



## Article

# Development of a Method for Estimating the Angle of Lumbar Spine X-ray Images Using Deep Learning with Pseudo X-ray Images Generated from Computed Tomography

Ryuma Moriya <sup>1</sup>, Takaaki Yoshimura <sup>2,3,4,5</sup> , Minghui Tang <sup>5,6</sup>, Shota Ichikawa <sup>7,8</sup>  and Hiroyuki Sugimori <sup>4,5,9,\*</sup> <sup>1</sup> Graduate School of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan<sup>2</sup> Department of Health Sciences and Technology, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan<sup>3</sup> Department of Medical Physics, Hokkaido University Hospital, Sapporo 060-8648, Japan<sup>4</sup> Global Center for Biomedical Science and Engineering, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan<sup>5</sup> Clinical AI Human Resources Development Program, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan<sup>6</sup> Department of Diagnostic Imaging, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Sapporo 060-8638, Japan<sup>7</sup> Department of Radiological Technology, School of Health Sciences, Faculty of Medicine, Niigata University, Niigata 951-8518, Japan<sup>8</sup> Institute for Research Administration, Niigata University, Niigata 950-2181, Japan<sup>9</sup> Department of Biomedical Science and Engineering, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan

\* Correspondence: sugimori@hs.hokudai.ac.jp; Tel.: +81-11-706-3410

**Abstract:** Background and Objectives: In lumbar spine radiography, the oblique view is frequently utilized to assess the presence of spondylolysis and the morphology of facet joints. It is crucial to instantly determine whether the oblique angle is appropriate for the evaluation and the necessity of retakes after imaging. This study investigates the feasibility of using a convolutional neural network (CNN) to estimate the angle of lumbar oblique images. Since there are no existing lumbar oblique images with known angles, we aimed to generate synthetic lumbar X-ray images at arbitrary angles from computed tomography (CT) images and to estimate the angles of these images using a trained CNN. Methods: Synthetic lumbar spine X-ray images were created from CT images of 174 individuals by rotating the lumbar spine from 0° to 60° in 5° increments. A line connecting the center of the spinal canal and the spinous process was used as the baseline to define the shooting angle of the synthetic X-ray images based on how much they were tilted from the baseline. These images were divided into five subsets and trained using ResNet50, a CNN for image classification, implementing 5-fold cross-validation. The models were trained for angle estimation regression and image classification into 13 classes at 5° increments from 0° to 60°. For model evaluation, mean squared error (MSE), root mean squared error (RMSE), and the correlation coefficient (r) were calculated for regression analysis, and the area under the curve (AUC) was calculated for classification. Results: In the regression analysis for angles from 0° to 60°, the MSE was 14.833 degree<sup>2</sup>, the RMSE was 3.820 degrees, and r was 0.981. The average AUC for the 13-class classification was 0.953. Conclusion: The CNN developed in this study was able to estimate the angle of an lumbar oblique image with high accuracy, suggesting its usefulness.

**Keywords:** deep learning; angle estimation of lumbar; regression model; classification model



**Citation:** Moriya, R.; Yoshimura, T.; Tang, M.; Ichikawa, S.; Sugimori, H. Development of a Method for Estimating the Angle of Lumbar Spine X-ray Images Using Deep Learning with Pseudo X-ray Images Generated from Computed Tomography. *Appl. Sci.* **2024**, *14*, 3794. <https://doi.org/10.3390/app14093794>

Academic Editor: Adam Konefal

Received: 30 March 2024

Revised: 21 April 2024

Accepted: 27 April 2024

Published: 29 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Low back pain is a common health problem worldwide, with numerous studies finding that it is a leading cause of disability and has a significant socioeconomic impact, reducing

the quality of life for many people [1–3]. Lumbar spondylolysis and spondylolisthesis are among the most common causes of low back pain [4–6]. Magnetic resonance imaging and X-ray images are used to identify lumbar spondylolysis and spondylolisthesis, and lumbar spine radiography is often used first [7,8]. Among these, lumbar oblique radiography is often used to determine the presence of the aforementioned lumbar spondylolysis and the morphology of the intervertebral joints, and the images obtained change depending on the angle at which the body is tilted [9,10]. Therefore, since the optimal angle varies depending on the joint or area to be observed and the suspected pathological condition, if the angle can be instantly estimated from the obtained image, it is easier to determine whether the image is what is required at that time. In other words, there is a need for instantaneous estimation of whether the oblique angle is appropriate for post-imaging evaluation and whether re-imaging is necessary.

Obtaining an appropriate lumbar oblique image is essential for accurate diagnosis and treatment decisions. However, it is difficult even for skilled radiological technologists to obtain the optimal angle in a single shot, and multiple shots are sometimes required. This leads to an increase in patient radiation exposure and prolongation of examination time, so improvements are needed. In addition, images taken at inappropriate angles may lead to a decrease in diagnostic accuracy and oversight, which may adversely affect the patient's treatment plan.

In recent years, deep learning techniques have been used in the medical field [11–17], and there has been a lot of research related to the spine [18]. Deep learning has the potential to automate the estimation of optimal imaging angles by modeling the experience and knowledge of skilled radiologists. The development of such technology is expected to reduce the burden on radiological technologists and improve the efficiency of examinations, thereby contributing to improvements in the quality of medical care.

To further elucidate the methodology of our study, it is crucial to explain the rationale behind employing two distinct convolutional neural network (CNN) models: one specialized for classification and another for regression. These models interact by initially using the classification model to categorize the images based on certain predefined criteria, which then guides the regression model to accurately estimate the imaging angles.

In this study, we wanted to investigate the possibility of using a convolutional neural network (CNN) to estimate the angle of a lumbar oblique image. Since big data is needed to create a CNN model, in the case of this study, lumbar oblique X-ray images from various angles with known acquisition angles are required. In this study, however, we created lumbar spine X-ray images of arbitrary angles pseudo-created from computed tomography (CT) images, and we used these images to create the model and evaluate its accuracy. This makes it possible to generate a large number of images from various angles. This will ensure a large dataset necessary for training the CNN model.

The purpose of this study is to create lumbar spine X-ray images of arbitrary angles pseudo-rectified from CT images and to estimate the angles of lumbar spine oblique images using a CNN trained by those images. The results of this study are expected to contribute to the optimization and efficiency of lumbar oblique imaging, thereby reducing the burden on patients and improving the quality of medical care. In addition, the method proposed in this study can be applied to images other than lumbar oblique images, and it is expected to contribute to the development of medical image analysis.

The significance of this study is that it proposes a method for automatically estimating the imaging angles of lumbar oblique images using deep learning. This method will enable the determination of the imaging angle based on objective data, which until now has relied heavily on the radiological technologist's experience and intuition. This is expected to improve the reproducibility and consistency of imaging and lead to improved diagnostic accuracy. The findings obtained in this study can also be applied to radiographic imaging of other parts of the body and are expected to contribute widely to the optimization of radiographic examinations.

Furthermore, this study is also significant in terms of the application of artificial intelligence (AI) in the medical field. AI is being applied to various aspects of medicine, but it has not yet been fully utilized in the field of radiological examinations. This study shows one way to improve the quality of radiological examinations using AI and expands the possibilities for medical applications of AI. As described above, this study proposes a new approach to the global health problem of low back pain using deep learning. The results of this study are expected to contribute to the optimization and efficiency of lumbar oblique radiography, thereby reducing the burden on patients and improving the quality of medical care. Furthermore, this research demonstrates the potential for medical applications of AI, and it is expected to help promote the use of AI in the medical field.

