Agricultural Field Extraction with Deep Learning Algorithm and Satellite Imagery

Article *in* Journal of the Indian Society of Remote Sensing · January 2022

DOI: 10.1007/s12524-021-01475-7

4 authors, including:

Alireza Sharifi
Shahid Rajaee University
76 PUBLICATIONS **1,217** CITATIONS
SEE PROFILE

Hadi Mahdipour
Spain
29 PUBLICATIONS **85** CITATIONS
SEE PROFILE

Aqil Tariq
Mississippi State University
127 PUBLICATIONS **2,319** CITATIONS
SEE PROFILE

Journal of the Indian Society of Remote Sensing
https://doi.org/10.1007/s12524-021-01475-7

RESEARCH ARTICLE

Agricultural Field Extraction with Deep Learning Algorithm and Satellite Imagery

Alireza Sharifi[1] · Hadi Mahdipour[2] · Elahe Moradi[1] · Aqil Tariq[3]

Received: 19 May 2021 / Accepted: 2 December 2021

Abstract
Automatic detection of borders using remote sensing images will minimize the dependency on time-consuming manual input. The lack of field border data sets indicates that current methods are ineffective. This article seeks to promote the detection of field borders from satellite images with general process based on a multi-task segmentation model. ResUNet-a is a convolutional neural network with a completely linked UNet backbone that supports sprawling and conditional inference. The algorithm will significantly increase model efficiency and its generalization by re-constructing connected outputs. Then individual field segmentation can be accomplished by post-processing model outputs. The model was extremely exact in field mapping, field borders, and thus individual fields using the Sentinel-2 and Landsat-8 images as inputs. The multitemporal images replacement with a single image similar to the composition time decreased slightly. The proposed model is able to reliably identify field borders and remove irrelevant limits from the image to acquire complex hierarchical contextual properties, thus outstriking classical edge filters. Our method is supposed to promote individual crop field extraction on a scale, by minimizing overfitting.

Keywords Convolutional neural networks · Edge detection · Field borders · Sentinel-2 · Remote sensing

Introduction

Many of the automated agriculture promises support farmers in monitoring agricultural fields during the season of growth. With specific field border, a requirement for field measurement has become essential and farmers are often required to provide precise digital records of their limits while they are signed up with the service provider (Sharifi, 2018). This procedure is mainly manual and time consuming and discourages people from doing so. Crop yield forecasting and food safety control are also used for predicting areas of crop yields through Earth Observation programs (Paudel et al., 2021). The recurring extraction of field borders through broad areas will benefit greatly in such applications. Automation not only promotes the inclusion of farmers and thus encourages more extensive use of automated agricultural services, but also allows for the provision by means of remote sensing (Tariq et al., 2021). Several techniques were developed to extract field borders from satellite images, which regularly and globally cover high-resolution cropping areas. These methods are divided into three models: edge-based models (Khan et al., 2019), region-based models (Meher et al., 2019), and hybrid models (Shi et al., 2020).

Edge-based methods are dependent on filters in an epoch in which pixel values shift quickly to detect discontinuities. Each filter describes a certain kernel that is converted into emphasizing edges with an input image (Scharr, Sobel and Canny operator are typical examples) (Kumar, Afzal, and Afzal 2020). There are a variety of problems while dealing with edge operators because their susceptibility in high frequency sometimes produces misrepresentatives and their parameterization is subjective and important to unique

✉ Alireza Sharifi
a_sharifi@sru.ac.ir

[1] Department of Surveying Engineering, Faculty of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran

[2] Chief Innovation Office, Sinenta Corp., La Cañada, 04120 Almeria, Spain

[3] State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China

Published online: 16 January 2022

circumstances (Rabbi et al., 2020). Thresholds that are post-processed and adapted locally can solve problems that resulting in better established, closed borders. Region-based methods are groups of adjacent entity pixels' dependent on a certain criteria of homogeneity. The quest for optimum segmentation for regional methods remains a test and error procedure that is expected to produce inappropriate results. For example, objects with low parameterization might cease expanding, creating sliver polygons and moving the extracted borders into them, before hitting actual borders (Burdick et al., 2018). Methods dependent on region often aim to oversegment areas with strong internal variability and slight neighboring undersegment areas (IdBenIdder & Laachfoubi, 2015; Sharifi, 2020). Any of these adverse reactions can be mitigated by oversegmenting images intentionally and determining whether neighboring features are fused into machine learning (Park et al., 2021). While edge-based and region-based methods are available, the user group seems to have no use of these methods and suggests a lack of fitness to do so. For example, from crowdsourced, manually digitized polygons, the only global field map is found (Ciobanu et al., 2019; S. Wang et al., 2019).

