Automatic Detection of Plant Rows for a Transplanter in Paddy Field Using Faster R-CNN

SHAOMIN LIN1, YU JIANG2, XUESHEN CHEN1, ASIM BISWAS3, SHUAI LI1, ZIHAO YUAN1, HAILIN WANG4, AND LONG QI5

1Department of Agricultural Engineering, College of Engineering, South China Agricultural University, Guangzhou 510642, China
2Modern Educational Technology Center, South China Agricultural University, Guangzhou 510642, China
3School of Environmental Sciences, University of Guelph, Guelph, ON N1G 2W1, Canada
4College of Electronic Engineering, South China Agricultural University, Guangzhou 510642, China

Corresponding authors: Long Qi (qilong@scau.edu.cn), Hailin Wang (hlwang@scau.edu.cn), Xueshen Chen (chenxs@scau.edu.cn), and Asim Biswas (biswas@uoguelph.ca)

This work was supported in part by the Key Research and Development Program of Guangdong under Grant 2019B020221002; in part by the National Key Research and Development Program of China under Grant 2018 YFD0200303, in part by the Natural Science Foundation of China under Grant 51875217, in part by the Science Foundation of Guangdong for Distinguished Young Scholars under Grant 2019B151502056, in part by the Earmarked Fund for Modern Agro-Industry Technology Research System under Grant CARS-01-43, in part by the National Science Foundation for Young Scientists of China under Grant 31801258, and in part by the Science and Technology Research Program of Guangzhou under Grant 201803020021.

ABSTRACT Uniform plant row spacing in a paddy field is a critical requirement for rice seedling transplanting, as it affects subsequent field management and the crop yield. However, current transplanters are not able to meet this requirement due to the lack of accurate navigation systems. In this study, a plant row detection algorithm was developed to serve as a navigation system of a rice transplanter. The algorithm was based on the convolutional neural network (CNN) to identify and locate rice seedlings from field images. The agglomerative hierarchical clustering (AHC) was used to group rice seedlings into seedling rows which were then used to determine the navigation parameters. The accuracies of the navigation parameters were evaluated using test images. Results showed that the CNN-based algorithm successfully detected rice seedlings from field images and generated a reference line which was used to determine navigation parameters (lateral distance and travel angle). Compared with mean absolute errors (MAE) test results, the CNN-based algorithm resulted in a deviation of 8.5 mm for the lateral distance and 0.50° for the travel angle, over the six intra-row seedling spacings tested. Relative to the test results, the CNN-based algorithm had 62% lower error for the lateral distance and 57% lower error for the travel angle when compared to a classical algorithm. These results demonstrated that the proposed algorithm had reasonably good accuracy and can be used for the rice transplanter navigation in real-time.

INDEX TERMS Agglomerative hierarchical clustering, convolutional neural network, image, navigation, rice, seedling, transplanting.

I. INTRODUCTION

Rice is the staple food for more than half of the global population. Transplanting rice seedlings is one of the most popular methods of rice production. A critical requirement for transplanting is to have straight plant rows and uniform row spacings in rice fields. Uniform seedling row spacing is favorable to increased rice yields and minimize plant damage in the subsequent field operations, such as weeding, fertilization, spraying, and harvest. Compared to manually driving transplanters, automatic transplanters have the potential to achieve more uniform plant row spacings. However, it is challenging to navigate a transplanter to autonomously maintain a desired plant row spacing. This study addressed this challenge by developing an algorithm for automatic detection of rice seedlings for real-time navigation.

The navigation technology based on GPS or computer vision has been primarily used for agricultural automatic vehicles. Two-dimensional LiDAR was used to detect corn plant rows [1], [2]. The main advantage was the short distance target location but the shape of the plants was not considered. Moreover, this method may not be suitable for
rice, as the seedlings are much smaller (only about 30 cm in height at the time of transplanting) with weak morphological features compared to corn plants. Although the predefined route planning of GPS-based vehicles could be a solution for automatic transplanting [3]–[5], relatively high cost for equipment, operation, and maintenance limited the use of this technology in small scale agriculture. Alternatively, a real-time control based on image analysis and predefined algorithms requires less sophisticated equipment, and provides a cost-effective solution for transplanter navigation [6], [7].

The aim of this study was to develop an image-based algorithm to detect rice plant rows in real-time and to supply a transplanter with precise adjustable navigation parameters. An overview of the vision-based crop row detection methods and the reasons for developing a new algorithm are introduced in the section below. Then, the new algorithm about how to detect crop rows and how to determine navigation parameters are presented. Next, the accuracies of the navigation parameters are assessed with test results, and finally, the conclusions are drawn.