To provide a clear understanding of the research presented in this article, the structure is organized as follows. Following this introduction, Section 2 describes the datasets used, the preprocessing steps, and the deep learning models developed. Section 3 presents the findings of the CNN's performance in estimating the angles of lumbar spine oblique images. Subsequently, Section 4 interprets these results, comparing them with existing methods, and explores the implications for clinical practice. Finally, Section 5 summarizes the study's contributions and outlines future research directions.

## 2. Materials and Methods

### 2.1. Subjects and Research Environments

In this study, CT data of the mediastinum and abdomen were obtained from "A new 2.5D representation for lymph node detection in CT" published in TCIA (The Cancer Image Archive) [7]. This dataset contained CT images from a variety of patients and was deemed suitable for the purpose of this study. Of the data acquired, those that did not include all of the lumbar spine were excluded, and finally, CT data from 174 patients were used. Table 1 shows the specifications of the computer used in this study. MATLAB (2022a; The MathWorks, Inc., Natick, MA, USA) was used as the analysis software.

**Table 1.** Software and specifications of the computer used in this study.

Environment	Contents
Software	MATLAB 2022a (Mathworks)
OS	Windows 10
CPU	Intel Core i9-10980XE 3.5 GHz
RAM	DDR4 2666 Mhz 64 GB
GPU	NVIDIA RTX P5000 16 GB × 4

### 2.2. Data Preprocessing—The Creation of Supervised Data

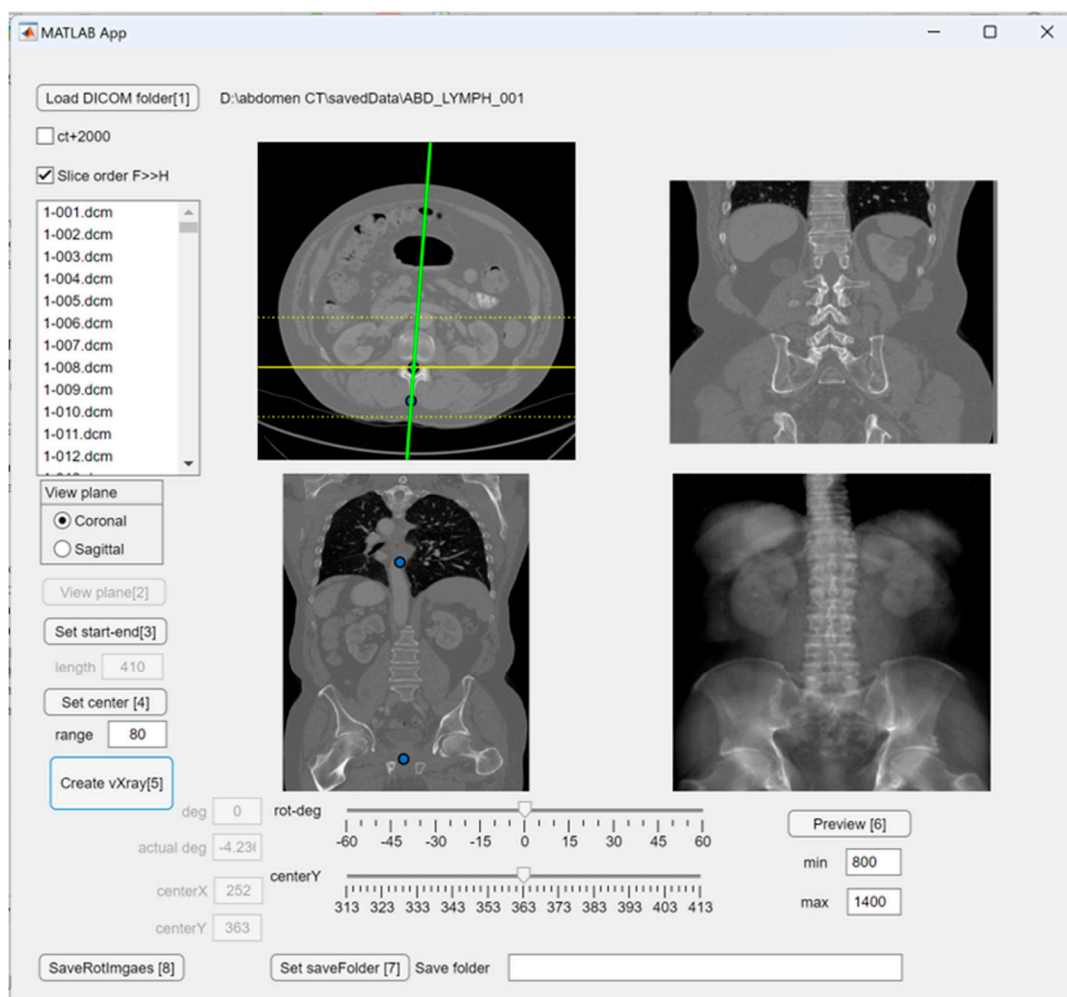
A large amount of supervised data is required to create a CNN model [19]; however, the dataset of lumbar spine oblique X-ray images at various angles needed for this study was not publicly available. Therefore, in this study, lumbar spine X-ray images of arbitrary angles were pseudo-fabricated from CT images, and these images were used as the supervised data.

The procedure for creating supervised data using the software shown in Figure 1 is summarized below:

1. Specify the center of the spinal canal and the spinous process in the axial section image of the third lumbar vertebra, thereby creating a reference line for the angle;
2. The axial section image is converted to a coronal section image, and the range is set so that the first to the fifth lumbar vertebrae are included in the image;
3. Cut out the set area in the plane perpendicular to the created reference line;
4. Create a pseudo X-ray image containing three-dimensional information by adding together the images cut out in step 3 to an arbitrary extent;
5. Visually adjust the maximum and minimum pixel values of the image to create an image with contrast similar to the actual X-ray image;

6. Create and save an image according to the procedure described up to step 5, rotated in  $5^\circ$  increments from  $-60^\circ$  to  $60^\circ$  relative to the reference angle.

