IEEE Geoscience and Remote Sensing Letters

Productive Crop Field Detection: A New Dataset and Deep Learning Benchmark Results

Journal:	*Geoscience and Remote Sensing Letters*
Manuscript ID	GRSL-00951-2023.R1
Manuscript Type:	Special Stream on Pattern Recognition and Image Processing for Geoscience andRemote Sensing of the SIBGRAPI 2023 – 36th Conference on Graphics, Patterns andImages.
Sub-topic:	Image Processing, Analysis and Classification
Date Submitted by the Author:	n/a
Complete List of Authors:	Garcia do Nascimento, Eduardo; UFSCar - Campus Sorocaba, Computer Science Just, John; Iowa State University, Computer Science; John Deere, ISG - Data Science and Analytics Almeida, Jurandy; Universidade Federal de São Carlos - Campus Sorocaba, Department of Computing Almeida, Tiago; UFSCar - Campus Sorocaba, Computer Science
Key Words:	Computational Infrastructure and GIS, Electromagnetics and Remote Sensing, Vegetation and Land Surface < Methodologies and Applications to, Optical Data < Processing, Sensors and Systems for

We thank the associate editor and reviewers for carefully analyzing our paper. In the following, we present the explanations for the comments and detail how we have addressed each raised issue. To facilitate tracking of changes made, we have referenced the figures and tables in this letter with the same number as in the manuscript.

Associate Editor

Reviewer's comment. *Dear authors, to ensure the quality and impact of your research, we kindly request that you carefully address the reviewers' comments and make the necessary revisions to your manuscript. Please provide clear and detailed responses to each comment, explaining how you have addressed the concerns raised.*

Response. We thank the associate editor and reviewers for the careful review of our paper and the excellent suggestions.

To address the issues raised by the reviewers, we have carefully revised the paper and provided significant improvements to the text, the description of the dataset, and better clarification of the proposed methodology and experiments. We have also fixed inconsistencies and minor errors.

Reviewer #1

Reviewer's comment. *The manuscript was written in good-quality English. However, some paper presentation choices are, as far as I am concerned, mistaken. For example, precious space had been wasted addressing related works which add little relevant information to the article in terms of content for a paper that brings a benchmark. Considering the small space available in GRSL, I recommend sacrificing section II in favor of improving sections III and IV. For* *instance, it would be important to bring more details about what data you produce for a given hexagon. In addition, it is necessary to explain in depth the data presented as input to the classifiers. Another point that requires more details is what are the network architectures employed. Without such sort of information, it would be better to study your code than to read your manuscript.*

Response. We want to express our gratitude for taking the time to review our paper. We appreciate your feedback and valuable suggestions for improving the manuscript. We have carefully considered your comments and revised the manuscript to address the issues raised. Regarding the concern about the space allocation in the paper, we agree that the presentation of related works, despite our efforts to demonstrate a robust background, could have been more concise, considering the limited space available in GRSL.

Changes. We have considered your suggestion and made significant revisions to reduce Section II, removing information that is not essential to this work or can be consulted in the bibliography, allowing us to allocate more space to Sections III and IV. These sections now provide complete information about the data generation process for each hexagon, including additional details on the input data used for the classifiers. Furthermore, we have provided in-depth explanations of the input data in Figure 1 (Revised manuscript – page 3) and network architectures employed in the research in Figure 5 (Revised manuscript – page 4), ensuring readers may understand our approach comprehensively. Below we present the main changes in the manuscript to clarify the raised issues.

Table I summarizes the dataset and its fields. The dataset includes 17 productive corn crop fields in the US within the same geographical area. It has 106,735 rows, each representing a hexagon captured on a specific date and time, followed by the 12 Sentinel-2 band values. From these, 75,990 rows are labeled as positive (i.e., productive crop fields hexagons tracked over the period) and 30,745 as negative (i.e.,

non-productive crop fields hexagons tracked over the period). The amount of positive hexagons representing geolocations is 5,066, while the number of negative hexagons is 2,050 resulting in a slightly unbalanced dataset. The number of images for each hexagon collected on different dates varies from 15 to 38, and the averages for each field are represented in the last column.

Table 1: Productive Fields Dataset Description

	Fields: 17				
	Hexes (samples)		Hex images collected in the period		
Field	positive	negative	positive	negative	average
F01	558	210	9,860	3,539	17.67 ±0.48
F02	734	151	14,242	2,863	19.40 ±0.74
F03	59	85	1,075	1,528	18.22 ±0.65
F04	350	150	8,326	3,518	23.79 ±0.48
F05	75	134	1,815	3,001	24.20 ±0.90
F06	288	60	4,999	1,044	17.36 ±0.68
F07	418	140	9,409	2,950	22.51 ±0.70
F08	425	154	8,584	3,084	20.20 ±0.82
F09	147	127	2,993	2,563	20.36 ±0.71
F10	369	176	6,973	3,194	18.90 ±0.40
F11	228	57	4,813	1,172	21.11 ±0.59
F12	91	96	1,701	1,738	18.69 ±0.53
F13	151	86	3,271	1,840	21.66 ±0.49
F14	275	216	4,871	3,825	17.71 ±0.46
F15	370	198	7,369	3,913	19.92 ±0.66
F16	271	25	5,030	413	18.56 ±0.69
F17	281	17	5,099	297	18.15 ±0.97

(Revised manuscript – page 3)

Figure 1 illustrates how we create the samples used as the input for the learning models.

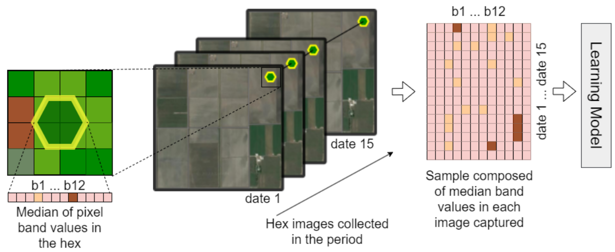


Figure 1: Sequence to build a sample based on the hexagon time-series

(Revised manuscript – page 2)

We have implemented the DL-based neural networks using 4 convolutional layers with a kernel size of 3 and stride of 1, as demonstrated in Figure 5.

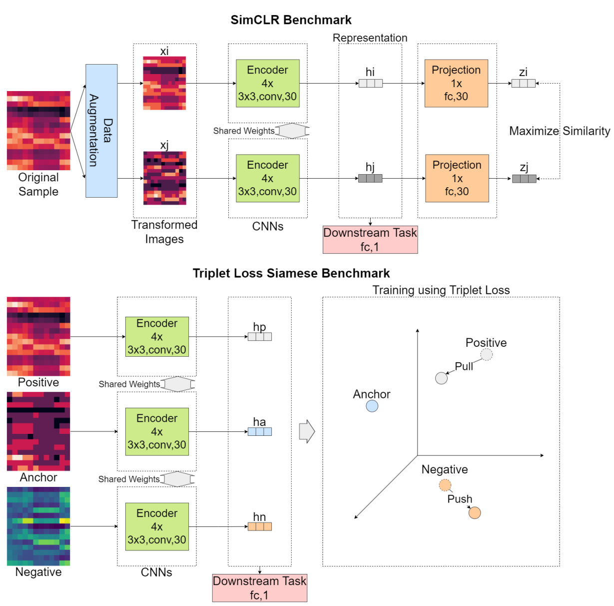


Figure 5: Neural network architectures employed in the benchmark testing

(Revised manuscript – page 4)

Reviewer #2

Reviewer's comment. *Comments to the Author: In this letter, the authors propose a dataset and present benchmark results for automatically detecting productive crop fields using machine learning and satellite imagery. The aim is to address the challenges posed by limited labeled data in precision agriculture. The proposed dataset offers valuable resources for further research and comparison in this field. However, to enhance understanding, I suggest reworking and resubmitting the letter. I have some doubts and remarks regarding the dataset, which I believe should be addressed.*

Response. We thank your valuable feedback on our paper. We appreciate your positive recognition of the proposed dataset and its potential for further research and comparison in precision agriculture. We also acknowledge your suggestion to rework and resubmit the manuscript to enhance its understanding. We appreciate your engagement with our work and your willingness to provide specific remarks and doubts regarding the description of the dataset and benchmark experiments. All the raised comments were addressed and the manuscript was updated accordingly.

Major points

Reviewer's comment. *Issue 1: In the letter, the authors state that "manually identifying productive fields is often a time-consuming and error-prone task." This statement could be misleading, as it may imply that machine learning approaches are error-free. I recommend rephrasing this assertion in the abstract and introduction to avoid confusion.*

Response. We thank the reviewer for commenting on how machine learning methods can contribute to this scenario and the limitations of these methods.

Changes. We have changed the text to clarify and avoid any possible misleading statements. Below we present the changes in the abstract and introduction to clarify the raised issues.

However, manually identifying productive fields is often time-consuming, costly, and subjective. Previous studies explore different methods to detect crop fields using advanced machine learning algorithms to support the specialists' decisions, but they often lack good quality labeled data. (Revised manuscript – Abstract)

Despite the expert's ability and knowledge, finding crop fields and drawing their boundaries has been a major, expensive, and time-consuming challenge [6]. (Revised manuscript – Introduction)

Reviewer's comment. *Issue 2: There is missing information about the dataset, such as the area of interest where the fields are located. It would be helpful to include a figure depicting the distribution of fields across the entire target region. Additionally, clarification is needed regarding the statement: "The amount of unique positive hexagons is 5,066, while the number of negative hexagons is 2,050, resulting in a slightly unbalanced dataset."*

Response. Unfortunately, due to data protection rules and customer confidentiality, we cannot share some details about the dataset, specifically the area of interest. After extensive work with John Deere's legal department, it has been agreed that we cannot provide the requested information per the limitations imposed by customer contracts and the potential harm in disclosing their crop operation patterns. However, it is important to highlight that the absence of such information does not affect the value of the dataset, mainly regarding its full potential for training accurate prediction models.

Changes. Below we present the changes in the manuscript to clarify the raised issues about the dataset. To avoid any confusion and clarify the description of the dataset, we first improved the presentation of Table I. Column names, as well as the content, were modified to reflect the most accurate representation of the data shared by this work.

