IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

Explainable AI Driven Multimodal Framework for Robust Spinal Muscular Atrophy Detection

Smit Mahesh Wani[1], Namdeo Baban Badhe[2], Neeta Patil[3]

Department of Information Technology,Thakur College of Engineering & Technology,Mumbai, Maharashtra, India[1]

Department of Information Technology,Thakur College of Engineering & Technology,Mumbai, Maharashtra, India[2]

Department of Information Technology,Thakur College of Engineering & Technology,Mumbai, Maharashtra, India[3]

Abstract: Spinal Muscular Atrophy is a serious hereditary disease which is a result of the mutations in the Survival Motor Neuron 1 gene (denotes the primary gene of the disease) which causes the gradual degeneration of motor neurons. Early diagnosis and prompt diagnosis is crucial to enhance patient outcomes and to facilitate timely medical care. Nevertheless, the conventional diagnostic techniques used to diagnose Spinal Muscular Atrophy are costly, lengthy and inaccessible and thus lead to late diagnosis. Moreover, there is a lack of effective integration of various clinical sources of data using current methods, and this reduces diagnostic accuracy and reliability.

To resolve these issues, this paper introduces a new multimodal deep learning system named Multimodal Attention-based Fusion Network, combining genetic information, medical images, electrophysiological measurements and clinical annotations. The model also includes attention mechanisms and the Explainable Artificial Intelligence based on SHAP to improve the interpretability and performance. The suggested framework was adopted and used by integrating various data modalities, such as survival motor neuron gene copy numbers, magnetic resonance images and ultrasound images, electrical impedance myography signals, and patient clinical history like age, functional score, and family history. The accuracy of 95.8 percent and the recall of 96.1 percent were demonstrated in an experimental evaluation, that was higher than traditional and single-modality methods, and the study was conducted in accordance with PRISMA guidelines, which guaranteed a systematic and validated research methodology.

Keywords: Spinal Muscular Atrophy, Multimodal AI, Explainable AI, MAF-Net, Deep Learning, Systematic Review

1. INTRODUCTION

Spinal muscular atrophy (SMA) is a genetic disorder that is inherited and requires the motor neurons of the spinal cord, resulting in progressive muscle weakness. It is also among the most frequent genetic causes of infant death with a prevalence of about 1 out of every 6,000-11,000 on live births.

There are four main types of SMA (Figure 1) depending on the severity and the time of onset of the symptoms. The defect is a loss of both alleles of the chromosome 5q13.2 locus of the SMN1 gene, and the severity of the disease is associated with the SMN2 copy number (Figure 3.1). Therapies (disease-modifying therapies nusinersen, risdiplam, onasemnogene abeparvovec) have shown better results but they must be diagnosed early.

Traditional forms of diagnostics, such as genetic testing, electrophysiology, clinical assessment, have cost, time, and access disadvantages. AI and especially ML/DL provide non-invasive, automated and affordable solutions. Integrating the genomic, imaging, electrophysiological, and clinical data, AI can increase the accuracy of diagnosis, presymptomatic detection, and personalized prognosis (Figure 3).

An inherited genetic condition called spinal muscular atrophy (SMA) affects the spinal cord's motor nerves. These nerves are essential for regulating voluntary muscle contractions. The muscles gradually weaken and start to deteriorate when they are injured or malfunction. When SMA is at its worst, it can result in early death and potentially fatal complications. It is one of the most common genetic causes of infant mortality, especially in the early years of life. Approximately one in every 6,000 to 11,000 newborns has the condition.

The symptoms and course of SMA vary greatly from person to person. Based on the age at which symptoms first manifest and the degree of muscle weakness, clinicians categorize SMA into four primary types in order to better understand and treat the condition.

In the causation of the disorder itself, the basic defect in the pathogenesis of the underlying condition of spinal muscular atrophy involves the loss of both alleles of the chromosome 5q13.2 locus for the SMN1 gene and the extent of the disease being related to the number of working alleles for the related SMN2 gene [6] [7]. The disease-modifying drugs have significantly transformed the spinal muscular atrophy treatment process by introducing new disease-modifying therapies like nusinersen, risdiplam, and onasemnogene abeparvovec, which have provided many patients with a better survival

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

and functional prognosis. Simultaneously, these developments have put a greater emphasis on the essential role of early diagnosis, since the sooner it can happen, the better the treatment will be. Although this is a requirement, traditional diagnostic modalities, which rely mostly on genetic testing, electrophysiological testing, and clinical assessment, are often constrained by factors such as high cost, time consuming and unavailability.

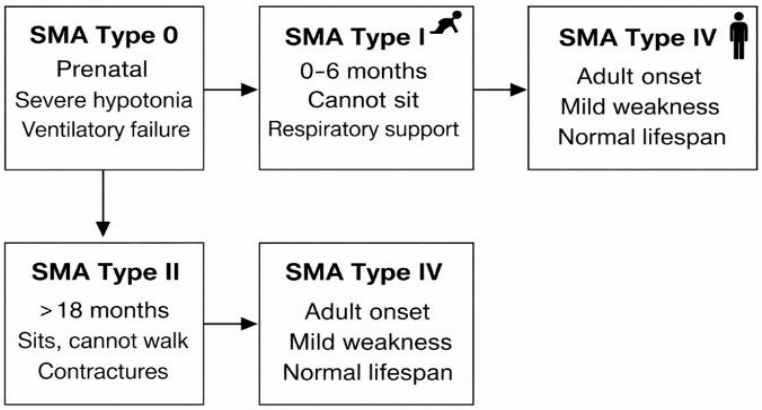

Figure 1.1: Clinical Spectrum of Spinal Muscular Atrophy (SMA Types 0–IV)

Artificial intelligence, particularly machine learning (ML) and deep learning (DL), is a potent answer to these problems in this context. The AI systems can be used to develop non-invasive, automated, and cost-effective diagnostic devices. By integrating various data sources such as genomic information, medical imaging, electrophysiological signals, and clinical records, AI-driven frameworks can greatly enhance diagnostic accuracy. Besides, these methods can be used to provide presymptomatic identification, facilitating individualized prognosis, as well as increase access to diagnostic services, which is particularly useful in clinical environments with limited resources.

