



Identifying Dike-Pond System Using an Improved Cascade R-CNN Model and High-Resolution Satellite Images

Yintao Ma ^{1,2,†}, Zheng Zhou ^{3,†}, Xiaoxiong She ^{1,2}, Longyu Zhou ^{1,2}, Tao Ren ^{1,2}, Shishi Liu ^{1,2,*} and Jianwei Lu ^{1,2}

- ¹ School of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China; mayt@webmail.hzau.edu.cn (Y.M.); shexiaoxiong@webmail.hzau.edu.cn (X.S.); zhoulongyu@webmail.hzau.edu.cn (L.Z.); rentao@mail.hzau.edu.cn (T.R.); lunm@mail.hzau.edu.cn (J.L.)
- ² Key Laboratory of Arable Land Conservation (Middle and Lower Reaches of Yangtze River), Ministry of Agriculture, Wuhan 430070, China
- ³ Ecological Environment Monitoring and Scientific Research Center, Yangtze River Basin Ecological Environment Supervision and Administration Bureau, Ministry of Ecological Environment, Wuhan 430014, China; zhouzheng@cjjg.mee.gov.cn
- * Correspondence: sslu@mail.hzau.edu.cn
- † Yintao Ma and Zheng Zhou contributed equally to this study.

Abstract: The dike-pond system (DPS) is the integration of a natural or man-made pond and crop cultivation on dikes, widely distributed in the Pearl River Delta and Jiangnan plain in China. It plays a key role in preserving biodiversity, enhancing the nutrient cycle, and increasing crop production. However, DPS is rarely mapped at a large scale with satellite data, due to the limitations in the training dataset and traditional classification methods. This study improved the deep learning algorithm Cascade Region Convolutional Neural Network (Cascade R-CNN) algorithm to detect the DPS in Qianjiang City using high-resolution satellite data. In the proposed mCascade R-CNN, the regular convolution layer in the backbone was modified into the deformable convolutional layer, which was more suitable for learning the features of DPS with variable shapes and orientations. The mCascade R-CNN yielded the most accurate detection of DPS, with an average precision (AP) value that was 2.71% higher than Cascade R-CNN and 11.84% higher than You Look Only Once-v4 (YOLOv4). The area of oilseed rape growing on the dikes accounted for 3.42% of the total oilseed rape planting area. This study demonstrates the potential of the deep learning methods combined with high-resolution satellite images in detecting integrated agriculture systems.

Keywords: dike-pond detection; high-resolution satellite; deep learning algorithm; Cascade R-CNN; YOLOv4



Citation: Ma, Y.; Zhou, Z.; She, X.; Zhou, L.; Ren, T.; Liu, S.; Lu, J. Identifying Dike-Pond System Using an Improved Cascade R-CNN Model and High-Resolution Satellite Images. *Remote Sens.* **2022**, *14*, 717. <https://doi.org/10.3390/rs14030717>

Academic Editor: Giuseppe Modica

Received: 3 December 2021

Accepted: 27 January 2022

Published: 3 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The dike-pond system (DPS) is the integration of agriculture and aquaculture. It is characterized by a natural or man-made pond and dikes on which crop, vegetables, or fruit trees are cultivated [1]. The DPS is the traditional agriculture system in the low-lying and watery areas in South Asia [2,3]. In China, DPSs are concentrated in the Pearl River Delta, Yangtze River Delta, and bank regions of great lakes [4]. The Huzhou Mulberry-dike and Fish-pond system in China was designated a globally important agriculture systems project (GIAHS) by the Food and Agriculture Organization of the United Nations (FAO) in 2017. The DPS plays a key role in preserving biodiversity [5], enhancing the nutrient cycle [6], and increasing crop production [2]. Accurately identifying the DPS and mapping their spatial distributions are significant for understanding the environmental impacts of the integrated agricultural systems.

Remote sensing provides a unique alternative for mapping the spatial distribution of DPSs at large scales. However, DPS is usually classified as aquaculture ponds or wetland, derived from optical remote sensing images or radar images [7–10]. Very few studies have

been devoted to identifying DPS from satellite images, because the small sizes and the complex compositions of crops and water bodies make them difficult to map with medium- or coarse-resolution remote sensing images and conventional classifiers. Li et al. [11] analyzed the trends of DPS between 1978 and 2016 in Shunde district of South China, using the time series of Landsat images and declassified intelligence satellite photographs from before the 1980s. Liu and Li [12] mapped DPS dynamics during 1949–2020 in the Guangdong-Hong Kong-Macao Greater Bay Area, using topographic maps from 1949, Landsat images, and high-resolution satellite images. Both studies focused on analyzing the spatial-temporal dynamics of DPS. The classification of DPS has mainly relied on medium-resolution satellite data and the object-oriented classification method, ignoring the relation between the dike and the pond.

Mapping DPS with remote sensing techniques requires identification of a pond with a regular or irregular shape and crops grown on the dikes as a whole target. This special landscape is difficult to automatically identify using conventional classifiers, such as random forest (RF) or support vector machine (SVM) algorithms. The conventional classifiers have limitations in detecting objects. Firstly, these classifiers need several moving windows with varying sizes to locate the target in the image, resulting in redundant windows and low efficiency. Secondly, these classification methods cannot effectively extract deep-level features and identify complex objects from remote sensing images. DPS is an integrated agricultural system, in which the pond and vegetation have contrast spectral characteristics but have a spatial connection. The conventional classification methods conducted on pixels or objects struggle to extract the features of the spatial relationship between water and vegetation and identify DPS as a target.