Table 1: Productive Fields Dataset Description

	Fields: 17				
	Hexes (samples)		Hex images collected in the period		
Field	positive	negative	positive	negative	average
F01	558	210	9,860	3,539	17.67 ±0.48
F02	734	151	14,242	2,863	19.40 ±0.74
F03	59	85	1,075	1,528	18.22 ±0.65
F04	350	150	8,326	3,518	23.79 ±0.48
F05	75	134	1,815	3,001	24.20 ±0.90
F06	288	60	4,999	1,044	17.36 ±0.68
F07	418	140	9,409	2,950	22.51 ±0.70
F08	425	154	8,584	3,084	20.20 ±0.82
F09	147	127	2,993	2,563	20.36 ±0.71
F10	369	176	6,973	3,194	18.90 ±0.40
F11	228	57	4,813	1,172	21.11 ±0.59
F12	91	96	1,701	1,738	18.69 ±0.53
F13	151	86	3,271	1,840	21.66 ±0.49
F14	275	216	4,871	3,825	17.71 ±0.46
F15	370	198	7,369	3,913	19.92 ±0.66
F16	271	25	5,030	413	18.56 ±0.69
F17	281	17	5,099	297	18.15 ±0.97

Table I summarizes the dataset and its fields. The dataset includes 17 productive corn crop fields in the US within the same geographical area. It has 106,735 rows, each representing a hexagon captured on a specific date and time, followed by the 12 Sentinel-2 band values. From these, 75,990 rows are labeled as positive (i.e., productive crop fields hexagons tracked over the period) and 30,745 as negative (i.e., non-productive crop fields hexagons tracked over the period). The amount of positive hexagons representing geolocations is 5,066, while

the number of negative hexagons is 2,050 resulting in a slightly unbalanced dataset. The number of images for each hexagon collected on different dates varies from 15 to 38, and the averages for each field are represented in the last column. (Revised manuscript – page 3)

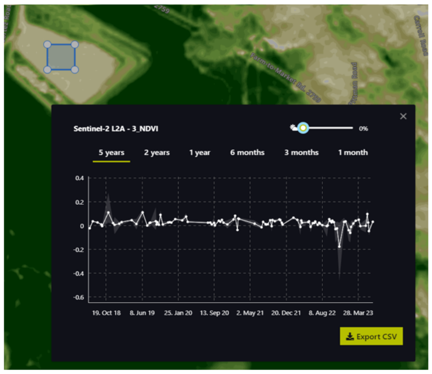
Reviewer's comment. *Issue 3: The selection of negative hexagons appears to be limited to the edges of positive hexagons. Could you explain the rationale behind this choice? From my understanding, it is also possible to have non-productive fields (as mentioned in the text, "typical farmers are usually interested in monitoring productive areas rather than all possible fields since many are not fertile or suitable for any crop"), which could be considered as negative samples. Why was this case not considered?*

Issue 4: Building upon the previous issue, the authors mention in the related works that "Most aforementioned studies were devised for detecting or segmenting all possible fields, regardless of whether they are productive." How does the proposed dataset address this issue if the negative samples are selected only around the positive fields and not within fields that are not productive?

Response. We appreciate your thoughtful question and the opportunity to provide further clarification.

The rationale behind our choice to select negative hexagons primarily from the edges of positive hexagons is based on the knowledge that the surroundings of farm fields are usually, however not always, non-field areas (e.g., forests, lakes, constructed areas, and others), which is enough for the algorithm to calculate the network weights properly. In other words, these areas train the algorithms to recognize diverse non-productive areas where the patterns of band values over time are distinctly different than productive cropped areas. A few examples with NDVI are demonstrated in Figure 7 with the help of the software called Sentinel Hub[1] .

[1]Sentinel Hub EO Browser. 2023. Freely available service-oriented satellite imagery tool.



(a) Non-productive field

(b) Road

(c) Productive field

Figure 7: Samples of different types of areas and their respective NDVI patterns observed with Sentinel Hub

Additionally, by selecting areas next to the fields, we are forcing the machine learning models to separate the most complex conditions turning the classification of non-productive areas far from crop fields into a simpler task.

As a consequence, the dataset works well for identifying non-productive fields patterns because the areas around productive fields (e.g., ditches full of grass/weeds/dirt, roads paved/dirt, trees, canals) are represented by time series of imagery patterns that much more closely match non-productive fields than productive fields. A fallow field with low vegetation (Figure 7(a)) will show

Available at: `https://www.sentinel-hub.com/`

patterns of band values over time that are closer to the patterns of a road, as demonstrated in Figure 7(b), while a productive field will appear in a more distant pattern, as shown in Figure 7(c). Even if grass or weeds grow in a non-productive field seasonally, the pattern will appear more like the ditches around productive fields.

In conclusion, by training an algorithm with the negative hexagons, we allow the prediction model to distinguish productive areas from non-productive, even if they visually look like one which is. To our knowledge, it is a great differential from previous studies.

Changes. To address the issue, we have added a brief explanation of the NDVI pattern that the method relies on to differentiate productive from non-productive fields. Moreover, we have included the Figure 3 (Revised manuscript – page 3) with an example of a satellite image with a region with the exact shape of a productive field, however properly classified as non-productive, thanks to the patterns learned with the negative samples provided by this dataset. Below we present the changes in the manuscript to clarify the raised issues.

One noteworthy aspect of the proposed method is its ability to distinguish non-productive fields that share the same shape as productive ones. Even a highly skilled individual would struggle to make this distinction. The key reason behind this lies in the trained model's ability to recognize patterns of band values over time from non-productive areas as distinctly different than productive cropped areas. This approach leads to significantly more accurate results. Figure 3 shows an example of a non-productive field correctly identified.

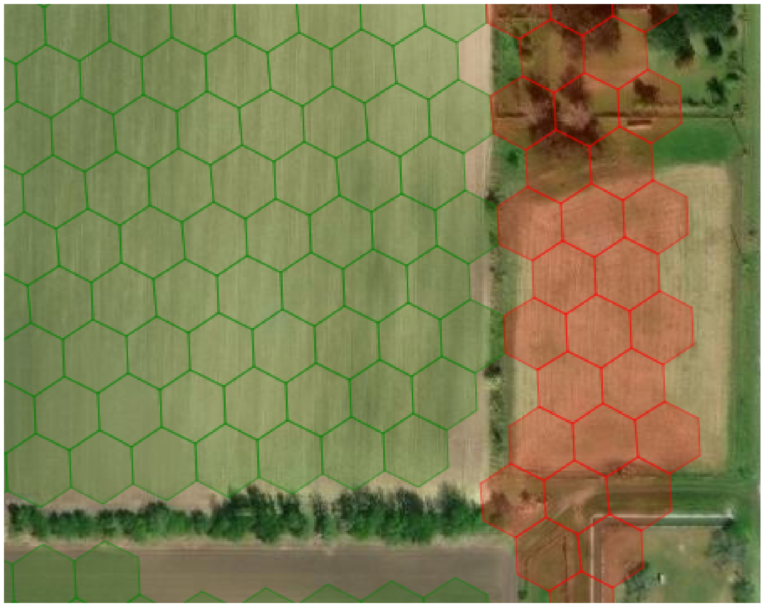


Figure 3: A non-productive field properly classified with red hexagons

(Revised manuscript – page 3)

Reviewer's comment. *Issue 5: In the related work section, the authors mention that "Evaluated on rural area maps, it detected 99% of total acreage, but it achieved poor performance close to urban areas." This statement seems irrelevant, as the authors do not differentiate between rural and urban regions when describing the dataset, methodology, and results.*

Response. We appreciate the opportunity to provide clarification and address your concern.

Changes. When describing the dataset, methodology, and results, we do not differentiate between rural and urban regions. The statement you quoted just briefly references a previous study's findings regarding the performance of a different approach in different settings. Our initial intent was to highlight the pros and cons of each related work. In the revised manuscript, we decided to remove that statement but keep that study since it has both connection and relevant contributions that we based on to build our research.

Minor points

Reviewer's comment. *Issue 1: Please verify the publication years. The text states, "In the same year, Kussul et al. [4] ..." referring to the same year as when "Garcia-Pedrero et al. [2] applied the U-Net..." This is incorrect; Kussul was published in 2017, while Garcia-Pedrero was published in 2019.*

Response. We thank you for bringing that to our attention. Upon verification, we can confirm that Kussul et al. [4] was indeed published in 2017, while Garcia-Pedrero et al. [2] was published in 2019. The mistake happened because the authors of the latest publication had two papers: one in 2017 and the other in 2019.

Changes. To address the issue we revised this part of the text associating Kussul et al. [4] with the right work being García-Pedrero et al. [3] and not Garcia-Pedrero et al. [2]. Below we present the changes in the manuscript to clarify the raised issue.

For instance, García-Pedrero et al. [3] achieved an accuracy of 92% using an ensemble algorithm called RUSBoost [5] to merge super-pixels and group blocks of the image, which were part of the same field. This study demonstrated how ML is a promising alternative and a precursor to further development using DL [3]. In the same year, Kussul et al. [4] applied DL to segment and classify crop types. They demonstrated the advantages of deeper networks by comparing a Multilayer Perceptron (MLP) with a CNN, where the latter achieved higher accuracies. Later, Garcia-Pedrero et al. [2] applied the U-Net, a convolutional neural network (CNN) for semantic im-

age segmentation, to perform the same task with improved performance [2, 3, 1]. (Revised manuscript – pages 2)

Reviewer's comment. *Issue 2: Please provide a description of the columns (nfield, f uniq, nf uniq) in Table I.*

Changes. We have revised Table I to clarify the columns' names and ensure we share appropriate data with no redundancy. The word 'unique' was removed because it confused this and other parts of the paper. In addition, we removed the total column, which could be inferred by summing up field and non-field columns. Finally, we have added the standard deviation of the number of different dates for each field to obtain more granularity of this information. As a consequence of these changes, we improved the overall quality of the information shared. Moreover, we made it more concise by removing the boxplot in the sequence since the standard deviations were better detailed by field in each line.

Reviewer's comment. *Issue 3: Please provide details on how the final classification was performed in the Triplet Loss Siamese network and SimCLR methods.*

Response. In both neural networks, once the convolutional layers of the encoders are adequately trained, they are connected to a single-layer dense classifier where the output is a binary value for whether the sample is part of a productive field. We agree with the reviewer this was not clear enough, requiring the reader to consult other documents and source code to obtain accurate information.

Changes. We agree that further clarification about the architectures of the networks employed in the deep learning methods is beneficial to this work. With that in mind, we have added Figure 5 (Revised manuscript – page 4) to provide details of both, SimCLR and Triplet Loss Siamese architectures and how the final classification is performed after training the encoder.

We have implemented the DL-based neural networks using 4 convolutional layers with a kernel size of 3 and stride of 1, as demonstrated in Figure 5.