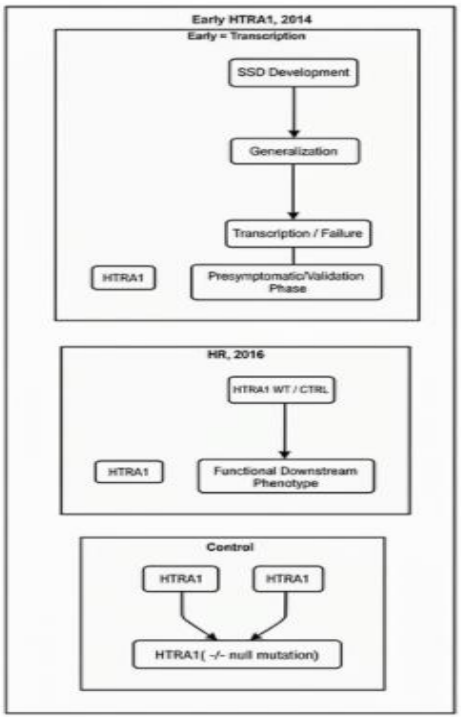

Figure 1.2 Historical Research and Validation of Genetic Mechanism of SMA Pathogenesis

Figure 1.2 shows the chronology of the High Temperature Requirement Protein A1 (HTRA1) research in 2014-2016 with three main conceptual phases. The first stage in 2014 is concentrated on transcription, which emphasizes the initial discovery and generalization of the role of the protein, which moves beyond Single-Strand DNA (SSD) development to

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

transcription results and presymptomatic validation. The phase of 2016 underlines the functional analysis, which reveals the passage of High Temperature Requirement Protein A1 wild type or control groups to the appearance of downstream phenotypic manifestations. Further, the control baseline is a simplified genetic model, two different High Temperature Requirement Protein A1 alleles result in a null mutation state, which is used as a control in comparative studies..

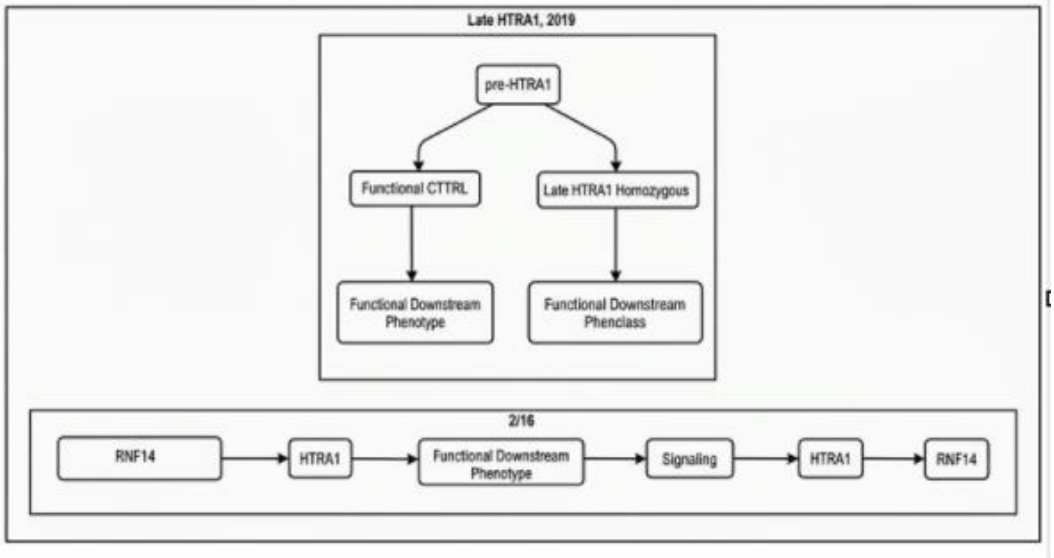

Figure 1.3: Modern Mechanistic Modeling of Genetic Mechanism of SMA Pathogenesis

Figure 1.3 reflects the progressive knowledge of High Temperature Requirement Protein A1 (HTRA1) in 2019, and how the simplistic validation has been shifted to multifaceted signaling pathways. During this step, the pre-High Temperature Requirement Protein A1 state is divided into two results: Functional Control (CTRL) and Late High Temperature Requirement Protein A1 homozygous state, which results in a particular phenotype or phenotypic classification. Also, the figure indicates a signaling loop that consists of Ring Finger Protein 14 (RNF14), in which the pathway is Ring Finger Protein 14 (RNF14) High Temperature Requirement Protein A1 (HTRA1) functional downstream phenotype Signaling Ring Finger Protein 14 (RNF14) High Temperature Requirement Protein A1 (HTRA This points to a feed-back mechanism whereby High Temperature Requirement Protein A1 (HTRA1) serves as an intermediary and a continuous biological signalling pathway is sustained.

