



Providing Choice & Value

Generic CT and MRI Contrast Agents



CONTACT REP

AJNR

Severity Classification of Pediatric Spinal Cord Injuries Using Structural MRI Measures and Deep Learning: A Comprehensive Analysis Across All Vertebral Levels

Zahra Sadeghi-Adl, Sara Naghizadehkashani, Devon Middleton, Laura Krisa, Mahdi Alizadeh, Adam E. Flanders, Scott H. Faro, Ze Wang and Feroze B. Mohamed

This information is current as of July 31, 2025.

AJNR Am J Neuroradiol published online 7 April 2025
<http://www.ajnr.org/content/early/2025/04/06/ajnr.A8770>

Severity Classification of Pediatric Spinal Cord Injuries Using Structural MRI Measures and Deep Learning: A Comprehensive Analysis Across All Vertebral Levels

Zahra Sadeghi-Adl, Sara Naghizadehkashani, Devon Middleton, Laura Krisa, Mahdi Alizadeh, Adam E. Flanders, Scott H. Faro, Ze Wang, and Feroze B. Mohamed

ABSTRACT

BACKGROUND AND PURPOSE: Spinal cord injury (SCI) in the pediatric population presents a unique challenge in diagnosis and prognosis due to the complexity of performing clinical assessments on children. Accurate evaluation of structural changes in the spinal cord is essential for effective treatment planning. This study aims to evaluate structural characteristics in pediatric patients with SCI by comparing cross-sectional area (CSA), anterior-posterior (AP) width, and right-left (RL) width across all vertebral levels of the spinal cord between typically developing (TD) and participants with SCI. We employed deep learning techniques to utilize these measures for detecting SCI cases and determining their injury severity.

MATERIALS AND METHODS: Sixty-one pediatric participants (ages 6-18), including 20 with chronic SCI and 41 TD, were enrolled and scanned using a 3T MRI scanner. All SCI participants underwent the International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) test to assess their neurological function and determine their American Spinal Injury Association (ASIA) Impairment Scale (AIS) category. T2-weighted MRI scans were utilized to measure CSA, AP width, and RL widths along the entire cervical and thoracic cord. These measures were automatically extracted at every vertebral level of the spinal cord using the SCT toolbox. Deep convolutional neural networks (CNNs) were utilized to classify participants into SCI or TD groups and determine their AIS classification based on structural parameters and demographic factors such as age and height.

RESULTS: Significant differences ($p < 0.05$) were found in CSA, AP width, and RL width between SCI and TD participants, indicating notable structural alterations due to SCI. The CNN-based models demonstrated high performance, achieving 96.59% accuracy in distinguishing SCI from TD participants. Furthermore, the models determined AIS category classification with 94.92% accuracy.

CONCLUSIONS: The study demonstrates the effectiveness of integrating cross-sectional structural imaging measures with deep learning methods for classification and severity assessment of pediatric SCI. The deep learning approach outperforms traditional machine learning models in diagnostic accuracy, offering potential improvements in patient care in pediatric SCI management.

ABBREVIATIONS: SCI = Spinal Cord Injury, TD = Typically Developing, CSA = Cross-Sectional Area, AP = Anterior-Posterior, RL = Right-Left, ASIA = American Spinal Injury Association, AIS = American Spinal Injury Association, CNN = Convolutional Neural Network.

Received month day, year; accepted after revision month day, year.

From the Department of Electrical and Computer Engineering (Z.S.), Temple University, Philadelphia, PA, USA; Department of Radiology (Z.S., S.N., D.M., L.K., M.A., A.E.F., S.H.F., F.B.M.), and Department of Neurosurgery (M.A.), Thomas Jefferson University, Philadelphia, PA, USA; Department of Diagnostic Radiology & Nuclear Medicine (Z.W.), University of Maryland School of Medicine, Baltimore, MD, USA.

The authors declare no conflicts of interest related to the content of this article.

Please address correspondence to Zahra Sadeghi-Adl, M.S., Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, USA, Zahra.sadeghi.adl@temple.edu.

SUMMARY SECTION

PREVIOUS LITERATURE: Pediatric spinal cord injury (SCI) is a rare yet debilitating condition that significantly impacts motor and sensory function. Previous studies have primarily focused on adult SCI, with limited research on pediatric populations. Structural MRI metrics, such as cross-sectional area (CSA), anterior-posterior (AP) width, and right-left (RL) width, have been explored as potential biomarkers for SCI severity. Traditional assessments rely on clinical examinations like the ASIA Impairment Scale (AIS), but objective imaging-based methods remain underdeveloped. Deep learning approaches have shown promise in medical image analysis, but their application to pediatric SCI classification and severity assessment has not been extensively investigated.

KEY FINDINGS: This study demonstrates that MRI-derived structural parameters, combined with a custom convolutional neural network (CNN), can accurately classify pediatric SCI and predict AIS severity. The CNN achieved 96.59% accuracy in distinguishing SCI from typically developing (TD) participants and 94.92% accuracy in AIS classification, outperforming traditional machine learning models.

KNOWLEDGE ADVANCEMENT: This study establishes a novel deep learning-based framework for pediatric SCI classification and severity assessment, leveraging MRI-derived structural metrics. The results highlight the feasibility of using imaging biomarkers for objective SCI evaluation. The approach offers potential clinical applications, aiding in diagnosis, prognosis, and personalized treatment planning for pediatric SCI patients.

