

ARTICLE

A hybrid 3D CNN–Transformer architecture for high-precision automated diagnosis of spinal disorders from magnetic resonance imaging volumes

Janardhanarao Addanki* and Bala Krishnasamy

Department of Computer Science and Engineering, School of Computing, Bharath Institute of Science and Technology (BIST), Chennai, Tamil Nadu, India

Abstract

The automated and precise interpretation of volumetric magnetic resonance imaging (MRI) for spinal cord anomalies is crucial for timely intervention and patient care planning. Deep learning in MRI volumetry is hampered by the need to disentangle mixed-resolution volumetric and MRI-slice data. To mitigate these challenges, this study introduces an original approach for automated multi-label classification of spinal MRI abnormalities. The proposed dual-scale spatiotemporal network (DS-STNet) processes volumetric 3D MRI data directly to achieve precise, automatic classification of spinal abnormalities. It combines dual-scale 3D convolutional feature extraction with a global attention mechanism implemented via a Transformer network, capturing both local anatomical features and inter-slice relationships. The model addresses MRI volumetric data in their raw states without computerized segmentation and feature extraction. DS-STNet was evaluated on MRI volumetric data from the Kaggle datasets (RSNA LumbarDISC and SpineWeb), where spinal abnormalities were classified into five categories. Experimental results showed an overall classification accuracy of 98.61%, with high precision, recall, F1 score, and area under the curve. State-of-the-art deep learning competitors, such as 3D convolutional neural networks (CNNs), 3D ResNets, Vision Transformers, and CNN–long short-term memory hybrids, were compared with DS-STNet. The proposed DS-STNet significantly outperformed these existing deep learning models. The findings confirm the effectiveness of incorporating dual-scale spatiotemporal feature learning with Transformer global attention for dependable identification of spinal abnormalities. The proposed DS-STNet offers a scalable, clinically feasible, and high-accuracy solution for automating MRI spinal diagnostics.

*Corresponding author:

Janardhanarao Addanki
(janardhan.addanki@gmail.com)

Citation: Addanki J, Krishnasamy B. A hybrid 3D CNN–Transformer architecture for high-precision automated diagnosis of spinal disorders from magnetic resonance imaging volumes. *Int J Systematic Innovation*, 10(5): 026240153. [https://doi.org/10.6977/IJoSI.202610_10\(5\).0016](https://doi.org/10.6977/IJoSI.202610_10(5).0016)

Received: June 12, 2026**Revised:** July 26, 2026**Accepted:** August 13, 2026**Published online:** September 18, 2026**Copyright:** © 2026 Author(s).

This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Keywords: Spinal magnetic resonance imaging classification; Hybrid 3D convolutional neural network–Transformer; Dual-scale spatiotemporal network (DS-STNet); Spatiotemporal feature learning; Medical image analysis

1. Introduction

Issues with the spinal cord and vertebral column primarily affect an individual's mobility, cognitive function, and overall quality of life. Cognitive issues and abnormalities

of the spinal cord and its surrounding structures are interconnected and, if left unaddressed, may contribute to a cascade of negative consequences, including progressive cognitive dysfunction. Neurological disorders, including multiple sclerosis, are associated with spinal cord lesions and gray matter loss in the brain, highlighting the need for precise and accurate spinal imaging and analysis during the early stages of disease (Rocca *et al.*, 2023). Spinal disorders in children and young people can be particularly challenging to diagnose because of developmental changes and subtle alterations in spinal morphology, necessitating sensitive imaging techniques (Mezzapesa *et al.*, 2004; Suleman *et al.*, 2022). In addition, cervical spinal cord atrophy has been associated with mobility impairment, further underscoring the clinical relevance of accurate spinal imaging (Valsasina, Aboulwafa, *et al.*, 2018).

The spinal canal houses the spinal cord, a delicate neural structure that is susceptible to injury and disease. Neurological disorders may manifest as progressive atrophy of the cervical and lumbar spinal cord segments. Such atrophy has been studied and measured through voxel-based morphometry and spatial normalization techniques on high-resolution 3D magnetic resonance imaging (MRI) (Valsasina, Horsfield, *et al.*, 2012). Because MRI provides superior soft-tissue contrast resolution without exposure to ionizing radiation, it is the modality of choice for evaluating spinal abnormalities (Thompson *et al.*, 2018). Advances in imaging techniques have enabled neurologists and radiologists to detect lesions and subtle pathological changes. However, the time-consuming nature of these tasks and variability in interpretation limit the practicality and reliability of manual analysis.

To accurately identify spinal abnormalities on MRI, spinal cord tissues are subjected to careful segmentation and detailed feature analysis. Although traditional strategies based on manual and semi-automated tracing can provide reliable segmentation, they are inefficient and impractical for large-scale clinical datasets. To address these limitations, fully automated deep learning-based segmentation approaches have emerged. Cascaded 3D convolutional neural networks (CNNs) are a recent advancement in the automated delineation of spinal lesions and have demonstrated improved segmentation accuracy compared with traditional methods (Valverde *et al.*, 2017). Similarly, semi-automated methods have been developed for rapid spinal cord segmentation from MRI, offering clinical benefits while maintaining high segmentation accuracy (Horsfield *et al.*, 2010).

Analysis of spinal cord structures has also enabled the introduction of analytical cross-sectional area measures, such as the mean upper cervical cord area, which has shown

considerable reproducibility and robustness to lesions. However, variability in the scanner and software remains, which can limit measurement consistency and the broader implementation of automated approaches (Weeda *et al.*, 2019). In the pediatric population, spinal cord analysis is particularly challenging because of heterogeneity in growth patterns and lesion distribution. MRI studies performed in children and adolescents have shown that spinal cord lesions are not only common but also provide diagnostic value during the early stages of disease (Debska *et al.*, 2017; Verhey *et al.*, 2010).

Techniques focused on morphometric analyses of the upper cervical spinal cord cross-sectional area have demonstrated value for monitoring disease progression in relapsing-remitting inflammatory disorders (Mann *et al.*, 2007). These measures are increasingly being used for prognostication, particularly in children, for whom early recognition is critical for improving long-term functional outcomes (Margoni, Preziosa, *et al.*, 2022). Recent neuroimaging studies have also demonstrated the role of microstructural changes in the subventricular zone in the pathophysiology of pediatric-onset disorders, highlighting the importance of assessing brain subregions using high-resolution volumetric imaging (Margoni, Storelli, *et al.*, 2025). In addition to neurodegenerative disorders, spinal injuries and age-related spinal degeneration are increasingly relevant to the growing global aging population. Evidence-based clinical practice guidelines emphasize the importance of early diagnosis, continuous rehabilitation, and efficient imaging-based monitoring to maintain functionality and mitigate the progression of neurological deficits in elderly individuals with spinal cord injury (Seijas *et al.*, 2024). Despite these advances, fully automated and accurate diagnostic classification of volumetric spinal MRI data remains challenging, particularly for complex tasks. These challenges include complex spinal anatomy, variable tissue contrast, partial-volume effects, and contextual dependencies between adjacent slices.

