Tiller estimation method using deep neural networks

Rikuya Kinose
Osaka Prefecture University
Yuzuko Utsumi (yuzuko@omu.ac.jp)
Osaka Metropolitan University
Masakazu Iwamura
Osaka Metropolitan University
Koichi Kise
Osaka Metropolitan University

Method Article

Keywords: Tiller number estimation, Deep neural network (DNN), Pretext task, Self-supervised learning, Regression

Posted Date: April 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1552723/v1

METHODOLOGY

Tiller estimation method using deep neural networks

Rikuya Kinose[1†]
, Yuzuko Utsumi[2*†]
, Masakazu Iwamura[2]
and Koichi Kise[2]

[*]Correspondence: yuzuko@omu.ac.jp
[2]Department of Core Informatics, Graduate School of Informatics, Osaka Metropolitan University, 1–1, Gakuencho, Naka, Sakai, 599–8531 Osaka, Japan
Full list of author information is available at the end of the article
[†]Equal contributor

Abstract

Background: A tiller is a branch on a grass plant, and the number of tillers is one of the most important determinants of yield. Traditionally, the tiller number is usually counted by hand, and so an automated approach is necessary for high-throughput phenotyping. Conventional methods use heuristic features to estimate the tiller number. Based on the successful application of DNNs in the field of computer vision, the use of DNN-based features instead of heuristic features is expected to improve the estimation accuracy. However, as DNNs generally require large volumes of data for training, it is difficult to apply them to estimation problems for which large training datasets are unavailable. In this paper, we use two strategies to overcome the problem of insufficient training data: the use of a pretrained DNN model and the use of pretext tasks for learning the feature representation. We extract features using the resulting DNNs and estimate the tiller numbers through a regression technique.

Results: We conducted experiments using a dataset of Setaria viridis. Experiments show that the proposed methods using a pretrained model and specific pretext tasks achieve better performance than the conventional method. The best mean absolute error between the hand-labeled and estimated tiller numbers by the proposed method is 0.57.

Conclusions: We realized applying DNN methods to tiller number estimation methods by using pretext tasks. The proposed method outperformed the conventional approach.

Keywords: Tiller number estimation; Deep neural network (DNN); Pretext task; Self-supervised learning; Regression

1 Background

A tiller is a branch of a grass plant. For members of the grass family such as rice and wheat, the number of tillers is one of the most important determinants of yield [1, 2]. Therefore, it is one of the traits that is targeted for phenotyping. Destructive surveys have commonly been used to count the number of tillers, because they are hard to count visually; leaves and tillers look similar, and the density of tillers tends to be highest at the base of the plant. However, destructive surveys present a bottleneck to phenotyping tasks because they are time-consuming and labor-intensive, making it impossible to trace the growth of the plants. To achieve nondestructive and automatic tiller number estimation, several image-based methods have been proposed [3, 4].

However, their estimation accuracy is generally poor because they are based on hand-made features. To estimate the tiller number, image-based approaches use hand-crafted features such as the area and aspect ratio of a plant within an image and the output of the Frangi filter [5] for linear regression. As these methods only use a few heuristic features of the plants' appearance, they do not take full advantage of the information contained in the images. The recent development of image recognition techniques using features learned by deep neural networks (DNNs) surpasses the performance of conventional hand-crafted feature-based methods [6, 7, 8, 9]. DNNs learn image features directly from the image appearance. Thus, the features learned by DNNs take full advantage of the plants' appearance. This motivates us to use DNNs to learn features as a means of realizing high-accuracy tiller number estimation.

DNNs requires large volumes of training data, consisting of pairs of an image and the corresponding ground-truth, to achieve their true recognition ability. For example, ImageNet [10] is a commonly used dataset for object recognition consisting of more than 14 million images and their ground-truths. The existing image dataset of grass plants [11] contains only around 600 images with the corresponding ground-truth tiller numbers. This is because the operation of counting the tiller numbers is time-consuming and labor-intensive, as mentioned above. Therefore, it is difficult to prepare sufficient training data for DNNs, making it almost impossible to apply DNN-based methods for tiller number estimation.

As a lack of training data is commonly encountered in the field of computer vision and pattern recognition, several methods have been developed to enable DNNs to be used with small-scale data. For example, transfer learning [12] transfers the network learning to another dataset, semi-supervised learning [13] uses partly labeled data for learning, and self-supervised learning [14] uses self-generating labels. Some self-supervised learning methods that learn features by solving other tasks have achieved comparable performance to supervised methods [14, 15, 16]. These other tasks are called "pretext tasks," and they can be applied to problems in which large numbers of unlabeled data are available.