INTRODUCTION

Spinal cord injury (SCI) represents a profound challenge in pediatric medicine, occurring amidst ongoing central nervous system development¹. In children with SCI, over 50% suffer complete injuries with total motor and sensory deficits below the injury level, with younger children being more susceptible to experiencing complete injuries². Despite advances in clinical and imaging techniques, accurately assessing SCI severity in children remains challenging, particularly when traditional methods fall short^{3,4}.

The International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) is the most widely accepted framework for assessing SCI severity⁵. This classification system provides a standardized approach for evaluating neurological function and injury severity. The American Spinal Injury Association (ASIA) Impairment Scale (AIS), which is part of the ISNCSCI, categorizes injury severity from AIS A (complete loss of motor and sensory function) to AIS D (incomplete, with significant motor function preserved), with intermediate levels B and C indicating partial recovery of sensory or motor function. However, the ISNCSCI relies heavily on clinical examinations, which can be challenging to administer accurately in young children^{6,7}. For children under six years of age, participating in and accurately performing the ISNCSCI examination can be particularly problematic, potentially leading to inaccuracies in assessing the true extent of their injury⁶. Furthermore, conditions such as spinal cord injury without MR (SCIWOMR) abnormality pose additional diagnostic challenges, as traditional imaging techniques may not reveal the precise location or extent of the injury⁸.

High-resolution MRI has significantly advanced our ability to visualize and quantify structural changes in the spinal cord^{9,10}. By measuring parameters such as cross-sectional area (CSA), anterior-posterior (AP) width, and right-left (RL) width along the spinal cord, high-resolution MRI provides critical insights into the extent of spinal cord damage¹¹. Studies have shown that changes in cervical spinal cord CSA are closely linked to functional outcomes, with smaller CSA correlating with poorer motor and sensory functions and reduced functional independence¹²⁻¹⁴. For instance, research has demonstrated that the AP width of the spinal cord correlates with sensory function, while the RL width aligns more closely with motor function¹³. Additionally, baseline measurements of AP width have been associated with lower limb motor scores at follow-up in acute SCI cases¹². However, much of this research has focused on adult populations, with limited studies addressing the morphological changes of the cervical spinal cord in children following SCI. Recent findings indicate that quantitative measures like CSA and AP width at the C2/3 level can serve as objective biomarkers for assessing neurological function in pediatric thoracolumbar spinal cord injury (TLSCI)¹⁴. Notably, CSA and RL width in the AIS A/B (motor complete) group were significantly lower compared to the typically developing (TD) group and the AIS C/D (motor incomplete) group, with RL width being the most sensitive biomarker for differentiating AIS A/B from AIS C/D. These findings underscore the potential of quantitative spinal cord measurements as biomarkers for assessing injury severity and predicting neurological recovery.

Despite these advancements, predicting SCI severity using individual structural measurements has not been thoroughly explored, particularly in pediatric populations. Traditional machine learning approaches, such as support vector machines (SVM) and random forests, have been used to analyze MRI data and have demonstrated some efficacy in classification tasks¹⁵. However, these methods often require manual feature extraction and may not fully capture complex, high-dimensional patterns in imaging data. In contrast, deep learning techniques, particularly convolutional neural networks (CNNs), offer an advantage by automatically learning hierarchical features from large-scale imaging datasets¹⁶. To leverage these advantages, we developed a custom CNN model specifically for analyzing pediatrics with chronic SCI. This model incorporates structural measurements, including CSA, AP width and RL width, alongside demographic factors such as height and age in its final layer. These demographic features are critical as they influence spinal cord structure and development, impacting the severity of SCI.

The goal of this paper is twofold: first, to develop and apply deep learning techniques to classify SCI and predict the severity based on these cross-sectional measurements, an innovative approach that to the best of our knowledge has not been explored in the pediatric context; and second, to demonstrate the differences in these measurements at all vertebral levels in pediatric SCI compared to TD controls, a comparison that has not been previously conducted with such a number of TD participants. By integrating high-resolution MRI imaging with advanced computational methods, this study aims to enhance both the understanding and assessment of pediatric SCI, offering new insights into early diagnosis and treatment planning. The importance of this work lies not in replacing clinical evaluations but in demonstrating the feasibility of using imaging-derived biomarkers as objective and reproducible tools for SCI assessment. While distinguishing SCI from TD participants may not hold direct clinical significance, it establishes the capability of MRI to serve as a rich source of quantitative information. More critically, the ability to predict AIS categories with high accuracy addresses a pressing need for objective severity measures that complement or even augment traditional methods. By reducing reliance on subjective testing, these tools have the potential to standardize clinical workflows, improve inter-rater reliability, and enable consistent monitoring of SCI progression.

