Study of Generalization Performance on Non-Contact Estimation of Lumbar Load Using Webcam Image by Deep Learning for Stationary Standing Forward Bending Posture

Riku Nishimoto
Kochi University of Technology
Miyanokuchi 185, Tosayamada, Kami, Kochi, Japan
Kochi, Japan
email: nishimoto.riku19990917@gmail.com

Kyoko Shibata
Kochi University of Technology
Miyanokuchi 185, Tosayamada, Kami, Kochi, Japan
Kochi, Japan
email: shibata.kyoko@kochi-tech.ac.jp

Abstract - **To prevent lumbago, it is effective to have a system that enables people to improve their habitual bad posture. Therefore, we will develop a method of estimating body load without user burden for constant observation of posture. Hence, this study proposes the use of a web camera, which everyone has and can acquire images on a daily basis without any burden, as a non-contact sensing method, and the use of deep learning as a means of estimating body load from web images. Deep learning models are created by deriving body load values using musculoskeletal analysis based on skeletal position coordinates extracted from posture images and labeling the images with these as true values. Thus, if a pre-trained deep learning model is created in advance, body load can be estimated from images alone, without the use of specialized software or cloud communication. If it is possible to easily visualize one's own body load in daily life, the system can be developed to provide feedback on posture evaluation and improvement plans based on the estimated body load. We consider that this will further increase the users' awareness of improvement and lead to the maintenance and promotion of health. In this paper, as the first step, a deep learning model is created for a stationary standing forward bending posture, and the accuracy of the lumbar load estimation by the deep learning model is evaluated. The results of individual learning using untrained data allowed us to estimate the lumbar load with high accuracy. Hence, the possibility of applying the proposed method to certain individuals is indicated. The other is, the results of ensemble learning confirmed models with high and low accuracy. Hence, the deep learning models that estimated untrained participants showed large variations in accuracy and insufficient generalization performance. Discussion of the results confirms that data bias is a contributing factor to the accuracy loss and indicates the possibility of obtaining generalization performance by improving data bias.**

Keywords- Deep learning; Single camera; Estimation; Musculoskeletal model simulation; Lumbar load.

I. INTRODUCTION

Many people of all ages and genders experience lumbago, and lumbago has become a social problem because of its potential impact on daily life. One of the causes of lumbago is the habit of a broken posture, which is very demanding on the body. From this, to prevent lumbago, it is useful to constantly observe posture in daily life. In cases where posture is out of balance, it is effective to have a system that allows people to improve their posture by themselves. To achieve this, this research group has been considering the quantitative estimation of the load on the lower back in order to determine whether the posture is good or bad. In the past, the lumbar region has been measured using optical motion capture, wearable inertial sensors, and bending sensors to non-invasively estimate lumbar load using biomechanics and statistics [2] [3] [4] [5] [6]. These estimation results showed qualitatively similar trends to the measured lumbar load ratios of Nachemson et al. [7] and Wilke et al. [8] and confirmed the usefulness of the estimation method. However, since specialized equipment and analysis are required, and users are burdened during measurement and estimation, it is difficult to apply this method to the observation, estimation, and evaluation of posture in daily life.

Therefore, in this study, as a way to reduce the burden on the user during measurement, consider using an easily accessible, non-wearing sensing device. Muto et al. [9] evaluated the posture of an elderly person using Kinect v2 for Windows (Microsoft), a depth camera, as a non-contact sensing device. However, the depth cameras essential to this research are not widely available to the public. On the other hand, several systems have been commercialized to evaluate posture based on the skeletal position that is detected by AI from 2D images that lack depth information (e.g., Posen [10]). Although joint angles and other factors are visualized in these systems, however, the loads applied to the body are not quantified. Hence, this study proposes a method for estimating body load using AI from a single camera image, which is a readily available device, as a method of constantly observing posture by self and quantitatively estimating body load [1]. If the proposed method can be realized, by creating a deep learning model in advance using specialized software, it will then be possible to visualize one's own body load in daily life simply by inputting posture images, without going through the cloud. In addition, it can be developed into a system that provides feedback on posture evaluation and improvement plans and evaluations based on this, it will enable posture condition to be evaluated without burden and lead to the prevention of lumbago.

In this paper, as the first step in creating the proposed system, the deep learning model is created with only the lumbar load as the body load and the posture as a static standing forward-bending posture. After that, the proposed method will be evaluated by the accuracy of the lumbar load estimation using the deep learning model created. In the previous paper [1], a deep learning model was created using multiple experimental participants as training data and estimation using untrained data from participants used for training data. The results were shown to be useful, as high correlations and small errors were identified. However, scope of application remains unclear since the estimation of training participants by a deep learning model was created using several participants. Therefore, this paper verifies scope of application of the proposed method. First, it is verified that the proposed method is applicable to specific individual. A deep learning model is created using specific participant as training data, and its accuracy is confirmed by estimation using untrained data from the same participant. Furthermore, to verify the generalization performance of the proposed method. A deep learning model is created using multiple participants as training data, and the accuracy is confirmed by performing estimation on untrained participants.

The rest of this paper is organized as follows. Section II describes the methods of the lumbar load estimation system proposed in this paper. Section III determines the criteria for evaluating the accuracy of the deep learning model. Section IV provides the methods and conditions for creating the lumbar load estimation system proposed in this paper, and discusses the experimental results based on the evaluation criteria identified in Section III. Section V discusses the results of Section IV. Finally, the conclusions close the paper.

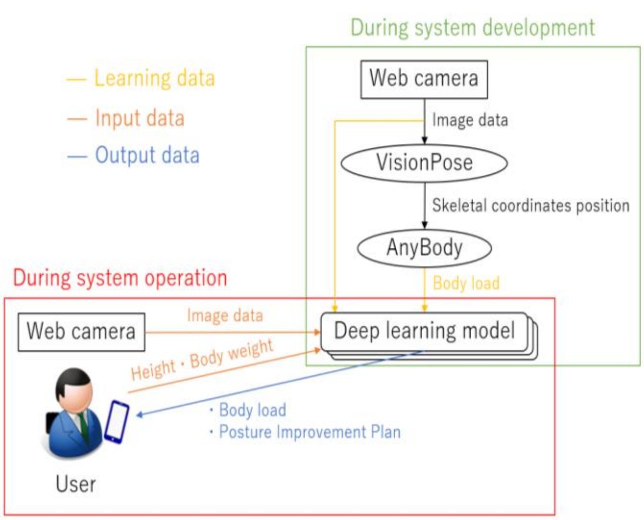

Figure 1. Overview of the proposed body load estimation system. [1]

II. LUMBAR LOAD ESTIMATION SYSTEM

This section describes the system proposed in this study. Figure 1 shows an overall view of the system to be developed in this study, as proposed in the previous report [1]. During system operation, user inputs an image of his/her posture, along with his/her height and weight, into the system, which estimates the body load and outputs improvement plans based on this. To achieve this, a deep learning model is created in advance during development.

