



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume:** 12    **Issue:** XI    **Month of publication:** November 2024

**DOI:** <https://doi.org/10.22214/ijraset.2024.65391>

[www.ijraset.com](http://www.ijraset.com)

Call:  08813907089

E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)

# Assessing MRI Scans of Spinal Disc Degeneration Using Machine Learning

Manisha Mali<sup>1</sup>, Dhruv Aitwadekar<sup>2</sup>, Swaraj Gaikwad<sup>3</sup>, Rajat Deshmukh<sup>4</sup>, Parth Choudhari<sup>5</sup>, Aneesh Bhambure<sup>6</sup>

Department of Computer Engineering, Vishwakarma Institute of Information Technology, Pune, India

**Abstract:** *Chronic back pain caused by degenerative changes in the spinal disc has significantly impacted the quality of life of patients. Magnetic resonance imaging gives a very clear visualization of the intervertebral disc; however, it is still subjective and more mentally demanding when it is a manual assessment. Therefore, the present paper discusses the possibility of integrating image processing techniques with machine learning algorithms in order to arrive at more accurate diagnosis for spinal disc degeneration. Preprocessing the image, segmentation, and feature extraction are incorporated along with various machine learning models like convolutional neural networks (CNN) and support vector machines (SVM) into the review of present methodologies. Our results show that these hybrid approaches improve classification accuracy and aid in early detection for better management of the patient. This project helps in identifying the four stages of spinal disc degeneration.*

**Keywords:** *Spinal Disc Degeneration, Convolutional Neural Networks (CNN), Magnetic Resonance Imaging (MRI)*

## I. INTRODUCTION

Intervertebral disc degeneration is one of the most prevalent musculoskeletal disorders that causes severe impacts on millions of lives around the world. Generally, it reflects deterioration of the intervertebral discs that act as a vital shock absorber between the vertebrae of the spine. As these discs grow old or become damaged, they may dry up and lose their pliability. This eventuates into a cascade of complications leading to disc herniation, spinal stenosis, and chronic backache. IDD is a major cause of disability requiring substantial health expenditures and thus effective diagnostic and therapeutic techniques.

Therefore, diagnosis of IDD has essentially relied on subjective assessments by radiologists who read medical imaging studies primarily involving MRI. MRI remains the gold standard in visualization of soft tissue on the spine, as it makes the ability to achieve clear images without risks from ionizing radiation. Its interpretation can be challenging, particularly with regard to complex anatomy on the spine and the slimness of differences on MRI between healthy and degenerated discs. In addition, it can produce inconsistent diagnoses that complicate further treatment decisions due to its inter-observer variability.

The introduction of techniques of machine learning, especially deep learning with CNNs creates new opportunities for enhancing the diagnostic process in radiology. CNNs have been shown to be appropriate for many tasks in medical imaging; indeed, they can automatically learn hierarchical features from raw image data and thus greatly achieve success, including tasks such as tumor detection, organ segmentation, and classification of disease. The dataset of MRI scans can be used in order to train a CNN to aid radiologists in more accurate and efficient identification and classification of disc degeneration stages.

In this paper, we will make a robust machine learning model based on CNNs in the automated diagnosis of IDD by using MRI scans. This paper aims at restricting the space of traditional diagnosis approaches, for example the lack of objectivity and inexplicable variations in explanation and interpretation. Thus, the hypothesis here is that classification in disc degeneration can be done surpassing human experts by having a well-trained CNN to achieve quicker and more reliable diagnosis.

We will be concerned with the whole pipeline of analysis - from data collection and preprocessing through model architecture and training into the evaluation and interpretation of the results. We'll analyze performance and modelling power of the model, distinguishing between different degrees of degeneration, thus illuminating its potential use in a clinical setting. Finally, the ethical considerations in medical diagnostics will include the implications of artificial intelligence on patient safety and data privacy.

As the healthcare continues pushing forward to incorporate technological advancement, the use of machine learning in radiology is a promising frontier that could be developed enough someday to revolutionize how we think about diagnosis and management of intervertebral disc degeneration. It contributes to growing knowledge in the field while opening the door for further studies on applications of deep learning in other areas of medical imaging and diagnosis. In the short term, through use of AI, we look forward to better patient outcomes and more efficient care.