In this paper, we describe the use of self-supervised learning and transfer learning to estimate the tiller number, even though there are relatively few training data [17]. To the best of our knowledge, this is the first attempt to estimate tiller numbers using DNN-based features. We apply transfer learning to the estimation task and examine how the features learned from other data affect the estimation. We also set some pretext tasks for learning DNNs and evaluate how the pretext tasks enhance the estimation performance. Experimental results show that the proposed method outperforms the conventional method and that the pretext tasks enhance the estimation accuracy.

1.1 Related work

We introduce some of the conventional research on tiller counting and DNN-based phenotyping. We also describe some previous studies on transfer learning and self-supervised learning, which we apply to develop an estimation method for the tiller number.

1.1.1 Tiller number estimation

Automated methods of counting and estimating the tiller number have been studied for many years. Several researchers have used remote sensing to estimate the tiller number (see Table 1 of [18]). The normalized difference vegetation index has been widely used to estimate the tiller numbers of wheat using remote sensing [18, 19, 20]. As remote sensing can acquire information over vast areas, it is suitable for estimating plant growth in large-scale fields. However, this approach cannot count the tiller numbers of individual plants.

To estimate individual tiller numbers, image-based tiller number estimation methods have been proposed [3, 4]. Fahlgren et al. [3] used the area and aspect ratio of a plant within an image as the dependent variables and established a tiller-counting linear regression model for Setaria viridis. The area and aspect ratio are automatically calculated using open-source software called PlantCV[1]. Boyle et al. [4] applied the Frangi filter [5] to wheat images and used linear regression of the Frangi filter output as the dependent variable. Because the features used for these methods are heuristic, it is not clear whether they are suitable for estimation. Moreover, because the features are only based on certain aspects of the appearance of the plants, these methods do not fully utilize the appearance information in estimating the tiller numbers. Unlike heuristic features, DNN-based features are learned from images. Thus, the DNN-based features express the appearance information of plants better than the heuristic features, and so DNN-based methods are expected to outperform conventional methods.

DNN-based tiller number estimation techniques have already been proposed [21, 22]. Deng et al. [21] applied DNN-based image detection to stubble images as a means of counting the tillers. However, this method requires a destructive survey, making it difficult to track the growth traits of the plants. The idea of counting tillers proposed by Wu et al. [22] is almost the same as that developed by Deng et al., except that the images are obtained using micro-CT. Unfortunately, micro-CT is too expensive to be widely used. Different from these methods, the proposed approach requires only an RGB image to estimate the tiller numbers. Therefore, it is suitable for easy and high-throughput phenotyping.

1.1.2 Image-based plant phenotyping using DNNs

To achieve high-throughput phenotyping, computer vision and pattern recognition techniques have been used to measure individual plant traits. The high recognition performance of DNNs has made it possible to measure various traits from images.

The most common task for DNN-based individual phenotyping is leaf counting. This task became very popular when leaf segmentation and counting challenges were included in computer vision problems at several plant phenotyping workshops. In conjunction with these challenges, an image dataset of Arabidopsis thaliana was released [23]. This dataset has since been used in the development of many methods [24, 25, 26]. However, the dataset has few image data in which the number of leaves is identified. Therefore, techniques that artificially increase the number of data using data synthesis based on plant models have been proposed, enabling DNNs to be applied to small sets of labeled data [25, 26]. This data synthesis approach

[1] `https://plantcv.danforthcenter.org`

cannot be easily applied to tiller number estimation because the structure of grass plants is too complicated to model.

In addition to counting the leaves of Arabidopsis thaliana, many traits have been estimated using DNNs. Roots are another typical subject for trait estimation using DNN-based image analysis. For example, segmentation algorithms for root regions [27, 28, 29] and root structure analysis based on the characterization of roots [30, 31] have been proposed. Certain traits of wheat, which is a member of the grass plant family, have also been estimated, such as the number of spikes and spikelets [32] and the emergence and biomass [33].

1.1.3 Pretext tasks

To apply DNNs to labeled training data, several methods using pretext tasks have been developed. Usually, these pretext tasks do not relate to the object tasks, and the labels of the pretext tasks are generated automatically. The model solves the pretext tasks instead of the object task, and is then transformed to solve the object task. Training DNNs using pretext tasks achieves comparable performance to supervised training.

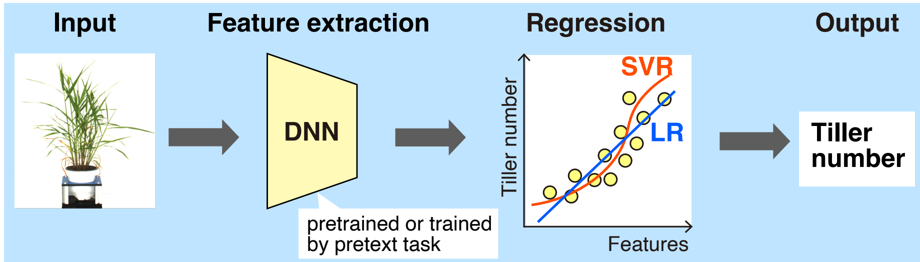
Various pretext tasks have been proposed. For example, colorizing images [34], solving jigsaw puzzles [35], predicting image rotations [14], and counting the number of objects within an image [16] have been used for representation learning. The learned representations are used for image segmentation, image recognition, and object recognition.

