WILEY | Hindawi

Research Article

A New Artificial Intelligence-Based Model for Amyotrophic Lateral Sclerosis Prediction

A. Khuzaim Alzahrani ,[1] **Ahmed A. Alsheikhy** ,[2] **Tawfeeq Shawly** ,[3] **Mohammad Barr** ,[2] **and Hossam E. Ahmed** [2,4]

[1]*Department of Medical Laboratory Technology, Faculty of Medical Applied Science, Northern Border University, Arar 91431, Saudi Arabia*
[2]*Department of Electrical Engineering, College of Engineering, Northern Border University, Arar 91431, Saudi Arabia*
[3]*Department of Electrical Engineering, Faculty of Engineering at Rabigh, King Abdulaziz University, Jeddah 21589, Saudi Arabia*
[4]*Electrical Engineering Department, Faculty of Engineering, Benha University, Benha, Egypt*

Correspondence should be addressed to Ahmed A. Alsheikhy; aalsheikhy@nbu.edu.sa

Currently, amyotrophic lateral sclerosis (ALS) disease is considered fatal since it affects the central nervous system with no cure or clear treatments. This disease affects the spinal cord, more specifically, the lower motor neurons (LMNs) and the upper motor neurons (UMNs) inside the brain along with their networks. Various solutions have been developed to predict ALS. Some of these solutions were implemented using different deep-learning methods (DLMs). Nevertheless, this disease is considered a tough task and a huge challenge. This article proposes a reliable model to predict ALS disease based on a deep-learning tool (DLT). The developed DLT is designed using a UNET architecture. The proposed approach is evaluated for different performance quantities on a dataset and provides promising results. An average obtained accuracy ranged between 82% and 87% with around 86% of the *F*-score. The obtained outcomes can open the door to applying DLMs to predict and identify ALS disease.

1. Introduction

Amyotrophic lateral sclerosis disease occurs due to the gradual deficit of motor neurons either in the brain or the spinal cord [1–4]. The development of unknown genes or pathophysiological processes is considered the main cause of the disease [3, 5, 6]. ALS is a complex disorder since it affects the whole body and causes paralysis. This disease is very rare and unfortunately being diagnosed lately. Physicians rely on various syndromes to identify the disease in its early stages, such as behavioral deficits or cognitive dysfunctions [1, 7–10]. If the disease is diagnosed behind time, then it could affect the treatment plan negatively [2, 5]. The efficient ways to predict and diagnose ALS disease are to look for related biomarkers and perform robust clinical evaluations using biological data [1, 11, 12]. Physicians have found that genes play a substantial role since it is believed to be a cause [3, 13–16]. In addition, the disease can be developed or occur from composite interrelations between various factors, such as genes, age, and sex [3, 16, 17].

ALS disease affects the UMNs and LMNs networks which results in dysfunctions in the bulbar, thoracic, and cervical segmentations [1, 3, 6]. These dysfunctions cause an increasing weakness in the skeletal muscles, which are involved in limb movements [3, 18–20]. Bulbar onset, spinal onset, and cervical onset are multiple phenotypes of ALS [3, 5]. Patients who are diagnosed with ALS suffer from the loss of speaking memory. These patients are not likely to face a neurologist at the beginning of the diagnostic phase since it is hard to predict it early if no proper clinical evaluations are performed [1–4]. The clinical evaluations should spot signs of dysfunctions in the bulbar, thoracic, and cervical segmentations [2, 3].

ALS is a rare disease, which occurs globally and is common among people aged between 40 and 70 [21]. It is found that 5%–10% of positive diagnosed cases occurred due to mutations in C9orf72, SOD1, and FUS genes, while the remaining were sporadic [21, 22]. This disease affects people of all ethnicities and races [21, 22]. Numerous signs and symptoms can be associated with ALS disease, such as muscle weakness, twitching, atrophy, and cramps [22]. In addition, difficulty in speaking and swallowing, hyperreflexia, emotional, and cognitive changes, and respiratory symptoms are common signs of being affected by this disease [21]. Physicians use various clinical assessments to diagnose ALS disease and these assessments are electromyography, nerve conduction analyses, magnetic resonance imaging (MRI), blood and urine tests, lumbar puncture or spinal tap, genetic testing, and muscle biopsy [22–24].

Increasing age, genetics, environmental factors such as exposure to pesticides, herbicides, lead, and mercury, smoking tobacco, physical trauma, and medical conditions such as primary lateral sclerosis, autoimmune diseases, and frontotemporal dementia are considered risk factors for ALS [23, 24]. Healthcare providers and physicians apply different methods, such as medications like riluzole, baclofen, and tizanidine to manage symptoms, physical and occupational therapy, speech and swallowing therapy, breathing support, nutrition support, psychological and emotional support, hospice, and palliative care to treat ALS disease [21, 22, 24]. Currently, physicians and researchers face several challenges, which can affect patients directly. These challenges are complications in timely diagnosis, speedy progression, dearth of a cure and inadequate treatment alternatives, difficulty of care, multifaceted genetics, inadequate research funding, and narrow access to clinical trials and rehabilitation services [24]. Table 1 provides a piece of general medical information about ALS.

Currently, various articles have been published using artificial intelligence (AI) technologies to address ALS prediction and the stratifications of patients [2–4]. These articles have provided favorable outcomes; however, using these approaches in healthcare facilities is limited due to some unseen challenges, such as generalizing the methods to work with unseen subjects [1–5]. It is crucial to have a suitable method that can be applied or deployed on any dataset. Magnetic resonance imaging (MRI) is considered one of the technologies that are used in clinical evaluations to diagnose ALS as stated in Table 1. The biggest challenge in diagnosing ALS is the limited availability of datasets [2]. This research offers a new deep-learning approach using a developed UNET architecture to predict ALS.