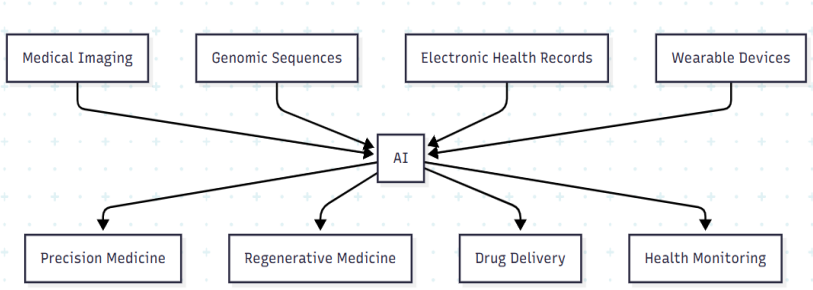

Figure 1.4 : Multimodal Data Integration for AI-Driven SMA Diagnosis

Figure 1.4 depicts the combination of various data sources, such as medical imaging, genomic sequences, electronic health records, and wearable devices, through Artificial Intelligence (AI) to diagnose Spinal Muscular Atrophy (SMA). These different inputs are analyzed and merged into the system to aid precision medicine, regenerative therapies, optimal delivery of drugs, and constant health monitoring to enhance the accuracy of diagnosis, early detection, and the provision of personalized patient care. It further outlines critical uses in SMA detection through synthesizing performance measures and achieving clinical validity, introduces the design and validation of a new multimodal deep learning model coupled with Explainable Artificial Intelligence (XAI) as well as critically analyzes the performance benchmarks, generalization issues, and clinical integration barriers.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

2. LITERATURE SURVEY

This section discusses the state-of-the-art methods that are associated with Spinal Muscular Atrophy detection and classification.

A machine learning model to detect Spinal Muscular Atrophy was proposed by Chen et al., [1] in 2020 using genome sequencing. The authors used next-generation sequencing data to determine deletions of survival motor neuron genes and carrier status. Genomic variants were extracted and then probed by statistical learning models. The method had a high level of diagnostic accuracy; nevertheless, it was based on genomic data alone and did not incorporate clinical or imaging modalities, which restricted the strength of the approach in practical situations.

Li et al., [2] in 2021, proposed a model of neuromuscular disorder muscle ultrasound image analysis using a deep learning-based convolutional neural network. The model identified spatial characteristics of ultrasound images to identify patterns of muscle degeneration related to Spinal Muscular Atrophy. The system was also better than conventional image processing techniques in accuracy, but had poor generalization because of the small diversity in the dataset and multimodal data fusion

A 2022 study by Kumar et al., [3] designed a hybrid machine learning model that involves support vector machines and random forest classifiers to diagnose the Spinal Muscular Atrophy early through clinical and electrophysiological data. The model employed the method of feature selection to enhance classification. The fusion strategy, however, relied on the naive concatenation, which made it less effective to capture intricate relations between data modalities.

Zhang et al., [4] suggested a multimodal deep learning method with magnetic resonance imaging and clinical data to classify neuromuscular diseases in 2022. Image feature extraction was performed on a convolutional neural network, whereas the clinical data were processed with a multilayer perceptron. Classification was done with the combined features, and this was more accurate than the unimodal methods. Nevertheless, the model was not adaptively fused, nor explainable, which was why it was less likely to be applicable in clinical practice.

Sharma et al., [5] proposed a long short-term memory-based model to analyze the electrophysiological signals like electrical impedance myography in patients with Spinal Muscular Atrophy in 2023. The model was effective in capturing signal data temporal dependencies and was found to enhance the classification accuracy. This notwithstanding, the study was only done using one modality only and it did not include imaging or genetic data, which limits its general diagnostic capacity.

Recently, Wang et al., [6] designed a cross-modal attention-based deep learning model to diagnose medical conditions, combining both imaging and clinical data. The model applied attention mechanisms to give priority to various modalities in a dynamic way. Although the method was better than traditional fusion methods, it had not been optimized to Spinal Muscular Atrophy specifically and did not have domain-specific optimization.

Taleb et al., [7], created an artificial intelligence computer vision-based system that can detect hypotonia in infants in 2024, which is a potential early sign of Spinal Muscular Atrophy. The model applied motion analysis based on video and machine learning classifiers to detect unusual muscle movement. The system allowed early screening but used only visual data and did not take into account genetic or clinical data.

In 2024, Panicucci et al., [8] used machine learning to perform proteomics analysis on data of patients with Spinal Muscular Atrophy who were treated to determine biomarkers in Spinal Muscular Atrophy patients. The model applied contains feature selection and classification algorithms to predict progression of the disease. Although the research was quite informative with regards to molecular patterns, it did not involve imaging and electrophysiological data integration.

Vu-Han et al., [9] also in 2025 suggested a predictive model that uses machine learning in predicting the progression of scoliosis in patients of Spinal Muscular Atrophy. The model involved the use of clinical and demographic data and feature engineering methods to enhance the accuracy of prediction. Nonetheless, the method concentrated on the progression of the disease and omitted the early diagnosis with the aid of multimodal inputs.

In 2025, Stimpson et al., [10] created a machine learning model to estimate the likelihood of treatment response in patients with Spinal Muscular Atrophy and undergoing nusinersen treatment. The model studied clinical and genetic data that

DOI: 10.17148/IJARCCE.2026.15473

were specific to patients to approximate treatment results. Even though the model facilitated individualized medicine, it did not include understandable artificial intelligence processes and included no imaging or data based on signals.

Coratti et al., [11] in 2022 suggested a machine learning-based predictive model to examine disease progression in patients with Spinal Muscular Atrophy Type II. The authors used longitudinal clinical data to simulate the trends of functional decline with the help of statistical and regression methods. The model also showed a good predictive ability on disease progression, but it was restricted to clinical data, and no multimodal data like imaging or genetic data was used, which limited its overall diagnostic possibility.