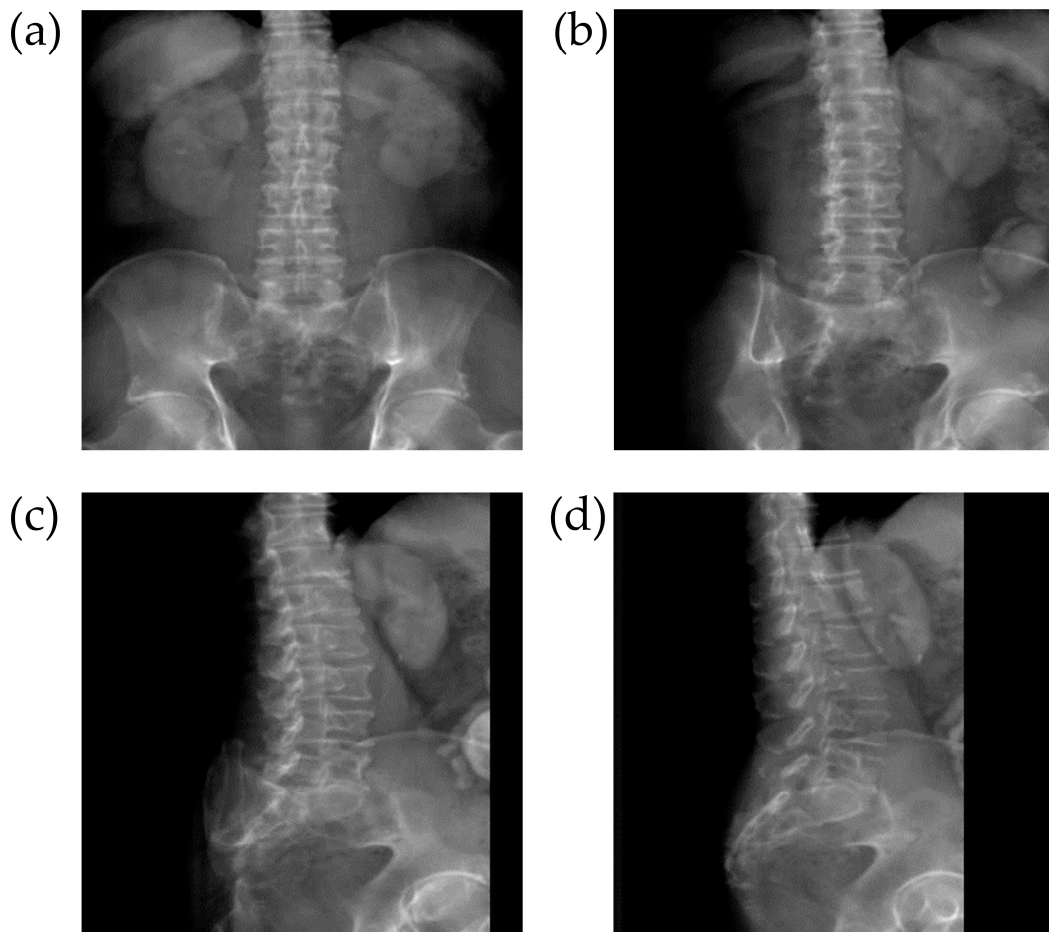
The selection of the range from  $-60^\circ$  to  $60^\circ$  relative to the reference angle ( $0^\circ$ ) for generating lumbar spine X-ray images was primarily based on the typical clinical angles used in lumbar oblique radiography. This range was chosen to encompass the angles most commonly utilized for diagnostic accuracy, as angles beyond  $60^\circ$  often result in a lateral view, which was not the focus of this study. Consequently, the study limited its scope to  $60^\circ$  to maintain relevance to standard oblique imaging practices. The above procedure was applied to all CT images, and 25 lumbar spine X-ray images were created per CT image per patient. An example of the images produced is shown in Figure 2.



**Figure 1.** Example of implementation for creating a supervised image. The green line represents the corrected  $0^\circ$  based on the CT image. The yellow dashed lines indicate the range used for generating the pseudo X-ray image, while the solid yellow line denotes the center of this range.

In creating the supervised data, a line connecting the center of the spinal canal and the spinous process was used as the reference line, and the imaging angle of the pseudo X-ray image was defined by how many degrees it was tilted from that line. This enabled the generation of a large amount of supervised data with clear imaging angles.

Since this method created a pseudo X-ray image from a CT image, there was a possibility that there may be some differences from the actual X-ray image. However, considering that the purpose of this study was to estimate the shooting angle, the pseudo X-ray images were considered to be sufficiently useful. In addition, this method allowed the generation of a large amount of supervised data, which was necessary for the training of the CNN model.



**Figure 2.** Example of lumbar spine pseudo X-ray image; (a) upper left:  $0^\circ$ , (b) upper right:  $20^\circ$ , (c) lower left:  $40^\circ$ , and (d) lower right:  $60^\circ$ .

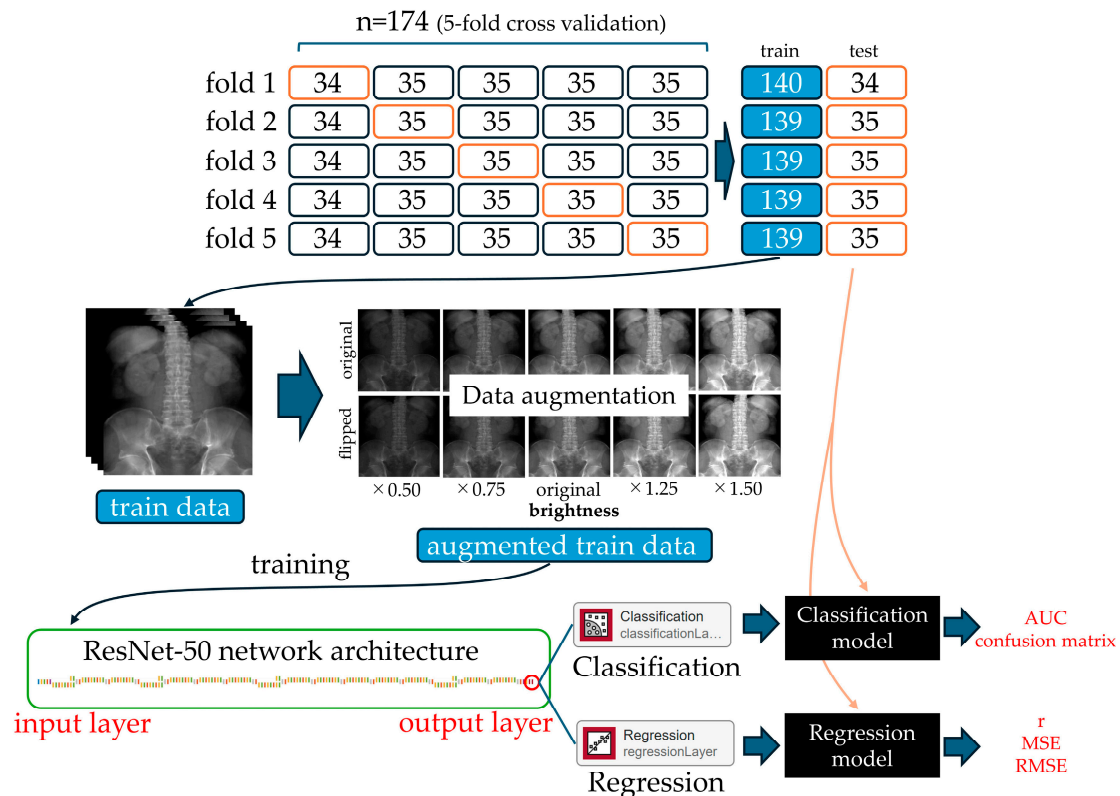
A dedicated program was developed to generate the supervised data. This program read CT images and automatically generated pseudo X-ray images at specified angles. The implementation screen of the program is shown in Figure 1. The program was developed using MATLAB. The development of the program enabled the efficient creation of a large amount of supervised data.

### 2.3. Training Dataset

The CNN model was trained using the created supervised data under the hyperparameters shown in Table 2. For data augmentation, the images were first flipped left and right, doubling the dataset size. Then, the luminance of each image was adjusted to five different levels (0.5, 0.75, 1.0, 1.25, and 1.5) by multiplying the pixel values of the original images by these coefficients, effectively increasing the dataset size by five. Collectively, these augmentations increased the diversity of the supervised data by a factor of ten, thereby enhancing the model's generalization performance.

Next, by integrating the  $\pm$  variations of the supervised images to encompass angles from  $0^\circ$  to  $60^\circ$ , using the expanded supervised data, two CNN models were created: the first was an image classification model that classified images into 13 classes of  $5^\circ$  each from  $0^\circ$  to  $60^\circ$ , and the second was a regression model that estimated angles from lumbar oblique images from  $0^\circ$  to  $60^\circ$ . Both models utilized ResNet50 [20] pretrained with ImageNet as the base architecture, with modifications to replace the original classification layer with a regression layer for the purpose of performing both classification and regression tasks. ResNet50 is known for its excellent performance in image recognition tasks and was considered suitable for the purposes of this study.