MATERIALS AND METHODS

Participants

Sixty-eight pediatric participants enrolled in this study between 2013 and 2017, comprising 41 TD with no history of spinal cord pathology or injury and 27 with SCI. Participants with SCI were recruited via an SCI registry and during routine visits to a pediatric hospital. TD participants were recruited via an institutional review board–approved pamphlet in the hospital. Inclusion criteria for participants with SCI required that they demonstrate an unchanged neurological examination and clinical status over the past 3 months (i.e., a stable neurological status) and be at least 12 months post-injury. All AIS impairment levels (AIS A-D) were enrolled. Seven SCI participants were excluded from the study due to issues related to metal distortion and poor MRI quality. The SCI participants had a mean age at the time of injury of 5.59 years (SD = 5.19, range [0.5–13.58]) and a mean age at scanning of 12.10 years (SD = 3.27, range [7.03–16.99]). The SCI group comprised 12 males and 8 females. In contrast, the 41 TD participants had a mean age at scanning of 11.71 years (SD = 3.06, range [6.37–16.78]) with a sex distribution of 17 males and 24 females. Statistical comparisons of age between the two groups showed no significant differences ($p = 0.66$). Recruitment of all participants was conducted using an Institutional Review Board (IRB)-approved brochure. All participants and their guardians provided informed assent and consent, respectively, in accordance with the IRB-approved protocol. The

Table 1: Demographic information for SCI participants. NA indicated that the measure was not available due to an incomplete ISNCSCI examination.

ID	Sex	Age at the Time of Injury (years)	Age at Time of Scanning (years)	Time From Injury (years)	AIS Grade	Neurological Level of Injury	Abnormal Signal
115	M	1.75	9.2	7.45	B	C6	-**
118	F	12.33	15.15	2.82	B	T10	T9-T12
142	F	10.75	13.96	3.21	C	L2	-
201	M	4.75	11.12	6.37	A	C6	C5-T5
202	F	12	15.13	3.13	NA	L2	-
206	F	3.67	8.07	4.4	D	C1	-
207	M	0.5	13.72	13.22	D	C5	-
208	M	2.08	8.08	6	A	C8	C6-T1
210	M	1.4	12.37	10.97	D	C1	-
211	F	1.58	12.60	11.02	NA	NA	T9-T11
212	M	8.4	12.68	4.28	A	T12	T10-T12
213	M	3.4	8.04	4.64	A	T7	T6-T8
214	M	0.67	8.66	7.99	NA	NA	-
215	M	13.58	14.96	1.38	A	C8	C6-T2
216	F	1.33	7.03	5.7	D	L2	T12-L1
217	M	15	16.99	1.99	B	C4	C4-C6
219	F	13.08	14.88	1.8	D	T4	T4-T6
220	M	2	16.94	14.94	D	L3	T7-T11
221	F	0.67	14.13	13.46	B	T3	C7-T5
222	M	2.91	8.15	5.24	NA	T4	T2-T4

* International Standards for Neurological Classification of Spinal Cord Injury

**Cases where no abnormal signal was detected, the spinal cord signal appeared normal, or the assessment was inconclusive due to artifacts or image distortion.

age range for all participants was 6–18 years. Demographic and injury information for SCI participants is summarized in Table 1. For some SCI participants, the AIS grade or level of injury is not provided in table due to incomplete ISNCSCI testing. This was due to testing limitations, including the patient’s inability to cooperate, non-testable dermatomes or myotomes due to a separate injury (e.g., brachial plexus injury), or unwillingness/inability to complete all parts of the examination (e.g., anorectal examination).

Data Acquisition

All participants underwent MRI scanning using a Siemens Verio (Erlangen, Germany) 3 Tesla MRI scanner. A 3D turbo spin echo (TSE) T2-weighted isotropic SPACE sequence was employed, with the following imaging parameters: a repetition time (TR) of 1500 ms, an echo time (TE) of 122 ms, a voxel size of $1 \times 1 \times 1 \text{ mm}^3$, a bandwidth of 751 Hz/px, a field of view (FOV) of $256 \times 256 \text{ mm}^2$, an acquisition matrix of $256 \times 256 \times 30$, and an acquisition time of 3 minutes and 21 seconds. Each subject underwent two acquisitions: one covering the cervical to upper thoracic cord, and another covering the upper to lower thoracic regions. This ensured coverage from at least the C1 through the T12-L1 disc, with overlap to facilitate the effective stitching of the two slabs. All MRI scans were reviewed by a board-certified pediatric neuroradiologist with over 30 years of experience, who assessed the presence of abnormal spinal cord signal changes. While many SCI cases exhibited clear abnormalities, some cases had no detectable abnormal signal or were deemed normal, while others were indeterminate due to artifacts or image distortion. In addition to undergoing imaging procedures, all participants diagnosed with SCI underwent a comprehensive evaluation using the ISNCSCI examination to determine their AIS category⁶. These findings are detailed in Table 1.

Preprocessing and Feature Extraction

The initial step in the image processing involved stitching two slabs using the scanner provided software to generate a comprehensive spinal cord image that covered vertebral levels C1 to T11 for all participants. This approach provided a continuous and detailed representation of the spinal cord necessary for accurate segmentation and subsequent analysis (Figure 1.a).

Following the stitching, the Spinal Cord Toolbox (SCT) was utilized for the segmentation of the spine and the labeling of vertebral levels for each subject¹⁷. This process effectively isolated the spinal cord from surrounding tissues and enabled precise identification and labeling of vertebral levels (Figure 1.b,c). Each segmented and labeled image then underwent a rigorous quality check performed by research fellows trained overseen by a pediatric board certified pediatric neuroradiologist. Images with metal distortion or those deemed to have insufficient quality were excluded from the study to ensure the reliability of the data. Any discrepancies identified during the quality check were manually corrected to maintain the integrity of the dataset.

For each axial slice of the spine, the CSA, RL width, and AP width were automatically extracted from this data set across all participants. The final measurements for each vertebral level from C1 to T11 were determined by averaging over the slices within each level (Figure 1.d).