1.1. Research Problem. Various methods were implemented using artificial intelligence technologies (AITs), such as in [6, 7, 9, 12, 13, 15, 16, 18] to predict ALS or its development rate. These technologies range from deploying diffusion tensor imaging (DTI) to convolutional neural networks (CNNs). The implemented methods achieved accuracy from 68% to nearly 95%. Thus, increasing accuracy is highly required.

1.2. Research Motivations and Contributions. To be consistent with the Saudi Vision 2030 and provide a reliable diagnosis tool to predict ALS are the motivations of this research. This study aims to predict ALS disease using the UNET architecture on a utilized dataset. The following points list the contributions of this research:

(1) Develop a new deep-learning approach based on the UNET model to predict ALS disease and its development rate.

(2) The developed approach is integrated with some data preprocessing tools to robust the outcomes.

(3) The implemented model is evaluated on a dataset using various characteristics.

This article is organized as follows: the related work is given in Section 2, and the suggested approach is described in Section 3. Section 4 provides a deep evaluation and its discussion. Section 5 concludes the article.

2. Literature Review

Interested researchers have developed various solutions to either identify ALS disease or estimate its progression rate. In this section, several works will be covered and discussed.

In [6], Pancotti et al. explored the advantages of using deep-learning methods to predict the ALS development rate. The authors performed the investigation on a dataset using three architectures. These architectures were a feed-forward neural network (FFNN), a convolutional neural network (CNN), and a recurrent neural network (RNN). In the first architecture, the authors used three hidden levels with a dropout regularization layer. The utilized hidden layers took their inputs from selected static and longitudinal features. In addition, a linear activation function was deployed, and the mean squared error (MSE) was evaluated as the loss function. For the second architecture, the inputs were divided into two parts, the longitudinal and a statical residual. 11×3 was the size of every input for the convolutional neural network. The last architecture was used for the longitudinal data only. The authors evaluated their models on a dataset from the Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT) database using two parameters, which were the root mean squared deviation (RMSD) and Pearson correlation coefficient (PCC). On the other hand, the proposed approach uses the same dataset on the UNET architecture to predict the disease and evaluate its progression rate. Various performance quantities are utilized for evaluation purposes. The proposed method reached an acceptable level of accuracy, which was found to range from 82% to 87%.

Faghri et al. in [7] applied supervised, semisupervised, and supervised machine-learning models on ALS patients to find the number of ALS subtypes to better understand this disease and study its heterogeneity. The authors obtained data from ALS patients in Italy between 1995 and 2015. In total, 2858 records were studied. Uniform manifold approximation and projection (UMAP) was the unsupervised model and neural network UMAP was the semisupervised

Table 1: General medical information about ALS [21–24].

Epidemiology	Rare disease. Occurs globally, common among people aged 40–70, 5–10% of the cases are familial due to mutations in C9orf72, SOD1, and FUS genes, while others are sporadic. Affects people of all ethnicities and races
Sign and symptoms	Muscle weakness, twitching, atrophy, and cramps. Difficulty in speaking and swallowing, hyperreflexia, emotional and cognitive changes, and respiratory symptoms
Diagnosis	Clinical assessments, electromyography, nerve conduction analyses, MRI, blood and urine tests, lumbar puncture or spinal tap, genetic testing, and muscle biopsy
Risk factors	Increasing age, genetics, environmental factors such as exposure to pesticides, herbicides, lead and mercury, smoking tobacco, physical trauma, medical conditions such as primary lateral sclerosis, autoimmune diseases, and frontotemporal dementia
Treatments	Medications such as riluzole, baclofen, and tizanidine to manage symptoms, physical and occupational therapy, speech and swallowing therapy, breathing support, nutrition support, psychological and emotional support, hospice, and palliative care
Challenges and gaps	Complications in timely diagnosis, speedy progression, dearth of a cure and inadequate treatment alternatives, difficulty of care, multifaceted genetics, inadequate research funding, narrow access to clinical trials and rehabilitation services

method, while an ensemble learning based on LightGBM was the supervised model. This method identified subtypes and provided useful insight into the ALS substructure, while the proposed approach in this study is able to determine whether a patient has ALS or not. Moreover, the estimation of the ALS development rate exists.

In [9], Huang et al. developed a model to predict ALS using a pattern analysis method. This model was implemented based on comorbidities and indicators of electronic medical records (EMRs). The authors analyzed these EMRs and later performed a comparison with healthy controls to find the associated comorbidities and select them. These selected associated comorbidities were used to build a machine-learning model and construct a new Weighted Jaccard Index (WJI) to develop a prediction system using two levels of comorbidities, which were single disease codes and clustered codes. The authors used WJI in four differentmachine-learning methods to predict ALS disease. These four models achieved 83.7% accuracy. In addition, other performance indicators were evaluated as well. The authors used a dataset from NHIRD in Taiwan. These data were collected between 1996 and 2013. The authors defined two groups as follows: positive and negative to represent ALS patients and healthy people. The healthy people were used to select healthy control parameters to build the prediction approach. The negative records were collected according to the matching gender and age attributes based on the selected healthy controls. Various experiments were performed to select the associated comorbidities and applied statistical analysis on these associated comorbidities to find the best healthy controls to implement the prediction model. The developed model categorized 162 ALS patients accurately. The proposed approach in this article used the UNET architecture to extract features from the utilized dataset and compute weight for every characteristic to predict the ALS disease and measure its progression rate.

The suggested method achieved a good accuracy between 82% and 87%, which is better than what was achieved by the method in [9].