While 3D CNNs can effectively capture local spatial features, they may have limited ability to model long-range dependencies between slices. Transformers can address this limitation through self-attention; however, they generally lack strong local feature inductive biases. As a result, hybrid CNN–Transformer models combine the complementary strengths of both architectures. Because MRI data are volumetric and contain long-range spatial relationships across slices, developing models that effectively balance local feature extraction with global contextual modeling is particularly relevant for accurate spinal image analysis. Such hybrid architectures may

therefore improve the identification of subtle anatomical features while preserving broader structural relationships within volumetric MRI data.

To meet these requirements, we propose a new model, the dual-scale spatiotemporal network (DS-STNet), which combines dual-scale 3D convolutional feature extraction with Transformer-based attention mechanisms to enable robust multi-class classification of spinal abnormalities from volumetric MRI data. The model was trained directly on unprocessed 3D MRI images, without the need for manual segmentation or handcrafted features, and was evaluated on the Radiological Society of North America (RSNA) Lumbar Degenerative Imaging Spine Classification (LumbarDISC) and SpineWeb spinal MRI datasets, which include studies representing five diagnostic categories: normal spine, disc degeneration, spinal stenosis, spondylolisthesis, and herniated disc. DS-STNet demonstrated strong classification performance across multiple evaluation metrics, including precision, recall, F1 score, and area under the curve (AUC), highlighting its potential for automated classification of spinal abnormalities from volumetric MRI data.

The main contributions of this study are as follows. A novel model (DS-STNet) is proposed, integrating dual-scale 3D convolutional feature extraction with Transformer-based global attention for volumetric spinal MRI analysis. It performs end-to-end multi-class classification directly from raw 3D MRI volumes without requiring handcrafted features or explicit image segmentation. Furthermore, an attention-guided refinement module with class-weighted optimization is incorporated to improve the discrimination of visually similar spinal disorders. Extensive experimental evaluation demonstrated that the proposed DS-STNet consistently outperformed conventional CNN-, ResNet-, Vision Transformer-, and hybrid deep learning-based approaches, highlighting its potential for accurate and reliable automated classification of spinal abnormalities from MRI data.

The rest of this paper is organized as follows. Section 2 discusses previous studies on automated analysis and classification of spinal MRI abnormalities, including both traditional and deep learning techniques. Section 3 describes the proposed DS-STNet framework, including the dataset, preprocessing procedures, dual-scale 3D convolutional architecture, Transformer module, attention-guided feature refinement, and model training strategy. Section 4 presents comprehensive experimental evaluations, including quantitative performance assessments, confusion matrix analyses, comparative analyses with benchmark models, model ablation studies, performance evaluations across multiple training and

testing runs, and analyses of model complexity. Section 5 concludes with a discussion of the experimental findings.

2. Related work

Magnetic resonance imaging provides excellent soft-tissue contrast and does not use ionizing radiation, making it well-suited for visualizing neural structures. Its clinical applications include the diagnosis, tracking, and prognosis of a wide range of spinal cord disorders. Early investigations employed MRI to assess spinal cord involvement in multiple sclerosis, spinal cord injuries, and congenital spinal cord defects. Verhey *et al.* (2010) examined MRI features of the spinal cord in children with multiple sclerosis and noted the importance of early identification of lesions and spinal anatomical alterations in determining disease prognosis. Mann *et al.* (2007) demonstrated the potential for improved disease monitoring and prognostication through qualitative and quantitative MRI analyses of the upper cervical spinal cord cross-sectional area, including several measurable metrics for tracking disease progression. In early multiple sclerosis, Margoni, Preziosa, *et al.* (2022) highlighted the role of advanced neuroimaging techniques, while Margoni, Storelli, *et al.* (2025) used high-resolution MRI to characterize microstructural features of the brain and spinal cord and further investigate their relationship. Collectively, these studies demonstrate the value of MRI for assessing the spinal cord and provide a foundation for developing automated approaches to the analysis of complex spinal disorders.

Other research has examined rehabilitation from spinal injuries and the mechanisms of functional recovery. Seijas *et al.* (2024) provided evidence-based recommendations for rehabilitation of older adults with spinal cord injury, emphasizing the importance of early diagnosis and appropriate clinical management. Shinozaki *et al.* (2021) studied the mechanisms of stem cell therapy for spinal cord injury and highlighted the importance of characterizing the extent and location of neural tissue damage when evaluating efficacy. Chalif *et al.* (2024) analyzed functional recovery following epidural spinal cord stimulation and, in their review of techniques that aid functional recovery, provided evidence of the value of advanced diagnostic imaging in neuromodulation therapies. These findings highlight the clinical need for accurate, automated spinal MRI assessment systems. Hernandez-Navarro *et al.* (2025) further analyzed non-invasive cerebral and spinal stimulation techniques for improving motor and gait recovery in patients with incomplete spinal cord injury, underscoring the importance of accurate assessment in the evaluation and application of these interventions.

There has also been extensive research on pediatric

spinal trauma and spinal cord injury without radiographic abnormality (SCIWORA) using MRI. Freigang *et al.* (2022) reported midterm outcomes of true SCIWORA on MRI and noted the ability of MRI to reveal subtle spinal cord alterations that may not be identified by traditional radiographic methods. Wang *et al.* (2022) investigated acute hyperextension myelopathy in children and identified MRI-based radiographic predictors associated with clinical improvement, emphasizing the potential value of volumetric spinal imaging for prognostic assessments. Kirshblum *et al.* (2020) provided international standards for the neurological classification of spinal cord injury that heavily incorporate MRI-based structural assessments. Liang *et al.* (2022) reviewed the risk factors and the prognosis of SCIWORA in children and reconfirmed that timely MRI-based diagnosis significantly shapes long-term neurological outcomes.

From an embryological and congenital perspective, Sergeenko *et al.* (2024) examined the relationship between abnormal spinal cord development and vertebral anomalies, providing insights into developmental abnormalities detectable on MRI. Amelot *et al.* (2020) described spinal dysraphism as a clinical entity of V.A.C.T.E.R.L. (vertebral, anal, cardiac, tracheoesophageal, renal, and limb anomalies) syndrome, further confirming the utility of MRI for evaluating complex congenital spinal anomalies. These anatomical and clinical investigations have advanced understanding of spinal pathologies across diverse and complex aspects of anatomy, physiology, and embryology, underscoring the need for sophisticated computational approaches to support automated analysis and interpretation of spinal imaging data.

As specialization in medical imaging artificial intelligence expands, deep learning has emerged as the first and only viable tool for automating MRI analysis. Geetha *et al.* (2025) proposed a hybrid deep learning framework for automated spinal tumor diagnosis from MRI scans. Although tumor diagnosis was only a specific application in their work, their findings demonstrate the potential of hybrid deep learning approaches for automated analysis of spinal pathologies using MRI. The complexity of spinal pathology, however, underscores the need for robust models capable of capturing diverse and subtle abnormalities in MRI data.