A 5-fold cross validation was used to train the model. This method divided the data set into five folds, four of which were used as the training set and the remaining one as the test set. This method reduced the bias of the dataset and allowed for an accurate evaluation of the generalization performance of the model. Figure 3 provides a comprehensive overview of the training process, including the implementation of 5-fold cross-validation to ensure the robustness and reliability of our models.



**Figure 3.** An overview of 5-fold cross validation and classification and regression model training and evaluation.

**Table 2.** Hyperparameters for training.

	Parameters
architecture	ResNet50
mini batch size	128
number of epochs	10
optimizer	SGDM <sup>a</sup>
momentum	0.9
learn rate drop factor	0.1
initial learning rate	0.0001
L2 regularization	0.0001

<sup>a</sup> Stochastic gradient descent with momentum.

## 2.4. Evaluation of Created Models

### 2.4.1. Evaluation of Regression Models

The accuracy of the created CNN models was evaluated. The correlation coefficient (r), mean squared error (MSE), and root mean squared error (RMSE) were used to evaluate the regression models, where r was the correlation coefficient between the actual angle and the estimated angle and is a measure of the strength of the correlation between the two. These indices were calculated using the following equations:

$$MSE = \frac{\sum_{i=1}^n (y_{obs,i} - y_{pred,i})^2}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{obs,i} - y_{pred,i})^2}{n}}$$

where  $y_{obs,i}$  is the actual angle,  $y_{pred,i}$  is the estimated angle, and  $n$  is the number of samples.

Lower values of MSE and RMSE indicated better performance of the model, as it suggested that the predicted values were closer to the actual values. In our study, we used these metrics to compare the accuracy of our proposed method to assess its effectiveness in estimating oblique angles from lumbar spine X-ray images.

#### 2.4.2. Evaluation of Classification Models

The performance of our classification model was assessed using the confusion matrix and the area under the curve (AUC) metric. The confusion matrix serves as a representation of the alignment between the classification model's predictions and the actual classes. It does so by displaying the distribution of predictions across different actual classes, providing a clear picture of the model's accuracy in classification tasks.

**Confusion Matrix:** This matrix is crucial for visualizing the performance of the classification model. It shows the numbers of correct and incorrect predictions divided into each category by the model, allowing us to evaluate the model's performance across various class labels. The confusion matrix was particularly useful in identifying how the model performed for each actual angle versus the predicted angle, enhancing our understanding of its precision and areas where improvement was needed.

**AUC:** AUC is a metric derived from the receiver operating characteristic (ROC) curve, which illustrates the trade-off between the true positive rate and the false positive rate of a classification model at various threshold settings. The AUC metric, ranging between 0 and 1, serves as a summary of the model's ability to correctly classify the instances across all possible thresholds. A value closer to 1 indicates higher model performance. In our study, the AUC was used to quantify the overall effectiveness of the classification model in distinguishing between different classes, providing a single measure of performance irrespective of any particular threshold.

Utilizing these evaluation methods allowed for a comprehensive assessment of the classification model's performance. The confusion matrix provided detailed insight into the model's accuracy for each class, and the AUC offered a holistic measure of its ability to discriminate between classes. Combined, these approaches enabled a robust analysis of the model's performance, with the level of precision and reliability evident.

### 3. Results

#### 3.1. Evaluation of Regression Model

The regression model's performance is summarized in Table 3, where we evaluated the model across five different folds (A to E) to ensure robustness and reliability. The metrics presented are MSE, RMSE, and the correlation coefficient ( $r$ ). The results highlighted an average MSE of 14.833 degree<sup>2</sup>, indicating the variance between the predicted values and the actual values across all folds. The RMSE, providing a sense of the average error margin, stood at 3.820 degrees. The correlation coefficient, averaging at 0.981, underscored a highly positive correlation between the predicted and actual values, suggesting the model's effectiveness in capturing the underlying relationship. In examining the performance across the five folds in Table 3, it was observed that the values for folds 2 and 5 exhibited greater MSE and RMSE values compared to folds 1, 3, and 4. Notably, fold 2 presented the highest MSE at 20.035 degree<sup>2</sup> and RMSE at 4.476 degrees, followed by fold 5 with an MSE of 18.83 degree<sup>2</sup> and RMSE of 4.339 degrees. Despite this, the correlation coefficient for fold 2 was lower than that of fold 5, indicating a variance in prediction accuracy and reliability across folds.

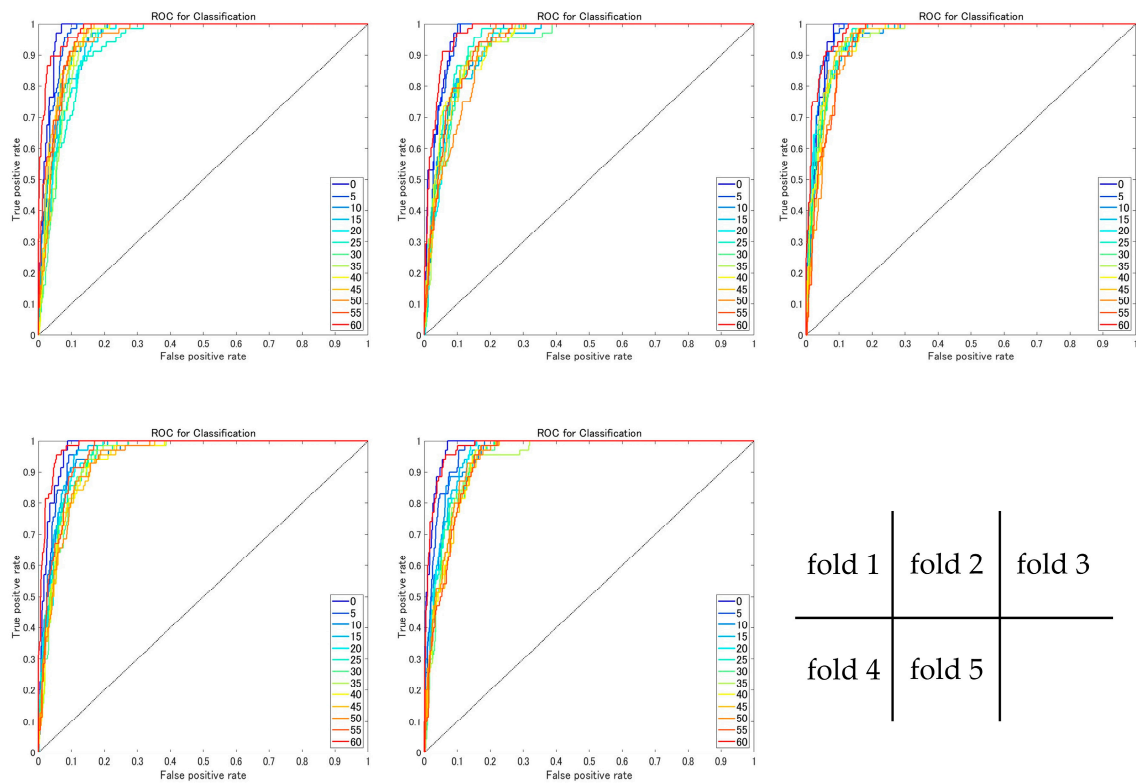
**Table 3.** Summary of regression model performance metrics across five folds.

