International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 11 | Nov - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Explainable Multimodal AI-Based Clinical Decision Support System for Neuromuscular Disorder Diagnosis (NMD-CDSS)

Mrs. Suchethana H.C[1], Deeksha Ganesh[1], Divya P[1], Shubrath Hegde K.B[1], and Suvas R Gowda K.S[1]

[1]Department of Information Science and Engineering, Jawaharlal Nehru New College of Engineering, Shivamogga, Karnataka, India – 577201.

Abstract—This paper proposes NMD-CDSS, an explainable multimodal clinical decision support system for analysing neuromuscular disorders. The system uses three different Deep Learning based modal specific pipelines to classify the data. The MRI pipeline uses the ResNet-18 model to classify the image slices into either healthy or muscular dystrophy. The EMG pipeline uses the CNN-BiLSTM model to classify the EMG signals from a needle into healthy, myopathy or neuropathy. The clinical pipeline uses a fine-tuned BioBERT model to classify the clinical text into one of the five diagnostic categories. Individual predictions are visualised using GradCAM, temporal gradient and token attribution. The web interface developed using Flask provides individual predictions, confidence scores and an overall summary report. The results obtained for the various datasets used were 88.50% (±7.21% on patient data) accuracy on the MRI data (87.06% (±5.20%) accuracy on slice data), 96.08% validation accuracy on EMG data (macro-F1 of 0.9532) and 94.74% test accuracy on clinical data (macro-F1 of 0.9463). An attention based fusion module is used for experimentation purposes only as the datasets are not available in a patient paired format. None of these results can be used as evidence of effectiveness in a clinical setting. Real data and clinical validation will be required before being used for diagnostic purposes.

***Index Terms*—neuromuscular disorders, multimodal learning, explainable AI, ResNet-18, CNN-BiLSTM, BioBERT, Grad-CAM, clinical decision support.**

I. Introduction

Neuromuscular disorders (NMD) affect how nerves, muscles, and the neuromuscular junction work together causing muscle weakness, fatigue, muscle wastage, and motor coordination issues. Common disorders include amyotrophic lateral sclerosis (ALS),myasthenia gravis, muscular dystrophy and peripheral neuropathies. Accurate and timely diagnosis is challenging due to the over lapping early symptoms many of these disorders' share [1].Diagnosis is often performed using EMG, MRI, nerve conduction studies, lab testing and clinical evaluation. Often data is analyzed individually which can slow down diagnosis and leave some of the data uninterpreted. Manual interpretation of data also relies heavily on expert knowledge. By using Deep Learning techniques (CNN, LSTM, Transformer, Biomedical Language Models) to extract discriminatory patterns from signals, images and text, it has motivated multimodal decision support systems. However, models must be interpretable for clinical use. The aim of this work is to develop an explainable modular web-based research prototype called NMD-CDSS. This processes the data using individual preprocessing steps and models for the MRI, EMG and clinical data. Explanations are provided for each prediction based on the modality. The three outputs produced are shown individually. An experimental fusion module is provided.

Overall contributions are:

- Three models trained and evaluated individually (MRI, EMG, clinical) with cross validated results for the MRI data at a patient level.
- GradCAM, temporal gradient and token attributions are provided for each modality.
- A full web application is provided using Flask with input validation, inference and report generation.
- A clear distinction between validated individual workflow and experimental fusion.

