to b[4] and b[5] (No computation required)

2.3 b[3], b[4], b[5]: Compute FFT outputs assuming x[7] = 0

2.4 Compute temporary FFT outputs X0[k] (Intermediate results)

**Step 3: Re-Calculation Phase (After x[7] is Received)**

3.1 Collect x[7] and update FFT computations

3.2 Compute final outputs by adding x[7]* W[7][k] to precomputed values:

X[k]=X0[k]+x[7] * W[7][k] (for each butterfly module)

**Step 4: Generalizing Pre-Calculation for Other Inputs**

4.1 Set x[6]= 0 and precompute b[4], b[5] upon receipt of x[4]

4.2 Perform re-calculation once x[6] arrives using x[6] W[3][k]

4.3 Repeat pre-calculation and recalculation steps for subsequent inputs

**Step 5: Output Final FFT Coefficients**

5.1 Return computed FFT values X[0]-X[7]

**End Algorithm**

The choice of parameters for the Infinite Impulse Response (IIR) filter, including the cut-off frequency and filter order, was carefully determined to balance signal clarity, computational efficiency, and biological relevance in DNA sequence analysis. The cut-off frequency of 0.3607 radians per second was selected to preserve key periodic components of DNA sequences, particularly the 3-base periodicity in protein-coding regions, while effectively attenuating high-frequency noise. This value was likely determined through empirical testing and spectral analysis to ensure optimal signal retention. The filter order of 10 was chosen as a compromise between sharp frequency selectivity and computational stability, as higher-order filters provide steeper roll-off but may introduce phase distortions and numerical instability. A 10th-order filter ensures effective noise reduction while maintaining essential DNA sequence features. The use of an IIR filter was justified by its computational efficiency, requiring fewer coefficients than an FIR filter for a similar response, making it suitable for large-scale DNA analysis. These parameter choices directly impact the results by enhancing periodic signal detection, reducing noise interference, and improving the accuracy of downstream biological interpretations.

The analytical procedure involved transforming encoded DNA sequences using Fourier transform, subsequently applying the IIR low-pass filter. Periodogram analysis was then conducted to detect signal peaks and corresponding frequencies. This protocol was systematically applied to both

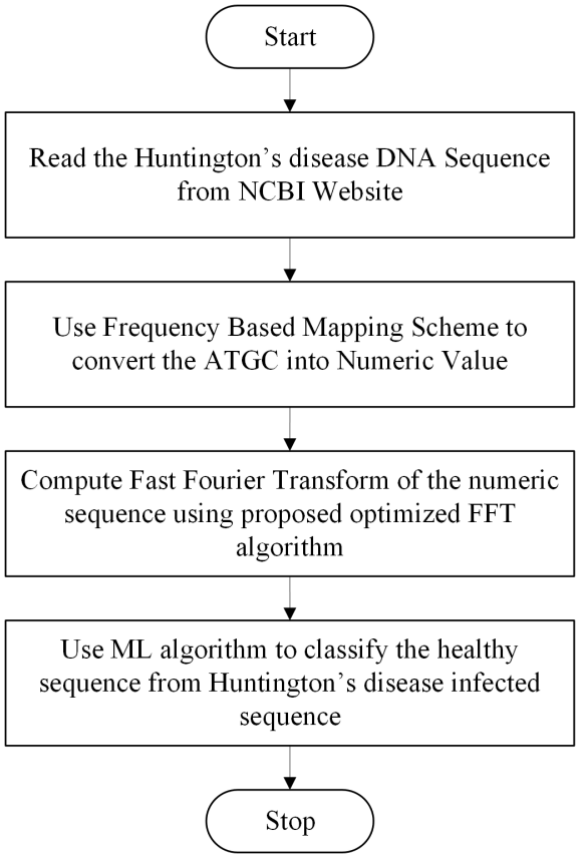



**FIGURE 2. Flowchart of the Process**

Huntington's Disease-associated and normal healthy DNA sequences. For each DNA sample, peak values and their corresponding frequencies were recorded, generating a comprehensive dataset. The experimental workflow is graphically represented in FIGURE 2 to facilitate methodological comprehension.

