

Article

Not peer-reviewed version

Crop-Type Recognition in Street-level Images with Convolutional Neural Networks

[Fernando Orduna-cabrera](#)^{*}, Marcial Sandoval-Gastelum, [Ian MCCALLUM](#), [Linda SEE](#), [Steffen Fritz](#), Santosh Karanam, Tobias Sturn, Valeria Javalera-Rincon, [F. Fernando González-Navarro](#)

Posted Date: 12 July 2023

doi: 10.20944/preprints202307.0724.v1

Keywords: crop type recognition; deep learning; crowdsourcing; street-level imagery



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Crop-Type Recognition in Street-Level Images with Convolutional Neural Networks

Fernando Orduna-Cabrera ^{1,*}, Marcial Sandoval-Gastelum ¹, Ian McCallum ¹, Linda See ¹, Steffen Fritz ¹, Santosh Karanam ¹, Tobias Sturm ¹, Valeria Javalera ¹ and Felix F. Gonzalez-Navarro ²

¹ International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

² Universidad Autónoma de Baja California, Mexicali, México

* Correspondence: orduna@iiasa.ac.at; Tel.: + 43(0) 2236 807 227

Abstract: The creation of crop-type maps from satellite data has proven challenging, often impeded by a lack of accurate in-situ data. This paper aims to demonstrate a method for crop-type (ie. Maize, Wheat and Other) recognition based on Convolutional Neural Networks using a bottom-up approach. We trained the model with a highly accurate dataset of crowdsourced labelled street-level imagery. Classification results achieved an AUC of 0.87 for wheat, 0.85 for maize and 0.73 for other. Given that wheat and maize are the two most common food crops globally, combined with an ever-increasing amount of available street-level imagery, this approach could help address the need for improved crop-type monitoring globally. Challenges remain in addressing the noisy aspect of street-level imagery (ie. buildings, hedgerows, automobiles, etc.), where a variety of different objects tend to restrict the view and confound the algorithms.

Keywords: crop type recognition; deep learning; crowdsourcing; street-level imagery

1. Introduction

The spatial extent of cropland areas has been mapped extensively since the mid-eighties, as increasing numbers of satellites have been launched and higher spatial and temporal resolution imagery has become available. For example, cropland is provided as a land cover class in many global land cover products such as GLC-2000 [1], MODIS land cover [2] and ESA-CCI [3], among many others. More recently, a time series of higher resolution cropland extent products has been produced using Landsat [4] at a 30m resolution while Sentinel-2 is now also being used for land cover mapping, including cropland extent at a 10 m resolution [5]. However, for monitoring food security at national to global scales, spatially explicit crop-type maps are urgently needed [6]. This advance from cropland to crop-type has proven challenging due to among other things, a lack of training data, suitable imagery and advanced algorithms.

With recent advances in analytical methods, data infrastructure and the availability of higher resolution imagery, several recent studies have applied machine-learning techniques for crop-type recognition. Some of the most successful include Support Vector Machines (SVMs) [7–9], Random Forests [9–12], Decision Trees [12–14], the maximum likelihood classifier (MLC) [11], [15], [16], Artificial Neural Networks (ANNs)[11], [17] and Minimum Distance (MD) [11]. Another example is the work undertaken by Mou et al. in [16], where the authors proposed a Deep Recurrent Neural Network (RNN) for hyperspectral image classification. The RNN model effectively analyzed the hyperspectral pixels as sequential data and determined the information categories via network reasoning. The specific application of CNNs in remote sensing for crop-type recognition has also shown excellent performance [13], [16], [18]–[24].

Several recent research studies have achieved higher accuracy in the learning phase because of the CNN implementation. For instance, Cai et al. in [18] introduced a methodology for the cost-effective and in-season classification of field-level crop types using Common Land Units (CLUs) from the United States Department of Agriculture (USDA) to aggregate spectral information based on a time series. The authors built a deep-learning classification model based on Deep Neural Networks

(DNN). The research aimed to understand how different spatial and temporal features affected the classification performance. The experiments also evaluated which input features are the most helpful in training the model and how various spatial and temporal factors affect crop-type classification. Castro et al. in [19] explored three approaches to improve the classification performance for land cover and crop type recognition in tropical areas; they used the image stacking approach as a baseline and AEs and a CNN for two Deep Learning approaches. Their findings outperformed the traditional system based on image stacking in overall and class accuracy.