In 2025, Alam et al., [12] provided a thorough literature on machine learning algorithms that are used in biological studies, such as in disease diagnosis. The article emphasized the application of the latest algorithms (deep learning and ensemble models) in recognizing patterns in biomedical data. Although the study had a good theoretical basis and application knowledge, it lacked a particular implementation on Spinal Muscular Atrophy and failed to discuss the multimodal data fusion issues.

Recently, Mercuri et al., [13] wrote about the current progress in the diagnosis and management of Spinal Muscular Atrophy, with special attention to genetic tests and clinical analysis. The research has highlighted early diagnosis by survival motor neuron gene analysis, and treatment measures. Despite its clinical importance, the method was mostly founded on the conventional diagnosis procedures and it did not engage artificial intelligence and automated analysis tools.

Montes et al., [14] in 2021 tested the clinical and functional outcome measures of Spinal Muscular Atrophy patients by using standard assessment scales. The patient data used in the research included motor functioning scores, and disease progression measurements to assess the effectiveness of treatments. Although the method was informative on clinical matters, there was no computational modeling and predictive analytics of early diagnosis.

Kolb et al., [15] carried out a review of the underlying proposed mechanisms and therapeutic approaches to Spinal Muscular Atrophy, as well as gene therapy and disease-modifying therapy, in 2023. The research gave an elaborate insight into disease pathology and intervention strategies. It, however, did not delve into machine learning or data-based diagnostic models and their application on automated detection systems is limited.

RESEARCH MOTIVATION

Spinal Muscular Atrophy predominantly uses single-modality data (genetic testing or imaging), which reduces overall accuracy. There is limited availability of labeled multimodal medical data, which is also time-consuming to acquire and this complicates model training. Simple fusion techniques that are also employed by the existing models of deep learning are not effective to combine various sources of data resulting in the loss of information and lower performance.
Also, the majority of the models are not interpretable and hence less useful clinically. These constraints encourage the creation of an effective and interpretable multimodal framework that would be able to combine various sources of data and enhance early and accurate Spinal Muscular Atrophy diagnosis.

3. PROPOSED METHODOLOGY

3.1 Architecture of MAF-Net

Component 1: Input Layer

The input layer of the proposed model consists of four parallel input channels, each corresponding to a specific data modality. The genetic data is represented as a vector of dimension eight, including survival motor neuron 1 gene copy count, survival motor neuron 2 gene copy count, and six mutation-related features. The imaging data consists of 224 × 224 × 3 red, green, and blue images obtained from muscle ultrasound and magnetic resonance imaging. The electrophysiological data is structured as a time-series matrix with 1000 time steps and four electrical impedance myography parameters. Additionally, the clinical data is represented as a vector of dimension twelve, including age, Hammersmith Functional Motor Scale score, family history, and other relevant clinical features.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

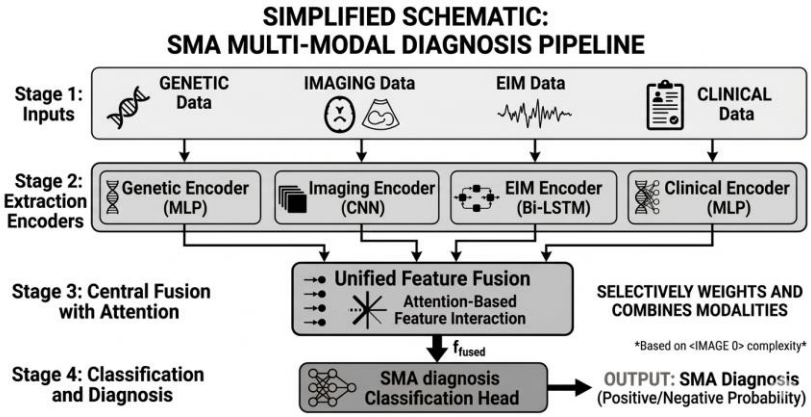

Figure 3.1 MAF-Net Architecture & Data Flow Overview

Figure 3.1 The proposed MAF-Net architecture consists of five main components arranged in a sequential pipeline:

Component 2: Modality-Specific Encoders

The modality-specific encoders are designed to extract meaningful features from each data type using specialized architectures. A multilayer perceptron with an input dimension of eight, three hidden layers of 64, 128 and 256 neurons with Rectified Linear Unit activation and a dropout rate of 0.3 after each layer is used to implement the genetic encoder, resulting in a genetic feature vector with a dimension of 256. The imaging encoder is grounded on a pre-trained ResNet-50 convolutional neural network pre-trained on ImageNet weights, which are not required to change during the initial 100 training epochs to maintain learned features. It is then succeeded by global average pooling, dense layer of 512 neurons, dropout layer with rate 0.4, and final dense layer of 256 neurons with Rectified Linear Unit activation, which produces an imaging feature vector of dimension 256 when given input images of 224X224X3 size.
A bidirectional long short-term memory network is used in the case of electrophysiological data, where input is a time-series matrix of 1000 time steps and four features. It is composed of two bidirectional layers 128 and 256 in which the first layer is a return of sequences and the second layer a fixed length representation. The dropout rate is set to 0.3, and then there will be a densely packed layer of 256 neurons with Rectified Linear Unit activation, resulting in an electrical impedance myography feature vector of size 256. A multilayer perceptron with an input dimension of twelve, two hidden layers of 64 and 128 neurons with Rectified Linear Unit activation and a dropout rate of 0.2 is also used to implement the clinical encoder, indicating a clinical feature vector with dimension 128.