A method for deriving the body load to be learned by the deep learning model is described. Tagawa et al. [11] proposed a device to visualize the dynamic load of various body parts from video alone using skeletal detection software. In this system, the body load is calculated using the Newton-Euler method based on the coordinates of the detected skeletal position. However, it is not suitable for estimating static posture, which is the participant of this study, because no acceleration occurs. In addition, it is difficult to obtain an accurate body load from an estimation based on skeletal position alone, because muscle activity and other factors cannot be considered. Therefore, in this study, body load is derived using AnyBody [12]. AnyBody is a musculoskeletal analysis software that can derive various human body information by creating a virtual human body model from skeletal positions. Also, AnyBody can be obtained the account the amount of muscle activity and other factors to determine the force, moment (torque), and muscle tension applied to a region. In the field of healthcare, much research has been conducted that make effective use of AnyBody. Previous research used AnyBody to analyze the effects of age and height on the lumbar region during manual material handling [13] and the effects of lumbar disc herniation on spin loading characteristics [14]. However, the input data for AnyBody are the skeletal coordinate positions of the human body. An optical motion capture camera is generally used, although this device cannot obtain information from the images. Therefore, AI skeletal detection software is used to detect skeletal coordinates from images. The skeletal coordinates detected using such software are used as input data to AnyBody. Thus, in this study, the skeletal position coordinates are detected from images using VisionPose [15]. VisionPose is one of the AI skeleton detection software, a highly accurate AI posture estimation engine that can detect skeletons from 2D camera images without using markers or depth sensors. VisionPose detects a total of 30 skeletal positions, including the hip and shoulder joints shown in Figure 2. Hence, in this study, the load applied to the body using AnyBody is derived from the skeletal position coordinates detected using VisionPose from the images. Out of the 30 locations detected by VisionPose shown on the right in Figure 2, the 15 locations in deficit are used to derive the load by AnyBody. After that, a deep learning model is created by labeling this as the true value with the image. As described above, during system development, specialized software such as AnyBody and VisionPose is used to create deep learning models. Then, based on this, when the system is used by the user, a structure will be built to enable estimation using only AI applications, without the need for several specialized software. These are the primary characteristics of the proposed system. It is a novel approach to combine AnyBody with AI.

In this paper, since we focus on lumbago, we use the lumbar load as the body load. In this process, a deep learning model is created for each body load that is appropriate for the posture to be observed because AnyBody can derive the load that occurs in any region from single image data. That is, many body loads can be estimated and visualized by the proposed system from single image data.

Next, the lumbar loading used in this paper is described. According to previous research [16], positive agreement was observed between in vivo measurements of disc compression forces between L4L5 and the values derived by AnyBody, demonstrating the suitability of the AnyBody model. Based on these results, this paper uses the compression force of the intervertebral disc between L4L5 derived from AnyBody as the lumbar load. In a previous report [17], as a preliminary step in creating a deep learning model, we evaluated the compression force of intervertebral disc between L4L5 derived by AnyBody using the skeletal position coordinates detected by VisionPose from web images of standing forward bending posture. This result showed an increasing trend of disc compression force with forward bending of the upper body, as measured by Nachemson et al. [7]. In response to this result, as one of the training data for the compression force of the deep learning model, the intervertebral disc between L4L5 derived by AnyBody is used as the true value in this paper.

III. Establishment of Criteria for Evaluating the Accuracy of Deep Learning Models

In deriving the true value of the disc compression force for the proposed method, the relationship between the error in skeletal detection by VisionPose and the anterior tilt angle and the disc compression force will be clarified. This determines the evaluation criteria for the accuracy of the deep learning model to be created.

A. *Experiment*

Three male participants (age 21 ± 1.00, height 1.70 ± 0.02 [m], weight 67.0 ± 1.70 [kg]) agreed to participate in the experiment in advance after obtaining approval from the University Ethics Committee and explaining the experimental details to the participants. A standing static posture image is acquired for them to obtain the anterior tilt angle and lumbar load by skeletal detection with VisionPose. One webcam (StreamCam: logicool) is used to get video. The camera is placed at the distance of 3 [m] from the center of the participant's body and at a height of 0.85 [m] from the floor. The movies are shot at 1080p/30fps. Three pictures are taken in each of the following conditions using the webcam: upright posture (0 degrees), 10 degrees, 20 degrees, and 30 degrees of forward tilt angle of the upper body. The angle ofc forward bend is determined by pressing the board against the lower back and measuring with a digital angle meter.

Based on the obtained images, one is, the skeletal position coordinates indicating the body center shoulder and hip positions detected by VisionPose are used to calculate the anterior tilt angle using a trigonometric function. After that, the error is calculated from the results of the measured and calculated values. The other is, based on the images, the skeletal position coordinates are detected from each image using VisionPose, and the skeletal position coordinates are input to AnyBody to derive the compression force of intervertebral disc between L4L5. In this process, the height and weight in the human body model in AnyBody are standardized to the participant's average in order to eliminate differences in the participant's physique in the derived values.