Multitemporal image features appear redundant when the borders of fields have been extracted from well-targeted one-date images in certain situations. In addition, it is impossible to produce a constant sequence of times in places such as the tropics because of continuous cloud coverage (Sudmanns et al., 2020). While multitemporal data would likely improve their accuracy, particularly in highly dynamic systems, we believe that large-scale field border extraction and uptake are needed for models with minimal preprocessing and parameter setting. Deep neural networks have new ways of extracting field borders because they need no designed features and since their design is highly adaptable to new problems (Mohammadi & Sharifi, 2021). Convolutional neural networks (CNNs) are used more and more in image processing so they can use hierarchical characteristics, from local to global images (Krizhevsky et al., 2017). Their depth is based on a deep network that uses convolutional operations. While filters are hand-made in an edge-based technique, these filters can learn from CNNs. These networks were originally developed for natural images and have been adapted for use in remote sensing applications such as road extraction (Gao et al., 2019), cloud detection (Chai et al., 2019), crop identification (Wu et al., 2021), and river and water body extraction (Guo et al., 2020). Thus, CNNs appear to be particularly well-suited for extracting field borders based on their size, though this has not been empirically demonstrated.

The objective of this study is to use ResUNet-a, for field border extraction. We also implemented the post-processing methods, because CNNs can produce discontinuous borders, that utilize their outputs to produce better borders and find different field. By using a composite of Sentinel-2 images with geometrical accuracy, we carried out a series of experiments that show that overfitting has been minimized without re-calibration to be implemented in a variety of conditions.

Method and Materials

Field borders are extracted by labeling each pixel with these classes: 'border' or 'not 'border'. If groups of interest are predictable, then training signals for similar learning tasks that could contribute to the accuracy of the initial task are ignored. Similar tasks can help a model better generalize the initial task. In the last layer of the architecture, these can also be combined to further restrict the inference of the initially task stabilize graduated changes and thereby increase model accuracy instead of forecasting all associated tasks concurrently and independently (Kosari et al., 2020). Therefore, field borders extraction was developed as a semantic segmentation problem in which several class labels are predicted. The four related tasks are to map areas, define borders, approximate the distance to a closest border and reconstruct images in the fields. Extracting field borders entails classifying each pixel in a multispectral image into one of two categories: "border" or "not border." Although single-tasking can produce satisfactory results, it lacks training signals from similar learning tasks (Waldner & Diakogiannis, 2020).

Model Architecture

The encoder compresses the contents of the knowledge of an arbitrarily large image. The decoder increases the encoded functionality progressively up to the main resolution and locates the categories of interest specifically. Two simple architectures have been introduced that vary in the encoder–decoder's layers. In the encoder there are six rest blocks in the ResUNet-a D6, and then in the PSP Pooling layer, then in the encoder there are seven building blocks in the ResUNet-a D7.

Border Extraction

Both make use of multiple semantic segmentation outputs and are data-driven, enabling the configuration of a small sample of reference data to be automated. In the watershed process, the three output masks are segmented using a seeded watershed segmentation algorithm. When catchment basins hit other catchments, they stop growing; their borders are lines that divide neighboring catchment basins.

By comparison, the centers of each field (seed) can be identified and developed prior to them colliding with the borders of other fields or pixels that do not belong to the field (background). The seed is identified by the distance mask's threshold, and the topographical surface is identified by the extent mask's threshold.