	MSE <sup>a</sup> (degree <sup>2</sup> )	RMSE <sup>b</sup> (degree)	Correlation Coefficient (r)
fold 1	11.146	3.339	0.985
fold 2	20.035	4.476	0.970
fold 3	11.175	3.343	0.983
fold 4	12.982	3.603	0.980
fold 5	18.830	4.339	0.985
mean	14.833	3.820	0.981

<sup>a</sup> Mean squared error; <sup>b</sup> root mean squared error.

### 3.2. Evaluation of Classification Model

The classification model’s performance is depicted through the ROC curve in Figure 4, showcasing the model’s ability to discriminate between classes effectively. Table 4 details the area under the curve (AUC) for each class across all folds, with a mean AUC of 0.953. This high level of AUC across varying degrees indicated the model’s consistent and reliable classification capability.



**Figure 4.** ROC curve for each fold. ROC curves represent different angle classes ranging from 0° to 60° in 5° increments. The color of each ROC curve changes gradually from cool to warm colors, with 0° represented by blue and 60° represented by red.

**Table 4.** AUC for each angular classification at each fold.

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean
0°	0.978	0.971	0.972	0.975	0.984	0.976
5°	0.968	0.965	0.965	0.964	0.969	0.966
10°	0.956	0.940	0.964	0.955	0.959	0.955
15°	0.946	0.935	0.961	0.959	0.955	0.951



Table 4. Cont.

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean
20°	0.940	0.933	0.961	0.951	0.952	0.947
25°	0.932	0.945	0.959	0.950	0.946	0.946
30°	0.943	0.931	0.957	0.943	0.944	0.944
35°	0.942	0.936	0.955	0.942	0.939	0.943
40°	0.954	0.939	0.956	0.939	0.941	0.946
45°	0.958	0.934	0.954	0.938	0.940	0.945
50°	0.949	0.928	0.945	0.939	0.945	0.941
55°	0.962	0.938	0.947	0.953	0.939	0.948
60°	0.983	0.974	0.978	0.984	0.982	0.980
mean	0.955	0.944	0.960	0.953	0.953	0.953

Additionally, the confusion matrix (Figure 5) visually represented the model’s classification accuracy regarding actual angles versus predicted angles. The matrix employed a color-coded scheme where red cells (value 1) signified complete accuracy in classification, while blue cells (value 0) indicated discrepancies. The granular breakdown provided by the confusion matrix revealed the model’s strengths and areas for improvement, offering valuable insight into its performance across different angle classifications.

In summary, both the regression and classification models demonstrated high degrees of accuracy and reliability in their respective tasks. The detailed metrics and visualizations, such as the ROC curve and confusion matrix, provided a comprehensive overview of the models’ performance, highlighting their strengths in predicting and classifying angles with high precision. Future work could explore further optimization techniques to enhance model performance, particularly at angles where classification accuracy could be improved.

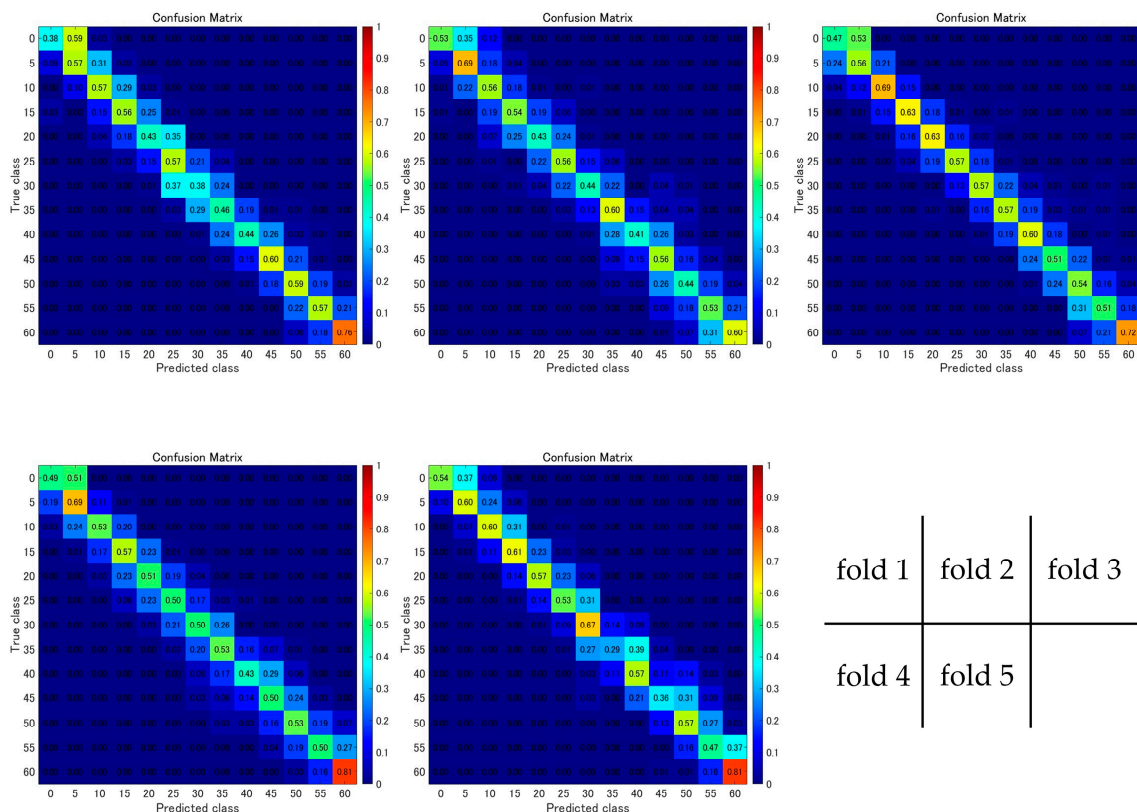


Figure 5. Confusion matrix for each fold.

## 4. Discussion

### 4.1. Study Results and Model Accuracy

The RMSE was 3.820 degrees and the correlation coefficient  $r$  was 0.981. The RMSE was used as an index to evaluate the error between the actual angle of the lumbar oblique image and the estimated angle output by the angle estimation model created in this study. The smaller the RMSE, the better the accuracy of the created model. On the other hand, the correlation coefficient is a statistical index that numerically indicated the strength and direction of the relationship between two variables, with a value closer to 1 indicating a stronger correlation. However, we compared the accuracy with that of other studies in similar fields as a reference.