Earlier imaging studies represented important advancements in spinal assessment but remained largely dependent on manual inspection, conventional statistical analyses, and basic morphometric measurements. While these approaches were scientifically useful, they were subject to observer variability, had limited scalability, and lacked sufficient sensitivity for detecting subtle features

across multiple classes of degenerative disorders. More recent hybrid deep learning approaches have enabled greater automation; however, many existing models remain 2D, pathology-specific, or constrained to single-task diagnosis. In addition, a number of existing systems struggle to simultaneously preserve fine-grained local anatomical features while capturing long-range contextual information and interslice dependency in volumetric MRI data. In direct response to these limitations, volumetric MRI has motivated the use of 3D CNNs, which can learn spatial features across depth, height, and width. While 3D CNNs improve volumetric representation, their receptive fields may remain limited, restricting their ability to capture long-range dependencies across distant slices and spinal levels. Transformer-based architectures can address this limitation by using self-attention to model long-range inter-slice relationships and global contextual dependencies across volumetric MRI data. However, Transformers generally lack the strong locality-inductive bias of CNNs, which is particularly important for medical image analysis involving fine-grained anatomical structures.

The aforementioned limitations motivate the development of the hybrid CNN–Transformer architecture, in which CNNs extract local features while Transformers capture and integrate global dependencies. This architecture is well suited to various medical image analysis tasks because it combines fine-grained spatial information with broader contextual information. In spinal MRI classification, however, the application of such hybrid architectures to fully automated multi-class classification of volumetric 3D MRI data remains limited, particularly in settings that do not require explicit image segmentation or handcrafted features.

While some hybrid CNN–Transformer models have shown promising performance in medical image analysis, they are often limited to disease-specific classification or segmentation tasks. Moreover, most of them do not jointly integrate dual-scale volumetric feature extraction, Transformer-based dependency modeling, attention-guided feature refinement, and explainability within a single framework for multi-class spinal MRI diagnosis. Furthermore, clinical interpretability and comprehensive evaluation of these models remain underexplored. In this regard, the study proposes an end-to-end hybrid 3D CNN–Transformer model, DS-STNet, for multi-class spinal disorder classification. The proposed model aims to provide accurate, explainable, and automated analysis of volumetric spinal MRI data.

The proposed DS-STNet was compared with recent CNN–Transformer-based methods published from 2024 to 2026 (Table 1). The existing solutions mainly rely on

Table 1. Novelty comparison of the proposed DS-STNet with recent hybrid CNN–Transformer models (2024–2026)

Model	Year	Application	Limitation	Advantage
SymTC (Chen <i>et al.</i> , 2024)	2024	Lumbar spine MRI segmentation	Segmentation only; no disease classification	End-to-end multi-class spinal disorder classification from volumetric MRI
ResMT (Cui <i>et al.</i> , 2024)	2024	Brain glioma grading	Limited to brain MRI	Designed specifically for spinal MRI with dual-scale feature learning
Hybrid CNN–Transformer (Geetha <i>et al.</i> , 2025)	2025	Spinal tumor classification	Single disease classification	Supports multiple spinal disorders within a unified framework
CAFR-Net (Ma <i>et al.</i> , 2025)	2025	Spinal MRI segmentation	Segmentation-focused; limited interpretability	Integrates classification, attention refinement, and Grad-CAM explainability
Proposed DS-STNet	2026	Multi-class spinal MRI diagnosis	–	Dual-scale 3D CNN, Transformer-based global attention, attention refinement, and explainable end-to-end classification

Abbreviations: CNN: Convolutional neural network; Grad-CAM: Gradient-weighted class activation mapping; MRI: Magnetic resonance imaging.

segmentation or classification for a single disease and fail to account for dual-scale feature extraction, global dependency learning via Transformers, attention-guided refinement, and explainability, all within a single framework. The proposed DS-STNet combines these complementary components into an end-to-end architecture for multi-class automated spinal MRI classification, leading to better diagnostic performance and improved clinical interpretability.

3. Proposed methodology

The proposed methodology is an automated multi-class classification framework for spinal cord abnormalities from volumetric MRI, based on dual-scale spatiotemporal feature learning using a hybrid 3D CNN–Transformer framework. Unlike traditional approaches that depend on handcrafted features or explicit image segmentation, the proposed approach performs end-to-end learning, directly mapping raw 3D MRI volumes to diagnostic class labels. The dual-scale convolutional design enables efficient extraction of local anatomical features at different spatial scales, while the Transformer module captures global contextual information and long-range interslice dependencies. The 3D CNNs process volumetric MRI data across depth, height, and width, while the Transformer component models dependencies across the entire volume. Hence, the proposed framework is designed to distinguish visually similar spinal disorders that may be challenging for conventional CNN-based classifiers.

Figure 1 shows the proposed architecture of the DS-STNet system for automated spinal abnormality

classification from volumetric MRI data. It comprises a dataset of spinal MRI scans, which is fed to a data stream manager; the resulting data streams are then presented to a segmentation module for the classification task. First, the input MRI volume is preprocessed with z-score normalization, volumetric resampling, and 3D data augmentation to improve image quality and increase data diversity. The enhanced volume is then fed into a dual-scale 3D CNN, comprising a local branch that learns fine-grained anatomical structures and a global branch that captures the broader spatial context. Feature maps are extracted from both branches and combined channel-wise to form a complete volumetric representation via feature fusion. These fused features are then converted to patch tokens and fed into a Transformer encoder, which uses multi-head self-attention to learn long-range spatial relationships and the global context. A weighted cross-entropy loss is used during training to compensate for class imbalance and optimize feature representation. The learned representations are further refined with a channel-spatial attention module that highlights the discriminative spinal regions while filtering out irrelevant background information. Finally, the features are aggregated to a global average and fed into a fully connected layer, followed by a Softmax classifier, to obtain the final prediction of the spinal abnormality.

3.1. Dataset description

The proposed DS-STNet framework was tested on spinal MRI datasets from two public sources: RSNA LumbarDISC and SpineWeb. The combined dataset comprised 1,850 patients and 3,700 MRI examinations,

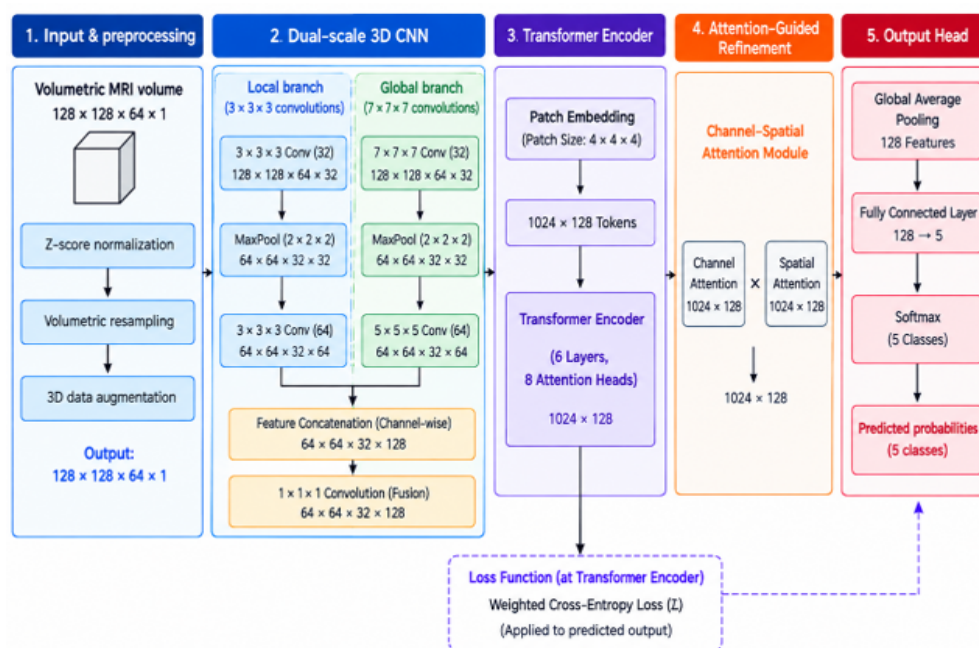


Figure 1. Proposed dual-scale spatiotemporal network (DS-STNet) hybrid 3D convolutional neural network (CNN)–Transformer framework for spinal abnormality classification.