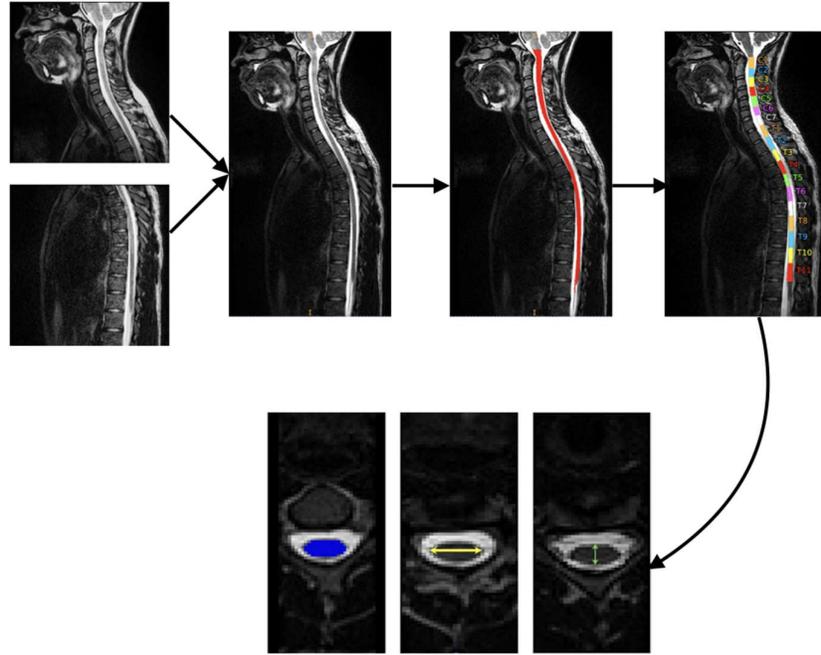


FIG 1. Workflow of the image processing steps. The initial stitching of two slabs to generate a continuous spinal cord image covering vertebral levels C1 to T11 (a). The segmentation (b) and labeling of vertebral levels (c) using the Spinal Cord Toolbox (SCT), isolating the spinal cord from surrounding tissues and accurately identifying vertebral levels. The measurement of cross-sectional area (CSA), right-left (RL) width, and anterior-posterior (AP) width for each vertebral level, averaged over the axial slices.

Statistical Analyses

To evaluate the structural differences between pediatric SCI and TD participants, we examined the distribution of each structural parameter (CSA, AP width, RL width) across all vertebral levels (C1 to T11) to determine whether the data followed a normal distribution. These parameters represent measurable morphological changes in spinal cord dimensions and were selected as quantitative indicators of spinal cord volume, potentially reflecting atrophy or denervation associated with SCI. Based on this assessment, we selected the appropriate statistical tests: independent t-tests were applied to vertebral levels where all structural parameters followed a normal distribution, while the Mann-Whitney U test was used for parameters that did not meet the normality assumption. Some vertebral levels exhibited normal distributions for all structural parameters and were analyzed using t-tests (c6-T3), whereas others had a mix of normally and non-normally distributed parameters, requiring a combination of t-tests and Mann-Whitney U tests accordingly. A significance level of $p < 0.05$ was adopted to identify statistically significant differences between the SCI and TD groups.

Additionally, Spearman's rank correlation coefficients were calculated to explore the relationship between each structural measure and the AIS categories. The AIS scale, while categorical, was treated as ordinal due to its ranked nature, making Spearman's correlation appropriate for assessing monotonic relationships. This non-parametric test was selected due to its effectiveness in analyzing non-normally distributed data and its ability to assess monotonic relationships. Correlations were evaluated across all vertebral levels to understand how structural changes correlate with the severity of SCI, as categorized by the AIS group. Results were deemed statistically significant if p-values were less than 0.05, indicating a meaningful difference or association between the parameters under investigation.

Model Architecture and Training

The model architecture designed for this study is a custom CNN developed to address the classification of SCI and TD participants, as well as to categorize the severity of the injury using the AIS grade (Figure 2). The model was initialized with random weights and was trained from scratch to learn relevant features from the input data. This architecture comprises two primary components: a CNN for feature extraction and a fully connected network for final classification. Efficiently capturing information from each spinal level, the CNN component considers both the above and below levels. This allows for the extraction of spatial features along the axial direction of the spinal cord, effectively encompassing the anatomical context surrounding each level. The model includes three convolutional layers, each with 32, 64, and 128 filters, respectively, with a filter size of 3. Each convolutional layer is followed by a ReLU activation function and a max-pooling layer with a pool size of 2 to reduce the spatial dimensions. The final convolutional layer's output is flattened and then combined with demographic features, height and age, before passing through two fully connected layers of 128 and 64 neurons, respectively. Additionally, dropout layers with a rate of 0.2 were added after each fully connected layer to reduce overfitting. The model is primed for two principal classification tasks: distinguishing between TD participants and those with SCI and categorizing injury severity according to the AIS classification.