II. RELATED WORK
With the advancements of technological and computational powers, computer vision, also referred as image-based plant or plant row detection, has received a lot of attention in developing navigation systems for various cultivation practices. Image-based plant detection algorithms are based on several methods, including image pixel value distribution, stereo vision, and machine learning. In the image pixel value distribution method, an excess green index was proposed to enhance green features of plant canopy and reduce the color noise of soil in the images [8]. Plants were located at the higher summation of the pixel values of banding image. Multi-segments of image pixel values were proposed to separate green plants from soil and similar colored weeds [9].

To test the effects of illumination intensity of field environment, a plant row detection procedure was developed based on the H component of field images in an HSI color space model [10]. The accuracy of the plant identification and row detection is not accurate, as the distribution of image pixel value is irregular due to the morphological characteristic of plants. This often limits the applications of the pixel value distribution method.

The stereo vision method involves extractions of three-dimensional information from digital images. In this method, feature matching is achieved using different algorithms. An algorithm based on Scale-Invariant Feature Transform was capable of matching with the multi-thread computing technology, and a normalized distance method was used to eliminate inaccurate feature points in detecting plant rows for tractor navigations [11]. Another feature matching algorithm was based on Speeded-up Robust Features, involving rotation invariant to match feature points and to remove mismatching points [12]. A feature matching algorithm based on Census transformation was combined with principal component analyses to detect plant rows [13]. While the stereo vision method shows promise, its accuracy is limited due to the mismatching of feature points, particularly for working in bumpy paddy fields and for rice seedlings with weak features.

In contrast, the machine learning-based technique showed a stronger promise for diverse situations. For example, in a complex field environment, field images were segmented using image processing and vector machines to study excess green indices [14]. A method based on vector machines could identify plants with green spectral components masked and unmasked [15]. An Expert System has been developed for greenness identification based on the Fuzzy Clustering with four color indices [16]. With the advancement of deep learning technologies, which have good performance in objective location and detection under various environmental conditions, convolutional neural networks (CNN) have been used in agricultural applications, especially in fruit recognition [17], [18]. Some researchers adopted the deep learning method in weed and crop recognitions. For example, a CNN-based semantic segmentation of crop fields was proposed to separate sugar beet plants, weeds, and background solely based on RGB data [19]. In order to address various problems in field area, a CNN-based algorithm was proposed to detect the rice disease [20]. A network that is based on machine learning was also proposed to detect weeds in row crops from an unmanned aerial vehicle [21].

In general, the learning-based technique has shown excellent performance in identifying green plants from field environmental images. Regardless of detection methods, the random nature of plant posture in images influences the accuracy of the location of the plant detected. This, in turn, affects the accuracy of navigation.

The goal of this study was to provide navigation parameters to rice transplanters to achieve uniform plant distributions. The specific objectives were to a) develop a learning-based algorithm based on a pre-trained CNN model to detect rice seedlings during transplanting, b) determine the reference seedling row and navigation parameters, and c) evaluate the accuracy of the navigation parameters using test results.

III. ALGORITHM
The algorithm developed had several components from image acquisition to navigation parameter determination. The overall procedure is illustrated in Fig. 1, and it included the following steps: A) image acquisition from a rice field; B) detecting and locating rice seedlings on the images with a pre-trained convolutional neural network (CNN); C) clustering detected rice seedlings into rows using the agglomerative hierarchical clustering (AHC) method; D) extracting the cluster of the reference row; E) fitting the centerline of the reference row to obtain the reference line (RL); and F) deriving the navigation parameters relative to the RL. These steps are further described in the following sections.