II. Related Work

EMG Diagnosis Jose et al. combined the lifting wavelet transform, fractal dimension and local binary pattern features with an MLP and majority voting to classify intramuscular EMG signals into myopathy and neuropathy. This achieved 99.87% accuracy on a limited dataset of 250 signals [2]. Mani and Nagaraj optimised the Bi-LSTM model using the Grey Wolf Optimizer to achieve 95% accuracy on the NinaPro DB1 dataset. Validation was mainly performed on a single public dataset [3]. A randomised trial found that using AI to aid EMG reporting did not improve performance when compared to physician only reporting. However, issues were found integrating the AI into the workflow [4]. Imaging and Explainability Hosny et al., proposed an EfficientViT explainable model using GradSHAP for diagnosing brain abnormalities using MRI data. Although this shows the importance of having explanations, it is computationally expensive and relied on public datasets [5]. Parameters of diffusion tensor imaging such as fractional anisotropy have been used to quantify motor neuron damage in ALS [6]. Neuromuscular ultrasound has been investigated as a non-invasive indicator of nerve and muscle degeneration [7]. A review looking at AI tools for NF1 related tumours highlights the lack of large standardised datasets for rare diseases [8]. Multimodal and Sensor Based Wearable multimodal sensors (IMU, plantar pressure and ECG) have been used with gradient boosting to classify peripheral neuropathy. EEG spectral power and coherence during the pre-motor phase has been proposed as a biomarker for ALS [9]. A review of 89 studies on AI enabled multimodal analysis for motor neuron disease shows the importance of having unified pipelines using imaging, omics and clinical biomarkers. However, there was limited interpretability, incomplete data integration and poor cross validation practices [10]. Clinical Next generation and exome sequencing achieves a high diagnostic yield for undiagnosed paediatric and consanguineous families [11] [12]. Surveys highlight that fatigue, weakness and reduced physical activity are the most troublesome symptoms in NMD [1]. Myotonometry, wearable assistive devices (Fabric Muscle Shoulder Exosuit, BCI Mouse), respiratory support and conventional/complementary therapies have been investigated. However, these were mostly performed on small cohorts [13] [14] [15] [16] [17].Overall, most AI systems are unimodal and often using small or single source datasets. There is limited explainability. NMD-CDSS focuses on creating a modular and explainable multi-input system.

III. Proposed System Architecture

NMD-CDSSNMD-CDSS has a modular design. Each modality has its own preprocessing stage, model and explanation. The MRI uses the ResNet18 [16] model, the EMG uses the CNN-BiLSTM model and the clinical data uses the BioBERT [17] model. The backend is written in Flask and handles data validation, model loading and inference. The frontend communicates with the backend using HTML, CSS and JavaScript. The class probabilities and explanations are shown. Any combination of the three modalities can be sent. As there are no verified patients across the datasets and different labels are used, the predictions are reported individually and never averaged. The experimental fusion module is separate.

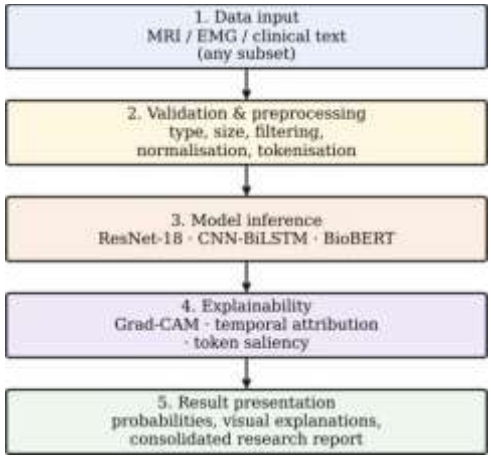


Fig. 2. shows the overall process of inputting, preprocessing, inference, explanations and presenting the results.

