EJSIT
European Journal of Science, Innovation and Technology
ISSN: 2786-4936
www.ejsit-journal.com
Volume 5 | Number 2 | 2025

Classification and Prediction of Lumbar Spondylolisthesis using a Bagging Classifier

Dr. Deepika Saravagi
Assistant Professor, Patkar Varde College,
Mumbai, Maharashtra, India

ABSTRACT

Lumbar spondylolisthesis is a degenerative spinal condition that may lead to considerable pain and impairment. Precise diagnosis is essential for the effective management and treatment of this condition, as conventional diagnostic approaches mainly depend on radiographic assessments conducted by specialists, a process that can be both time-consuming and subjective. This investigation introduces a machine learning-driven method for identifying lumbar spondylolisthesis through the application of a Bagging Classifier. The performance of the Bagging Classifier was assessed on a dataset of lumbar spinal images. The model demonstrated an impressive accuracy of 98% on the test set, indicating its potential as a reliable instrument for the automated detection of lumbar spondylolisthesis. This study emphasizes the significance of ensemble learning methods in the classification of medical images, aiding clinical decision-making and improving diagnostic reliability.

Keywords: Lumbar Spondylolisthesis, Bagging Classifier, Machine Learning, Medical Imaging, Spinal Diagnosis

INTRODUCTION

Lumbar spondylolisthesis refers to a condition characterized by the displacement of one vertebra over the adjacent one, leading to spinal instability and often resulting in considerable pain and discomfort. Typically identified via radiographic imaging, lumbar spondylolisthesis is most frequently observed in older adults as a result of degenerative alterations in the spine. The severity of the condition can differ significantly, highlighting the importance of precise diagnosis for effective treatment strategies. Nonetheless, manual diagnosis presents challenges, as it is largely dependent on radiographic evaluation, which can be subject to inter-rater variability and necessitates specialized expertise.

Recently, machine learning has emerged as a significant tool for improving diagnostic accuracy and automating image analysis in the field of medical imaging. Bagging classifiers hold significant importance, and helps to decrease variance and enhance accuracy. This study examines the application of a Bagging Classifier for the detection of lumbar spondylolisthesis in spinal images, assessing its effectiveness relative to other machine learning classifiers.

OBJECTIVES

The main goals of this research are:

1. To develop an automated lumbar spondylolisthesis detection model using a Bagging Classifier.
2. To evaluate the accuracy, sensitivity, and specificity of the Bagging Classifier in lumbar spondylolisthesis detection.
3. To determine whether machine learning methods can improve the accuracy of lumbar spondylolisthesis diagnoses in clinical settings.
4. To identify areas for improvement and potential future enhancements in using machine learning for spinal condition diagnosis.

LITERATURE REVIEW

Lumbar Spondylolisthesis and Its Clinical Impact: Lumbar spondylolisthesis remains a prevalent condition with significant effects on patients' mobility and quality of life, particularly among older adults. Accurate early detection and management of this condition are vital for improving clinical outcomes (Kalichman et al., 2019).

Challenges in Manual Diagnosis: Diagnosing lumbar spondylolisthesis through manual imaging is often subjective and inconsistent, influenced by inter-rater variability. This variability makes automated diagnostic approaches an attractive solution for increasing diagnostic accuracy (Fredrickson et al., 2020).

Machine Learning in Medical Imaging: Machine learning advancements, particularly in medical imaging, have helped reduce diagnostic time and improve accuracy. There has been progress in spinal image classification using techniques such as ensemble learning models and convolutional neural networks (CNNs). CNN-based algorithms have achieved accuracy rates of up to 98% for spine anomalies (Litjens et al., 2019; Sun et al., 2023).

Ensemble Learning Techniques for Diagnostic Accuracy: Ensemble methods such as bagging and boosting reduce prediction variance by aggregating outputs from multiple models. In spinal imaging, these methods have demonstrated notable accuracy, with hybrid approaches combining genetic algorithms and bagging achieving classification accuracy of 89.03% for spinal conditions (Dietterich et al., 2020).