A. IMAGE ACQUISITION
Rice seedling images were obtained from a paddy field in Jiangmen, Guangdong province, China (22°30′26.97″ N,
113°05'45.54" E) during March and April 2018. The soil in paddy field has been mashed by a mechanical masher and smoothed by a soil mechanical grader. Commonly, these processes would be taken to support rice seedlings with a homogeneous soil texture and a suitable growing paddy field. The image acquisition system included a color camera (Logitech® webcam, C920) with a resolution of 640 × 480 pixels and a computer (Core i7 2.9 GHz, 8 GB RAM, Nvidia Quadro M1200) (Fig. 2a). The operating system of the computer was Ubuntu 16.04. The camera was mounted on the side of a transplanter (Model: SPV-6CMD, Kubota, Japan, the travelling speed was 0~5.94km/h) at 1.30 m above the ground facing forward and downward at an angle of 10° relative to the vertical line. Unlike traditional agricultural equipment, the camera was mounted at the side of the transplanter to capture the images on-the-go. This is because the transplanted rice seedlings were used as the reference for the operation and were located at the side. The length of camera field of vision (FOV) on this vertical line was 1.05 m. Thus, the processing time of the real-time navigation system was <0.64 seconds per picture. Thus, the camera FOV covered every rice seedling. Fig. 2a shows a transplanted area and a non-transplanted area. This information was used in developing the plant detection algorithm. A seedling image captured the paddy field with half of the field area transplanted and half non-transplanted (Fig. 2b). The seedling row that separated the transplanted area from non-transplanted area was defined as the reference row and was used to extract navigation parameters.

Images were collected at different times of the day, including morning, noon, and dusk, to capture images under various illumination and lighting conditions. After the initial sample screening, a total of 2,904 images containing 43,721 seedlings were obtained (Table 1).

**TABLE 1.** Units for magnetic properties.

| Lighting condition | Image | Seedling |
|--------------------|-------|----------|
| Morning            | 981   | 13,615   |
| Noon               | 1,053 | 15,897   |
| Dusk               | 870   | 14,209   |
| Total              | 2,094 | 43,721   |

Faster R-CNN [22] consisted of a region proposal network (RPN) and a detection network. To achieve optimum performance, three shared CNN networks were tested including ZF Net [23], VGG_CNN_M_1024 Net, and VGG_16 Net [24]. They were separately embedded in Faster R-CNN and pre-trained with the field images using two training methods: Alternating Training method and Approximate Joint Training method. Three shared CNN networks and the two training methods gave six combinations which were the six pre-trained models tested in this study. Then, the best combination was selected for the algorithm.

2) DETECTION OF RICE SEEDLINGS
The trained Faster R-CNN model was used to detect seedlings and their locations. Fig. 3 shows the flow of the model. Briefly, an original image from the paddy field was input into a shared CNN, which provided a series of convolutional
feature maps. A region proposal network (RPN) was developed by sliding a $3 \times 3$ spatial window with 9 anchors (3 scales and 3 aspect ratios) over each convolutional feature map. The classification scores and boundary for each anchor box were simultaneously predicted for a total of 300 region proposals of rice seedlings with non-maximum suppression (NMS). Then, a detection network was developed and the input region proposals on the convolutional feature maps were classified. Each region proposals gave an output with a score and a location coordinate. Those region proposals were the recognized rice seedlings. A threshold score of 0.8 was used to identify those seedlings. Finally, the detected seedlings were marked on the output image.

3) CLUSTERING RICE SEEDLINGS INTO ROWS
The seedling proposals on the output image marked with a rectangular bounding-box were identified with coordinates $t = (tx_{\text{min}}, ty_{\text{min}}, tx_{\text{max}}, ty_{\text{max}})$. Each bounding-box represented a seedling. The center of the bounding-box $(\bar{x}_t, \bar{y}_t)$ was assumed to be the center of the seedling. This assumption caused some errors, due to the random nature of seedling postures in fields. The center of the seedling was calculated by the following equations:

$$\bar{x}_t = tx_{\text{min}} + \frac{1}{2} \times (tx_{\text{max}} - tx_{\text{min}})$$

$$\bar{y}_t = ty_{\text{min}} + \frac{1}{2} \times (ty_{\text{max}} - ty_{\text{min}})$$

Then, detected seedlings in the image were grouped into seedling rows using the agglomerative hierarchical clustering (AHC) [25]. This method was based on the distance of the center coordinates of the bounding-box, partitioning $n$ observations (seedling proposals) into $k$ clusters (each seedling row was one cluster) in such a way that each observation belongs to a cluster with the nearest distance to the cluster. As each image contained three seedling rows, the terminating condition for categorization of clusters was set as $k = 3$.

4) EXTRACTING THE REFERENCE ROW
The reference row was the seedling row that separated the transplanted and non-transplanted area. Fig. 2 shows that the transplanted area is on the right side of the transplanter, while the transplanted area would be on the left side for the next transplanting path. As a result, the reference row would switch between the left and right sides of the field images as well. This left and right switching needed to be considered in developing the plant row classification. In order for the transplanter to identify the current side of the reference row, the average value of the x-coordinate of each cluster was calculated using Eq. (3), and the results were arranged in order of size as $(\bar{X}_{\text{cluster1}}, \bar{X}_{\text{cluster2}}, \bar{X}_{\text{cluster3}})$.