Imamura et al. in [12] implemented an artificial intelligence-based approach to diagnose ALS using induced pluripotent stem cells (iPSCs). The authors used images of spinal motor neurons (SMNs) to develop the model and analyze it using a convolutional neural network (CNN). This method reached 97% of the area under the curve, which was the main performance indicator. The authors trained their model using a VGG-16 neural network. This approach nearly achieved an average of 84% accuracy, while the proposed technique in this article utilized an artificial intelligence-based method to predict ALS using the UNET structure and reached an accuracy between 82% and 87%. This range is better than what was reached in [12].

3. The Proposed Algorithm

3.1. Problem Statement. Various solutions to identify the ALS disease or predict its development rate based on artificial intelligence technologies were developed, such as in [6, 7, 9, 12, 13]. These works were either to identify the disease or predict its progression rate. None proposed both. In addition, some works provided no information about accuracy. Due to these reasons, this article proposes a model to identify ALS and predict its progression rate using an artificial intelligence solution based on the UNET structure.

3.2. Dataset. The utilized dataset in this study was obtained from the GitHub repository [25]. This dataset contains over 1,500 records of ALS patients and healthy people. These records were split into more than 30 columns. The columns refer to various information, such as the patients' IDs, gender, time of visits and diagnosis, and laboratory results. Many data were

missing, so the dataset was cleaned and preprocessed before being utilized in the proposed approach. Several tables were constructed in the training, testing, and analysis stages. Table 2 provides details about the used data in this research and the number of data that were assigned for training, validation, and testing.

3.3. The Proposed Methodology. This part provides a full explanation of the proposed approach. This approach takes its inputs from the constructed tables and performs some preprocessing operations to prepare data to be completely utilized. Figure 1 presents a block diagram of the proposed model.

The block diagram shows that the developed method consists of three main phases, which are the preprocessing phase, the neural network, i.e., UNET, and the evaluation of the implemented method by finding the performance quantities. An internal architecture of the developed UNET is shown in Figures 2 and 3, respectively. Initially, input data are segmented as shown in Figure 2 and then the segmented data are processed to extract features and categorize results to produce outputs as illustrated in Figure 3.

The Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised (ALSFRS-R) is the most common technique worldwide to evaluate ALS disease. It measures 12 daily activities based on scores from 0 to 4, where 0 refers to complete loss of being able to perform an activity and 4 represents normal ability. This scale is used in this study to predict the development rate of ALS. Since the scale ranges from 0 to 4 for each activity, then the maximum value is 48. The sum of all activities represents a score on the scale. Various characteristics are used, such as the number of visits after being diagnosed with ALS, onset type, and onset age. Eighteen features, also referred to as characteristics, are utilized in this research to categorize data and determine the progression rate of ALS. These features include the twelve measured activities and other factors in the utilized dataset. Several statistical parameters are used in this study, such as maximum, minimum, mean, variance, covariance, and standard deviation. These parameters were applied to healthy people to determine the healthy controls (parameters) in the proposed method. These parameters are compared with ALS patients in the developed method.

As shown in Figure 1, data in the dataset go through several operations in the preprocessing phase to prepare data to be used without any issues and to avoid overfitting, which could occur due to high dimensionality. The utilized data are divided into 70% for training, 10% for validation, and the rest for testing purposes to evade unfairness. During the training session, the 5-fold cross-validation technique was deployed to speed up the process, confirm the model solidity, and optimize the hyperparameters of the proposed approach. 7,500 bootstraps were applied to compute the confidence intervals. Table 3 lists the applied settings of hyperparameters in the proposed method.

After the required tables were constructed, the remaining clean and useful data were divided into two classes. One class was allocated for ALS patients and another class for healthy people. These two classes underwent data incorporation to produce complete sets of medical records.

In addition, a statistical analysis was performed on different disease codes after counting them one by one to support developing the proposed method. A threshold to represent the minimum number of ALS patients was set. During the segmentation stage, as shown in Figure 2, the proposed model measures a weight for each characteristic and feeds these weights to the categorization stage to predict ALS and compute its progression rate. Figure 4 illustrates a distribution of the ALSFRS-R score within a year of the training set. This distribution represents a slope versus the counted score. This slope is utilized to evaluate the progression rate of ALS.

Features with high weights get higher attention and are inserted into a group called importance characteristics. This group is used in the validation and testing sets to predict the disease. Table 4 shows a sample of the obtained weights for 5 records. The first column refers to patients' IDs and the second column to the calculated weights.

Various performance indicators were utilized to evaluate the developed approach. These performance indicators were accuracy, precision, sensitivity, *F*-score, cross-entropy loss (CEL), Dice, and Jaccard. In addition, four performance metrics were required to compute the previous performance indicators. True positive (TP), true negative (TN), false positive (FP), and false negative (FN) were the required metrics. The following equations show how the performance indicators are determined in the proposed model.

$$\text{CEL} = -\sum_{i=1}^{N} q\, x \log(P). \tag{1}$$

N refers to the number of classes being evaluated in this study, $N = 2$. q represents a binary indicator, which is computed in the proposed system, and P is the probability. This quantity provides a clear sight of how far the proposed model is from the needed results. Hence, the lower the value, the better results are achieved.

(1) Precision (PRC) is computed as displayed in the following equation:

$$\text{PRC} = \frac{\text{TP}}{(\text{TP} + \text{FP})}. \tag{2}$$

(2) Sensitivity (SEN) is evaluated as shown in the following equation:

$$\text{SEN} = \frac{\text{TP}}{(\text{TP} + \text{FN})}. \tag{3}$$

(3) Accuracy (ACR) is computed using the following equation:

$$\text{ACY} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FN} + \text{FP})}. \tag{4}$$

(4) *F*-score is determined via the following equation:

$$F - \text{Score} = 2 \times \left[\frac{(\text{PRC} \times \text{SEN})}{(\text{PRC} + \text{SEN})} \right]. \tag{5}$$

TABLE 2: Details of the utilized data.