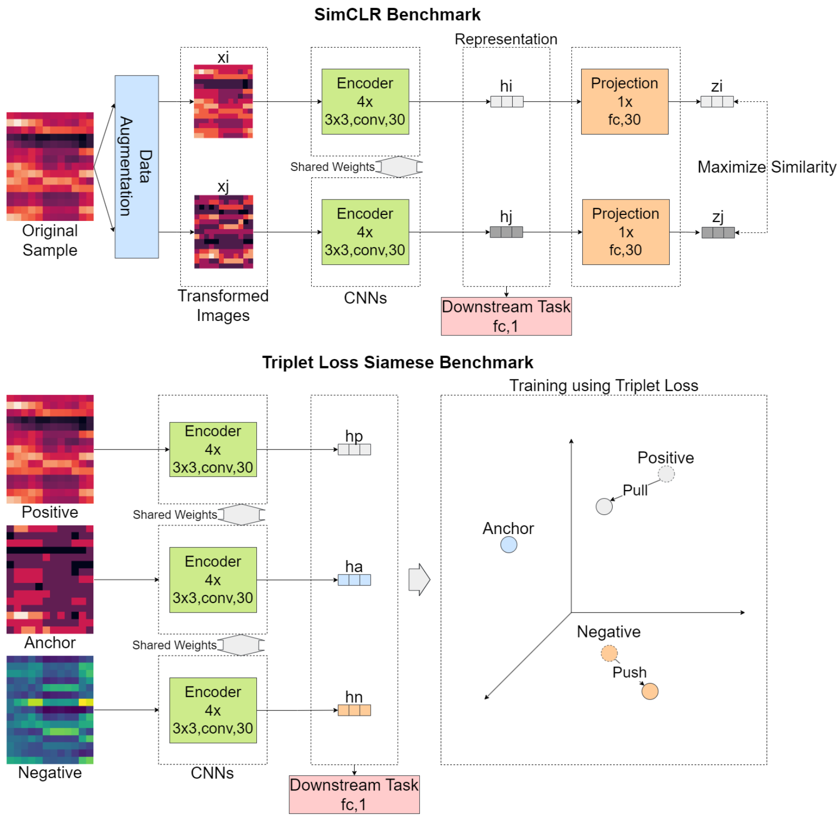


Figure 5: Neural network architectures employed in the benchmark testing

(Revised manuscript – page 4)

Reviewer's comment. *Issue 4: Since the authors did not use SegNet, UNet, or any other form of FCN, the following sentence does not provide valuable information to the reader: "In SegNet, only the pooling indices are transferred to the expansion path from the compression path, using less memory. In con-*

trast, in UNet, entire feature maps are transferred from the compression path to the expansion one, demanding much more memory. Both techniques attained remarkable performance."

Response. We understand your concern about the value and relevance of the provided information.

Changes. We acknowledge that in the context of the paper's specific content and the methods employed, this sentence may not contribute directly to the reader's understanding of the research findings. It could be seen as a digression from the main focus of the paper. This comment summed up to the extension of the related work section motivated us to remove these details but still keep this reference.

References

[1] Garcia-Pedrero, A., Gonzalo-Martín, C., Lillo-Saavedra, M., Rodríguez-Esparragón, D., 12 2018. The outlining of agricultural plots based on spatiotemporal consensus segmentation. Remote Sensing 10, 1991.

[2] Garcia-Pedrero, A., Lillo-Saavedra, M., Rodriguez-Esparragon, D., Gonzalo-Martin, C., 2019. Deep learning for automatic outlining agricultural parcels: Exploiting the land parcel identification system. IEEE Access 7, 158223–158236.

[3] García-Pedrero, A., Gonzalo-Martín, C., Lillo-Saavedra, M., 4 2017. A machine learning approach for agricultural parcel delineation through agglomerative segmentation. International Journal of Remote Sensing 38, 1809–1819.

[4] Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 03 2017. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci. Remote Sens. Lett. PP, 1–5.

[5] Seiffert, C., Khoshgoftaar, T. M., Hulse, J. V., Napolitano, A., 2010. Rusboost: A hybrid approach to alleviating class imbalance. IEEE Trans. Syst., Man, Cybern. A 40, 185–197.

[6] Wagner, M. P., Oppelt, N., 4 2020. Extracting agricultural fields from remote sensing imagery using graph-based growing contours. Remote Sensing 12, 1205.

IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 13, NO. 9, NOVEMBER 2023

Productive Crop Field Detection: A New Dataset and Deep Learning Benchmark Results

Eduardo Nascimento, John Just, Jurandy Almeida, and Tiago Almeida

***Abstract*—In precision agriculture, detecting productive crop fields is an essential practice that allows the farmer to evaluate operating performance separately and compare different seed varieties, pesticides, and fertilizers. ~~However, manually identifying productive fields is often a time-consuming, and error-prone task. Previous studies explore different methods to detect crop fields using advanced machine learning algorithms to bring significant improvements in both aspects, but they often lack good quality labeled data.~~ However, manually identifying productive fields is often time-consuming, costly, and subjective. Previous studies explore different methods to detect crop fields using advanced machine learning algorithms to support the specialists' decisions, but they often lack good quality labeled data. In this context, we propose a high-quality dataset generated by machine operation combined with Sentinel-2 images tracked over time. As far as we know, it is the first one to overcome the lack of labeled samples by using this technique. In sequence, we apply a semi-supervised classification of unlabeled data and state-of-the-art supervised and self-supervised deep learning methods to detect productive crop fields automatically. Finally, the results demonstrate high accuracy in Positive Unlabeled learning, which perfectly fits the problem where we have high confidence in the positive samples. Best performances have been found in Triplet Loss Siamese given the existence of an accurate dataset and Contrastive Learning considering situations where we do not have a comprehensive labeled dataset available.**

***Index Terms*—Crop field detection, precision agriculture, machine learning.**

I. Introduction

FOOD production needs to grow by 70% to meet the demands of the expected world population by 2050 [1]. Motivated by this challenge, agriculture has adopted technologies to improve and optimize input returns while preserving natural resources. Integrating these technologies promotes a farming management concept known as *precision agriculture* [2]. The main goal of precision agriculture is to provide tools for allowing the farmer to observe, measure, and respond to field variability in crops facilitating faster and better decisions. In addition, these techniques are generic enough to be applied to various crops, including but not limited to corn, soy, coffee, sugarcane, beans, and even pastures [3, 4].

To efficiently organize and manage large crops, farmers leverage remote sensing to divide their land into smaller observation units that this work will refer to as *agricultural* or *crop fields*. A field shape is designed based on topography and mechanization planning. For example, fields are built around contour farming and the ideal length to fill the wagon capacity in a sugarcane crop. The same logic is followed for grains in the Brazilian South and Southeast. Nonetheless, the most important variable is the harvester capacity in the Midwest, where the topography is often plain [5, 6, 7]. A *productive crop field* is an area consistently used for the cycle of growth and harvest of a crop, typically yearly but more often in some regions with favorable soil and weather.

Thanks to the availability of a massive amount of labeled data, the development of artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has allowed for acceleration and improvement in many areas of agriculture [8, 9, 10, 11, 12]. Despite all the advances, having an adequate amount of labeled data to train such methods can be costly and not always affordable for precision farming.

Usually, farmers rely on experts to build the field boundaries in dedicated Geographic Information Systems (GIS) software. Despite the expert's ability and knowledge, finding crop fields and drawing their boundaries has been a major, ~~error-prone~~ expensive, and time-consuming challenge [13]. The larger the customer's land is, the more significant the number of fields to be created. For instance, there are farms in Brazil with more than eighteen thousand fields that require regular updates to reflect their accurate states. Finding productive fields is even more challenging because it requires previous knowledge of these areas or the analysis of satellite images over time.

Aiming to facilitate access to data ready to support ML and DL models for productive crop field detection, we introduce a new dataset of productive fields based on agriculture machine's operation combined with Sentinel-2 images collected over a period of time. Firstly, this information is aggregated in geospatial L12 ~~hexes~~ hexagons, establishing a common ground for employing different satellite sources with different resolutions. Secondly, typical farmers are usually interested in monitoring productive areas rather than all possible fields since many are not fertile or suitable for any crop. For this purpose, the proposed dataset contains highly accurate positive samples (i.e., productive crop fields) and inferred negative ones, making it well-suited for positive and unlabeled learning [14]. Moreover, recent advances suggest that self-supervised methods (e.g., contrastive learning) may also provide a promising alternative even for cases where labeled samples are scarce [15]. Finally, we offer state-of-the-art ML and DL benchmark results for further comparison, including classification methods with different training strategies (i.e., supervised learning, self-supervised learning, and semi-

E. Nascimento, T. Almeida and J. Almeida are with the Department of Computer Science, Federal University of São Carlos, São Paulo, Brazil (e-mail: egnascimento@gmail.com; talmeida@ufscar.br; jurandy.almeida@ufscar.br)

J. Just is with the Department of Data Science and Analytics, John Deere, Iowa, USA (e-mail:johnpjust@gmail.com)

Manuscript received May 18, 2023; revised ~~September 17, 2014~~ July 6, 2023.

supervised classification of unlabeled data).

II. Related Work

Automatically detecting productive crop fields is a challenge that intersects multiple areas, and its solution could fill important gaps in agriculture management and environment control [7]. In this context, Crommelinck et al. [16] has worked on boundary detection of agricultural fields to provide automated recording of land rights, a slightly different purpose despite sharing common problems. However, despite the proven benefits of fully automated boundary detection and the research already accomplished, it is still considered an open problem with significant innovation opportunities, especially when sufficient training data is unavailable [12, 17].

North et al. [18] used classical computer vision techniques to detect the boundaries of farm fields in New Zealand. However, their technique resulted in only 59% accuracy, which is insufficient for most applications. Moreover, their strategy is highly impacted by variations caused by the season in the images extracted. Still based on classic computer vision, Evans et al. [19] used a region-based technique called canonically-guided region growing to segment multi-spectral images and, despite being more accurate than previous work, it was applied to a very small and manually annotated set of images requiring further investigation in a more broad collection of samples.

Since classical computer vision techniques alone are often insufficient to achieve good results, further research evaluated ML models to detect crop boundaries. For instance, García-Pedrero et al. [20] achieved an accuracy of 92% using an ensemble algorithm called RUSBoost [21] to merge super-pixels and group blocks of the image, which were part of the same field. This study demonstrated how ML is a promising alternative and a precursor to further development using DL [20]. In the same year, Kussul et al. [22] applied DL to segment and classify crop types. They demonstrated the advantages of deeper networks by comparing a Multilayer Perceptron (MLP) with a CNN, where the latter achieved higher accuracies. Later, Garcia-Pedrero et al. [8] applied the U-Net, a convolutional neural network (CNN) for semantic image segmentation, to perform the same task with improved performance [8, 20, 23].