A machine learning approach is employed to distinguish between infected and normal DNA by constructing a classification model. This involves using a comprehensive training dataset with pre-labeled classes, agreeing the model to absorb the distinctive characteristics of each class. Four different classification schemes are utilized: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Naive Bayes, and Neural Networks. In the field of machine learning, the evaluation of model performance is critical for ensuring the reliability and effectiveness of predictive systems. Metrics such as Area Under the Curve (AUC), Classification Accuracy (CA), F1 Score, Recall, Precision, and Matthews Correlation Coefficient (MCC) provide comprehensive insights into how well a model classifies data.

The AUC represents the degree of separability between classes, quantifying the model's ability to distinguish positive from negative classes across various thresholds. An AUC of 0.5 indicates no discriminative ability, while an AUC of 1.0 signifies perfect classification. Meanwhile, CA simply measures the proportion of correctly predicted instances among the total instances, serving as a straightforward gauge of overall performance. However, in imbalanced datasets, CA can be misleading; hence, metrics like Recall and Precision become vital. Recall, also known as True Positive Rate, measures the proportion of actual positives that were correctly identified by the model, thereby highlighting the model's ability to capture relevant instances. In contrast, Precision assesses the accuracy of the positive predictions made by the model, reflecting the proportion of true positives among all predicted positives. The F1 Score harmonizes these two metrics by providing a single score that considers both Recall and Precision, offering a balanced measure particularly useful in scenarios where positive class instances are rare. This score is especially pertinent in applications like healthcare diagnostics or fraud detection, where the cost of false negatives can be significantly higher than false positives. Furthermore, the Matthews Correlation Coefficient (MCC) stands out as a robust metric that considers all four confusion matrix categories. MCC is particularly useful in evaluating binary classifiers, especially with imbalanced datasets, as it provides a clearer view of the model's predictive power than accuracy alone. Collectively, these metrics offer a multidimensional assessment of machine learning models. For instance, while a model may achieve high accuracy, its low recall or precision could render it inadequate for sensitive applications.

Therefore, when designing a machine learning system, it's vital to consider the specific context and goals of the application to select appropriate metrics that align with the desired outcomes. Real-world applications may call for prioritizing metrics that minimize false negatives, such as Recall, in critical fields like medical imaging or risk assessment in finance. Conversely, in applications where the consequences of false positives can be dire, such as spam detection, Precision might take precedence. In summary, AUC, CA, F1 Score, Recall, Precision, and MCC are instrumental in dissecting model performance, providing a nuanced understanding that enables practitioners to tailor their models to specific tasks and datasets. An informed choice of these metrics, guided by the nature of the data and the implications of prediction errors, can significantly enhance the utility of machine learning solutions across various domains. Therefore, a comprehensive analysis involving multiple performance metrics is not only advisable but essential for the valid assessment of any machine learning model's efficacy and reliability.

## IV. RESULTS AND DISCUSSION

In this study, the computational efficiency of the proposed Fast Fourier Transform (FFT) technique is compared with the traditional FFT method across a dataset comprising 2300 distinct genomic sequences. The formula to calculate the speed of calculation for Fast Fourier Transform, in percentage for the proposed system compared to the reference system is

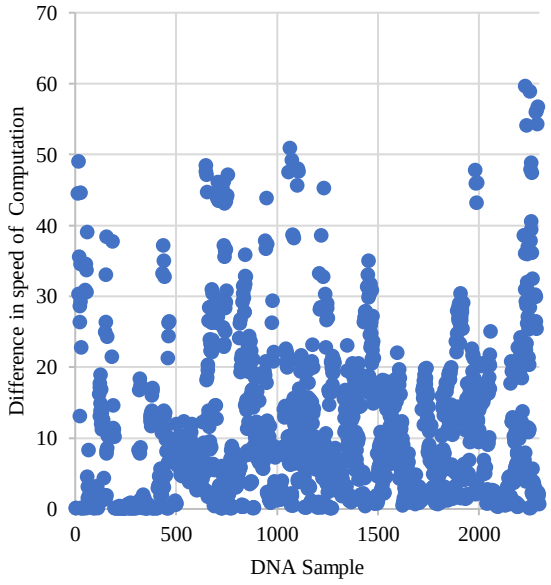



**FIGURE 3. Scatter Plot depicting the Improvement in Computation Speed**

given in Eq. (7) [3] below.