Recent advances in deep learning algorithms have provided great opportunities for automatically identifying targets on high-resolution remote sensing images [13]. Deep learning is a hierarchical feature learning method that uses multi-layer neural networks. Convolutional neural networks (CNNs) are one of the most successful network architectures in deep learning methods through end-to-end learning. CNNs have demonstrated competitive abilities in classifying agricultural landscapes from remote sensing images at the pixel or object level [14–17]. However, few relevant studies have evaluated CNN-based object detection methods in agricultural applications, because of the complex properties of agricultural targets and the lack of annotated datasets, such as ImageNet, to meet the requirements of deep learning methods. Li et al. (2020) [18] and Chen et al. (2021) [19] detected agricultural greenhouses from high-resolution satellite data using the You Look Only Once-v3 (YOLO-v3) and CNN, respectively.

Despite the limited number of studies, deep learning methods have great potential for directly recognizing complex agricultural landscapes as targets. In addition to extracting complicated features in DPS, the algorithm used to identify DPS also needs to deal with irregular shapes and different orientations. The objective of this study was to develop a new architecture based on the state-of-art cascade region-based convolutional neural network (R-CNN) to detect the DPS from high-resolution satellite images. The novelty of the proposed method is that it is adaptive to the irregular shapes of DPSs and can provide a more accurate bounding box in the DPS detection. Based on the derived DPS map, we analyzed the spatial distribution of DPSs and quantified the area of oilseed rape growing on dikes, which is usually overlooked in remote mapping of cropland or the statistical data. This study was conducted in Qianjiang City, Hubei Province, China, where DPSs are widely distributed. The DPS in Qianjiang is characterized by a combination of winter oilseed rape growing on dikes and an aquaculture pond.

2. Materials and Methods

2.1. Study Area

The study area was Qianjiang City, a sub-prefecture-level city in South-Central Hubei province, China that covers an area of 200,400 ha. Qianjiang is located on the Jiangnan Plain, and has abundant water resources, including rivers, lakes, and ponds (Figure 1). In

total, 6 lakes are scattered throughout the city, with a total area of 1800 ha. Qianjiang has a humid subtropical climate, with an annual (1988–2017) temperature of 16.6 °C and annual precipitation of 1162 mm [20].

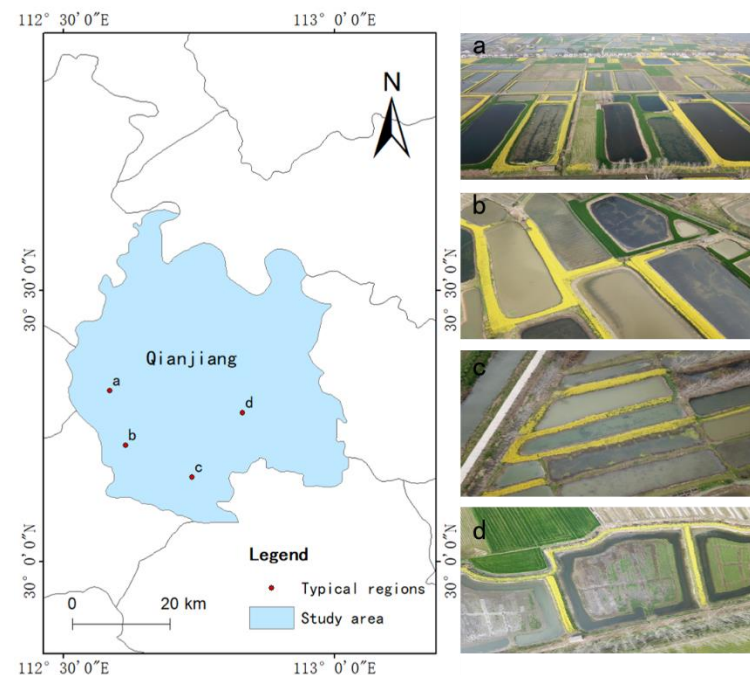


Figure 1. Location of the study area and photos of dike-pond systems (DPSs).

Aquaculture plays a very important role in the economy of Qianjiang. The total aquaculture area was 9195 ha in 2019. Oilseed rape (*Brassica napus* L.) is the main winter crop in Qianjiang. It is widely grown on the dikes of aquaculture ponds, and in spring, when oilseed rape blossoms, DPSs are easier to identify (Figure 1).

2.2. Data

In total, 5 high-resolution satellite images from Gaofen-1 (GF-1) and Gaofen-2 (GF-2) that covered the entire study area with cloud cover less than 10% were downloaded from China Centre For Resources Satellite Data and Application (<http://www.cresda.com/CN/>, accessed on 12 August 2021). GF-1 and GF-2 were launched by the China National Space Administration on 26 April 2013 and 19 August 2014, respectively. GF-1 carries 2 panchromatic (PAN) and multispectral (MS) cameras, with a spatial resolution of 2 and 8 m for the PAN and MS bands, respectively. GF-2 also employs 2 PAN and MS cameras, capable of collecting images with a spatial resolution of 0.81 and 3.24 m at nadir in the PAN and MS bands, respectively. Approximately 80% of the study area was covered by 1 scene from GF-1 obtained on 8 March 2020, and the rest of the area was covered by 4 scenes from GF-2 obtained on 27 March 2018 due to the limited data availability.

