

A deep learning-based object detection method for the recognition of anatomical structures in pterional neurosurgical images

Original

A deep learning-based object detection method for the recognition of anatomical structures in pterional neurosurgical images / Morello, A., Nitti, F., Dolci, L., Molinari, F., Garbossa, D., Salvi, M., Cofano, F.. - In: BRAIN AND SPINE. - ISSN 2772-5294. - 5:(2025). (EANS2025) [10.1016/j.bas.2025.105132].

Availability:

This version is available at: 11583/3003912 since: 2025-10-13T14:49:25Z

Publisher:

Elsevier

Published

DOI:10.1016/j.bas.2025.105132

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

specific training experience for neurosurgery residents.

Methods

Thirty-four neurosurgery residents at the University of Milan participated in a structured training session where next-generation, 3D-printed brain with tumor models were used. These models, derived from real patient imaging data, accurately replicated glioblastomas (GBMs), low-grade gliomas, and extra-axial tumors (meningioma-like). The GBM models fluoresced under 5-aminolevulinic acid (5-ALA), enabling realistic simulation of intraoperative tumor visualization. The eco-sustainable models closely mimicked the consistency and resectability of real brain tumors and were ultrasound-scannable, producing echogenic images identical to real intraoperative conditions. Training was conducted using standard operating room equipment, including ultrasound imaging, the Orbeye exoscope, fluorescence-guided surgery, and an ultrasonic aspirator (CUSA Clarity), ensuring a highly realistic experience. After the session, residents completed a Likert-scale satisfaction questionnaire assessing the models' realism, utility, and impact on surgical confidence.

Results

Residents reported a 45% increase in confidence in tumor resection procedures, particularly in identifying and excising fluorescent tumor margins. The models were rated highly realistic in terms of tactile feedback, anatomical accuracy, and compatibility with neurosurgical instruments. The ultrasound-scannable feature was widely appreciated before, during, and after tumor resection, for its role in improving intraoperative imaging skills.

Conclusions

These 3D-printed models represent a breakthrough in neurosurgical training, providing a safe, affordable, standardized, and ethical alternative to traditional patient-based learning. Their high fidelity, eco-sustainability, and compatibility with different surgical technologies make them a promising new paradigm for neurosurgical training, filling the gap between virtual simulation and real-life surgery.

Optional Image

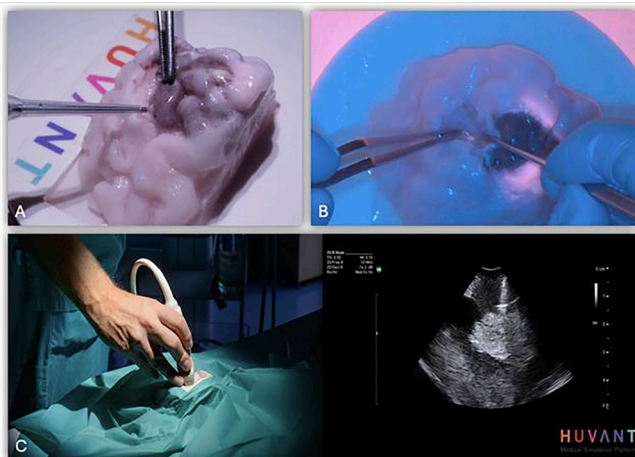


Fig. 1: HUVAANT MODEL. Figure 1A shows the 3D model representing an intracerebral glial lesion. Figure 1B demonstrates the visualization of the lesion using 5-ALA fluorescence. Figure 1C displays the ultrasound imaging of the lesion.*

<https://doi.org/10.1016/j.bas.2025.105132>

1261

Beyond Facts: The Role of Empathy in AI Models Responding to Patient Queries about Lumbar Spinal Stenosis