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

where $x_i$ was the x-coordinate of the cluster and $\bar{X}$ was the average. When the $\bar{X}$ was at the left side of the image, the third cluster with the maximum x-coordinate value was the cluster of the reference row. In contrast, when the $\bar{X}$ was at the right side of image, the first cluster with the minimum x-coordinate value was the cluster of the reference row.
5) DETERMINING THE REFERENCE LINE (RL)
Reference line (RL) was the center line of the reference seedling row. It was obtained from the linear least-squares regression of the center points of the seedlings in the reference row, i.e., the RL was the best-fitted straight line of the centers of the seedlings. The RL was expressed by the regression equation: \( y = bx + a \), where \( b \) was the slope and \( a \) was the intercept. The offsets of a point \( d \) from a straight line was obtained as:

\[
 d = \frac{|y - (a + bx)|}{\sqrt{1 + b^2}} \quad (4)
\]

The expression to the sum of offsets \( E \) was given by:

\[
 E = \sum_{i=1}^{n} d_i = \sum_{i=1}^{n} \frac{|y_i - (a + bx_i)|}{\sqrt{1 + b^2}} \quad (5)
\]

where \( n \) was the total number of seedlings in the reference row. The RL was obtained by minimizing \( E \) in Eq. (5).

6) DETERMINING THE NAVIGATION PARAMETERS
Navigation parameters were determined based on the position of transplanter relative to the RL. In the field, the current locations of both transplanter and RL were captured in an image taken by the camera. The location of transplanter was described by its coordinate and travel direction (angle). The camera was located on the transplanter in a way that the midpoint on the boundary of the image represented the current location of the transplanter, and the vertical line indicates the current travel direction of the transplanter, as shown in Fig. 5. The current starting position of transplanter had a lateral distance of \( \Delta x \) and a travel angle of \( \theta \), relative to the RL, which were used as the navigation parameters. With the previously obtained regression equation of the RL, these two navigation parameters were calculated using Eq. (6) and (7):

\[
 \Delta x = -\frac{a}{b} - \frac{1}{2b}x_i \quad (6)
\]

\[
 \theta = \arctan \left( \frac{1}{b} \right) \quad (7)
\]

C. EVALUATIONS OF THE ALGORITHM
The source of errors of the algorithm was identified in the process of determining the RL. As mentioned, the center of seedling detected was assumed to be the center of the bounding-box. As illustrated in Fig. 6a, the center of the bounding-box may not represent the true center of the seedling, due to the fact that individual seedlings have random posture in field conditions. Therefore, the true center of seedling, defined as the center of the seedling stem (Fig. 6b), was calibrated. The images with calibrated seedling centers were used for testing the accuracy of seedling detection of the algorithm.

1) COLLECTION OF TESTING IMAGES
Testing images were collected from a paddy field. The field was transplanted at six intra-row seedling spacings: 100, 120, 140, 160, 180, and 210 mm. Images were taken using the same image acquisition system under different intra-row spacings, illuminations and lighting conditions. First, an image was taken at the field location as before, containing three rows of seedlings. Then, the seedlings in the
FIGURE 6. Field images for algorithm evaluation: (a) a seedling detected showing the source of error regarding the center of seedling, (b) defining the seedling center using the seedling stem, (c) a black pole image, and (d) binary image of black poles.

TABLE 2. Set of testing images.

| Lighting condition | Seedling image | Auxiliary image |
|--------------------|----------------|-----------------|
| Morning            | 42             | 42              |
| Noon               | 56             | 56              |
| Dusk               | 42             | 42              |
| Total              | 120            | 120             |

reference row were replaced by black poles, and an image was taken again at the same location using the same camera angle (the transplanter and camera locations were not changed). The information extracted from the black pole image (Fig. 6c) was used to compare the information extracted from the seedling image to evaluate the accuracy of the algorithm. In total, 20 images of seedlings and 20 images of black poles were taken for each intra-row spacing, giving a total of 240 images for six intra-row seedling spacings (Table 2).

2) ANALYSES OF THE BLACK POLE IMAGE
The images with black poles were analyzed to extract the theoretical navigation parameters which were used to evaluate the CNN-based model results of navigation parameters. Analyses were performed in OpenCV on Visual Studio 2010 to extract the centerline of black poles on the images, and the steps are summarized as follows:
1) Segmenting the original image to differentiate black poles and the background;
2) Finding contours from the binary image (Fig. 6d);
3) Extracting the black pole contours with an area filter;
4) Calculating the center coordinates of the black pole contours;
5) Fitting the straight line with the center coordinates of black poles;
6) Calculating the navigation parameters.