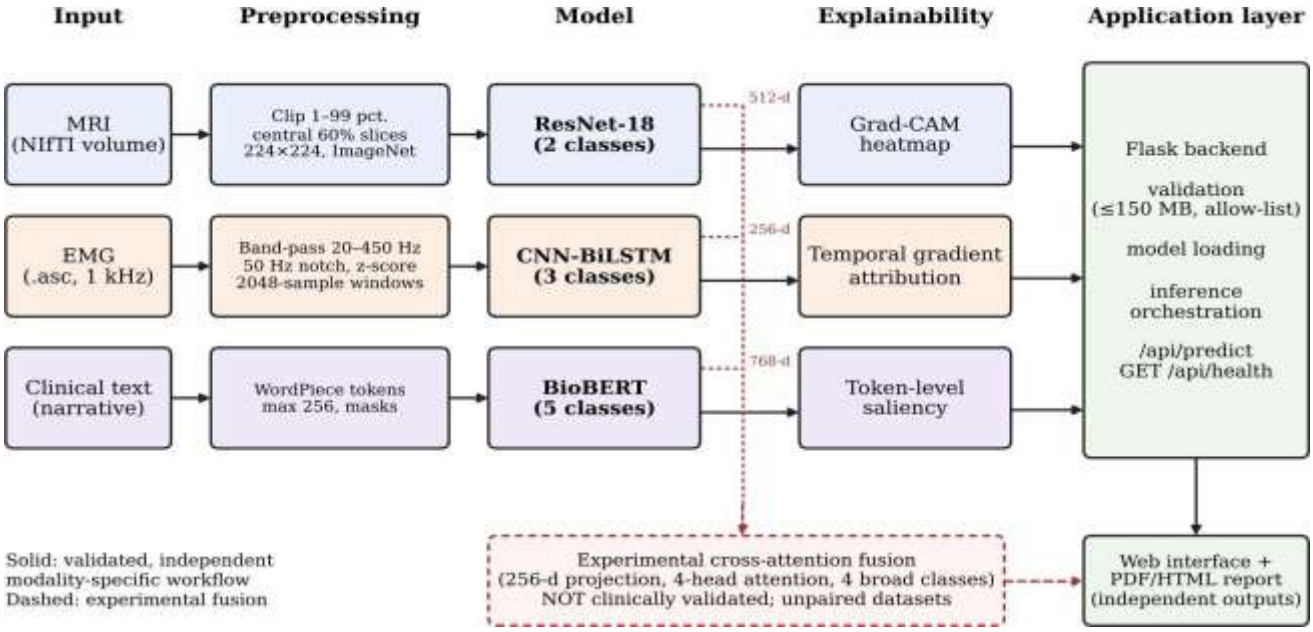


Fig. 1. shows the architecture of NMD-CDSS. The solid lines show the individual process carried out in the application. The dashed red elements show the experimental fusion (feature dimensions - 512, 256, 768) which has not been clinically validated.

IV. Materials and Methods

A. MRI

The MRI uses the MMThigh muscle MRI dataset (HuashanMyo, Folkhälsan). The dataset is stored in .nii / .nii.gz files. The implementation used 1117 slices (191 Healthy, 926 Muscular Dystrophy). The control group were mapped to healthy and the limb-girdle, myotonic and tibial muscular dystrophy, inclusion body myositis and HMERF groups were mapped to muscular dystrophy. Volumes were converted into 2D using clipping (1-99th percentile) (Eq. 1). Representative axial slices were taken from the centre 60% of the volume. The slices were resized to 224 x 224, replicated to 3 channels and ImageNet normalized. Up to 15 slices were processed at inference and the softmax averaged (Eq. 2).

$$x' = min\left(max\left(x, p_1\right), p_{99}\right) \quad (1)$$

$$P_{vol} = \frac{1}{K}\sum_{k=1}^{K} softmax\left(z_k\right), K \leq 15 \quad (2)$$

where p1 is the 1st percentile, p99 is the 99th percentile and zk is the logits for slice k.

B. EMG

The EMG data comes from a dataset on Mendeley Data. The dataset contains 400 ASCII recordings using concentric needle electrodes from the deltoid and biceps brachii muscles at a sampling rate of 1000Hz. The data was split into healthy, myopathy and neuropathy. The header information was removed and DC offset subtracted. The data was filtered using a 4 th order zero phase Butterworth bandpass filter (20-450 Hz) and a 50Hz notch filter (Q = 30).

The data was then z-score normalised (Eq. 3) and split into windows of 2048 data points (approx. 2.05s) with a stride of 1024 data points (50% overlap).

$$x_t = \frac{x_t - \mu}{\sigma} \quad (3)$$

The CNN-BiLSTM contains 3 convolutional layers (32, 64, 128 filters with kernel sizes 7, 5, 3). Each convolutional layer has batch normalisation, ReLU activation and max pooling. The BiLSTM contains 2 layers with 128 hidden units in each direction (256 dimension output). Dropout was used and the final layer was fully connected with 3 outputs. Cross entropy loss function was used during training along with Adam optimizer (lr = 10^{-3}). Early stopping was used.

The window probabilities were averaged to obtain a final prediction. The window overlap was 50% so the metrics could be optimistic. Unless the data is split into whole recordings/patients.

