European Spine Journal
https://doi.org/10.1007/s00586-025-09167-3

RESEARCH

Fully automated pedicle screw manufacturer identification in plain radiograph with deep learning methods

Rattapoom Waranusast[1] · Panomkhawn Riyamongkol[1] · Santi Weerakul[2] · Nattharut Chaibhuddanugul[2] · Artit Laoruengthana[2] · Akaworn Mahatthanatrakul[2]

Received: 18 January 2025 / Revised: 4 July 2025 / Accepted: 14 July 2025

Abstract

Introduction Pedicle screw manufacturer identification is crucial for revision surgery planning; however, this information is occasionally unavailable. We developed a deep learning-based algorithm to identify the pedicle screw manufacturer from plain radiographs.

Methods We collected anteroposterior (AP) and lateral radiographs from 276 patients who had thoracolumbar spine surgery with pedicle screws from three international manufacturers. The samples were randomly assigned to training sets (178), validation sets (40), and test sets (58). The algorithm incorporated a convolutional neural network (CNN) model to classify the radiograph as AP and lateral, followed by YOLO object detection to locate the pedicle screw. Another CNN classifier model then identified the manufacturer of each pedicle screw in AP and lateral views. The voting scheme determined the final classification. For comparison, two spine surgeons independently evaluated the same test set, and the accuracy was compared.

Results The mean age of the patients was 59.5 years, with 1,887 pedicle screws included. The algorithm achieved a perfect accuracy of 100% for the AP radiograph, 98.9% for the lateral radiograph, and 100% when both views were considered. By comparison, the spine surgeons achieved 97.1% accuracy. Statistical analysis revealed near-perfect agreement between the algorithm and the surgeons.

Conclusion We have successfully developed an algorithm for pedicle screw manufacturer identification, which demonstrated excellent accuracy and was comparable to experienced spine surgeons.

Keywords Machine learning · Pedicle screw · Computer-Assisted radiographic image interpretation · Deep learning

Introduction

Spinal fusion is the standard treatment for various spinal diseases, including degeneration, trauma, deformity, tumors, and infection. The number of spinal instrumented fusions has been rapidly increasing in the past decades, with the most common instrument utilized being the pedicle screw fixation system [1, 2]. The pedicle screw plate system was used by Roy-Camille in 1963 and was later popularised in the 1990 s [3]. Instrumented spinal fusion usually results in good to excellent outcomes. However, revision rates remain high at 13–32%, caused by complications such as pseudarthrosis, adjacent segment disease, postoperative junctional kyphosis, and infection [4–7]. More concerning is the increasing trend in revision rates over time [8, 9]. Lang et al. [8] revealed that studies published after 2014 demonstrate an increase in reoperation rates compared to the earlier studies. The pedicle screw-rod system was developed

✉ Akaworn Mahatthanatrakul
akawornm@nu.ac.th

Rattapoom Waranusast
rattapoomw@nu.ac.th

Panomkhawn Riyamongkol
panomkhawnr@nu.ac.th

Santi Weerakul
jswe2000@yahoo.com

Nattharut Chaibhuddanugul
beth_na@hotmail.com

Artit Laoruengthana
artitlao@gmail.com

1 Department of Electrical and Computer Engineering, Faculty of Engineering, Naresuan University, Phitsanulok, Thailand

2 Department of Orthopaedics,Faculty of Medicine, Naresuan University, Phitsanulok, Thailand

with variations of screws, rods, and nuts designs. There are dozens of international pedicle screw manufacturers and many local manufacturers, with each manufacturer developing its system for pedicle screw insertion and removal. The different removal equipment designs indicate the need to identify the pedicle screw manufacturer and specific pedicle screw model before the revision surgery. Failure to correctly identify the model of pedicle screws may result in difficulty or being incapable of removing the existing instruments. Usually, the models of pedicle screws used are recorded in the operative note. Still, it is not uncommon that this information may be unavailable at the time of revision, especially with the obsolete instrument or when the index operation was done at another hospital or in another country. Surgeons are usually accustomed to pedicle screws from a few manufacturers and may be unfamiliar with others, so pre-operative visual identification from plain radiographs is challenging.

Artificial intelligence (AI) is being developed worldwide to assist humans in performing multiple tasks [10, 11]. Computer vision is one of the AI branches that has rapidly evolved in recent years. With the advances of deep learning and neural networks, many models were developed for the segmentation, classification, and interpretation of medical images [12]. There are currently only a few studies [13–17] using these technologies to identify the pedicle screw manufacturer. Although the results were quite satisfactory, several parts of the algorithms need improvement.

We developed an algorithm for fully automated pedicle screw manufacturer identification in the plain thoracolumbar radiograph using multiple deep learning methods to help spine surgeons identify and prepare the correct equipment for the upcoming revision surgery.

Materials and methods

Overview of the study design

This retrospective single-center study was designed following the Checklist for Artificial Intelligence in Medical Imaging (CLAIM) guideline [18]. The study protocol was conducted in agreement with the Declaration of Helsinki and approved by our institutional review board (IRB No. P3-0075-2565).

The study was designed to evaluate the proposed algorithm for pedicle screw manufacturer classification. The radiographs were randomly divided into a training set (65%), a validation set (15%), and a test set (20%). After the training, the trained algorithm classifies the screw manufacturer in each radiograph in the test set. Two spine surgeons with 5 and 20 years of experience with these pedicle screws independently classify the pedicle screw manufacturer of each patient using AP and lateral radiographs. The surgeons were given six example images from the training set, consisting of AP and lateral radiographs of each manufacturer. The interrater correlation between the algorithm and spine surgeons was calculated.