~~In the same year, //citet kussul2017 applied DL to segment and classify crop types. They demonstrated the advantages of deeper networks by comparing a Multilayer Perceptron (MLP) with a CNN, where the latter achieved higher accuracies.~~ Concurrently, Persello et al. [9] designed a DL approach using SegNet, a CNN commonly used for semantic segmentation in Very High Resolution (VHR) images. ~~In SegNet, only the pooling indices are transferred to the expansion path from the compression path using less memory. In contrast, in U-Net, entire feature maps are transferred from the compression path to the expansion one, demanding much more memory. Both techniques attained remarkable performance. However,~~ In this work, some significant mistakes were generally associated with poor-quality training data [9]. The authors extend their work to medium-resolution images from Sentinel-2 ~~using two different DL architectures, SRC-Net and MD-FCN. Ultimately,~~ , however, they observed a severe limitation during training since ~~these~~ the employed methods comprised many convolutional layers resulting in a time-consuming task [9, 10].

Wagner and Oppelt [13] designed a modified version of the growing snakes active contour model based on graph theory concepts [13]. ~~This graph-based growing contour technique can extract complex boundary networks in agricultural landscapes requiring little supervision. Evaluated on rural area maps, it detected 99% of total acreage, but it achieved a poor performance close to urban areas.~~ In sequence, Wagner and Oppelt [24] integrated a DL approach into their already proven graph-based model. ~~For this, an MLP was used to make a pixel-wise prediction of whether it was a boundary.~~ The method achieved high results in rural areas but required further investigation of the missed boundaries, especially in areas closer to the cities [24, 13].

Waldner and Diakogiannis [11] compared DL methods, starting with ResUnet-a and followed by an improvement named Fractal-ResUNet, a network designed for semantic segmentation of agricultural images. ~~It employs hierarchical watershed segmentation, a method that relies on graph theory, suggesting this is a successful approach to address the challenge of segmenting cropland images.~~ The best accuracies were achieved with Sentinel-2 images, and it is considered state-of-the-art in boundary detection of agricultural fields. However, it depends on a large amount of training data with high-quality labeling to achieve consistent results [11, 12].

Most aforementioned studies were devised for detecting or segmenting all possible fields, regardless of whether they are productive. Only a few consider productive fields, but even these disregard their eventual changes over time.

III. Productive Fields Dataset

The dataset is stored in tabular format, where each row corresponds to a hexagon, along with a timestamp and satellite band values. The hexagon format was used for geo analysis since it facilitates storing and modeling spatiotemporal data from multiple disparate sources while avoiding using rasters. We used the H3 Python library[1] to create such hexagons with a level of L12. At this level, each hexagon has edges of approximately 10 meters, resulting in a hexagonal area that covers 307 square meters, slightly bigger than three Sentinel-2 pixels. This resolution offers a balanced trade-off between granularity and computational efficiency while training the models. Moreover, hexagons of this size can capture fine-grained variations in satellite data within a manageable dataset size, enabling efficient processing and analysis.

The timestamps are associated with the date and time the Sentinel-2 images were captured, providing temporal information. There are images from 2018-10-29 17:04:21 to 2019-08-15 16:59:01. Finally, the twelve band values correspond to the pixel median values in the area covered by that particular H3 hexagon. With this, we can summarize the satellite data for each hexagon in a compact and interpretable format, facilitating downstream ML analyses. Figure 1 illustrates how

[1]The H3 Python library is available at https://pypi.org/project/h3/. Access on May 4, 2023.

we create the samples used as the input for the learning models.

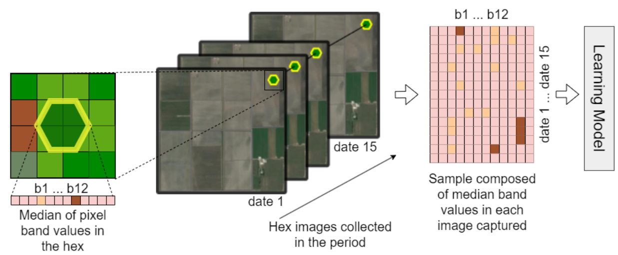


Fig. 1. Sequence to build a sample based upon the hexagon time-series

Table I summarizes the dataset and its fields. The dataset includes 17 productive corn crop fields in the US within the same geographical area. It has 106,735 rows, each representing a hexagon captured on a specific date and time, followed by the 12 Sentinel-2 band values. From these, 75,990 rows are labeled as positive (i.e., productive crop fields hexagons tracked over the period) and 30,745 as negative (i.e., non-productive crop fields hexagons tracked over the period). The amount of positive hexagons representing geolocations is 5,066, while the number of negative hexagons is 2,050 resulting in a slightly unbalanced dataset. The number of images for each hexagon collected on different dates varies from 15 to 38, and the averages for each field are represented in the last column.

TABLE I
PRODUCTIVE FIELDS DATASET DESCRIPTION

	Fields: 17				
	Hexes (samples)		Hex images collected in the period		
Field	positive	negative	positive	negative	average
F01	558	210	9,860	3,539	17.67 ±0.48
F02	734	151	14,242	2,863	19.40 ±0.74
F03	59	85	1,075	1,528	18.22 ±0.65
F04	350	150	8,326	3,518	23.79 ±0.48
F05	75	134	1,815	3,001	24.20 ±0.90
F06	288	60	4,999	1,044	17.36 ±0.68
F07	418	140	9,409	2,950	22.51 ±0.70
F08	425	154	8,584	3,084	20.20 ±0.82
F09	147	127	2,993	2,563	20.36 ±0.71
F10	369	176	6,973	3,194	18.90 ±0.40
F11	228	57	4,813	1,172	21.11 ±0.59
F12	91	96	1,701	1,738	18.69 ±0.53
F13	151	86	3,271	1,840	21.66 ±0.49
F14	275	216	4,871	3,825	17.71 ±0.46
F15	370	198	7,369	3,913	19.92 ±0.66
F16	271	25	5,030	413	18.56 ±0.69
F17	281	17	5,099	297	18.15 ±0.97

~~begin figure [!htb] Distribution of the number of samples by satellite image date. end figure~~

We have labeled each hexagon as positive or negative based on the productive agriculture machine operations, such as planting or harvesting. Specifically, hexagons containing evidence of these operations were labeled positive, while neighboring hexagons within a three-layer radius were labeled as negative, as shown in Figure 2. Labeling neighboring samples as negative is particularly interesting because they are often the hardest for prediction models to classify.

Fig. 2. Productive crop field filled with positively labeled ~~hexes~~ hexagons in green surrounded by a three ~~hexes~~ hexagons layer of inferred negative samples.

One noteworthy aspect of the proposed method is its ability to distinguish non-productive fields that share the same shape as productive ones. Even a highly skilled individual would struggle to make this distinction. ~~The key reason behind this lies in the trained model's ability to recognize patterns in the Normalized Difference Vegetation Index (NDVI) over time rather than relying solely on static images~~ The key reason behind this lies in the trained model's ability to recognize patterns of band values over time from non-productive areas as distinctly different than productive cropped areas. This approach leads to significantly more accurate results. Figure 3 shows an example of a non-productive field correctly identified.

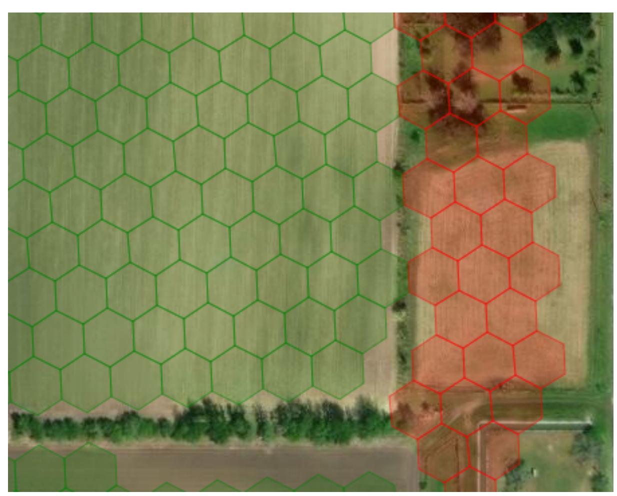


Fig. 3. A non-productive field properly classified with red ~~hexes~~ hexagons

Some false negatives may occur where productive fields are not detected due to the absence of agriculture machines equipped with satellite receptors or, for many reasons, when data are not shared. However, the dataset was manually curated to minimize the number of false negatives, specifically targeting the hexagonal regions between crop and non-crop fields. Despite these efforts, some false negatives may still be present. Even so, the proposed dataset is a valuable resource for future remote sensing research since, to our knowledge, it is the first to offer high-quality labeled productive crop fields.

IV. Benchmark Results

This section offers state-of-the-art ML and DL benchmark results to facilitate further comparison. We have analyzed the semi-supervised classification of unlabeled data and deep neural networks with different training strategies (i.e., supervised and self-supervised learning). We can use these models to predict previously unseen ~~hexes~~ hexagons or pixels based on the precision required for the output image. Both the dataset and implementation notebooks are publicly available at https://github.com/egnascimento/productivefieldsdetection.

A. Data processing

We processed and grouped the data to construct bidimensional time series samples. Each comprises 15 randomly selected image dates, as shown in Figure 4. A total of 8,342 grouped multitemporal time series samples resulted. Selecting sparse dates throughout the year is the best strategy for a comprehensive collection comprising different weather seasons and crop stages.

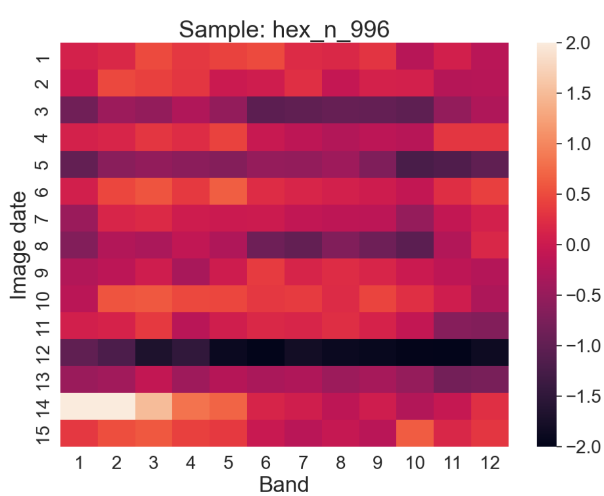


Fig. 4. Multitemporal time series sample visualization with 15 rows (i.e., each row corresponding to a unique date a satellite image was captured of the ~~hex~~ hexagon location) and 12 columns (i.e., satellite band values).