Abbreviation: MRI: Magnetic resonance imaging.

including T1-weighted, T2-weighted, and short-tau inversion recovery (STIR) sequences, acquired on 1.5 T and 3.0 T MRI scanners. The public datasets were obtained from several clinical sites, and the manufacturer of these scanners and detailed acquisition parameters were not always provided for all examinations. The data consist of five diagnostic classes: normal, disc degeneration, herniated disc, spinal stenosis, and spondylolisthesis. The patient split was 70:15:15 for train:validation:test. All MRI volumes were preprocessed before training to ensure consistent image quality: N4 bias field correction, intensity normalization, contrast-limited adaptive histogram equalization (CLAHE) enhancement, non-local means filtering, and resizing to $128 \times 128 \times 64$.

The distribution of patients and MRI volumes across the five diagnostic classes is summarized in Table 2. Despite the moderate class imbalance, this work addressed it by using class-weighted cross-entropy loss and 3D data augmentation to enhance classification performance for the underrepresented spinal disorders.

The process began with creating a volumetric input. Each MRI scan was transformed into a 3D tensor that captures the spatial and contextual data of adjacent slices. To achieve stability across scanners and patients, the volumetric input was z-score-normalized to standardize voxel intensities, thereby improving convergence and

reducing domain shifts. This is represented mathematically as:

Table 2. Class distribution of the combined spinal magnetic resonance imaging (MRI) dataset

Class	Number of patients	Number of MRI volumes
Normal	380	760
Disc degeneration	500	1,000
Herniated disc	400	800
Spinal stenosis	330	660
Spondylolisthesis	240	480
Total	1,850	3,700

$$I_{norm} = \frac{I - \mu}{\sigma} \quad (1)$$

where I represents the initial voxel intensity, and μ and σ are the averages and standard deviations of the intensities, respectively. After the data volume was normalized, the first step in the processing was to resample the data to ensure all input data had the same spatial resolution. This ensured that the hybrid network did not receive inputs of different sizes and thus could compute the embedding representation without geometric distortions. The next processing step

after volume resampling was data augmentation, in which the volume was augmented across three spatial dimensions using rotation, elastic deformation, and random cropping to enhance network generalization, simulate anatomical variability, and alleviate the class imbalance often associated with medical datasets.

The next step after data preparation was to extract features using a 3D CNN with a dual-scale design. This is the first stage of the proposed framework. The dual-scale approach extracts features at different scales, capturing fine features such as the boundaries of discs and nerve roots, and coarse features that are more global or holistic, such as the alignment of the vertebrae and the overall curvature of the spine. In the local-scale branch of the model, smaller convolution kernels are used to extract and preserve fine, detailed spatial information, and in the global-scale branch, larger kernels are used to capture feature context and compose fine details with that context. The standard method of conducting a convolution in 3D was used for feature extraction in each convolutional layer and is defined as:

$$F_l = \sigma(W_l * F_{l-1} + b_l) \quad (2)$$

where F_{l-1} and F_l denote the input and output feature maps at layer l ; W_l and b_l represent the learnable kernel weights and bias terms, respectively; $*$ indicates 3D convolution; and $\sigma(\cdot)$ denotes the nonlinear activation function. The extracted features from both branches were fused via channel-wise concatenation, followed by pointwise convolution for feature compression. This fusion enables complementary integration of micro-anatomical texture cues and macro-structural shape information, forming a unified dual-scale spatiotemporal representation of the spinal volume.

To mitigate the limitation of convolutional networks in modeling long-range dependencies, the fused dual-scale features were then passed to a transformer-based global context modeling module. The current processed 3D feature volume was presented to the module as a sequence of non-overlapping 3D patches, which were then flattened and their token embeddings projected. These tokens were then provided as input to the transformer encoder, allowing the network to examine the interplay among distant anatomical regions across the entire spinal structure. The main component of this module is multi-head self-attention, which computes joint relations among all patch tokens, regardless of their relative positions in space. The self-attention operation is expressed as follows.

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V} \quad (3)$$

where $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$ denote the query, key, and value matrices, respectively, derived from the token embeddings, and d is the embedding dimensionality. This operation enables the network to selectively focus on diagnostically significant spinal regions while suppressing irrelevant background responses. Through stacked Transformer encoder layers, the model progressively refined its understanding of inter-slice dependencies, anatomical symmetry, and pathological spread patterns.

Attention-enhanced features were submitted to the attention-enhanced global context modeling technique for the next stage, which adaptively calibrates spatial and channel-wise responses through attention-guided refinement. This refinement process enhances the channel and feature responses to abnormal activations and reduces the noise and redundancy. The representation of refined features was globally pooled and then passed to the fully connected output layers of the classifier. The Softmax function is given by:

$$\hat{y} = \text{Softmax}(\mathbf{W}_c \cdot \mathbf{F}_{\text{refined}} + \mathbf{b}_c) \quad (4)$$

where W_c and b_c denote the classifier weights and bias parameters, respectively, F_{refined} represents the refined global feature descriptor, and $\hat{Y}$ corresponds to the predicted probability distribution over all spinal abnormality classes.

To ensure stable optimization under class imbalance and overlapping pathological signatures, the proposed framework employs a class-aware weighted cross-entropy loss, which assigns a higher penalty to underrepresented abnormality categories while preserving discrimination among dominant classes. The loss function is defined as:

$$\mathcal{L} = -\sum_{c=1}^C w_c y_c \log(\hat{y}_c) \quad (5)$$

where C is the number of spinal abnormality classes, y_c is the ground-truth label, $\hat{Y}_c$ is the predicted probability, and w_c denotes the class-specific weighting factor. This formulation ensures that the learning process remains balanced and sensitive to rare but clinically significant spinal disorders.

In summary, the proposed hybrid model is a unified deep architecture combining dual-scale spatiotemporal feature extraction, Transformer-based global reasoning, and attention-based feature refinement. The model's fully automated and scalable classification of spinal cord abnormalities is free from reliance on handcrafted heuristics and manual segmentation. Volumetric learning, combined with long-range dependencies and class-specific model optimization, drastically improves the model's overall dependability and generalization, especially across varying MRI protocols. The system is an intelligent,

dependable, clinically applicable decision support system.