Proposed algorithm

The proposed algorithm integrates multiple deep learning techniques to fully automate the identification of pedicle screw manufacturers from plain radiographs. It consists of several steps, each designed to improve the accuracy and reliability of the classification process. The system takes as input a pair of radiographs, one from the AP view and one from the lateral view of a patient, and follows a multi-step approach to predict the manufacturer of the pedicle screws used in the surgery. Figure 1 summarizes the proposed algorithm from the inputs to the output.

Input radiographs

The algorithm begins by accepting a pair of X-ray images, one from the AP view and one from the lateral view. These two views provide different perspectives of the pedicle screws, which allow the system to capture different structural details important for manufacturer identification.

View classification

Before analyzing the screws, the algorithm uses the *view classifier model* mentioned in the previous section to ensure that the input radiographs are correctly identified as either AP or lateral views. This step is essential because the subsequent models are view-specific.

Screw detection

After the views are classified, the *screw detector model* is applied to both the AP and lateral radiographs to detect and extract the pedicle screws in the images. This step isolates the pedicle screws from the surrounding anatomy, ensuring that only the relevant parts of the image are used in the next stages.

Manufacturer classification

Once the screws are detected and localized, the system extracts each screw from the image. The *screw classifier model* is then employed to classify the manufacturer of each pedicle screw based on the extracted images. Separate

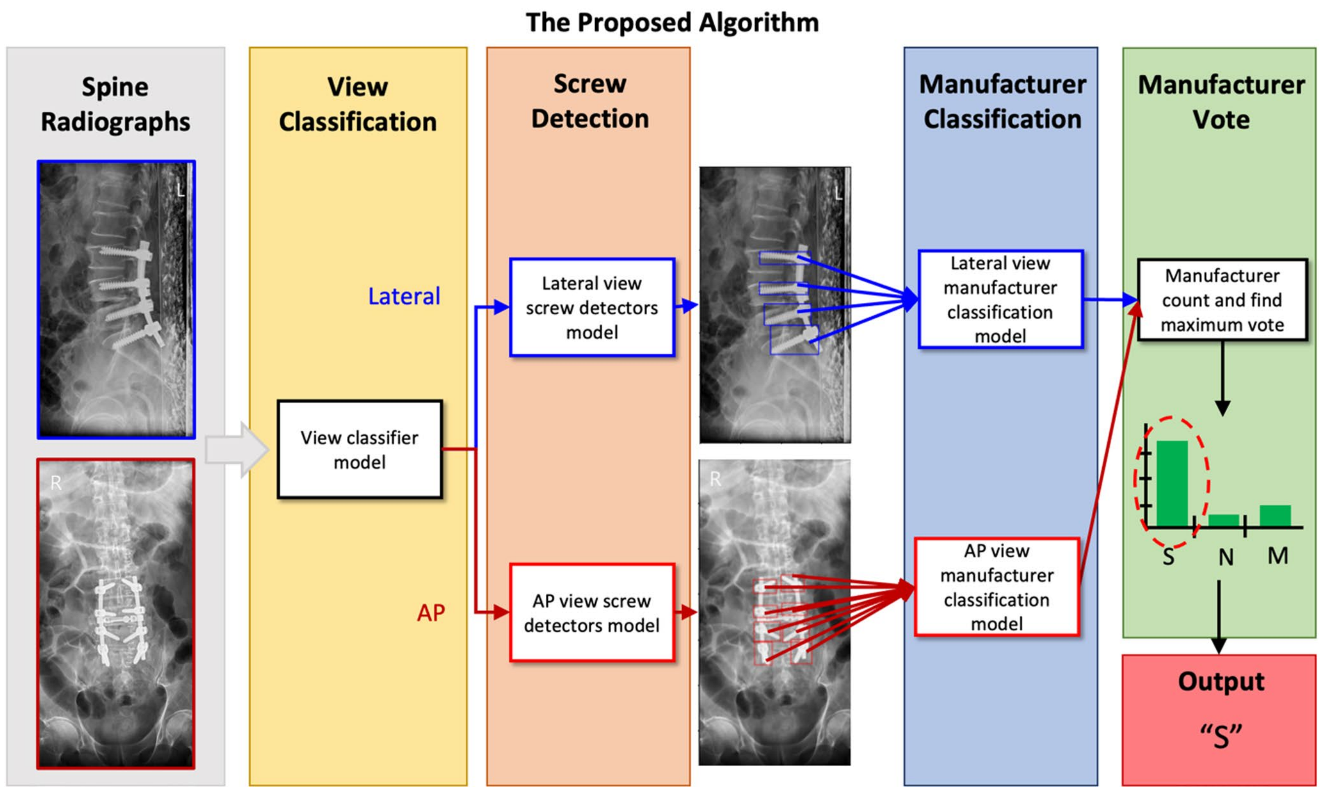

Fig. 1 The overview of the proposed algorithm

classifiers are trained for AP and lateral views to optimize accuracy for each perspective.