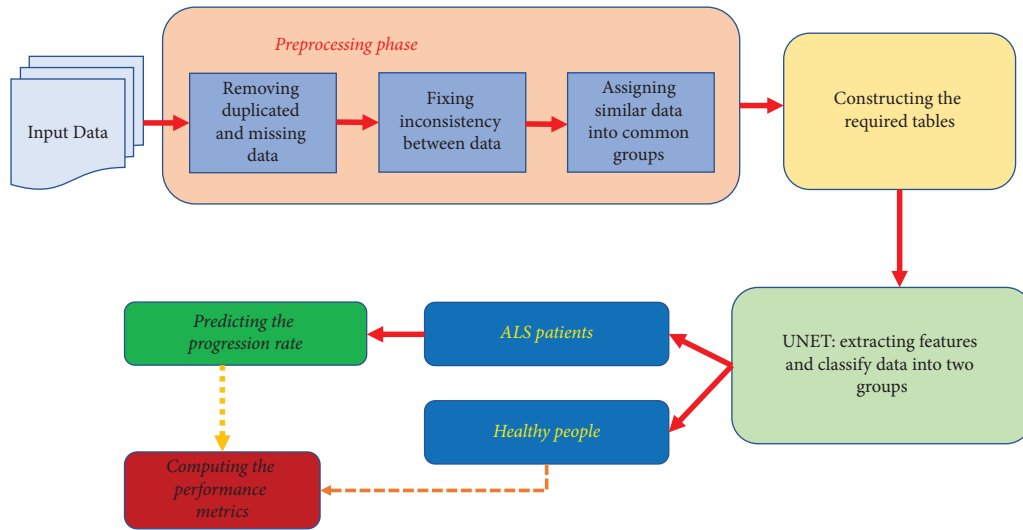| Number of records					
Male			Female		
ALS patients	Healthy		ALS patients		Healthy
354	286		471		428
Training		Validation		Testing	
ALS	Healthy	ALS	Healthy	ALS	Healthy
782	449	67	41	147	53

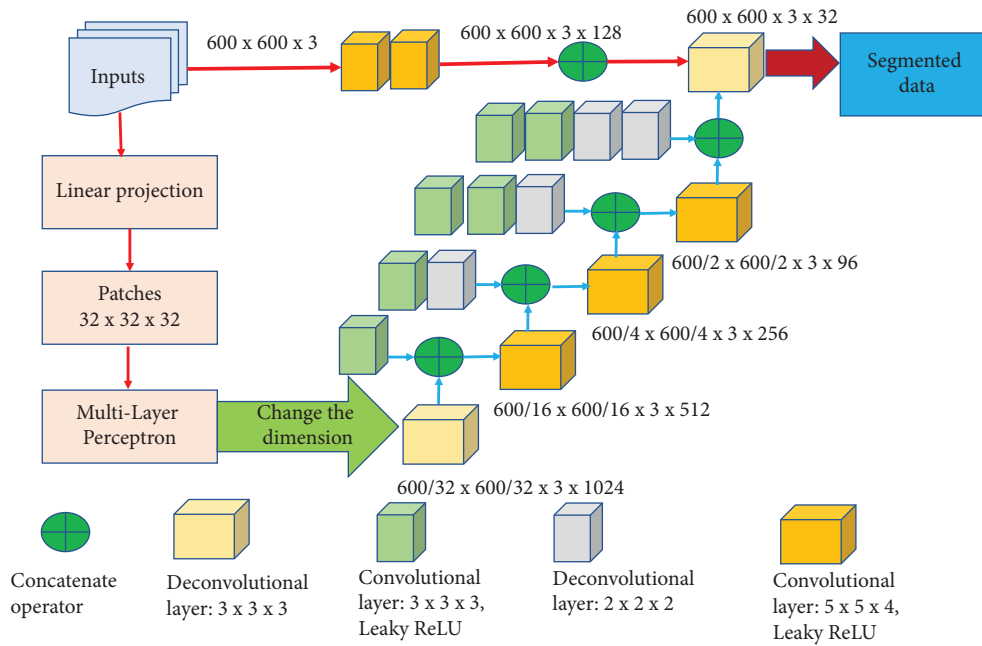FIGURE 1: The flowchart of the developed approach.

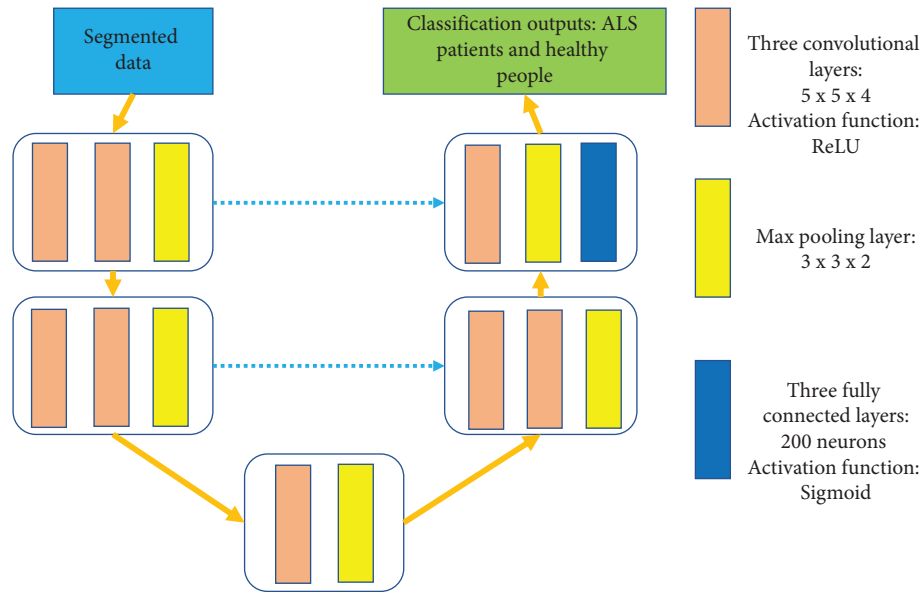FIGURE 2: The internal structure of the segmentation phase.