The navigation parameters extracted from black poles were considered as true values. Against these values, the navigation parameters from the CNN-based algorithm were evaluated for its accuracy.

3) COMPARISON OF THE CNN-BASED ALGORITHM WITH A CLASSICAL ALGORITHM
The navigation parameters from CNN-based algorithm were further compared with an existing classical algorithm [26]–[28]. The classical algorithm (ExG-Otsu) was developed based on color indices 2G-R-B (ExG) to differentiate background and green plants in images. It included the following main steps: a) extracting the reference area from original image as the region of interest (ROI) image; b) separating the rice seedlings and background in the ROI image with the ExG method; c) segmenting the ROI image into binary image with the Otsu threshold number; d) cutting the binary ROI image into 20 strip images along Y-axis; e) projecting the white pixels (rice seedlings) along X-axis and Y-axis in each strip image; f) recording the center location as (x, y) which is the maximum projecting coordinate value of each axis projection; g) mapping all the center locations from the strip images into the original image; h) fitting the center line in those locations with the least square method; and finally, i) extracting the navigation parameters.

IV. RESULTS

A. SELECTED TRAINING MODEL, BATCH SIZE AND LEARNING RATE
Among the six combinations of three shared CNNs and two training methods, the mean average precision and the detection rate of these pre-trained models were compared. The Approximate Joint Training method showed better performance than the Alternate Training method, regardless of the shared CNNs (Table 3). The models, VGG_CNN_M_1024 Net and VGG_16 Net trained in Approximate Joint Training method had higher precisions and detection rates than the other combinations. The former model was faster (6.4 frames per second) in rice seedling detection and responded to the real-time characteristic of the navigation system, whereas the latter model had a better precision of 88.7%. Considering the overall performance, the model with VGG_CNN_M_1024 Net and Approximate Joint Training method was used for further training tests.

The batch size of each training iteration and the learning rate affected the model performance. Their effects on the
TABLE 3. Mean average precision and rate of detection in pre-training of network.

| Shared CNN | Training method            | Precision (%) | Rate (fps) |
|------------|---------------------------|---------------|------------|
| ZF         | Approximate Training      | 79.3          | 5.8        |
| ZF         | Alternating Training      | 75.0          | 8.3        |
| VGG_CNN_M_1024 | Approximate Training   | 86.9          | 6.4        |
| VGG_CNN_M_1024 | Alternating Training    | 76.3          | 7.9        |
| VGG_16     | Approximate Training      | 88.7          | 2.6        |
| VGG_16     | Alternating Training      | 72.1          | 2.2        |

TABLE 4. Mean average precision of trained models.

| Batch size | Learning rate | Precision (%) |
|------------|---------------|---------------|
| 32         | 0.0005        | 78.8          |
| 32         | 0.001         | 87.2          |
| 32         | 0.005         | 79.0          |
| 64         | 0.0005        | 79.0          |
| 64         | 0.001         | 88.1          |
| 64         | 0.005         | 88.9          |
| 128        | 0.0005        | 79.0          |
| 128        | 0.001         | 87.9          |
| 128        | 0.005         | 89.6          |
| 256        | 0.0005        | 79.3          |
| 256        | 0.001         | 88.2          |
| 256        | 0.005         | 89.8          |

Mean average precision were examined under the batch sizes of 32, 64, 128, and 256 and learning rates of 0.0005, 0.001, and 0.005. Results showed that the model had higher precisions when the batch size and learning rate were increased (Table 4). However, a larger batch size required a higher computing capacity. Using the high learning rates sometimes caused computing convergence problems. The best parameter combination that gave the highest precision (89.8%) without causing convergence problems was the batch size of 256 and learning rate of 0.005. Thus, the model trained with this parameter combination was further used to detect seedlings and their locations.

B. PERFORMANCE OF THE ALGORITHM

1) NAVIGATION PARAMETERS

The CNN-based algorithm was able to achieve all functions desired. Seedling images were detected, located, and grouped into seedling rows (Fig. 7a). The reference row was successfully identified, and the RL was generated through regression. For the corresponding image with black poles, the true center line of the seedlings in the reference row is shown in Fig. 7b. Compared with the true center line, deviations from the RL could be seen when using the CNN-based algorithm. The errors on the navigation parameters are described in the following sections.