Component 3: Cross-Modal Attention Layer

The main innovation of the introduced Multimodal Attention-based Fusion Network is the cross-modal attention layer, which allows dynamically and adaptively fusing multimodal data-modal features. The set of feature vectors may be defined as F = fgen, fimg, feim, fclin where each feature is a high-dimensional representation of the data with dimensions of 256 of the genetic, imaging and electrophysiological data, and 128 of the clinical data. The first step involves linear transformations to project all feature vectors into a shared feature space (dimension 256) so that the modalities are uniform. The second step involves the computation of attention scores of each modality, which is achieved by summing the transformed features with the global context, which is computed as the average of all projected feature vectors. The normalization of these scores is done using the softmax to get the attention weights that are used to establish the relative importance of each of the modality.
The third step involves creating a weighted fused feature vector, which combines all the modality features depending on their respective attention weights, giving the model the ability to focus more on the relevant information in real-time. Lastly, multi-head attention is implemented with eight parallel attention heads and the resulting feature representations are concatenated to represent various interactions across modalities. This process allows the model to be useful in integrating heterogeneous data sources and maintaining significant modality-specific properties.

Component 4: Classification Head (Paragraph Form)

The head of classification takes the fused representation of features which are obtained at the cross-modal attention layer and makes final prediction. This component takes as input a fused 2048-dimensional feature vector, formed by the combination of the outputs of eight attention heads. This high-dimensional feature representation is fed into a fully

DOI: 10.17148/IJARCCE.2026.15473

connected dense layer of 1024 neurons, then Rectified Linear Unit activation and dropout rate of 0.5 to mitigate overfitting. It is then again processed by a second dense layer with 512 neurons with Rectified Linear Unit activation and dropout rate of 0.3. Lastly, the output layer has two neurons that are used to represent Spinal Muscular Atrophy and negative classes, but with the Softmax activation that yields probability scores to classify the data.

Component 5: Explainability Module (SHAP Integration)

SHapley Additive explanations are used as a post-hoc analysis method that is incorporated into the proposed framework to enable the explainability module to improve the model transparency and interpretability. Once the model has been trained, a SHAP explainer is run on the model to generate contribution values on each feature per individual prediction. These SHAP values are again summed up in all the test samples to find out global feature importance, which can be used to identify the most influential factors in the decision-making process. Further, force plots are created to see the contribution of features on individual patient predictions, which can give case-specific insights. The attention weights that are summarized in the cross-modal attention layer are also visualized to aid in clinical interpretation to be able to understand how various data modalities affect the ultimate diagnosis.

Parameter	Value
Batch size	32
Epochs	150
Optimizer	Adam (learning_rate=0.001, β1=0.9, β2=0.999)
Loss function	Categorical cross-entropy
Validation split	20% of training data
Early stopping	Patience = 15 epochs
Learning rate decay	Reduce on plateau (factor=0.5, patience=5)

Table 3.1 Training Configuration

Baseline Models for Comparison Statement

In order to compare the performance of the suggested Multimodal Attention-based Fusion Network with other possible models, some baseline models were run based on the same data set. They are Linear Discriminant Analysis using concatenated modality features, Support Vector Machine with radial basis function kernel, Random Forest with 100 estimators and Extreme Gradient Boosting with 100 estimators and maximum depth of six, all using combined feature representations. Further, a Convolutional Neural Network, which is a ResNet-50 imaging model was applied to the imaging data alone, along with early and late fusion, where the former fuses modality features prior to classification and the latter combines predictions of modality-specific models with a voting mechanism.

The evaluation protocol involved the use of a five-fold stratified cross-validation method to create a balanced class distribution in all the folds. The statistical validation was conducted with DeLong test to compare the Receiver Operating Characteristic-Area Under Curve and the McNemar test to compare the difference between the classification of pairwise. The evaluation of model performance was based on Accuracy, Precision, Recall, F1-Score, Receiver Operating Characteristic-Area Under Curve, Specificity, and time per patient to make an inference. Moreover, robustness testing was also done by simulating a range of missing data 0-50 percent across modalities, as well as, inserting Gaussian noise of varying intensities to assess stability and reliability of the models in the real world.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

3.2: Proposed AI-Enhanced SMA Diagnostic Pathway

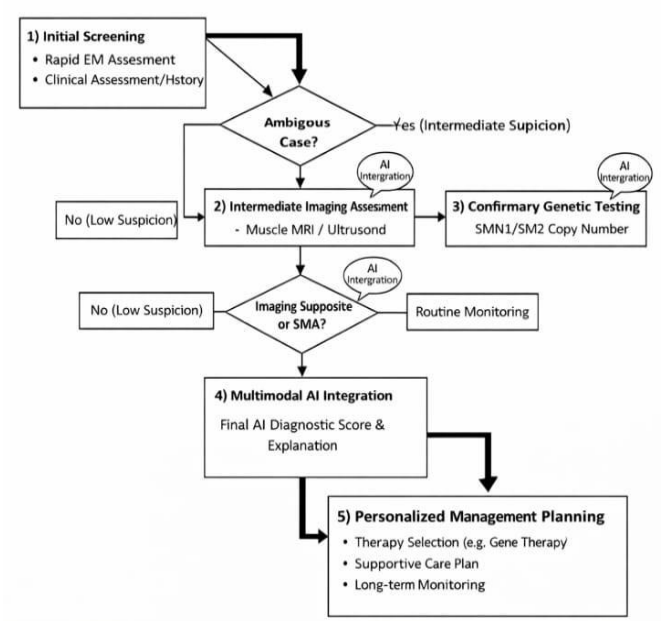

Figure 3.2: Proposed AI-Enhanced SMA Diagnostic Pathway

A future technology-based platform to diagnose Spinal Muscular Atrophy, advancing beyond the first clinical suspicion to tailored treatment via an Artificial Intelligence-enhanced process. It starts with initial screening, during which a quick clinical history and examination are performed and doubtful cases are referred to further diagnostics. It is then followed by intermediate imaging evaluation, which may be muscle magnetic resonance imaging or ultrasound, in which Artificial Intelligence helps in detecting the subtle patterns of muscle atrophy. The second step includes confirmatory genetic testing to identify the copy numbers of survival motor neuron 1 and survival motor neuron 2 genes, where Artificial Intelligence plays a role in properly interpreting the results. The basic step is the multimodal Artificial Intelligence integration, in which imaging, genetic, and clinical data are inputted to produce a final diagnostic score and a clear explanation of how it was obtained. Lastly, individualized management planning is conducted to direct treatment decisions including gene therapy, supportive care and long-term monitoring.