$$SI\ (\%) = [(Time_{Prop} - Time_{Ref}) / Time_{Prop}] * 100 \quad (7)$$

Here, Speed Improvement is denoted as SI, Time taken by the reference system to perform the calculation is denoted as $Time_{Ref}$ and Time taken by the proposed system to perform the calculation is denoted as $Time_{Prop}$. This formula quantifies the percentage reduction in computation time achieved by the proposed system relative to the reference system. The scatter plot presented in FIGURE 3 illustrates the findings: out of the 2300 sequences, 240 experienced a decrease in computational speed. This reduction is largely attributed to the shorter lengths of these sequences. The reduction in computational efficiency is primarily attributed to the shorter lengths of these sequences, which, contrary to expectations, require more execution time. This is because shorter sequences necessitate more frequent recalculations during the processing steps, as there is less pre-calculation data available to optimize the execution. On the other hand, longer sequences, despite containing more data, benefit from the pre-calculation process, which allows for more efficient feature extraction and reduces the overall execution time. The pre-calculation of certain metrics or transformations for longer sequences results in faster processing, as the computational effort is distributed more evenly across the data. Therefore, longer sequences can

leverage these optimizations, leading to shorter execution times compared to their shorter counterparts.

Conversely, an impressive 90% of the samples demonstrated enhanced processing speed. This significant improvement underscores the effectiveness of the proposed algorithm, warranting its designation as the "Optimized Genomic FFT Algorithm." The results indicate that the algorithm excels in handling genomic data, particularly larger datasets, thus greatly reducing computation time.

The enhanced speed performance is attributed to the algorithm's ability to minimize processing overhead while maximizing efficiency in spectral analysis. This makes it particularly beneficial for tasks in computational genomics where rapid data processing is crucial.

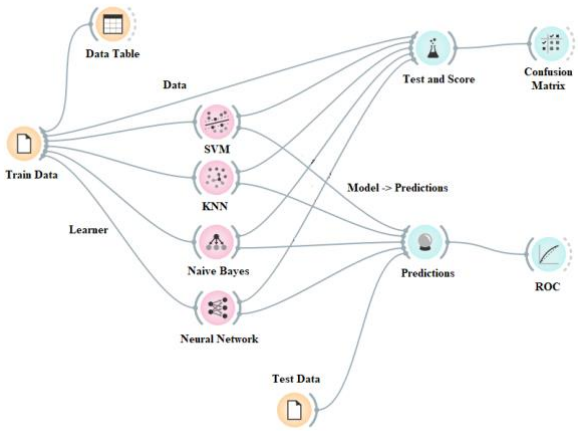



**FIGURE 4. Modelling in Orange Data Mining Tool**

**TABLE 1**
**AUC and CA of Proposed Model**

| Model | Method | AUC | CA |
|---|---|---|---|
| This Study | SVM | 1.000 | 0.990 |
| This Study | KNN | 1.000 | 1.000 |
| This Study | Naïve Bayes | 0.973 | 0.909 |
| This Study | Neural Network | 1.000 | 0.994 |
| Deepak [24] | Triad | 0.89 | 0.85 |
| Ghofrani [25] | Neuro Imaging | 0.89 | 0.83 |
| Ouwerkerk [26] | Recurrent Neural Network | 0.90 | 0.85 |
| Cheng [27] | Decision Tree | 0.99 | 0.97 |
| Odish [28] | EEG Bio Marker | 0.90 | 0.83 |

Data preprocessing and feature extraction involved applying Fast Fourier Transform followed by Infinite Impulse Response (IIR) filtering to precisely identify spectral characteristics and dominant frequency components. The extracted parameters served as critical predictive features for developing a diagnostic classification model for detecting HD-related sequence variations.

The dataset used in this analysis was sourced from the National Center for Biotechnology Information (NCBI) and contains genetic information from individuals affected by Huntington's disease as well as healthy individuals for comparison. It comprises 2,300 different data points, which include sequences of DNA from both affected and unaffected individuals. These sequences provide insights into potential genetic variations associated with Huntington's disease, making the dataset highly relevant for genetic research and disease diagnostics.