The selected GF-1 and GF-2 images were orthorectified and projected onto the Albers equal-area conic projection. The MS images were registered to the PAN images using polynomial warping with automatically generated tie points. The red, green, and blue (RGB) bands of the MS data were used with the corresponding PAN images using the nearest-neighbor diffusion-based pan-sharpening algorithm [21]. All the RGB composites were resampled to a spatial resolution of 2 m using the cubic convolution resampling method.

To train and validate the deep learning model, the regions of interest (ROIs) of 416×416 pixels containing the DPSs were cropped into tiles with 50% overlap. We labeled 1006 sample tiles, containing a total of 5903 targets. Eighty percent of the samples were used to train the deep learning models, and the remaining samples were used to validate the models.

2.3. Methods

Object detection methods can be categorized into two types: two-stage methods and one-stage methods. For two-stage methods, object detection is treated as a multi-task learning problem that combines classification and bounding box regression. On the one hand, two-stage methods typically require a heavy computational load. On the other hand, one-stage methods require only a single pass through the neural network and predict all the bounding boxes in one run. One-stage methods have recently become popular, mainly because of their computational efficiency. In this study, we improved the two-stage algorithm Cascade R-CNN to provide a more accurate detection of the DPS. The Cascade R-CNN based on feature pyramids network (FPN) and ResNet-101 backbone and popular one-stage algorithm YOLOv4 were applied for comparisons. After detecting DPSs in the study area, the bounding boxes were converted to vector data, and the number of bounding boxes represented the number of DPS.

2.3.1. Modified Cascade R-CNN

Cai and Vasconcelos proposed Cascade R-CNN, a multi-stage extension of the R-CNN [22]. Cascade R-CNN incorporates high-quality object detectors to improve the detection accuracy by beating the overfitting problem at training and quality mismatch at inference. A study showed that Cascade R-CNN, based on the ResNet-101 and FPN backbone, was observed to have outperformed several two-stage (e.g., Faster R-CNN) and one-stage detectors (e.g., YOLOv2) on the MS-COCO2017 dataset [22].

In this study, we used a Cascade R-CNN based on ResNeXt-101 and FPN backbone. In a remotely sensed image, DPSs are variable in shape and position. To improve the ability to learn deformable features, we modified ResNeXt-101 by replacing the regular convolutional layer with the deformable ConvNet v2 (DCNv2). DCNv2 is developed from DCNv1, which allows the grid sampling locations to swim with respect to the feature map when learning a spatial offset. However, DCNv1 suffers from the problem of irrelevant image content. DCNv2 is adaptive to an object's structure and is more powerful in focusing on pertinent image regions than DCNv1 [23]. ResNeXt-101+DCNv2 extracts the features of four different scales. The FPN recursively fuses features from higher levels to the current level.

The fused features are divided into four stages: one Region Proposal Network (RPN) and three detectors. The sampling of the first detection stage followed the procedures by Ren et al. [24]. In the following stages, resampling was implemented by simply using the regressed bounding boxes from the previous stage [22]. These 3 detectors were trained with an intersection over union (IoU) thresholds of 0.5, 0.6, and 0.7, respectively, to find a good set of close false positives for training the next stage. At each stage, the Cascade R-CNN included a classifier and a regressor optimized for the IoU threshold. The architecture of the modified Cascade R-CNN (mCascade R-CNN) is illustrated in Figure 2.

2.3.2. YOLOv4

In addition, the one-stage algorithm YOLOv4 was applied to detect DPSs for comparison, because it is one of the most popular target detection methods with high speed and accuracy. YOLOv4, an evolution of the YOLOv3, is a real-time object detection algorithm that recognizes different objects in a single frame. YOLOv4 generally includes three parts, namely the backbone, neck, and head networks. The backbone network is mainly used to extract image features, and the neck network can enhance the image features. The head network conducts classifications and regressions based on the features derived from the backbone and neck networks.

The image features were extracted using the CSPDarknet53 module in YOLOv4. CSPDarknet53 uses DenseNet and Cross Stage Partial connection (CSP) to enhance the learning ability of CNN and reduce model calculations and memory costs while maintaining accuracy. The RGB sample tiles with a size of $416 \times 416 \times 3$ were used as the input, and 3 outputs were generated after passing through CSPDarknet53. The sizes of the 3 feature outputs were $76 \times 76 \times 256$, $38 \times 38 \times 512$, and $19 \times 19 \times 1024$. The neck network in

YOLOv4 used Spatial Pyramid Pooling (SPP) and PANet to generate feature pyramids. SPP used 3 sliding kernels, namely 5×5 , 9×9 , and 13×13 , to convolve the candidate images, and then applied multi-scale max pooling to obtain the same dimensions of the feature map. PANet extracted and integrated features at various scales. The feature maps of different scales output by PANet were spliced, and after the convolution operation, three heads of the different scales were obtained. Classifications and regressions were applied to the three heads to predict the bounding box and the confidence level. The architecture of YOLOv4 used in this study is presented in Figure 3.