The original dataset was split into 80% training, 10% test, and 10% validation data keeping the balance between positive and negative labeled data.

B. Machine learning and deep learning methods

Based on the dataset's characteristics and training strategies, three state-of-the-art methods on ML and DL were chosen:

- *Positive Unlabeled (PU)* [14] is a semi-supervised classification approach of unlabeled data that is particularly suitable for the current study, given the availability of accurate positive data, the lack of negative one, and the ease of obtaining unlabeled data. PU learning relies on supervised methods, and we selected Support Vector Machines (SVM) for this evaluation.
- *Triplet Loss Siamese network* [25] is designed to learn representations that effectively discriminate similar and dissimilar samples in a latent space. The Triplet Loss function compares the distance between an anchor sample and a positive one against the distance between the anchor sample and a negative one.
- *Contrastive Learning* [26] is especially recommended in scenarios of limited data availability (e.g., regions with precision agriculture limited or nonexistent). Contrastive learning methods are employed to address the shortage of samples by employing data augmentation. This approach has been widely used to tackle diverse issues with significant potential for detecting crop fields. We have implemented the well-known SimCLR [27]. As Contrastive Learning is highly dependent on good data augmentation, a random jitter of up to $\pm 10\%$ was applied to all bands to create augmented samples.

We have implemented the DL-based neural networks using 4 convolutional layers with a kernel size of 3 and stride of 1, as demonstrated in Figure 5.

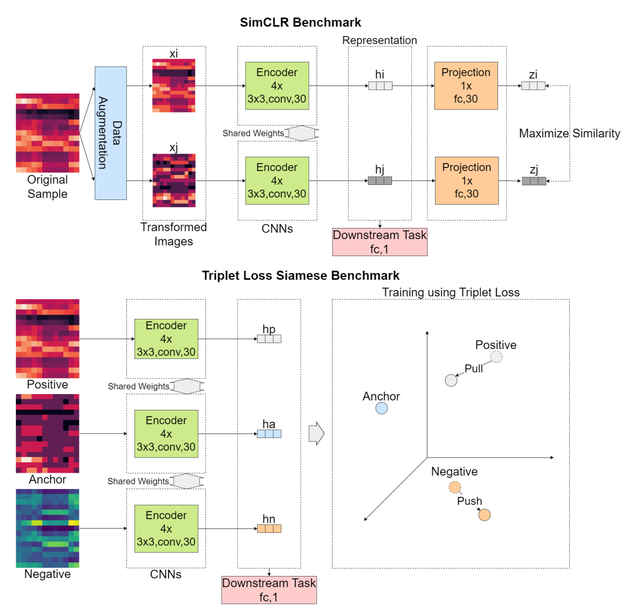


Fig. 5. Neural network architectures employed in the benchmark testing

C. Results

Table II presents the results, including metrics suitable for evaluating this problem [28]. The accuracies of PU Learning, Contrastive Learning, and Triplet Loss Siamese were 95.36%, 97.05%, and 97.47%, respectively. PU learning achieved accuracies slightly below DL methods, a remarkable performance that can still be significantly improved by changing the underlying method to a DL neural network. The F1 score weighted average was used since the dataset is imbalanced. Additionally, we included individual F1 scores by class to demonstrate that the best results came from the more reliable samples labeled by machine data. Finally, the analysis of the Matthews correlation coefficient (MCC), appropriate to imbalanced datasets, supported the conclusion that the Triplet Loss Siamese method achieved the best overall results.

TABLE II
BENCHMARK RESULTS

Models	Metrics				
	Accuracy	F1-weighted	F1-Neg	F1-Pos	MCC
DL Baseline	96.62%	96.61%	94.26%	97.61%	0.92
PU SVM	95.36%	95.36%	92.27%	96.68%	0.89
C. Learning	97.05%	97.03%	94.96%	97.91%	0.93
Triplet Loss	**97.47%**	**97.47%**	**95.77%**	**98.19%**	**0.94**

Analyzing the results, we have observed that all models missed, in most cases, the prediction of samples located between field and non-field areas, as presented in Figure 6. These mistakes, however, should not be relevant when performing classification at the pixel level.

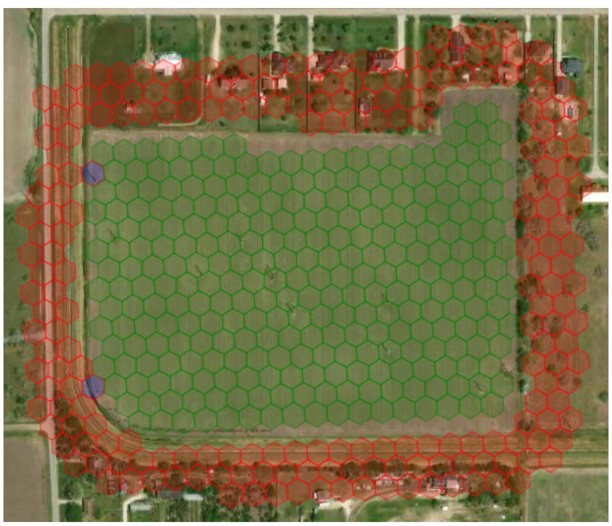


Fig. 6. Blue hexagons represent prediction mistakes, usually found between productive and non-field areas.

V. CONCLUSION

The dataset proposed in this paper represents an important alternative to training highly accurate prediction models that automatically detect productive crop fields. The main advantage is that it does not rely on manual work, which ensures, although not perfect, high-quality labeled samples in large volume since they are being continuously produced by machines equipped with precision agriculture devices. In sequence, we evaluated three different and complementary state-of-the-art ML and DL algorithms well-suited to the problem of detecting productive crop fields. The benchmark results present a performance high enough for most real applications involving detecting and delineating productive crop fields. For future work, we suggest collecting more data, including different crops. We also suggest improvements in the data augmentation when contrastive learning is used and evaluating an instance segmentation algorithm, possibly a graph-based one, which should be connected to the end of the pipeline to produce shape files ready-to-use by precision agriculture systems.

ACKNOWLEDGMENT

We would like to express our gratitude to John Deere for generously sharing the data. This partnership has reinforced the company's unwavering commitment to advancing scientific research while prioritizing the confidentiality and security of its customers' data. We deeply appreciate this collaboration and recognize the crucial role that technology companies like John Deere play in supporting scientific innovation.

REFERENCES

[1] G. Nelson, M. Rosegrant, A. Palazzo, I. Gray, C. Ingersoll, R. Robertson, S. Tokgoz, T. Zhu, T. Sulser, C. Ringler, S. Msangi, and L. You, "Food security, farming, and climate change to 2050: Scenarios, results, policy options," *Food Security, Farming, and Climate Change to 2050: Scenarios, Results, Policy Options*, 2 2010.

[2] N. Zhang, M. Wang, and N. Wang, "Precision agriculture - a worldwide overview," *Comput. Electron. Agric.*, vol. 36, pp. 113–132, 11 2002.

[3] D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps," *Biosystems Engineering*, vol. 114, pp. 358–371, 2013.

[4] I. Bhakta, S. Phadikar, and K. Majumder, "State-of-the-art technologies in precision agriculture: a systematic review," *Journal of the Science of Food and Agriculture*, vol. 99, pp. 4878–4888, 8 2019.

[5] M. Spekken, J. P. Molin, and T. L. Romanelli, "Cost of boundary manoeuvres in sugarcane production," *Biosystems Engineering*, vol. 129, pp. 112–126, 1 2015.

[6] L. M. Griffel, V. Vazhnik, D. S. Hartley, J. K. Hansen, and M. Roni, "Agricultural field shape descriptors as predictors of field efficiency for perennial grass harvesting: An empirical proof," *Comput. Electron. Agric.*, vol. 168, p. 105088, 2019.

[7] Édson Luis Bolfe, L. A. de Castro Jorge, I. D. Sanches, A. L. Júnior, C. C. da Costa, D. de Castro Victoria, R. Y. Inamasu, C. R. Grego, V. R. Ferreira, and A. R. Ramirez, "Precision and digital agriculture: Adoption of technologies and perception of brazilian farmers," *Agriculture (Switzerland)*, vol. 10, pp. 1–16, 12 2020.

[8] A. Garcia-Pedrero, M. Lillo-Saavedra, D. Rodriguez-Esparragon, and C. Gonzalo-Martin, "Deep learning for automatic outlining agricultural parcels: Exploiting the land parcel identification system," *IEEE Access*, vol. 7, pp. 158 223–158 236, 2019.

[9] C. Persello, V. A. Tolpekin, J. R. Bergado, and R. A. de By, "Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping," *Remote Sensing of Environment*, vol. 231, p. 111253, 9 2019.

[10] K. M. Masoud, C. Persello, and V. A. Tolpekin, "Delineation of agricultural field boundaries from sentinel-2 images using a novel super-resolution contour detector based on fully convolutional networks," *Remote Sensing*, vol. 12, 1 2020.

[11] F. Waldner and F. I. Diakogiannis, "Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network," *Remote Sensing of Environment*, vol. 245, 8 2020.

[12] F. Waldner, F. I. Diakogiannis, K. Batchelor, M. Ciccotosto-Camp, E. Cooper-Williams, C. Herrmann, G. Mata, and A. Toovey, "Detect, consolidate, delineate: Scalable mapping of field boundaries using satellite images," *Remote Sensing*, vol. 13, pp. 1–24, 6 2021.

[13] M. P. Wagner and N. Oppelt, "Extracting agricultural fields from remote sensing imagery using graph-based growing contours," *Remote Sensing*, vol. 12, p. 1205, 4 2020.

[14] F. Denis, R. Gilleron, and F. Letouzey, "Learning from positive and unlabeled examples," *Theoretical Computer Science*, vol. 348, no. 1, pp. 70–83, 2005, aLT 2000.

[15] R. Güldenring and L. Nalpantidis, "Self-supervised contrastive learning on agricultural images," *Computers and Electronics in Agriculture*, vol. 191, p. 106510, 12 2021.

[16] S. Crommelinck, M. Koeva, M. Y. Yang, and G. Vosselman, "Application of deep learning for delineation of visible cadastral boundaries from remote sensing imagery," *Remote Sensing*, vol. 11, p. 2505, 11 2019.

[17] R. Yang, Z. U. Ahmed, U. C. Schulthess, M. Kamal, and R. Rai, "Detecting functional field units from satellite images in smallholder farming systems using a deep learning based computer vision approach: A case study from bangladesh," *RSASE*, vol. 20, p. 100413, 2020.

[18] H. C. North, D. Pairman, and S. E. Belliss, "Boundary delineation of agricultural fields in multitemporal satellite imagery," *IEEE IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, pp. 237–251, 1 2019.

