

# Identification of L5 vertebra on lumbar spine radiographs using deep learning

Jeoung Kun Kim<sup>1,\*</sup>, Min Cheol Chang<sup>2,\*</sup> ,  
Wook Tae Park<sup>3</sup> and Gun Woo Lee<sup>3</sup> 

## Abstract

**Objective:** Deep learning is an advanced machine-learning approach that is used in several medical fields. Here, we developed a deep learning model using an object detection algorithm to identify the L5 vertebra on anteroposterior lumbar spine radiographs, and assessed its detection accuracy.

**Methods:** We retrospectively recruited 150 participants for whom both anteroposterior whole-spine and lumbar spine radiographs were available. The anteroposterior lumbar spine radiographs of these patients were used as the input data. Of the 150 images, 105 (70%) were randomly selected as the training set, and the remaining 45 (30%) were assigned to the validation set. YOLOv5x, of the YOLOv5 family model, was used to detect the L5 vertebra area.

**Results:** The mean average precisions 0.5 and 0.75 of the trained L5 detection model were 99.2% and 96.9%, respectively. The model's precision was 95.7% and its recall was 97.8%. Furthermore, 93.3% of the validation data were correctly detected.

**Conclusion:** Our deep learning model showed an outstanding ability to identify L5 vertebrae.

## Keywords

Deep learning, lumbar vertebrae, algorithms, X-rays, YOLOv5, machine learning

Date received: 6 July 2023; accepted: 14 December 2023

<sup>1</sup>Department of Business Administration, School of Business, Yeungnam University, Gyeongsan-si, Republic of Korea

<sup>2</sup>Department of Physical Medicine and Rehabilitation, Yeungnam University College of Medicine, Daegu, Republic of Korea

<sup>3</sup>Department of Orthopaedic Surgery, Yeungnam University College of Medicine, Daegu, Republic of Korea

\*These authors contributed equally to this work.

### Corresponding author:

Gun Woo Lee, Department of Orthopaedic Surgery, Yeungnam University Hospital, Yeungnam University College of Medicine, 170 Hyonchung-ro, Nam-gu, Daegu 42415, Republic of Korea.

Email: gwlee1871@gmail.com



## Introduction

Accurate and reliable spine numbering is critical for pathological diagnosis and pre-procedural or pre-operative planning. However, it can be difficult in patients with lumbosacral transitional vertebrae, especially when whole-spine images are not available.<sup>1</sup> There are two types of lumbosacral transitional vertebrae: 1) the sacralization of L5, a congenital spinal anomaly in which an elongated transverse process of the last lumbar vertebra fuses with the “first” sacral segment in varying degrees;<sup>1-3</sup> and 2) the lumbarization of S1, a spinal anomaly in which the first and second segments of the sacrum are not fused, and the lumbar spine consequently appears to have six (rather than five) vertebrae.<sup>1-4</sup>

The lumbar vertebrae can be accurately numbered using whole-spine radiography or whole-spine magnetic resonance imaging (MRI) by counting down from the C2 to the sacrum.<sup>5,6</sup> In cases without available full-spine imaging, it is frequently difficult to accurately number the lumbar vertebrae. Up to 32% of neurosurgeons report that they have performed spinal surgery at the wrong level at least once in their careers.<sup>7</sup> Therefore, prior to a spinal procedure or operation, clinicians should perform whole-spine imaging in patients with spine disorders to accurately identify the lumbar vertebral level.

Deep learning techniques have recently emerged as powerful methods to automatically learn feature representations from data.<sup>8-10</sup> In particular, these techniques can substantially improve object detection.<sup>11</sup> Object detection refers to determining whether there are any instances of objects from specified categories in an image; if there are, the spatial location and extent of each object instance are shown via a bounding box.<sup>11</sup> Object detection, which is the foundation of image comprehension and computer vision, also

serves as the basis for solving complex or high-level vision tasks such as segmentation, object tracking, image captioning, activity recognition, and event detection.<sup>12</sup> In the medical field, object detection is widely used to identify various pathologies, such as malignancy, pneumonia, and bony fractures, on medical images.<sup>13,14</sup> We hypothesized that object detection may similarly be applied to identify lumbar vertebral levels.