### 4.2. Comparison with Other Studies

As for angle prediction using CNNs in the vertebral field, many studies [21–23] have been conducted to estimate the Cobb angle in patients with scoliosis. The Cobb angle represents the angle of posterior curvature of the spine and generally refers to the angle that indicates the degree of posterior curvature of the thoracic spine. This angle is usually quantified by measuring the curvature between two specific vertebrae of the thoracic spine. With reference to these results, it can be said that the accuracy of the regression model created in this study in predicting the angle at which lumbar oblique image was taken was sufficiently high. The reason why is that the accuracy of the present study was higher than that of another study [24]. According to the paper, the RMSE for predicting the Cobb angle was 5.48 degrees on average, and the correlation coefficient between the actual Cobb angle and the predicted Cobb angle was 0.91. This is thought to be because the study used images of a human bent forward taken from multiple angles to predict the Cobb angle, and the images used were not as clear as X-ray images. On the other hand, when the accuracy of this study was compared with that of another study [25] that estimated the Cobb angle using a model similar to this study, which was created using X-ray images of the vertebral body, the accuracy of this study was relatively lower. In the external validation dataset of the model for predicting the Cobb angle, the RMSE value was  $1.5^\circ$ , and although there was no direct mention of the  $r$  value, the agreement between the model and the Cobb angle was shown to be very high, with an intraclass correlation coefficient (ICC) of 0.996. This may be due to the difference in the number of data used, but it is also thought to be due to the difference in the dimension of the angle being predicted. Specifically, this is because the Cobb angle estimation estimates a two-dimensional angle in terms of curvature between two specific vertebrae, whereas the lumbar oblique image angle estimation estimates an angle that can be described as a “degree of rotation” rather than a degree of vertebral flexion, etc., in other words, a three-dimensional angle. We believe this difference is the reason why the accuracy of this study was relatively low compared to the accuracy of other studies that estimated Cobb angles. Similar to other regression methods using deep learning in medical imaging [26,27], the significant positive correlation obtained suggested that the predictor was reliable. In Table 3, the observed variations in MSE and RMSE values for folds 2 and 5 invite cautious speculation regarding their cause. These disparities might reflect distinctive characteristics of the data subsets in these folds or specific interactions between the model and these subsets. Notably, while fold 2 shows larger error magnitudes, as indicated by the higher MSE and RMSE values, it has a lower correlation coefficient compared to fold 5, suggesting that errors in fold 5, although smaller in magnitude, could be more variable or less consistent. It is plausible that such differences are attributed to the stochastic nature of the data or to unique patterns not yet fully understood. Increasing the number of cases might enable a more comprehensive evaluation of the model’s performance across varying data conditions. Our study effectively demonstrates the utility of a specific pre-trained CNN model for medical imaging via transfer learning. However, reliance on this singular CNN architecture, while efficient, limits the exploration of newer models that may offer enhanced performance with fewer resources [28,29]. Future research should consider a broader array of models to improve generalizability and performance in diverse imaging

contexts [30,31]. This study lays the groundwork for further exploration into the capabilities of CNNs and their alternatives in medical imaging.

The image classification model, classifying images into 13 classes of  $5^\circ$  each from  $0^\circ$  to  $60^\circ$ , showed high accuracy with an average AUC of 0.953. To clarify the analysis of our results, we examined why the classification accuracy varied significantly across different angles. Specifically, we discussed the mechanisms that may contribute to higher AUC scores at certain angles such as  $0^\circ$  and  $60^\circ$ , considering factors like the unique characteristics of these images and their representation in the training dataset. Despite the confusion matrix appearing inaccurate, the high AUC was due to the difference between the target and output angles being within  $\pm 5^\circ$  for most test data. For instance, for the  $30^\circ$  target, 67 out of 68 images were classified as  $25^\circ$  to  $35^\circ$ . However, for the  $0^\circ$  target, more images were classified as  $5^\circ$  than  $0^\circ$ , resulting in lower accuracy. Further, we considered the reasons behind the notably lower classification accuracy for most angles except the  $60^\circ$  classes. This was likely because the number of learned images for  $0^\circ$  was half that of other angles, as  $+x^\circ$  and  $-x^\circ$  were used as  $x^\circ$ , resulting in twice the number of images for angles other than  $0^\circ$ . To address the concerns regarding the data volume of the  $0^\circ$  category, we should analyze in a future study whether the reduced number of images for this angle was solely responsible for its decreased accuracy or if it also contributed to lower accuracy across other categories. Consequently, it can be speculated that the model did not learn enough  $0^\circ$  lumbar spine X-ray images, which may have led to the observed decrease in accuracy.

#### 4.3. Limitations of the Study

Although the results of this study show a sufficiently high degree of accuracy, several limitations exist. First, it is unclear whether the regression and classification models created look at the vertebrae and determine their angles in the same manner as humans. Not only pseudo-created lumbar spine X-ray images such as those used in this study but also X-ray images contain a variety of information in addition to the vertebral bodies, so it is unclear which part of the body the models are looking at to estimate the angle. Next, while this study used lumbar spine X-ray images pseudo-generated from CT images for model creation, we acknowledge the importance of incorporating real-world X-ray images for a more comprehensive evaluation of the model's performance. Future studies should validate the accuracy of the model using actual lumbar spine X-ray images to confirm its applicability in clinical settings.

As in other studies using CNNs [11–14,16,17], we believe that the accuracy of the angle estimation model may be improved by increasing the number of training images and by using other models and parameters. In particular, in terms of increasing the number of training images, by using the method of pseudo-creating lumbar spine X-ray images from CT images in this study, it is possible to obtain lumbar spine X-ray images from any dataset as long as the lumbar spine is included in the CT image range. We believe that this will solve one of the major problems in CNN research, namely, the lack of training data. However, it is not easy to collect a large number of datasets that include the lumbar spine in the CT scan area. In addition, the lumbar spine X-ray images pseudo-generated from CT images do not exactly match the actual X-ray images, so there are limitations to learning using pseudo images. Effective data collection and learning methods need to be considered while taking these points into account.

The significance of this study lies in the fact that the CNN was trained using pseudo X-ray images of the lumbar spine taken at arbitrary angles from CT images. We have not been able to confirm any studies that have created CNNs using pseudo X-ray images, indicating the uniqueness of this study. In addition, no previous study has been confirmed that regressively analyzes the lumbar spine according to angle or classifies the lumbar spine into classes by angle. Therefore, we believe that the present results suggest the usefulness of the CNN created using lumbar spine X-ray images pseudo-generated from CT images for angle estimation in lumbar spine oblique images using deep learning techniques. However, this study is only an initial investigation, and the aforementioned limitations exist. In

the future, while taking these limitations into account, it will be necessary to verify the method using a larger dataset to explore other models and parameters and to evaluate the method using actual X-ray images. It is also important to further validate the usefulness of the method proposed in this study by applying it to X-ray images of other sites. To further enhance our research perspective, incorporating advanced techniques in image processing, segmentation, and simulation of spine or bone is essential. Notable studies such as CSR-Net for scaphoid fracture segmentation [32], wearable technologies for monitoring spinal movement [33], adversarial models for bone segmentation in CT images [34], and neuroprosthetic approaches for managing hemodynamics post-spinal injury [35] provide valuable insight for future implementations.

The clinical significance of this study is that it is expected to optimize the angle and improve reproducibility in lumbar oblique imaging. Lumbar oblique radiography plays an important role in the diagnosis of lumbar spondylolysis and spondylolisthesis, but it must be performed at an appropriate angle. However, at present, determination of the imaging angle relies heavily on the radiologist's experience and intuition, and it lacks reproducibility and consistency. The CNN-based angle estimation proposed in this study will enable objective and quantitative angle determination, which is expected to lead to improved diagnostic accuracy. It is also expected to contribute to the reduction of patient radiation exposure by reducing the need for re-imaging. However, further verification and improvements are needed to introduce the system proposed in this study into actual clinical practice. In particular, evaluation using actual X-ray images and comparative studies with radiologists are important. In addition, issues such as the impact on the work of radiologists and costs associated with the introduction of the system should also be considered.

The main limitations of this study include the aforementioned problems with the dataset and the imprecise basis for model decisions. We used a ResNet pre-trained on ImageNet, but we must also consider using a model pre-trained on medical images [36] to learn our dataset in this study. The other several limitations of this study should be acknowledged. First, it is unclear whether the regression and classification models created look at the vertebrae and determine their angles in the same manner as humans. Both pseudo-created and actual X-ray images contain various information besides vertebral bodies, so it is uncertain which parts the models used to estimate angles. While this study used pseudo-generated lumbar spine X-ray images for model creation, future studies should validate the accuracy of the model using real-world X-ray images to confirm its applicability in clinical settings. Additionally, the study focused only on lumbar oblique images and did not examine the applicability to other sites or imaging methods. Furthermore, the study concentrated solely on the accuracy of angle estimation, but in actual clinical practice, factors other than angle, such as the patient's body shape, pathological condition, and characteristics of the equipment used, may also influence the optimization of imaging. A more comprehensive study that takes these factors into account is needed.