In establishing the proposed method, we set some pretext tasks for tiller number estimation according to these previous methods. The application of pretext task means that tiller number estimation can be conducted using DNNs, even if few labeled data are available.

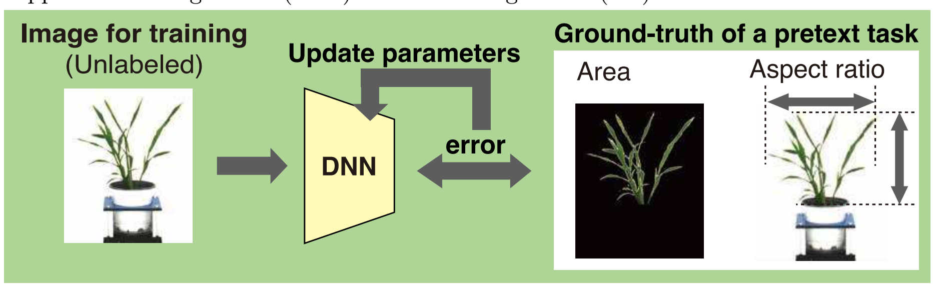
2 Methods

We explain how the proposed method estimates the tiller number from an image. We adopt regression-based estimation for tiller counting, as in conventional image-based tiller number estimation methods [3, 4]. This is because regression-based estimation is more practical than the detection-based method. Tillers have a similar appearance to leaves, and so it is hard to detect tillers from images. Moreover, the tillers become too dense to detect as the plant grows. Therefore, we adopt the regression-based method.

Figure 1a shows an overview of the proposed tiller number estimation method. This research takes the tiller number as the dependent variable and the image features extracted by the DNN model as the independent variables. As mentioned in Section 1, there are few images in existing grass plant datasets in which the number of tillers is labeled. Therefore, it is difficult to train DNN models directly for the tiller estimation task. This paper proposes the use of a pretrained model and a model trained on pretext tasks to extract image features, whereupon the tiller numbers can be estimated using the extracted features. Image resources and processing, the pretrained model, pretext tasks, and regression models are now described in detail.



(a) The proposed tiller estimation method. A plant image taken from the side is used for estimating the number of tillers. Features are extracted from the input image using a DNN. The DNN model is a pretrained model or a trained model by a pretext task. The tiller numbers are estimated by regression, and the extracted features are used as independent values for regression. The proposed method uses two regression models: support vector regression (SVR) and Linear regression (LR).



(b) Training the DNN by the pretext tasks. The pretext task in the proposed method is estimating the area and aspect ratio of the plant within the input image. The image for training is unlabeled with the tiller numbers. The ground-truth of the area and aspect ratio of the plant within the image is calculated using image processing beforehand. The DNN model is trained by updating the parameters as reducing the error between the output of the DNN and the ground-truth.

Figure 1: Overview of the proposed method and training DNN by a pretext task to acquire feature representations.

2.1 Image resources and processing

We used the dataset that appears in [11]. The first row in Fig. 2 shows some examples of the dataset. The dataset contains 25,570 images of potted Setaria taken from the side in a controlled laboratory environment. The images are in RGB color, and the image resolution is $2,454 \times 2,056$ pixels. In the dataset, 576 images have tiller numbers that were counted manually. Thus, there are 24,994 unlabeled images that have no tiller number. Many of the unlabeled images were taken at the same time as the labeled images. To avoid mixing unlabeled data that are similar to the labeled data, we only used the 22,110 unlabeled images that were not taken at the same time as the labeled images.

We normalized the images before the experiments. The magnification of the images was artificially determined according to the plant growth degree. As the first row of Fig. 2 shows, the pot size and the background differ depending on the plant size. If such images were used for learning, the network may learn features that focus on changes in the pots and backgrounds. To avoid the network focusing on parts unrelated to the plants, we normalized the images. Because all plants were in pots of the same size, the images were resized so that the pot size was the same.

Figure 2: Examples of the images used for the experiments. The first row shows original images taken at different magnifications. The second row shows normalized images in which the pot is the same size.

After removing the background, the images were translated to place the pot in the middle of the image. To make the images square for input to the network, the images were padded with white pixels. Finally, we resized the images to 224×224 pixels. All procedures were performed using OpenCV, and we used the bicubic method for pixel interpolation when the images were resized.