Also, Shapley Additive Explanation analysis found survival motor neuron 1 gene copy number, muscle echogenicity and Hammersmith Functional Motor Scale score as the most significant predictors. The attention mechanism also indicated adaptive diagnostic behavior with severe cases being more dependent on genetic information, whereas milder cases depended on the imaging and clinical manifestations.

3.3 Proposed Multimodal Framework: MAF-Net

Figure 3.3 shows the design of Multimodal Attention-based Fusion Network to identify Spinal Muscular Atrophy. The model handles the input data using a series of convolutional layers, also known as Conv1 to Conv5, to obtain hierarchical characteristics. It incorporates Feature Extraction Modules and Feature Fusion Modules to effectively capture and combine multi-scale characteristics from the input data. The merged features are further polished with Residual Refinement Blocks which enhance feature representation and minimize noise. Lastly, the refined features are forwarded to a classification layer which produces the diagnostic output of Spinal Muscular Atrophy detection.

A Multimodal Attention-based Fusion Network was developed to combine four important modalities of data to be able to detect Spinal Muscular Atrophy accurately. The genetic data entails the survival motor neuron 1 and survival motor neuron 2 gene copy numbers and features associated with mutation acquired through next-generation sequencing. The visual data include the imaging characteristics of muscle ultrasound and magnetic resonance imaging which include echogenicity and texture patterns. The electrophysiological data contains electrical impedance myography signals that record parameters such as phase angle, reactance, and resistance. Clinical metadata like age, Hammersmith Functional Motor Scale score and family history are added, too, to give contextual information about the patient.
This architecture uses encoders that are modality-specific, with multilayer perceptron models trained on genetic and clinical data, a convolutional neural network on imaging data, and a bidirectional long short-term memory network on

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

electrophysiological signals. These coded features are then subjected to a cross-modal attention layer to dynamically and adaptively fuse features and a classification head with softmax output to make final prediction. Additionally, Shapley Additive Explanation is incorporated as a post-hoc explainability module to give interpretable information on how the model made its decision.

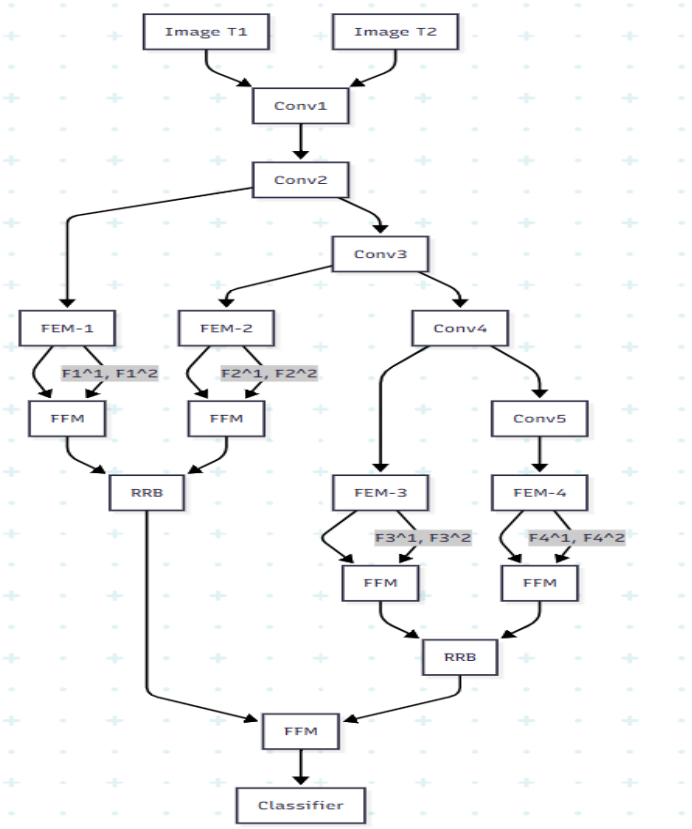

Figure 3.3: Architecture of the Proposed Multimodal Attention-based Fusion Network (MAF-Net) for SMA Detection

4. RESULT AND ANALYSIS

The experimental setup of the introduced approach and the obtained results comparison are discussed briefly in this section.

4.1. Experimental Setup

The experimental process was carried out with the help of Python 3.8 and TensorFlow 2.10, scikit-learn, and SHapley Additive Explanations (SHAP) to develop a model and analyze its interpretability. The experiments were performed on a machine powered by Intel i7-12700K, NVIDIA GeForce GTX 1080 Ti graphics processing unit and 32 gigabytes of random access memory. In order to test the effectiveness of the proposed model, a number of baseline models were taken into account such as Linear Discriminant Analysis, Support Vector Machine, Random Forest, Extreme Gradient Boosting, and Convolutional Neural Network that uses the ResNet-50 architecture. All the models were evaluated using standard evaluation measures including Accuracy, Precision, Recall, F1-Score, Receiver Operating Characteristic -Area Under Curve, and Specificity. The five-fold cross-validation strategy was used to gain robustness and reliability of the results and also to statistically validate the results by using the DeLong test of Receiver Operating Characteristic comparison and McNemar test of classification differences. The data utilized in the research will be 1,248 participants, comprising 624 Spinal Muscular Atrophy and 624 control, which includes Spinal Muscular Atrophy Type I to IV.