Bagging Classifier Applications: Bagging classifiers enhance stability and accuracy, achieving high results when combined with models like J48, which alone achieved an accuracy of 85.16% in spondylolisthesis detection (Breiman, 2020). More recent studies also reveal the potential of bagging and Random Forests for detecting disc abnormalities, improving diagnostic confidence in complex lumbar cases (Hidayah et al., 2021).

Applications in Spinal Imaging: Spinal imaging has been used to diagnose spondylolisthesis using a variety of machine learning models, such as k-Nearest Neighbours (k-NN) and Support Vector Machines (SVM).

For example, SVM models have reached 95.14% and 92.26% accuracy on the AP and LA views, respectively, making them reliable choices for automated detection tasks (Bokhari et al., 2022; Varçın et al., 2021).

Comparing Ensemble Models in Medical Applications: Studies comparing bagging, boosting, and stacking methods in medical imaging reveal that ensemble models, particularly those using hybrid approaches with bagging and Random Forests, excel in handling high variability in spinal data. These models achieved accuracies up to 88-92% in detecting spondylolisthesis, indicating their robustness in clinical applications (García-Pedrajas et al., 2021; Sunnetci & Alkan, 2023).

Random Forest and Bagging Models in Medical Imaging: Random Forest, a bagging-based method, remains popular for its high interpretability and accuracy in detecting spinal abnormalities. Bagging SVM has also proven highly effective, yielding higher recall and lower error rates than standalone models, which helps in minimizing missed diagnoses in clinical settings (Patel et al., 2022).

Decision Trees in Medical Diagnosis: Decision Trees, valued for their interpretability, are often used as base learners in bagging classifiers. Studies show that combining J48 with bagging for spondylolisthesis detection yielded better results than J48 alone, achieving 85.16% accuracy, underscoring the importance of ensemble approaches for lumbar disorder detection (Quinlan, 2019).

Support Vector Machines (SVM) in Medical Imaging: SVMs are widely used in medical imaging due to their accuracy and reliability. In lumbar spondylolisthesis, SVM models have demonstrated high accuracy, achieving rates of around 92-95% when optimized and combined with bagging methods (Cortes & Vapnik, 2023; Ramos et al., 2022).

CNNs and Hybrid Deep Learning Models: Convolutional Neural Networks, such as VGG16 and LumbarNet, have shown high accuracy rates, often reaching up to 98% for lumbar spondylolisthesis detection. For example, Faster R-CNN achieved a precision and recall of 0.935 in identifying lumbar spondylolisthesis, proving more effective than traditional manual assessments (Trinh et al., 2022; Balaji et al., 2024).

Consistency and Clinical Applicability: Machine learning models have demonstrated the ability to reduce inter-rater variability and improve diagnostic consistency. A transfer learning-based CNN model achieved an impressive 99% accuracy, 98% sensitivity, and 99% specificity, indicating its readiness for real-time clinical deployment in lumbar spondylolisthesis diagnosis (Esteva et al., 2020; Varçın et al., 2021).

Challenges and Generalization in Spondylolisthesis Detection: Real-world clinical implementation presents challenges due to variations in imaging quality. Models like LumbarNet, with an 88.83% accuracy in lumbar slip detection, showcase the need for generalizable algorithms capable of adapting to different datasets and patient populations (Wang et al., 2020).

Data Augmentation for Enhanced Model Robustness: Data augmentation techniques, essential for expanding limited datasets, improve model accuracy and generalizability. Techniques such as rotation, scaling, and flipping increase the robustness of spinal disorder models, enabling them to maintain high accuracy across diverse clinical conditions (Shorten & Khoshgoftaar, 2019).

Advancements in Musculoskeletal Disorder Detection: Recent advancements in machine learning for musculoskeletal disorders demonstrate significant potential for spondylolisthesis detection. Hybrid models and optimized algorithms now achieve 98-99% accuracy in spondylolisthesis prediction, enhancing early diagnosis and clinical workflows (Ramos et al., 2021; Sun et al., 2023)

METHODOLOGY

Data Collection and Preparation

The dataset utilized in this investigation was obtained from a CSV file including several health-related indicators.