The post-processing system thresholds are installed in two stages: the first stage optimizes the scope of fields collected (the scale mask's threshold); the second stage optimizes the form and size of fields (thresholds on the border and distance masks). This technique therefore needs reference dataset for optimal threshold values to be available. In the first place, the degree threshold is defined by the maximization of the correlation coefficient of the Matthew (Ghaderizadeh et al., 2021). The MCC is calculated as below.

$$\mathrm{MCC} = \frac{\mathrm{TP.TN} - \mathrm{FP.FN}}{\sqrt{(\mathrm{TP}+\mathrm{FN})(\mathrm{TP}+\mathrm{FP})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{1}$$

where true positive, true negative, false positive, and false negative rates are denoted by TP, TN, FP, and FN, respectively. MCC values range from -1 (complete disagreement) to + 1 (complete agreement), with 0 indicating that the matter is correct.

Second, border and distance mask borders are optimized together to reduce the inappropriate separation into smaller (oversegmentation) of larger objects and an improper aggregation into larger neighboring small objects (undersegmentation). In the reference fields (T) and in the extracted field (E) data using Eqs. (2) and (3) (Wang & Wang, 2012):

$$S_{\mathrm{over-seg}} = 1 - \frac{|T_i \cap E_j|}{|T_i|} \tag{2}$$

$$S_{\mathrm{under-seg}} = 1 - \frac{|T_i \cap E_j|}{|E_j|} \tag{3}$$

Dataset

We have chosen Sentinel-2 images, which acquired from some agricultural fields from Sarab County in East Azerbaijan Province, Iran. This dataset increases the probability of cloud-free images acquired while also data which support multi-temporal composites with 10-m resolution. The sentinel- 2 images are given for large-scale applications free of charge. Additionally, we obtained Landsat-8 images in the visible and near-infrared bands that were cloud-free. Preprocessing consisted of two steps: subtraction of the blue, green, red, and near infrared bands from the satellite images and standardization of pixel values to ensure that each band had a mean of zero and a standard deviation of one. Standardization is important because the input variable size of the neural network is responsive and gradient optimization approaches (Mancino et al., 2020).

Also, field borders were generated by digitizing manually all fields across the study area using google earth images. Field polygons were rasterized at 10 m, matching the grid of the Sentinel-2 and supplying a collection of reference pixels for validation. Pixels of "non-field" have been set to 0. Each image examined (representing various study areas) was divided into a training and testing area in a 3:2 ratio. Due to the limited GPU memory, each Sentinel-2 tile has been partitioned into a series of smaller images (256,256 pixels) as input images. There was only one preprocessing step: standardizing pixel values to have a mean of zero and a standard deviation of one for each band. Standardization is required because neural networks are sensitive to the scale of their input variables and gradient optimization methods converge faster when features have a mean of zero and a variance of one.

Evaluation Method

A frequent problem in classification is determining the appropriate threshold for class assignment given a continuous classifier. The threshold value selected is highly dependent on the user requirements and their tolerance for false positives and negatives (Mahdipour et al., 2016). The Receiver Operating Characteristic (ROC) curve is frequently used to assess a continuous classifier's ability to correctly identify a binary array, in this case a map of "boundary" and "not boundary" pixels. The true positive rate (the fraction of mapped boundary pixels that are correctly classified as "boundary") is plotted against the false positive rate for a range of classifier threshold values (the fraction of mapped not boundary pixels that are incorrectly classified as "boundary"). As the threshold is lowered from a value where all pixels are classified as "not boundary" to one where all pixels are classified as "boundary," a good classifier will identify true positives more quickly than it accepts false positives. The ROC curve plots the true positive rate against the false positive rate, with more accurate classifiers resulting in a curve that is closer to the plot's upper left corner (Mahdipour et al., 2020). Thus, the area under the curve can be used to quantify a classifier's overall performance (AUC). The ROC plots as a straight line between (0,0) and (1,1) with an AUC of 0.5 for a random classification surface where each pixel has a 50% chance of being classified as boundary or not boundary. The AUC of a classifier is expected to be between 0.5 and 1.0. (Fig. 1).

