Performance Comparison of GAN-Augmented and Traditional CNN Models for Spinal Cord Tumor Detection

Dr. Bharati Ainapure[1], *Dr. Shrikaant Kulkarni* [2], *Dr. Midhunchakkaravarthy*[3]
[1] *Research scholar (Post Doc), Associate Professor, Vishwakarma University* ; [2] *Research Professor, Sanjivani University, India* ; [3] Professor, Lincoln University College
Email ID : pdf.bharati@lincoln.edu.my

Abstract: Detecting spinal cord tumors using MRI scans is a very challenging issue within clinical radiology. The availability of high-quality, expertly annotated datasets which can be useful for spinal cord tumor detection are limited since the intricate anatomy of the spinal cord complicates image interpretation and this tumor is comparatively rare. Convolutional Neural Network (CNNs) have emerged as reliable solution for automating this detection process; however, they often tend to fail when they are provided with limited or imbalanced datasets. This limitation can be overcome using Generative Adversarial Networks (GANs) by generating realistic synthetic MRI images in order to enhance dataset diversity and effectively mitigate class imbalance. In this paper, we have provided one-qualitative comparison of GAN-augmented CNNs versus traditional CNNs for spinal cord tumor detection. We have tried to demonstrate how GAN-based augmentation improves model accuracy, generalization, robustness and overall classification metrics. We have also discussed current limitations, particularly the difficulty of training stable GANs and the underutilization of full 3D volumetric data. Moving forward, we have mentioned challenges and key directions for future work.

Keywords: Machine Learning; Deep Learning; Spinal Cord Cancer; Convolutional Neural Networks; Generative Artificial Intelligence; Diagnostic Accuracy; Predictive Analytics; Generative Adversarial Networks;

Introduction

Spinal cord tumors represent a rare subset of central nervous system malignancies, but their consequences can be severe. Even a small lesion can potentially lead to permanent paralysis or loss of sensation if they're diagnosed late. Magnetic Resonance Imaging (MRI) is the well know tool used in healthcare industry for spinal cord tumors detection. This tool is widely used as this is non-invasive procedure, images are captured from multiple angles and provides excellent contrast for soft tissues [1]. Though this traditional system of disease identification is beneficial, but this has disadvantages also 1.It is time-consuming, 2. Difficult and 3. Accuracy of the diagnosis is directly dependent on the radiologist's level of expertise and experiences in their clinical practice, a radiologist's capacity to correctly diagnose spinal tumors is primarily dependent on their experience and skill level. As a result, there is a growing demand for automated, data-driven diagnostic systems that can produce reliable and consistent results.

Medical image analysis is emerging field as Deep Learning (DL) and especially Convolutional Neural Networks (CNNs) [2] playing crucial role in image feature identification. CNNs are meant to extract important features from raw pixel data. Figure 1 shows general architecture of CNN. These networks are able to learn significant features like edges and contours from the images. Then gradually network learns to identify more complex patterns like anatomical structure and diseases. But training of the CNN needs huge and well labeled dataset to get optimal performance.