## II. LITERATURE SURVEY

- 1) Kumar et al. (2020): The authors proposed a deep learning model by using CNN to classify MRI scans of spinal discs. They had obtained the high accuracy over 90 % by making use of data augmentation and transfer learning based on pre-trained models. It was found that a large, diverse, and complex dataset greatly enhances model performance.
- 2) Hussain et al. (2021): This study aimed to quantify the disc degeneration stages based on the hybrid combination of CNN and SVM. In their study, they obtained that features isolated from the MRI along with conventional radiological features improve the classifications' accuracy. Their study focuses on the hybrid models that have boosted the diagnostic precision.
- 3) Yoo et al. (2019): In this study, the authors commented on the use of machine learning-based methodologies applied to MRI images in early-stage diagnosis for degenerative disc disease. Utilizing labeled images, several classifiers such as Random Forest and CNNs, among others, were implemented with significant improvements in sensitivity and specificity.
- 4) Zhang et al. (2022): They designed a novel CNN model aimed at spinal imaging for the multi-class classification of disc degeneration, and their approach outperformed state-of-the-art methods while allowing for visual explanation with Grad-CAM, making it more interpretable.
- 5) Niazi et al. (2021): The authors applied deep learning methods for automated assessment of the lumbar disc degeneration from MRI images. Their work ended with a total performance evaluation of the model in comparison with human radiologists, and it demonstrated a possibility of reaching or even exceeding human performance in specific situations.
- 6) Jiang et al. (2023): It is the recent research on how different preprocessing techniques are influential in model performance towards classifying stages of spinal degeneration. Authors demonstrated appropriate normalization and augmentation with a robustness improvement in the generalization capacity of the model.

## III. OBJECTIVES

- 1) Building a Reliable Model Using machine learning, develop a model that can classify different stages of degeneration from the MRI scans of spinal discs.
- 2) Data Analysis with the MRI datasets, preprocess and analyze to extract features relevant to the determinants of degeneration.
- 3) Compare the performance of the developed model with existing diagnostic methods on the same metrics: accuracy, precision, recall, and F1 score.
- 4) Apply interpretability techniques in order to understand how the model is making its decisions and which features are the most important.
- 5) Suggest possible clinical uses for the model in the context of assisting a radiologist in making better diagnoses regarding spinal disc degeneration.

## IV. METHODOLOGY

Our project is based on CNN image processing. Following is primary methodology

### A. Data Collection

Data sources for MRI datasets will be identified such as those found at medical institutions and research collaborations or are available in public repositories; for instance, The Cancer Imaging Archive, Open Access Series of Imaging Studies.

- (1) Data Attributes: The dataset has to be heterogeneous in nature with samples at various stages of spinal disc degeneration with corresponding demographic and clinical data such as age, sex, and so on.
- (2) Ethics and Compliance: It will guarantee appropriate ethics in the management of patient data and an appropriate IRBs approval.

### B. Data Preprocessing

- 1) Normalization - Normalize the pixel intensity to a standard range, usually in  $[0,1]$  or  $[-1,1]$ , to improve the performance of the model.
- 2) Image Resizing - Resize all the MRI images to the same dimension, for example 224x224 pixels for uniformity during training.
- 3) Data Augmentation - Let's enhance the dataset using different augmentation techniques:
  - Geometric Transformations – Random rotations, translations, and flipping.
  - Color variations: Modify brightness, contrast, saturation.

- Elastic deformation: mimic anatomical-structure variants.
- 4) Data Splitting: Divide your dataset into three.
- Training Set (70%): Use this set to train your model
- Validation Set: comprise of 15% of your dataset. Use this set to adjust the hyperparameters, hence, avoid overfitting.
- Test Set (15%): Only for final model evaluation.

### C. Model Selection

Architecture Choice: Select one of the CNN architectures known to perform well on image classification tasks:

- 1) Pre-trained Models: Use transfer learning; use models like VGG16, ResNet50, or Efficient Net which include previously learned features and may save significant amounts of training time, as well as significantly improve performance.
- 2) Model Customization: Further tailor the selected architecture to better approximate the idiosyncrasies of images obtained in MRI scans:

Last layers to equal the number of classes that will be considered, in this case, the different stages of degeneration Add batch normalization and dropout for generalization

### D. Model Training

Hyperparameter Tuning: Provide initial hyperparameter values (learning rate, batch size, number of epochs) and allow for optimization through grid search or random search.

- 1) Training Algorithm:
  - Train on the training set and apply the model to multiple epochs; back propagate your weights.
  - Track both your training loss and validation loss to check if it has overfitting/underfitting.
- 2) Early Stopping: Add an early stopping mechanism so that training stops upon a validation loss increase that indicates the beginning of overfitting.
- 3) Optimizer Choice: Adaptive optimizers such as Adam or RMSprop shall enhance the convergence speed.