2.2 Feature extraction

2.2.1 Pretrained model

If there are insufficient labeled data for training, DNNs cannot achieve high performance. One solution is to use a pretrained model. In general, it is supposed that a pretrained model will achieve better performance when the target image domain is similar to that of the pretrained model. As we are estimating tiller numbers from natural images, a pretrained model using natural images is preferable for our task. Therefore, we use the pretrained VGG-16 model [6] which was trained using a vast number of natural images from the ImageNet dataset [10]. The VGG-16 model performed very well in the ImageNet ILSVRC2014 classification and localization task[2]. This model is popular because it extracts good feature representations despite having a simpler structure than other networks.

2.2.2 Pretext tasks

As mentioned in Section 1, it is impossible to learn the feature expression from the tiller number estimation task directly because of the shortage of labeled training data. Thus, we use pretext tasks to learn the feature expression, and estimate the tiller number using the learned features.

The VGG-16 model [6] is trained using pretext tasks that predict appearance-related values acquired automatically from a plant image. As shown in Fig. 1b, we set two pretext tasks: estimating the area of a plant within an image and estimating the aspect ratio of a plant. We consider the area and aspect ratio because they were used as the dependent variables for estimating the tiller numbers in a previous

[2] `http://www.image-net.org/challenges/LSVRC/2014/results`

study [3] and are expected to provide good feature expressions for tiller number estimation.

We investigate two methods of estimating the values in the pretext tasks: the values themselves and the discretized values. When estimating the value itself, the network is trained so that the output is the area or aspect ratio. We call the pretext task that estimates the value itself the "regression task," because in this case, the pretext task can be regarded as a regression task with the image as the independent value and the continuous values of area and aspect ratio as the dependent values. In the case of estimating discrete values, instead of outputting a numerical value, the network predicts the discretized values of the area and aspect ratio of the plants in the input image. Therefore, predicting discrete values is equivalent to classification. We call the pretext task estimating the discretized values the "classification task."

We conducted network training on the pretext tasks using the normalized images. The ground-truth of the pretext tasks was calculated automatically using image processing. Following the "Single plant RGB image workflow" in the PlantCV tutorial[3], the normalized images were translated into HSV and Lab images, and thresholding was applied to the saturation component of the HSV images and the a and b components of the Lab images. The plant area was then segmented by taking the logical sum of the threshold results. The area and aspect ratio were calculated from the segmented plant area. All processes were conducted using PlantCV [4]. The images were divided into four or eight classes in the classification task according to the area and aspect ratio values, respectively. The images were divided as the number of images in each class became the same.

The network was trained to predict the class to which the input image belongs. In the regression task, the network was trained to predict the area and aspect ratio of the input images. Both tasks used 80% of the images for training and 20% of the images for testing. The network used for training was the VGG-16 model pretrained by ImageNet. We trained the network 12 times and adopted the model that gave the lowest training error for tiller number estimation. We used the Keras TensorFlow2 backend to execute the training process.

2.3 Regression models

We use two regression models to estimate the tiller numbers, namely support vector regression (SVR) and linear regression (LR).

SVR involves the application of a support vector machine to regression. The most significant advantage of SVR is that it deals with nonlinear regression problems through the same framework as linear SVR. In SVR, a feature space can be mapped to a space of much higher dimension using a kernel function. When the kernel function is nonlinear, SVR can deal with nonlinear regression problems. Moreover, SVR can learn from small-sized datasets. Hence, we apply SVR to tiller estimation. Specifically, we extract features from labeled images using the models described in Sections 2.2.1 and 2.2.2, and then apply SVR.

We also use linear regression (LR) for the estimation task. LR is one of the simplest regression methods and is equivalent to a fully connected neural network without a

[3] https://plantcv.readthedocs.io/en/stable/tutorials/vis_tutorial/

[4] https://plantcv.danforthcenter.org/

Table 1: MAE of estimation results when using SVR and LR.

		Pretext task					
		Area			Aspect ratio		
Reg. model	Pretrained	4 cls.	8 cls.	Reg.	4 cls.	8 cls.	Reg.
SVR	0.80	0.74	0.78	0.91	0.73	0.73	1.00
LR	0.79	0.74	0.71	0.57	0.96	1.06	0.62

hidden layer. Because it is easy to implement LR with methods using DNN-based features, we apply LR for the estimation task. As with SVR, we learn the LR model using the features extracted from labeled images.

We estimated the tiller number using the features extracted by the trained and VGG-16 ImageNet pretrained models. In the case of SVR, we used scikit-learn [5] for the implementation, which is one of the most popular machine learning libraries for Python. The radial basis function was used as the kernel. The cost parameter C and parameter ϵ were set to 100 and 1.0, respectively, and default values were used for the other parameters. LR was implemented by adding a fully connected layer to the VGG-16 model. We then trained only the added layer while freezing VGG-16.

3 Results

3.1 Tiller number estimation

We used six-fold cross-validation to calculate the accuracy of the tiller number estimation. That is, the images were divided into six groups and the regression models were trained with five groups and validated with the remaining group. This process was repeated until all groups had been used for validation. The accuracy of the model was calculated by taking the average of each of the six cross-validation tasks. We adopt the mean absolute error (MAE) to evaluate the accuracy of the proposed method.