Dr. Sami Ridwan¹, Alan H. Daniels, Prof.², Dr. med Stefan Motov^{3,4}, Rory K.J. Murphy, Ass. Prof.⁵, Dr. med Michael Schwake, Prof.⁶, MD Michael Selby⁷, Martin N. Stienen, Prof.^{3,4}, Dr. med Patrick Schuss, Prof.⁸

¹ Department of Neurosurgery, Klinikum Ibbenbueren, Ibbenbueren, Germany;

² Department of Orthopaedic Surgery, Warren Alpert Medical School, Brown University, Providence, United States; ³ Spine Center of Eastern Switzerland, Kantonsspital St. Gallen & Medical School of St. Gallen, St. Gallen, Switzerland;

⁴ Department of Neurosurgery, Kantonsspital St. Gallen & Medical School of St. Gallen, Switzerland; ⁵ Department of Neurosurgery, Barrow Neurological Institute,

St. Joseph's Hospital and Medical Center, Phoenix, United States; ⁶ Department of Neurosurgery, University Hospital Münster, Münster, Germany; ⁷ Adelaide Spine

and Brain, North Adelaide, Australia; ⁸ Department of Neurosurgery, BG Klinikum Unfallkrankenhaus Berlin, Germany

Advancement in AI and robotics in neurosurgery (Trends & Innovation Morning Session), Lehar 2, October 8, 2025, 7:00 AM - 8:00 AM

Patients are increasingly turning to Large Language Models (LLMs) for medical information, yet the reliability and patient-centeredness of their responses remain debated. This study assesses the performance of LLMs in addressing patient queries related to lumbar spinal stenosis (LSS).

Methods

Most common patient questions about LSS were gathered from six categories using Google's "People Also Ask" function. These were submitted to three LLMs: ChatGPT-4o, Gemini Advanced 2.0, and DeepSeek V3. Spine experts evaluated the responses based on completeness, clarity, empathy, and appropriateness. They also assessed overall quality and whether LLM-generated answers could sufficiently replace professional medical advice. Multivariate analysis was conducted to compare LLM performance, and logistic regression was used to identify factors influencing the likelihood of LLMs being perceived as viable substitutes for experts.

Results

The three LLMs provided generally high-quality responses, with no significant differences in content, communication, empathy, or length. DeepSeek produced the highest proportion of "excellent" ratings (24%), followed by Gemini Advanced (21%) and ChatGPT-4o (13%). Gemini Advanced was most frequently considered a sufficient replacement for expert advice (28%), followed by DeepSeek (24%) and ChatGPT-4o (17%). Regression analysis revealed that higher empathy scores significantly influenced the perception of LLM reliability as an expert substitute ($p=0.04$), whereas accuracy alone was not a decisive factor.

Conclusions

As AI models continue to improve and become more accessible, regular scientific evaluation of their responses is essential. While LLM-generated answers were generally rated as above average, none of the models consistently served as an adequate substitute for expert guidance. Empathy emerged as a critical determinant of perceived reliability, suggesting that factual accuracy alone is insufficient to establish trust in AI-generated medical advice.

<https://doi.org/10.1016/j.bas.2025.105131>

1623

A deep learning-based object detection method for the recognition of anatomical structures in pterional neurosurgical images

Dr. Alberto Morello¹, Dr. Francesco Nitti², Dr. Lorenzo Dolci¹, Filippo Molinari, Prof.², Diego Garbossa, Prof.¹, Massimo Salvi, Prof.², Fabio Cofano, Prof.¹

¹ Neurosurgery Unit, Department of Neuroscience "Rita Levi Montalcini", "Città della Salute e della Scienza" University Hospital, University of Turin, Turin, Italy;

² Biolab, PoliToBIOMed Lab, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy

Advancement in AI and robotics in neurosurgery (Trends & Innovation Morning Session), Lehar 2, October 8, 2025, 7:00 AM - 8:00 AM

The use of Artificial Intelligence (AI) has already changed neurosurgical approach to various pathologies. The application of machine vision algorithms relates to real-time anatomical recognition based on live image feedback from the surgical device and has the potential to address these disadvantages and to serve as a reliable tool for intraoperative orientation. With this motivation, we present a method for live image-only guidance by exploiting a large dataset of annotated neurosurgical videos, with the potential to be a tool to aid in the training of residents, particularly in learning neurosurgical anatomy.