B. *Estimation Results*

Table I shows the error between the calculated and measured values of each forward tilt angle of the upper body, and the compression force of the intervertebral disc between L4L5 derived by AnyBody. Table I shows that the average absolute error of the forward tilt angle of the upper body detected by VisionPose is 3.20 [°]. Furthermore, the mean of standard deviation of the derived disc compression force between L4L5 was 13.1 [N], which is approximately 2.00 [%] of the mean body weight, indicating a high accuracy with little variation between participants of the data. In addition, from the derivation results shown in Figure 3, it can be read that the L4/L5 intervertebral disc

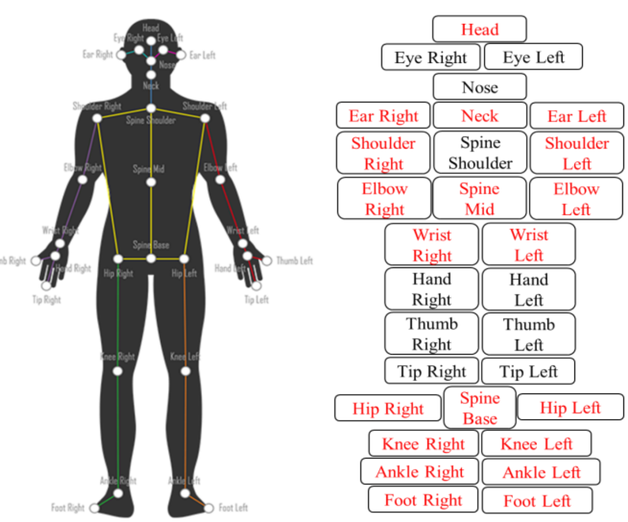

Figure 2. Skeletal position coordinates to be detected by VisionPose. Skeletal positions to be used for AnyBody are shown in red. (Source [15] on the left of the image)

TABLE I. MEAN ABSOLUTE ERROR BETWEEN CALCULATED AND MEASURED VALUES AND COMPRESSION FORCE OF INTERVERTEBRAL DISC BETWEEN L4L5 AT EACH UPPER BODY FORWARD TILT ANGLE

Upper body forward tilt angle [°]	0	10	20	30	Mean value
Mean absolute error of angle [°]	2.17 ± 0.3	2.58 ± 1.1	4.58 ± 0.9	3.46 ± 4.3	3.20 ± 0.924
L4/L5 intervertebral disc load [N]	365 ± 16.4	584 ± 6.50	798 ± 10.3	998 ± 19.2	

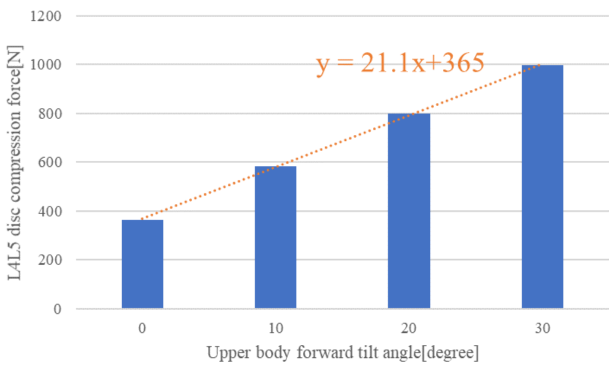

Figure 3. Compression force of intervertebral disc between L4L5 derived using AnyBody for each angle of forward tilt of upper body.

load increases almost linearly in 10-degree increments. The estimation results captured the trend of increased lumbar loading due to forward tilt of the upper body, as described in orthopedic clinical practice. Therefore, the lumbar load derived by AnyBody from the skeletal position coordinates detected by VisionPose can be used as the true value of the training data for the deep learning model. Following the results, the slope of the linear function of the approximate line was calculated to obtain an average change in compressive force per unit angle of 21.1 [N]. These results will be used to evaluate the deep learning models to be created in subsequent sections.

IV. ESTIMATION OF LUMBAR LOAD USING DEEP LEARNING MODEL

This section describes the experimental methods used to collect training and validation data, the preprocessing applied to the measured data, and the training conditions for CNN. Furthermore, the created deep learning model is used to estimate the lumbar load and confirm its accuracy.

A. *Experiment*

Three male participants (age 23.2 ± 0.748, height 1.73 ± 3.49 [m], weight 67.2 ± 4.35 [kg]) agreed to participate in the experiment in advance after obtaining approval from the University Ethics Committee and explaining the experimental details to the participants. To efficiently acquire posture images for use in the creation and accuracy validation of the deep learning model, video is captured for the forward bending motion of standing posture. The equipment used and camera locations are the same as in the experiment in Section III. The body gradually bends from an upright standing posture to about 30 degrees in 2 seconds, then the body gradually raises in 2 seconds to an upright standing posture. This is taken as one trial, and 5 trials are obtained. A total of three videos are obtained for each participant.

B. *Estimation Methods*

In this estimation, it is desirable to obtain the posture load at a specific point in time, so frame-by-frame images should be used for learning and estimation, rather than processing with video that includes time information. Hence, the video obtained by the experiment for a total of 15 trials for 5 using the Python module OpenCV (image processing library). The video of each participant is converted to an image at each frame rate, generating 1800 images per participant for a total of 9000 images. In addition, the lumbar load to be used as the true value is obtained by using VisionPose to detect the skeletal position coordinates from the videos of 15 participants in the trials. The first through fourth images of each trial as training data and the fifth image as validation data. The training data for each model are 1440 images and the of 5 participants in the same experiment. The derived disc compression force is normalized by dividing it by height and weight to eliminate differences due to body size. The normalized values are labeled as the true values for each frame of training data to create a deep learning model. After that, using the deep learning model created, estimation of lumbar load is performed on the validation data, and the normalized values are converted to disc compression force [N] by multiplying by height and weight.

TABLE II. CNN LEARNING CONDITIONS

		Set value
Batch size		64
Classes		100
Epochs		200
Dropout		0.2
Convolution layer	Filter size1	32
	Filter size2	64
	Stride	1
Pooling layer	Size	(2, 2)
Fully connected layer		64