All of these applications require training data on the presence of different crop types, where the lack of training data is one challenge identified in many research papers. Many studies collect field-based data as part of the development of a training data set (see e.g., [25]), which is costly and often not shared with the broader remote sensing community. Other sources of training data that have yet to be fully exploited include geo-tagged photographs such as those taken at street level. For example, in the study by Wang et al. [20], local farmers contributed by utilizing a mobile application to take pictures and assign a label according to the crop type as training data for crop type mapping. Another approach is where citizens take pictures using mobile phones and then provide a classification label [26], [27] using a tool such as Picture Pile. These approaches [20], [26] are unique since they use images taken from the field. There are also ever-increasing amounts of free and open crowdsourced imagery that are becoming available online, e.g., via Mapillary and Google Street View, so it is important to develop methodologies to exploit these available resources to develop training data sets for crop type mapping using remote sensing.

In this paper we use geotagged crowd-classified street-level photographs in a bottom-up approach to train a CNN utilising a novel deep residual learning architecture [28], referred to here as the Maize-Wheat-Other CNN (MWO CNN). In terms of crop species of global importance to food security, both maize and wheat (and related wheat crops) are crucial to meeting global food demand [29]. The MWO CNN uses the geotagged photographs from the Earth Challenge Food Insecurity crowdsourcing campaign [27]. We refer to these pictures as noisy images because in addition to Maize and Wheat, many objects such as cars, streets, buildings, and people make crop-type classification more complex. Finally, we present results from the CNN model regarding the performance in predicting crop type and discuss the potential of this approach for substantially increasing the amount of training data available for crop type mapping in the future. The model is open access on <https://github.com/iiasa/CropTypeRecognition.git>.

2. Materials and Methods

Crowdsourced Street-level Imagery

A total of 10,776 street-level photographs were selected for this study, the majority of which were taken from Google Street View, including a small amount from Mapillary. The bulk of the images are from France (Figure 1). These images were then placed into the Picture Pile rapid image classification app [30], and then classified by volunteers. The set of images is unique because of the quality of the images, which ranges from poor to excellent for the purposes of crowdsourcing crop species. For example, while some images contain very clear, unobstructed pictures of roadside crops, others contain objects such as cars, houses, etc., in addition to a crop field (Figure 2). **Error! Reference source not found.** lists the total number of images used in this study along with the number of images used in the model training, test, and validation data sets.

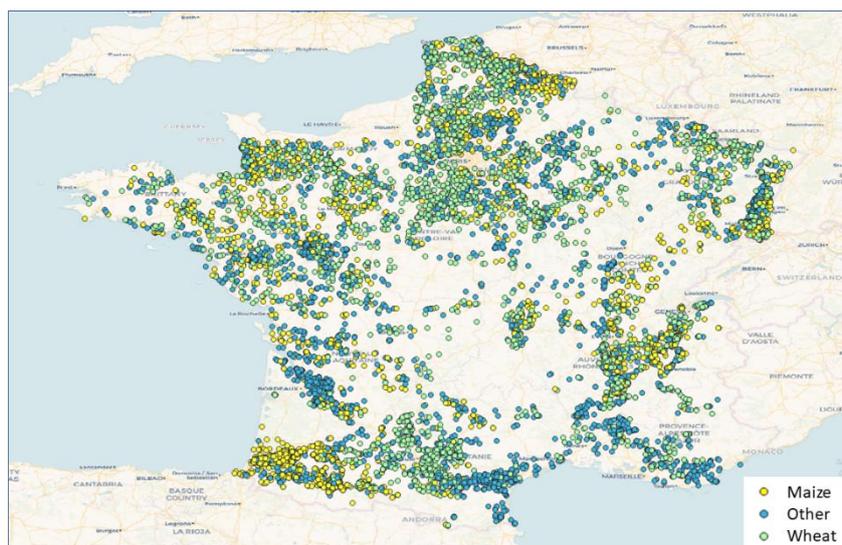


Figure 1. Locations of the majority of the 10,776 street-level images classified via crowdsourcing in Picture Pile, as either Maize, Wheat or Other, over France.