Voting system for final prediction

After classifying each screw in both the AP and lateral images, the final manufacturer prediction for a patient is determined by a voting mechanism. This process aggregates the predictions from all screws in both views using a technique similar to a hard voting ensemble method, also known as majority voting. This process involves counting the predictions made by each classifier and selecting the class that receives the most votes as the final prediction. That is, the manufacturer with the most votes from the individual screw predictions across both radiographs is selected as the final output. This step increases the overall accuracy by combining the information from multiple viewpoints and screws. Fig. S1 illustrates the algorithm as a flowchart.

Dataset and preprocessing

Digital Imaging and Communications in Medicine (DICOM) files of plain radiographs of 300 patients who underwent thoracolumbar pedicle screw insertion from January 2017 to December 2020 were collected from the picture archiving and communication system (PACS). All DICOM images were checked and anonymized by manually removing all metadata from the file using Bee DICOM Viewer (SinoUnion Healthcare Inc., Beijing). The images were then converted to JPG format. The inclusion criteria are: (1) Age more than 20 years old (2) Underwent thoracolumbar pedicle screws insertion using one of the three pedicle screw system that were most commonly used in our hospital (CD Horizon Legacy, Medtronic, Minnesota; SpheRx, NuVasive, California; Xia II, Stryker, Michigan). (3) Has a record of the manufacturer of pedicle screws in the operative note. Exclusion criteria (1) Low-quality image (2) Iliac screw or S2 alar-iliac screw. (3) Broken screws (4) Unable to incorporate every pedicle screw in a single radiograph, e.g., patients with long-level thoracolumbar fusion. 6) Patient with pedicle screws from multiple manufacturers.

Ground truth

The ground truth and data labeling were derived from actual surgical records, ensuring accurate identification of the pedicle screw manufacturers used in each case. The images were labeled as AP or lateral, and pedicle screws from which manufacturer were used. A spine surgeon with 10 years of experience did the label. This labeling was used for both training and evaluating the machine learning models. Additionally, this ground truth was used for benchmarking the algorithm's performance, as it allowed direct comparison between the model's predictions and the identifications made by the spine surgeons.

Models training and postprocessing

Model training for the view classifier

To effectively classify the input radiographic images as either AP or lateral views, we trained a classifier using the VGG-16 architecture [19]. Since the pedicle screws appear differently when viewed from AP and lateral views, it is important to develop a classifier that could accurately distinguish between these views. By correctly identifying the view, the algorithm can select the appropriate, view-specific model for implant manufacturer identification, thereby improving classification accuracy with a relatively small dataset.

A total of 124 radiographic images were randomly selected for training the view classifier, divided into 84 images for training, 20 for validation, and 20 for testing. We employed data augmentation techniques to artificially increase the diversity of the training data and improve model generalization [20]. The augmentations included: randomly horizontal flip (flip right-to-left), rotation by ± 15 degrees, translation by ± 10% of the image size, scaling by ± 10%, adjustments to brightness and contrast by ± 30%, and random application of Gaussian blur and sharpening. We used a real-to-augmented data ratio of 1:3, chosen based on experimental comparisons of model performance across different augmentation ratios, including 1:0, 1:3, 1:5, and 1:7. The model was trained for 200 epochs, using a pre-trained VGG-16 model as the base, with the augmented images enhancing the training process.

Model training for the screw detector

We employed YOLOv5 [21] to detect and crop the pedicle screws from the images, ensuring that the CNN models in the later stages would only process the relevant regions. Using 124 images with a total of 820 labeled screws, the YOLOv5 model was trained.

Model training for the screw classifier

After obtaining the screw detection model, we developed a program to automatically load all the radiographic images in the training set, detect, and crop out the pedicle screw sections. These cropped screw images were then organized into separate folders based on the manufacturer. We trained a classifier to identify the manufacturer using the VGG-16 model based on these screw images. From a total of 695 screw images, we got 425 images from AP views and 270 images from lateral view. The VGG-16 model was used for training, with the input image size set to 224 × 224 pixels. To improve the model's performance, we applied the same data augmentation technique as the view classifier training with the addition of a random vertical flip. To enhance model performance, we applied the same data augmentation techniques used in view classifier training, with the addition of random vertical flipping. A real-to-augmented data ratio of 1:5 was employed to further improve generalization. We trained the model for 200 epochs using transfer learning from VGG-16 with Batch Normalization, trained on the ImageNet dataset [19].

Statistical analysis

Parametric data were reported with mean ± standard deviation. Data were compared between the training set, validation set, and test set by ANOVA for continuous data and the Pearson Chi-square test for categorical data. Precision, recall, specificity, accuracy, and F1 score of the algorithm and surgeons were calculated and reported using weighted averages between classes. The agreement between classification by algorithm and surgeons was analyzed using Cohen's kappa. Statistical analyses were performed using IBM SPSS version 29.0 (IBM Corp., Armonk, NY), with statistically significant results at $p < 0.05$.

Results

After exclusion, 276 anteroposterior (AP) and 276 lateral radiographs were included in the dataset (Fig. 2). The mean age was 59.5 ± 9.3 years. Total number of pedicle screws was 1,887 screws. The screws were inserted in the thoracic vertebra with 119 screws, the lumbar vertebra with 1,581 screws, and the sacrum with 187 screws. The radiographs were divided into a training set ($n = 178$), validation set ($n = 40$), and test set ($n = 58$) without statistically significant differences in demographic data, except the validation set has a smaller percentage of thoracic screw compared to the other sets (Table 1).