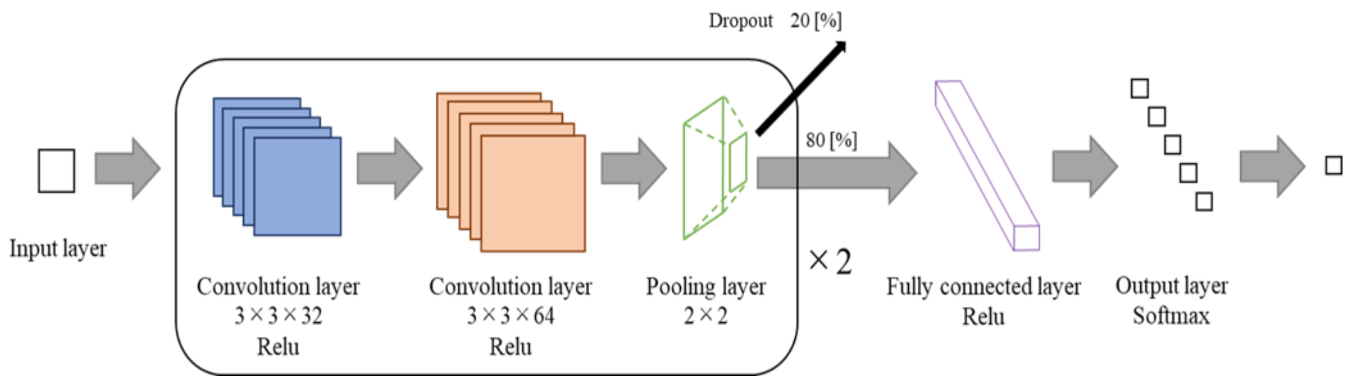

Figure 4. The structure of the CNN learning conditions.

In this paper, we use Convolutional Neural Network (CNN) for deep learning to estimate lumbar load. Numerical estimation is based on images, we consider that the same mechanism can be used for estimation as in the classification problem. The values of each parameter are shown in Table II. Figure 4 shows the structure of the CNN learning conditions used in this report. The CNN structure in the deep learning model consists of an input layer, followed by two convolutional layers, one pooling layer, two convolutional layers with dropout to prevent overfitting, one pooling layer, smoothing to prevent dropout to prevent overfitting, and output to an output layer after passing through all coupled layers one layer. The Relu function is used as the activation function in the convolution layer, the Relu function is used in the all-coupling layer, the Softmax function is used in the output layer, and Adam [18] is used for optimization. Keras Documentation [19] was used to create the above structure in Python. Keras.Callbacks.EarlyStopping is used as the termination condition, with the training error used as the monitor and auto as the mode.

C. *Individual Learning*

In this section, a deep learning model created for an individual confirms the applicability of the proposed method to a specific individual. A total of five deep learning models (Models A, B, C, D, and E) are created for each validation data, which are 360 images. The accuracy of the lumbar load estimated by the deep learning model is evaluated from each of the 360 images of the validation data. Figure 5 plots the estimates for each angle of forward tilt. The anterior tilt angle is calculated using a trigonometric function with the skeletal position coordinates indicating the body-centered shoulder and hip positions detected by VisionPose, based on the upright posture as 0 [°]. All model estimation results captured the trend of increased lumbar load due to upper body forward tilt as described in orthopedic clinical practice. Then, the lumbar load derived from the same verification data using AnyBody is compared to the estimated value as the true value. Figure 6 plots the estimated values from the deep learning model and the true values derived by AnyBody. Table III shows the Pearson's correlation coefficient and mean absolute error for each deep learning model. Pearson's correlation coefficients were 0.993 at maximum, 0.978 at minimum, and 0.987 ± 0.00770 at mean, indicating a high correlation in all models. In addition, the mean absolute error between the deep learning estimates and the true values derived by AnyBody was a maximum of 28.8 [N], a minimum of 22.5 [N], and an average of 26.3 ± 5.22 [N]. This is approximately 3.91 [%] of the average weight. Further, based on the results of the experiment described in Section III-B, the compression force changes by 21.1 [N] per 1 [°] of forward tilt angle. The average absolute error of the results of this experiment is equivalent to an error of 1.25 [°] of forward tilt angle, which is smaller than the average detection error of 3.20 [°] for the forward tilt angle in VisionPose. Therefore, the error is small.

In response to this result, the deep learning model for the standing forward bending posture that was created was able to estimate the lumbar load of the participant with high accuracy, and the proposed method is applicable as a lumbar load estimation method for the individuals used in the training. Thus, the user can check the posture change by own self by creating a deep learning model specialized for own self in advance.

TABLE III. PEARSON'S CORRELATION COEFFICIENT AND MEAN ABSOLUTE ERROR FOR EACH DEEP LEARNING MODEL CREATED FOR AN INDIVIDUAL

	Model A	Model B	Model C	Model D	Model E
Pearson's correlation coefficient	0.983 ± 0.0101	0.994 ± 0.00172	0.993 ± 0.00271	0.978 ± 0.0178	0.990 ± 0.00617
Mean absolute error [N]	28.4 ± 4.05	25.1 ± 3.79	26.8 ± 5.91	28.8 ± 7.39	22.5 ± 4.94

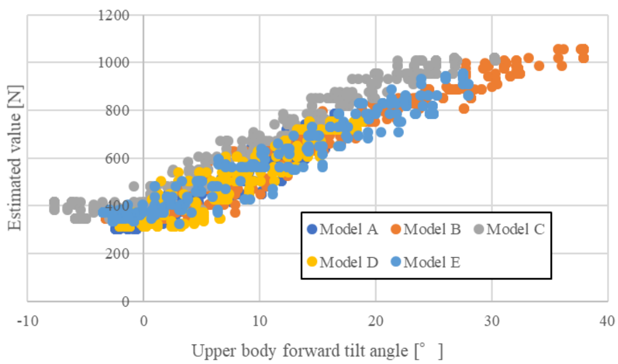

Figure 5. Intervertebral disc compression forces estimated for untrained data by a deep learning model created for an individual at each anterior tilt angle. For example, Model A plots the results of a deep learning model created for participant a using the first through fourth images of each trial as training data, and the fifth image of each trial is estimated as validation data.

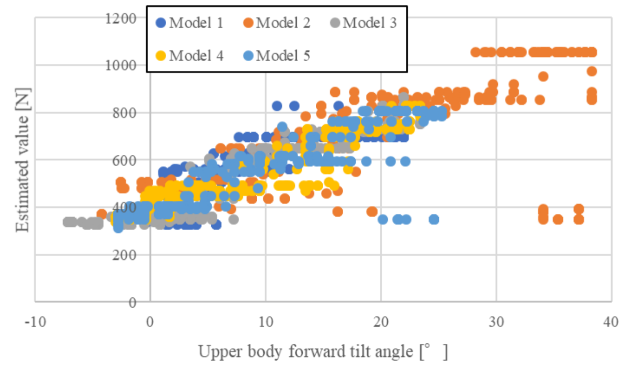

Figure 7. Intervertebral disc compression forces estimated for untrained persons by deep learning models trained on multiple people for each anterior tilt angle. For example, Model 1 plots the results of a deep learning model created with participants B, C, D, and E as training data, with participant A estimated as validation data.