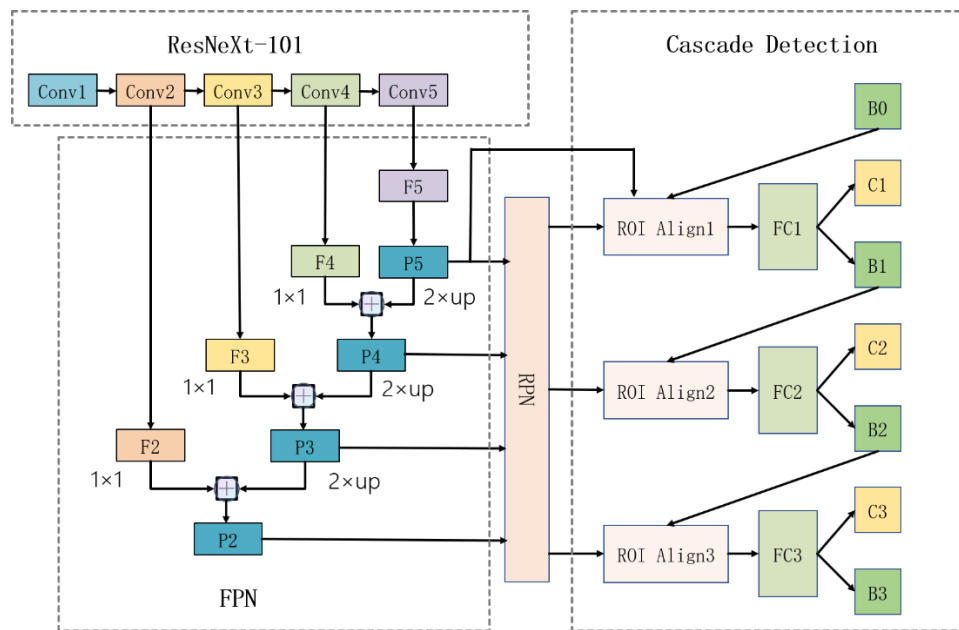


Figure 2. Architecture of the modified Cascade region-based convolutional neural network (mCascade R-CNN). “Conv” is backbone convolutions, “F” is feature image, “P” is feature image fused by feature pyramid network (FPN), “FC” is fully connected layer, “B” is bounding box, and “C” is classification.

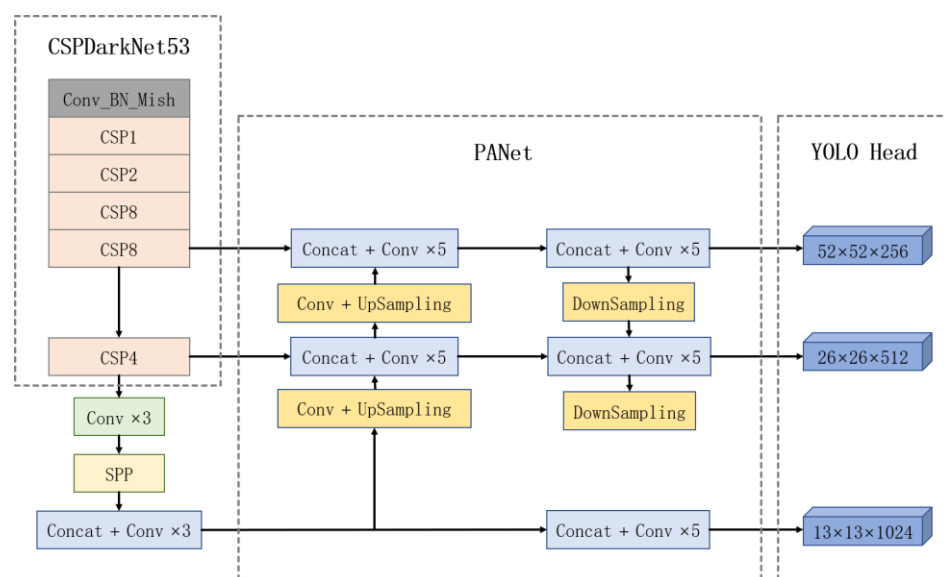


Figure 3. Architecture of the You Look Only Once v4 (YOLOv4) algorithm. “CSP” is the Cross Stage Partial connection, “Concat” means concatenate, and “Conv” is the convolution layer.

2.3.3. Evaluation of Model Performance

To evaluate the performance of mCascade R-CNN, Cascade R-CNN, and YOLOv4, we calculated the mean average precision based on the validation dataset. The mean average precision value is the area under the precision–recall curve of all classes. In this study, we only identified the DPS, and hence, the mean was not necessary. Average precision (AP) was calculated as follows:

$$AP = \int_0^1 P(R) dR \quad (1)$$

where P represents the precision rate and R represents the recall rate.

The precision rate is the proportion of predicted positives that are actually positive, and the recall rate is the proportion of observed positive samples that are correctly predicted as positive. Precision and recall are expressed as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

where TP is the number of real positive samples, FP is the number of false positive samples, and FN is the number of false negative samples.

2.4. Classification of Oilseed Rape

SVM was applied to classify pixels into oilseed rape and other land cover types. The SVM model was trained using a Gaussian radial basis function. Blue, green, red, and near-infrared bands were used as the inputs. Oilseed rape pixels were easily identified during the flowering stage. To train the model, 23,300 samples were used, and to validate the model, 9988 samples were used.

2.5. Kernel Density Estimation

Kernel density estimation (KDE) is a non-parametric estimation of probability density. It generates a smooth density probability surface, and provides a clear visualization of the spatial distribution of sample points (Brunsdon, 1995). The built-in kernel density tool in ArcGIS 10.5 was applied to calculate the density probability of the DPSs at a resolution and bandwidth of 100 m and 1 km, respectively.

3. Results

3.1. DPS Identification Accuracy

Figure 4 compares the accuracy of DPS detection using the 3 deep-learning algorithms. The AP value of the mCascade R-CNN is 2.71% higher than that of Cascade R-CNN and 11.84% higher than that of YOLOv4, indicating that mCascade R-CNN provides a more accurate identification of DPS, whereas YOLOv4 has a weaker detection effectiveness.