In the current study, we developed a deep learning model with an object detection algorithm to identify the L5 vertebra on lumbar spine radiographs, and evaluated its accuracy.

## Methods

### Participants

Participants aged  $\geq 20$  years who visited the spine center of Yeungnam University Hospital (Daegu, Republic of Korea) between June 2016 and December 2021 were retrospectively recruited. Both anteroposterior whole-spine and lumbar spine radiographs from each included participant were stored in the hospital picture archiving and communication system. The study protocol was approved by the Institutional Review Board of Yeungnam University Medical Center (IRB no. 2023-02-005) and followed the Helsinki Declaration of 1975 as revised in 2013. The requirement for written informed consent was waived because of the study's retrospective nature. The authors have de-identified all patient details, and the reporting of this study conforms to the Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) guidelines.<sup>15</sup>

### Input data

Each anteroposterior lumbar spine radiograph was randomly assigned to the training or validation set using the scikit-learn

package (French Institute for Research in Computer Science and Automation, Rocquencourt, France) (Figure 1). Table 1 lists the dataset details. On each patient’s anteroposterior lumbar spine radiograph, the L5 vertebra was identified by counting down from C2 to the sacrum on an antero-posterior whole-spine radiograph.

To label the L5 vertebra, LabelImg (Heartex, San Francisco, CA, USA; <https://github.com/heartexlabs/labelImg>), an open-source software that is widely used to annotate and label bounding boxes on images, was used. On each patient’s antero-posterior lumbar spine radiograph, the L5 vertebra was identified by counting down from C2 to the sacrum using the anteroposterior whole-spine radiograph as a guide. Rectangular bounding boxes were created to cover all L5 areas with a class named “L5.” All annotation files were exported in YOLO format.

Deep learning model development

Python 3.9.12 (<https://www.python.org/>) and PyTorch 1.12.0 (<https://pytorch.org/>) were used to develop a deep learning

model for detecting L5 vertebrae on lumbar spine radiographs. YOLOv5x, of the YOLOv5 family model (Ultralytics, Los Angeles, CA, USA), was used to detect the L5 vertebra area (Figure 2). YOLOv5 is a family of state-of-the-art object-detection architectures that have been pretrained on the COCO dataset. Our L5 detection model detected only one class of objects. The YOLOv5x model with the highest accuracy was therefore used because accuracy was more important than detection speed.<sup>15</sup>

We used a stochastic gradient descent optimizer and generalized intersection over union loss function to optimize our model. Data augmentation techniques, such as image flip, mixup, and mosaic, improved

Table 1. Dataset configuration.

	Sample size, <i>n</i> (no metal implants)	Sample size, <i>n</i> (metal implants)	Total, <i>n</i> (%)
Training set	97	8	105 (70%)
Validation set	41	4	45 (30%)
Total	138	12	150 (100%)

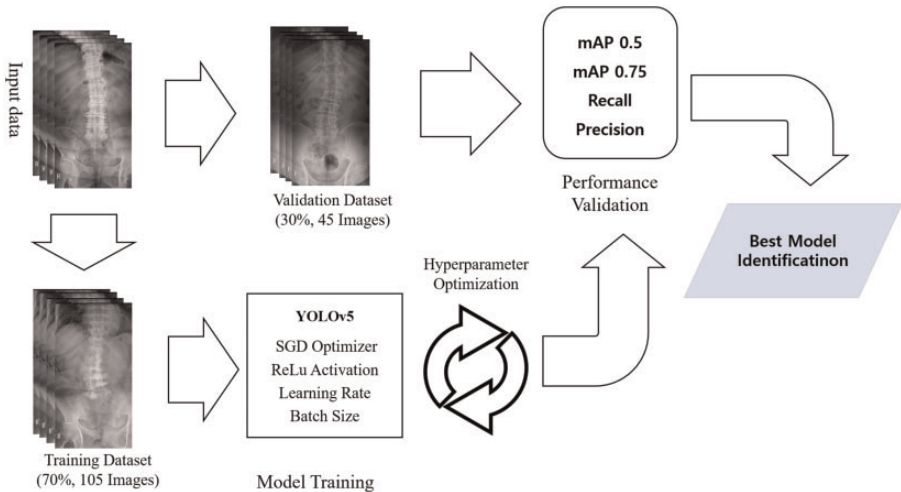


Figure 1. Diagram of the process of the L5 detection model training. mAP, mean average precision; SGD, stochastic gradient descent.