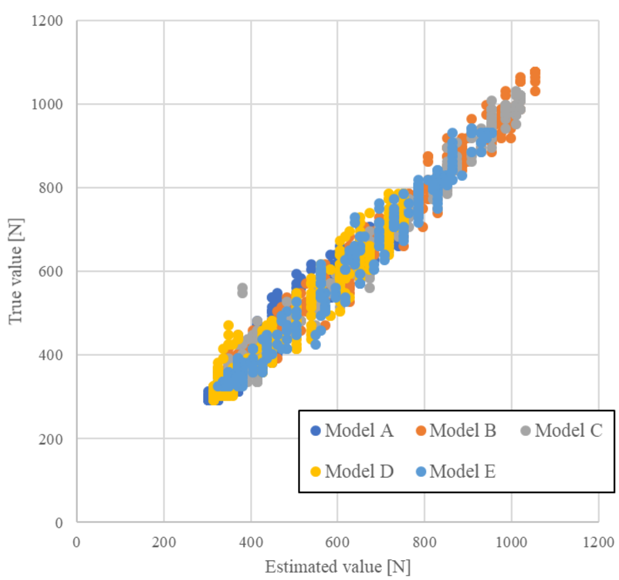

Figure 6. Comparison of disc compression forces estimated for untrained data by a deep learning model trained on an individual with the true values derived by AnyBody.

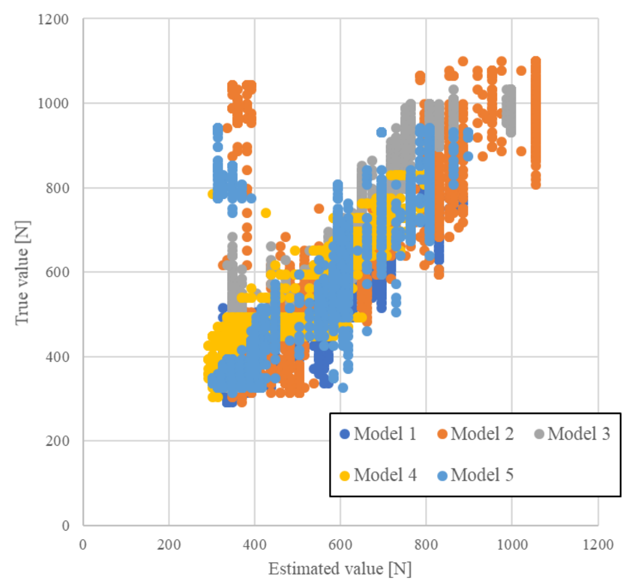

Figure 8. Comparison of disc compression forces estimated for untrained participants by a deep learning model trained on multiple people with the true values derived by AnyBody.

D. Ensemble Learning

In this section, a deep learning model created for several participants is used to verify the generalization performance of the proposed method. A total of five deep learning models (Models 1, 2, 3, 4, and 5) are created by cross-validating four of the five participants as training data and one as validation data. The training data for each model are 7200 images and the validation data are 1800 images. Evaluate the accuracy of the lumbar load estimated by the deep learning model from each of the 1800 images of the validation data. Figure 7 plots the estimated values for each angle of forward tilt. Although there were some outliers, all models captured the trend of increased lumbar load due to forward tilt of the upper body, as described in orthopedic clinical practice. Then, the lumbar load derived from the same verification data using AnyBody is compared to the estimated value as the true value. Figure 8 plots the estimated values from the deep learning model and the true values derived by AnyBody. Table IV shows the Pearson's correlation coefficient and mean absolute error for each deep learning model. Pearson's correlation coefficient was 0.966 at maximum, 0.800 at minimum, and 0.888 ± 0.0274 at the mean, indicating a high correlation, although inferior to individual learning. However, Figure 8 shows that

TABLE IV. PEARSON'S CORRELATION COEFFICIENT AND MEAN ABSOLUTE ERROR FOR EACH DEEP LEARNING MODEL WITH MULTIPLE PEOPLE TRAINED

	Model 1	Model 2	Model 3	Model 4	Model 5
Pearson's correlation coefficient	0.947 ± 0.0226	0.801 ± 0.0357	0.966 ± 0.00509	0.921 ± 0.0237	0.803 ± 0.0498
Mean absolute error [N]	54.2 ± 10.4	89.4 ± 9.44	126 ± 9.49	57.0 ± 6.13	69.5 ± 13.8

outliers were observed when the estimated values were between 300 and 400 [N]. In addition, the mean absolute error between the deep learning estimates and the true values derived by AnyBody was 126 [N] at maximum, 54.2 [N] at minimum, and 79.2 ± 9.84 [N] on average. This is approximately 11.9 [%] of the average weight. Based on the results of the experiment described in Section III-B, the compression force changes by 21.1 [N] per 1 [°] of forward tilt angle. The average absolute error of the results of this experiment is 3.75 [°] of forward tilt angle, which is a higher value than the average detection error of 3.20 [°] for the forward tilt angle in VisionPose. Therefore, the error is large.

However, Model 1 with the smallest average absolute error has an error equivalent to a forward tilt angle of 2.57 [°], which is smaller than the average detection error of 3.20 [°] for the VisionPose's forward tilt angle, and thus can be estimated with a small error.

The deep learning models created for the standing forward bending posture target did not show sufficient generalization performance in estimation for untrained participants, due to variations in accuracy caused by some models satisfying the evaluation criteria and others not.

V. CONSIDERATION

The deep learning model created in Section IV-3 was used to estimate the lumbar load of untrained participants in several participants. The results showed that the estimation accuracy varied and the proposed method did not demonstrate sufficient generalization performance. This section considers the causes of this result and offers prospects for improving the accuracy of deep learning models.