View classification can classify AP and lateral views with 100% accuracy. AP view screw detectors can detect 406 screws out of 405 screws (1 error of detection) (Fig. 3a). Lateral view screw detectors model can detect only 230 out of 405 screws due to the parallel of the screws. (Figure 3b and c). However, when compared to the labeled data, the screw detectors model achieved a mean Average Precision (mAP@0.5 IoU) of 95.7%. AP screw classification model of each screw can classify the manufacturer with an

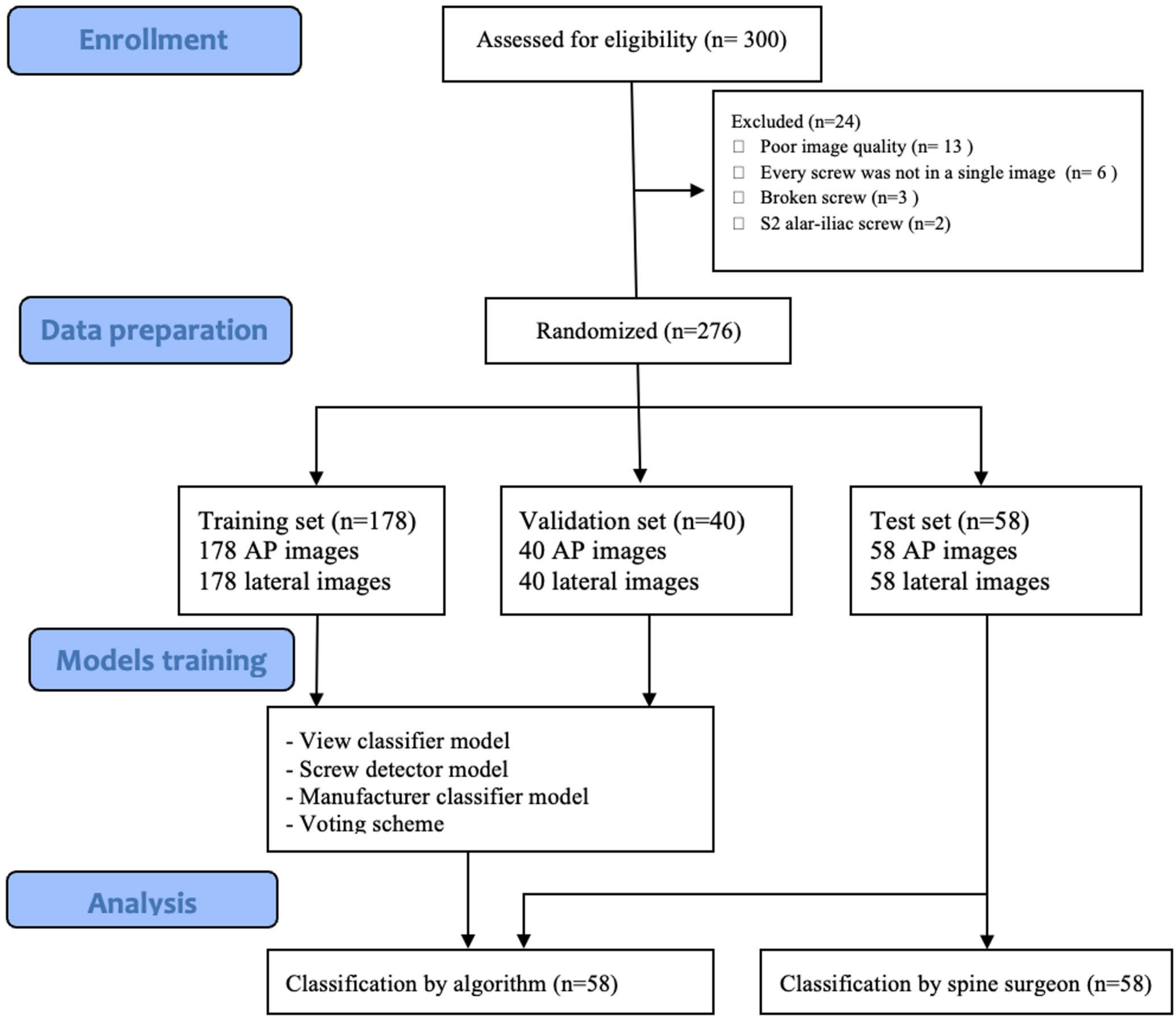

Fig. 2 Flow diagram of data enrollment

accuracy of 96.6% and F1 score of 0.95 (Fig. 4a; Table 2) Meanwhile, the lateral screw classification model has an accuracy of 95.1% and F1 score of 0.93. After the manufacturer vote, the algorithm can identify the manufacturer with 100% accuracy in the AP view and 98.9% in the lateral view. The accuracy is 100% in mixed view (Fig. 4b) Spine surgeons can classify the manufacturer of the pedicle screw with an accuracy of 97.1% (Fig. 4c). Surgeons tended to have more accuracy on pedicle screws that they were more familiar with. The algorithm and the surgeons had a near-perfect agreement with kappa = 0.92 (95% CI = 0.84 to 1.00, $p < 0.001$).