The CNN model was implemented using Python 3.9 and PyTorch 1.12.1. Training and inference were conducted on a cloud-based

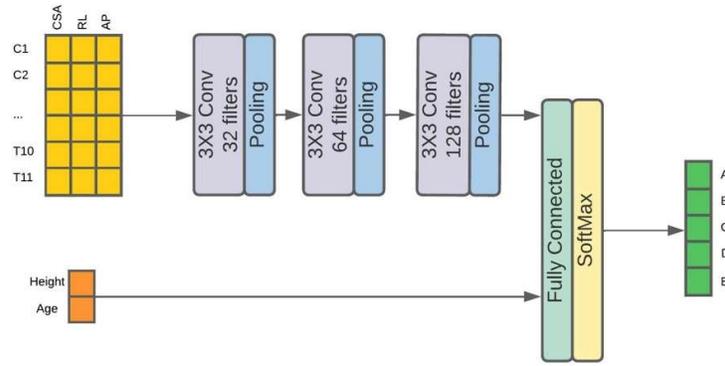


FIG 2. Architecture of the custom CNN developed for classifying pediatric SCI and assessing injury severity using the AIS. The network consists of three convolutional layers with ReLU activation functions, followed by max-pooling layers to progressively reduce spatial dimensions while retaining key features. The extracted features are then augmented with demographic information, including height and age, before being passed through fully connected layers. The final layer utilizes a softmax activation function to predict class probabilities for both SCI/TD classification and ASIA Impairment Scale categories.

high-performance computing environment equipped with an NVIDIA A100 GPU, an Intel CPU, and 16 GB of available RAM. The model was trained using a categorical cross-entropy loss function for both classification tasks, with the Adam optimizer set at an initial learning rate of 0.001. Training was conducted with a batch size of 32 over 100 epochs. A validation split of 20% was used to monitor the model’s performance, and early stopping was implemented with a patience of 5 epochs based on validation loss to prevent overfitting.

To ensure robust performance evaluation, 30 models were trained with different 80/20 splits of the dataset into training and testing sets. The training dataset was shuffled before each epoch to ensure a balanced representation of classes within each batch. The accuracies of these models were averaged to obtain a more reliable estimate of model performance. For each of the 30 trained models, we computed precision, recall, F1 score. These were averaged across all models to assess the overall performance of the CNN in distinguishing AIS categories.

In addition to the custom CNN, traditional machine learning models, such as random forest and SVM, were also employed for comparison. The random forest model was implemented using the scikit-learn library with a maximum tree depth of 10, and the SVM model used a radial basis function kernel. Both models were trained on the same feature set to benchmark the performance of the deep learning approach. The results of the traditional models were compared against the CNN’s performance to evaluate the relative effectiveness and robustness of the deep learning techniques.

RESULTS

Significant structural differences ($p < 0.05$) were observed between pediatric SCI and TD participants across all measured parameters—CSA, AP width, and RL width—at multiple vertebral levels from C1 to T11 (Figure 3). For CSA, SCI participants exhibited significantly reduced values compared to TD controls at all vertebral levels. Similarly, RL width showed significant differences at all vertebral levels. For AP width, significant reductions were observed at most vertebral levels, except for C2, C3, and T5, where the differences did not reach statistical significance (Figure 4).

Correlation analysis revealed moderate to strong relationships between structural measures and AIS categories (Figure 5). Particularly at levels C6 and C7, CSA showed moderate correlations with the AIS scale, with coefficients ranging from 0.37 to 0.51, suggesting an association with injury severity. The AP width also demonstrated moderate correlations, with values reaching up to 0.49, indicating its meaningful relationship with the severity of injury. The RL width showed the strongest correlations, with coefficients up to 0.56.

The custom CNN-based models demonstrated high accuracy in distinguishing between TD and SCI participants. The CNN achieved an accuracy of 96.59% on the test set (95% confidence interval (CI): 94.50% - 98.68%). Integrating structural parameters with deep learning models enabled accurate prediction of AIS categories. The model achieved an overall accuracy of 94.92% in predicting the AIS category of the test set (95% CI: 92.10% - 97.74%). In comparison, traditional models showed lower accuracy: the random forest model achieved 74.00% accuracy (95% CI: 69.00% - 79.00%), and the SVM model achieved 68.89% accuracy (95% CI: 63.00% - 74.00%). Table 2 presents the results for both the TD/SCI classification and the AIS category determination tasks across all models. The model exhibited strong precision, recall, and F1 scores across the AIS categories, with nearly perfect precision for AIS A and B categories (Table 3).

Table 2: Classification Accuracy of SVM*, Random Forest, and CNN models.

Task	SVM	Random forest	CNN
TD/SCI Classification	0.7635	0.7842	0.9559
AIS Classification	0.6889	0.7400	0.9492

* Support Vector Machine

Table 3: CNN model performance for classification of AIS grades.

Category	Precision	Recall	F1-Score
A	0.99	0.95	0.97
B	0.99	0.93	0.94
D	0.95	0.94	0.94
TD	0.88	0.98	0.93