Figure 5 provides a subset of the DPS identification results to visually assess the detection performance of the 3 methods. Cascade R-CNN and mCascade R-CNN identified more DPSs than YOLOv4. When zooming in on the bounding boxes, we found that the bounding box of mCascade R-CNN had less overlap and fit to the DPS more than the Cascade R-CNN. A more accurate bounding box facilitates the extraction of the oilseed rape growing on the dikes. However, all three methods showed difficulties in identifying ponds full of aqua plants and ponds with narrow dikes, and ponds with oilseed rape growing on only one side. As mCascade R-CNN provided the most accurate identification of the DPS, the following analyses were conducted on the results generated by the mCascade R-CNN.

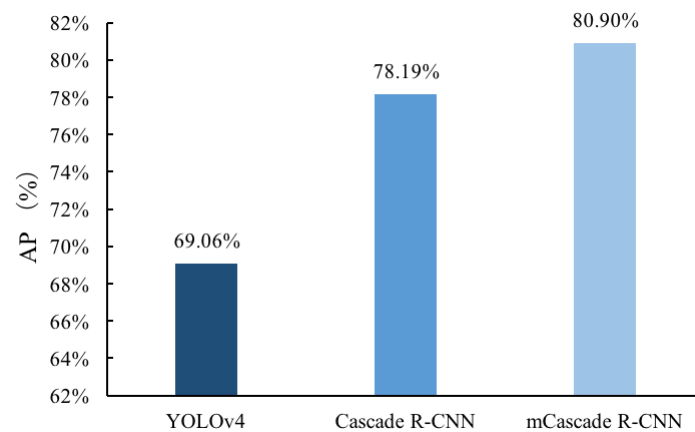


Figure 4. Comparison of the average precision (AP) in the results of You Look Only Once v4 (YOLOv4), Cascade region-based convolutional neural networks (Cascade R-CNN), and modified Cascade region-based convolutional neural networks (mCascade R-CNN).

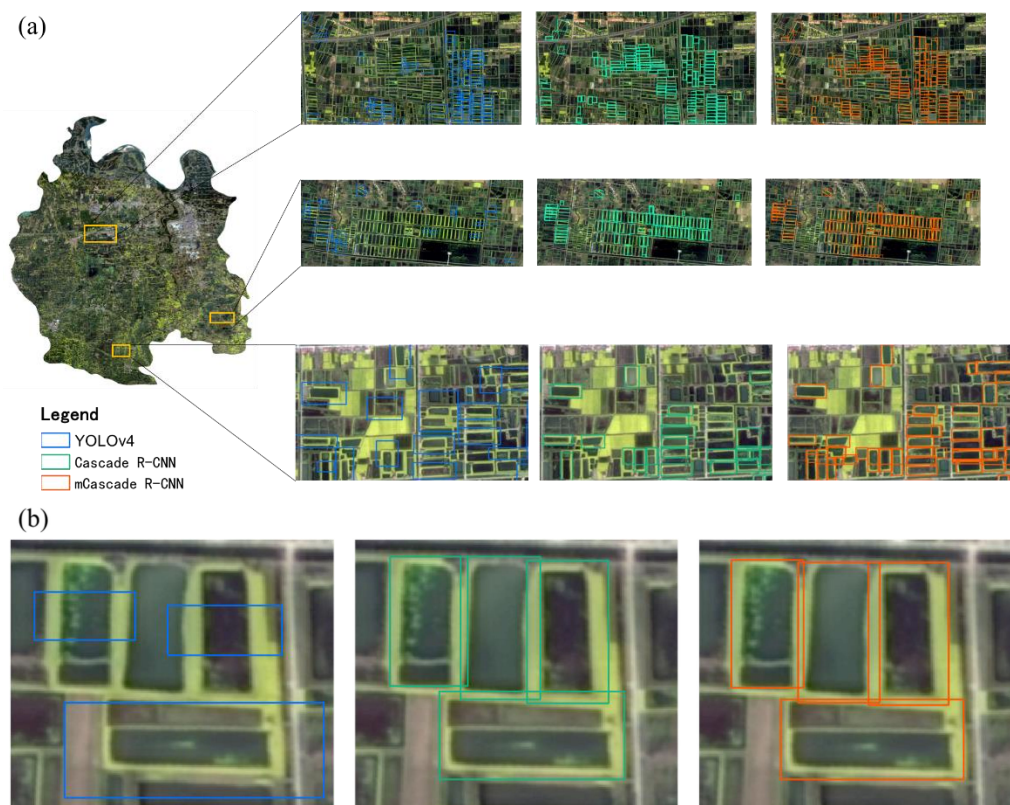


Figure 5. Dike-pond systems (DPSs) detected by You Look Only Once v4 (YOLOv4), Cascade region-based convolutional neural networks (Cascade R-CNN), and modified Cascade region-based convolutional neural networks (mCascade R-CNN). Image subset of the DPS detection results (a) and the zoom-in showing the accuracy of the bounding boxes (b).

3.2. Spatial Distribution of DPS

The mCascade R-CNN detected 2975 DPSs based on the number of bounding boxes in the study area. Figure 6 illustrates the KDE of the detected DPS. The KDE values range from 0 to 10, indicating an uneven spatial distribution of the DPSs. Hotspots are distributed in the middle and southwestern parts of Qianjiang City, distant from the urban area.