First, Figures 9 and 10 show the error rate and accuracy of Model 1. Model 1 has a relatively good Pearson's correlation coefficient and mean absolute error, with few outliers, among all deep learning models. Figure 9 shows the error rate per epoch during training for the Model 1 deep learning model. In the training data, the loss function decreases as the number of epochs increases. However, in the validation data, the loss function increases after a certain point. Furthermore, Figure 10 shows the accuracy per epoch during training for the Model 1 deep learning model. In the training data, the percentage of the accuracy increases as the number of epochs increases. However, in the validation data, there is no change in the accuracy value at a certain point in time. These figures suggest that overfitting has occurred. Citation [20] has been validated using MNIST and states that one of the causes of overfitting is lack of data. The 7200 training data for the deep learning model created in this paper are extremely small compared to the 50000 training data for CIFAR-100, a data set with the same number of classifications. This indicates insufficient training data. Hence, the deep learning model is expected to be improved by increasing the training data. However, collecting huge amounts of data through experiments is costly and labor intensive. Therefore, one idea is to artificially edit image data through data expansion, as in previous research [21], to increase the training data without involving any actual experiments. Thus, it is expected to lead to an improvement in the accuracy of deep learning models.

Second, Figure 11 shows the distribution of the image data acquired from the experiment for per normalized disc compression force. Figure 12 shows the distribution of

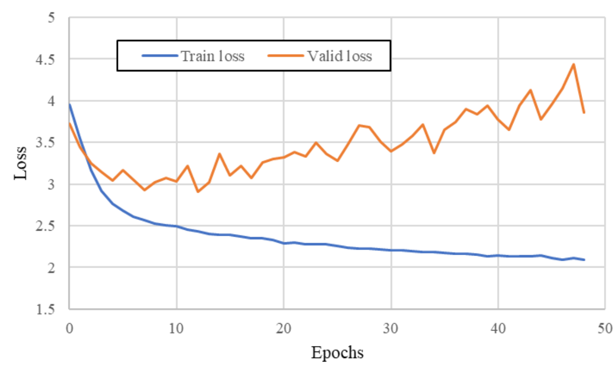

Figure 9. Error rates for training and validation data for Model 1.

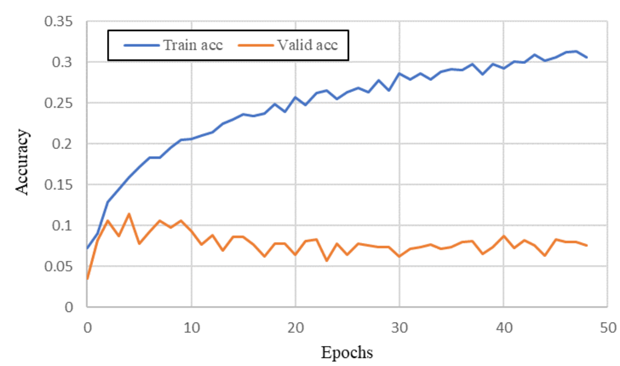

Figure 10. Accuracy for training and validation data for Model 1.

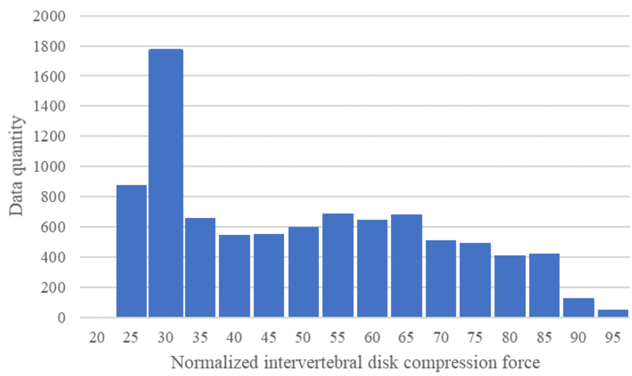

Figure 11. Distribution of the number of image data acquired in the experiment per normalized disc compression force.

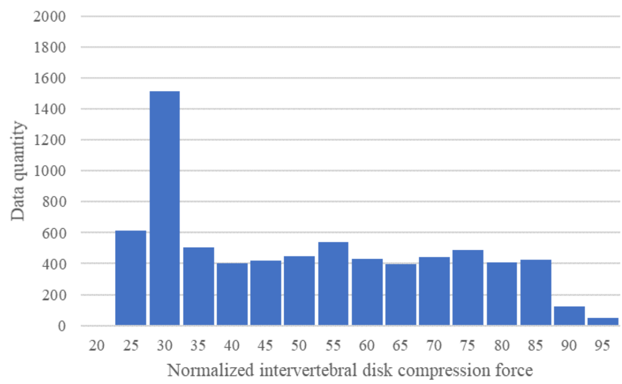

Figure 12. Distribution of the number of images used as training data for Model 1 per normalized disc compression force.

image data per normalized disc compression force used to train Model 1. Figure 12 shows that the amount of image data in the range of 30 to 35 normalized compressions used to train Model 1 was more than the other compressions. Further, a similar trend can be seen in all other models. Hence, the occurrence of outliers due to bias in the number of training data is another factor thought to reduce accuracy. Therefore, to verify the occurrence of outliers due to the bias in the number of training data, the number of image data is randomly deleted so that the number of image data between 30 to 35 becomes 800, the same level as the other range. A deep learning model was created using training data that had been adjusted to reduce bias in a simplified manner by this process, and the accuracy of the model was verified. Figure 13 plots the estimated values for each angle of forward tilt. All results of model estimation captured the trend of increased lumbar load due to forward tilt of the upper body as described in orthopedic clinical practice. Figure 14 plots the estimated values by the deep learning model using the training data after bias adjustment and the true values derived by AnyBody. Table V shows the Pearson's correlation coefficient and mean absolute error for each deep learning model created using the training data after bias adjustment. Pearson's correlation coefficients were 0.971 at maximum, 0.843 at minimum, and 0.936 ± 0.0361 at the mean, reducing the outliers seen in Figure 8 when the estimates were 300~400 [N], and showing a higher correlation than the results in Section IV-3, in which the training data were biased. In addition, the mean absolute error between the deep learning estimates and the true values derived by AnyBody was 126 [N] at maximum, 44.7 [N] at minimum, and 80.1 ± 12.37 [N] on average. Next, the bias adjustment is made to confirm what changes occur in each of the deep learning models. For Model 2 and Model 5,

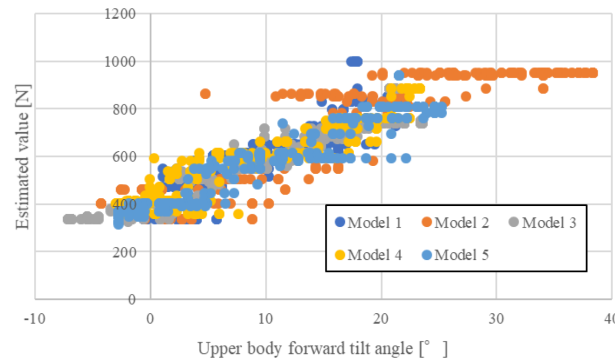

Figure 13. Intervertebral disc compression forces estimated for untrained persons by deep learning models trained on multiple people using bias-adjusted training data for each anterior tilt angle.