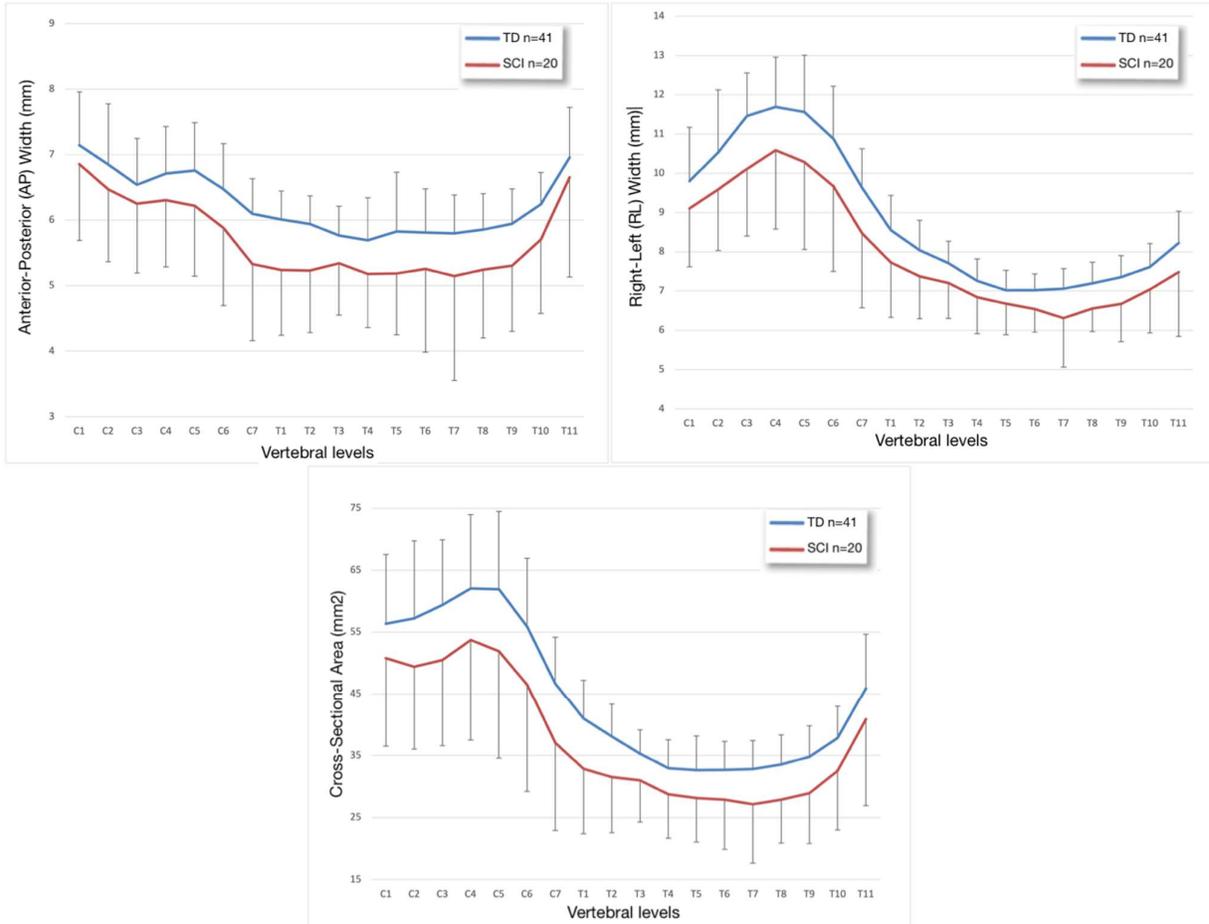


FIG 3. Cross-sectional measurements of cross-sectional area (CSA), right-left (RL) width, and anterior-posterior (AP) width averaged for all typically developing (TD) and spinal cord injury (SCI) participants across vertebral levels C1 to T11. The figure displays mean values with standard deviation bars, highlighting differences between TD and SCI groups at each vertebral level.

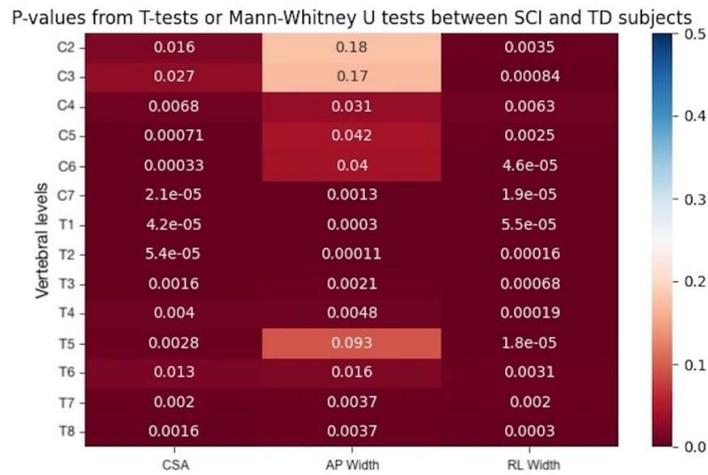


FIG 4. Structural differences between pediatric spinal cord injury (SCI) and typically developing (TD) participants across measured parameters—cross-sectional area (CSA), anterior-posterior (AP) width, and right-left (RL) width—at vertebral levels C1 to T11.

Correlation of Structural Measures with ASIA Impairment Scale

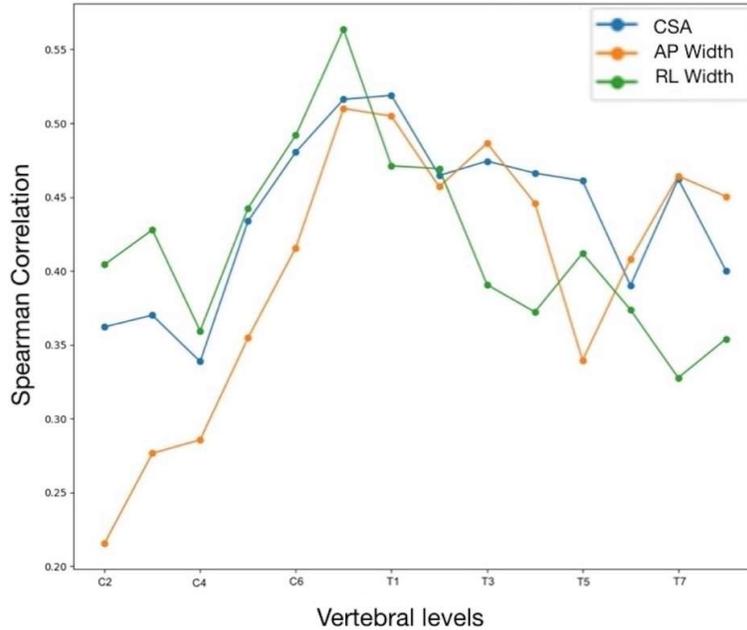


FIG 5. Correlation analysis between structural measures (CSA, AP width, RL width) and AIS categories, showing moderate to strong relationships.