Table 1 presents the MAE when using SVR and LR to estimate the tiller numbers.

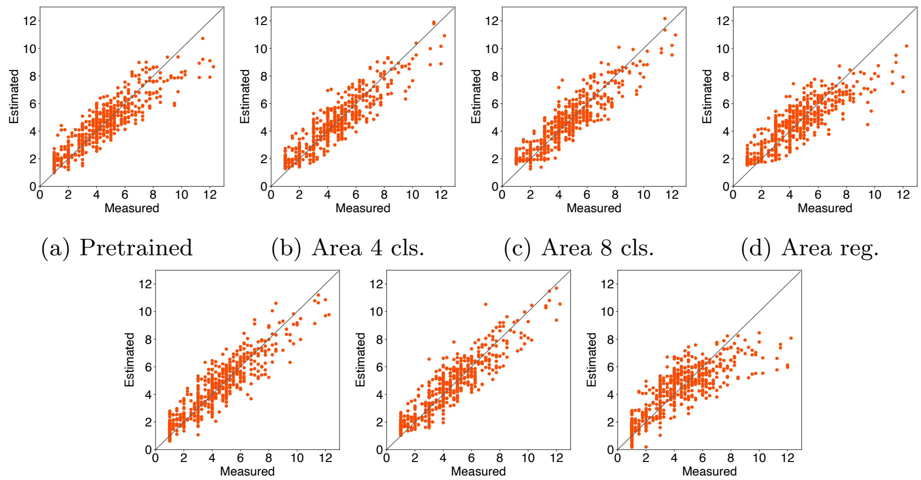
3.2 Individual estimation results

The measured tiller number (horizontal axis) and estimated tiller number (vertical axis) are compared in Figs. 3 and 4 for the cases using SVR and LR for tiller number estimation, respectively. In the figures, black lines indicate when the measured and estimated tiller numbers are the same, and the red and blue dots denote the samples. Thus, the closer the dots are to the black lines, the more accurate are the estimation results.

4 Discussion

This proposed method is the first attempt to apply self-supervised learning using pretext tasks for plant phenotyping, as far as we know. Plant datasets have insufficient labeled data for applying DNNs. The proposed semi-supervised method for estimating the number of tillers requires only a few labeled data. Therefore, the proposed method would be one of the solutions to use DNNs for estimating plant phenotyping.

[5] `https://scikit-learn.org/stable/`



(a) Pretrained (b) Area 4 cls. (c) Area 8 cls. (d) Area reg.

(e) Asp. ratio 4 cls. (f) Asp. ratio 8 cls. (g) Asp. ratio reg.

Figure 3: Experimental results using SVR for tiller number estimation. Horizontal and vertical axes represent the measured and estimated tiller numbers, respectively. Each red dot denotes a sample for the estimated tiller number. The black line indicates the case where the measured and estimated data match. Therefore, the closer the points are to the black line, the more accurate is the estimation.

The proposed method show good estimation accuracy. The best MAE of 0.57 is achieved when the area regression is used for the pretext task and LR is used to predict the tiller numbers. For reference, the MAE reported by Fahlgren et al. [3] for estimating the tiller number using image features with the same data under different conditions was 0.92. Note that it is not possible to make a general comparison because of the different image usage conditions and because Fahlgren et al. [3] used a different number of images to that in the dataset [11]. However, it appears that the proposed method achieves good accuracy.

To clarify the effect of feature learning in the pretext tasks, we compared the accuracy of the pretext tasks and pretrained models. Many of the pretext tasks resulted in higher accuracy than using the pretrained model. Therefore, learning features using pretext tasks contributes to improving the accuracy of estimating tiller numbers.

The tiller number estimation accuracy depends on the trait estimated in the pretext task. The tiller number estimation accuracy is better when the area is used in the pretext task than when the aspect ratio is used. Therefore, the features learned in the pretext task using the area are more effective for tiller estimation than those learned from the aspect ratio.

The pretext task that gives the better tiller number estimation accuracy also depends on the tiller number estimation method. When SVR is used, the application of classification in the pretext task results in better accuracy than regression. In contrast, when LR is used, the application of regression in the pretext task achieves better accuracy than classification.