Failed cases analysis

Examples of incorrectly classified cases are presented in Fig. 5 The reason for failed classification includes non-convergence of the screws in AP view resulting in decreased features for the algorithm to determine the manufacturer (Fig. 5a and b). Another reason for failed classification in the upper-thoracic level lateral radiograph is that the screw detector model can detect only three screws out of ten. The lateral classifier model also wrongly identifies two out of three detectable screws. (Figure 5c and d)

Table 1 Demographic data

Parameters	Training set (*n*=178)	Validation set (*n*=40)	Test set (*n*=58)	*p*-value
Female (%)	119 (67%)	28 (70%)	40 (69%)	0.35
Age (year), mean (±SD)	59.5±9.1	59.6±9.1	59.3±10.1	0.38
Multilevel fusion (%)	145 (81%)	35 (88%)	44 (76%)	0.1
Number of pedicle screws	1202	280	405	0.06
Thoracic level	73	10	36	0.02
Lumbar level	1007	239	335	0.65
Sacral level	122	31	34	0.46
Manufacturer				0.67
NuVasive	61	14	20	
Medtronic	57	13	18	
Stryker	60	13	20	

SD: Standard deviation p-value demonstrates statistical difference between the training set, validation set, and test set

Discussion

Computer vision is increasingly utilized to assist doctors in interpreting medical images. Several machine learning models assist radiologists in interpreting radiographs, computed tomography, and magnetic resonance imaging [22, 23]. Implant identification is also another important application of computer vision [13]. Several studies have used machine learning approaches to identify the pedicle screw manufacturers. Yang et al. [14] developed a model using a conventional transfer learning algorithm to identify pedicle screws from five manufacturers in patients who underwent single-level lumbar spinal fusion. After training with 2894 images, they achieved 98% sensitivity and specificity. However, the generalization may be limited due to the inclusion of only single-level surgery. Anand et al. [15] developed another deep learning model to classify pedicle screws from five manufacturers using 396 patients who underwent posterior thoracolumbar instrumentation, including those with multi-level fusion. The authors used the KAZE algorithm to extract features for model training and classification. The performance of their algorithm is outstanding for the classification of two or three manufacturers (91% and 82%, respectively), but the performance drops to 66% for 5-class identification. More recently, Yao et al. [16] used YOLOv5 to crop the radiograph and then trained the model with EfficientNet [24] to identify the brand of pedicle screw from 7 manufacturers with an accuracy of 94%.