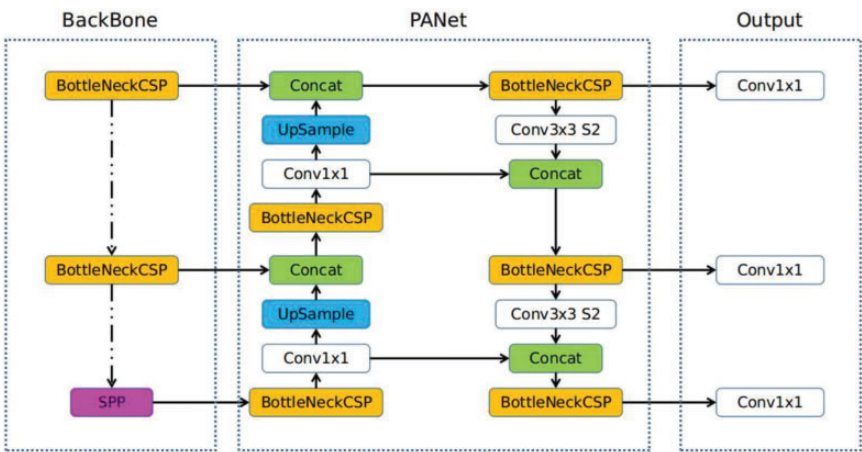


Figure 2. Overview of the YOLOv5 model.

Table 2. Model hyperparameters for training.

Parameter	Value	Remark
Optimizer	SGD	
Loss function	GIOU	
lr0	0.01	initial learning rate
lrf	0.1	final OneCycleLR learning rate ( $lr0 \times lrf$ )
Weight_decay	0.0005	optimizer weight decay $5e^{-4}$
momentum	0.937	SGD momentum/Adam beta1
hsv_h	0.015	image HSV-Hue augmentation (fraction)
hsv_v	0.4	image HSV-Value augmentation (fraction)
flipr	0.5	image flip left-right (probability)
mosaic	1.0	image mosaic (probability)
mixup	0.1	image mixup (probability)

GIOU, generalized intersection over union; HSV, hue, saturation, value; lr, learning rate; lr0, initial learning rate; lrf, final learning rate; SGD, stochastic gradient descent.

the model’s generalization performance. Table 2 lists the details of the model’s experimental hyperparameters.

Performance analysis

The mean average precisions (mAPs) 0.5 and 0.75 were used to measure the performance of the L5 detection model. The mAP\_0.5 and mAP\_0.75 metrics were considered to indicate “good detection” when the intersections over unions, which are an indicator of the degree to which two areas

overlap between the areas of the detected and real objects, were  $\geq 50\%$  and  $\geq 75\%$ , respectively.

Failure analysis was also performed to provide additional information, to better understand the causes of incorrectly detected cases as well as the strengths and weaknesses of the deep learning algorithm.<sup>16</sup>

Results

We recruited 150 patients (mean age,  $66.86 \pm 12.97$  years; 84 men and 66 women)

with anteroposterior whole-spine and lumbar spine radiographs. Of the 150 anteroposterior lumbar spine images, 105 (70%) were randomly selected as the training set, whereas the remaining 45 (30%) were assigned to the validation set.

The mAP<sub>0.5</sub> and mAP<sub>0.75</sub> metrics of the trained L5 detection model were 99.2% and 96.9%, respectively, indicating its outstanding performance.<sup>17</sup> The precision and recall were 95.7% and 97.8%, respectively. Figure 3 shows the changes in mAP<sub>0.5</sub> and mAP<sub>0.75</sub> metrics by training epoch. Of the 45 validation data points, the L5 vertebra was correctly detected with  $\geq 90\%$  confidence in 42 (93.3%; Figure 4).