[19] C. Evans, R. Jones, I. Svalbe, and M. Berman, "Segmenting multispectral landsat tm images into field units," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, pp. 1054–1064, 2002.

[20] A. García-Pedrero, C. Gonzalo-Martín, and M. Lillo-Saavedra, "A machine learning approach for agricultural parcel delineation through agglomerative segmentation," *International Journal of Remote Sensing*, vol. 38, pp. 1809–1819, 4 2017.

[21] C. Seiffert, T. M. Khoshgoftaar, J. V. Hulse, and A. Napolitano, "Rusboost: A hybrid approach to alleviating class imbalance," *IEEE Trans. Syst., Man, Cybern. A*, vol. 40, pp. 185–197, 2010.

[22] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. PP, pp. 1–5, 03 2017.

[23] A. Garcia-Pedrero, C. Gonzalo-Martín, M. Lillo-Saavedra, and D. Rodríguez-Esparragón, "The outlining of agricultural plots based on spatiotemporal consensus segmentation," *Remote Sensing*, vol. 10, p. 1991, 12 2018.

[24] M. P. Wagner and N. Oppelt, "Deep learning and adaptive graph-based growing contours for agricultural field extraction," *Remote Sensing*, vol. 12, p. 1990, 6 2020.

[25] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 07-12-June, 2015, pp. 815–823.

[26] A. Jaiswal, A. R. Babu, M. Z. Zadeh, D. Banerjee, and F. Makedon, "A survey on contrastive self-supervised learning," *Technologies*, vol. 9, 2020.

[27] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," *37th ICML 2020*, vol. PartF168147-3, 2020.

[28] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, 2018.

IEEE Geoscience and Remote Sensing Letters (GRSL)

Cover Letter

Manuscript Title: Productive Crop Field Detection: A New Dataset and Deep Learning Benchmark Results

Authors' Names/Affiliations:

1.) Eduardo Garcia do Nascimento (**Corresponding author**)

Federal University of São Carlos (UFSCar)

Department of Computer Science

Rod. João Leme dos Santos, (SP-264), Km 110, Sorocaba, Sao Paulo, Brazil

2.) John Just

John Deere

Intelligent Solutions Group - Department of Data Science and Analytics

9505 Northpark Dr, Johnston, IA 50131, USA

3.) Jurandy Almeida

Federal University of São Carlos (UFSCar)

Department of Computer Science

Rod. João Leme dos Santos, (SP-264), Km 110, Sorocaba, Sao Paulo, Brazil

4.) Tiago A. Almeida

Federal University of São Carlos (UFSCar)

Department of Computer Science

Rod. João Leme dos Santos, (SP-264), Km 110, Sorocaba, Sao Paulo, Brazil

Relevance of this Manuscript to the IEEE Geoscience and Remote Sensing Letters

This paper introduces a new dataset of productive fields based on the agriculture machine's operation combined with Sentinel-2 images collected over a period. The proposed dataset contains highly accurate positive samples (i.e., productive crop fields) and inferred negative ones, making it well-suited for positive and unlabeled learning. Moreover, we offer state-of-the-art machine learning and deep learning benchmark results for further comparison, including classification methods with different training strategies (i.e., supervised learning, self-supervised learning, and semi-supervised classification of unlabeled data). Considering that typical farmers are usually interested in monitoring productive areas rather than all possible fields since many are not fertile or suitable for any crop, this paper is especially relevant to the IEEE Geoscience and Remote Sensing Letters because it fills an important gap by offering a new research avenue toward productive field detection. The presented dataset is accurate and publicly available, figuring as a valuable resource for promoting large-scale classification within the remote sensing community.

Issues relating to Journal Policies

There are no issues relating to this journal policies.

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article. The authors also declare that there is no conflict of interest with any reviewers.

Approval of the Manuscript

All authors read and approved this submitted manuscript.

Originality of the Manuscript

The authors inform that this manuscript represents original work that has not been published in any other journal or conference proceedings, and it is not currently being considered by another journal or conference.

Signature

Eduardo Garcia do Nascimento

(Corresponding author on behalf of all authors)

Submission date: May 18, 2023.

Revision date: July 6, 2023.

IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 13, NO. 9, NOVEMBER 2023

Productive Crop Field Detection: A New Dataset and Deep Learning Benchmark Results

Eduardo Nascimento, John Just, Jurandy Almeida, and Tiago Almeida

***Abstract*—In precision agriculture, detecting productive crop fields is an essential practice that allows the farmer to evaluate operating performance separately and compare different seed varieties, pesticides, and fertilizers. However, manually identifying productive fields is often time-consuming, costly, and subjective. Previous studies explore different methods to detect crop fields using advanced machine learning algorithms to support the specialists' decisions, but they often lack good quality labeled data. In this context, we propose a high-quality dataset generated by machine operation combined with Sentinel-2 images tracked over time. As far as we know, it is the first one to overcome the lack of labeled samples by using this technique. In sequence, we apply a semi-supervised classification of unlabeled data and state-of-the-art supervised and self-supervised deep learning methods to detect productive crop fields automatically. Finally, the results demonstrate high accuracy in Positive Unlabeled learning, which perfectly fits the problem where we have high confidence in the positive samples. Best performances have been found in Triplet Loss Siamese given the existence of an accurate dataset and Contrastive Learning considering situations where we do not have a comprehensive labeled dataset available.**

***Index Terms*—Crop field detection, precision agriculture, machine learning.**

I. Introduction

FOOD production needs to grow by 70% to meet the demands of the expected world population by 2050 [1]. Motivated by this challenge, agriculture has adopted technologies to improve and optimize input returns while preserving natural resources. Integrating these technologies promotes a farming management concept known as *precision agriculture* [2]. The main goal of precision agriculture is to provide tools for allowing the farmer to observe, measure, and respond to field variability in crops facilitating faster and better decisions. In addition, these techniques are generic enough to be applied to various crops, including but not limited to corn, soy, coffee, sugarcane, beans, and even pastures [3, 4].

To efficiently organize and manage large crops, farmers leverage remote sensing to divide their land into smaller observation units that this work will refer to as *agricultural* or *crop fields*. A field shape is designed based on topography and mechanization planning. For example, fields are built around contour farming and the ideal length to fill the wagon capacity in a sugarcane crop. The same logic is followed for grains in the Brazilian South and Southeast. Nonetheless, the most important variable is the harvester capacity in the Midwest, where the topography is often plain [5, 6, 7]. A *productive crop field* is an area consistently used for the cycle of growth and harvest of a crop, typically yearly but more often in some regions with favorable soil and weather.

Thanks to the availability of a massive amount of labeled data, the development of artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has allowed for acceleration and improvement in many areas of agriculture [8, 9, 10, 11, 12]. Despite all the advances, having an adequate amount of labeled data to train such methods can be costly and not always affordable for precision farming.

Usually, farmers rely on experts to build the field boundaries in dedicated Geographic Information Systems (GIS) software. Despite the expert's ability and knowledge, finding crop fields and drawing their boundaries has been a major, expensive, and time-consuming challenge [13]. The larger the customer's land is, the more significant the number of fields to be created. For instance, there are farms in Brazil with more than eighteen thousand fields that require regular updates to reflect their accurate states. Finding productive fields is even more challenging because it requires previous knowledge of these areas or the analysis of satellite images over time.

Aiming to facilitate access to data ready to support ML and DL models for productive crop field detection, we introduce a new dataset of productive fields based on agriculture machine's operation combined with Sentinel-2 images collected over a period of time. Firstly, this information is aggregated in geospatial L12 hexagons, establishing a common ground for employing different satellite sources with different resolutions. Secondly, typical farmers are usually interested in monitoring productive areas rather than all possible fields since many are not fertile or suitable for any crop. For this purpose, the proposed dataset contains highly accurate positive samples (i.e., productive crop fields) and inferred negative ones, making it well-suited for positive and unlabeled learning [14]. Moreover, recent advances suggest that self-supervised methods (e.g., contrastive learning) may also provide a promising alternative even for cases where labeled samples are scarce [15]. Finally, we offer state-of-the-art ML and DL benchmark results for further comparison, including classification methods with different training strategies (i.e., supervised learning, self-supervised learning, and semi-supervised classification of unlabeled data).

II. Related Work

Automatically detecting productive crop fields is a challenge that intersects multiple areas, and its solution could fill

E. Nascimento, T. Almeida and J. Almeida are with the Department of Computer Science, Federal University of São Carlos, São Paulo, Brazil (e-mail: egnascimento@gmail.com; talmeida@ufscar.br; jurandy.almeida@ufscar.br)

J. Just is with the Department of Data Science and Analytics, John Deere, Iowa, USA (e-mail:johnpjust@gmail.com)

Manuscript received May 18, 2023; revised July 6, 2023.

important gaps in agriculture management and environment control [7]. In this context, Crommelinck et al. [16] has worked on boundary detection of agricultural fields to provide automated recording of land rights, a slightly different purpose despite sharing common problems. However, despite the proven benefits of fully automated boundary detection and the research already accomplished, it is still considered an open problem with significant innovation opportunities, especially when sufficient training data is unavailable [12, 17].

North et al. [18] used classical computer vision techniques to detect the boundaries of farm fields in New Zealand. However, their technique resulted in only 59% accuracy, which is insufficient for most applications. Moreover, their strategy is highly impacted by variations caused by the season in the images extracted. Still based on classic computer vision, Evans et al. [19] used a region-based technique called canonically-guided region growing to segment multi-spectral images and, despite being more accurate than previous work, it was applied to a very small and manually annotated set of images requiring further investigation in a more broad collection of samples.

Since classical computer vision techniques alone are often insufficient to achieve good results, further research evaluated ML models to detect crop boundaries. For instance, García-Pedrero et al. [20] achieved an accuracy of 92% using an ensemble algorithm called RUSBoost [21] to merge super-pixels and group blocks of the image, which were part of the same field. This study demonstrated how ML is a promising alternative and a precursor to further development using DL [20]. In the same year, Kussul et al. [22] applied DL to segment and classify crop types. They demonstrated the advantages of deeper networks by comparing a Multilayer Perceptron (MLP) with a CNN, where the latter achieved higher accuracies. Later, Garcia-Pedrero et al. [8] applied the U-Net, a convolutional neural network (CNN) for semantic image segmentation, to perform the same task with improved performance [8, 20, 23].

Concurrently, Persello et al. [9] designed a DL approach using SegNet, a CNN commonly used for semantic segmentation in Very High Resolution (VHR) images. In this work, some significant mistakes were generally associated with poor-quality training data [9]. The authors extend their work to medium-resolution images from Sentinel-2 , however, they observed a severe limitation during training since the employed methods comprised many convolutional layers resulting in a time-consuming task [9, 10].