Before analysis, several preprocessing steps were undertaken to prepare the data for accurate and meaningful analysis. First, the raw DNA sequences were cleaned to remove any missing or incomplete data, ensuring that only high-quality sequences were used. Noise and irrelevant segments, such as non-coding regions that do not contribute to genetic disease markers, were filtered out to focus on the portions of the sequences most likely to reveal disease-related information. The sequences were then aligned to ensure consistency and remove any discrepancies in length or structure across the dataset. Additionally, the sequences were standardized to a uniform format, making them compatible with various computational tools used for DNA sequence analysis, such as those employed for filtering and feature extraction. By performing these preprocessing steps, the dataset was refined to ensure its quality and relevance for the analysis, allowing for reliable identification of potential genetic markers linked to Huntington's disease. Out of 2,300 total sequences, 1,840 sequences are strategically allocated for model training using the Orange Data Mining Tool as in FIGURE 4, ensuring robust algorithm development. The remaining 460 sequences were systematically reserved for independent model validation and performance assessment, facilitating an unbiased evaluation of the proposed classification approach. The simulation is carried out in Dell Computer with i7 processor and 8GB RAM.

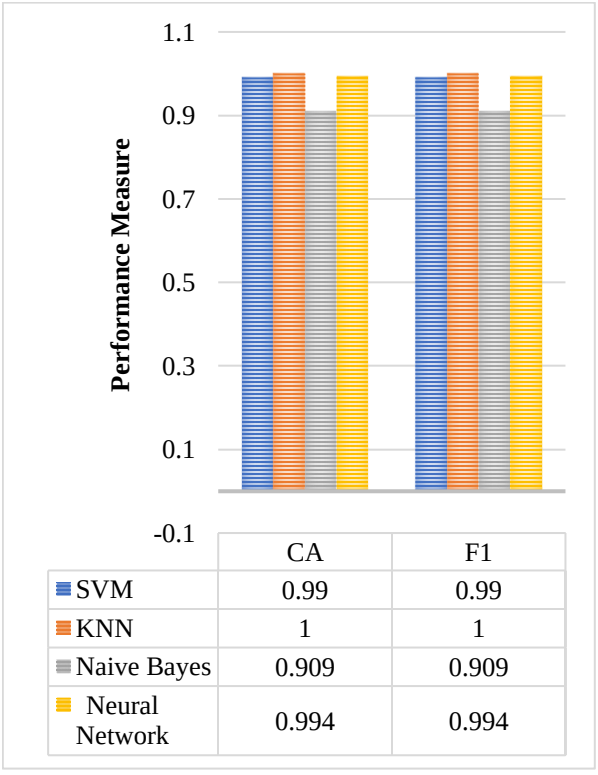
The evaluation of classification models using various metrics, including AUC, Classification Accuracy (CA), F1 Score, Precision (Prec), Recall, and Matthews Correlation Coefficient (MCC). Classification accuracy measures the overall correctness of the model by calculating the proportion of correctly classified sequences. However, since accuracy alone may not be sufficient in cases of class imbalance, additional metrics were used for a more reliable evaluation. Precision evaluates the proportion of correctly identified positive cases, which is crucial in minimizing false positives and ensuring reliable classification of disease-affected individuals. Recall, or sensitivity, assesses the model’s ability

to correctly identify actual positive cases, which is essential for detecting all instances of Huntington's disease within the dataset. F1-score, the harmonic mean of precision and recall, provides a balanced evaluation, particularly useful when false positives and false negatives carry different consequences. Matthews correlation coefficient (MCC) was incorporated as it considers all four elements of the confusion matrix (true positives, true negatives, false positives, and false negatives), making it a more robust metric, especially for imbalanced datasets. By leveraging these diverse metrics, the evaluation framework ensures a well-rounded understanding of model performance, allowing for a more accurate comparison of different classification approaches.

From TABLE 1, the K-Nearest Neighbors (KNN) classifier emerged as the top-performing model, achieving a perfect score of 1.000 across all metrics. This indicates flawless classification of the dataset, demonstrating KNN's exceptional capability to accurately differentiate between classes without any errors. Hyperparameter tuning was performed to optimize model performance.