For each of the three detection methods: black pole image analyses, the CNN-based algorithm, and classical ExG-Otsu algorithm, the RL of the reference row was determined; then, navigation parameters: lateral distance ($\Delta x$) and travel angle ($\theta$) were calculated for each of the six intra-row seedling spacings. Results showed that the value of $\Delta x$ did not vary with the intra-row spacing of seedling regardless of the detection method (Fig. 8a). The value of $\Delta x$ depended on the initial positions of the transplanter. In this case, the average value of $\Delta x$ was 328.2 pixels (715.3 mm) over all methods and intra-row spacings. Results of $\Delta x$ were quite stable as indicated by their low standard deviations. As for $\theta$ both positive and negative values were obtained. Fig. 8b presents results of $\theta$ in absolute values varying from 2.04 to 4.62°. This navigation parameter was highly variable, as indicated by the magnitudes and standard deviations. Similar to the results of $\Delta x$, the intra-row seedling spacing did not affect the $\theta$, regardless of the detection method.
2) ACCURACY OF THE CNN-BASED ALGORITHM

The CNN-based algorithm was evaluated in terms of its accuracy in determining the navigation parameters. The accuracy was assessed against the results from the black pole image analyses, using mean absolute errors (MAE). For comparison purposes, MAE of the classical (ExG + Otsu) algorithm relative to the black pole results were also determined. Errors from the CNN-based algorithm were compared with those from ExG + Otsu algorithm. Results showed that the errors of the CNN-based algorithm in determining the \( \Delta x \) were quite low (Fig. 9a). The errors ranged from 2.92 to 5.59 pixels (6.38 and 12.4 mm) with an average of 3.90 pixels (8.5 mm). When compared to the ExG + Otsu algorithm, the CNN-based algorithm showed lower errors in determining the \( \Delta x \) at all intra-row seedling spacings. In determining the \( \theta \), the CNN-based algorithm also had lower errors (Fig. 9b). The errors were quite stable over different intra-row seedling spacings, and the average error was only 0.50°.

The results show that the accuracy of the CNN-based algorithm proposed in this study was higher compared to the ExG + Otsu algorithm. The overall error of the CNN-based algorithm was 62% lower for the lateral distance and 57% lower for the travel angle.

Each group was composed of images captured under various illumination. The adaptability of the CNN-based algorithm and the classical algorithm to illumination change was assessed, using root mean square errors (RMSE). Results showed that the RMSE of the CNN-based algorithm in determining the \( \Delta x \) ranged from 1.51 to 3.31 pixels (3.30 and 7.24 mm) with an average of 2.69 pixels (5.88 mm) (Fig. 10a).

When compared to the ExG + Otsu algorithm, the CNN-based algorithm showed stable results in determining the \( \Delta x \) under various illumination and lighting condition. In determining the \( \theta \), the RMSE of the CNN-based algorithm ranged from 0.25° to 0.36° with an average of 0.32° (Fig. 10b). The RMSE were more stable than the results of the ExG + Otsu algorithm at different illuminations.

The ExG + Otsu algorithm worked well when the crops had strong green color indices. However, the rice seedlings’ color feature was weak and their reflection on water also contributed to similar colors. This definitely challenged the identification of rice seedlings accurately. In that case, the precision in detecting the location of rice seedlings was negatively affected. In contrast, the algorithm that was based on Faster R-CNN located rice seedlings by their morphology features, and therefore, the effects of soil and water were minimized in seedling detection.

The proposed method performed well for rice fields under various illumination conditions. The detection speed of the algorithm was approximately 5 fps, and it was fast enough to respond to the control of the navigation of transplanter. However, the proposed method identified the rice seedlings with the canopies of seedlings in the image. The canopy features were highly variable, depending on postures of plants, especially long leave crops, like rice seedlings. This is primarily caused deviations of the seedling locations. Therefore, the effects of rice seedling leaf shape and postures need to be studied in future researches.
FIGURE 10. Root mean square error (RMSE) values of two detection algorithms for different intra-row seeding spacings: (a) lateral RMSE, and (b) angular RMSE.