C. Clinical Text

The clinical text contains 250 structured constructed narratives of patient demographics, complaints, duration of symptoms, weakness, sensory and reflex findings. The text was sanitised (minimum length of 15 characters), tokenised using the BioBERT WordPiece tokeniser (Maximum 256 tokens with padding, truncating and attention masks). dmis-lab/biobert-base-cased-v1.1 was used to classify the text. The text was fine tuned with a 5-output head using the Hugging Face Trainer. AdamW optimizer was used with (lr = 2 x 10^{-5}). Weight decay of 0.01 was used. The batch size was 16 and linear learning rate decay was used. 175 records were used for training, 37 for validation and 38 for testing.

D. Experimental Multimodal Fusion Pathway

The ResNet-18 (512 dimensional), BiLSTM (256 dimensional) and BioBERT (768 dimensional) feature vectors were mapped into a common 256-dimensional space. The three tokens were passed through a 4 head attention layer and final 4 class head (healthy, muscular dystrophy, myopathy, neuropathy). This is just for demonstration purposes and has not been trained or validated on real paired patients.

V. Model Architecture and Mathematical Formulation

Each model maps the input x to logits z. The probabilities are calculated using softmax. The models were trained using the crossentropy loss function to minimise the loss. The loss function was:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c} \log \frac{e^{z_{i,c}}}{\sum_{j=1}^{C} e^{z_{i,j}}} \quad (4)$$

VII. Experimental Results and Discussion

The application can be used to view the class probabilities, explanations and summary report for each input modality. In the example given in the project, the MRI model predicted muscular dystrophy (99.9%). The EMG model predicted myopathy (97.9%). The clinical model predicted healthy (45.8%). These are just examples of outputs from a research prototype and are not diagnoses. The outputs are not consistent as they are designed to be individual.

MRI. The ResNet-18 [16] model was initialised with ImageNet weights. The weights up to layer3 were frozen. Layer4 and a new head was fine tuned. The head contains dropout (p = 0.4) and a fully connected layer with 512 to 2 outputs. The model was cross validated using 5 folds of StratifiedGroupKFold to ensure all slices from the same patient were in the same fold. The weights used in deployment were retrained on all 1117 slices for 2 epochs.

EMG. The CNN was used to extract local patterns from the waveform and the BiLSTM was used to capture the temporal patterns in the window. The window probabilities were averaged to give a final result for the recording. Where Nw is the number of valid windows.

$$P_{\text{rec}} = \frac{1}{N_w}\sum_{w=1}^{N_w} p_w \quad (5)$$

Fusion (Experimental) Each feature vector fm was mapped into a 256-dimensional vector using hm = Wmfm + bm ∈ ℝ 256. The three tokens were passed into the multi-head attention (4 heads). The multi-head attention was calculated using:

$$Attn(Q, K, V) = softmax\left(\frac{QK^{\top}}{\sqrt{d_k}}\right)V \quad (6)$$

Where Q, K, V are linear projections of H and dk = 64 for each head. The attention weights do not show any clinically meaningful cross modal relationships.

VI. Explainable AI Methodology

MRI - Grad-CAM [18] The gradient of the class score yc with respect to the feature maps Ak from the last convolutional block was globally averaged across all Z spatial locations to get the channel weights. The weighted sum was rectified to produce a heatmap. This was rescaled and overlayed onto the slice.

$$\alpha_k^c = \frac{1}{Z}\sum_{i,j}\frac{\partial y^c}{\partial A_{ij}^k} \quad (7)$$

$$L_{\text{Grad-CAM}}^c = ReLU\left(\sum_k \alpha_k^c A^k\right) \quad (8)$$

EMG - Temporal Gradient The gradient of the predicted class score was taken with respect to each sample in the input. This was smoothed and normalised to produce a saliency map. This can be seen on the waveforms.

$$a_t = norm\left(smooth\left(\left|\frac{\partial y^c}{\partial x_t}\right|\right)\right) \quad (9)$$

Clinical - Token Attributions The gradient with respect to the token embeddings was calculated. Each token was given a score estimating how much it influenced the choice of class. This showed the words that had an influence on the decision.

These do not show that the area, signal or phrase is abnormal in a clinical sense. They just show what the model was paying attention to.