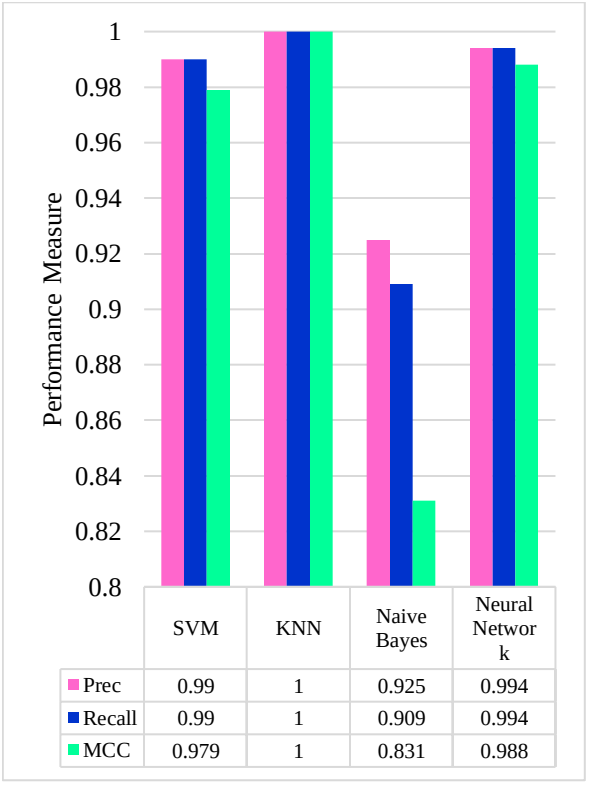
The Support Vector Machine (SVM) also delivered excellent results, achieving an AUC of 1.000 and scoring 0.990 for CA as in FIGURE 5, F1, Precision, and Recall. The MCC value of 0.979 further emphasizes its strong correlation between predictions and actual outcomes.

While slightly below KNN in terms of overall performance, SVM remains highly effective and accurate, making it a robust choice for classification tasks. Its consistency across all metrics highlights its ability to generalize well on the given dataset.




| | CA | F1 |
|---|---|---|
| SVM | 0.99 | 0.99 |
| KNN | 1 | 1 |
| Naive Bayes | 0.909 | 0.909 |
| Neural Network | 0.994 | 0.994 |

**FIGURE 5. Classification Accuracy and F1 Score of Proposed Model**




| | SVM | KNN | Naive Bayes | Neural Network |
|---|---|---|---|---|
| Prec | 0.99 | 1 | 0.925 | 0.994 |
| Recall | 0.99 | 1 | 0.909 | 0.994 |
| MCC | 0.979 | 1 | 0.831 | 0.988 |

**FIGURE 6. Comparison of Precision, Recall and MCC values of Proposed Model**

The Neural Network model achieved near-perfect results, with an AUC of 1.000 and CA, F1, Precision, and Recall scores of 0.994. Its MCC value of 0.988 as in FIGURE 6, reflects a strong predictive relationship, underscoring its reliability in handling complex datasets. Although its performance is marginally lower than KNN, the Neural Network remains a powerful model, capable of capturing intricate patterns within the data. This positions it as a competitive alternative to KNN and SVM for high-accuracy classification tasks.

In contrast, the Naïve Bayes classifier exhibited comparatively lower performance, with an AUC of 0.973 and CA, F1, Precision, and Recall scores of 0.909. Its MCC value of 0.831, while adequate, is significantly lower than the other models, indicating a moderate predictive relationship. These results suggest that Naïve Bayes, while less effective than KNN, SVM, and Neural Networks, still has merit in scenarios requiring simple implementation and computational efficiency. Overall, the analysis underscores the superiority of KNN, SVM, and Neural Networks for high-accuracy applications, with Naïve Bayes being suitable for less demanding use cases.

While the proposed methodology demonstrates promising results, certain limitations must be acknowledged. One of the potential limitation is the presence of biases in the dataset, as the data sourced from NCBI may not fully represent the genetic diversity of all individuals affected by Huntington's

disease. This could lead to model over fitting, where the trained models perform well on the given dataset but may not generalize effectively to new, unseen data. Additionally, the relatively small dataset size as 2,300 sequences may limit the robustness of deep learning or highly complex machine learning models, potentially affecting the statistical significance of the findings.

## V. CONCLUSION

This study presents an optimized Fast Fourier Transform (FFT) method for genomic data analysis, addressing challenges in sequential input processing and computational efficiency. By integrating FFT with Infinite Impulse Response (IIR) filtering, the approach improves accuracy and efficiency, particularly for large datasets. Machine learning models further enhance the methodology, enabling precise classification of Huntington's disease (HD)-related DNA sequences. K-Nearest Neighbors (KNN) achieved a perfect score of 1 across AUC, Classification Accuracy, Precision, Recall, F1 Score, and Matthews Correlation Coefficient, while Support Vector Machine (SVM) and Neural Networks also performed strongly. On a dataset of 2,300 sequences, the optimized FFT algorithm processed 90% of sequences faster than traditional methods, making it valuable for large-scale genomic analysis. The ability to quickly and accurately identify genetic markers enhances its application in medical diagnostics. Despite its effectiveness, dataset biases from NCBI and the small sample size (2,300 sequences) may impact model generalization and statistical significance. Future research should adapt this approach for other genetic disorders like Alzheimer's and cystic fibrosis, and explore its potential in real-time clinical diagnostics, advancing precision medicine.