#### *4.4. Future Research Directions*

In future work, it would be beneficial to explore model optimization algorithms and feature selection techniques that could not be applied in this study. Adapting the performance of algorithms through other state-of-the-art meta-heuristic algorithms [37] and feature selection methods [38] could provide further insight into their effectiveness. Moreover, investigating the impact of different parameter settings and incorporating adaptive mechanisms could potentially enhance the robustness and efficiency of the proposed model in this study.

The results, significance, and limitations of this study are discussed above. This study proposed a new approach of angle estimation by CNN using lumbar spine X-ray images pseudo-generated from CT images and suggested its usefulness. Future work should take into account the limitations of this study while conducting a larger-scale, multifaceted validation.

## 5. Conclusions

In this study, a CNN-based regression model and a classification model were created using lumbar spine X-ray images pseudo-generated from CT images. As a result, the regression model achieved a high accuracy of 3.820 degrees for the RMSE and 0.981 for the correlation coefficient  $r$ , and the classification model obtained an excellent result with an average AUC of 0.953. These results suggested that the CNN using lumbar spine X-ray images pseudo-generated from CT images was useful for angle estimation of lumbar spine oblique images. The significance of this study lies in the fact that we proposed a new approach of creating pseudo-created lumbar spine X-ray images from CT images and used them to estimate angles using a CNN. Conventionally, a large number of actual lumbar spine X-ray images are required for angle estimation of lumbar spine oblique images using a CNN. However, by using the method proposed in this study, it was possible to create pseudo lumbar spine X-ray images from CT images, thereby ensuring the large amount of image data necessary for CNN training. This will greatly contribute to the development of angle estimation of lumbar spine oblique images using deep learning techniques.

This study proposes a new approach to optimize and standardize lumbar oblique radiography and is expected to contribute to the development of radiological examination. In the future, further validation and improvement of this study should be conducted while taking into account the limitations of this study. In addition, it is expected that the findings obtained in this study will be applied to other sites and imaging methods, thereby contributing widely to the optimization and standardization of radiological examinations.

**Author Contributions:** R.M. contributed to the data analysis, algorithm construction, and writing and editing of the manuscript. T.Y., M.T. and S.I. reviewed and edited the manuscript. H.S. proposed the idea and contributed to the data acquisition, performed supervision and project administration, and reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The created models in this study are available upon request from the corresponding author. The source code of this study is available at <https://github.com/MIA-laboratory/LumberAngleEST> (accessed on 30 March 2024).

**Acknowledgments:** The authors would like to thank the laboratory members of the Medical Image Analysis Laboratory for their help.

**Conflicts of Interest:** The authors declare that no conflicts of interest exists.

## References

1. Chen, S.; Chen, M.; Wu, X.; Lin, S.; Tao, C.; Cao, H.; Shao, Z.; Xiao, G. Global, Regional and National Burden of Low Back Pain 1990–2019: A Systematic Analysis of the Global Burden of Disease Study 2019. *J. Orthop. Transl.* **2022**, *32*, 49–58. [[CrossRef](#)] [[PubMed](#)]
2. Ferreira, M.L.; De Luca, K.; Haile, L.M.; Steinmetz, J.D.; Culbreth, G.T.; Cross, M.; Kopec, J.A.; Ferreira, P.H.; Blyth, F.M.; Buchbinder, R.; et al. Global, Regional, and National Burden of Low Back Pain, 1990–2020, Its Attributable Risk Factors, and Projections to 2050: A Systematic Analysis of the Global Burden of Disease Study 2021. *Lancet Rheumatol.* **2023**, *5*, e316–e329. [[CrossRef](#)] [[PubMed](#)]
3. Yang, Y.; Lai, X.; Li, C.; Yang, Y.; Gu, S.; Hou, W.; Zhai, L.; Zhu, Y. Focus on the Impact of Social Factors and Lifestyle on the Disease Burden of Low Back Pain: Findings from the Global Burden of Disease Study 2019. *BMC Musculoskelet. Disord.* **2023**, *24*, 679. [[CrossRef](#)] [[PubMed](#)]
4. Chung, C.C.; Shimer, A.L. Lumbosacral Spondylolysis and Spondylolisthesis. *Clin. Sports Med.* **2021**, *40*, 471–490. [[CrossRef](#)] [[PubMed](#)]
5. Eisenstein, S.M.; Ashton, I.K.; Roberts, S.; Darby, A.J.; Kanse, P.; Menage, J.; Evans, H. Innervation of the Spondylolysis “Ligament”. *Spine* **1994**, *19*, 912–916. [[CrossRef](#)]
6. Nordstrom, D.; Santavirta, S.; Seitsalo, S.; Hukkanen, M.; Polak, J.M.; Nordsletten, L.; Konttinen, Y.T. Symptomatic Lumbar Spondylolysisneuroimmunologic Studies. *Spine* **1994**, *19*, 2752–2758. [[CrossRef](#)]