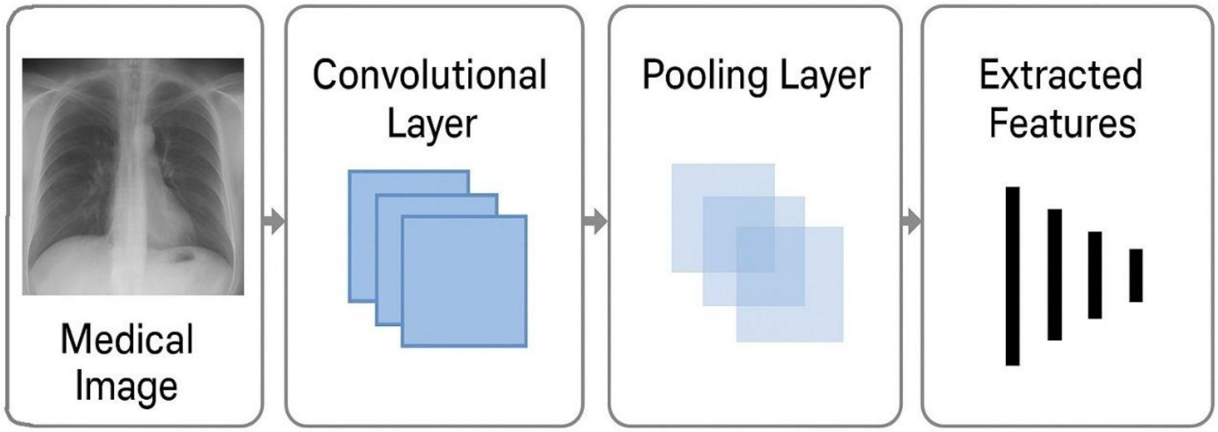

Figure 1: Architecture of CNN

In the field of spinal cord images there is always scarcity of good images to create prediction models. To overcome this difficulty, Generative Adversarial Networks (GANs) have emerged as a promising solution. These networks are very help to create pool of images for CNN training. GAN contains two networks, 1. Generate and 2. Discriminator. These two subnetworks compete each other to produce high quality synthetic medical images that closely reflect real medical images [3]. By augmenting existing datasets with GAN-generated images, researchers can significantly expand training dataset volume as well as diversity to create more robust and accurate classifiers. This is very useful in medical domains where getting enough data is often limited by ethical concerns, patient privacy laws, and the rarity of certain conditions [4]–[6].

In this paper we propose comparative study focusing on spinal cord tumour detection using traditional CNN models and GAN-enhanced CNN architectures. This work reviews datasets, key machine learning methods and performance metrics used in existing research, while outlining the benefits of adding GANs such as higher accuracy, better handling of imbalanced data, and more interpretable results. We have concluded by outlining future scope involving 3D GAN-CNN hybrids, explainable AI, and federated learning frameworks for multi-institutional collaboration.

Background and Related Work

a) Spinal Cord Tumor Detection: Challenges and Imaging Considerations

Spinal cord tumors represent a relatively rare and account for approximately 10–15% of all central nervous system (CNS) neoplasms. Their diagnosis is clinically significant because of its anatomical position because even small delays in diagnosis or errors in detection can cause irreversible neurological damage like permanent damage to a patient's ability to move or feel sensation. While MRI remains the standard tool for examining spine but automated spinal MRI analysis is still very challenging. It can be difficult to consistently segment and classify tumors due to their wide variations in shape, texture, intensity distribution, and spatial orientation [7], [8].

These complexities are escalated by the limited availability of large, high quality and well-balanced datasets for spinal tumor research. On the contrary, brain or lung imaging has robust, publicly available datasets that researchers worldwide can access easily[6] . There's no widespread repository for spinal tumor imaging. Moreover, the ones that are available are typically small and unbalanced.

b) Deep Learning in Medical Image Analysis

DL has completely changed the way we approach medical image analysis. Its ability to learn complex patterns directly from raw imaging data without needing manual feature engineering is really commendable. When it comes to medical imaging tasks like MRIs, CT scans, and ultrasounds, CNNs have proven especially effective. These networks are able to identify both specific details and more general patterns in medical images because they are actually modelled after how our own visual system functions. Due to this, they are being widely applied in various problems such as brain tumor segmentation, diabetic retinopathy grading, organ delineation and other diagnostic imaging challenges. CNNs are excellent at interpreting visual information because they can learn hierarchical features, ranging from simple edges to complex anatomical structures.

However, CNNs face two key limitations in medical domains: (1) they require large and well annotated dataset which are difficult to obtain, and (2) when trained on limited or imbalanced datasets, CNNs are prone to overfitting, resulting in models that perform well on training data but fail to generalize to unseen clinical cases. Traditional augmentation techniques such as geometric transformations, intensity variations, and spatial perturbations cannot reproduce the realistic anatomical diversity seen in real patient anatomy or pathology. In order to fill this gap, Generative Adversarial Networks (GANs) have been used to create artificial but clinically relevant model training samples.