Algorithm 1. DS-STNet training for spinal abnormality classification

```

Input: 3D MRI volumes  $V$  with labels  $Y$ 
Output: Trained DS-STNet model
1.   Initialize Dual-Scale 3D CNN, Transformer, Attention
    Module, and Classifier
2.   for each epoch do
3.     for each batch ( $V_b$ ,  $Y_b$ ) do
4.       Normalize  $V_b$  using z-score
5.       Apply 3D data augmentation
6.        $F_{\text{local}} \leftarrow \text{Local 3D CNN}(V_b)$ 
7.        $F_{\text{global}} \leftarrow \text{Global 3D CNN}(V_b)$ 
8.        $F_{\text{fused}} \leftarrow \text{Concat}(F_{\text{local}}, F_{\text{global}})$ 
9.       Tokens  $\leftarrow \text{PatchEmbedding}(F_{\text{fused}})$ 
10.       $F_{\text{ctx}} \leftarrow \text{Transformer}(\text{Tokens})$ 
11.       $F_{\text{ref}} \leftarrow \text{AttentionRefinement}(F_{\text{ctx}})$ 
12.      Pred  $\leftarrow \text{Softmax}(\text{Classifier}(F_{\text{ref}}))$ 
13.      Compute weighted cross-entropy loss
14.      Update network weights using AdamW
15.    end for
16.  end for
17.  Return trained model

```

4. Results and discussion

The spinal cord MRI datasets are a collection of volumetric MRI scans of the spinal column annotated by medical professionals and acquired from a variety of clinical settings. This collection is available on RSNA LumbarDISC and SpineWeb for public download. The design and functionality of the proposed DS-STNet hybrid 3D CNN-Transformer model were evaluated on a combined dataset, with the collected MRI volumetric scans serving as the training set. The dataset contained volumetric MRI scans from patients of diverse backgrounds, acquired with different MRI protocols and machines. This is ideal for testing classifier robustness. The dataset contained five major classes: normal spine, disc degeneration, spinal stenosis, spondylolisthesis, and herniated disc. A total of 3,700 3D MRI volumes were used in this study. As model training is sensitive and performance is affected by variations in volumetric 3D scans, all volumes were adjusted to a fine $128 \times 128 \times 64$ voxel resolution. 70% of this dataset was allocated to the training data, 15% to the validation data, and the remaining 15% to the testing data. This approach, known as stratified sampling, ensures that all three data subsets are proportionally populated with samples from all abnormality classes in the dataset.

To avoid scanner-related intensity variability and enhance convergence stability, all MRI volumes were

z-score normalized prior to model training. To improve model generalization and reduce overfitting, 3D data augmentation techniques, such as random rotations, elastic deformations, intensity scaling, and spatial cropping, were applied. Due to volumetric memory constraints, the DS-STNet model was trained using the AdamW optimizer with a learning rate of 0.0001 and a batch size of 4. To avoid overtraining, the maximum number of training epochs was capped at 150, and early stopping was implemented based on the validation loss. In the classification head, a dropout regularization rate of 0.3 was applied. To account for the natural imbalance across spinal abnormality categories, a class-weighted cross-entropy loss was implemented.

The trainable parameters, floating-point operations (FLOPs), graphics processing unit (GPU) memory size, and model size were used to measure the computational complexity of the proposed DS-STNet. The network had 32.8 million trainable parameters and required 58.6 GFLOPs to process a $128 \times 128 \times 64$ MRI volume, as shown in Table 3. The model used 9.4 GB of GPU memory during training and 125.3 MB of storage. To ensure practicality for clinical use, the average inference time was 0.12 s per MRI volume.

Table 3. Computational complexity of the proposed DS-STNet

Metric	Value
Trainable parameters	32.8 million
Floating-point operation (FLOPs; input: $128 \times 128 \times 64$)	58.6 GFLOPs
Graphics processing unit memory	9.4 GB
Model size	125.3 MB

The proposed DS-STNet was implemented in Python 3.11 using PyTorch 2.2. We used an Intel Core i9 processor workstation with 64 GB of RAM and a 4090 GPU (24 GB of VRAM), running Windows 11 (64-bit), for model training and evaluation. The AdamW optimizer was used, with an initial learning rate of 1×10^{-4} and a batch size of 4.

4.1. Results

Table 4 presents the classification results of the DS-STNet model on previously unseen test samples. For each of the five anomaly classes, the model reported high precision, recall, and F1-score, even for visually similar classes that are difficult to discriminate. The model achieved an overall

precision of 98.58%, recall of 98.20%, F1-score of 98.39%, and AUC of 98.78%, all indicative of high classification quality.

Table 4. Performance of the proposed DS-STNet model on the test dataset

Class	Precision (%)	Recall (%)	F1-score (%)	AUC (%)
Normal spine	99.21	98.74	98.97	99.34
Disc degeneration	98.63	98.12	98.37	98.85
Spinal stenosis	98.44	98.05	98.24	98.66
Spondylolisthesis	97.92	97.68	97.80	98.12
Herniated disc	98.71	98.40	98.55	98.94
Overall average	98.58	98.20	98.39	98.78

Abbreviation: AUC: Area under the curve.

As presented in Figure 2, the confusion matrix for the DS-STNet model shows that test instances belonging to Class NP were occasionally misclassified as Class HER. This misclassification may be attributed to the visual similarity between these two degenerative anomalies on MRI. Incorporating a global inter-slice dependency model decreased this confusion compared with CNN-based models by enabling the network to capture long-range spatial relationships across MRI slices.

To validate the proposed DS-STNet model, it was

compared with leading deep learning methodologies used in spinal MRI analysis and/or medical image classification, including the standard 2D CNN, standard 3D CNN, 3D ResNet, 3D DenseNet, Vision Transformer, and a hybrid 3D CNN + long short-term memory (LSTM). All models were trained and evaluated under the same test setting using the same data and metrics. The results of the comparison are shown in Table 5. The equations for the measures are as follows:

(i) Accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total Number of Samples}} \quad (6)$$

(ii) Precision:

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \quad (7)$$

(iii) Recall:

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} \quad (8)$$

(iv) F-score (harmonic mean of precision and recall):

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

Based on the comparison in Table 5, we can conclude that the DS-STNet model consistently outperformed all baseline and state-of-the-art deep learning models. More specifically, the proposed model outperformed the traditional 2D CNN approach by approximately 9%, the baseline 3D CNN model by over 5%, the Vision

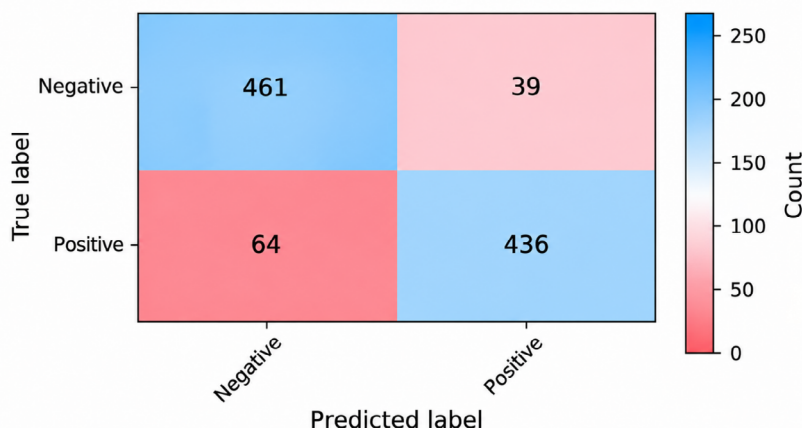


Figure 2. Confusion matrix of the proposed DS-STNet model for multi-class spinal abnormality classification

Transformer approach by almost 3%, and the hybrid 3D CNN + LSTM model by around 2%. The above results suggest that the 3D DS-STNet model can learn a richer, more discriminative representation of spinal abnormalities than any single architecture.