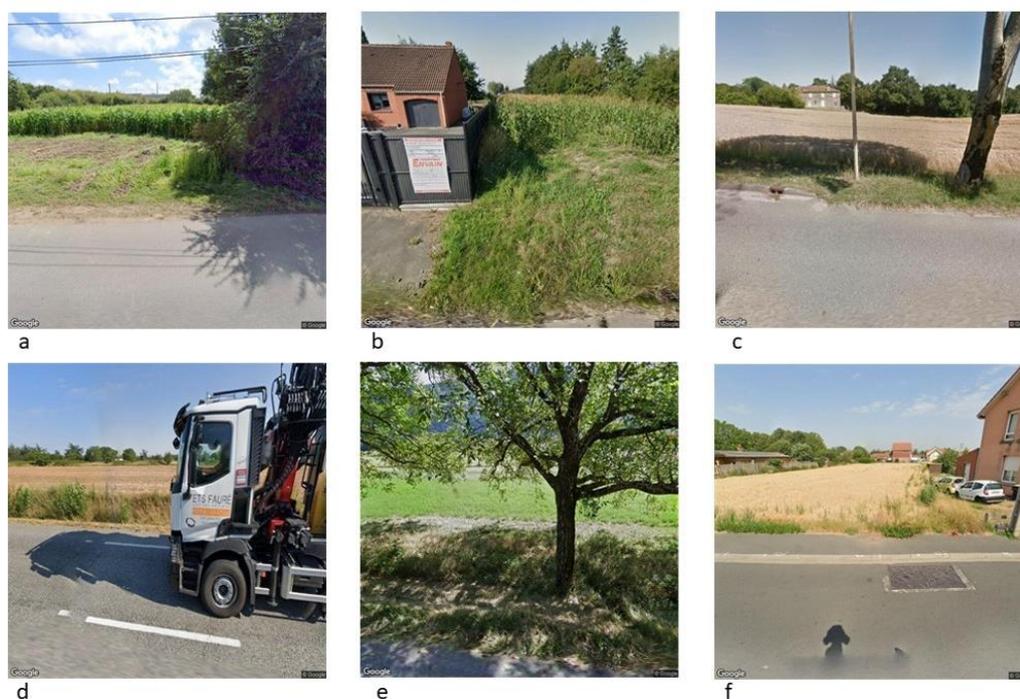


Figure 2. Typical noisy street-level images containing crop species and additional non-crop objects such as roads, buildings, vehicles and trees (a-f). These images were classified by the crowd as a) maize; b) maize; c) wheat; d) wheat; e) other and f)wheat.

Table 1. The total number of classified street-level photographs used in the study, separated by crop type and usage by the CNN.

Crop type	Total images	Test	Training	Validation
Maize	3592	359	2873	360
Wheat	3592	359	2873	360
Other	3592	359	2873	360

In order to ensure accuracy of the crowdsourced image classifications, we created a set of 867 control point images. Each of these images was classified between 5 and 8 times by different

individuals. If a minimum of 5 classifications agreed then we marked that image as a crowdsourced control image. Specifically for France, at the end of the campaign, we compared the crowdsourced results with the Land Parcel Information System (LPIS) of France.