A. MRI Classification Results

Figure 3 shows the pooled 5-fold confusion matrix for slice data. The accuracy at a patient level and macro-F1 for each fold. The results show that 132 of the 191 healthy slices were correctly classified. 59 were incorrectly classified as muscular dystrophy. 843 of the 926 muscular dystrophy slices were correctly classified. 83 were incorrectly classified as healthy. The accuracy was 87.06% (±5.20%) (macro-F1 0.7850 (±0.0470)).

Slice-level The accuracy for the slice data should not be interpreted as the accuracy for the patient data. Figure 3(C) shows the loss curves for the training and validation data for fold 5. The training loss decreases steadily. The validation accuracy fluctuates throughout the epochs. This should not be interpreted as a validation metric.

Training was stopped after 53 epochs using early stopping. The training and validation loss decreased. The validation accuracy reached ~96% in the latter epochs. The main confusion was between myopathy and neuropathy. 172 myopathy windows were predicted as neuropathy.



Metric	Slice-level (pooled)	Patient-level (CV)
Accuracy	87.08% ± 5.20%	88.50% ± 7.21%
Macro-F1	0.7850 ± 0.0470	0.8336 ± 0.1010
Weighted-F1	0.8710	0.8872
Precision (macro)	0.77	0.85
Recall (macro)	0.80	0.88
Specificity (macro)	0.80	0.86
Evaluated support	1,117 slices	88 subjects (5 folds)

Fig. 3. MRI (ResNet-18). (A) Pooled 5-fold confusion matrix (total K≤151 slices). (B) Accuracy and macro-F1 for patient data. (C) Loss curves showing training and validation loss for fold 5. (D) Summary metrics from the project report.

B. EMG

Figure 4 shows the validation confusion matrix (K≤152 windows) and the training history. The model predicted 6262 healthy windows, 2933 myopathy windows and 2997 neuropathy windows correctly. This achieved a validation accuracy of 96.08%. The macro-F1 was 0.9532. The F1 for each class was: Healthy - 0.983; Myopathy - 0.936; Neuropathy - 0.940.

The windows overlap by 50% and the data may not have been split into whole recordings/patients. This means the values could be optimistic. It does not show that the model generalises to unseen patients.

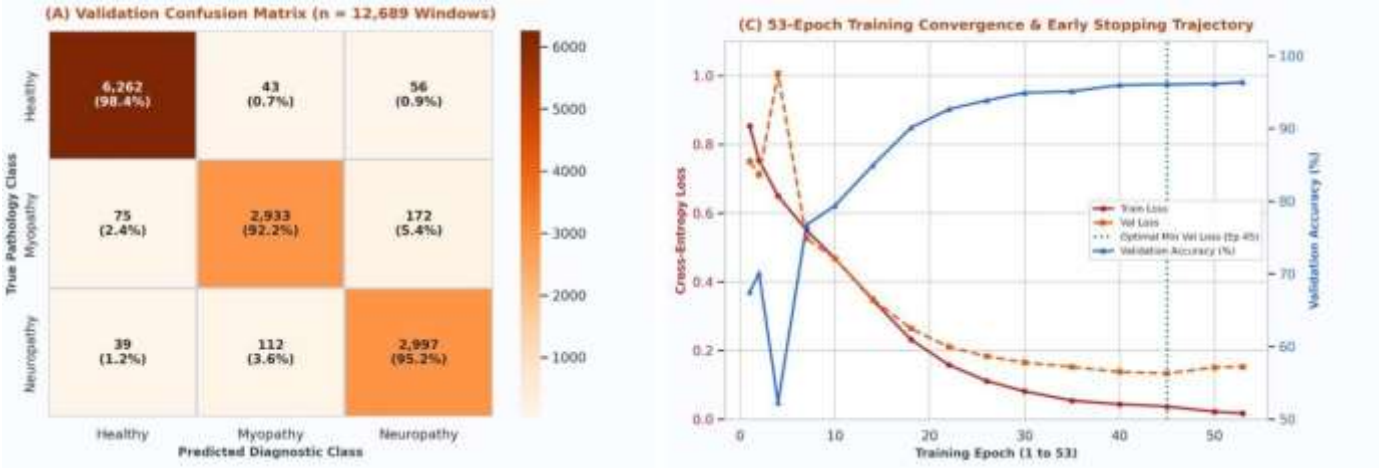


Fig. 4. EMG (1D CNN + BiLSTM). (A) Validation confusion matrix (K≤153 windows). (C) Loss curves showing training and validation loss over 53 epochs.