Wagner and Oppelt [13] designed a modified version of the growing snakes active contour model based on graph theory concepts [13]. In sequence, Wagner and Oppelt [24] integrated a DL approach into their already proven graph-based model. The method achieved high results in rural areas but required further investigation of the missed boundaries, especially in areas closer to the cities [24, 13].

Waldner and Diakogiannis [11] compared DL methods, starting with ResUnet-a and followed by an improvement named Fractal-ResUNet, a network designed for semantic segmentation of agricultural images. The best accuracies were achieved with Sentinel-2 images, and it is considered state-of-the-art in boundary detection of agricultural fields. However, it depends on a large amount of training data with high-quality labeling to achieve consistent results [11, 12].

Most aforementioned studies were devised for detecting or segmenting all possible fields, regardless of whether they are productive. Only a few consider productive fields, but even these disregard their eventual changes over time.

III. Productive Fields Dataset

The dataset is stored in tabular format, where each row corresponds to a hexagon, along with a timestamp and satellite band values. The hexagon format was used for geo analysis since it facilitates storing and modeling spatiotemporal data from multiple disparate sources while avoiding using rasters. We used the H3 Python library[1] to create such hexagons with a level of L12. At this level, each hexagon has edges of approximately 10 meters, resulting in a hexagonal area that covers 307 square meters, slightly bigger than three Sentinel-2 pixels. This resolution offers a balanced trade-off between granularity and computational efficiency while training the models. Moreover, hexagons of this size can capture fine-grained variations in satellite data within a manageable dataset size, enabling efficient processing and analysis.

The timestamps are associated with the date and time the Sentinel-2 images were captured, providing temporal information. There are images from 2018-10-29 17:04:21 to 2019-08-15 16:59:01. Finally, the twelve band values correspond to the pixel median values in the area covered by that particular H3 hexagon. With this, we can summarize the satellite data for each hexagon in a compact and interpretable format, facilitating downstream ML analyses. Figure 1 illustrates how we create the samples used as the input for the learning models.

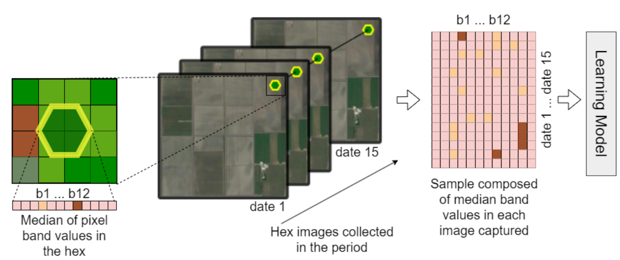


Fig. 1. Sequence to build a sample based upon the hexagon time-series

Table I summarizes the dataset and its fields. The dataset includes 17 productive corn crop fields in the US within the same geographical area. It has 106,735 rows, each representing a hexagon captured on a specific date and time, followed by the 12 Sentinel-2 band values. From these, 75,990 rows are labeled as positive (i.e., productive crop fields hexagons tracked over the period) and 30,745 as negative (i.e., non-productive crop fields hexagons tracked over the period). The amount of positive hexagons representing geolocations is 5,066, while the number of negative hexagons is 2,050 resulting in a slightly unbalanced dataset. The number of images for each hexagon

[1]The H3 Python library is available at https://pypi.org/project/h3/. Access on May 4, 2023.

collected on different dates varies from 15 to 38, and the averages for each field are represented in the last column.

TABLE I
PRODUCTIVE FIELDS DATASET DESCRIPTION

	Fields: 17				
	Hexes (samples)		Hex images collected in the period		
Field	positive	negative	positive	negative	average
F01	558	210	9,860	3,539	17.67 ±0.48
F02	734	151	14,242	2,863	19.40 ±0.74
F03	59	85	1,075	1,528	18.22 ±0.65
F04	350	150	8,326	3,518	23.79 ±0.48
F05	75	134	1,815	3,001	24.20 ±0.90
F06	288	60	4,999	1,044	17.36 ±0.68
F07	418	140	9,409	2,950	22.51 ±0.70
F08	425	154	8,584	3,084	20.20 ±0.82
F09	147	127	2,993	2,563	20.36 ±0.71
F10	369	176	6,973	3,194	18.90 ±0.40
F11	228	57	4,813	1,172	21.11 ±0.59
F12	91	96	1,701	1,738	18.69 ±0.53
F13	151	86	3,271	1,840	21.66 ±0.49
F14	275	216	4,871	3,825	17.71 ±0.46
F15	370	198	7,369	3,913	19.92 ±0.66
F16	271	25	5,030	413	18.56 ±0.69
F17	281	17	5,099	297	18.15 ±0.97

We have labeled each hexagon as positive or negative based on the productive agriculture machine operations, such as planting or harvesting. Specifically, hexagons containing evidence of these operations were labeled positive, while neighboring hexagons within a three-layer radius were labeled as negative, as shown in Figure 2. Labeling neighboring samples as negative is particularly interesting because they are often the hardest for prediction models to classify.

Fig. 2. Productive crop field filled with positively labeled hexagons in green surrounded by a three hexagons layer of inferred negative samples.

One noteworthy aspect of the proposed method is its ability to distinguish non-productive fields that share the same shape as productive ones. Even a highly skilled individual would struggle to make this distinction. The key reason behind this lies in the trained model's ability to recognize patterns of band values over time from non-productive areas as distinctly different than productive cropped areas. This approach leads to significantly more accurate results. Figure 3 shows an example of a non-productive field correctly identified.

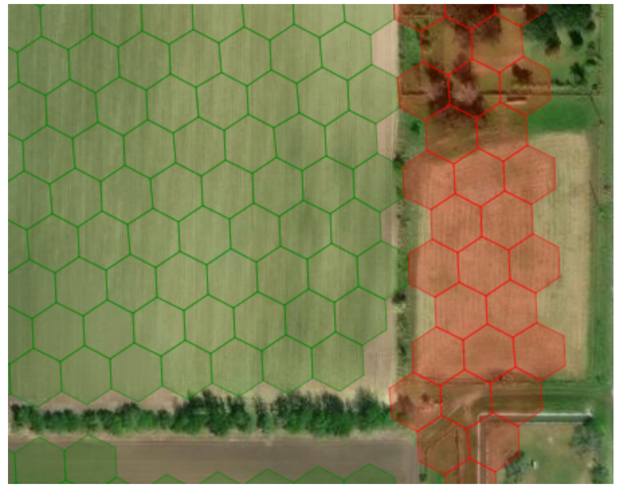


Fig. 3. A non-productive field properly classified with red hexagons

Some false negatives may occur where productive fields are not detected due to the absence of agriculture machines equipped with satellite receptors or, for many reasons, when data are not shared. However, the dataset was manually curated to minimize the number of false negatives, specifically targeting the hexagonal regions between crop and non-crop fields. Despite these efforts, some false negatives may still be present. Even so, the proposed dataset is a valuable resource for future remote sensing research since, to our knowledge, it is the first to offer high-quality labeled productive crop fields.

IV. BENCHMARK RESULTS

This section offers state-of-the-art ML and DL benchmark results to facilitate further comparison. We have analyzed the semi-supervised classification of unlabeled data and deep neural networks with different training strategies (i.e., supervised and self-supervised learning). We can use these models to predict previously unseen hexagons or pixels based on the precision required for the output image. Both the dataset and implementation notebooks are publicly available at https://github.com/egnascimento/productivefieldsdetection.

A. Data processing

We processed and grouped the data to construct bidimensional time series samples. Each comprises 15 randomly selected image dates, as shown in Figure 4. A total of 8,342 grouped multitemporal time series samples resulted. Selecting sparse dates throughout the year is the best strategy for a comprehensive collection comprising different weather seasons and crop stages.

The original dataset was split into 80% training, 10% test, and 10% validation data keeping the balance between positive and negative labeled data.

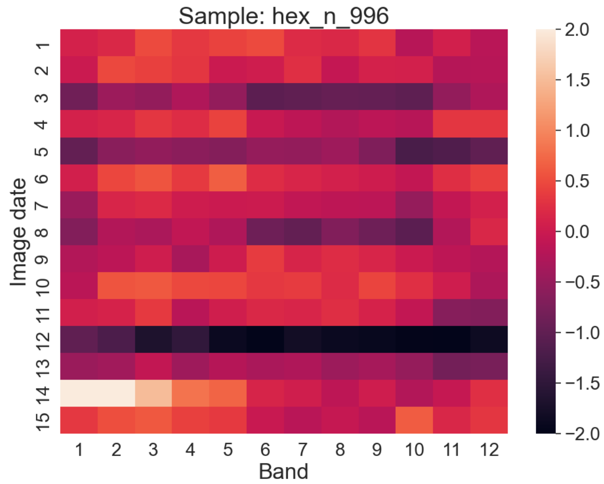


Fig. 4. Multitemporal time series sample visualization with 15 rows (i.e., each row corresponding to a unique date a satellite image was captured of the hexagon location) and 12 columns (i.e., satellite band values).

B. Machine learning and deep learning methods

Based on the dataset's characteristics and training strategies, three state-of-the-art methods on ML and DL were chosen:

- *Positive Unlabeled (PU)* [14] is a semi-supervised classification approach of unlabeled data that is particularly suitable for the current study, given the availability of accurate positive data, the lack of negative one, and the ease of obtaining unlabeled data. PU learning relies on supervised methods, and we selected Support Vector Machines (SVM) for this evaluation.
- *Triplet Loss Siamese network* [25] is designed to learn representations that effectively discriminate similar and dissimilar samples in a latent space. The Triplet Loss function compares the distance between an anchor sample and a positive one against the distance between the anchor sample and a negative one.
- *Contrastive Learning* [26] is especially recommended in scenarios of limited data availability (e.g., regions with precision agriculture limited or nonexistent). Contrastive learning methods are employed to address the shortage of samples by employing data augmentation. This approach has been widely used to tackle diverse issues with significant potential for detecting crop fields. We have implemented the well-known SimCLR [27]. As Contrastive Learning is highly dependent on good data augmentation, a random jitter of up to $\pm 10\%$ was applied to all bands to create augmented samples.

We have implemented the DL-based neural networks using 4 convolutional layers with a kernel size of 3 and stride of 1, as demonstrated in Figure 5.