V. CONCLUSION

This study proposed a new algorithm to provide navigation parameters for transplanting rice seedlings in paddy fields. The trained convolutional neural network (CNN) showed the best performance when using the combination of VGG_CNN_M_1024 Net and Approximate Joint Training method, when compared with six other combinations. The highest model precision was 89.8%, occurred at the batch size of 256 and learning rate of 0.005. The algorithm performed well in extracting the reference lines of reference seedling rows, regardless of the intra-row seedling spacings. The navigation parameters determined from the reference lines had lower errors (MAE): 8.5 mm on average for lateral deviation and 0.32° on average for angular deviation. These were 62% and 57% lower when compared with the corresponding errors from a classical algorithm developed based on color indices. The adaptability to illumination change assessed from each group of testing data was stable (RMSE): 5.88 mm on average for lateral deviation and 0.32° on average for angular deviation. These were 69% and 60% lower when compared with the corresponding errors from a classical algorithm developed based on color indices. While the proposed machine learning algorithm showed strong promise in developing a real-time navigation system for rice transplanting, further testing needs to be carried out in different field conditions and rice varieties. Effects of leaf orientations and postures need further research as well. Furthermore, although the rice seedling row detection algorithm developed in high accuracy, the navigation system for transplanter required a strategy-based control system to stably navigate the transplanter according to navigation parameters.

REFERENCES

[1] D. Reiser, M. Vázquez-Arellano, D. S. Parafos, M. Garrido-Izard, and H. W. Griepentrog, “Iterative individual plant clustering in maize with assembled 2D LiDAR data,” Comput. Ind., vol. 99, pp. 42–52, Aug. 2018.
[2] U. Weiss and P. Biber, “Plant detection and mapping for agricultural robots using a 3D LiDAR sensor,” Robot. Auto. Syst., vol. 59, no. 5, pp. 265–273, May 2011.
[3] H. W. Griepentrog, M. Nørremark, H. Nielsen, and B. S. Blackmore, “Seed mapping of sugar beet,” Precis. Agricult., vol. 6, no. 2, pp. 157–165, Apr. 2005.
[4] H. Moussazadeh, “A technical review on navigation systems of agricultural autonomous off-road vehicles,” J. Terramechanics, vol. 50, no. 3, pp. 211–232, Jun. 2013.
[5] M. Pérez-Ruiz, D. C. Slaughter, C. J. Gliever, and S. K. Upadhyaya, “Automatic GPS-based intra-row weed knife control system for transplanted row crops,” Comput. Electron. Agric., vol. 80, pp. 41–49, Jan. 2012.
[6] C. Kunz, J. F. Weber, G. G. Peteinatos, M. Sökofeld, and R. Gerhards, “Camera steered mechanical weed control in sugar beet, maize and soybean,” Precis. Agricult., vol. 19, no. 4, pp. 708–720, Aug. 2018.
[7] K. H. Choi, S. K. Han, S. H. Han, K.-H. Park, K.-S. Kim, and S. Kim, “Morphology-based guidance line extraction for an autonomous weeding robot in paddy fields,” Comput. Electron. Agric., vol. 113, pp. 266–274, Apr. 2015.
[8] M. Kejun, Z. Ping, Z. Yun, Y. Qingxing, and Z. Qufa, “The study for the navigation information about automatic navigation of unmanned rice mechanical transplanter based on machine vision,” J. Agricult. Mechanization Res., vol. 5, pp. 63–66, May 2005.
[9] M. Montalvo, G. Pajares, J. M. Guerrero, J. Romeo, M. Guijarro, and A. Ribeiro, “Automatic detection of crop rows in maize fields with high weeds pressure,” Expert Syst. Appl., vol. 39, no. 15, pp. 11889–11897, Nov. 2012.
[10] G. Guoqin and L. Ming, “Navigating path recognition for greenhouse mobile robot based on K-means algorithm,” Trans. Chin. Soc. Agricult. Eng., vol. 30, no. 07, pp. 451–456, Jan. 2014.
[11] T. Guangzhaoo, A. Qiu, J. Changying, G. Baoxing, W. Haiqing, and Z. Jiandong, “Real-time motion detection for intelligent agricultural vehicle based on stereo vision,” Trans. Chin. Soc. Agricult. Machinery, vol. 44, no. 07, pp. 210–215, Jul. 2013.
[12] Z. Zhbin, Z. Shuailing, L. Xiwen, and W. Fengqi, “Matching method of green crops based on SURF feature extraction,” Trans. Chin. Soc. Agricult. Eng., vol. 31, no. 14, pp. 172–178, Jul. 2015.
[13] Z. Zhiqiang, Z. Zhongxiang, D. Yuefeng, Z. Shuo, and M. Enrong, “Method for detecting crop rows based on binocular vision with Census transformation,” Trans. Chin. Soc. Agricult. Eng., vol. 32, no. 11, pp. 205–213, Jun. 2016.
[14] S. Qiaolin and A. Qiu, “Image segmentation of green plants in corn field based on neighborhood histogram,” Acta Agricult. Jiangxi, vol. 23, no. 5, pp. 126–128, May 2011.
[15] J. M. Guerrero, G. Pajares, M. Montalvo, J. Romeo, and M. Guijarro, “Support vector machines for crop/weeds identification in maize fields,” Expert Syst. Appl., vol. 39, no. 12, pp. 11149–11155, Sep. 2012.
[16] J. Romeo, G. Pajares, M. Montalvo, J. M. Guerrero, and M. Guijarro, “A new Expert System for greenness identification in agricultural images,” Expert Syst. Appl., vol. 40, no. 6, pp. 2275–2286, May 2013.
[17] M. Stein, S. Bargoti, and J. Underwood, “Image based mango fruit detection, localisation and yield estimation using multiple view geometry,” Sensors, vol. 16, no. 11, p. 1915, Nov. 2016.
[18] S. Bargoti and J. Underwood, “Deep fruit detection in orchards,” in Proc. IEEE Int. Conf. Robot. Automat. (ICRA), May-Jun. 2017, pp. 3626–3633.
[19] A. Milioni, P. Lottes, and C. Stachniss, “Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs,” in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), May 2018, pp. 2229–2235.
[20] G. Zhou, W. Zhang, A. Chen, M. He, and X. Ma, “Rapid detection of rice disease based on CNN-RM and faster R-CNN fusion,” IEEE Access, vol. 7, pp. 143190–143206, 2019, doi: 10.1109/ACCESS.2019.2943454.
[21] M. Bah, A. Hafiane, and R. Canals, “Deep learning with unsupervised data labeling for weed detection in line crops in UAV images,” Remote Sens., vol. 10, no. 11, p. 1690, Oct. 2018.