DISCUSSION

This study demonstrates the potential of combining high-resolution MRI structural measurements with deep learning techniques for assessing chronic spinal cord injuries in pediatric populations. The results underscore the effectiveness of deep learning techniques in capturing complex patterns in high-dimensional MRI data and highlight significant differences in spinal cord structure between SCI and TD participants.

The significant differences observed in CSA, AP width, and RL width between SCI and TD participants corroborate previous studies that have noted structural alterations associated with SCI¹⁴. The reduction in CSA, AP width, and RL width in SCI participants is consistent across all vertebral levels from C1 to T11, emphasizing that structural changes are widespread and not limited to specific segments. These findings support the utility of cross-sectional measurements as biomarkers for assessing SCI and underline the importance of comprehensive spinal cord imaging in pediatric SCI evaluations.

The custom CNN demonstrated exceptional performance in distinguishing between SCI and TD participants, achieving an accuracy of 96.59%. This performance highlights the CNN's ability to effectively leverage spatial features along the spinal cord, incorporating information from adjacent levels to enhance classification accuracy. Moreover, the integration of structural parameters with deep learning models facilitated accurate prediction of AIS categories, with the CNN model achieving an overall accuracy of 94.92% in this task.

The robustness of the model is further highlighted by its high precision, recall, and F1 scores across the different AIS categories. The nearly perfect precision for AIS grade A and B categories suggests the CNN is particularly reliable in correctly identifying severe cases of spinal cord injury. Additionally, the high recall for TD participants indicates the model's proficiency in recognizing typically developing individuals. However, the slightly lower precision for the TD group suggests a minor tendency to classify other categories as TD, possibly due to overlapping structural features between less severe SCI cases and TD participants.

Traditional machine learning models, including random forest and SVM, were evaluated as benchmarks against the CNN. The random forest model achieved an average accuracy of 74.00%, and the SVM model achieved 68.89%. These results indicate that while traditional models provide useful insights, they do not match the performance of deep learning techniques in this context.

In our dataset, 35% of SCI participants had either no detectable abnormal signal or were indeterminate due to artifacts or image distortion, as assessed by a board-certified pediatric neuroradiologist. Despite this, our CNN-based model achieved high classification accuracy, demonstrating that MRI-derived structural parameters (CSA, AP width, and RL width) can effectively differentiate SCI from TD participants and predict AIS severity, even in cases where clear abnormal signal was not identified. This suggests that subtle structural changes, which may not always be visually apparent, still provide meaningful diagnostic information that deep learning models can leverage.

While distinguishing SCI from TD participants may not hold direct clinical significance, this classification serves as a critical first step, demonstrating that MRI contains rich, quantitative information that can be leveraged for automated analysis. The ability to predict AIS categories is particularly relevant, as it provides an objective alternative to current methods that depend heavily on subjective evaluations. This capability could reduce inter-clinician variability, standardize clinical workflows, and enable consistent monitoring of SCI progression. In complex or ambiguous cases, such as unresponsive patients or those with unclear trauma histories, imaging-based tools could serve as valuable adjuncts to clinical assessments. Moreover, as these models improve, they may enhance patient stratification for

treatment planning and serve as a foundation for developing personalized therapeutic strategies.

Acute pediatric SCI cases were not included in this study due to the challenges in recruiting these patients, as acute cases are less commonly available for research participation. As a result, this study focused on chronic SCI, where recruitment was more feasible, and structural changes had stabilized, allowing for reliable imaging and analysis. However, if acute-phase data were available, our current analysis pipeline would be applicable for segmentation and evaluation. Future studies incorporating acute-phase pediatric SCI imaging would provide valuable insights into early-stage structural changes and their potential role in predicting long-term outcomes.

In this study, we used 3D T2-weighted imaging to extract spinal cord structural parameters, ensuring consistent and high-resolution segmentation across vertebral levels. However, the same analysis applies to 2D T2-weighted images, as CSA, AP width, and RL width are calculated from axial slices and averaged within each vertebral level. This makes 2D imaging a viable option for structural assessment if slice positioning and spacing are carefully controlled. The main limitation of 2D imaging is its sensitivity to variations in slice positioning within a vertebral level. While 3D imaging provides isotropic resolution for more precise averaging, 2D acquisitions depend on pre-defined slice thickness and spacing, which could introduce variability if not standardized. Despite this, 2D sequences can still provide meaningful structural insights if slices are carefully selected for consistent anatomical coverage. Additionally, while this study did not include metal-suppression techniques, surgical hardware in some SCI cases may introduce susceptibility artifacts that affect image quality. Future studies could explore advanced MRI acquisition methods including view-angle-tilting (VAT) and slice encoding for metal artifact correction (SEMAC) to mitigate metal-induced distortions, improving spinal cord visualization in post-surgical SCI cases and enhancing the accuracy and reliability of structural analyses. The future application of the deep learning approach used in this study to 2D and VAT/SEMAC sequences has potential to improve the generalizability of these results and allow for utilization in a broader variety of imaging scenarios.