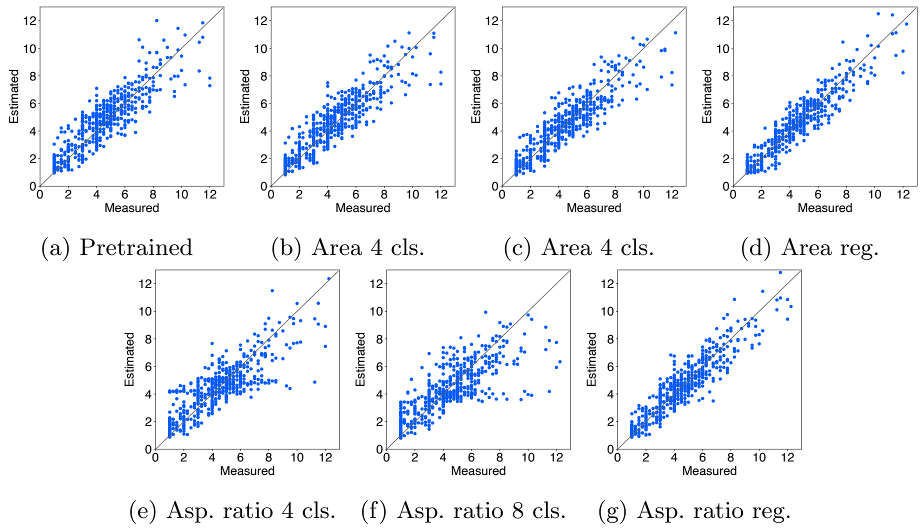


Figure 4: Experimental results using LR for tiller number estimation. The contents of the graph are the same as in Fig. 3

There is clearly a different tendency when SVR and LR are used for tiller number estimation. As shown in Fig. 2a, when the pretrained model features are used with SVR, the estimated tiller number is substantially underestimated when the measured tiller number is significant. When the features learned by classification tasks are used, the accuracy of the tiller number estimation improves for samples with larger tiller numbers, as shown in Figs. 2b, 2c, 2e and 2f, compared with pretrained model features. Thus, using classification for the pretext tasks improves the tiller number estimation accuracy. However, when using regression for the pretext tasks, the estimation accuracies are worse than those with the pretrained model. In particular, as shown in Figs. 2d and 2g, the estimation results for samples with larger tiller numbers are worse than those using the pretrained model.

When regression and LR were used for the pretext task and tiller number estimation, respectively, the estimation accuracy improves for all samples, as shown in Figs. 4d and 4g. In particular, comparing the pretrained model with the regression pretext task, the accuracy is enhanced for samples with large measured tiller numbers. In contrast, when the features learned by the classification task are used, the estimation accuracy is the same or worse than that of the pretrained model, as shown in Figs. 4b, 4c, 4e, and 4f. When the aspect ratio is used for the classification task, the estimation accuracy becomes worse, with the estimated tiller numbers consistently lower than the measured values.

In future work, we will use other pretext tasks to learn the feature representations. The mechanisms of the pretext tasks remain obscure, and it is not known what kinds of pretext tasks are most effective for a given object task. Therefore, we will attempt to determine the most appropriate pretext task for the object task by trial and error.

We also plan to apply the proposed method to other grass plant family such as wheat and rice. In this paper, we used a Setaria viridis dataset for evaluation. As most DNN-based methods have good generalization ability, the proposed methods is likely to be effective for other plants.

Additionally, we will apply the proposed method to other plant phenotyping tasks. The proposed method assumes that few labeled training data are available. This is typically true of plant phenotyping tasks because many appearance traits are measured manually. We expect that the proposed method will be helpful in automating the measurement of various traits.

5 Conclusion

This paper has proposed a DNN-based tiller estimation method that achieves improved performance compared with conventional methods. The proposed method uses two separate models for feature extraction: a pretrained VGG-16 model and a model produced by solving pretext tasks. We considered both SVR and LR to estimate the tiller numbers. Experimental results show that the pretrained model and the model based on pretext tasks allow the proposed method to outperform the conventional approach.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Availability of data and materials
The dataset analyzed during the current study are available in the figshare repository,
`https://figshare.com/articles/dataset/DDPSC_Phenotyping_Manuscript_1_Files/1272859` [11].

Competing interests
The authors declare that they have no competing interests.

Funding
This work was supported by JSPS KAKENHI Grant Number JP18K18074.

Authors' contributions
YU and MI contributed to the conception and design of the study. RK performed the statistical analysis. KK prepared the materials for the research. YU wrote the first draft of the manuscript. All authors contributed to manuscript revision, and have read and approved the submitted version.

Acknowledgments
We thank Stuart Jenkinson, PhD, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Author details
[1]Department of Computer Science and Intelligent Systems, Graduate School of Engineering, Osaka Prefecture University, 1–1, Gakuencho, Naka, Sakai, 599–8531 Osaka, Japan. [2]Department of Core Informatics, Graduate School of Informatics, Osaka Metropolitan University, 1–1, Gakuencho, Naka, Sakai, 599–8531 Osaka, Japan.