Table 5. Comparative performance analysis with existing methods

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC (%)
2D CNN	89.34	88.92	88.41	88.66	90.12
3D CNN	92.78	92.35	91.88	92.11	93.05
3D ResNet	94.62	94.09	93.71	93.90	94.88
3D DenseNet	95.14	94.88	94.33	94.60	95.41
Vision Transformer	96.02	95.77	95.31	95.54	96.26
3D CNN + LSTM	96.74	96.17	95.84	96.00	96.92
Proposed DS-STNet	98.61	98.58	98.20	98.39	98.78

Abbreviations: AUC: Area under the curve; CNN: Convolutional neural network; LSTM: Long short-term memory.

The proposed DS-STNet model has high computational performance and still trains and performs inference at turbo-unit speed. Inference on a single 3D MRI volume took less than 0.12 s, and the average training time per epoch was 2.8 min on a single GPU workstation. This indicates that the model performs well for near-real-time clinical applications and for healthy triple-digit hospital screenings.

Figure 3 presents the performance of the various deep learning models tested on spinal MRI data. The models included 2D CNN, 3D CNN, 3D ResNet, 3D DenseNet, Vision Transformer, and the proposed DS-STNet. Of all the models, the DS-STNet model achieved the highest accuracy (98.61%). The 2D CNN model achieved the lowest score because it could not analyze 3D data. 3D CNN models with and without residual connections performed better than the 2D CNN model, although the difference in performance is not substantial. Vision Transformer and 3D CNN + LSTM models performed better than other existing models, though they did not match the performance of DS-STNet, which captures data in 3D. All models in this case performed dual-scale spatiotemporal capture with spaced global attention.

The proposed DS-STNet model outperformed the other models because it handles data more efficiently. It has the most data-handling versatility and confidence in differentiating multiple pathologies. Other models cannot match it in accurately classifying spinal disorders. The experiments indicate that the proposed DS-STNet model can achieve clinically relevant and reliable triage and classification of spinal cord abnormalities from volumetric MRI scans. The model's ability to identify even subtle pathologies is due to refined dual-scale spatiotemporal feature extraction for global context modeling via attention in 3D MRI volumes. The results show that the proposed DS-STNet outperformed the baseline methods in classification performance for automated spinal MRI analysis.

The proposed DS-STNet was evaluated using the ablation study model (Table 6). The components individually improved classification accuracy. The highest accuracy of 98.61% was achieved by the complete model, which includes a dual-scale CNN, a Transformer, an

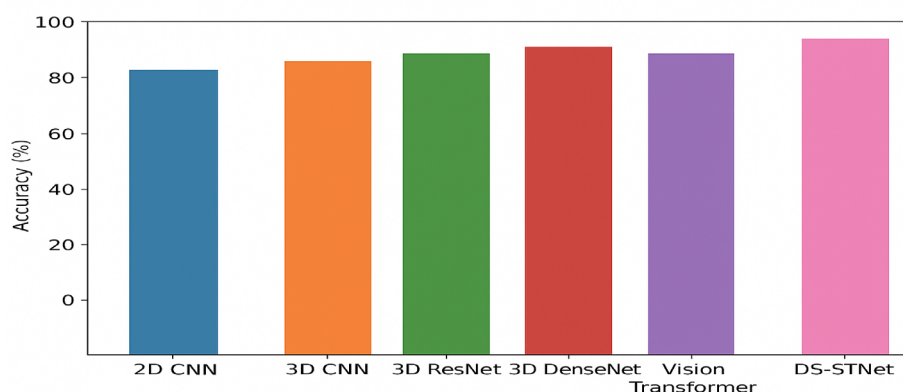


Figure 3. Accuracy comparison of the proposed DS-STNet model with existing methods
Abbreviation: CNN: Convolutional neural network.

attention module, and a weighted loss, demonstrating the efficiency of each module.

The proposed DS-STNet model was explainable using the gradient-weighted class activation mapping (Grad-CAM) method (Figure 4 and Table 7). The model's activation scores and localization accuracy suggest that it consistently identifies clinically relevant spinal abnormalities, thereby enhancing the interpretability of predictions.

Table 6. Ablation study of the proposed DS-STNet

Model configuration	Accuracy (%)	F1-score (%)
Baseline 3D CNN	94.12	93.78
Dual-scale CNN	95.84	95.42
Transformer	97.26	97.01
Attention module	98.04	97.82
Weighted loss (Proposed DS-STNet)	98.61	98.43

Abbreviation: CNN: Convolutional neural network.

4.3. Discussion

The proposed DS-STNet model achieved an overall accuracy of 98.61%, higher than that of previous deep

learning models. The fusion of a dual-scale 3D CNN and a Transformer has been shown to effectively capture both local anatomical and global contextual information, thereby improving classification accuracy.

This study has limitations. A publicly available dataset was used to evaluate the proposed framework; as a result, it may not capture the clinical variability encountered in real-world settings. Furthermore, visually similar spinal disorders were frequently misclassified as one another, indicating that distinguishing between closely related abnormalities remains challenging.

Table 7. Quantitative explainability analysis of the proposed DS-STNet

Disease class	Mean activation score	Lesion localization accuracy (%)
Normal	0.18	98.4
Disc degeneration	0.88	97.2
Herniated disc	0.93	98.1
Spinal stenosis	0.90	97.5
Spondylolisthesis	0.89	97.0