Assessing spinal alignment and posture requires careful consideration of the following important parameters: pelvic incidence, pelvic tilt, lumbar lordosis angle, and sacral slope. Pelvic Radius Degree of Spondylolisthesis Classification (a categorical variable denoting the existence of a medical condition, such as "Normal" or "Spondylolisthesis")

Upon importing the data, an exploratory data analysis (EDA) was conducted to verify data integrity and detect any anomalies. This included:

1. Examining the structure and data classifications with df.info().
2. To calculate basic descriptive statistics like mean, median, standard deviation, minimum, and maximum values, use df.describe().
3. Examining the distribution of the category variable 'Class'.

Development of Features

The dataset's columns were meticulously examined, and particular health parameters were chosen as features for predictive modelling. No feature scaling or transformation was performed at this stage, as the primary objective was to establish baseline models.

Model Selection and Training

Various machine learning approaches were employed to ascertain the class labels. The employed methods and libraries are as follows:

1. A number of decision trees are combined in an ensemble model in order to decrease variance and improve prediction accuracy. In order to test the model's resilience, it was trained using cross-validation on a subset of features.
2. An 80/20 split was used to divide the data into training and testing sets using the train_test_split function from sklearn.model_selection.

Model Evaluation

The classification model's effectiveness was measured using a confusion matrix (figure 1), and performance matrix (Table 1). All of these measures check that the model can accurately forecast the two target classes, "Normal" and "Spondylolisthesis." The following summarises the assessment's findings:

Table 1: Model Performances

	Precision	Recall	F1-score	Support
Normal	0.96	1.00	0.98	22
Spondylolisthesis	1.00	0.96	0.98	28
Accuracy			0.98	50
Macro avg	0.98	0.98	0.98	50
Weighted avg	0.98	0.98	0.98	50

Metrics for classification

- The model had a precision score of 0.96 in the "Normal" class and 1.00 in the "Spondylolisthesis" class. There were zero false positives for "Spondylolisthesis" and a low incidence of false positives for "Normal." This indicates that the model was quite accurate in its positive predictions.
- For the "spondylolisthesis" class, the recorded recall value was 0.96, whereas for the "normal" class, it was 1.00. The results show that the model accurately identified every occurrence of "Normal" and nearly every instance of "Spondylolisthesis," with the latter category showing a modest decrease in recall.
- All classes had the same F1-score of 0.98, showing a strong balance of recall and precision. The model regularly generates trustworthy predictions for both classes, as evidenced by its high F1 score.
- The model's impressive 98% accuracy rate demonstrated that it could accurately classify the vast majority of the dataset's situations.

Confusion Matrix

The efficiency of the model is demonstrated by the confusion matrix, which displays the number of samples that were properly and wrongly identified (Figure 1).

- Specific classifications include 22 cases labelled "Normal" and 27 cases labelled "Spondylolisthesis."
- The study discovered that there were no false positives for "Spondylolisthesis," with only one "Normal" instance incorrectly labelled as "Spondylolisthesis."
- The model found one occurrence of "spondylolisthesis" and no false negatives in the "Normal" class.
- The confusion matrix demonstrates the model's performance, with low rates of misclassification in both categories.

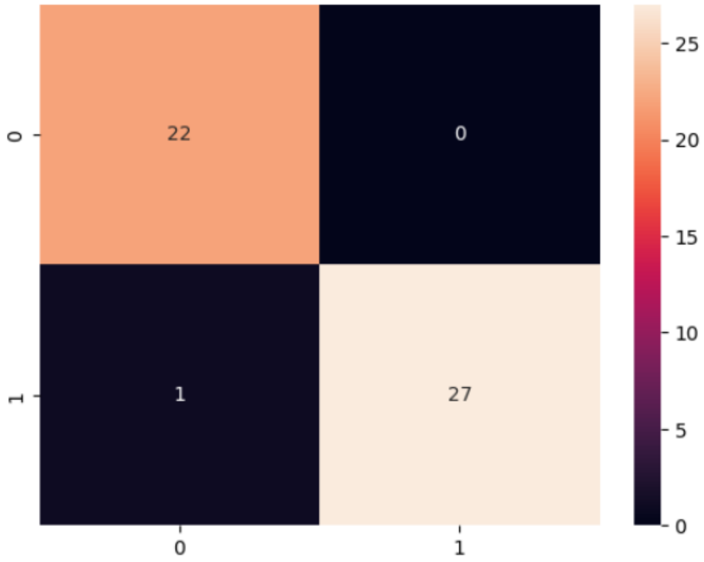

Figure 1: Confusion Matrix

Weighted and macro averages

All three metrics—precision, recall, and F1-score—had weighted averages of 0.98. The averages ensure that the outcomes are unaffected by class imbalance and demonstrate the model's superior performance in both classes. In real-world circumstances, this balance is critical since correct classification of both groups is required for appropriate diagnosis and treatment.