The performance of various methods was evaluated using two widely used metrics, namely the overall accuracy

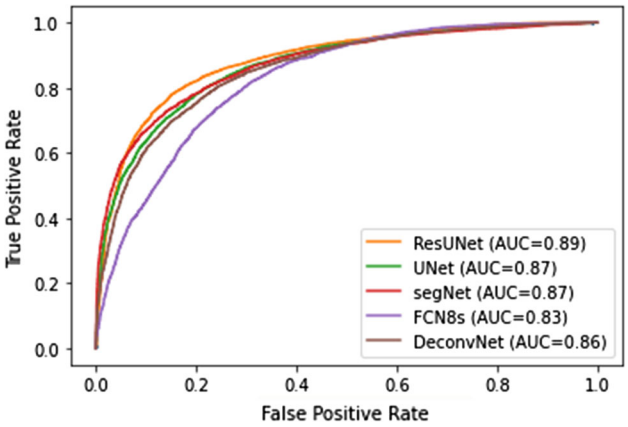

Fig. 1 The overall performance of models by the area under the curve (AUC)

(OA) and the mean average precision (mAP). For comparisons, four cutting-edge FCN models were used: SegNet (Badrinarayanan, Kendall, and Cipolla 2017), DeconvNet (Hyeonwoo Noh, Hong, et al., 2015; Noh, Seunghoon, et al., 2015), FCN8s (Shelhamer et al., 2017), and U-Net (Ronneberger et al., 2015a). These methods were chosen because they have all been demonstrated to be effective in semantic labeling for remote sensing images and are all open source and simple to implement.

Result and Discussion

The model was implemented on Python platform using a computer with a 1.2 GHz quad-core CPU, 6 GB of memory, and a Linux 4.4 based Raspbian operating system. We trained ResUNet-a D6 and D7 models (it took 73 s); We trained ResUnet-a D6 and D7 models using Adam as the optimizer. Adam calculates individual adaptive learning rates for different parameters using estimates of the gradient's first and second moments. Adam achieves faster convergence than other alternatives, as demonstrated empirically (Kingma & Ba, 2015). The weight decay (WD) approach incorporates a weight decay parameter into the learning rate in order to incrementally decrease the weight and bias values of the neural network during each training iteration. This technique is equivalent to adding an L2 regularization term to the loss function and can accelerate convergence. For 200 epochs, we trained models with different weight decay parameters (10^{-4}, 10^{-5}, and 10^{-6}.). For 200 epochs, the network is trained. We chose this large number of epochs to ensure that each network performs optimally and to save the optimal set of weights for each network to avoid overfitting. The first three iterations were parameterized with weight decay (WD) values, while the final iteration was configured interactively. After each epoch, the loss function and MCC were calculated for the test set. We concluded that it was unlikely that models would increase accuracy over more than 200 epochs because training curves showed overfit signs in seventy epochs.

Ground truth map were created using digitizing of Google Earth images. The maps were then converted to a binary file including the farm border (one) and the foreground (zero). Finally, using evaluation metrics, comparisons were made and numerical results were obtained according to Table 1. The average accuracy is 85% and the MCC is 74%. The cropland grade was marginally lower than the non-cropland grade by 85 percent. Tuning the map scale threshold to optimize MCC only resulted in small deviations in comparison with a 50% defect threshold. However, we maintained the optimized threshold because we tried to reach the grading with the most balanced accuracy for croplands and non-croplands. Also, the results suggest that temporal data must also be taken into account in certain situations.

As can be seen in Table 2, our model defined borders with considerably higher sharpness and less noise than an edge-based benchmark tool (more than 0.86 for hit rates). This shows, by using hierarchical contextual knowledge, that convolutional neural networks reduce the value of temporal information. Current neural networks are invariant of size. The model functions are object-based, which means they don't depend on pixel values but rather on the fact that they all belong to the same object. This process is not subject to a specific resolution.

We contrasted the mask found on the borders of the Canny filter with our ResUNet-a (Fig. 2). The edge-based approach has generated slightly less and noisier interfaces. Inside pixel values were by average higher with traditional rim detection and wider than those achieved with ResUNet-a. The retracted edges become clearer as the neural network learns to be adaptive to specific edge shapes. As compared to a monthly composite, feeding the method with single data has no effect on the model's performance. For example, oversegmentation and undersegmentation

Table 1 Pixel-based assessment of proposed model for the single image with different thresholds, multitemporal images, and resampled Sentinel-2 image with specific thresholds