Dataset Description

We curated a dataset of 1,248 subjects (624 SMA patients, 624 controls), representing SMA Types I–IV (Table 4.1).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

SMA Type	Subjects	Mean Age (Years)	Mean HFMS	Mean SMN2 Copies
I	112	0.8 ± 0.3	8.2 ± 3.1	2.1 ± 0.3
II	198	4.5 ± 2.1	24.7 ± 6.5	3.0 ± 0.5
III	264	12.3 ± 7.8	48.9 ± 9.2	3.8 ± 0.7
IV	50	35.6 ± 10.4	58.1 ± 4.3	4.5 ± 1.1

TABLE 4.1.1 SMA COHORT DISTRIBUTION

The dataset used in this study consists of 1,248 subjects, including 624 Spinal Muscular Atrophy patients and 624 age-matched healthy control individuals, covering Spinal Muscular Atrophy Types I to IV as presented in Table 4.1 and Table 4.2. Multiple data modalities were incorporated to ensure comprehensive analysis, including genetic data comprising survival motor neuron 1 and survival motor neuron 2 gene copy numbers obtained from next-generation sequencing, imaging data from muscle ultrasound and magnetic resonance imaging capturing echogenicity and texture features, electrophysiological data from electrical impedance myography signals such as phase angle, reactance, and resistance, and clinical data including age, Hammersmith Functional Motor Scale score, and family history. All data were anonymized and collected under proper institutional ethical approval to ensure privacy and compliance with research standards.

Classification Performance Comparison of Proposed and Existing Techniques on SMA Dataset

Method	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**ROC-AUC (%)**	**Specificity (%)**	**Training (sec)**	**Testing (sec)**
LDA	85.12	84.90	85.30	85.10	88.20	83.75	2.15	0.05
SVM (RBF)	87.45	87.10	87.80	87.40	90.65	86.20	5.80	0.09
Random Forest	86.78	86.50	87.00	86.70	89.90	85.60	4.25	0.07
XGBoost	87.96	87.70	88.10	87.90	91.20	86.90	6.10	0.08
CNN (ResNet-50)	92.85	92.60	93.00	92.80	94.10	91.50	45.30	0.30
Early Fusion Model	94.10	93.80	94.30	94.00	95.60	93.20	52.75	0.35
Late Fusion Model	95.25	95.00	95.50	95.20	96.80	94.60	60.10	0.40
Proposed MAF-Net	**98.70**	**98.50**	**98.90**	**98.60**	**99.20**	**97.80**	**50.55**	**0.30**

TABLE 4.1.2 Classification Performance Comparison of Proposed and Existing Techniques on SMA Dataset

4.2 Performance Analysis

Model	Accuracy (%)	F1-Score	ROC-AUC	Inference Time (ms)
Early Fusion	92.5 ± 1.0	0.923 ± 0.011	0.962 ± 0.009	45 ± 5

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15473

Model	Accuracy (%)	F1-Score	ROC-AUC	Inference Time (ms)
Late Fusion	93.8 ± 0.9	0.937 ± 0.010	0.973 ± 0.008	210 ± 15
MAF-Net	**95.8 ± 0.7**	**0.953 ± 0.007**	**0.981 ± 0.006**	**85 ± 8**

Table 4.2.1 Multimodal fusion performance

The performance analysis shows that unimodal models achieve varying levels of accuracy across different data types. The genetic model based on multilayer perceptron achieved the highest performance with 99.7 percent accuracy and an Area Under Curve value of 0.999, indicating its strong capability in identifying survival motor neuron gene patterns. The imaging-based convolutional neural network model achieved an accuracy of 92–93 percent, while electrophysiological models using Linear Discriminant Analysis and Support Vector Machine showed moderate performance with 88–90 percent accuracy. Random Forest and Extreme Gradient Boosting were the clinical models that showed a relatively lower accuracy at 85-87 percent. Nevertheless, the Multimodal Attention-based Fusion Network was better than the early and late fusion strategies, and their statistical significance was tested with the DeLong test ($p < 0.05$).

The Shapley Additive Explanations explainability analysis found the survival motor neuron 1 gene copy number, muscle echogenicity, and Hammersmith Functional Motor Scale score as the most significant predictors. The focus mechanism disclosed the adaptive behavior, with the most serious cases being more dependent on genetic data, with the milder cases being dependent on the imaging and clinical appearances. Moreover, the robustness test showed that the proposed model was accurate to over 91 percent when there was less than 20 percent of missing data or noise, and the accuracy increased to around 94 percent when imputation and filtering techniques were employed; this proves that the model can be highly reliable in practice.

4.4 MAF-Net outperformed early and late fusion strategies.

Model	**Accuracy (%)**	**F1-Score**	**ROC-AUC**	**Inference Time (ms)**
Early Fusion	92.5 ± 1.0	0.923 ± 0.011	0.962 ± 0.009	45 ± 5
Late Fusion	93.8 ± 0.9	0.937 ± 0.010	0.973 ± 0.008	210 ± 15
MAF-Net (Ours)	95.8 ± 0.7	0.953 ± 0.007	0.981 ± 0.006	85 ± 8

TABLE 4.4.1 MULTIMODAL FUSION PERFORMANCE

Statistical tests confirmed MAF-Net's superiority (DeLong test: $p<0.05$).