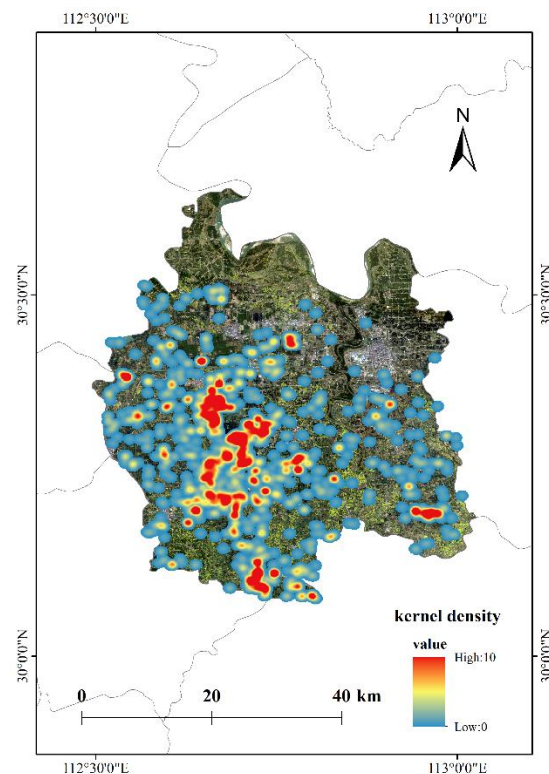


Figure 6. Kernel density estimation (KDE) of the dike-pond systems (DPSs) in Qianjiang City.

3.3. Oilseed Rape Planting Area in the DPS

Figure 7 demonstrates the winter oilseed rape map derived from the high-resolution satellite image using the SVM classification method. The overall classification accuracy was 99%. The total oilseed rape planting area was 14,433 ha in the study area. According to the results of mCascade R-CNN, the area of oilseed rape growing on the dikes of ponds was 493 ha, accounting for approximately 3.42% of the total oilseed rape area.

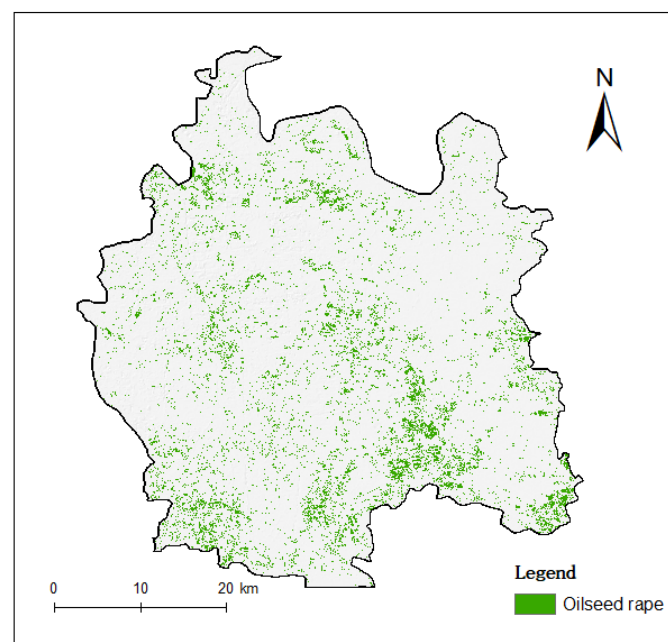


Figure 7. Oilseed rape map derived from the high-resolution satellite image using the Support Vector Machine (SVM) classification method.

4. Discussion

DPS is a typical eco-agricultural landscape that is distributed in plains or deltas covered by dense waterways. There is a lack of studies devoted to mapping DPS at a large scale using satellite data due to the complex combination of ponds and crops growing on dikes. The spatial distribution, ecological function, and environmental impact of DPSs cannot be quantitatively evaluated without an accurate map of DPSs. This study developed the mCascade R-CNN to identify the DPS as a target, achieving an AP value of 80.90%. In previous studies, the overall accuracy of the DPS classification reached 90% and even higher [11,12]. However, the overall accuracy cannot be compared with AP. The accuracy evaluates the performance of the classifier across all classes when all classes are of equal importance. However, this study aimed to identify the DPS, and thus the AP was used instead of the overall accuracy. The AP value not only assesses the accuracy of the target detections, but also takes into account the accuracy of the bounding box. Moreover, the accuracy assessment of the DPS classification studies was based on a small sample size, which is not comparable to the over 1000 samples used for the validation in this study.

The improvements of the mCascade R-CNN over the baseline were not only in the accuracy of the target detection, but also in the accuracy of the bounding box. A more accurate bounding box facilitates better estimation of the crop area within the DPS. However, we found that ponds with very narrow dikes are difficult to identify because the features may be weakened in such cases. The performance could be improved in several ways, such as by increasing the sample sizes, replacing the horizontal bounding box with an oriented bounding box, or testing other advanced deep learning methods. For example, non-maximum suppression is an integral part of the object detection algorithm, but it leads to a missed detection when the bounding boxes significantly overlap with each other. In the detection of DPS, we noticed that one or two were missed occasionally in a row of DPSs. The soft-non-maximum suppression decays the detection scores of all other objects as a continuous function of their overlap with the detection box [25], and may improve the accuracy in rows of DPS.