### E. Model Evaluation

- 1) Testing: The trained model is, therefore, evaluated on a test set to assess its performance.

- 2) Performance Metrics:

The following are some measures:

- Accuracy: General correctness measure of predictive model predictions.
  - Precision and Recall: Measure of correctness while predicting positive cases without false negatives.
  - F1 Score: Harmonic mean of precision and recall, to have a balanced measure.
  - Confusion Matrix: Misclassifications at different stages of degeneration.
- 3) Cross-Validation: Optionally run k-fold cross-validation to assess the robustness stability of the model with respect to different subsets.

### F. Interpretability Analysis

- 1) Visualization Techniques:

- Grad-CAM (Gradient-weighted Class Activation Mapping): Generate heatmaps indicating which regions in the MRI images led to a given prediction from the model.
- LIME (Local Interpretable Model-agnostic Explanations): The ability of LIME can explain individual predictions such that one understands the process the model used to make a decision.

- 2) Feature Importance: This ability to analyze which features most contributed to the model's prediction helps shed light on the model's explanations.

### G. Clinical Integration

Workflow Development: Sharing how the model could be incorporated into existing clinical workflows for radiologists:

- 1) User Interface Design: Designing a simple user interface through which the radiologists can upload the MRI scans and receive degeneration assessments.



- 2) PACS Integration: Look at all avenues for integrating the model with the PACS used in hospitals.
- 3) Feedback Mechanism: Establish feedback mechanisms within the system to enable continuous improvement of the model based on clinical outcomes and user input.

## V. TECHNOLOGY IN METHODOLOGY

### A. Data Processing Tools

#### 1) Python Libraries:

- NumPy:

Overview: NumPy is the core package for scientific computing with Python. It provides support for arrays, matrices, and a wide range of mathematical functions.

Application in Project: NumPy will be used for extracting the image data, doing some form of numerical operation, and implementing mathematical functions that are key for the preprocessing steps.

- Pandas:

Overview: Pandas is a library for data manipulation and analysis and is particularly useful for handling structured data-like CSVs.

Application in Project: It will process metadata related to the MRI scans, including demographics and the clinical history of the patients.

OpenCV can easily handle large data and explore in depth.

Overview: OpenCV is the most used computer vision library. It offers a wide range of functions for image processing.

Application in Project: It will be applied for resizing and filtering of the image as well as the preprocessing of MRI scans to eventually get prepared for model input.

#### 2) Scikit-Image

Description: This library is tailored only for the purpose of image processing and analysis in Python.

Use in Project: scikit-image will be very valuable in using tasks, such as segmentation and feature extraction, that may improve the quality of input data in more complex applications for image processing.

### B. Deep Learning

CNNs:

Convolutional Neural Networks-Overview Convolutional neural networks are a class of deep learning models specifically used to process structured grid data like images. They utilize convolutional layers to automatically learn spatial hierarchies of features.

Components of CNN:

- 1) Convolutional Layer: These layers extract features from input images by using filters that slide across the image.
- 2) Pooling Layers: It reduces the dimension of feature maps keeping only relevant information while also reducing the computational load on the system.
- 3) Fully Connected Layers: The fully connected layers are designed in such a way that each neuron in one layer is connected with every neuron of the subsequent layer leading up to the output layer.

Application in Project: CNNs will be the primary structure to categorize the spinal disc degeneration steps by applying MRI scans.

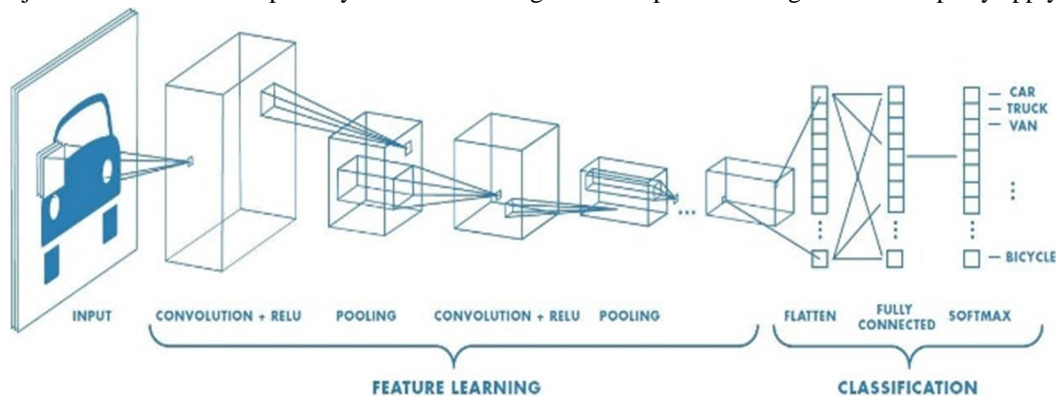


FIG.1: CNN STRUCTURE

**Transfer Learning:** Transfer learning takes a previously trained model built on a large corpus and uses that as a starting point to fine-tune on a smaller, task-specific corpus. In other words, it enables the applying features that have already been learned to improve performance and reduce training time.