CNN Methodology & Architecture

Convolutional Neural Networks (CNN) are a popular data mining technique for image recognition, first introduced by Fukushima [31]. CNNs consist of a hierarchical multi-layer network of nodes with various connections between the nodes to adjoining layers. LeCun [32] introduced a backpropagation algorithm to compute the gradient to improve CNN performance. The use of CNNs for object classification has been implemented in many domains, achieving a high efficiency and accuracy [33]–[35]. In addition, there has been a considerable amount of research to support the improvement of CNNs, e.g., the study published by Chen et al. [36], where they introduced the concept of Deep Learning into hyperspectral data classification. They verified the eligibility of stacked Autoencoders (AEs) by following classical spectral information based on clustering, focused on spatially dominated information. They also proposed a Deep Learning framework to merge the two features, resulting in high classification accuracies. Chen et al. [37] introduced a 3-D CNN-based model for feature extraction to extract efficient spectral-spatial features for hyperspectral image classification. Another contribution can be found in Makantasis et al. [38], who proposed a deep learning-based classification method that hierarchically constructed high-level features by encoding the spatial and spectral information of the pixels, and a Multi-layer Perceptron (MLP), which was responsible for the classification task.

Network depth

For the proposed CNN architecture used in this study, we applied a specific configuration for the convolution unit block (CUB); the architecture includes a convolution layer, an activation layer, and a subsampling layer. We tested different filters (or kernels) for each CUB to evaluate the output size. We performed a set of iterations to find the network depth that yielded the best results. Considering the computational cost, we used a CUB that was from 2 to 4 layers deep. After several iterations, we set the architecture to have a depth of 2 CUBs.

Layer kernel and stride size

The set of pictures used contains details such as the size of the objects, the colors, different shapes, and other variables. The experiments undertaken considered several variables such as the width, the size of the kernel, and the stride size. A kernel refers to a small matrix of learnable weights that is applied to the input image to extract its main features during the convolution operation. The stride hyperparameter refers to the number of pixels the filter or kernel is shifted or moved at each step when performing convolution. The experiments gave better results with a small kernel and a smaller stride size. The kernel size is reduced to a new two-by-two shape, and the stride is set up to advance two steps per axis during the convolution operation.

Dense layer hidden units

The units in a dense layer are individual artificial neurons connected to all the neurons in the preceding layer that produce a single output value. This output value is calculated using a weighted sum of the inputs, where the weights are the parameters learned in the kernels by the neural network during the training process. For our architecture, we set the total number of hidden units in the dense layer to double the output size of the last convolution layer. Figure 3 shows the first convolutional output layer with 16 units. Hence, the output layer is followed by a second convolutional layer with 32 units.

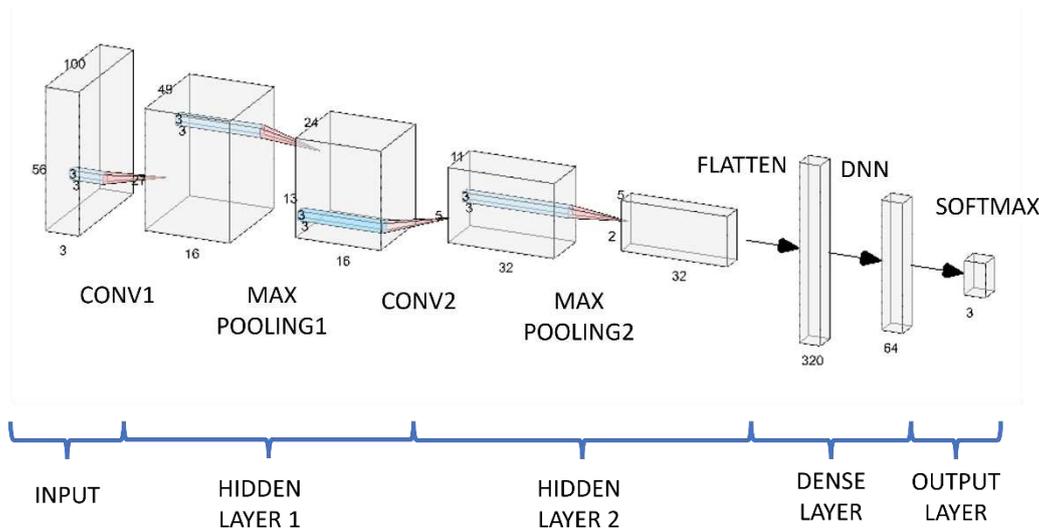


Figure 3. The CNN architecture used in the Maize-Wheat-Other (MWO) model.