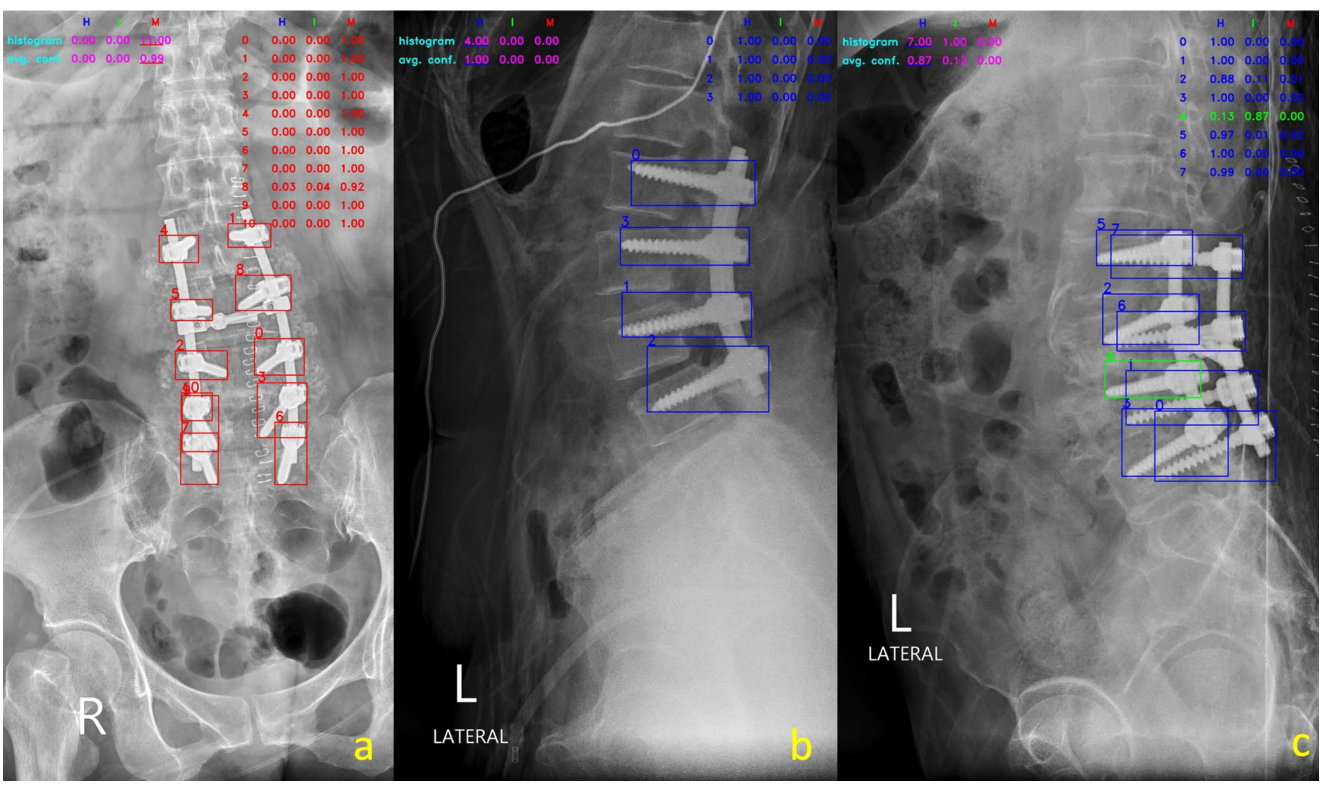

Fig. 3 Bounding boxes show classifier results identifying screw manufacturers. Legend: H (Blue color)=Stryker, I (Green color)=NuVasive, M (Red color)=Medtronic. Numbers (0.00–1.00) indicate confidence in the prediction. (a) Error of screw detector model in the anteroposterior radiograph. The screw detector model identifies 11 screws instead of the actual ten screws. (b) The error of screw detector model in a lateral radiograph. (c) Correct detection of the pedicle screw in a lateral radiograph

Screw classifier: Anteroposterior view		Predicted implants		
		Stryker	Nuvasive	Medtronic
Actual implants	Stryker	113	1	4
	Nuvasive	5	135	10
	Medtronic	0	1	136

Screw classifier: Lateral view		Predicted implants		
		Stryker	Nuvasive	Medtronic
Actual implants	Stryker	65	1	1
	Nuvasive	2	79	2
	Medtronic	0	11	69

a

Algorithm: Anteroposterior view		Predicted implants		
		Stryker	Nuvasive	Medtronic
Actual implants	Stryker	20	0	0
	Nuvasive	0	20	0
	Medtronic	0	0	18

Algorithm: Lateral view		Predicted implants		
		Stryker	Nuvasive	Medtronic
Actual implants	Stryker	20	0	0
	Nuvasive	0	20	0
	Medtronic	0	1	17

Algorithm: Both views		Predicted implants		
		Stryker	Nuvasive	Medtronic
Actual implants	Stryker	20	0	0
	Nuvasive	0	20	0
	Medtronic	0	0	18

b

Spine Surgeon		Predicted implants		
		Stryker	Nuvasive	Medtronic
Actual implants	Stryker	20	0	0
	Nuvasive	3	17	0
	Medtronic	0	0	18

c

Fig. 4 Confusion matrix of screw classifier model (a), confusion matrix of the algorithm (b), and confusion matrix of spine surgeons

Table 2 Accuracy of the algorithm for pedicle screw manufacturer prediction using AP, lateral, and both views compare to surgeons

Metrics	Screw: AP	Screw: Lateral	Vote: AP	Vote: Lateral	Vote: Both	Surgeons
Precision (%)	95.1	93.1	100.0	98.4	100.0	96.0
Recall (%)	94.9	93.0	100.0	98.3	100.0	95.7
Specificity (%)	97.5	96.1	100.0	99.1	100.0	97.8
Accuracy (%)	96.6	95.1	100.0	98.9	100.0	97.1
F1 Score	0.95	0.93	1.00	0.98	1.00	0.96

Screw: Screw classifier model for each screw, Vote: Result after the vote, AP: Anteroposterior

Initially, we attempted direct classification from the full radiographic images, similar to the published studies [14, 15]. However, this approach resulted in an unsatisfactory accuracy of 58% for AP views and 75% for lateral views. The primary issues were the limited dataset size and the VGG-16 model's small input size limitation (224 × 224 pixels), which caused significant information loss during image resizing. Additionally, random cropping often resulted in non-implant regions being used for training, as the implants constitute only a small portion of the image. To address this challenge, we observed that key distinguishing features, such as screw angles, pitch distance, and thread length, are found on the pedicle screws. Therefore, we developed a multi-step approach focusing on detecting and isolating the screws before classification resulting in improved accuracy. The main difference between our approach and previous algorithm is that while other models are trained on whole images with multiple pedicle screws and crosslinks.