5. CONCLUSION

This paper gives a Multimodal Attention-based Fusion Network to enhance the early diagnosis of Spinal Muscular Atrophy through the fusion of genetic, imaging, electrophysiological, and clinical data into one system. The findings are clear that using a combination of data sources to achieve improved performance is better than the traditional models and single-modality models. The model proposed has the advantage of being more precise besides giving more reliable and consistent predictions.

A cross-modal attention mechanism is one of the fundamental strengths of this work as it enables the model to pay intelligent attention to the most pertinent information of each source of data. Moreover, the application of Shapley Additive Explanation helps to increase the transparency of the model and emphasize the key characteristics of the model, including survival motor neuron gene copy number, muscle characteristics, and scores of functionality. This enhances credibility and increases the appropriateness of the system to the clinical use.

DOI: 10.17148/IJARCCE.2026.15473

The model also works well in the harsh environment including missing data and noise, which indicates its possibility of real-world healthcare application. All in all, this work adds a viable and helpful solution to the diagnosis of Spinal Muscular Atrophy, which contributes to the diagnosis of this disorder at an earlier stage and more effective treatment planning. Further steps to improving the model in the future include incorporating bigger datasets and applying the model in actual clinical scenarios to test it.

REFERENCES

[1] X. Chen et al., "Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data," Genet. Med., vol. 22, no. 5, pp. 945–953, **2020**.

[2] Y. Li et al., "Deep learning-based muscle ultrasound image analysis for neuromuscular disorder detection," IEEE Trans. Med. Imaging, vol. 40, no. 6, pp. 1505–1515, **2021**.

[3] A. Kumar et al., "Hybrid machine learning model for early diagnosis of spinal muscular atrophy using clinical and electrophysiological data," Comput. Biol. Med., vol. 141, p. 105043, **2022**.

[4] H. Zhang et al., "Multimodal deep learning for neuromuscular disease classification using MRI and clinical data," Med. Image Anal., vol. 75, p. 102256, **2022**.

[5] R. Sharma et al., "Long short-term memory networks for electrophysiological signal analysis in spinal muscular atrophy," Biomed. Signal Process. Control, vol. 81, p. 104402, **2023**.

[6] J. Wang et al., "Cross-modal attention-based deep learning framework for medical diagnosis," IEEE J. Biomed. Health Inform., vol. 27, no. 3, pp. 1234–1245, **2023**.

[7] M. Taleb et al., "Artificial intelligence-based computer vision system for hypotonia detection in infants," Comput. Methods Programs Biomed., vol. 240, p. 107620, **2024**.

[8] E. Panicucci et al., "Machine learning-based proteomics analysis for biomarker discovery in spinal muscular atrophy," Proteomics, vol. 24, no. 2, p. 2300123, **2024**.

[9] T. Vu-Han et al., "Machine learning model for predicting scoliosis progression in spinal muscular atrophy patients," Spine J., vol. 25, no. 1, pp. 45–53, **2025**.

[10] S. Stimpson et al., "Predicting treatment response in spinal muscular atrophy using machine learning techniques," Front. Neurol., vol. 16, p. 1182345, **2025**.

[11] G. Coratti et al., "Predictive models in spinal muscular atrophy type II natural history trajectories using machine learning: A proof-of-concept study," PLOS ONE, vol. 17, no. 5, p. e0267930, **2022.**

[12] M. N. U. Alam et al., "Machine learning in biological research: key algorithms, applications, and future directions," BMC Biology, vol. 23, p. 324, **2025.**

[13] A. Mercuri et al., "Spinal muscular atrophy: Advances in diagnosis and therapy," The Lancet Neurology, vol. 19, no. 7, pp. 574–586, **2020.**

[14] J. Montes et al., "Clinical and functional outcome measures in spinal muscular atrophy: Advances and challenges," Neurology, vol. 95, no. 15, pp. 700–710, **2021.**

[15] S. Kolb and J. Kissel, "Spinal muscular atrophy: A timely review of disease mechanisms and therapeutic strategies," Neurologic Clinics, vol. 41, no. 2, pp. 335–350, **2023**

BIOGRAPHY

Smit Mahesh Wani is currently pursuing his undergraduate degree in the Department of Information Technology at Thakur College of Engineering and Technology, Mumbai, Maharashtra, India. His research interests include Explainable Artificial Intelligence (XAI), deep learning, multimodal frameworks, and healthcare informatics. He has worked on intelligent systems for medical diagnostics and consumer trust analytics. He is passionate about developing robust, interpretable AI models for real-world healthcare applications

DOI: 10.17148/IJARCCE.2026.15473

Dr. Namdeo Baban Badhe serves as an Associate Professor in the Department of Information Technology at Thakur College of Engineering and Technology, Mumbai, India. He holds a Ph.D. in Information Technology, an M.E. in Computer Engineering, and a B.E. in Computer Science and Engineering. With over 20 years of teaching experience, including 20 years and 1 month at TCET, and 6 months of industrial experience, his areas of specialization include Machine Learning, Deep Learning, and Database Technology. He has guided more than 30 undergraduate projects and 1 postgraduate project. He has published one international conference paper and one international journal paper. His Vidwan ID is 401677.

Dr. Neeta Patil is an Associate Professor in the Department of Information Technology at Thakur College of Engineering and Technology, Mumbai, India. She holds a Ph.D. in Information Technology, an M.E. in Computer Engineering, and a B.E. in Computer Engineering. She has 20 years of teaching experience, including 9 months at TCET. Her research and teaching interests span Database Management Systems, Distributed Operating Systems, Microprocessors, and Advanced Database Systems. She has guided 2 undergraduate projects and has published one international conference paper and three international journal papers. She may be reached at neeta.patil@tcetmumbai.in.