

BP-01

A Machine Learning Model for Predicting Surgical Intervention in Mild Degenerative Cervical Myelopathy

Sung Bae Park¹

¹Seoul National University Borama h., Neurosurgery, Seoul, South Korea

Background: The patients with mild Degenerative Cervical Myelopathy (DCM) may require surgical intervention within 5 years to prevent further functional decline.

Research Question: Can a ML model with fine-tuned vision transformers reliably predict the need for surgical intervention in patients with mild DCM?

Design: retrospective study

Methods: Data from 294 patients diagnosed with mild DCM were collected. MRI images underwent pre-processing. The dataset was split into training (n=236) and testing (n=58) subsets. Model development began with ResNet18(CNN) optimized through CB_focal, mix-up data augmentation, and multiple input combinations of original and segmentation images. Fine-tuned foundation models (BioMedCLIP) with multi-task learning frameworks. Performance metrics included AUC, sensitivity, and F1 score.

Results: The ResNet18 model achieved a high AUC (>0.9), demonstrating reliable differentiation between surgical and non-surgical cases. Sensitivity was maximized to ensure minimal omission of patients requiring surgery. Mix-up data augmentation improved the model's ability to predict cases with borderline probabilities, further refining decision-making thresholds. Grad-CAM visualizations provided insight into the model's focus on critical regions, such as areas of canal and cord compression, aligning with clinical relevance. Fine-tuned BioMedCLIP and multi-task learning enhanced generalizability but did not surpass the ResNet18 model's performance. Ensemble methods, averaging predictions across multiple MRI slices, further strengthened prediction stability.

Discussion: By leveraging small datasets with advanced modeling techniques, the ResNet18 model achieved clinically actionable predictions, facilitating earlier surgical interventions for patients at risk of deterioration.

Fig. 1

Results

- Fine-tuning Foundation Models – BioMedCLIP

Test	F1	Accuracy	Specificity	Sensitivity	AUC	Optimal threshold
Image 3	0.4000	0.7414	0.3571	0.8636	0.7419	

- Ensemble

Test	F1	Accuracy	Specificity	Sensitivity	AUC	Optimal threshold
Image 1 + Blank 2 w/ CB_focal	0.7692	0.8966	0.7143	0.9545	0.8734	0.430114

- Multi-task Learning – ConvNext & Image Resampling

Test	F1	Accuracy	Specificity	Sensitivity	AUC	Optimal threshold
Image 1 + Canal mask 1 + Blank 1 w/ CB_focal	0.6400	0.8448	0.5714	0.9318	0.8880	

Fig. 2