c) Convolutional Neural Network (CNN)-Based Models for Tumor Detection

Early studies employed AlexNet and VGGNet architectures to classify MRI slices, while later works integrated deeper networks such as ResNet, DenseNet, and InceptionNet for improved feature extraction. Thus, CNNs have shown real promise for diagnosing spinal and brain tumors. For instance, authors in [9], [10] developed a multi-scale CNN for lung and medical image classification and found it outperformed traditional region-growing methods when measuring how well the segmentation matched expert annotations. Author in [11] took a different approach, using transfer learning with ResNet50 to classify spinal MRIs and managed to achieve over 90% accuracy even with relatively small datasets. These techniques demonstrate CNNs' ability to capture spatial correlations, but they also show how much their performance consistency depends on large labelled datasets.

d) Generative Adversarial Networks (GANs) for Medical Image Augmentation

GAN were first introduced by Goodfellow et al. (2014), are made up of two competing neural networks: a discriminator that distinguishes between real and synthetic samples and a generator that creates images. Through this competitive training process, GANs acquire the ability to capture complex, high-dimensional data distributions, ultimately enabling the generation of medical images that closely resemble actual clinical scans [12].

Several GAN variants have been adapted for medical imaging [13], [14]:

- DCGAN (Deep Convolutional GAN): Generates high-quality grayscale or MRI-like textures.
- CycleGAN: Performs domain-to-domain translation (e.g., T1 ↔ T2 MRI).

- Conditional GAN (cGAN): Generates images conditioned on specific labels or modalities.
- StyleGAN: Enables fine-grained control over anatomical and texture details.

Authors in [15] showed that expanding training datasets with synthetic lesions produced by GAN can significantly improve the performance of CNN-based classification models in medical imaging tasks, including those involving spinal and brain MRI data. Zhou et al. (2022) took this approach further by using a Deep Convolutional GAN specifically designed for spinal cord MRI images. Their model performed significantly better when tested on fresh scans which were never seen before and observed a 6–8% improvement in classification accuracy [16]. Together, these studies provide strong evidence that GAN-generated images improve the reliability of models in practical clinical settings.

e) Hybrid GAN–CNN Architectures

Recently, researchers have been investigating how to pair CNNs and GANs together such that we can obtain advantage of each method. In these frameworks, generally GANs act as data generators, while CNNs serve as classifiers or segments. For example:

- Han et al. (2021) proposed a GAN-based data augmentation pipeline followed by a CNN classifier for glioma detection, achieving higher F1-scores compared to standard augmentation [17].
- R. Polattimur et al. (2025) used a conditional GAN to generate labeled MRI slices for spinal tumor segmentation, improving Dice and IoU scores [18].
- Berman et al. (2023) introduced an adversarial trained CNN that learns to resist noise, enhancing robustness against MRI intensity variations [19].

Collectively, these hybrid models demonstrate that GAN-augmented CNNs consistently outperform traditional CNNs in data-scarce environments.

Comparative Analysis

This section reviews published studies comparing traditional CNN and GAN-augmented CNN models for spinal and related tumor detection in MRI.

a) Evaluation Parameters

Table 1: Evaluation metrics

Metric	Definition / Purpose
Accuracy (ACC)	Overall percentage of correctly classified samples.
Precision (P)	Proportion of predicted tumor regions that are truly positive.
Recall (R) / Sensitivity	Fraction of actual tumors correctly detected.
F1-Score	Harmonic mean of Precision and Recall, balancing false positives and negatives.
Dice Coefficient (DSC)	Measures spatial overlap for segmentation.
AUC-ROC	Evaluates discrimination ability across thresholds.

The metrics shown in Table 1 are consistently reported in most comparative studies.

b) Qualitative Insights

Comparative analysis of traditional CNN models versus GAN-augmented CNN models is presented in Table 2

Table 2: Qualitative insights of GAN and CNN models

Aspect	Traditional CNN Models	GAN-Augmented CNN Models
Data Dependence	Require large labeled datasets; prone to overfitting on small data.	Benefit from synthetic data, improving generalization.
Feature Localization	Broader focus; may miss subtle tumor boundaries.	Grad-CAM shows sharper attention on tumor regions.
Noise Robustness	Sensitive to MRI artifacts and intensity variations.	Adversarial training improves noise tolerance.
Model Stability	Relatively stable training.	Sensitive to GAN quality; requires careful tuning.
Clinical Relevance	Limited by data imbalance.	Higher recall → fewer missed tumor cases.