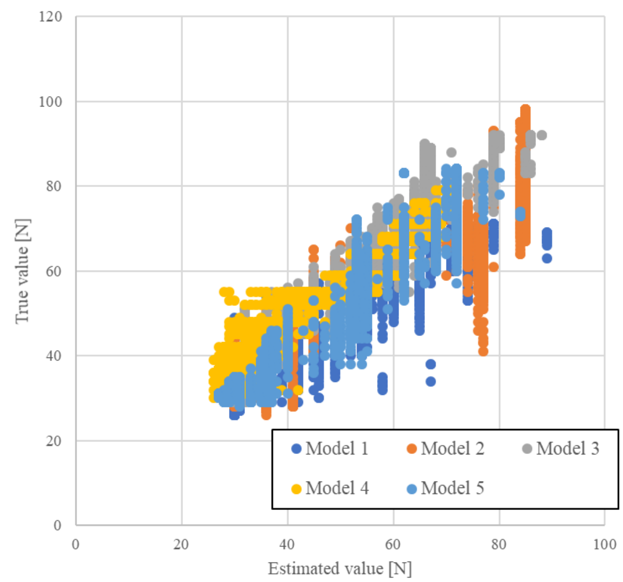

Figure 14. Comparison of disc compression forces estimated for untrained participants by a deep learning model trained on multiple people using bias-adjusted training data with the true values derived by AnyBody.

TABLE V. PEARSON'S CORRELATION COEFFICIENT AND MEAN ABSOLUTE ERROR FOR EACH DEEP LEARNING MODEL WITH MULTIPLE PEOPLE TRAINED USING BIAS-ADJUSTED TRAINING DATA

	Model 1	Model 2	Model 3	Model 4	Model 5
Pearson's correlation coefficient	0.843 ± 0.159	0.938 ± 0.00524	0.972 ± 0.00439	0.960 ± 0.00674	0.968 ± 0.00590
Mean absolute error [N]	80.0 ± 35.2	73.8 ± 3.18	126 ± 9.47	80.3 ± 6.65	44.7 ± 7.32

where many outliers can be seen in Figures 7 and 8, the bias adjustment significantly improved the outliers. Accompanying this improvement was an increase in accuracy in both Pearson's correlation coefficient and absolute mean error. However, Model 1 after bias adjustment was less accurate than Model 1 before bias adjustment for both Pearson's correlation coefficient and the absolute value of the mean error. Otherwise, Model 4 after bias adjustment improved the correlation coefficient and worsened the mean absolute error. Model 3 confirmed no change in accuracy due to bias adjustment.

In the two indices used to evaluate the accuracy of this paper, the accuracy by average of all models was lower than the results in Section IV-3, in which the training data were biased. In some of these cases, the correlation coefficient improved and the mean absolute error worsened, while in other cases both Pearson's correlation coefficient and mean absolute error worsened. Although, the outliers are eliminated in all models. Figures 15 and 16 show the error rate and the accuracy per epoch during training for the Model 1 deep learning model after bias adjustment. Although a minute change, the results of the validation data track the results of the training data, indicating that overfitting can be prevented by correcting the bias in the data. Hence, the possibility of improving the accuracy of the deep learning model was observed by homogenizing the training data. Furthermore, in this study, the training data was acquired through continuous repetition of forward bending movements, thus there is probably room for improvement with respect to this approach.

If these problems can be improved and applied to untrained users, it will be possible to estimate lumbar load with a deep learning model prepared in advance, without having created a deep learning model specific to the individual in advance. That is, the scope of application can be expanded to include untrained user for general use. As a result, the system will not only improve the posture of users who habitually have bad posture, but also enable healthy users to easily use the system as a preventive measure.

VI. CONCLUSIONS

To prevent lumbago, it is effective to constantly observe the posture of daily life. Therefore, we will develop a method to quantitatively estimate and visualize the load applied to the user's own body without any burden on the user. To achieve this, this study proposes a body load estimation method used on a deep learning model that uses web images and body load derived by AnyBody as training data. In this paper, as a preliminary step, we created a deep learning model using only lumbar load as the body load and assuming a stationary standing forward bending posture. After that, the accuracy of estimating lumbar load from web images using the created deep learning model was evaluated.

In the individual learning model, a high correlation was obtained between the estimates by the deep learning model and the true values derived by AnyBody, indicating that the errors were small. Therefore, it is possible to create a deep learning model in advance specifically for a specific user by using specialized software to create a deep learning model to be applied to that user. Thereafter the user to estimate the lumbar load in the target posture simply by inputting images, and can check the posture transition by own self. Hence, improvement of posture suited to individual will be possible without repeated visits to the hospital.

Figure 15. Error rates for training and validation data for Model 1 after bias adjustment.

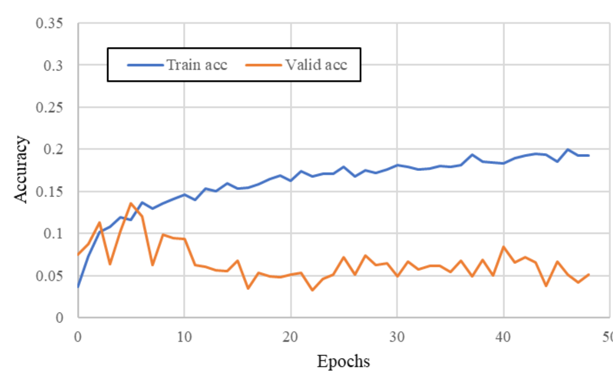

Figure 16. Accuracy of training and validation data for Model 1 after bias adjustment.