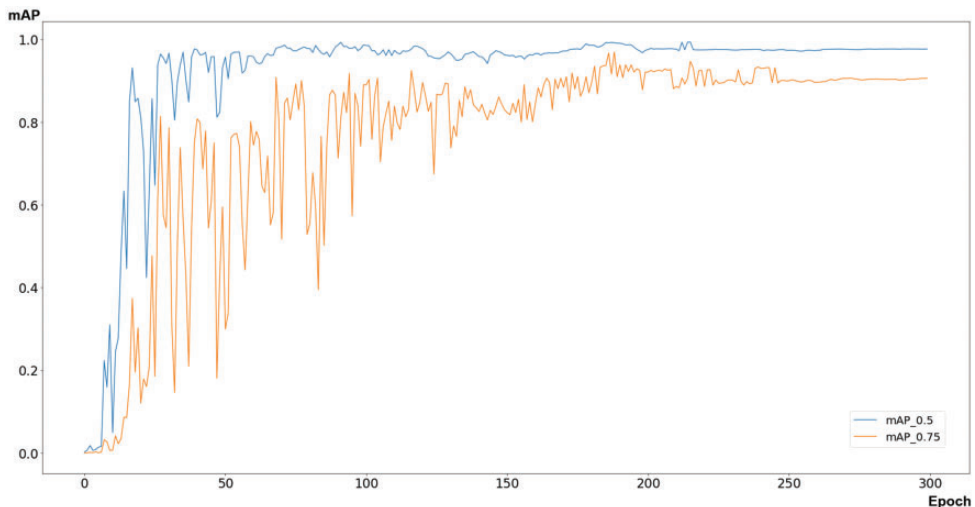
Failure analysis was performed to aid our understanding of the incorrectly detected cases. The L5 vertebra was completely misdetected in two (4.4%) of the 45 validation data points (Figure 5a, b). The metal implants in Figure 5b seem to have caused this misdetection. Figure 5c shows an incorrectly detected case. In our L5 detection model, two vertebrae were selected as L5. The vertebra in the upper bounding box was presented as L5 with 87% confidence, whereas that in the lower bounding box was

presented as L5 with 53% confidence; however, the vertebra in the lower bounding box was the actual L5 vertebra.