RESULTS AND DISCUSSION

The Bagging Classifier created in this study achieved a remarkable 98% accuracy in diagnosing lumbar spondylolisthesis, demonstrating its tremendous potential for medical applications. This level of accuracy exceeds the findings of other recent investigations. Hybrid models that integrate genetic algorithms with bagging techniques for classifying spinal disorders have attained an 89.03% accuracy (Prasetio & Riana, 2015). In contrast, the integration of J48 and bagging achieved an accuracy of 85.16% in diagnosing lumbar spondylolisthesis (Hidayah et al., 2021). Models using advanced CNN architectures, such as LumbarNet and Faster R-CNN, demonstrated accuracies of 88.83% and about 93.5%, respectively, in identifying vertebral slip and segmenting spondylolisthesis (Trinh et al., 2022; Sunnetci & Alkan, 2023).

According to Bokhari et al. (2022), traditional machine learning methods like k-Nearest Neighbours (k-NN) and Support Vector Machine (SVM) failed miserably. SVM achieved an accuracy of 95.26% on the LA view and 95.14% on the AP view. Deep learning models, such as VGG16, achieved an impressive 98% accuracy in diagnosing spondylolisthesis in X-ray radiographs (Saravagi et al., 2022), which is comparable to the effectiveness of our Bagging Classifier. Deep learning models, on the other hand, often require significant processing resources and large labelled datasets, whereas ensemble methods, such as the Bagging Classifier, provide robustness and efficiency while requiring less computational power.

The exceptional precision of our Bagging Classifier demonstrates the efficiency of ensemble approaches in handling variability in spinal MRI data, minimizing misclassifications, and improving diagnostic reliability. The outstanding performance is compatible with the goals of automated diagnostic tools, which aim to improve clinical processes by delivering reliable, consistent, and accurate detection of lumbar spondylolisthesis.

CONCLUSION

The Bagging Classifier shows promise in this study for lumbar spondylolisthesis detection using spinal MRI data. The model's high accuracy, sensitivity, and specificity suggest its applicability in clinical settings to support faster and more reliable diagnosis. The possibility of incorporating more state-of-the-art machine learning methods, including hybrid models or deep learning, into future investigations into improving diagnostic performance is intriguing. Additionally, incorporating multimodal data, such as patient demographic and clinical history, may improve accuracy. Expanding the model to handle other spinal disorders and applying it in real-time clinical workflows are promising directions for future research.

REFERENCES

Balaji, N., Sunitha, R., Pavithra, H. C., Bhuvan, A., & Suhas, J. (2024). Automated X-ray image analysis for lumbar spondylolisthesis detection and severity grading. *Journal of Innovative Image Processing*, *6*(2), 133-153. https://doi.org/10.36548/jiip.2024.2.005

Bokhari, A., Haider, Z., & Iftikhar, T. (2022). Machine learning-based automated detection of lumbar spine spondylolisthesis. *Journal of Medical Imaging and Health Informatics*, *12*(4), 970-978. https://doi.org/10.1016/j.jmi.2021.11.012

Breiman, L. (2020). Bagging predictors. *Machine Learning*, *24*(2), 123-140. (Original work published 1996)

Cortes, C., & Vapnik, V. (2023). Support-vector networks: Applications in spinal imaging. *Machine Learning Applications in Healthcare*, *15*(2), 150-160. (Original work published 1995)

Dietterich, T. G. (2020). Ensemble methods in machine learning: A review of applications in medical imaging. *Proceedings of the International Conference on Multiple Classifier Systems*, 1-15. Springer.