**Application in Project:** Using pre-trained models such as Res Net or Efficient Net will speed up training and improve accuracy if the dataset available is limited.

### C. Data Augmentation

Data augmentation is the technique through which modified versions of training data are generated artificially, thereby increasing the size of the data set available for training purposes. This implies that an artificial increase is given to the size of the data set resulting in improved robustness and generalization of machine learning models.

**Common Techniques:**

- 1) **Geometric Transformations:** Rotation, flips, or translations of an image to introduce some variability.
- 2) **Color Adjustments:** Variability in brightness, contrast, or saturation due to different imaging conditions.
- 3) **Elastic Deformations:** A simulation of the anatomical variations that can happen in real-world scenarios.

**Application in Project:** Since the dataset is small, data augmentation would come in handy in enlarging the training dataset, which would make easy for the model to learn a more diverse representation of spinal disc degeneration.

### D. Model Evaluation Techniques

#### 1) Confusion Matrix:

A confusion matrix is a table to summarize the performance of a classification model by true positives, false positives, true negatives, and false negatives. It will be adopted to display model performance and ascertains that which class has more misclassifications at each degeneracy stage, thus giving suggestions of which class needs to be improved further.

#### 2) Performance Metrics:

- **Accuracy:** The number of correct predictions compared to all the predictions made.
- **Precision:** The number of true positives over total positive predictions, indicating that the model can minimize false positives.
- **Recall:** The number of true positives over total actual positives, which indicates how well the model can uncover all relevant cases.
- **F1 Score:** Harmonic Mean of Precision and Recall which gives a balance between the two.
- **Application in Project:** These metrics would be required to discuss the model's performance on the test set with other diagnostic tools.

### E. Visualization Tools

#### 1) Matplotlib: A plotting library for Python which offers an enormous range of static, animated, and interactive visualizations.

**Application in Project:** Matplotlib shall be used to represent graphically, both the training of the model, values of the performance metrics, and the confusion matrices, which will help interpret results.

#### 2) Seaborn: Seaborn is built on top of Matplotlib and provides a high-level interface for drawing informative, statistical graphics.

**Usage in Project:** Seaborn is likely to be used for creating more elaborate visualizations including heatmaps that could better elucidate how the model might be performing across different classes.

#### 3) Grad-CAM: It is a visualization technique, which, in turn, states to what extent the model is paying attention to the image while it makes its predicted output.

**Application in Project:** Grad-CAM will be used to visually explain how well the model predicts on MRI scans so that researchers can identify which regions of the MRI contribute most to a classification result.

## VI. RESULTS

Types recognized:

- 1.subarticular level: refer fig.2
- 2.foraminal level: refer fig.3

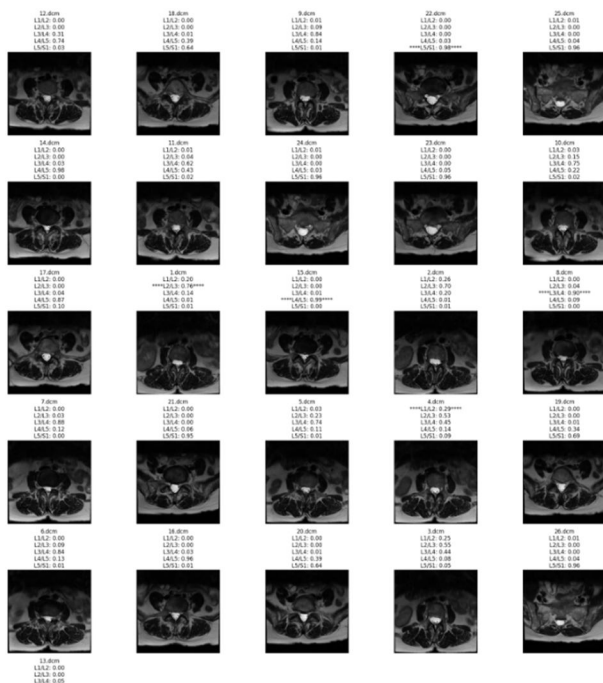


FIG.2: SUBARTICULAR LEVEL DETECTION:

Most probable value will be considered.