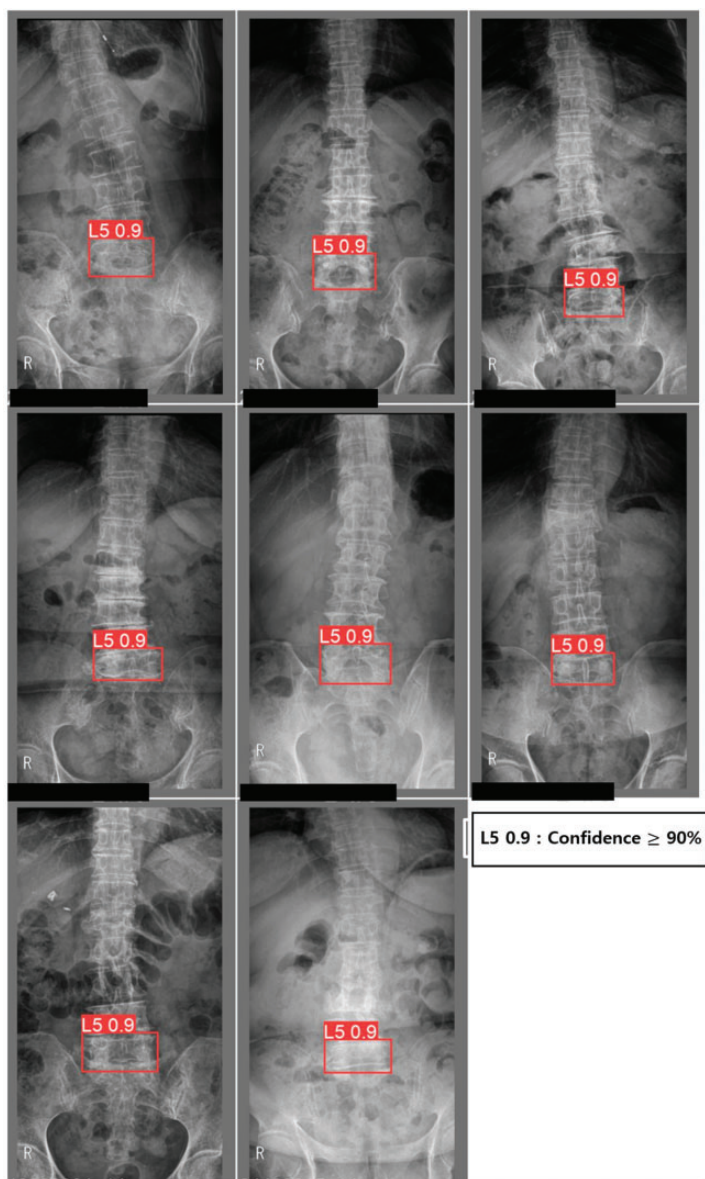
## Discussion

In the current study, we developed a deep learning model to detect the L5 vertebra using YOLOv5x. The mAP<sub>0.5</sub> metric of our developed model was 98.2%. The results of the failure analysis showed that 42 cases (93.3%) among the validation dataset were correctly identified, whereas only three (6.7%) were incorrectly detected. Given that that  $\geq 90\%$  accuracy in a deep learning classification model is classified as outstanding, our L5 vertebra detection model appears to have outstanding performance.<sup>17</sup>

We aimed to confirm the exact location of L5 using deep learning. The accurate determination of lumbar spine segment is of great importance.<sup>18</sup> Lumbar spine pathology and related clinical symptoms are strongly associated with specific lumbar nerve roots, which can be localized to the lumbar spine level.<sup>19</sup> The selection of incorrect lumbar spine levels increases the



**Figure 3.** Mean average precision (mAP) changes according to training epoch.

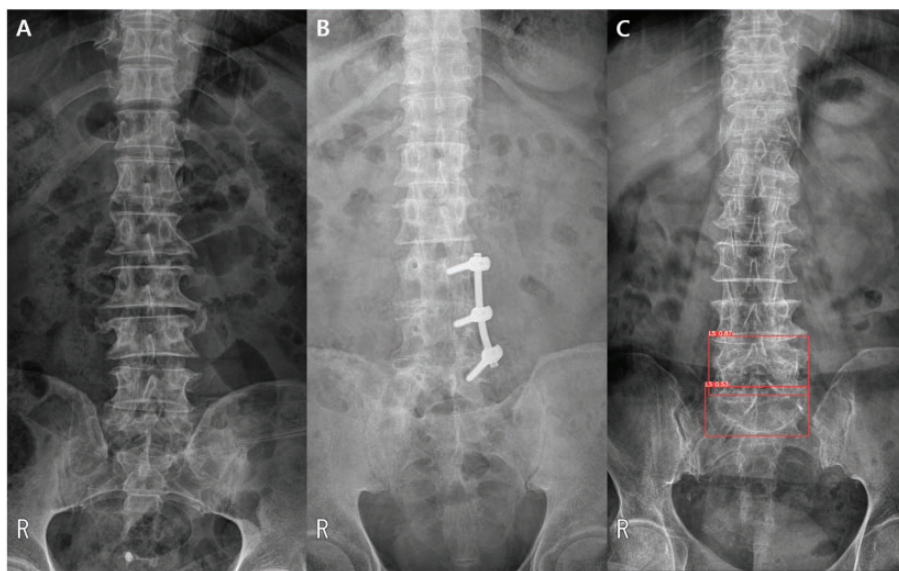


**Figure 4.** Examples of correctly detected cases.

likelihood of identifying the wrong lumbar nerve root, which can result in the treatment of incorrect lumbar segment levels during interventional or surgical procedures.<sup>20</sup> Historically, several methods for accurately identifying the lumbar spine level have been reported, including

a method using the 12th rib, iliolumbar ligament, kidney level, and craniocaudal level on whole-spine radiographs. However, conventional methods are not accurate;<sup>21–23</sup> spine physicians have therefore endeavored to develop further methods for accurately identifying lumbar spine segment numbers.