Threshold	OA	MCC	F_C	F_{NC}
Single (50%)	85.40	74.12	78.34	87.15
Single (35%)	86.19	73.75	78.52	86.83
Multitemporal (35%)	78.21	69.09	78.96	81.59
Resamples 30-m (35%)	75.83	71.33	73.41	79.20

Table 2 Object-based assessment of proposed model for the Sentinel-2 image, resampled Sentinel-2 image, and Landsat-8 image

Image	Sentinel-2	Sentinel-2 (30 m)	Landsat-8
Hit rate	0.952	0.879	0.865
Oversegmentation	0.893	0.797	0.781
Undersegmentation	0.897	0.795	0.778
Eccentricity	0.941	0.915	0.903
Shift (m)	5	7	8

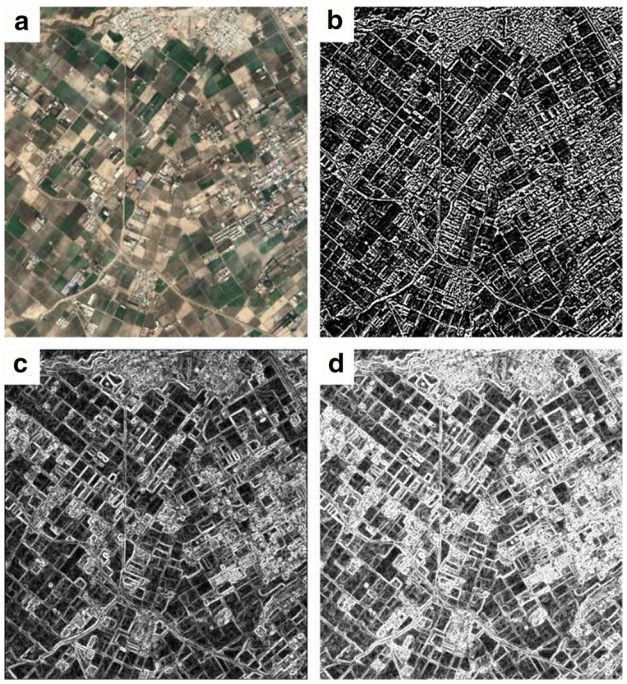

Fig. 2 (**a**) Agricultural fields in the study of area and comparison of the field borders extracted using (**b**) proposed method, (**c**) references borders, and (**d**) Canny filter

premiums fell, while the compensation remained unchanged.

The quantitative results obtained using various methods are summarized in Table 3. ResUNet produced the best result, with an overall accuracy of 85.60 percent and a mean AP score of 88.05 percent, when compared to all other methods utilizing encoder-decoder network architectures. SegNet outperformed the ResUNet model in terms of accuracy. ResUNet and SegNet are demonstrated to perform admirably. UNet outperforms DeconvNet and FCN8s in terms of stable classification performance across a variety of scenes.

Table 3 The quantitative results using the deep learning models

Model	mAP	OA
FCN8s	63.88	75.09
DeconvNet	65.22	77.87
UNet	69.21	81.11
SegNet	72.04	83.80
ResUNet	75.62	85.60

The statistically relevant variances were only for the sub-segmentation rate, which means that it is a valid alternative to derive field borders from single dated images. The model has been trained on Sentinel-2 images and attained a high degree of pixel and field accuracy. The same model was correctly created, without retraining. The same model was acquired by Sentinel-2 and Landsat-8 for single images of the same site. In areas where cloud coverage is continuous as in the tropics, the ability to remove borders from a one-date image is cost-efficient. It shows that by exploiting contextual knowledge at multiple levels, convolutional neural networks minimize the impact of spectral and temporal information.