Crops, vegetables, or fruit trees growing on dikes are usually overlooked because of their small areas and fragmented distribution. In the study area, oilseed rape growing on the dikes accounted for 3.42% of the total oilseed rape area. However, the actual growing area on the dikes is higher than 3.42% due to the uncertainty in the DPS identification and the accuracy of the bounding box. Until 2019, the average cultivated land area per farmer was approximately 0.35 ha [26]. With the increasing demand for agricultural production in China, integrated agriculture systems, such as the crop-fishery-(livestock) system, provide an effective way to balance the limited cultivated land and higher profits from fisheries. As machine learning and computer vision techniques have developed rapidly in recent years, identifying and quantifying crops growing in integrated agriculture systems is more feasible and accurate, which compensates our knowledge of the agricultural and economic conditions of smallholders.

Integrated agriculture systems have developed rapidly in the recent decade with the advancements in new agricultural technology, loss of labors to the cities, and rural revitalization policies, which will drive spatiotemporal change in the DPSs. This study focused on developing a deep learning method to identify DPS, and thus the analyses were conducted on one-date satellite images. Future studies will analyze the spatiotemporal change of the DPSs at the larger scale based on multi-temporal satellite images. In the study area, DPS is mainly composed of a pond and oilseed rape plant, but it has diverse compositions in other regions, such as the Pearl River Delta. The proposed model could be applied to other regions, but it needs large quantities of training samples for the model to learn the different compositions of DPSs. Furthermore, farm ponds are more vulnerable to pollution than larger water bodies [27]. The map of the DPS provides a basic dataset to evaluate the impact of the runoff from dikes on the ponds.

5. Conclusions

This study proposed the mCascade R-CNN algorithm to detect the DPS in Qianjiang City using high-resolution satellite data. The mCascade R-CNN modified the regular convolution layer in the backbone of Cascade R-CNN into a deformable convolutional layer, which was more suitable for learning the features of DPSs with variable shapes and orientations. The mCascade R-CNN algorithm yielded the most accurate detection of DPS, with an AP value that was 2.71% and 11.84% higher than Cascade R-CNN and YOLOv4, respectively. Based on the DPS map derived from the mCascade R-CNN model, KDE analysis illustrated that the DPS was distributed unevenly in the spatial area. The hotspots of the DPS were located in the middle and southwestern parts of the study area. The area of oilseed rape growing on the dikes accounted for 3.42% of the total oilseed rape planting area. This study demonstrates the potential of deep learning methods combined with high-resolution satellite images in detecting integrated agricultural systems. Mapping DPS at a large scale would facilitate the quantification of the environmental-economic benefits of integrate agriculture systems, and provide valuable data sources to support agricultural management of small holders.

Author Contributions: Conceptualization, S.L. and J.L.; methodology, Y.M. and Z.Z.; validation, X.S. and L.Z.; formal analysis, Y.M. and Z.Z.; data curation, X.S. and L.Z.; writing—original draft preparation, S.L., Z.Z. and Y.M.; writing—review and editing, Z.Z. and T.R.; visualization, Y.M., X.S. and L.Z.; funding acquisition, J.L., T.R. and S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by China Agriculture Research System of MOF and MARA (Grant No. CARS-12), Hubei Province Agriculture Research System (Grant No. HBHZD-ZB-2020-005) and the Fundamental Research Funds for the Central Universities (2662021ZH001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the conflict of interest with the on-going research of the corresponding author.

Acknowledgments: The authors are thankful for the valuable suggestions given by Qingfeng Guan from School of Geography and Information Engineering at China University of Geosciences (Wuhan).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Lo, C.P. Environmental Impact on the Development of Agricultural Technology in China: The Case of the Dike-Pond ('Jitang') System of Integrated Agriculture-Aquaculture in the Zhujiang Delta of China. *Agric. Ecosyst. Environ.* **1996**, *60*, 183–195. [\[CrossRef\]](#)
- Yuan, L.; Hu, Y.; Cheng, J. Review of Dike-Pond System, Guangdong. *Guangdong Agric. Sci.* **2014**, *41*, 38–41.
- Karim, M.; Little, D.C. The Impacts of Integrated Homestead Pond-Dike Systems in Relation to Production, Consumption and Seasonality in Central North Bangladesh. *Aquac. Res.* **2018**, *49*, 313–334. [\[CrossRef\]](#)
- Gu, X.; Lou, L.; Liu, M.; Min, Q. Review and Prospect of Studies on the Dyke-Pond System. *J. Nat. Resour.* **2018**, *33*, 709–720.
- Kadoya, T.; Akasaka, M.; Aoki, T.; Takamura, N. A Proposal of Framework to Obtain an Integrated Biodiversity Indicator for Agricultural Ponds Incorporating the Simultaneous Effects of Multiple Pressures. *Ecol. Indic.* **2011**, *11*, 1396–1402. [\[CrossRef\]](#)
- Chen, W.; He, B.; Nover, D.; Lu, H.; Liu, J.; Sun, W.; Chen, W. Farm Ponds in Southern China: Challenges and Solutions for Conserving a Neglected Wetland Ecosystem. *Sci. Total Environ.* **2019**, *659*, 1322–1334. [\[CrossRef\]](#) [\[PubMed\]](#)
- Vote, C.; Eberbach, P.; Inthavong, T.; Lampayan, R.M.; Vongthilard, S.; Wade, L.J. Quantification of an Overlooked Water Resource in the Tropical Rainfed Lowlands Using RapidEye Satellite Data: A Case of Farm Ponds and the Potential Gross Value for Smallholder Production in Southern Laos. *Agric. Water Manag.* **2019**, *212*, 111–118. [\[CrossRef\]](#)
- Mao, D.; Luo, L.; Wang, Z.; Wilson, M.C.; Zeng, Y.; Wu, B.; Wu, J. Conversions between Natural Wetlands and Farmland in China: A Multiscale Geospatial Analysis. *Sci. Total Environ.* **2018**, *634*, 550–560. [\[CrossRef\]](#) [\[PubMed\]](#)
- Stiller, D.; Ottinger, M.; Leinenkugel, P. Spatio-Temporal Patterns of Coastal Aquaculture Derived from Sentinel-1 Time Series Data and the Full Landsat Archive. *Remote Sens.* **2019**, *11*, 1707. [\[CrossRef\]](#)