## REFERENCES

[1] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," 1964.

[2] H. Ye et al., "An Efficient Conflict-Free 2-D FFT for FMCW Radar Imaging on DSP," in IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1-11, 2024, Art no. 8508711, doi: 10.1109/TIM.2024.3476605.

[3] Wen-Fang Yen, Shingchern D. You, Yung-Chao Chang,Real-time FFT with pre-calculation,Computers & Electrical Engineering, Volume 35, Issue 3,2009,Pages 435-440,https://doi.org/10.1016/j.compeleceng.2008.10. 002.

[4] Abhishek Chauhan, Ksh Milan Singh,Recursive sliding DFT algorithms: A review, Digital Signal Processing, Volume 127, 2022,https://doi.org/10.1016 /j.dsp.2022.103560.

[5] Mingyang Li, Wanzhong Chen, FFT-based deep feature learning method for EEG classification, Biomedical Signal Processing and Control, Volume 66,2021,https://doi.org/10.1016/j.bspc.2021.102492.

[6] Z. B. Akhtar and A. D. Gupta, "Advancements within Molecular Engineering for Regenerative Medicine and Biomedical Applications an Investigation Analysis towards A Computing Retrospective", j.electron.electromedical.eng.med.inform, vol. 6, no. 1, pp. 54-72, Jan. 2024.

[7] I. Nurcahyati, T. H. Saragih, A. Farmadi, D. Kartini, and M. Muliadi, "Classification of Lung Disease in X-Ray Images Using Gray Level Co-Occurrence Matrix Method and Convolutional Neural Network", j.electron.electromedical.eng.med.inform, vol. 6, no. 4, pp. 332-342, Aug. 2024.

[8] L. Ratié and S. Humbert, "A developmental component to Huntington's disease," May 01, 2024, Elsevier Masson s.r.l. doi: 10.1016/j.neurol.2024.04.001.

[9] A. McGarry, K. Biglan, and F. Marshall, "Huntington's disease," in Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric Disease, Elsevier, 2025, pp. 171–181. doi: 10.1016/B978-0-443-19176-3.00012-1.

[10] P. Gonzalez-Baez Ardisana, J. S. Solís-Mata, and J. D. Carrillo-Ruiz, "Neurosurgical therapy possibilities in treatment of Huntington disease: An update," Aug. 01, 2024, Elsevier Ltd. doi: 10.1016/j.parkreldis.2024.107048.

[11] E. Cubo et al., "The association between lifestyle factors and mortality in Huntington's disease," Neurologia, 2024, doi: 10.1016/j.nrl.2024.01.001.

[12] C. Gallezot et al., "Emotion expression through spoken language in Huntington disease," Cortex, vol. 155, pp. 150–161, Oct. 2022, doi: 10.1016/j.cortex.2022.05.024.

[13] H. Liu, Q. Bai, X. Wang, Y. Jin, X. Ju, and C. Lu, "Immune signature of gene expression pattern shared by autism spectrum disorder and Huntington's disease," IBRO Neurosci Rep, vol. 17, pp. 311–319, Dec. 2024, doi: 10.1016/j.ibneur.2024.09.004.

[14] M. Ghofrani-Jahromi et al., "Prognostic enrichment for early-stage Huntington's disease: An explainable machine learning approach for clinical trial," Neuroimage Clin, vol. 43, Jan. 2024, doi: 10.1016/j.nicl.2024.103650.

[15] M. Henry, "An ultra-precise Fast Fourier Transform," Measurement: Sensors, vol. 32, Apr. 2024, doi: 10.1016/j.measen.2024.101039.

[16] F. Ai and V. Lomakin, "Fast Fourier Transform periodic interpolation method for superposition sums in a periodic unit cell," Comput Phys Commun, vol. 304, Nov. 2024, doi: 10.1016/j.cpc.2024.109291.