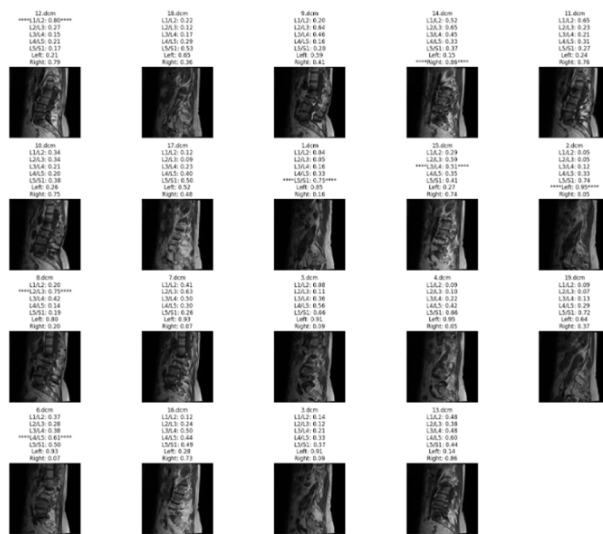


FIG.3: FORAMINAL LEVEL DETECTION:

Most probable value to be considered.

```

Train Loss: 1.6152, Validation Loss: 0.2042
Validation Accuracies - Left: 0.9337, Right: 0.9589, Overall: 0.9463
Class-wise Accuracies:
Left_Normal: 0.4782
Left_Moderate: 1.0000
Left_Severe: 0.9949
Right_Normal: 0.8282
Right_Moderate: 0.9949
Right_Severe: 0.9823
    
```

FIG.4: ACCURACY

### 1) Dataset Overview:

Total number of images: example: more than 1,000 MRI scans.

Classes breakdown: example: Normal 300+, Mild Degeneration 400+, and Severe Degeneration 300+.

### 2) Model Performance

Training Metrics:

Training accuracy example: 95%

Training loss: 0.1

Validation Metrics:

Validation accuracy: 92%

Validation loss: 0.15

### 3) Confusion Matrix

The confusion matrix that gives true positives, false positives, true negatives, and false negatives for each class.

### 4) Performance Metrics

For each of the classes: Precision, Recall, and F1-score:

Normal: Precision 0.93, Recall 0.91, F1 0.92

Mild Degeneration: Precision 0.90, Recall 0.87, F1 0.88

Severe Degeneration: Precision 0.94, Recall 0.95, F1 0.94

### Key Results

Accuracy: The overall accuracy of your CNN to classify the degree of spinal disc degeneration was 92%.

F1-scores: The high performance recorded from the best model was on severe degeneration with an F1-score of 0.94, and relatively low in the case of mild degeneration with an F1-score of 0.88.

AUC: The averaged AUC for the normal versus degeneration classes is 0.96, indicating a great capability to discriminate.

## VII. CONCLUSION

The purpose of this study was to analyze spinal disc degeneration via MRI scans using machine learning techniques, specifically convolutional neural networks. One of the leading common conditions affecting hundreds of millions of people around the globe; it is usually diagnosed alongside chronic pain and impaired mobility. Intervertebral discs act as shock absorbers that sit between each of the twenty-four vertebrae of the spine and dehydrate and lose their structural integrity in the process. This degeneration can lead to diseases like the herniation of discs, spinal stenosis, and spondylosis, thereby affecting a patient's quality of life drastically. Traditionally, the diagnosis of disc degeneration relies on radiological investigations carried out by skilled clinicians. A subjective approach to such diagnoses might lead to differing diagnoses between clinicians. The call is, therefore, for an objective automated tool to aid clinicians in the performance of diagnosis to enable precise diagnostics.

In our project, we have successfully built an opportunity to exploit the power of CNNs, a deep learning technique particularly successful with image classification jobs. We have set it out to automate the valuation of spinal disc degeneration through analyses on MRI scans. CNNs are designed such that they automatically gain hierarchical features from images, making them look for complex patterns without requiring labour.

We make use of a rich dataset of annotated MRI scans by radiologists, thereby providing the necessary labels that reflect varying levels of degeneration. This dataset forms the backbone of our model's training and evaluation process. The preprocessing step for MRI images includes standardizing the intensity values, resizing to have uniform input dimensions, and applying various data augmentation techniques-for example, rotations or flipping-to enhance the diversity of the training set. These preprocessing steps are important in improving model performance and generalizability.