Figure 3: The flowchart of the categorization stage.

Table 3: The utilized configurations of hyperparameters.

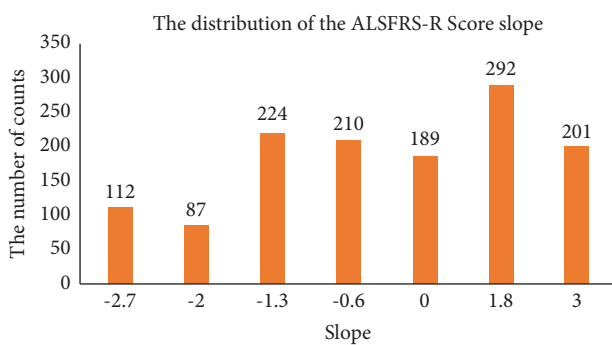| Name of the hyperparameters	Values
Learning rate (L)	0.01
Batch size	32
Dropout	0.05
Optimizer	Adam
Regression weight	0.001
Momentum	0.75
Activation functions	Leaky ReLU and Sigmoid
Number of iterations	6500
Number of epochs	45

Figure 4: The slope distribution of the ALSFRS-R score.

Table 4: Sample of the calculated weights.

Patient ID	Weight
1	0.51
2	0.761
3	0.87
4	0.56
5	0.032

(5) Dice (DIC) is calculated as shown in the following equation:

$$\text{DIC} = \frac{2\text{TP}}{(2\text{TP} + \text{FP} + \text{FN})}. \tag{6}$$

(6) Jaccard Index (JI) determines an overlap area between the detected fire area and the ground truth label. This quantity is computed as illustrated in the following equation:

$$\text{JI} = \frac{(\text{TL} \cap \text{PL})}{(\text{TL} \cup \text{PL})}, \tag{7}$$

where TL refers to the true label and PL represents the predicted label. Moreover, the nominator refers to the intersecting objects, while the dominator denotes the number of alliances between two groups.

4. Results and Discussion

This section provides an analysis to predict ALS disease and its development rate through several experiments. In addition, an evaluation of any signs that affect the proposed

method and its outcomes is provided. All performance indicators are assessed through one dataset. To confirm the association and correlation between data and their actual classes, the data were distributed evenly into three sets as shown in Table 1. The developed deep-learning-based approach was examined and evaluated using the MATLAB platform, which was installed on a machine. This machine was running with Windows Pro 11 using an Intel Core I7 8th Gen., 16 GB RAM, 64-bit operating system, 2 GHz.

4.1. Predicting Results. Due to the difficulty of finding the ALS dataset, we trust the utilized data and work on them with confidence. A hundred healthy people from the training and testing sets were selected as the control group in this study. All related data for the control group were identified and counted as well. The implemented method was training using 1231 records as listed in Table 2. A comparison of similarities between the two constructed groups was conducted using statistical analysis. This procedure shorted inputs of two groups by deleting unwanted values. The estimated average values of all considered performance indicators are shown in Table 5. The model accomplished 85.21% accuracy and 86.05% *F*-score, while precision and sensitivity were 84.86% and 84.43%, respectively. These outcomes were obtained using 6500 iterations; however, increasing the number of iterations enhanced the model's accuracy by nearly 6.8%. Moreover, the required processing time increased significantly, which is considered a side effect.

The developed approach calculated individual accuracy for the three main onsets, namely, spinal, bulbar, and limb. These results are illustrated in Figure 5. The implemented approach identified the bulbar type more than the other two types due to its data availability in the utilized dataset.

During the training stage, the running time was nearly 27 minutes, which was significantly higher as the proposed model went through three main stages. These stages were the preprocessing, segmentation, and identification. The last two stages consumed most of the running time. In order to minimize the execution time, the patch size of each segmented data was decreased partially by 30%–50% and the achieved running time was noticeably good. The execution time went down from 27 minutes to 18 minutes. Figure 6 reveals the maximum attained results of all the considered performance indicators.

Computing the running time of the developed approach to categorize input data in seconds, the number of applied variables within the method and the number of floating-point operations per second (FLOPS) were crucial; thus, they were measured and evaluated. These assessments express the calculation complexity of the presented model. Both FLOPS and the number of parameters were in millions. Table 6 shows these results. The approach created massive FLOPS and the number of variables due to the internal structure of the internal and the number of used characteristics. Nevertheless, the final outcomes were favorable and promising. The running time refers to the achieved time after shortening the patch size by nearly 45%.

TABLE 5: The results of the performance indicators.