Technical Challenges

Technical challenges to implement GAN based CNN models are listed in Table 3.

Table 3: List of technical challenges

Challenge	What It Means in Practice
Data Scarcity	There simply aren't enough annotated spinal MRI datasets available, which makes it hard to train both GANs and CNNs effectively. When you're working with small datasets, models tend to either memorize the training examples (overfitting) or struggle when they encounter new patients.
GAN Training Instability	GANs have a reputation for being tricky to train. You might run into problems like mode collapse—where the generator gets stuck creating similar images over and over—or vanishing gradients that prevent the network from learning properly. These issues can result in synthetic images that just aren't good enough to be useful.
Evaluating Synthetic Data Quality	There is no standardized way to measure whether GAN-generated MRIs are clinically realistic or not. Currently, researchers often rely on visual inspection, which is subjective by nature and can differ from person to person.
Computational Demands	Training GANs and deep CNNs require long time and powerful GPUs which makes them limited for smaller research teams or institutions that don't have access to high-end computing resources.
Ethical and Privacy Questions	There are no standard ethical rules on use and regulation of the synthetic medical data, even though patient privacy is maintained.
Missing Benchmark Datasets	The standardized datasets for spinal tumors that researchers can use to test their models are not easily available which makes it difficult to validate the results.
The Black Box Problem	Both GANs and CNNs operate as "black boxes" which means it's hard to understand how exactly they are working. Due to this lack of transparency, they are little unreliable and trustworthy.

Future Research Directions

The integration of GAN with CNN has made significant progress in spinal cord tumor detection, but following are several key research directions that could further advance this field:

- To better understand tumors as complete, three-dimensional objects rather than a stack of disconnected slices, future tools are shifting toward working with three-dimensional imaging data. This transition is expected to substantially improve both the detection and precise mapping of tumors in spinal cord imaging.
- Bringing together data from different sources, such as multiple MRI sequences, CT scans, and genetic data will help in making comprehensive conclusions with respect to each individual case. Moreover, in order to improve transparency where the clinicians can understand which features affect model decisions, new interpretability tools must be developed.
- Hospitals are increasingly using federated learning, which allows institutions to collaborate and train models without sharing patient private and sensitive information. Along with this, the requirement for establishing standards for evaluating the quality of synthetic images generated by GANs also becomes important.
- Researchers need access to large, shared datasets of spinal MRI scans with verified diagnoses to properly validate their models. In addition, the community needs to set up some standards, such as uniform metrics for performance evaluation, procedures for image generation and ways of measuring accuracy so everyone can compare their respective results.

Conclusion

In this study, we shave systematically examined the performance of conventional CNNs and GAN augmented CNNs for the crucial task of spinal cord tumor detection. While conventional model can identify meaningful patterns from the limited and uneven datasets common in rare spinal diseases, their performance eventually becomes unstable. The main problem is scarcity of annotated cases and real data. Due to this, it is difficult to generalize the model and convergence becomes unstable. GAN-based augmentation helps by generating additional synthetic MRI scans that look realistic and are anatomically plausible. These artificial images are generated such that they reflect to real patient images like differences in tumor size, shape, contrast, and spatial context that reflect the diversity. These synthetic images expand the training distribution and help the model to learn from a broader range of instances. The proposed study also discussed challenges in building such models. Training GANs is computationally intensive and can run into problems like mode collapse or training instability. Most importantly, the value of synthetic data depends entirely on its quality.