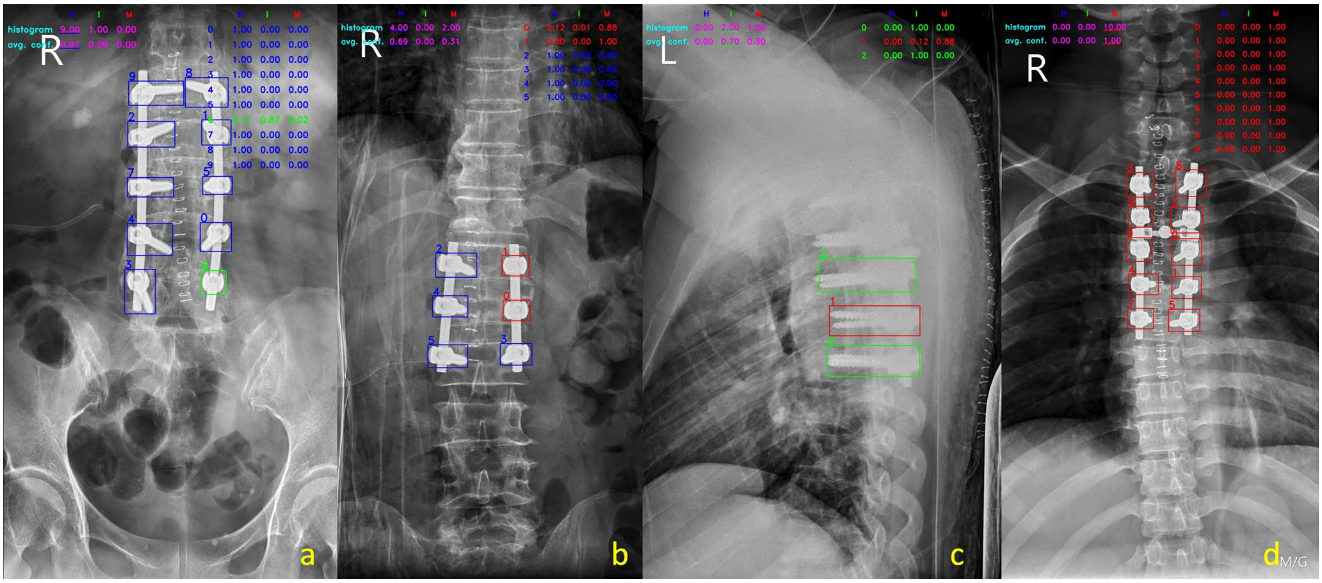

Fig. 5 Suboptimal screw convergence examples where only screw heads are visible without body and thread portions, leading to potential model misclassification due to missing anatomical features. (**a** and **b**). An error in the algorithm for the lateral radiograph (**c**). Only 3 out of 10 screws in T2-T6 levels were detected, and 2 of them were incorrectly classified. However, in the anteroposterior film of the same patient, an algorithm for anteroposterior radiographs can correctly detect and classify all ten screws (**d**)

In contrast, our classifier model trained specifically with each pedicle screw.

Our methodology offers several benefits. First, we can reduce the black-box problem of a deep neural network because the algorithm focuses on and extracts the features from each screw, resulting in an explainable process. Second, we can increase the sample by converting 178 radiographs into 1,200 individual pedicle screw images for model training, improving model robustness. Third, the algorithm is more user-friendly as there is no need to crop the image to a specific area, unlike other algorithms. This feature also accelerates the classification process when identifying a high volume of cases. Fourth, by detecting and classifying each pedicle screw, our model can be applied to patients who underwent multiple surgeries with implants from different manufacturers. However, in cases of patients with multiple implant systems, the voting mechanism might produce incorrect result because it will select only one manufacturer with the highest screw count, so it is crucial to consider the prediction probability of each screw in such cases.

The difference in pedicle screw numbers between training, validation, and test sets approaches statistical significance ($p = 0.06$). While this heterogeneity between datasets could potentially affect model performance, our results show that the model maintained high accuracy in the test sets. This suggests that our model successfully learned generalizable features rather than dataset-specific patterns, indicating good robustness despite the data heterogeneity.

Previous studies typically reported that lateral image classifiers perform better than the AP image classifiers [13–16]. However, in this study, we found the opposite result. Our AP classifier outperformed the lateral classifier. This difference can be explained by several technical factors. First, identifying the individual pedicle screws in lateral radiograph is more challenging due to the parallel orientation of the pedicle screws (Fig. 4b and c). Second, also due to the overlapping appearance of pedicle screws, we had fewer training samples for the lateral screws resulting in reduced performance of the lateral image classifiers.

Limitations and future study

The study was limited to cases from our institutions, and we have only three most commonly used instruments. Using computer vision for classification, we have to increase the training dataset when another class is added. Otherwise, the performance will decrease. To integrate this model into clinical practice, we need to extend this study to cases from other institutions, other manufacturers, and the inclusion of spinal instruments other than pedicle screws. Another limitation is that we included only thoracolumbar pedicle screws. Since posterior cervical screws have a different shape and size than thoracolumbar pedicle screws, we might need to train the model separately for cervical screw classification.

Conclusion

We developed a new algorithm for the identification of pedicle screw manufacturer from three manufacturers with an excellent accuracy and comparable to experienced spine surgeons.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00586-025-09167-3.

Acknowledgements This research was funded by the Faculty of Medicine, Naresuan University. The authors would like to thank the Clinical Epidemiology Unit at Naresuan University Hospital for scientific method advice, as well as the Radiology Department at Naresuan University Hospital for assistance with radiograph preparation.

Author contributions Conceptualization: RW, PR, NC, AL, AMData curation: RW, PR, SW, NCFormal Analysis: RW, SW, NC, AMVisualization: RW, AM Methodology: RW, SW, NC, AL, AMProject Administration: AL, AM Writing– Original Draft: RW, PRWriting– Review & Editing: RW, PR, SW, NC, AL, AM.

Data availability Data generated or analyzed during the study are available from the corresponding author by request.

Declarations

Competing interests The authors declare no competing interests.