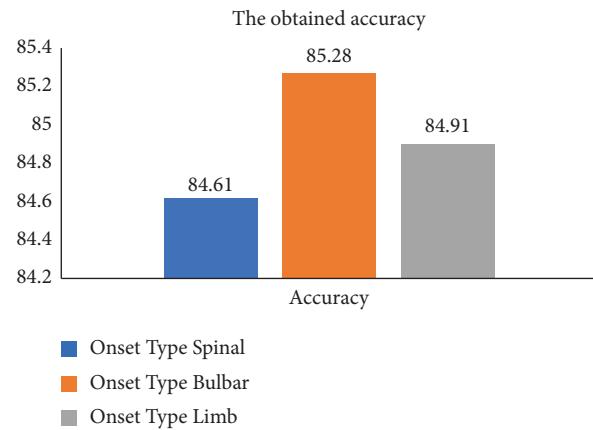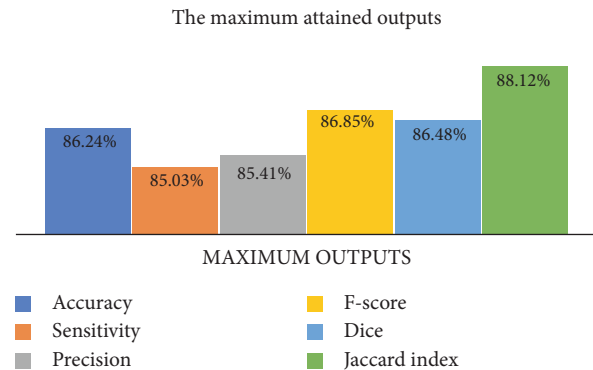| Performance indicators	Average results (%)
Accuracy	85.21
Sensitivity	84.43
Precision	84.86
F-score	86.05
Dice	85.88
Jaccard Index	87.19

FIGURE 5: Accuracy of different onsets.

FIGURE 6: The largest accomplished outputs.

TABLE 6: The assessment results of the computation complexity.

Running time (s)	FLOPS (M)	Number of parameters (M)
78	57.71	63.82

Figure 7 demonstrates two sample graphs of outcomes, which are a chart of cross-entropy and the receiver operating characteristic curve (ROC) for all the three sets, namely, training, validation, and testing.

Tables 7 and 8 reveal the yield grouping outcomes on the testing set and a comparison assessment between several developed models [6, 7, 9, 12, 13, 15, 18] and the proposed approach, respectively. The identification results are ALS and healthy. The comparative assessment evaluation involves the deployed tool, accuracy, *F*-score, and Dice. Table 8 shows that the presented algorithm in this study produces
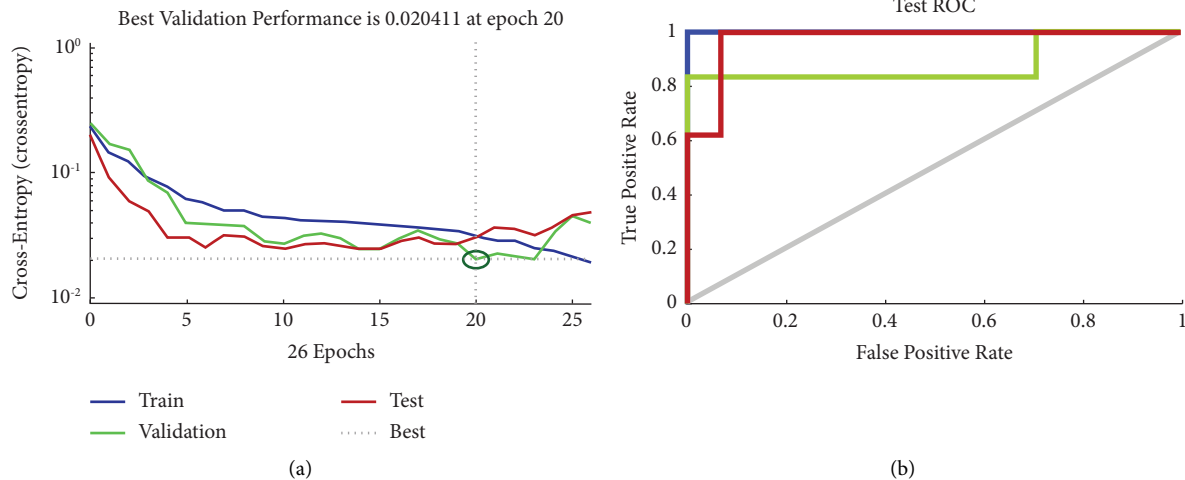
(a)

(b)

Figure 7: (a) The chart of the cross-entropy result. (b) The achieved ROC curve.

Table 7: Identification outputs.

Identified ALS patients	Identified healthy people
117	39

Table 8: The conducted assessment results.

References	Deployed technology	Accuracy	F-score	Dice
[6]	Standard feedforward neural network (FFNN), a convolutional neural network (CNN), and a recurrent neural network (RNN)	N/A	N/A	N/A
[7]	Unsupervised uniform manifold approximation and projection (UMAP) modeling, semisupervised (neural network UMAP) modeling, and supervised (ensemble learning based on LightGBM) modeling	93.24%	N/A	N/A
[9]	Comorbidities and associated indicators using electronic medical records (EMRs)	83.7%	73.95%	N/A
[12]	Induced pluripotent stem cells (iPSCs)	N/A	N/A	N/A
[13]	A multilayer neural network	96.63%	N/A	N/A
[15]	Stacked autoencoders	88%	N/A	N/A
[18]	A neural network	68.8%	N/A	N/A
[26]	XGBoost (XGB) and SHAP	70.7%	N/A	N/A
The presented approach	New UNET architecture	85.21%	86.05%	85.88%

promising results and surpasses some implemented methods in the literature. The attained results in Table 7 reveal that the suggested method identified nearly 84% of the data appropriately.