Methods

By manually annotating the common anatomical pathways of 34 pterional surgeries in the training set, we generated a neurosurgical roadmap. The mapping is modeled through the use of an autoencoder architecture. The data was split at video level into training (60%), validation (25%), and test (15%) sets. Approximately 100 frames per video were densely annotated at pixel level, resulting in a total of 1,300 labeled frames, representing about 5% of the key frames in the dataset. To evaluate the impact of this AI model in improving the performance of residents, 21 residents are recruited for the study.

Results

The dataset consists of 34 neurosurgical videos recorded during open craniotomy procedures at Molinette Hospital, Turin. The accuracy of the automatic recognition of anatomical structures was 89.3%. Moreover, this system could perform real-time automatic surgical phase recognition at 32 fps. In addition, when assisted by the AI model, the performance of the residents improved in a manner that was inversely proportional to their level of training.

Conclusions

In this paper, we propose a new approach to the neuronavigation based on deep learning. This practice and proposed application of AI is based on live images and uses segmentations of anatomical structures to locate on a common surgical roadmap. The method makes it possible to locate and predict anatomical structures that will be encountered back and forth along the surgical pathway, similar to a mapping application. In addition, this model can be used as tool for learning neurosurgical anatomy.

<https://doi.org/10.1016/j.bas.2025.105132>

56

Pilot Study: AI-Assisted Deep Analysis of Cervical Spinal Stenosis – Integrating Deep Learning and Radiomics to Enhance Diagnostic Precision

Dr. Pavlina Lenga¹, Vasilios Mougios¹, Alexandra Ertl², Jonas Lin¹, Daniela Becker¹, Klaus Maier Hein³, Peter Neher², Sandro Krieg¹, Robin Peretzke²

¹ Department of Neurosurgery, Heidelberg University Hospital, Heidelberg, Germany;

² Division of Medical Image Computing, German Cancer Research Center, Heidelberg, Germany; ³ Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany

Cervical degenerative (Spine Morning Session), Strauss 2-3, October 8, 2025, 7:00 AM - 8:00 AM

The prevalence of degenerative spinal disorders, including cervical spinal stenosis leading to cervical spondylotic myelopathy (CSM), is rapidly increasing in the geriatric population worldwide. Neurosurgeons must adeptly interpret MRI images to correlate complex imaging findings with patient symptoms, a task that is often intricate and nuanced. Furthermore, explaining the surgical necessity to patients can be daunting due to the complexity of these findings. Here, we present preliminary results from the development of an AI-based algorithm designed to accurately identify cervical stenosis in MRI scans, perform volumetric analysis, and automatically calculate a comprehensive range of metrics to enhance the validation and understanding of the disease.

Methods

Our system is designed as a multi-step workflow to facilitate rapid end-to-end analysis. First, it automatically identifies the stenosis localization on axial T2 sequences. Next, it performs automatic spine and vertebrae segmentation. This is followed by the computation of compression metrics, generated using the Spinal Cord Toolbox. Finally, the segmentations and localization data are used to calculate the extension of measures based on radiomics. Aiming to find biomarkers for stenosis quantification we correlated both the morphologic measures of the SCT as well as the radiomics based volumetric measures with the JOA pre OP score using the Pearson correlation. Figure 1 presents the workflow.