References

1. Reisener MJ, Pumberger M, Shue J, Girardi FP, Hughes AP (2020) Trends in lumbar spinal fusion-a literature review. J Spine Surg 6(4):752–761. https://doi.org/10.21037/jss-20-492
2. Rajaee SS, Bae HW, Kanim LE, Delamarter RB (2012) Spinal fusion in the united states: analysis of trends from 1998 to 2008. Spine (Phila Pa 1976) 37(1):67–76. https://doi.org/10.1097/BRS.0b013e31820cccfb
3. Roy-Camille R, Saillant G, Mazel C (1986) Internal fixation of the lumbar spine with pedicle screw plating. Clin Orthop Relat Res. 203:7–17. https://journals.lww.com/clinorthop/abstract/1986/02000/internal_fixation_of_the_lumbar_spine_with_pedicle.3.aspx
4. Cummins D, Hindoyan K, Wu HH, Theologis AA, Callahan M, Tay B et al (2022) Reoperation and mortality rates following elective 1 to 2 level lumbar fusion: A large state database analysis. Global Spine J 12(8):1708–1714. https://doi.org/10.1177/2192568220986148
5. Puvanesarajah V, Shen FH, Cancienne JM, Novicoff WM, Jain A, Shimer AL et al (2016) Risk factors for revision surgery following primary adult spinal deformity surgery in patients 65 years and older. J Neurosurg Spine 25(4):486–493. https://doi.org/10.3171/2016.2.SPINE151345
6. Martin BI, Mirza SK, Comstock BA, Gray DT, Kreuter W, Deyo RA (2007) Reoperation rates following lumbar spine surgery and the influence of spinal fusion procedures. Spine (Phila Pa 1976) 32(3):382–387. https://doi.org/10.1097/01.brs.0000254104.55716.46
7. Kim CH, Chung CK, Park CS, Choi B, Hahn S, Kim MJ et al (2013) Reoperation rate after surgery for lumbar spinal stenosis without spondylolisthesis: a nationwide cohort study. Spine J 13(10):1230–1237. https://doi.org/10.1016/j.spinee.2013.06.069
8. Lang Z, Li JS, Yang F, Yu Y, Khan K, Jenis LG et al (2019) Reoperation of decompression alone or decompression plus fusion surgeries for degenerative lumbar diseases: a systematic review. Eur Spine J 28(6):1371–1385. https://doi.org/10.1007/s00586-018-5681-2
9. Martin BI, Mirza SK, Comstock BA, Gray DT, Kreuter W, Deyo RA (2007) Are lumbar spine reoperation rates falling with greater use of fusion surgery and new surgical technology? Spine (Phila Pa 1976) 32(19):2119–2126. https://doi.org/10.1097/BRS.0b013e318145a56a
10. Lee S, Jung J-Y, Mahatthanatrakul A, Kim J-S (2024) Artificial intelligence in spinal imaging and patient care: A review of recent advances. Neurospine 21(2):474–486. https://doi.org/10.14245/ns.2448388.194
11. Hornung AL, Hornung CM, Mallow GM, Barajas JN, Espinoza Orias AA, Galbusera F et al (2022) Artificial intelligence and spine imaging: limitations, regulatory issues and future direction. Eur Spine J 31(8):2007–2021. https://doi.org/10.1007/s00586-021-07108-4
12. Compte R, Granville Smith I, Isaac A, Danckert N, McSweeney T, Liantis P et al (2023) Are current machine learning applications comparable to radiologist classification of degenerate and herniated discs and modic change? A systematic review and meta-analysis. Eur Spine J 32(11):3764–3787. https://doi.org/10.1007/s00586-023-07718-0
13. Dutt R, Mendonca D, Phen HM, Broida S, Ghassemi M, Gichoya J et al (2022) Automatic localization and brand detection of cervical spine hardware on radiographs using weakly supervised machine learning. Radiol Artif Intell 4(2):e210099. https://doi.org/10.1148/ryai.210099
14. Yang HS, Kim KR, Kim S, Park JY (2021) Deep learning application in spinal implant identification. Spine (Phila Pa 1976) 46(5):E318–E24. https://doi.org/10.1097/BRS.0000000000003844
15. Anand A, Flores AR, McDonald MF, Gadot R, Xu DS, Ropper AE (2023) A computer vision approach to identifying the manufacturer of posterior thoracolumbar instrumentation systems. J Neurosurg Spine 38(4):417–424. https://doi.org/10.3171/2022.11.SPINE221009
16. Yao YC, Lin CL, Chen HH, Lin HH, Hsiung W, Wang ST et al (2024) Development and validation of deep learning models for identifying the brand of pedicle screws on plain spine radiographs. JOR Spine 7(3):e70001. https://doi.org/10.1002/jsp2.70001
17. Huang KT, Silva MA, See AP, Wu KC, Gallerani T, Zaidi HA et al (2019) A computer vision approach to identifying the manufacturer and model of anterior cervical spinal hardware. J Neurosurg Spine 31(6):844–850. https://doi.org/10.3171/2019.6.SPINE19463
18. Tejani AS, Klontzas ME, Gatti AA, Mongan JT, Moy L, Park SH et al (2024) Checklist for artificial intelligence in medical imaging (CLAIM): 2024 update. Radiology: Artif Intell 6(4):e240300. https://doi.org/10.1148/ryai.240300
19. Simonyan K, Zisserman A (2014) doi Very Deep Convolutional Networks for Large-Scale Image Recognition.: https://doi.org/10.48550/arXiv.1409.1556
20. Perez L, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning. https://doi.org/10.48550/arXiv.1712.04621
21. Glenn Jocher AC, Stoken A et al (2022) ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation. Zenodo

22. Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ et al (2019) Deep Learn Med Imaging Neurospine 16(4):657–668. https://doi.org/10.14245/ns.1938396.198
23. Liawrungrueang W, Han I, Cholamjiak W, Sarasombath P, Riew KD (2024) Artificial intelligence detection of cervical spine fractures using convolutional neural network models. Neurospine 21(3):833–841. https://doi.org/10.14245/ns.2448580.290
24. Tan M, Le QV (2019) EfficientNet: rethinking model scaling for convolutional neural networks. https://doi.org/10.48550/arXiv.1905.11946