Activation and loss functions

We did a grid search using different loss and activation functions to find the combination that yields the best performance. The search method was performed by training the model based on the architecture depicted in **Error! Reference source not found. 3. Error! Reference source not found.** shows the pairs of activation and loss functions used in the model experiments.

Table 2. Different activation and loss functions for the experiments.

Activation function	Loss function
RELU	Mean Squared Error (MSE)
IDENTITY	POISSON
TANH	MEAN SQUARED LOGARITHMIC ERROR

Training epochs

Adjusting the hyper-parameters for the training epoch is mandatory for improving the performance results. If the training phase rates are low, more training iterations are required to stabilize the learning-time curve. After fifteen epochs, the learning curve stabilized; therefore, the maximum training epoch was initialized to that value.

Evaluation methods

The proposed approach was tested using the MWO CNN. The experiment's outcome for each trained model consists of assessing the accuracy and the Receiver Operating Characteristic (ROC) Curve [39] evaluating the multi-class model. The Precision [40] assesses the proportion of samples correctly classified. The F1-Score evaluates the accuracy and sensitivity using the harmonic mean [40]. The sensitivity and specificity evaluate the proportional classification for positive and negative samples [39]. We split the data into 80% for training, 10% for the test, and 10% for validation.

3. Results

Crowdsourcing

Using the crowdsourcing classifications and the parcel information from the official French 2016-2019 Land Parcel Information Systems (LPIS), we computed a confusion matrix to examine the performance of the crowd [41]. Since each image is labelled by more than one person, we selected those classifications from the database in which a minimum of 8 classifications per location were collected and there was a majority agreement, i.e., at least 5 classifications were of the same crop type. Based on this, we found an accuracy of 63.1%. We then discarded the non-cropland class, which resulted in an overall accuracy of 98.7%.

The final confusion matrix is shown in Table 3. We can justify discarding the non-cropland class because in this class, we included crops that have already been harvested and hence no crop could be identified from these photographs.

Table 3. Confusion matrix with land parcel information in columns and volunteer classifications as rows.

	Wheat type crop	maize	Sun-flower	Vine-yard	Sorghum	Olive trees	Other crop	Total
Wheat type crop	468	2	0	1	1	0	3	475
maize	2	589	0	0	0	0	0	591
Sun-flower	0	1	46	0	0	0	2	49
Vine-yard	0	1	0	939	0	0	0	940
Sorghum	0	4	0	0	1	0	0	5
Olive trees	0	0	0	0	0	0	0	0
Other crop	9	1	0	0	0	0	6	16
Total	484	598	46	941	2	0	11	0.986

MWO CNN

Table 4 shows the overall results for the architecture shown in Figure 1 for MWO CNN, using noisy street-level images and recognizing three kinds of pictures with an overall accuracy of 75.93%. Figure 4 depicts the ROC values for each class. The ROC curve shows the trade-off between the precision and the specificity. As a reference, the ideal classifier should have a high precision and low specificity. The area under the ROC curve (AUC) is a measure of the classifier's overall performance, where a value of 1 indicates a perfect classifier (e.g., no wrong classification in all the test samples) and a value of 0.5 indicates that the classifier performs no better than random chance. Figure 4 depicts the model outcome; this is more accurate when classifying pictures of wheat crops; which is the class with the highest AUC of 0.8722. The second-best performance corresponds to the class "Maize", with an AUC value of 0.8485. As can be seen, the model struggled to classify images of the "Other" class successfully.

Table 4. Results for the MWO CNN by the Maize, Wheat and Other classes.