The other is, ensemble learning models showed high correlations, but models with high and low accuracy were identified, with a large variation in accuracy, and they did not show sufficient generalization performance. However, it was confirmed that data bias was one of the contributing factors to the lower accuracy. Therefore, if the data bias is improved, the proposed method has the potential to be applicable as a lumbar load estimation method to untrained persons.

The deep learning model created in this paper only covers the estimation of lumbar load in the forward bending posture of the upper body. However, the AnyBody used to derive the lumbar load in this paper can derive various body loads on the body from single measured data. Therefore, various body loads are obtained from a single image data, and a new deep learning model is created using this as training data. That is, it is possible to estimate selected body loads in each posture by acquiring various loads applied to each body part from images of postures that are considered to have a large load on the lumbar, such as hunching back and warped back, including the upper body forward bending posture targeted in this paper, and learning them together with the images. In addition, the accuracy of the deep learning model is improved by optimizing the program through filtering and attention mechanisms.

If these methods can be used to estimate the body load of any posture with a high degree of accuracy, the system can be developed into a system that quantifies and presents the load based on the proposed methods, allowing users to observe their own posture without burden. Thus, raise awareness of improvement, prevent lumbago, and ultimately, maintain and promote health.

REFERENCES

[1] R. Nishimoto and K. Shibata, “Estimation of Lumbar Load from Webcam Images Using Convolutional Neural Network for Standing Forward Bending Stationary Posture,” GLOBAL HEALTH 2022, Lecture No. 70027, 2022.

[2] K. Shibata, Y. Inoue, Y. Iwata, J. Katagawa, and R. Fujii, “Study on Noninvasive Estimate Method for Intervertebral Disk Load at Lumbar Vertebrae,” Transactions of the JSME, vol. 78, No. 791, 2012, pp. 130-141. (in Japanese)

[3] Y. Tsuyoshi, K. Shibata, M. Sonobe, and Y. Inoue, “A Method to Estimate Lumbar Intervertebral Disk Load Using Inertial Sensors for Development of Postures Improvement Support System,” The Japan Society of Mechanical Engineers Chugoku-Shikoku Branch, Student Organization, The 56th Graduation Research Presentation by Students, Lecture No. 112, 2018. (in Japanese)

[4] T. Iituka, K. Shibata, and Y. Inoue, "Estimation of Radius of Curvature of Lumbar Spine Using Bending Sensor for Low Back Pain Prevention," Human-Computer Interaction – INTERACT 2015, vol. 9299 of the series, Lecture Notes in Computer Science, pp. 533-536, 2015.

[5] Y. Suzuki, K. Shibata, M. Sonobe, Y. Inoue, and H. Satoh, "Noninvasive estimation of lumber disk load during motion to improve the posture," AHFE 2017 International Conference on Safety Management and Human Factors, Advances in Intelligent Systems and Computing, vol. 604, pp. 578-588, Code 193719, 2017.

[6] H. Himeda, K. Shibata, and H. Satoh, "Estimation of Load on Lumbar Spine While Walking by Using Multiple Regression Analysis," Advances in Intelligent Systems and Computing, vol. 1205 AISC, pp. 282-288, 2020.

[7] B. J. G. Andersson, R. Örtengren, A. Nachemson, and G. Elfström, “Lumbar Disc Pressure and Myoelectric Back Muscle Activity during Sitting,” I. Studies on an Experimental Chair, Scand J Rehab Med 6, pp. 104-114, 1974.

[8] H.J. Wilke, P. Neef, M. Caimi, T. Hoogland, and L.E. Claes, “New In Vivo Measurements of Pressures in the Intervertebral Disc in Daily Life,” SPINE, vol. 24, No. 8, pp. 755-762, 1999.

[9] Y. Muto, M. Sugou, H. Ito, K. Tsumurai, Y. Hosono, and T. Muto, “Application Method of Kinect for Evoluation of Physical Distortion of Elderly Adult,” Human Interface Society, vol. 19, No. 3, 2017. (in Japanese)

[10] Posen, Inc., Available from: https://posen.ai/, September 2023.

[11] K. Tagawa, T. Kawan, and E. Matsuo, “Development of AI camera for visualizing physical workload,” The 63rd Conference of the Japan Human Factors and Ergonomics Society, No. 2D2-06, 2022. (in Japanese)

[12] AnyBody Technology A/S, Available from: https://www.anybodytech.com/, December, 2023.

[13] T. Chihara, K. Iwahara, and J. Sakamoto, “Effect of age-related muscle weakness and height on L5/S1 compression force during manual material handling,” Transactions of the JSME, vol. 85, No. 876, 2019, p. 19-00125. (in Japanese)

[14] S. Kuai, W. Liu, R. Ji, and W. Zhou, “The Effect of Lumbar Disc Herniation on Spine Loading Characteristics during Trunk Flexion and Two Types of Picking Up Activities,” Journal of Healthcare Engineering vol. 2017, doi: 10.1155/2017/6294503.

[15] NEXT-SYSTEM Co., Ltd., Available from: https://www.next-system.com/visionpose/, December, 2023.

[16] T. Bassani, E. Stucoviz, Z. Qian, M. Briguglio and F. Glbusera, “Validation of the AnyBody full body musculoskeletal model in computing lumbar spine loads at L4L5 level,” Journal of Biomechanics, vol. 58, pp. 89-96, 2017.

[17] R. Nishimoto and K. Shibata, “Estimation of Lumbar Load from Webcam Images Using Musculoskeletal Model Simulation,” The Japan Society of Mechanical Engineers Chugoku-Shikoku Branch, Student Organization, The 50th Graduation Research Presentation by Students, Lecture No. 05b2, 2022. (in Japanese)

[18] J.L. Ba and D.P. Kingma, “Adam: A method for stochastic optimization,” In Proceedings of International Conference on Learning Representations, 2015.

[19] Keras Documentation, Available from: https://keras.io/, August, 2023.

[20] K. Saito “Deep Learning from the Basic,” p. 190, O’Reilly Japan, 2016. (in Japanese)

[21] L. Taylor and G. Nitschke, "Improving Deep Learning with Generic Data Augmentation," 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, pp. 1542-1547, 2018.