C. Clinical Text Classification Results

On Figure 5 shows the results for the test data. 36 out of 38 records were correctly classified (94.74% accuracy, 0.9463 macro-F1). All of the records for the other 4 classes were correctly classified. 2 of the ALS records were incorrectly classified. One was predicted as myasthenia gravis and one was predicted as peripheral neuropathy. (Recall - 0.75, F1 - 0.857).

The validation loss decreased from ~1.46 after 10 epochs. The validation accuracy and macro-F1 increased to ~97.3% and 0.9713. The lower accuracy on the test data shows the importance of having a test set. The data is 250 constructed narratives and only 38 records were used for testing.

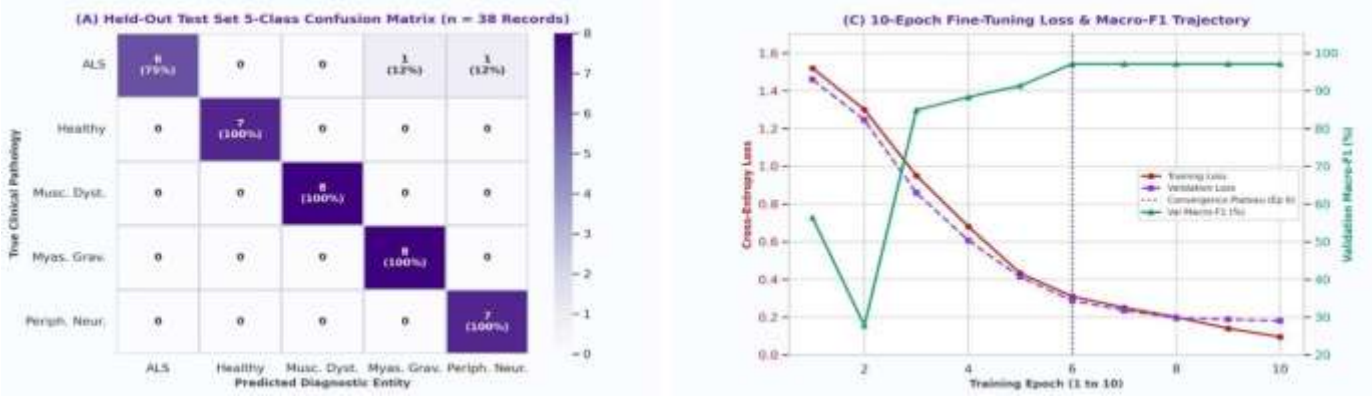


Fig. 5. Clinical (Fine-tuned BioBERT). (A) Test confusion matrix (K≤154 records). (C) Loss curves showing training and validation loss over 10 epochs.

D. Interpretation of Modality-Specific Performance

Table I shows the results for each modality. It is difficult to compare as they perform different tasks (2, 3, 5 classes), use different data types and the evaluation is performed at a different level (patient/slice, window, record). The MRI had the lowest performance but the evaluation was more conservative as it was cross validated based on patients (varied between folds). The EMG had a high performance, but the evaluation was performed at a window level. This could be affected by leakage as there was 50% overlap. The text had a low performance due to the dataset being small and constructed. None of the results should be ranked.

TABLE I. SUMMARY OF REPORTED MODALITY-LEVEL RESULTS (NOT MUTUALLY COMPARABLE).

Modality / model	Evaluation unit	n	Accuracy	Macro-F1
MRI / ResNet-18	Patient (5-fold CV)	88 subjects	88.50 ± 7.21%	0.8336 ± 0.1010
MRI / ResNet-18	Slice (pooled)	1,117 slices	87.06 ± 5.20%	0.7850 ± 0.0470
EMG / CNN-BiLSTM	Window (validation)	12,689 windows	96.08%	0.9532
Text / BioBERT	Record (held-out test)	38 records	94.74%	0.9463

VIII. LIMITATIONS, ETHICAL CONSIDERATIONS, AND FUTURE WORK

Data The results for the MRI are reported at both slice and patient level. This should not be confused. The patient data is split into small folds (17-18 subjects). The EMG results are reported at a window level. Due to the windows overlapping by 50% the results could be optimistic, unless the data is split into whole recordings/patients. The text results are from a small constructed dataset.