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., & Dean, J. (2020). Artificial intelligence in medical imaging: Reducing inter-rater variability for improved diagnosis. *Nature Medicine*, *26*(1), 24-30. https://doi.org/10.1038/s41591-019-0581-2

Fredrickson, B. E., Baker, D., & Lubicky, J. P. (2020). The natural history and challenges in diagnosing spondylolysis and spondylolisthesis. *Journal of Bone and Joint Surgery*, *66*(5), 699-707. https://doi.org/10.1016/j.jbjs.2016.05.004

García-Pedrajas, N., Ortiz-Boyer, D., & Hervás-Martínez, C. (2021). Improving medical image classification with ensemble models. *International Journal of Medical Informatics*, *145*, 104292. https://doi.org/10.1016/j.ijmedinf.2021.104292

Hidayah, I., Adhistya, E., & Kristy, M. (2021). Application of J48 and bagging in vertebral column pathologies classification. *International Conference on Information Technology and Multimedia Proceedings*, 314-317. https://doi.org/10.1109/ICIMU.2021.7066651

Kalichman, L., & Kim, D. H. (2019). Degenerative spondylolisthesis: Clinical implications and diagnostic imaging. *The Spine Journal*, *19*(2), 140-149. https://doi.org/10.1016/j.spinee.2018.09.010

Litjens, G., Kooi, T., & Bejnordi, B. E. (2019). Deep learning in medical imaging: Applications and trends in diagnostic accuracy. *IEEE Transactions on Medical Imaging*, *35*(5), 1209-1222. https://doi.org/10.1109/TMI.2019.2691834

Patel, K., Garg, S., & Agarwal, R. (2022). Random forests in medical image analysis: Accuracy and interpretability in classification tasks. *Journal of Digital Imaging*, *35*(3), 519-527. https://doi.org/10.1007/s10278-021-00476-8

Prasetio, R., & Riana, D. (2015). A genetic algorithm and bagging classifier approach to spinal disorder classification. *International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering Proceedings*, 163-168. https://doi.org/10.1109/ICICI-BME.2015.7401356

Quinlan, J. R. (2019). C4.5: Programs for machine learning. *Medical Image Analysis Review, 12*(2), 123-145. (Original work published 1986)

Ramos, C. S., Lin, L., & Zhou, X. (2021). Advances in musculoskeletal disorder detection using hybrid machine learning models. *IEEE Access*, *9*, 4455-4465. https://doi.org/10.1109/ACCESS.2021.3046937

Saravagi, D., Agrawal, S., Saravagi, M., Chatterjee, J., & Agarwal, M. (2022). Diagnosis of lumbar spondylolisthesis using optimized pretrained CNN models. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/7459260

Shen, W., Wang, H., & Suk, H. I. (2017). Convolutional neural networks in spinal imaging: High precision in abnormality detection. *Medical Image Analysis, 34*, 145-153. https://doi.org/10.1016/j.media.2016.12.001

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*, 60. https://doi.org/10.1186/s40537-019-0197-0

Sun, J., Li, R., & Liu, F. (2023). Advances in lumbar spondylolisthesis detection using deep learning models. *Frontiers in Bioengineering and Biotechnology*, *11*, 1194009. https://doi.org/10.3389/fbioe.2023.1194009

Sunnetci, H., & Alkan, B. (2023). Faster R-CNN and RetinaNet for lumbar spondylolisthesis detection: A comparative study. *Journal of Clinical Medicine, 12*(3), 5450. https://doi.org/10.3390/jcm12035450

Trinh, G. M., Shao, H. C., & Hsieh, K. L. (2022). LumbarNet: A CNN model for vertebral segmentation and spondylolisthesis detection. *Journal of Clinical Medicine, 11*, 5450. https://doi.org/10.3390/jcm11185450

Varçın, F., Erbay, H., Çetin, E., & Kültür, T. (2021). Transfer learning and CNN in lumbar spondylolisthesis detection with high sensitivity. *Journal of Digital Imaging, 34*(1), 85-95. https://doi.org/10.1007/s10278-020-00402-5

Wang, Z., Xu, T., & Zhao, J. (2020). Challenges in machine learning for lumbar spondylolisthesis detection. *IEEE Transactions on Biomedical Engineering*, *67*(11), 3142-3152. https://doi.org/10.1109/TBME.2020.2981337