References

1. J. L. R. de Paiva et al., “The Role of MRI in the Diagnosis of Spinal Cord Tumors,” Seminars in Ultrasound, CT and MRI, vol. 44, no. 5. W.B. Saunders, pp. 436–451, 01-Oct-2023.
2. B. Ainapure, S. Kulkarni, and M. Chakkaravarthy, “TriDx: a unified GAN-CNN-GenAI framework for accurate and accessible spinal metastases diagnosis,” Eng. Res. Express, vol. 7, no. 4, p. 045241, Oct. 2025.
3. B. Ainapure, D. S. Kulkarni, and D. Midhunchakkaravarthy3, “Generative Adversarial Networks in Medical Imaging: A Review of Architectures, Applications, and Challenges,” SGS - Eng. Sci., vol. 1, no. 4, p. LGPR, Oct. 2025.

4. Z. Qin, Z. Liu, P. Zhu, and Y. Xue, “A GAN-based image synthesis method for skin lesion classification,” Comput. Methods Programs Biomed., vol. 195, p. 105568, Oct. 2020.
5. A. Iqbal, M. Sharif, M. Yasmin, M. Raza, and S. Aftab, “Generative adversarial networks and its applications in the biomedical image segmentation: a comprehensive survey,” Int. J. Multimed. Inf. Retr., vol. 11, no. 3, pp. 333–368, Jul. 2022.
6. J. J. Jeong, A. Tariq, T. Adejumo, H. Trivedi, J. W. Gichoya, and I. Banerjee, “Systematic Review of Generative Adversarial Networks (GANs) for Medical Image Classification and Segmentation,” J. Digit. Imaging, vol. 35, no. 2, pp. 137–152, 2022.
7. K. T. Hong et al., “Lumbar Spine Computed Tomography to Magnetic Resonance Imaging Synthesis Using Generative Adversarial Network: Visual Turing Test,” Diagnostics, vol. 12, no. 2, 2022.
8. K. Margetis, J. M. Das, and P. D. Emmady, “Spinal Cord Injuries,” StatPearls, Jun. 2025.
9. O. Sarkar et al., “Multi-Scale CNN: An Explainable AI-Integrated Unique Deep Learning Framework for Lung-Affected Disease Classification,” Technologies, vol. 11, no. 5, Sep. 2023.
10. Q. Han et al., “DM-CNN: Dynamic Multi-scale Convolutional Neural Network with uncertainty quantification for medical image classification,” Comput. Biol. Med., vol. 168, p. 107758, Jan. 2024.
11. A. Al-Kubaisi and N. N. Khamiss, “A transfer learning approach for lumbar spine disc state classification,” Electron., vol. 11, no. 1, Dec. 2022.
12. G. Cohen and R. Giryes, “Generative Adversarial Networks,” in Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook, Third Edition, vol. 3, no. January, Neural information processing systems foundation, 2023, pp. 375–400.
13. S. Karthika and M. Durgadevi, “Generative Adversarial Network (GAN): A general review on different variants of GAN and applications,” in Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021, 2021.
14. A. Borji, “Pros and cons of GAN evaluation measures: New developments,” Comput. Vis. Image Underst., vol. 215, p. 103329, Jan. 2022.
15. I. S. A. Abdelhalim, M. F. Mohamed, and Y. B. Mahdy, “Data augmentation for skin lesion using self-attention based progressive generative adversarial network,” Expert Syst. Appl., vol. 165, p. 113922, Mar. 2021.
16. D. N. Sindhura, R. M. Pai, S. N. Bhat, and M. Pai M. M, “Deep learning-based automated spine fracture type identification with Clinically validated GAN generated CT images,” Cogent Eng., vol. 11, no. 1, Dec. 2024.
17. M. M. E. Yurtsever, Y. Atay, B. Arslan, and S. Sagiroglu, “Development of brain tumor radiogenomic classification using GAN-based augmentation of MRI slices in the newly released gazi brains dataset,” BMC Med. Inform. Decis. Mak., vol. 24, no. 1, p. 285, Dec. 2024.
18. R. Polattimur, E. Dandıl, M. S. Yıldırım, and U. Şenol, “Fully Automated Segmentation of Cervical Spinal Cord in Sagittal MR Images Using Swin-Unet Architectures,” J. Clin. Med., vol. 14, no. 19, p. 6994, Oct. 2025.
19. L. Weninger, “Deep Learning-based Analysis of Diffusion MRI Image Data from Multicenter Studies and from Glioma Patients.” .