Conclusion

This study demonstrates that solving the issue of the field border extraction of satellite images with a convolutional neural network as multiple semantic segmenting tasks delivers excellent efficiency at pixel and object stage. ResUNet-a as to extract field borders from rougher artifacts can be associated, but this needs to be confirmed experimentally, through the use of several convolutions to classify features on various scales. We observed a reduced hit rate and geometric accuracy up to 11% while we were applying the model to resampled Sentinel-2 to 30-m and Landsat-8 images. The main reason was that smaller fields in the 30-m image were not resolved. Differences between the reference resolution and extracted maps may lead to artificial distortions in the precise evaluation. The difference between the results of resampled Sentinel-2 to 30-m and Landsat-8 images is lack of accuracy is partially due to differences in spectral and spatial responses in these satellite images. The results showed that semantic segmentation in combination with multitasking is used to increasing the potential to recover individual areas from satellite image. One of the main limitation of this study was the computational density of the proposed model, which requires significantly more FLOPS than similar models EfficientNets. ResNets are typically run on GPUs and they are computationally heavy. Also, the maximum pixel-to-spatial level accuracy of segmentation which our proposed model can achieve was one of our limitation. Many suppliers of automated agriculture services urgently require the ability to automatically detect field borders from remote sensing images. In order to forecast the likelihood for any pixel to belong to the field, to the border of the field and to

predict the distance to the closest border, our model was based on multitasking and conditioning. These projections were then post-processed such that individual fields were segmented and extracted. The proposed method demonstrated suitable efficiency in field border detection and strong generalization skills over time, space and sensors.

Declarations

Conflict of interest The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

1. Alireza Sharifi
2. Hadi Mahdipour
3. Elahe Moradi
4. Aqil Tariq

References

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(12), 2481–2495.

Burdick, J., Marques, O., Weinthal, J., & Furht, B. (2018). Rethinking skin lesion segmentation in a convolutional classifier. *Journal of Digital Imaging*. https://doi.org/10.1007/s10278-017-0026-y

Chai, D., Newsam, S., Zhang, H. K., Qiu, Y., & Huang, J. (2019). Cloud and cloud shadow detection in landsat imagery based on deep convolutional neural networks. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2019.03.007

Ciobanu, R. I., Negru, C., Pop, F., Dobre, C., Mavromoustakis, C. X., & Mastorakis, G. (2019). Drop computing: Ad-Hoc dynamic collaborative computing. *Future Generation Computer Systems*. https://doi.org/10.1016/j.future.2017.11.044

Gao, L., Song, W., Dai, J., & Chen, Y. (2019). Road extraction from high-resolution remote sensing imagery using refined deep residual convolutional neural network. *Remote Sensing*. https://doi.org/10.3390/rs11050552

Ghaderizadeh, S., Abbasi-Moghadam, D., Sharifi, A., Zhao, Na., & Tariq, A. (2021). Hyperspectral image classification using a hybrid 3D–2D convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2021.3099118

Guo, H., He, G., Jiang, W., Yin, R., Yan, L., & Leng, W. (2020). A Multi-scale water extraction convolutional neural network (MWEN) method for GaoFen-1 remote sensing images. *ISPRS International Journal of Geo-Information*. https://doi.org/10.3390/ijgi9040189

Hadi, M., Khademi, M., & Yazdi, H. S. (2016). Efficient land-cover segmentation using meta fusion. *Journal of Information Systems and Telecommunication*. https://doi.org/10.7508/jist.2016.03.003

Hadi, M., Shabanian, M., & Kazerouni, I. A. (2020). Vectorized kernel-based fuzzy C-means: A method to apply KFCM on crisp and non-crisp numbers. *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*. https://doi.org/10.1142/S0218488520500270

IdBenIdder, H., & Laachfoubi, N. (2015). Unsupervised multi-level thresholding method for weather satellite cloud segmentation. *International Journal of Computer Applications*. https://doi.org/10.5120/20826-3544

Khan, S., Lee, D. H., Khan, M. A., Gilal, A. R., & Mujtaba, G. (2019). Efficient edge-based image interpolation method using neighboring slope information. *IEEE Access*. https://doi.org/10.1109/ACCESS.2019.2942004

Kingma, D.P., and Ba, J. (2015). "Adam: A method for stochastic optimization." *CoRR* abs/1412.6980.