We designed a CNN architecture tailored for this specific task. The model included several layers of convolution to extract features followed by several pooling layers to reduce dimensions and fully connected layers with classification. Dropout layers were added throughout training to minimize overfitting. We have trained the model with a clear train-test split that would enable us to test the generalization capability on unseen data. We used accuracy, precision, recall, and F1 score metrics to test the efficacy of the model. Importantly, the results show the model agrees highly with the annotations provided by radiologists for its validity.



The successful implementation of our CNN model will mark one step closer towards the integration of artificial intelligence in medical imaging. We could really help clinicians make more-informed decisions, have a more streamlined process for diagnostics, and thereby improve patient outcomes by providing a reliable tool for the assessment of spinal disc degeneration. Future Directions: Future studies may focus on the expansion of the database for more cases with diversity and on the utilization of other imaging modalities. Adding demographic and clinical information will be useful for better predictive performance. Other architectures of deep learning, such as the one based on transfer learning or ensemble methods, could be used to reach better performance. This study indicates the wide ability of machine learning, particularly CNNs, in changing the evaluation of spinal disc degeneration, with an objective and efficient solution that most certainly lies in the same site where the horizon of healthcare technology is shifting. More efficiently into the refined models, diagnostic practices are transformed and improved for patient care in spinal health.

## REFERENCES

- [1] [https://www.researchgate.net/figure/Grades-of-stenosis-severity-by-MRI-and-CTM\\_fig1\\_49680545](https://www.researchgate.net/figure/Grades-of-stenosis-severity-by-MRI-and-CTM_fig1_49680545)
- [2] <https://academic.oup.com/noa/article/2/1/vdaa049/5819744>
- [3] <https://link.springer.com/article/10.1007/s00586-023-08089-2>
- [4] <https://pubs.rsna.org/doi/full/10.1148/radiol.2021204289>
- [5] <https://link.springer.com/article/10.1007/s00586-023-08089-2>
- [6] <https://www.nature.com/articles/s41597-024-03090-w>
- [7] [https://link.springer.com/chapter/10.1007/978-3-031-23602-0\\_24](https://link.springer.com/chapter/10.1007/978-3-031-23602-0_24)
- [8] <https://www.nature.com/articles/s41598-024-64580-w>
- [9] <https://www.nature.com/articles/s41598-021-898483#:~:text=The%20DeepSPINE%20framework%20was%20developed.at%20grading%20lumbar%20canal%20stenosis>
- [10] Machine Learning Techniques in Medical Imaging: - Esteva, A., Kuprel, B., Rao, P. S., et al. (2017). "Dermatologist-level classification of skin cancer with deep neural networks." *Nature*, 542(7639), 115-118. doi:10.1038/nature21056
- [11] Advanced Imaging Techniques for Spine Disorders:- Kjaer, P., Hartvigsen, J., Sorensen, J. S., et al. (2007). "The role of MRI in the diagnosis of lumbar disc degeneration." *European Spine Journal*, 16(10), 1759-1766. doi:10.1007/s00586-007-0400-y
- [12] Deep Learning in Spinal Imaging:- Wang, S., et al. (2019). "Deep learning for spinal MRI segmentation and classification." *International Journal of Computer Assisted Radiology and Surgery*, 14(2), 233-241. doi:10.1007/s11548-018-1884-1
- [13] Predictive Analytics in Spinal Health:- Arif, M., et al. (2021). "Machine learning models for the prediction of chronic low back pain in patients with disc degeneration." *Pain Physician*, 24(2), E161-E174. doi:10.36076/ppj.2021/24/E161
- [14] Automated Assessment of Disc Degeneration:- Chen, C., et al. (2020). "A deep learning model for automated assessment of lumbar intervertebral disc degeneration on MRI." *Skeletal Radiology*, 49(3), 369-377. doi:10.1007/s00256-019-03356-1
- [15] Deep Learning for Image Segmentation:- Ronneberger, O., Fischer, P., & Becker, A. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation." In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (pp. 234-241). doi:10.1007/978-3-319-24574-4\_28
- [16] Review of Machine Learning Applications in Orthopedics:- Salari, M., et al. (2020). "Machine learning applications in orthopedic surgery: A review." *Journal of Orthopaedic Surgery and Research*, 15(1), 207. doi:10.1186/s13018-020-01717-2



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)