10. Xia, Z.; Guo, X.; Chen, R. Automatic Extraction of Aquaculture Ponds Based on Google Earth Engine. *Ocean Coast. Manag.* **2020**, *198*, 105348. [\[CrossRef\]](#)
11. Li, F.; Liu, K.; Tang, H.; Liu, L.; Liu, H. Analyzing Trends of Dike-Ponds between 1978 and 2016 Using Multi-Source Remote Sensing Images in Shunde District of South China. *Sustainability* **2018**, *10*, 3504. [\[CrossRef\]](#)
12. Liu, G.; Li, J. Tracking Dike-Pond Landscape Dynamics in a Core Region of the Guangdong-Hong Kong-Macao Greater Bay Area Based on Topographic Maps and Remote Sensing Data during 1949–2020. *Aquaculture* **2022**, *549*, 737741. [\[CrossRef\]](#)
13. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep Learning in Environmental Remote Sensing: Achievements and Challenges. *Remote Sens. Environ.* **2020**, *241*, 111716. [\[CrossRef\]](#)
14. Sidike, P.; Sagan, V.; Maimaitijiang, M.; Maimaitiyiming, M.; Shakoor, N.; Burken, J.; Mockler, T.; Fritsch, F.B. DPEN: Deep Progressively Expanded Network for Mapping Heterogeneous Agricultural Landscape Using WorldView-3 Satellite Imagery. *Remote Sens. Environ.* **2019**, *221*, 756–772. [\[CrossRef\]](#)
15. Zhang, M.; Lin, H.; Wang, G.; Sun, H.; Fu, J. Mapping Paddy Rice Using a Convolutional Neural Network (CNN) with Landsat 8 Datasets in the Dongting Lake Area, China. *Remote Sens.* **2018**, *10*, 1840. [\[CrossRef\]](#)
16. Zhang, D.; Pan, Y.; Zhang, J.; Hu, T.; Zhao, J.; Li, N.; Chen, Q. A Generalized Approach Based on Convolutional Neural Networks for Large Area Cropland Mapping at Very High Resolution. *Remote Sens. Environ.* **2020**, *247*, 111912. [\[CrossRef\]](#)
17. Thorp, K.R.; Drajat, D. Deep Machine Learning with Sentinel Satellite Data to Map Paddy Rice Production Stages across West Java, Indonesia. *Remote Sens. Environ.* **2021**, *265*, 112679. [\[CrossRef\]](#)
18. Li, M.; Zhang, Z.; Lei, L.; Wang, X.; Guo, X. Agricultural Greenhouses Detection in High-Resolution Satellite Images Based on Convolutional Neural Networks: Comparison of Faster R-CNN, YOLO v3 and SSD. *Sensors* **2020**, *20*, 4938. [\[CrossRef\]](#)
19. Chen, W.; Xu, Y.; Zhang, Z.; Yang, L.; Pan, X.; Jia, Z. Mapping Agricultural Plastic Greenhouses Using Google Earth Images and Deep Learning. *Comput. Electron. Agric.* **2021**, *191*, 106552. [\[CrossRef\]](#)
20. Xu, Q.; Wang, Q.; Chen, Z. Climatic Ecological Characteristics and Climate Risk of Crayfish Breeding in Qianjiang. *J. Agric.* **2019**, *9*, 73–77.
21. Sun, W.; Chen, B.; Messinger, D.W. Nearest-Neighbor Diffusion-Based Pan-Sharpening Algorithm for Spectral Images. *Opt. Eng.* **2014**, *53*, 013107. [\[CrossRef\]](#)
22. Cai, Z.; Vasconcelos, N. Cascade R-CNN: Delving Into High Quality Object Detection. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6154–6162. [\[CrossRef\]](#)
23. Zhu, X.; Hu, H.; Lin, S.; Dai, J. Deformable ConvNets v2: More Deformable, Better Results. *arXiv* **2018**, arXiv:1811.11168.
24. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Bodla, N.; Singh, B.; Chellappa, R.; Davis, L.S. Soft-NMS—Improving Object Detection with One Line of Code. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 5562–5570. [\[CrossRef\]](#)
26. Hubei Provincial Statistics Bureau. *Statistical Year Book of Cities and States, Qianjiang City*; Hubei Provincial Statistics Bureau: Wuhan, China, 2020. Available online: <http://tjj.hubei.gov.cn/tjsj/> (accessed on 1 December 2021).
27. Huang, S.-L.; Lee, Y.-C.; Budd, W.W.; Yang, M.-C. Analysis of Changes in Farm Pond Network Connectivity in the Peri-Urban Landscape of the Taoyuan Area, Taiwan. *Environ. Manag.* **2012**, *49*, 915–928. [\[CrossRef\]](#)