Fusion There are no verified patients across the MRI, EMG and text data. The labels are also different. This means that trying to match unrelated records would not give valid examples for training. The fusion is therefore just for experimentation.

Ethical Issues NMD-CDSS is a research prototype and therefore is not an approved diagnostic tool. It does not replace clinical assessment. The explanations do not show causation, they show what areas the model was focusing on. The input data is validated (limited to 150MB, allowed file extensions, tidy up temporary files). This does not guarantee secure uploading. If deployed in a real situation, security measures would need to be put in place to protect the privacy, access control, have security checks and regulation. The benefits could be beneficial for screening but this is hypothetical until validated.

Future work. To improve the system, paired multimodal patient data with consistent reference diagnosis should be used. The EMG data should be validated at a recording level. The text data should be evaluated on real data. External validation with clinicians should be performed and the system integrated securely with health information systems.

IX. CONCLUSION

This paper proposes NMD-CDSS, a modular explainable multimodal research prototype. The system uses the ResNet-18 model for MRI, CNN-BiLSTM for EMG and BioBERT for clinical text. The web interface was created using Flask. The explanations were provided using Grad-CAM, temporal and token attributions. Each model provided promising results on their own data. (88.50% accuracy on patient data, 96.08% on window data, 94.74% on test data). However, the evaluation varied. The EMG and text results had issues with leakage and the size of the data. The fusion was just for experimentation and has not been validated as there was no paired patient data. Larger individual datasets and clinical validation is required for any practical use.

REFERENCES

[1] M. Mancuso, A. Colitta, M. Lavorato, P. Van den Bergh, J. Kirschner, and C. Kornblum, "The most bothersome symptoms in neuromuscular diseases: the ERN EURO-NMD survey," *Orphanet J. Rare Dis.*, vol. 20, no. 221, pp. 1–10, 2025, doi: 10.1186/s13023-025-03742-z.

[2] S. Jose, S. T. George, M. S. P. Subathra, V. S. Handiru, P. K. Jeevanandam, U. Amato, and E. S. Suviseshamuthu, "Robust

classification of intramuscular EMG signals to aid the diagnosis of neuromuscular disorders," *IEEE Open J. Eng. Med. Biol.*, vol. 1, pp. 235–242, 2020, doi: 10.1109/OJEMB.2020.3017130.

[3] R. K. Mani and B. Nagaraj, "Diagnosis and classification of neuromuscular disorders using Bi-LSTM optimized with grey wolf optimizer for EMG signals," *Sci. Rep.*, vol. 15, Art. no. 19274, 2025, doi: 10.1038/s41598-025-03766-2.

[4] A. Gorenshtein, Y. Weisblat, M. Khateb, G. Kenan, I. Tsirkin, G. Fayn, S. Geller, and S. Shelly, "AI-based EMG reporting: A randomized controlled trial," *J. Neurol.*, vol. 272, p. 586, 2025, doi: 10.1007/s00415-025-13261-3.

[5] M. Hosny, A. M. Elshenhab, and A. Maged, "Explainable AI-based method for brain abnormality diagnostics using MRI," *Biomed. Signal Process. Control*, vol. 100, p. 107184, 2025, doi: 10.1016/j.bspc.2024.107184.

[6] Y. Sang and J. Xu, "Evaluation of motor neuron injury in ALS by different parameters of diffusion tensor imaging," *IEEE Access*, vol. 8, pp. 72381–72392, 2020, doi: 10.1109/ACCESS.2020.2985064.

[7] S. H. Elwan, M. S. Hashem, M. M. Ghanem, S. M. Ahmed, H. S. Mohammed, and B. B. Elsayed, "Role of neuromuscular ultrasound in the evaluation of amyotrophic lateral sclerosis," *Egypt. J. Neurol. Psychiatr. Neurosurg.*, vol. 61, no. 20, pp. 1–9, 2025, doi: 10.1186/s41983-025-00951-y.