[17] R. R. Mantovani, G. L. Menezes, and J. R. R Dórea, "Predicting respiration rate in unrestrained dairy cows using image analysis and fast Fourier transform," 2024. [Online]. Available: http://creativecommons.org/ licenses/by/4.0/

[18] R. Priyadharsini and S. Sasipriya, "A novel hybrid fast Fourier transform processor in 5G+ and bio medical applications," Microprocess Microsyst, vol. 105, Mar. 2024, doi: 10.1016/j.micpro.2024.105022.

[19] A. Sharma et al., "Heart rate and blood pressure measurement based on photoplethysmogram signal using fast Fourier transform," Computers and Electrical Engineering, vol. 101, Jul. 2022, doi: 10.1016/j.compeleceng.2022.108057.

[20] S. Chandrasekaran and U. B. Nageswaran, "A mutated addition–subtraction unit to reduce the complexity of FFT," Applied Nanoscience (Switzerland), vol. 13, no. 4, pp. 2935–2944, 2023, doi: 10.1007/s13204-021-02278-5.

[21] N. H. Arif, M. R. Faisal, A. Farmadi, D. T. Nugrahadi, F. Abadi, and U. A. Ahmad, "An Approach to ECG-based Gender Recognition Using Random Forest Algorithm," Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 6, no. 2, pp. 107–115, Apr. 2024, doi: 10.35882/jeeemi.v6i2.363.

[22] K. Singh et al., "Emerging pharmacological approaches for Huntington's disease," Oct. 05, 2024, Elsevier B.V. doi: 10.1016/j.ejphar.2024.176873.

[23] J. S. Paulsen et al., "Detection of Huntington's disease decades before diagnosis: The Predict-HD study," J Neurol Neurosurg Psychiatry, vol. 79, no. 8, pp. 874–880, 2008, doi: 10.1136/jnnp.2007.128728.

[24] Deepak, S., Ojha, A., Acharjya, K. et al. A novel and proposed triad machine learning-based framework for the prognosis of Huntington's disease. Int. j. inf. tecnol. 16, 2603–2610 (2024). https://doi.org/10.1007/s41870-023-01719-4

[25] Ghofrani-Jahromi, M., Poudel, G. R., Razi, A., Abeyasinghe, P. M., Paulsen, J. S., Tabrizi, S. J., Saha, S., and Georgiou-Karistianis, N., "Prognostic enrichment for early-stage Huntington's disease: An explainable machine learning approach for clinical trial," NeuroImage: Clinical, vol. 43, 2024, Art. no. 103650.

[26] Ouwerkerk, J., Feleus, S., van der Zwaan, K.F. et al. Machine learning in Huntington's disease: exploring the Enroll-HD dataset for prognosis and driving capability prediction. Orphanet J Rare Dis 18, 218 (2023). https://doi.org/10.1186/s13023-023-02785-4

[27] Cheng, H. P. Liu, W. Y. Lin, and F. J. Tsai, "Identification of contributing genes of Huntington's disease by machine learning,"

BMC Med Genomics, vol. 13, no. 1, Dec. 2020, doi: 10.1186/s12920-020-00822-w.

[28] Odish, O.F.F., Johnsen, K., van Someren, P. et al. EEG may serve as a biomarker in Huntington's disease using machine learning automatic classification. Sci Rep 8, 16090 (2018). https://doi.org/10.1038/s41598-018-34269-y

[29] C. Saravanakumar and N. Usha Bhanu, "A Constructive approach to Numerical Mapping scheme of Nucleotides for Preprocessing in Machine Learning," Proceedings of the Confluence 2022 - 12th International Conference on Cloud Computing, Data Science and Engineering, pp. 482–486, 2022, doi: 10.1109/Confluence52989.2022 .9734186.

[30] A. Mugilan, A. Nivetha, P. Ponraj and S. Jemimah, "Prediction of Exons using Normalized Probability Parameters derived from Statistical Analysis of Coding Sequences," 2020 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2020, pp. 503-506, doi: 10.1109/ComPE49325.2020.9200154.

[31] Muhammad Nadim Mubaarok, Triando Hamonangan Saragih, Muliadi, Fatma Indriani, Andi Farmadi, and A. Rizal, "Comparison of the Adaboost Method and the Extreme Learning Machine Method in Predicting Heart Failure ", j.electron.electromedical.eng. med.inform, vol. 6, no. 3, pp. 253-263, Jul. 2024.