CROP	PRE	RECALL	F1	AUC
Maize	79.18	80.28	79.72	0.85
Wheat	77.67	86.94	82.04	0.87
Other	69.87	60.56	64.88	0.73

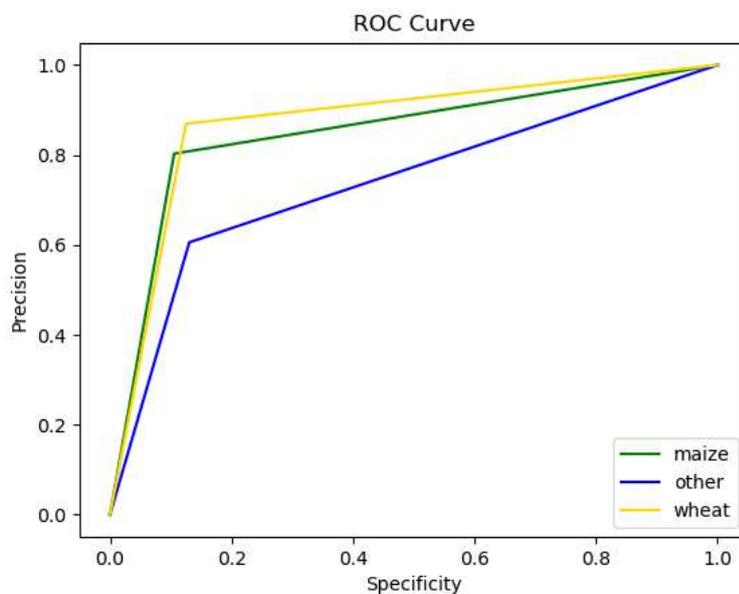


Figure 4. The Receiver Operating Characteristic (ROC) curve for the MWO CNN model for maize, wheat and other classes.

4. Discussion

As crop-type detection from satellite data has proven challenging, we are exploring alternative methods. Here we introduced a deep-learning architecture to classify noisy street-level images according to the following three classes: maize, wheat, and other objects. In addition to the crops of interest, street-level imagery may include objects such as cars, roads, buildings, trees, people, other crops, and more. Because of the nature of the viewing angle for street-level imagery, automatic classification can prove challenging as the above-mentioned objects often obscure the view.

However, there is a vast and growing archive of street-level imagery available across the globe, which offers great potential for gaining insight into crop-type mapping. Interesting to note is the high level of accuracy documented during the crowdsourcing approach (> 95%). While it was not possible to replicate this level of accuracy with the MWO CNN model, initial results are still promising (AUC of 0.87 for wheat and AUC of 0.85 for maize).

After analyzing the nature of the images corresponding to the other class, we noticed that many contain non-crop objects and crop-types different than maize or wheat. The similarity between the additional crops and those of interest (i.e., maize and wheat) make it difficult for the model to distinguish between them. As a result, we achieved a low performance of AUC of 0.73 for the other class.

Nonetheless these initial results are promising owing to the vast potential of this data as an in-situ dataset. With additional improvements, classified street-level imagery could provide a powerful training dataset for global satellite mapping. Crop-type information, combined with the image acquisition dates could be ingested into various global land products.

5. Conclusion

We have introduced the first fully open crop type recognition system based on a Convolutional Neural Network (CNN) architecture for a crop-type recognition application using deep learning to classify two specific crop-types in street-level images. The architecture demonstrates the application of CNN methods to recognize maize, wheat, and other classes in the images.

The MWO CNN model was trained using more than 8,000 crowdsourced street-level images from a Picture Pile campaign over France, where citizens contributed to labeling the images. The crowdsourced images were classified with an accuracy of > 95%. The MWO CNN model achieved an AUC of 0.87 for wheat and 0.85 for maize, the two most predominant global crops. The other class achieved an AUC of 0.73. Given the specific viewing angle of street-level imagery, various non-crop structures impeded the view confounding the algorithms. Nonetheless, this method holds great

potential to massively increase our ability to globally track important crop-types as the amount of street-level imagery continues to increase globally.

Funding: This study has been supported by the Open-Earth-Monitor Cyberinfrastructure project with funding from the European Union's Horizon Europe research and innovation program under grant agreement № 101059548.