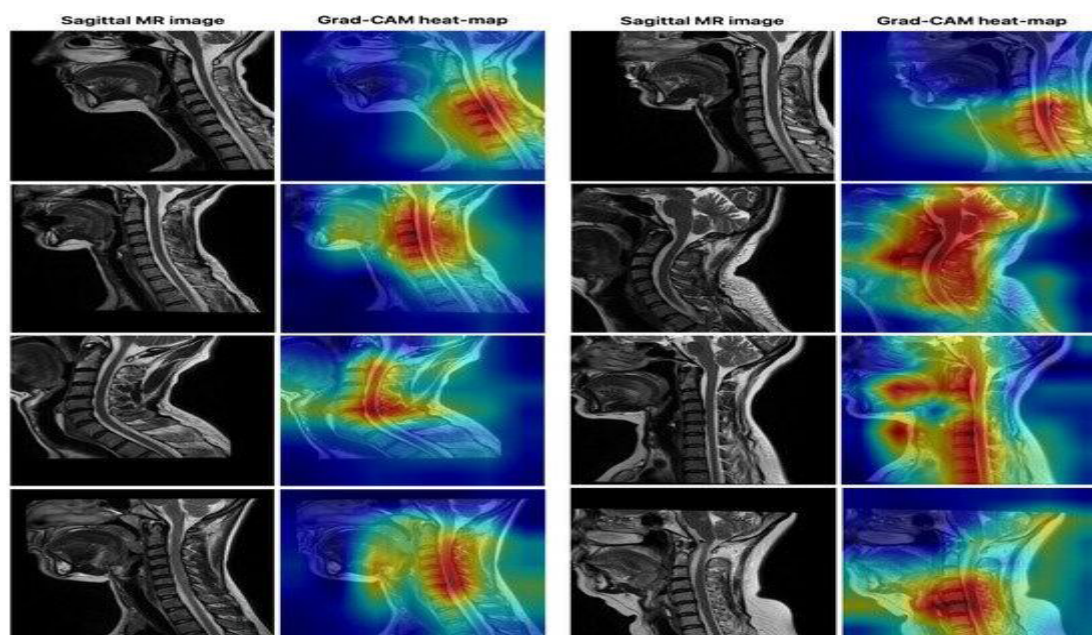


Figure 4. Representative Grad-CAM heatmaps produced by the proposed DS-STNet, highlighting clinically relevant regions in spinal MR images. The heatmaps show the regions with the highest (red) and lowest (blue) activations the model uses to predict its classification.

Abbreviations: Grad-CAM: Gradient-weighted class activation mapping; MR: Magnetic resonance.

The proposed model could assist radiologists by providing rapid and accurate automated classification of spinal MRI data. Given its high accuracy and low inference time, the model shows potential for integration into computer-aided diagnosis systems.

5. Conclusion and future work

In this paper, a hybrid 3D CNN–Transformer framework, DS-STNet, was developed for multi-class classification of spinal cord abnormalities. The framework integrates dual-scale spatiotemporal feature extraction, Transformer-based global dependency modeling, and attention-guided feature refinement within an end-to-end architecture. The proposed model aims to capture both fine-grained anatomical features and long-range inter-slice contextual relationships. Evaluation on a publicly available MRI dataset and comparison with state-of-the-art deep learning models demonstrated the effectiveness of the proposed framework. DS-STNet achieved an accuracy of 98.61% and demonstrated balanced precision, recall, F1, and AUC values across the spinal abnormality dimensions compared with the baseline methods. Confusion matrices and training curves further supported the model's classification performance and stable convergence. Future work should focus on multi-center clinical validation, real-time clinical deployment, lesion localization and segmentation, and privacy-preserving federated learning. Further refinement of the Transformer architecture could also improve model portability, computational efficiency, and energy efficiency, facilitating deployment on resource-constrained platforms. Large-scale evaluation across multiple clinical centers will be necessary to establish the generalizability and clinical utility of the proposed framework.

Acknowledgments

None.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Author contributions

Conceptualization: All authors

Formal analysis: Janardhanarao Addanki

Investigation: All authors

Methodology: Janardhanarao Addanki

Writing–original draft: Janardhanarao Addanki

Writing–review & editing: All authors

Ethics approval and consent to participate

Ethical approval is not required, as the study used publicly available datasets.

Consent for publication

Not applicable.

Availability of data

The datasets used in this study are publicly available and can be obtained from RSNA LumbarDISC (<https://www.kaggle.com/competitions/rsna-2024-lumbar-spine-degenerative-classification>) and SpineWeb (http://spineweb.digitalimaginggroup.ca/Index.php?n=Main.Datasets#Dataset_16.3A_609_spinal_anterior-posterior_x-ray_images). The datasets were accessed in accordance with the terms and conditions specified by the respective dataset providers.

Further disclosure

During the preparation of this manuscript, ChatGPT was used to improve language clarity, readability, grammar, and sentence structure, and to assist with coding error corrections.

References

- Amelot, A., Cretolle, C., de Saint Denis, T., Sarnacki, S., Catala, M., & Zerah, M. (2020). Spinal dysraphism as a new entity in V.A.C.T.E.R.L syndrome, resulting in a novel acronym V.A.C.T.E.R.L. *European Journal of Pediatrics*, 179(7), 1121–1129.
<https://doi.org/10.1007/s00431-020-03029-4>
- Chalif, J. I., Chavarro, V. S., Mensah, E., Johnston, B., Fields, D. P., Chalif, E. J., Chiang, M., Sutton, O., Yong, R., Trumbower, R., & Lu, Y. (2024). Epidural spinal cord stimulation for spinal cord injury in humans: A systematic review. *Journal of Clinical Medicine*, 13(4), 1090.
<https://doi.org/10.3390/jcm13041090>
- Chen, J., Qian, L., Ma, L., Urakov, T., Gu, W., & Liang, L. (2024). SymTC: A symbiotic Transformer-CNN net for instance segmentation of lumbar spine MRI. *Computers in Biology and Medicine*, 179, 108795.
<https://doi.org/10.1016/j.combiomed.2024.108795>
- Cui, H., Ruan, Z., Xu, Z., Luo, X., Dai, J., & Geng, D. (2024). ResMT: A hybrid CNN–Transformer framework for glioma grading with 3D MRI. *Computers & Electrical Engineering*, 120, 109745.
<https://doi.org/10.1016/j.compeleceng.2024.109745>
- Debska, J., Kotulska-Jozwiak, K., Pawliszak, P., Bilska, M.,