4.2. Estimation of the Progression Rate. The developed approach estimates the development rate of the ALS disease if a patient is predicted to be diagnosed with the disease. This is performed by drawing the slope of the ALSFRS-R score for only the predicted ALS patients. Figure 8 illustrates the slope diagram. It says that the survival rate probability decreases as time goes on. In addition, by the end of the first year, the survival rate becomes 30% and the death is ensured by the end of the coming years.

Exploring the effect of decreasing the number of utilized features was conducted in this research. The number of characteristics was reduced to seven features only. These features were selected based on the achieved values of the

ALSFRS-R score, which were Q1 speech, Q3 swallowing, Q4 handwriting, Q6 dressing, Q7 turning in bed, Q8 walking, and Q9 climbing. We noticed that the considered performance indicators went down dramatically by more than 40%. This shows that the number of utilized characteristics plays a considerable role.

4.3. Discussion. In this research, an artificial intelligence-based solution to predict ALS disease and its development rate is presented using one dataset from the GitHub repository. It is good to mention that this dataset does not represent a typical distribution. Nevertheless, it supported this study and provided favorable information. The presented algorithm generated promising outcomes since its accuracy lies in an acceptable range from 82% to 87%. This range is better than what was achieved in [9, 18]. In addition, the utilized features contributed to the prediction system and the estimation of the progression rate. The proposed
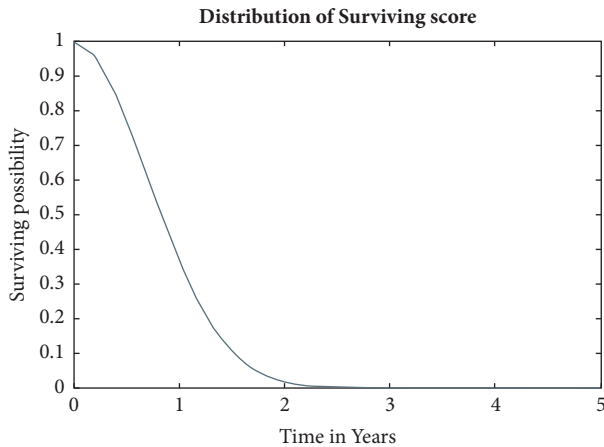
FIGURE 8: The slope of the survival rate.

method was compared with some developed state-of-the-art models in the literature and provided a good insight finding. However, no conclusive advantage was obtained for the assurance intervals.

It is noticed that using artificial intelligence (AI) solutions based on deep learning can give advantages and gain favorable outcomes in terms of accuracy and precision. Explaining and interpreting AI structures are difficult. However, these methods can be deployed and used to support and assist physicians in their diagnosis to provide good treatment plans. Identifying ALS disease and its progression rate were the main aims of this study. Various deep-learning technologies were applied. However, their results were undesired. Thus, these results were neglected. We believe that this occurred due to the limited data availability and how the methods were deployed and interacted with the used features. To prove the efficacy of the presented approach and its suitability, several statistical tools and performance indicators were applied and evaluated. Furthermore, the prediction algorithm was analyzed using different configurations. To improve the findings, the Adam optimizer tool was adopted and it showed a key role in enhancing accuracy by approximately 4.78% and reducing the execution time by less than 7%. Among the implemented works, the authors in [13] achieved the highest accuracy, while the proposed model attained moderate outputs but no specific solution could provide an absolute ALS diagnosis. The presented method in this study can be deployed to identify the considered disease early and it is cost-effective. However, the execution time is considered high and can be seen as a disadvantage.

5. Conclusion

In this article, an artificial intelligence-based algorithm to predict ALS and its development rate is presented. It is obvious that the system's accuracy is increased if the quality of utilized data is good enough to let the model pulls out features without any issues. The quality of data can be improved by performing some required operations, such as cleaning and removing all associated entries of missing

values. Increasing the number of used features in the prediction algorithm enhances its findings if these characteristics are trained well. Even though the applied dataset was small, the outputs of the prediction model are higher than 80%, which is acceptable. These outputs were compared with other AI solutions and showed promising conclusions. The presented approach is very cost-effective; however, its running time is a drawback, and this can be minimized by reducing the number of utilized layers and their associated parameters in the segmentation and learning phases. Moreover, the computed value of the false positive rate increases if the utilized dataset contains symptoms that are similar to ALS disease. The proposed algorithm shows that the detection of the disease in its early stage can be realized. This detection can provide a good plan for treatment and quality of life for diagnosed patients. In addition, the implemented approach can be applied by healthcare providers to support and aid physicians in diagnosing ALS properly.

Future work is projected to enhance the identification outputs and minimize the running time for the whole process. In addition, decreasing the complexity of the prediction algorithm is another intention of the projected future work.

Data Availability

The data used to support the findings of this study were obtained from the GitHub repository [21] and are available at the following link: https://github.com/yonghao206/Origent.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] E. Tavazzi, E. Longato, M. Vettoretti et al., "Artificial intelligence and statistical methods for stratification and prediction of progression in amyotrophic lateral sclerosis: a systematic review," *Artificial Intelligence in Medicine*, vol. 142, pp. 102588–102617, 2023.

[2] A. Behler, H. P. Muller, A. C. Ludolph, and J. Kassubek, "Diffusion tensor imaging in amyotrophic lateral sclerosis: machine learning for biomarker development," *International Journal of Molecular Sciences*, vol. 24, no. 3, pp. 1911–1916, 2023.

[3] E. L. Feldman, S. A. Goutman, S. Petri et al., "Amyotrophic lateral sclerosis," *The Lancet*, vol. 400, no. 10360, pp. 1363–1380, 2022.