**Figure 5.** Three examples of incorrectly detected cases. (a, b) L5 detection failure and (c) L5 vertebra was incorrectly detected. Our model presented two vertebrae as L5. The upper bounding box presented the L5 vertebra with 87% confidence, whereas the lower bounding box presented the L5 vertebra with 53% confidence. However, the vertebra in the lower bounding box was the actual L5 vertebra in this case.

The accurate selection of the L5 level is of particular importance in clinical settings because most lumbar pathologies that require intervention or surgery are located at L4 to L5 and L5 to S1.<sup>24</sup> However, it is difficult to identify the L5 vertebra using conventional methods for the following reasons. First, L5 is easily miscounted when sacralization or lumbarization occurs. Second, in the case of spondylolisthesis, it is easy to erroneously number the lumbar vertebrae when only an anteroposterior radiograph is examined. Third, in cases of severe osteoporosis, spinal or pelvic deformity, or congenital spinal pathologies, an accurate L5 selection is more difficult than usual. When spinal computed tomography or MRI is performed together with lumbar spinal radiography, the determination of the L5 vertebra is slightly easier; however, in clinical practice, spinal radiography is typically performed without spinal computed tomography or MRI. Therefore, during

spinal surgery or intervention, the L5 vertebra can sometimes be incorrectly determined by clinicians.

Regarding the determination of the L5 vertebra, many studies have examined patients with lumbosacral transitional vertebrae. There are two main methods for determining the location of the L5 vertebra. The first method identifies the L5 vertebra on the basis of characteristic vertebrae or structures. For example, clinicians can accurately identify L5 by counting down the number of vertebral bodies from C2. Moreover, identification of the T12 vertebra by identifying the vertebra to which the last rib is attached can reveal the L1 vertebra, and the fourth vertebra below the L1 is the L5. The L5 vertebra can also be identified by its transverse process using the Ferguson view. However, in typical clinical situations, whole-spine radiography is not performed. Furthermore, the method used to initially identify T12 is relatively

inaccurate because of the existence of dysplastic ribs, and the method using the Ferguson view has a sensitivity of just 76% to 84%.

The second method involves determining the L5 vertebra by identifying the iliac crest tangent sign in the coronal planes of MRI scans. However, this method has a sensitivity of just 81%. Compared with these traditional methods for counting vertebra number, our recently developed deep learning model showed high accuracy (93.3%) for detecting the L5 vertebra on anteroposterior lumbar spine radiographs. Deep learning models are characterized by multi-layer structures with multiple hidden layers; they thus have a better detection accuracy ability than traditional shallow learning models [8]. We therefore believe that our deep learning model extracts valuable features that differentiate the L5 vertebra from other vertebral levels.

In the present study, we developed a deep learning model for detecting the L5 vertebra using YOLOv5x that showed outstanding performance. We believe that our model may be useful in clinical practice. To our knowledge, our study offers the first deep learning model to detect the L5 vertebra. However, the present investigation had some limitations. First, it included a relatively small amount of imaging data. Second, we used a dataset from a single hospital. Third, we did not compare the accuracy of our developed model with that of humans. Fourth, we did not evaluate the performance of our developed model using data from external clinics or hospitals. To increase the accuracy and versatility of our deep learning model and measure its performance more accurately, these limitations should be addressed in future studies.

#### Declaration of conflict of interest

The authors declare that there is no conflict of interest.

#### Funding

This work was supported by a National Research Foundation of Korea grant funded by the Korean government (MSIT; NO.RS-2023-00219725), and by a 2021 Yeungnam University Research Grant.