- Chmielewski, D., Puzio, I., & Jurkiewicz, E. (2017). Spinal cord lesions in children and adolescents with multiple sclerosis—Magnetic resonance imaging. *Neurologia i Neurochirurgia Polska*, 51(1), 77–81.
<https://doi.org/10.1016/j.pjnns.2016.11.006>
- Freigang, V., Butz, K., Seebauer, C. T., Karnosky, J., Lang, S., Alt, V., & Baumann, F. (2022). Management and mid-term outcome after «real SCIWORA» in children and adolescents. *Global Spine Journal*, 12(6), 1208–1213.
<https://doi.org/10.1177/2192568220979131>
- Geetha, R., Kumar, E., Vasantharaj, A., & Mohamedyaseen, A. (2025). Hybrid deep learning model for spinal tumor diagnosis on MRI scans. *Technology and Health Care*. Advance online publication.
<https://doi.org/10.1177/09287329251385793>
- Hernandez-Navarro, A., Ros-Alsina, A., Yurtseven, M., Wright, M., & Kumru, H. (2025). Non-invasive cerebral and spinal cord stimulation for motor and gait recovery in incomplete spinal cord injury: Systematic review and meta-analysis. *Journal of NeuroEngineering and Rehabilitation*, 22(1), 53.
<https://doi.org/10.1186/s12984-025-01557-4>
- Horsfield, M. A., Sala, S., Neema, M., Absinta, M., Bakshi, A., Sormani, M. P., Rocca, M. A., Bakshi, R., & Filippi, M. (2010). Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: Application in multiple sclerosis. *NeuroImage*, 50(2), 446–455.
<https://doi.org/10.1016/j.neuroimage.2009.12.080>
- Kirshblum, S., Snider, B., Rupp, R., & Read, M. S. (2020). Updates of the international standards for neurologic classification of spinal cord injury: 2015 and 2019. *Physical Medicine and Rehabilitation Clinics of North America*, 31(3), 319–330.
<https://doi.org/10.1016/j.pmr.2020.03.005>
- Liang, J., Wang, L., Hao, X., Wang, G., & Wu, X. (2022). Risk factors and prognosis of spinal cord injury without radiological abnormality in children in China. *BMC Musculoskeletal Disorders*, 23(1), 428.
<https://doi.org/10.1186/s12891-022-05393-8>
- Ma, R., Dai, X., Yang, Z., Wei, Z., & Zhang, B. (2025). CAFR-Net: A transformer-contrastive framework for robust spinal MRI segmentation via global-local synergy. *PLOS ONE*, 20(7), e0327642.
<https://doi.org/10.1371/journal.pone.0327642>
- Mann, R. S., Constantinescu, C. S., & Tench, C. R. (2007). Upper cervical spinal cord cross-sectional area in relapsing remitting multiple sclerosis: Application of a new technique for measuring cross-sectional area on magnetic resonance images. *Journal of Magnetic Resonance Imaging*, 26(1), 61–65.
<https://doi.org/10.1002/jmri.20959>
- Margoni, M., Preziosa, P., Rocca, M. A., & Filippi, M. (2022). Pediatric multiple sclerosis: Developments in timely diagnosis and prognostication. *Expert Review of Neurotherapeutics*, 22(5), 393–403.
<https://doi.org/10.1080/14737175.2022.2064743>
- Margoni, M., Storelli, L., Pagani, E., Preziosa, P., Mistri, D., Gueye, M., Rubin, M., Moiola, L., Filippi, M., & Rocca, M. A. (2025). Subventricular zone microstructure in pediatric-onset multiple sclerosis. *Annals of Neurology*, 97(5), 979–992.
<https://doi.org/10.1002/ana.27180>
- Mezzapesa, D. M., Rocca, M. A., Falini, A., Rodegher, M. E., Ghezzi, A., Comi, G., & Filippi, M. (2004). A preliminary diffusion tensor and magnetization transfer magnetic resonance imaging study of early-onset multiple sclerosis. *Archives of Neurology*, 61(3), 366–368.
<https://doi.org/10.1001/archneur.61.3.366>
- Rocca, M. A., Valsasina, P., Meani, A., Gobbi, C., Zecca, C., Barkhof, F., Schoonheim, M. M., Strijbis, E. M., Vrenken, H., Gallo, A., Bisecco, A., Ciccarelli, O., Yiannakas, M., Rovira, A., Sastre-Garriga, J., Palace, J., Matthews, L., Gass, A., Eisele, P., ... Filippi, M. (2023). Spinal cord lesions and brain grey matter atrophy independently predict clinical worsening in definite multiple sclerosis: A 5-year, multicentre study. *Journal of Neurology, Neurosurgery & Psychiatry*, 94(1), 10–18.
<https://doi.org/10.1136/jnnp-2022-329854>
- Seijas, V., Schrepfer, L., Posada, A. M., Spir, M. A., Machado, B., Sigrist-Nix, D., Scheel-Sailer, A., Eriks-Hoogland, I., & Sabariego, C. (2024). Evidence-based recommendations for the rehabilitation and management of the ageing population with spinal cord injury: A systematic review of clinical practice guidelines. *European Journal of Physical and Rehabilitation Medicine*, 60(3), 433–444.
<https://doi.org/10.23736/S1973-9087.24.08244-3>
- Sergeenko, O. M., Savin, D. M., & Diachkov, K. A. (2024). Association of spinal cord abnormalities with vertebral anomalies: An embryological perspective. *Child's Nervous System*, 40(5), 1415–1425.
<https://doi.org/10.1007/s00381-024-06336-5>
- Shinozaki, M., Nagoshi, N., Nakamura, M., & Okano, H. (2021). Mechanisms of stem cell therapy in spinal cord injuries. *Cells*, 10(10), 2676.
<https://doi.org/10.3390/cells10102676>
- Suleman, S., Harwood, M., Kang, P., Narayan, R., & Hughes, J. N. (2022). Assessment of cervical spinal cord volume in pediatric-onset multiple sclerosis. *Clinical Imaging*, 90, 44–49.
<https://doi.org/10.1016/j.clinimag.2022.06.021>

- Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., Correale, J., Fazekas, F., Filippi, M., Freedman, M. S., Fujihara, K., Galetta, S. L., Hartung, H. P., Kappos, L., Lublin, F. D., Marrie, R. A., Miller, A. E., Miller, D. H., Montalban, X., ... Cohen, J. A. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology*, 17(2), 162–173.
[https://doi.org/10.1016/S1474-4422\(17\)30470-2](https://doi.org/10.1016/S1474-4422(17)30470-2)
- Valsasina, P., Aboulwafa, M., Preziosa, P., Messina, R., Falini, A., Comi, G., Filippi, M., & Rocca, M. A. (2018). Cervical cord T1-weighted hypointense lesions at MR imaging in multiple sclerosis: Relationship to cord atrophy and disability. *Radiology*, 288(1), 234–244.
<https://doi.org/10.1148/radiol.2018172311>
- Valsasina, P., Horsfield, M. A., Rocca, M. A., Absinta, M., Comi, G., & Filippi, M. (2012). Spatial normalization and regional assessment of cord atrophy: Voxel-based analysis of cervical cord 3D T1-weighted images. *American Journal of Neuroradiology*, 33(11), 2195–2200.
<https://doi.org/10.3174/AJNR.A3139>
- Valverde, S., Cabezas, M., Roura, E., González-Villà, S., Pareto, D., Vilanova, J. C., Ramió-Torrentà, L., Rovira, A., Oliver, A., & Lladó, X. (2017). Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. *NeuroImage*, 155, 159–168.
<https://doi.org/10.1016/j.neuroimage.2017.04.034>
- Verhey, L. H., Branson, H. M., Makhija, M., Shroff, M., & Banwell, B. (2010). Magnetic resonance imaging features of the spinal cord in pediatric multiple sclerosis: A preliminary study. *Neuroradiology*, 52(12), 1153–1162.
<https://doi.org/10.1007/s00234-010-0755-9>
- Wang, Y., Zeng, L., Zhu, F., Huang, G., Wan, Y., Yao, S., Chen, K., & Guo, X. (2022). Acute hyperextension myelopathy in children: Radiographic predictors of clinical improvement. *Spinal Cord*, 60(6), 498–503.
<https://doi.org/10.1038/s41393-021-00755-7>
- Weeda, M. M., Middelkoop, S. M., Steenwijk, M. D., Daams, M., Amiri, H., Brouwer, I., Killestein, J., Uitdehaag, B. M. J., Dekker, I., Lukas, C., Bellenberg, B., Barkhof, F., Pouwels, P. J. W., & Vrenken, H. (2019). Validation of mean upper cervical cord area (MUCCA) measurement techniques in multiple sclerosis (MS): High reproducibility and robustness to lesions, but large software and scanner effects. *NeuroImage: Clinical*, 24, 101962.
<https://doi.org/10.1016/j.nicl.2019.101962>