Data Availability Statement: The data presented in this study are available on <https://github.com/iiasa/CropTypeRecognition.git>

References

1. E. Commission, J. R. Centre, S. Fritz, E. Bartholomé, and A. Belward, *Harmonisation, mosaicing and production of the Global Land Cover 2000 database (Beta Version)*. Publications Office, 2004.
2. D. Sulla-Menashe, J. M. Gray, S. P. Abercrombie, and M. A. Friedl, "Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product," *Remote Sensing of Environment*, vol. 222, pp. 183–194, 2019.
3. S. Bontemps *et al.*, "Multi-year global land cover mapping at 300 m and characterization for climate modelling: achievements of the Land Cover component of the ESA Climate Change Initiative," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-7/W3, pp. 323–328, 2015.
4. P. Potapov *et al.*, "Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century," *Nature Food*, vol. 3, no. 1, pp. 19–28, 2022.
5. Z. S. Venter, D. N. Barton, T. Chakraborty, T. Simensen, and G. Singh, "Global 10 m Land Use Land Cover Datasets: A Comparison of Dynamic World, World Cover and Esri Land Cover," *Remote Sensing*, vol. 14, no. 16, 2022.
6. S. Fritz *et al.*, "A comparison of global agricultural monitoring systems and current gaps," *Agricultural Systems*, vol. 168, pp. 258–272, 2019.
7. R. Sonobe, H. Tani, X. Wang, N. Kobayashi, and H. Shimamura, "Discrimination of crop types with TerraSAR-X-derived information," *Physics and Chemistry of the Earth*, vol. 83–84, pp. 2–13, 2015.
8. J. Guo, P. L. Wei, J. Liu, B. Jin, B. F. Su, and Z. S. Zhou, "Crop Classification Based on Differential Characteristics of H α Scattering Parameters for Multitemporal Quad- and Dual-Polarization SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. 6111–6123, Oct. 2018.
9. S. Feng, J. Zhao, T. Liu, H. Zhang, Z. Zhang, and X. Guo, "Crop Type Identification and Mapping Using Machine Learning Algorithms and Sentinel-2 Time Series Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 9, pp. 3295–3306, Sep. 2019.
10. F. Vuolo, M. Neuwirth, M. Immitzer, C. Atzberger, and W. T. Ng, "How much does multi-temporal Sentinel-2 data improve crop type classification?," *International Journal of Applied Earth Observation and Geoinformation*, vol. 72, pp. 122–130, Oct. 2018.
11. Y. Zhan, S. Muhammad, P. Hao, and Z. Niu, "The effect of EVI time series density on crop classification accuracy," *Optik*, vol. 157, pp. 1065–1072, Mar. 2018.
12. R. Sonobe, H. Tani, X. Wang, N. Kobayashi, and H. Shimamura, "Random forest classification of crop type using multioral TerraSAR-X dual-polarimetric data," *Remote Sensing Letters*, vol. 5, no. 2, pp. 157–164, Feb. 2014.
13. H. McNairn, A. Kross, D. Lapen, R. Caves, and J. Shang, "Early season monitoring of corn and soybeans with TerraSAR-X and RADARSAT-2," *International Journal of Applied Earth Observation and Geoinformation*, vol. 28, pp. 252–259, 2014.
14. H. McNairn, J. Shang, X. Jiao, and C. Champagne, "The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 12, pp. 3981–3992, 2009.
15. Y. Chen *et al.*, "Mapping croplands, cropping patterns, and crop types using MODIS time-series data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 69, pp. 133–147, Jul. 2018.
16. B. K. Kenduiywo, D. Bargiel, and U. Soergel, "Higher Order Dynamic Conditional Random Fields Ensemble for Crop Type Classification in Radar Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4638–4654, Aug. 2017.
17. C. C. Júnior *et al.*, "Artificial Neural Networks and Data Mining Techniques for Summer Crop Discrimination: A New Approach," *Canadian Journal of Remote Sensing*, vol. 45, no. 1, pp. 16–25, Jan. 2019.
18. Y. Cai *et al.*, "A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach," *Remote Sensing of Environment*, vol. 210, pp. 35–47, 2018.
19. J. D. B. Castro, R. Q. Feitoza, L. C. La Rosa, P. M. A. Diaz, and I. D. A. Sanches, "A Comparative Analysis of Deep Learning Techniques for Sub-Tropical Crop Types Recognition from Multitemporal Optical/SAR