Kosari, A., Sharifi, A., Ahmadi, A., & Khoshsima, M. (2020). Remote sensing satellite's attitude control system: Rapid performance sizing for passive scan imaging mode. *Aircraft Engineering and Aerospace Technology, 92*(7), 1073–1083. https://doi.org/10.1108/AEAT-02-2020-0030

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*. https://doi.org/10.1145/3065386

Kumar, L., Afzal, M. S., & Afzal, M. M. (2020). Mapping shoreline change using machine learning: A case study from the Eastern Indian Coast. *Acta Geophysica*. https://doi.org/10.1007/s11600-020-00454-9

Mancino, G., Ferrara, A., Padula, A., & Nolè, A. (2020). Cross-comparison between landsat 8 (OLI) and landsat 7 (ETM+) derived vegetation indices in a mediterranean environment. *Remote Sensing*. https://doi.org/10.3390/rs12020291

Meher, B., Agrawal, S., Panda, R., & Abraham, A. (2019). A survey on region based image fusion methods. *Information Fusion*. https://doi.org/10.1016/j.inffus.2018.07.010

Mohammadi, M., & Sharifi, A. (2021). Evaluation of convolutional neural networks for urban mapping using satellite images. *Journal of the Indian Society of Remote Sensing*. https://doi.org/10.1007/s12524-021-01382-x

Noh, H, S Hong, and B Han. 2015. "Learning Deconvolution Network for Semantic Segmentation." In *2015 IEEE International Conference on Computer Vision (ICCV)*, 1520–28. https://doi.org/10.1109/ICCV.2015.178.

Noh, H., Seunghoon, H., and Bohyung, H. (2015). "Learning deconvolution network for semantic segmentation." In *Proceedings of the IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV.2015.178

Park, K., Chae, M., & Cho, J. H. (2021). Image pre-processing method of machine learning for edge detection with image signal processor enhancement. *Micromachines*. https://doi.org/10.3390/mi12010073

Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., & Athanasiadis, I. N. (2021). Machine learning for large-scale crop yield forecasting. *Agricultural Systems*. https://doi.org/10.1016/j.agsy.2020.103016

Rabbi, J., Ray, N., Schubert, M., Chowdhury, S., & Chao, D. (2020). Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network. *Remote Sensing*. https://doi.org/10.3390/RS12091432

Ronneberger, O., Fischer, P., & Brox, T. (2015a). U-Net: convolutional networks for biomedical image segmentation. *In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-24574-4_28

Sharifi, A. (2018). Estimation of biophysical parameters in wheat crops in golestan province using ultra-high resolution images. *Remote Sensing Letters, 9*(6), 559–568. https://doi.org/10.1080/2150704X.2018.1452058

Sharifi, A. (2020). Using sentinel-2 data to predict nitrogen uptake in maize crop. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13*, 2656–2662. https://doi.org/10.1109/JSTARS.2020.2998638

Shelhamer, E., Long, J., & Darrell, T. (2017). Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2016.2572683

Shi, J., Jin, H., & Xiao, Z. (2020). A novel hybrid edge detection method for polarimetric SAR images. *IEEE Access*. https://doi.org/10.1109/ACCESS.2020.2963989

Sudmanns, M., Tiede, D., Augustin, H., & Lang, S. (2020). Assessing global sentinel-2 coverage dynamics and data availability for operational earth observation (EO) applications using the EO-compass. *International Journal of Digital Earth*. https://doi.org/10.1080/17538947.2019.1572799

Tariq, A., Shu, H., Siddiqui, S., Munir, I., Sharifi, A., Li, Q., & Linlin, Lu. (2021). Spatio-temporal analysis of forest fire events in the margalla hills, islamabad, pakistan using socio-economic and environmental variable data with machine learning methods. *Journal of Forestry Research*. https://doi.org/10.1007/s11676-021-01354-4

Waldner, F., & Diakogiannis, F. I. (2020). Deep learning on edge: extracting field boundaries from satellite images with a convolutional neural network. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2020.111741

Wang, J., & Wang, X. (2012). VCells: simple and efficient superpixels using edge-weighted centroidal voronoi tessellations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2012.47

Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., & Chan, K. (2019). Adaptive federated learning in resource constrained edge computing systems. *IEEE Journal on Selected Areas in Communications*. https://doi.org/10.1109/JSAC.2019.2904348

Wu, F., Bingfang, Wu., Zhang, M., Zeng, H., & Tian, F. (2021). Identification of crop type in crowdsourced road view photos with deep convolutional neural network. *Sensors (switzerland)*. https://doi.org/10.3390/s21041165