References

1. Badshah, M.A., Naimei, T., Zou, Y., Ibrahim, M., Wang, K.: Yield and tillering response of super hybrid rice Liangyoupeijiu to tillage and establishment methods. Crop Journal **2**(1), 79–86 (2014). doi:10.1016/j.cj.2013.11.004
2. Cai, T., Xu, H., Peng, D., Yin, Y., Yang, W., Ni, Y., Chen, X., Xu, C., Yang, D., Cui, Z., Wang, Z.: Exogenous hormonal application improves grain yield of wheat by optimizing tiller productivity. Field Crops Research **155**, 172–183 (2014). doi:10.1016/j.fcr.2013.09.008
3. Fahlgren, N., Feldman, M., Gehan, M.A., Wilson, M.S., Shyu, C., Bryant, D.W., Hill, S.T., McEntee, C.J., Warnasooriya, S.N., Kumar, I., Ficor, T., Turnipseed, S., Gilbert, K.B., Brutnell, T.P., Carrington, J.C., Mockler, T.C., Baxter, I.: A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in Setaria. Molecular Plant **8**(10), 1520–1535 (2015). doi:10.1016/j.molp.2015.06.005
4. Boyle, R.D., Corke, F.M.K., Doonan, J.H.: Automated estimation of tiller number in wheat by ribbon detection. Machine Vision and Applications **27**(5), 637–646 (2016). doi:10.1007/s00138-015-0719-5
5. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, pp. 130–137 (1998). doi:10.1007/BFb0056195

6. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: Proceedings of International Conference on Learning Rresentations (2015)
7. Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E.: Squeeze-and-excitation networks. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7132–7141 (2018). doi:10.1109/CVPR.2018.00745. 1709.01507
8. Taigman, Y., Yang, M., Ranzato, M., Wolf, L.: DeepFace: Closing the gap to human-level performance in face verification. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1701–1708 (2014). doi:10.1109/CVPR.2014.220
9. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815–823 (2015). doi:10.1109/CVPR.2015.7298682
10. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: ImageNet: A large-scale hierarchical image database. In: Proceedings of Computer Vision and Pattern Recognition, pp. 248–255 (2009). doi:10.1109/CVPR.2009.5206848
11. Gehan, M., Fahlgren, N., Feldman, M., Wilson, M., Hill, S., Bryant, D., Ficor, T., Turnipseed, S., Warnasooriya, S., Shyu, C., Gilbert, K., Kumar, I., McEntee, C., Brutnell, T.P., Carrington, J.C., Mockler, T.C., Baxter, I.: A versatile phenotyping system and analytics platform reveals diverse temporal responses to water limitation in Setaria (2015). doi:10.6084/m9.figshare.1272859.v12. `https://figshare.com/articles/dataset/DDPSC_Phenotyping_Manuscript_1_Files/1272859`
12. Huang, Y., Cheng, Y., Bapna, A., Firat, O., Chen, D., Chen, M., Lee, H., Ngiam, J., Le, Q.V., Wu, Y., Chen, z.: Gpipe: Efficient training of giant neural networks using pipeline parallelism. In: Advances in Neural Information Processing Systems, vol. 32 (2019). https://proceedings.neurips.cc/paper/2019/file/093f65e080a295f8076b1c5722a46aa2-Paper.pdf
13. Miyato, T., Maeda, S.-I., Koyama, M., Ishii, S.: Virtual adversarial training: A regularization method for supervised and semi-supervised learning. IEEE Transactions on Pattern Analysis and Machine Intelligence **41**(8), 1979–1993 (2019). doi:10.1109/TPAMI.2018.2858821
14. Gidaris, S., Singh, P., Komodakis, N.: Unsupervised representation learning by predicting image rotations. In: Proceedings of International Conference on Learning and Representations, pp. 1–16 (2018)
15. Noroozi, M., Vinjimoor, A., Favaro, P., Pirsiavash, H.: Boosting self-supervised learning via knowledge transfer. In: Proceedings of Computer Vision and Pattern Recognition, pp. 9359–9367 (2018). doi:10.1109/CVPR.2018.00975
16. Noroozi, M., Pirsiavash, H., Favaro, P.: Representation learning by learning to count. In: Proceedings of Internatinal Conference on Computer Vision, pp. 5898–5906 (2017). doi:10.1109/ICCV.2017.628
17. Utsumi, Y., Nakamura, K., Iwamura, M., Kise, K.: DNN-Based Tiller Number Estimation for Coping with Shortage of Labeled Data. In: Proceedins of Computer Vision Problems in Plant Phenotyping, p. 1 (2019)
18. Scotford, I.M., Miller, P.C.H.: Estimating tiller density and leaf area index of winter wheat using spectral reflectance and ultrasonic sensing techniques. Biosystems Engineering **89**(4), 395–408 (2004). doi:10.1016/j.biosystemseng.2004.08.019
19. Dampney, P., Quegan, S., Meadows, P.: Advanced radar for measuring green area index (GAI), biomass and shoot numbers in wheat (Radwheat). Technical Report 252, HGCA Project Report (2001)
20. Taylor, J.C., Wood, G.A., Welsh, J.P., Knight, S.: Exploring management strategies for precision farming of cereals assisted by remote sensing. Aspects of Applied Biology (60), 53–60 (2000)
21. Deng, R., Jiang, Y., Tao, M., Huang, X., Bangura, K., Liu, C., Lin, J., Qi, L.: Deep learning-based automatic detection of productive tillers in rice. Computers and Electronics in Agriculture **177**(April), 105703 (2020). doi:10.1016/j.compag.2020.105703
22. Wu, D., Wu, D., Feng, H., Duan, L., Dai, G., Liu, X., Wang, K., Yang, P., Chen, G., Gay, A.P., Doonan, J.H., Niu, Z., Xiong, L., Yang, W.: A deep learning-integrated micro-CT image analysis pipeline for quantifying rice lodging resistance-related traits. Plant Communications **2**(2) (2021). doi:10.1016/j.xplc.2021.100165
23. Minervini, M., Fischbach, A., Scharr, H., Tsaftaris, S.A.: Finely-grained annotated datasets for image-based plant phenotyping. Pattern Recognition Letters **81**, 80–89 (2016). doi:10.1016/j.patrec.2015.10.013
24. Aich, S., Stavness, I.: Leaf Counting with Deep Convolutional and Deconvolutional Networks. In: Proceedings of ICCV 2017 Workshop on Computer Vision Problems in Plant Phenotyping, pp. 2080–2089 (2017). doi:10.1109/ICCVW.2017.244
25. Ubbens, J., Cieslak, M., Prusinkiewicz, P., Stavness, I.: The use of plant models in deep learning: An application to leaf counting in rosette plants. Plant Methods **14**(1), 1–10 (2018). doi:10.1186/s13007-018-0273-z
26. Ward, D., Moghadam, P., Hudson, N.: Deep Leaf Segmentation Using Synthetic Data. In: In Proceedings of the British Machine Vision Conference (BMVC) Workshop on Computer Vision Problems in Plant Pheonotyping (CVPPP) (2018)
27. Han, T.H., Kuo, Y.F.: Developing a system for three-dimensional quantification of root traits of rice seedlings. Computers and Electronics in Agriculture **152**, 90–100 (2018). doi:10.1016/j.compag.2018.07.001
28. Wang, T., Rostamza, M., Song, Z., Wang, L., McNickle, G., Iyer-Pascuzzi, A.S., Qiu, Z., Jin, J.: SegRoot: A high throughput segmentation method for root image analysis. Computers and Electronics in Agriculture **162**, 845–854 (2019). doi:10.1016/j.compag.2019.05.017
29. Gaggion, N., Ariel, F., Daric, V., Lambert, É., Legendre, S., Roulé, T., Camoirano, A., Milone, D.H., Crespi, M., Blein, T., Ferrante, E.: ChronoRoot: High-throughput phenotyping by deep segmentation networks reveals novel temporal parameters of plant root system architecture. GigaScience **10**(7), 1–15 (2021). doi:10.1093/gigascience/giab052
30. Yasrab, R., Atkinson, J.A., Wells, D.M., French, A.P., Pridmore, T.P., Pound, M.P.: RootNav 2.0: Deep learning for automatic navigation of complex plant root architectures. GigaScience **8**(11) (2019). doi:10.1093/gigascience/giz123
31. Wu, J., Wu, Q., Pagès, L., Yuan, Y., Zhang, X., Du, M., Tian, X., Li, Z.: RhizoChamber-Monitor: A robotic platform and software enabling characterization of root growth. Plant Methods **14**(1), 1–15 (2018).

doi:10.1186/s13007-018-0316-5
32. Pound, M.P., Atkinson, J.A., Wells, D.M., Pridmore, T.P., French, A.P.: Deep learning for multi-task plant phenotyping. In: Proceedings of 2017 IEEE International Conference on Computer Vision Workshops, pp. 2055–2063 (2017). doi:10.1109/ICCVW.2017.241
33. Aich, S., Josuttes, A., Ovsyannikov, I., Strueby, K., Ahmed, I., Duddu, H.S., Pozniak, C., Shirtliffe, S., Stavness, I.: DeepWheat: Estimating Phenotypic Traits from Crop Images with Deep Learning. In: Proceedings of 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 323–332 (2018). doi:10.1109/WACV.2018.00042
34. Zhang, R., Isola, P., Efros, A.A.: Colorful image colorization. In: Proceedings of European Conference on Computer Vision, pp. 649–666 (2016). doi:10.1007/978-3-319-46487-9_40
35. Noroozi, M., Favaro, P.: Unsupervised learning of visual representations by solving jigsaw puzzles. In: Proceedings of European Conference on Computer Vision, pp. 69–84 (2016). doi:10.1007/978-3-319-46466-4_5