[22] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[23] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” Comput. Sci., to be published.

[24] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” Comput. Sci., to be published.

[25] S. C. Johnson, “Hierarchical clustering schemes,” Psychometrika, vol. 32, no. 3, pp. 241–254, Sep. 1967.

[26] T. Bakker, H. Wouters, K. V. Asselt, J. Bontsema, L. Tang, J. Müller, and G. V. Straten, “A vision based row detection system for sugar beet,” Comput. Electron. Agricult., vol. 60, no. 1, pp. 87–95, Jan. 2008.

[27] G. Jiang, X. Wang, Z. Wang, and H. Liu, “Wheat rows detection at the early growth stage based on Hough transform and vanishing point,” Comput. Electron. Agricult., vol. 123, pp. 211–223, Apr. 2016.

[28] I. Vidović, R. Cupec, and Ž. Hocenski, “Crop row detection by global energy minimization,” Pattern Recognit., vol. 55, pp. 68–86, Jul. 2016.

**SHAOMIN LIN** received the M.Sc. degree from South China Agricultural University, in 2019. His research interests include machine vision, deep learning, and their applications in agriculture.

**YU JIANG** is currently a Senior Engineer with the Modern Educational Technology Center, South China Agricultural University. Her research interest includes image processing and analysis.

**XUESHEN CHEN** is currently an Associate Professor with the College of Engineering, South China Agricultural University. His research interests include control theory and mechanical engineering.

**ASIM BISWAS** is currently an Associate Professor with the School of Environmental Sciences, University of Guelph. His research interests include soil sensor and digital soil mapping.

**SHUAI LI** received the M.Sc. degree from South China Agricultural University, in 2019. His research interest includes control theory and engineering.

**ZIHAO YUAN** is currently pursuing the M.Sc. degree with the College of Engineering, South China Agricultural University. His research interest includes image processing and analysis in agriculture.

**HAILIN WANG** is currently the Dean of the College of Engineering, South China Agricultural University. His research interests include automobile and mechanical dynamics, and vehicle engineering.

**LONG QI** received the M.Sc. and Ph.D. degrees from Jilin University, in 2006 and 2009, respectively. He is currently a Professor with the College of Engineering, South China Agricultural University. His research interests include image processing, machine vision, deep learning, simulation, and their applications in agriculture. He has published more than 50 articles and holds numerous patents in his research areas. He got the Youth Science and Technology Award of CSAE, and was founded by the Science Foundation of Guangdong for Distinguished Young Scholars, in 2019.