- Image Sequences," in *Proceedings - 30th Conference on Graphics, Patterns and Images, SIBGRAPI 2017*, 2017, pp. 382–389.
20. S. Wang *et al.*, "Mapping Crop Types in Southeast India with Smartphone Crowdsourcing and Deep Learning," *Remote Sensing*, vol. 12, no. 18, 2020.
 21. Y. Wu, P. Wu, Y. Wu, H. Yang, and B. Wang, "Remote Sensing Crop Recognition by Coupling Phenological Features and Off-Center Bayesian Deep Learning," *Remote Sensing*, vol. 15, no. 3, 2023.
 22. H. Pei, T. Owari, S. Tsuyuki, and Y. Zhong, "Application of a Novel Multiscale Global Graph Convolutional Neural Network to Improve the Accuracy of Forest Type Classification Using Aerial Photographs," *Remote Sensing*, vol. 15, no. 4, 2023.
 23. G. Li *et al.*, "Multi-Year Crop Type Mapping Using Sentinel-2 Imagery and Deep Semantic Segmentation Algorithm in the Hetao Irrigation District in China," *Remote Sensing*, vol. 15, no. 4, 2023.
 24. F. Weilandt *et al.*, "Early Crop Classification via Multi-Modal Satellite Data Fusion and Temporal Attention," *Remote Sensing*, vol. 15, no. 3, 2023.
 25. J. Fowler, F. Waldner, and Z. Hochman, "All pixels are useful, but some are more useful: Efficient in situ data collection for crop-type mapping using sequential exploration methods," *International Journal of Applied Earth Observation and Geoinformation*, vol. 91, p. 102114, 2020.
 26. L. See *et al.*, "Lessons learned in developing reference data sets with the contribution of citizens: the Geo-Wiki experience," *Environmental Research Letters*, vol. 17, no. 6, p. 65003, May 2022.
 27. D. Fraisl *et al.*, "Demonstrating the potential of Picture Pile as a citizen science tool for SDG monitoring," *Environmental Science & Policy*, vol. 128, pp. 81–93, 2022.
 28. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." arXiv, 2015.
 29. O. Erenstein, J. Chamberlin, and K. Sonder, "Estimating the global number and distribution of maize and wheat farms," *Global Food Security*, vol. 30, p. 100558, 2021.
 30. O. Danylo *et al.*, "The Picture Pile tool for rapid image assessment: A demonstration using Hurricane Matthew," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 27–32, 2018.
 31. K. Fukushima, "Neocognitron: A Hierarchical Neural Network Capable of Visual Pattern Recognition," 1988.
 32. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998.
 33. D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, High Performance Convolutional Neural Networks for Image Classification."
 34. P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, vol. 2003-Janua, pp. 958–963, 2003.
 35. T. Wiesner-Hanks *et al.*, "Millimeter-Level Plant Disease Detection From Aerial Photographs via Deep Learning and Crowdsourced Data," *Frontiers in Plant Science*, vol. 10, 2019.
 36. Y. Chen, Z. Lin, X. Zhao, W. Gang, Y. Gu, and S. Member, "Deep Learning-Based Classification of Hyperspectral Data," vol. 7, no. June 2014, pp. 1–14, 2015.
 37. Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232–6251, 2016.
 38. K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2015-Novem, pp. 4959–4962, 2015.
 39. A. J. Bowers and X. Zhou, "Receiver Operating Characteristic (ROC) Area Under the Curve (AUC): A Diagnostic Measure for Evaluating the Accuracy of Predictors of Education Outcomes," *Journal of Education for Students Placed at Risk (JESPAR)*, vol. 24, no. 1, pp. 20–46, 2019.
 40. S. S. M. M. Rahman, F. B. Rafiq, T. R. Toma, S. S. Hossain, and K. B. B. Biplob, "Performance assessment of multiple machine learning classifiers for detecting the phishing URLs," in *Data Engineering and Communication Technology: Proceedings of 3rd ICDECT-2K19*, 2020, pp. 285–296.
 41. S. Fritz *et al.*, "Crowdsourcing In-Situ Data Collection Using Gamification," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 254–257.

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