#### ORCID iDs

Min Cheol Chang  <https://orcid.org/0000-0002-7629-7213>

Gun Woo Lee  <https://orcid.org/0000-0002-8441-0802>

#### References

1. Peckham ME, Hutchins TA, Stilwill SE, et al. Localizing the L5 vertebra using nerve morphology on MRI: an accurate and reliable technique. *AJNR Am J Neuroradiol* 2017; 38: 2008–2014.
2. Hughes RJ and Saifuddin A. Imaging of lumbosacral transitional vertebrae. *Clin Radiol* 2004; 59: 984–991.
3. Mahato NK. Redefining lumbosacral transitional vertebrae (LSTV) classification: integrating the full spectrum of morphological alterations in a biomechanical continuum. *Med Hypotheses* 2013; 81: 76–81.
4. Jancuska JM, Spivak JM and Bendo JA. A review of symptomatic lumbosacral transitional vertebrae: Bertolotti's syndrome. *Int J Spine Surg* 2015; 9: 42.
5. Paik NC, Lim CS and Jang HS. Numeric and morphological verification of lumbosacral segments in 8280 consecutive patients. *Spine (Phila Pa 1976)* 2013; 38: E573–578.
6. Paik NC, Lim CS and Jang HS. Numbering of vertebrae on MRI using a PACS cross-referencing tool. *Acta Radiol* 2012; 53: 785–789.
7. Jhavar BS, Mitsis D and Duggal N. Wrong-sided and wrong-level neurosurgery: a national survey. *J Neurosurg Spine* 2007; 7: 467–472.
8. Kim JK, Chang MC and Park D. Deep learning algorithm trained on brain magnetic resonance images and clinical data to predict motor outcomes of patients with corona radiata infarct. *Front Neurosci* 2022; 15: 795553.



9. Kim JK, Choo YJ, Choi GS, et al. Deep learning analysis to automatically detect the presence of penetration or aspiration in videofluoroscopic swallowing study. *J Korean Med Sci* 2022; 37: e42.
10. Kim JK, Choo YJ, Shin H, et al. Prediction of ambulatory outcome in patients with corona radiata infarction using deep learning. *Sci Rep* 2021; 11: 7989.
11. Liu L, Ouyang W, Wang X, et al. Deep learning for generic object detection: a survey. *Int J Comput Vis* 2020; 128: 261–318.
12. Li Y, Ouyang W, Wang X, et al. Vip-CNN: visual phrase guided convolutional neural network. In: *IEE Conference on Computer Vision and Pattern Recognition*, Hawaii, USA, 22–26 July 2017.
13. Bhandari A, Prasad PWC, Alsadoon A, et al. Object detection and recognition: using deep learning to assist the visually impaired. *Disabil Rehabil Assist Technol* 2021; 16: 280–288.
14. Yang R and Yu Y. Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. *Front Oncol* 2021; 11: 638182.
15. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147: 573–577.
16. Liu K, Tang H, He S, et al. Performance validation of YOLO variants for object detection. In *Proceedings of the 2021 International Conference on Bioinformatics and Intelligent Computing* 2021; 239–243.
17. Mongan J, Moy L and Kahn CE Jr. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. *Radiol Artif Intell* 2020; 2: e200029.
18. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010; 5: 1315–1316.
19. Kim GU, Park WT, Chang MC, et al. Diagnostic technology for spine pathology. *Asian Spine J* 2022; 16: 764–775.
20. Kasai Y, Paholpak P, Wisanuyotin T, et al. Incidence and skeletal features of developmental cervical and lumbar spinal stenosis. *Asian Spine J* 2023; 17: 240–246.
21. Kato S, Terada N, Niwa O, et al. Factors affecting incomplete L5/s posterior lumbar interbody fusion, including spinopelvic sagittal parameters. *Asian Spine J* 2022; 16: 526–533.
22. Tatara Y, Niimura T, Sekiya T, et al. Changes in lumbosacral anatomy and vertebral numbering in patients with thoracolumbar and/or lumbosacral transitional vertebrae. *JBJS Open Access* 2021; 6: e20.00167.
23. Tureli D, Ekinici G and Baltacioglu F. Is any landmark reliable in vertebral enumeration? A study of 3.0-Tesla lumbar MRI comparing skeletal, neural, and vascular markers. *Clin Imaging* 2014; 38: 792–796.
24. Zhou S, Du L, Liu X, et al. Quantitative measurements at the lumbosacral junction are more reliable parameters for identifying and numbering lumbosacral transitional vertebrae. *Eur Radiol* 2022; 32: 5650–5658.