## AUTHOR'S BIOGRAPHY

**SARAVANAKUMAR C** is an accomplished Assistant Professor in the Department of Electronics and Communication Engineering at SRM Valliammai Engineering College. With a robust academic foundation, he has earned his PhD from Anna University, a Post-Graduate degree from Thanthai Periyar Government Institute of Engineering and Technology, and an Undergraduate degree from Pallavan Engineering College. His extensive academic journey reflects his dedication to the field of engineering and his commitment to continuous learning and professional development. With nearly two decades of experience in academia and research, He has established himself as a specialist in several cutting-edge areas, including Signal Processing, FPGA (Field-Programmable Gate Array) implementation, Machine Learning, and Bioinformatics. His expertise in these domains has not only enriched his teaching but also significantly contributed to the advancement of research in these fields. His research contributions are noteworthy, with a publication record that includes 18 Scopus-indexed papers and 4 Web of Science documents. These publications highlight his active engagement in high-impact research and his ability to contribute valuable insights to the scientific community. Additionally, he has successfully guided a research project funded by the Tamil Nadu State Council for Science and Technology, demonstrating his capability to lead and execute research initiatives that address real-world challenges. He is a Fellow of the Institution of Electronics and Telecommunication Engineers (IETE) and a Life Member of the Indian Society for Technical Education (ISTE). These memberships underscore his significant contributions to the field of engineering and technology and his commitment to fostering innovation and excellence in the academic and professional spheres.

**MARIRAJAN S** commitment to professional development and his contributions to the field are further evidenced by his active memberships in several prestigious organizations. He is a Life Member of the Computer Society of India (CSI), the Institution of Electronics and Telecommunication Engineers (IETE), and the Indian Society for Technical Education (ISTE). These memberships highlight his dedication to staying at the forefront of technological advancements and his involvement in fostering a collaborative environment for innovation and learning. In addition to his research and professional affiliations, he is deeply committed to his role as an educator. His teaching philosophy emphasizes practical knowledge and hands-on experience, ensuring that his students are well-prepared to meet the challenges of the rapidly evolving field of electronics and communication engineering. His passion for teaching and his ability to inspire students make him a valuable asset to SRM Valliammai Engineering College. His work continues to influence both his students and the broader engineering community, driving forward advancements in VLSI Design and Signal Processing. His dedication to his profession and his contributions to research and education underscore his role as a leading figure in the field of electronics and communication engineering.

**PANDIAN A** serves as an Assistant Professor in the Department of Electronics and Communication Engineering at SRM Valliammai Engineering College. He has a strong academic foundation, having earned his post-graduate degree from SRM Easwari Engineering College and his undergraduate degree from Paavai Engineering College. With over 14 years of academic and research experience, He has established himself as a dedicated educator and researcher in the field of electronics and communication engineering. He specializes in several interdisciplinary domains, including Image Processing, VLSI (Very Large-Scale Integration) Design, and Machine Learning. His expertise in these areas has enabled him to contribute significantly to both teaching and research, bridging the gap between theoretical knowledge and practical applications. His research work is widely recognized, with 5 publications indexed in Scopus and 1 publication in Web of Science, reflecting his commitment to advancing knowledge in his field. In addition to his academic and research contributions, he is a Life Member of the Indian Society for Technical Education (ISTE), a testament to his dedication to the advancement of technical education and research. His involvement in professional organizations underscores his commitment to staying at the forefront of technological advancements and fostering a collaborative environment for innovation.

**DURGADEVI K** has received her bachelor's degree in Electrical and Electronics Engineering from KLN College of Information Technology and obtained her Master of Technology (M.Tech) in VLSI Design from Kalasalingam University. Currently, she is working as an Assistant Professor in the Department of Electronics and Communication Engineering at SRM Valliammai Engineering College, Kattankulathur, India. She has further advanced her academic credentials by earning her PhD degree from Anna University, demonstrating her commitment to academic excellence and research. Her work in these areas reflects her dedication to exploring innovative solutions and advancing knowledge in the field. She is a Life Member of the Indian Society of Technical Education (ISTE) and the Institution of Electronics and Telecommunication Engineers (IETE), highlighting her active involvement in professional organizations and her commitment to the advancement of technical education and research. These memberships also underscore her dedication to staying updated with the latest developments in her field and contributing to the broader engineering community. She inspires excellence in technical education and research.