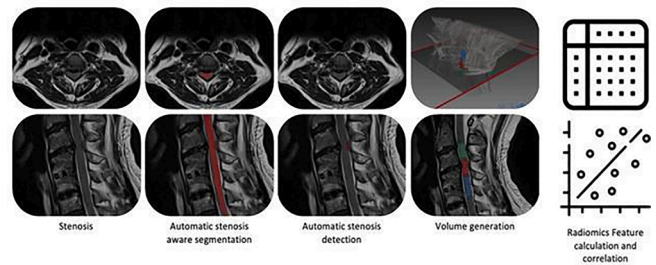
Results

Our dataset consists out of 42 subjects on which we trained our models for spinal segmentation and detection. Our nnU-Net based stenosis detection algorithm failed to make predictions in 2/9 test cases. The remaining 7 cases were predicted highly accurately with an absolute mean euclidean distance of 1.834 ± 1.209 . All predictions lie within a 5 mm tolerance. Consequently our model predicted the location of the stenosis correctly in 77.8 % of our test cases. All 5 morphologic measures and 107 radiomics features were compared to JOA preop using Pearson correlation. The significance level 0.05 was reduced to 0.00045 in respect to Bonferroni correction. We found a significant correlation ($p=0.0001$) to one radiomics features.

Conclusions

This AI-based approach to cervical stenosis detection demonstrates promising accuracy in identifying stenosis and extracting relevant features, with strong potential to enhance clinical decision-making. While the model shows good performance, our future goal is to expand the dataset, improving segmentation quality, and refining the system to achieve broader clinical applicability.

Optional Image



<https://doi.org/10.1016/j.bas.2025.105133>

516

Exploring Outcome Discrepancies between Patients Receiving and not Receiving Plates following Anterior Discectomy and Fusion

Ms Azra Gül^{1,2}, Ms Jessica Grace², Ms Daniela Limbania², Dr. Timothy R. Smith², Dr. Carmen Vleggeert-Lankamp^{1,2,3}

¹ Leiden University Medical Center, Leiden, Netherlands; ² Computational Neuroscience Outcomes Center, Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, United States; ³ Department of Neurosurgery, Spaarne Gasthuis, Netherlands

Cervical degenerative (Spine Morning Session), Strauss 2-3, October 8, 2025, 7:00 AM - 8:00 AM

Anterior Cervical Discectomy and Fusion (ACDF) for cervical radiculopathy is commonly performed with a cage and plate in the United States, whereas in Europe, standalone cages are frequently used instead. Knowledge about difference in clinical outcome is scarce. This study aims to enhance this knowledge to assist physicians in optimizing treatment choices.

Methods

A multi-centre retrospective cohort study was conducted in patients from the US and from the Netherlands who underwent a single-level ACDF procedure between 2015 and 2024. A total of 570 patients with cervical radiculopathy were included: 414 received a cage with plate (US), 54 received cages with integrated screws (US) and 102 received a standalone cage (52% Netherlands).

Clinical outcomes were assessed using measures of arm pain, neck pain, anxiety/-depression and pain interference at baseline, 3, 6, 12 and 24 months after inclusion, and analysed using linear mixed-effects models. Furthermore, to assess individual patient-level outcome, cut-off values for success were utilized (successful if arm pain ≤ 2.5 , neck pain ≤ 3.5)

Results

At baseline, the three groups were comparable. Linear mixed-effects models revealed no statistically significant differences in clinical outcome over time. However, when examining success outcome for neck and arm pain, the standalone cage group demonstrated the highest percentage of patients achieving successful outcomes for arm pain; 62.9% at 6 months, 79.4% at 12 months, and 85.7% at 24 months. In the cage with plate group, success rates stabilized at 59.8% at 6 months, 60.7% at 12 months, and 58.6% at 24 months. The cage with integrated screws group showed greater variability, with success rates of 56.0% at 6 months, 26.7% at 12 months, and 44.4% at 24 months. A Pearson Chi-Square test revealed a significant difference after 12 months ($p < 0.001$) and 24 months ($p=0.005$) of follow-up in successful outcome. A similar trend was observed for arm pain.

Conclusions

No significant differences were found in overall clinical outcomes between the three ACDF constructs. However, success rates for arm and neck pain appeared to be the highest in the standalone cage group.

Optional Image