[8] F. Hellmann, I. Ristow, L. Well, S. Lohse, M. Anokhin, M. Kuhlen, E. André, and A. Harder, "Artificial intelligence-based tools for precision diagnosis and treatment of neurofibromatosis type 1 associated peripheral and central glial tumors," *Orphanet J. Rare Dis.*, vol. 20, p. 551, 2025, doi: 10.1186/s13023-025-04093-5.

[9] M. Talha, M. Kyrarini, and E. Ali, "Physiological features for classification of different types of peripheral neuropathy using multimodal wearable sensors," *IEEE Access*, vol. 14, pp. 8660–8671, 2026, doi: 10.1109/ACCESS.2026.3652722.

[10] S. Bista, A. Coffey, M. Mitchell, A. Fasano, S. Dukic, T. Buxo, E. Giglia, M. Heverin, et al., "Abnormal EEG spectral power and coherence measures during pre-motor stage in amyotrophic lateral sclerosis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 33, pp. 232–242, 2025, doi: 10.1109/TNSRE.2024.3523109.

[11] V. Alevizos and G. A. Papakostas, "AI-enabled multimodal analysis enhances detection of motor neuron disease pathways," *Iran J. Comput. Sci.*, vol. 8, pp. 1331–1351, 2025, doi: 10.1007/s42044-025-00348-3.

[12] P. Kulsirichawaroj, M. Chanvanichtrakool, P. Wattanadilokchatkun, T. Pho-iam, J. Han, and S. Zuchner, "Next-generation sequencing for pediatric-onset neuromuscular disorders unresolved by conventional diagnostic methods," *Pediatr. Res.*, vol. 98, pp. 2195–2202, 2025, doi: 10.1038/s41390-025-04160-4.

[13] T. Aleem, M. Rashid, N. Ahmad, M. F. Asif, M. Tariq, N. A. Malik, and J. A. Poulter, "Exome sequencing reveals broad genetic heterogeneity for neuromuscular disorders in consanguineous Pakistani families," *Eur. J. Hum. Genet.*, vol. 33, pp. 1264–1270, 2025, doi: 10.1038/s41431-025-01915-9.

[14] K. Lukas, K. Gutschmidt, B. Schoser, and S. Wenninger, "Evaluation of myotonometry for myotonia, muscle stiffness and elasticity in neuromuscular disorders," *J. Neurol.*, vol. 270, pp. 5398–5407, 2023, doi: 10.1007/s00415-023-11867-z.

[15] S. J. Park, S. Jo, H.-I. Shin, E. Lee, J. H. Kim, H. Jung, J. M. Kim, Y. L. Lee, S. E. Hyun, W. H. Lee, and C. H. Park, "Soft exosuit based on fabric muscle to assist shoulder joint movements in patients with neuromuscular diseases," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 33, pp. 3866–3877, 2025, doi: 10.1109/TNSRE.2025.3613709.

[16] Y. Jiang, K. Li, Y. Liang, D. Chen, M. Tan, and Y. Li, "Daily assistance for amyotrophic lateral sclerosis patients based on a wearable multimodal brain–computer interface mouse," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 33, pp. 150–161, 2025, doi: 10.1109/TNSRE.2024.3520984.

[17] A. Chakraborti, S. Dutta, M. Das, M. Sonigra, M. S. Khot, and A. R. Wani, "Role of high flow nasal cannula in managing type I respiratory failure in patients with neuromuscular disorders," *BMC Pulm. Med.*, vol. 26, p. 150, 2026, doi: 10.1186/s12890-026-04198-y.

[18] L. Ä. Böhne, C. Wirner, B. Schoser, C. Schröter, and P. Baum, "Frequency and satisfaction of conventional and complementary or alternative therapies for neuromuscular disorders," *Neurol. Res. Pract.*, vol. 5, Art. no. 53, 2023, doi: 10.1186/s42466-023-00281-5.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778.

[20] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang, "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics*, vol. 36, no. 4, pp. 1234–1240, 2020.

[21] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 618–626.