[4] J. Thome, R. Steinbach, J. Grosskreutz, D. Durstewitz, and G. Koppe, "Classification of amyotrophic lateral sclerosis by brain volume, connectivity, and network dynamics," *Human Brain Mapping*, vol. 43, no. 2, pp. 681–699, 2022.

[5] V. Renga, "Brain connectivity and network analysis in amyotrophic lateral sclerosis," *Neurology Research International*, vol. 2022, Article ID 1838682, 20 pages, 2022.

[6] C. Pancotti, G. Birolo, C. Rollo et al., "Deep learning methods to predict amyotrophic lateral sclerosis disease progression," *Scientific Reports*, vol. 12, no. 1, pp. 13738–13810, 2022.

[7] F. Faghri, F. Brunn, A. Dadu et al., "Identifying and predicting amyotrophic lateral sclerosis clinical subgroups: a population-based machine-learning study," *The Lancet Digital Health*, vol. 4, no. 5, pp. e359–e369, 2022.

[8] H. Wang, "Phylogenetic analysis of MicroRNA biomarkers for amyotrophic lateral sclerosis," *Biocell*, vol. 45, no. 3, pp. 547–561, 2021.

[9] C. H. Huang, B. S. Yip, D. Taniar, C. S. Hwang, and T. W. Pai, "Comorbidity pattern analysis for predicting amyotrophic lateral sclerosis," *Applied Sciences*, vol. 11, 2021.

[10] F. Faghri, F. Brunn, A. Dadu, P. Errals, and E. Zucchi, "Identification and prediction of ALS subgroups using machine learning," *The Lancet Digital Health*, vol. 4, pp. 1–49, 2021.

[11] R. Khosla, M. Rain, S. Sharma, and A. Anand, "Amyotrophic lateral sclerosis (ALS) prediction model derived from plasma and CSF biomarkers," *PLoS One*, vol. 16, no. 2, Article ID e0247025, 2021.

[12] K. Imamura, Y. Yada, Y. Izumi et al., "Prediction model of amyotrophic lateral sclerosis by deep learning with patient induced pluripotent Stem Cells," *Annals of Neurology*, vol. 89, no. 6, pp. 1226–1233, 2021.

[13] P. Rusiya and N. S. Chaudhari, "Amyotrophic lateral sclerosis EEG classification using deep neural network and TLBO," in *Proceedings of the International Conference on Innovative Computing and Communications*, pp. 1–7, Delhi, India, June 2020.

[14] J. M. Meier, H. K. V. D. Burgh, A. D. Nitert, P. Bede, and S. C. D. Lange, "Connectome-based propagation model in amyotrophic lateral sclerosis," *Annals of Neurology*, vol. 85, no. 5, pp. 725–738, 2020.

[15] S. P. Menon, "Deep learning for prediction of amyotrophic lateral sclerosis using stacked auto encoders," *International Journal of Biomedical Data Mining*, vol. 1, no. 1, pp. 1–134, 2020.

[16] V. Grollemund, P. F. Pradat, G. Querin et al., "Machine learning in amyotrophic lateral sclerosis: achievements, pitfalls, and future directions," *Frontiers in Neuroscience*, vol. 13, pp. 135–228, 2019.

[17] C. Schuster, O. Hardiman, and P. Bede, "Survival prediction in Amyotrophic lateral sclerosis based on MRI measures and clinical characteristics," *BMC Neurology*, vol. 17, no. 1, pp. 73–10, 2017.

[18] H. K. van der Burgh, R. Schmidt, H. J. Westeneng, M. A. de Reus, L. H. van den Berg, and M. P. van den Heuvel, "Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis," *NeuroImage: Clinica*, vol. 13, pp. 361–369, 2017.

[19] P. Rong, Y. Yunusova, J. Wang, and J. R. Green, "Predicting early bulbar decline in amyotrophic lateral sclerosis: a speech subsystem approach," *Behavioural Neurology*, vol. 2015, Article ID 183027, 11 pages, 2015.

[20] M. Elamin, P. Bede, A. Montuschi, N. Pender, A. Chio, and O. Hardiman, "Predicting prognosis in amyotrophic lateral sclerosis: a simple algorithm," *Journal of Neurology*, vol. 262, pp. 1447–1454, 2015.

[21] E. Longinetti and F. Fang, "Epidemiology of amyotrophic lateral sclerosis: an update of recent literature," *Current Opinion in Neurology*, vol. 32, no. 5, pp. 771–776, 2019.

[22] G. Logroscino, D. Urso, and R. Tortelli, "The challenge of amyotrophic lateral sclerosis descriptive epidemiology: to estimate low incidence rates across complex phenotypes in different geographic areas," *Current Opinion in Neurology*, vol. 35, no. 5, pp. 678–685, 2022.

[23] H. Ilieva, M. Vullaganti, and J. Kwan, "Advances in molecular pathology, diagnosis, and treatment of amyotrophic lateral sclerosis," *BMJ*, vol. 383, Article ID e075037, 2023.

[24] M. Vidovic, L. H. Müschen, S. Brakemeier, G. Machetanz, M. Naumann, and S. Castro-Gomez, "Current state and future directions in the diagnosis of amyotrophic lateral sclerosis," *Cells*, vol. 12, no. 5, p. 736, 2023.

[25] Y. Hao, "Yonghao206," 2023, https://github.com/yonghao206/Origent.

[26] V. D. Stefano, F. Prinzi, M. Lugetti, M. Russo, and S. Tozza, "Machine learning for early diagnosis of ATTRv amyloidosis in non-endemic areas: a multicenter study from Italy," *Brain Sciences*, vol. 13, no. 805, pp. 1–10, 2023.