Despite the promising results, there are limitations to this study. The dataset, while substantial for this population, is limited to pediatric chronic SCI, and the models' performance may vary with different datasets or clinical settings. Future research should focus on validating these findings with larger, diverse datasets and exploring longitudinal analyses to track changes over time. Additionally, incorporating other imaging modalities and advanced techniques, such as transfer learning and multi-modal data integration, could further enhance model performance and clinical applicability. Because this study used imaging at the chronic stage of SCI, these results cannot be used directly for prognostic purposes, but the high accuracy is encouraging for future applications with acute imaging. To this point, longitudinal studies spanning from the acute to chronic stages are essential to comprehensively understand SCI progression and to develop predictive models capable of assessing both current severity and future functional outcomes based on structural changes over time.

CONCLUSIONS

This study highlights the significant potential of integrating high-resolution MRI structural measures with deep learning techniques for the classification of pediatric SCI. By quantitatively analyzing structural parameters such as CSA, AP width, and RL width, we observed significant differences between TD participants and those with SCI. These structural alterations were effectively captured and analyzed using convolutional neural network (CNN) models, achieving high accuracy in classifying TD and SCI cases, and classification of AIS categories. The CNN models demonstrated an accuracy of 96.59% in distinguishing between TD and SCI, and an accuracy of 94.92% in predicting the AIS categories. The ability to accurately classify SCI and AIS categories can significantly enhance diagnostic and prognostic capabilities, leading to more personalized and effective treatment strategies for pediatric SCI patients.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Neurological Disorders and Stroke (NINDS) of the National Institutes of Health (NIH) under award numbers R01 NS079635 (awarded to Drs. Mohamed and Mulcahey) and R01 NS111113-01A1 (awarded to Drs. Mohamed and Krisa).

REFERENCES

1. Cunha NS, Malvea A, Sadat S, et al. Pediatric spinal cord injury: a review. *Children* 2023;10:1456
2. DeVivo MJ, Vogel LC. Epidemiology of spinal cord injury in children and adolescents. *The Journal of Spinal Cord Medicine* 2004;27:S4-S10
3. Anjum A, Yazid MDi, Fauzi Daud M, et al. Spinal Cord Injury: Pathophysiology, Multimolecular Interactions, and Underlying Recovery Mechanisms. *International Journal of Molecular Sciences* 2020;21:7533
4. Yousefifard M, Hosseini M, Oraii A, et al. Severity assessment of impairment in spinal cord injury; a systematic review on challenging points about International Standards for Neurological Classification of Spinal Cord Injury. *Frontiers in Emergency Medicine* 2021
5. Kirshblum SC, Burns SP, Biering-Sorensen F, et al. International standards for neurological classification of spinal cord injury (revised 2011). *The journal of spinal cord medicine* 2011;34:535-546
6. Mulcahey MJ, Gaughan J, Betz RR, et al. The International Standards for Neurological Classification of Spinal Cord Injury: reliability of data when applied to children and youths. *Spinal Cord* 2007;45:452-459
7. Mulcahey MJ, Gaughan JP, Chafetz RS, et al. Interrater reliability of the international standards for neurological classification of spinal cord injury in youths with chronic spinal cord injury. *Archives of physical medicine and rehabilitation* 2011;92:1264-1269
8. Faro SH, Saksena S, Krisa L, et al. DTI of chronic spinal cord injury in children without MRI abnormalities (SCIWOMR) and with pathology on MRI and comparison to severity of motor impairment. *Spinal cord* 2022;60:457-464
9. Tuszynski MH, Gabriel K, Gerhardt K, et al. Human spinal cord retains substantial structural mass in chronic stages after injury. *Journal of neurotrauma* 1999;16:523-531
10. Curati W, Kingsley D, Kendall B, et al. MRI in chronic spinal cord trauma. *Neuroradiology* 1992;35:30-35
11. Trolle C, Goldberg E, Linnman C. Spinal cord atrophy after spinal cord injury—A systematic review and meta-analysis. *NeuroImage: Clinical* 2023;38:103372
12. Seif M, Curt A, Thompson AJ, et al. Quantitative MRI of rostral spinal cord and brain regions is predictive of functional recovery in acute spinal cord injury. *NeuroImage: Clinical* 2018;20:556-563

13. Lundell H, Barthelemy D, Skimminge A, et al. Independent spinal cord atrophy measures correlate to motor and sensory deficits in individuals with spinal cord injury. *Spinal Cord* 2011;49:70-75
14. Qi Q, Wang L, Yang B, et al. The relationship between the structural changes in the cervical spinal cord and sensorimotor function of children with thoracolumbar spinal cord injury (TLSCI). *Spinal Cord* 2024:1-7
15. Muller JJ, Wang R, Middleton D, et al. Machine learning-based classification of chronic traumatic brain injury using hybrid diffusion imaging. *Frontiers in Neuroscience* 2023;17:1182509
16. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik* 2019;29:102-127
17. De Leener B, Lévy S, Dupont SM, et al. SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data. *Neuroimage* 2017;145:24-43