7. Yamane, T.; Yoshida, T.; Mimatsu, K. Early Diagnosis of Lumbar Spondylolysis by MRI. *J. Bone Jt. Surg. Ser. B* **1993**, *75*, 764–768. [[CrossRef](#)]
8. Trinh, G.M.; Shao, H.C.; Hsieh, K.L.C.; Lee, C.Y.; Liu, H.W.; Lai, C.W.; Chou, S.Y.; Tsai, P.I.; Chen, K.J.; Chang, F.C.; et al. Detection of Lumbar Spondylolisthesis from X-Ray Images Using Deep Learning Network. *J. Clin. Med.* **2022**, *11*, 5450. [[CrossRef](#)] [[PubMed](#)]
9. Chou, R.; Deyo, R.A.; Jarvik, J.G. Appropriate Use of Lumbar Imaging for Evaluation of Low Back Pain. *Radiol. Clin. North Am.* **2012**, *50*, 569–585. [[CrossRef](#)]
10. DeVine, J.G. Commentary: Standardization of Dynamic Lumbar Imaging and Diagnostic Criteria for Discogenic Low Back Pain. *Spine J.* **2011**, *11*, 999–1001. [[CrossRef](#)]
11. Kawakami, M.; Hirata, K.; Furuya, S.; Kobayashi, K.; Sugimori, H.; Magota, K.; Katoh, C. Development of Combination Methods for Detecting Malignant Uptakes Based on Physiological Uptake Detection Using Object Detection With PET-CT MIP Images. *Front. Med.* **2020**, *7*, 616746. [[CrossRef](#)]
12. Asami, Y.; Yoshimura, T.; Manabe, K.; Yamada, T.; Sugimori, H. Development of Detection and Volumetric Methods for the Triceps of the Lower Leg Using Magnetic Resonance Images with Deep Learning. *Appl. Sci.* **2021**, *11*, 12006. [[CrossRef](#)]
13. Manabe, K.; Asami, Y.; Yamada, T.; Sugimori, H. Improvement in the Convolutional Neural Network for Computed Tomography Images. *Appl. Sci.* **2021**, *11*, 1505. [[CrossRef](#)]
14. Sugimori, H.; Shimizu, K.; Makita, H.; Suzuki, M.; Konno, S. A Comparative Evaluation of Computed Tomography Images for the Classification of Spirometric Severity of the Chronic Obstructive Pulmonary Disease with Deep Learning. *Diagnostics* **2021**, *11*, 929. [[CrossRef](#)] [[PubMed](#)]
15. Hirata, K.; Sugimori, H.; Fujima, N.; Toyonaga, T.; Kudo, K. Artificial Intelligence for Nuclear Medicine in Oncology. *Ann. Nucl. Med.* **2022**, *36*, 123–132. [[CrossRef](#)]
16. Yoshimura, T.; Hasegawa, A.; Kogame, S.; Magota, K.; Kimura, R.; Watanabe, S.; Hirata, K.; Sugimori, H. Medical Radiation Exposure Reduction in PET via Super-Resolution Deep Learning Model. *Diagnostics* **2022**, *12*, 872. [[CrossRef](#)]
17. Ichikawa, S.; Itadani, H.; Sugimori, H. Prediction of Body Weight from Chest Radiographs Using Deep Learning with a Convolutional Neural Network. *Radiol. Phys. Technol.* **2023**, *16*, 127–134. [[CrossRef](#)]
18. Galbusera, F.; Casaroli, G.; Bassani, T. Artificial Intelligence and Machine Learning in Spine Research. *JOR Spine* **2019**, *2*, e1044. [[CrossRef](#)] [[PubMed](#)]
19. Khadka, A.; Remagnino, P.; Argyriou, V. Synthetic Crowd and Pedestrian Generator for Deep Learning Problems. In Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing—Proceedings, Barcelona, Spain, 4–8 May 2020; pp. 4052–4056.
20. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [[CrossRef](#)]
21. Tu, Y.; Wang, N.; Tong, F.; Chen, H. Automatic Measurement Algorithm of Scoliosis Cobb Angle Based on Deep Learning. *J. Phys. Conf. Ser.* **2019**, *1187*, 042100. [[CrossRef](#)]
22. Jaremko, J.L.; Poncet, P.; Ronsky, J.; Harder, J.; Dansereau, J.; Labelle, H.; Zernicke, R.F. Genetic Algorithm-Neural Network Estimation of Cobb Angle from Torso Asymmetry in Scoliosis. *J. Biomech. Eng.* **2002**, *124*, 496–503. [[CrossRef](#)]
23. Caesarendra, W.; Rahmani, W.; Mathew, J.; Thien, A. AutoSpine-Net: Spine Detection Using Convolutional Neural Networks for Cobb Angle Classification in Adolescent Idiopathic Scoliosis. *Lect. Notes Electr. Eng.* **2022**, *898*, 547–556. [[CrossRef](#)]
24. Wu, C.; Meng, G.; Lian, J.; Xu, J.; Gao, M.; Huang, C.; Zhang, S.; Zhang, Y.; Yu, Y.; Wang, H.; et al. A Multi-Stage Ensemble Network System to Diagnose Adolescent Idiopathic Scoliosis. *Eur. Radiol.* **2022**, *32*, 5880–5889. [[CrossRef](#)]
25. Kokabu, T.; Kanai, S.; Kawakami, N.; Uno, K.; Kotani, T.; Suzuki, T.; Tachi, H.; Abe, Y.; Iwasaki, N.; Sudo, H. An Algorithm for Using Deep Learning Convolutional Neural Networks with Three Dimensional Depth Sensor Imaging in Scoliosis Detection. *Spine J.* **2021**, *21*, 980–987. [[CrossRef](#)]
26. Usui, K.; Yoshimura, T.; Tang, M.; Sugimori, H. Age Estimation from Brain Magnetic Resonance Images Using Deep Learning Techniques in Extensive Age Range. *Appl. Sci.* **2023**, *13*, 1753. [[CrossRef](#)]
27. Inomata, S.; Yoshimura, T.; Tang, M.; Ichikawa, S.; Sugimori, H. Estimation of Left and Right Ventricular Ejection Fractions from Cine-MRI Using 3D-CNN. *Sensors* **2023**, *23*, 6580. [[CrossRef](#)]
28. Salehi, A.W.; Khan, S.; Gupta, G.; Alabdullah, B.I.; Almjally, A.; Alsolai, H.; Siddiqui, T.; Mellit, A. A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope. *Sustainability* **2023**, *15*, 5930. [[CrossRef](#)]
29. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [[CrossRef](#)]
30. Burri, S.R.; Ahuja, S.; Kumar, A.; Baliyan, A. Exploring the Effectiveness of Optimized Convolutional Neural Network in Transfer Learning for Image Classification: A Practical Approach. In Proceedings of the 2023 International Conference on Advancement in Computation and Computer Technologies, InCACCT 2023, Gharuan, India, 5–6 May 2023; pp. 598–602.
31. Benavente, D.; Gatica, G.; González-Feliu, J. Balanced Medical Image Classification with Transfer Learning and Convolutional Neural Networks. *Axioms* **2022**, *11*, 115. [[CrossRef](#)]
32. Chen, C.; Liu, B.; Zhou, K.; He, W.; Yan, F.; Wang, Z.; Xiao, R. CSR-Net: Cross-Scale Residual Network for Multi-Objective Scaphoid Fracture Segmentation. *Comput. Biol. Med.* **2021**, *137*, 104776. [[CrossRef](#)]

33. Li, C.; Liu, D.; Xu, C.; Wang, Z.; Shu, S.; Sun, Z.; Tang, W.; Wang, Z.L. Sensing of Joint and Spinal Bending or Stretching via a Retractable and Wearable Badge Reel. *Nat. Commun.* **2021**, *12*, 2950. [[CrossRef](#)] [[PubMed](#)]
34. Chen, C.; Qi, S.; Zhou, K.; Lu, T.; Ning, H.; Xiao, R. Pairwise Attention-Enhanced Adversarial Model for Automatic Bone Segmentation in CT Images. *Phys. Med. Biol.* **2023**, *68*, 035019. [[CrossRef](#)] [[PubMed](#)]
35. Squair, J.W.; Gautier, M.; Mahe, L.; Soriano, J.E.; Rowald, A.; Bichat, A.; Cho, N.; Anderson, M.A.; James, N.D.; Gandar, J.; et al. Neuroprosthetic Baroreflex Controls Haemodynamics after Spinal Cord Injury. *Nature* **2021**, *590*, 308–314. [[CrossRef](#)]
36. Mei, X.; Liu, Z.; Robson, P.M.; Marinelli, B.; Huang, M.; Doshi, A.; Jacobi, A.; Cao, C.; Link, K.E.; Yang, T.; et al. RadImageNet: An Open Radiologic Deep Learning Research Dataset for Effective Transfer Learning. *Radiol. Artif. Intell.* **2022**, *4*, e210315. [[CrossRef](#)] [[PubMed](#)]
37. Xu, Z.; Heidari, A.A.; Kuang, F.; Khalil, A.; Mafarja, M.; Zhang, S.; Chen, H.; Pan, Z. Enhanced Gaussian Bare-Bones Grasshopper Optimization: Mitigating the Performance Concerns for Feature Selection. *Expert Syst. Appl.* **2023**, *212*, 118642. [[CrossRef](#)]
38. Xu, Y.; Huang, H.; Heidari, A.A.; Gui, W.; Ye, X.; Chen, Y.; Chen, H.; Pan, Z. MFeature: Towards High Performance Evolutionary Tools for Feature Selection. *Expert Syst. Appl.* **2021**, *186*, 115655. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.