C. Results

Table II presents the results, including metrics suitable for evaluating this problem [28]. The accuracies of PU Learning, Contrastive Learning, and Triplet Loss Siamese were 95.36%, 97.05%, and 97.47%, respectively. PU learning achieved accuracies slightly below DL methods, a remarkable performance that can still be significantly improved by changing the underlying method to a DL neural network. The F1 score weighted average was used since the dataset is imbalanced. Additionally, we included individual F1 scores by class to demonstrate that the best results came from the more reliable samples labeled by machine data. Finally, the analysis of the Matthews correlation coefficient (MCC), appropriate to imbalanced datasets, supported the conclusion that the Triplet Loss Siamese method achieved the best overall results.

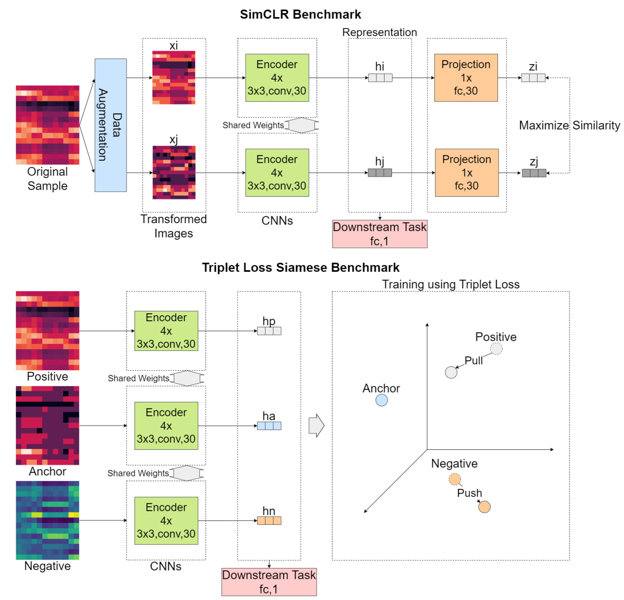


Fig. 5. Neural network architectures employed in the benchmark testing

TABLE II
BENCHMARK RESULTS

Models	Metrics				
	Accuracy	F1-weighted	F1-Neg	F1-Pos	MCC
DL Baseline	96.62%	96.61%	94.26%	97.61%	0.92
PU SVM	95.36%	95.36%	92.27%	96.68%	0.89
C. Learning	97.05%	97.03%	94.96%	97.91%	0.93
Triplet Loss	**97.47%**	**97.47%**	**95.77%**	**98.19%**	**0.94**

Analyzing the results, we have observed that all models missed, in most cases, the prediction of samples located between field and non-field areas, as presented in Figure 6. These mistakes, however, should not be relevant when performing classification at the pixel level.

V. Conclusion

The dataset proposed in this paper represents an important alternative to training highly accurate prediction models that automatically detect productive crop fields. The main advantage is that it does not rely on manual work, which ensures, although not perfect, high-quality labeled samples in large volume since they are being continuously produced by machines equipped with precision agriculture devices. In sequence, we evaluated three different and complementary state-of-the-art ML and DL algorithms well-suited to the problem

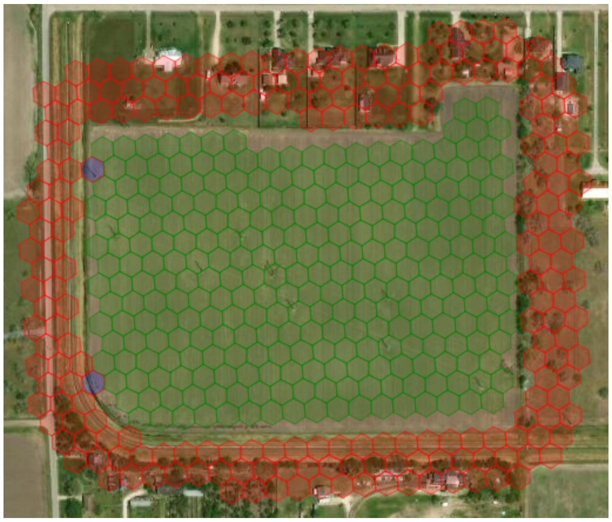


Fig. 6. Blue hexagons represent prediction mistakes, usually found between productive and non-field areas.

of detecting productive crop fields. The benchmark results present a performance high enough for most real applications involving detecting and delineating productive crop fields. For future work, we suggest collecting more data, including different crops. We also suggest improvements in the data augmentation when contrastive learning is used and evaluating an instance segmentation algorithm, possibly a graph-based one, which should be connected to the end of the pipeline to produce shape files ready-to-use by precision agriculture systems.

Acknowledgment

We would like to express our gratitude to John Deere for generously sharing the data. This partnership has reinforced the company's unwavering commitment to advancing scientific research while prioritizing the confidentiality and security of its customers' data. We deeply appreciate this collaboration and recognize the crucial role that technology companies like John Deere play in supporting scientific innovation.

References

[1] G. Nelson, M. Rosegrant, A. Palazzo, I. Gray, C. Ingersoll, R. Robertson, S. Tokgoz, T. Zhu, T. Sulser, C. Ringler, S. Msangi, and L. You, "Food security, farming, and climate change to 2050: Scenarios, results, policy options," *Food Security, Farming, and Climate Change to 2050: Scenarios, Results, Policy Options*, 2 2010.

[2] N. Zhang, M. Wang, and N. Wang, "Precision agriculture - a worldwide overview," *Comput. Electron. Agric.*, vol. 36, pp. 113–132, 11 2002.

[3] D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps," *Biosystems Engineering*, vol. 114, pp. 358–371, 2013.

[4] I. Bhakta, S. Phadikar, and K. Majumder, "State-of-the-art technologies in precision agriculture: a systematic review," *Journal of the Science of Food and Agriculture*, vol. 99, pp. 4878–4888, 8 2019.

[5] M. Spekken, J. P. Molin, and T. L. Romanelli, "Cost of boundary manoeuvres in sugarcane production," *Biosystems Engineering*, vol. 129, pp. 112–126, 1 2015.

[6] L. M. Griffel, V. Vazhnik, D. S. Hartley, J. K. Hansen, and M. Roni, "Agricultural field shape descriptors as predictors of field efficiency for perennial grass harvesting: An empirical proof," *Comput. Electron. Agric.*, vol. 168, p. 105088, 2019.

[7] Édson Luis Bolfe, L. A. de Castro Jorge, I. D. Sanches, A. L. Júnior, C. C. da Costa, D. de Castro Victoria, R. Y. Inamasu, C. R. Grego, V. R. Ferreira, and A. R. Ramirez, "Precision and digital agriculture: Adoption of technologies and perception of brazilian farmers," *Agriculture (Switzerland)*, vol. 10, pp. 1–16, 12 2020.

[8] A. Garcia-Pedrero, M. Lillo-Saavedra, D. Rodriguez-Esparragon, and C. Gonzalo-Martin, "Deep learning for automatic outlining agricultural parcels: Exploiting the land parcel identification system," *IEEE Access*, vol. 7, pp. 158 223–158 236, 2019.

[9] C. Persello, V. A. Tolpekin, J. R. Bergado, and R. A. de By, "Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping," *Remote Sensing of Environment*, vol. 231, p. 111253, 9 2019.

[10] K. M. Masoud, C. Persello, and V. A. Tolpekin, "Delineation of agricultural field boundaries from sentinel-2 images using a novel super-resolution contour detector based on fully convolutional networks," *Remote Sensing*, vol. 12, 1 2020.

[11] F. Waldner and F. I. Diakogiannis, "Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network," *Remote Sensing of Environment*, vol. 245, 8 2020.

[12] F. Waldner, F. I. Diakogiannis, K. Batchelor, M. Ciccotosto-Camp, E. Cooper-Williams, C. Herrmann, G. Mata, and A. Toovey, "Detect, consolidate, delineate: Scalable mapping of field boundaries using satellite images," *Remote Sensing*, vol. 13, pp. 1–24, 6 2021.

[13] M. P. Wagner and N. Oppelt, "Extracting agricultural fields from remote sensing imagery using graph-based growing contours," *Remote Sensing*, vol. 12, p. 1205, 4 2020.

[14] F. Denis, R. Gilleron, and F. Letouzey, "Learning from positive and unlabeled examples," *Theoretical Computer Science*, vol. 348, no. 1, pp. 70–83, 2005, aLT 2000.

[15] R. Güldenring and L. Nalpantidis, "Self-supervised contrastive learning on agricultural images," *Computers and Electronics in Agriculture*, vol. 191, p. 106510, 12 2021.

[16] S. Crommelinck, M. Koeva, M. Y. Yang, and G. Vosselman, "Application of deep learning for delineation of visible cadastral boundaries from remote sensing imagery," *Remote Sensing*, vol. 11, p. 2505, 11 2019.

[17] R. Yang, Z. U. Ahmed, U. C. Schulthess, M. Kamal, and R. Rai, "Detecting functional field units from satellite images in smallholder farming systems using a deep learning based computer vision approach: A case study from bangladesh," *RSASE*, vol. 20, p. 100413, 2020.

[18] H. C. North, D. Pairman, and S. E. Belliss, "Boundary delineation of agricultural fields in multitemporal satellite imagery," *IEEE IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, pp. 237–251, 1 2019.

[19] C. Evans, R. Jones, I. Svalbe, and M. Berman, "Segmenting multispectral landsat tm images into field units," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, pp. 1054–1064, 2002.

[20] A. García-Pedrero, C. Gonzalo-Martín, and M. Lillo-Saavedra, "A machine learning approach for agricultural parcel delineation through agglomerative segmentation," *International Journal of Remote Sensing*, vol. 38, pp. 1809–1819, 4 2017.

[21] C. Seiffert, T. M. Khoshgoftaar, J. V. Hulse, and A. Napolitano, "Rusboost: A hybrid approach to alleviating class imbalance," *IEEE Trans. Syst., Man, Cybern. A*, vol. 40, pp. 185–197, 2010.

[22] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. PP, pp. 1–5, 03 2017.

[23] A. Garcia-Pedrero, C. Gonzalo-Martín, M. Lillo-Saavedra, and D. Rodríguez-Esparragón, "The outlining of agricultural plots based on spatiotemporal consensus segmentation," *Remote Sensing*, vol. 10, p. 1991, 12 2018.

[24] M. P. Wagner and N. Oppelt, "Deep learning and adaptive graph-based growing contours for agricultural field extraction," *Remote Sensing*, vol. 12, p. 1990, 6 2020.

[25] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 07-12-June, 2015, pp. 815–823.

[26] A. Jaiswal, A. R. Babu, M. Z. Zadeh, D. Banerjee, and F. Makedon, "A survey on contrastive self-supervised learning," *Technologies*, vol. 9, 2020.

[27] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," *37th ICML 2020*, vol. PartF168147-3, 2020.

[28] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, 2018.