Deep Learning Classification for Crop Types in North Dakota

Ziheng Sun, Liping Di, *Senior Member, IEEE*, Hui Fang, and Annie Burgess

***Abstract*—Recently, agricultural remote sensing community has endeavored to utilize the power of artificial intelligence (AI). One important topic is using AI to make the mapping of crops more accurate, automatic, and rapid. This article proposed a classification workflow using deep neural network (DNN) to produce high-quality in-season crop maps from Landsat imageries for North Dakota. We use historical crop maps from the agricultural department and North Dakota ground measurements as training datasets. Processing workflows are created to automate the tedious preprocessing, training, testing, and postprocessing workflows. We tested this hybrid solution on new images and received accurate results on major crops such as corn, soybean, barley, spring wheat, dry beans, sugar beets, and alfalfa. The pixelwise overall accuracy in all three test regions is over 82% for all land types (including noncrop land), which is the same level of accuracy as the U.S. Department of Agriculture Cropland Data Layer. The texture of DNN maps is more consistent with fewer noises, which is more comfortable to read. We find DNN is better on recognizing big farmlands than recognizing the scattered wetlands and suburban regions in North Dakota. The model trained on multiple scenes of multiple years and months yields higher accuracy than any of the models trained only on a single scene, a single month, or a single year. These results reflect that DNN can produce reliable in-season maps for major crops in North Dakota big farms and could provide a relatively accurate reference for the minor crops in scattered wetland fields.**

***Index Terms*—Agricultural remote sensing, crop mapping, deep neural network (dnn), geoprocessing workflow, image classification, Landsat, North Dakota.**

I. Introduction

AGRICULTURE has been producing food, fibers, and energy to support the entire human society for thousands of years. To sustain and enhance crop production, farmers and stakeholders need to make informed decisions on many related activities. A lot of information is required for these decisions such as how much of what types of crops are planted in which areas—the crop map [1]–[6]. Crop maps are usually categorized into in-season maps (generated during the growing season) and after-season maps (generated after the growing season). The in-season maps are much more helpful for timely decision-making and have bigger financial benefits than the after-season maps.

The growing availability of data from recurring and long-term satellite observations has prompted a strong desire by the agricultural community to use satellite data in creating crop maps over large geographic areas [7]. The U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) combines the annually collected data by its field offices with satellite imagery to generate an annual crop map product, Cropland Data Layer (CDL). CDL is highly accurate on major crop types (e.g., ∼95% accuracy for corn and soybean) and seamlessly covers CONUS [8], [9]. Because CDL relies on ground truth and Landsat images of the entire growing season, it is only available internally around September of the current year and publicly around February of the next year. However, most crop stakeholders, such as farmers, agricultural departments, food market, and insurance companies, need in-season crop maps to assist them to make timely decisions [4], [10]. It has been a great challenge for the community to achieve accurate in-season maps. After observing the spectral features of crops over the growing season, it seems possible to generate in-season inventories with large-scale remote sensing imagery and limited field-observed data. The spatiotemporal and spectral variety in satellite images across the growing season should be able to help accurately recognize different types of crops (see Fig. 1).

There are two significant difficulties associated with operational remote sensing-based in-season crop mapping over large areas: 1) unavailability of in-season ground truths; and 2) complexity in managing the large-scale preprocessing and postprocessing workflows. Most existing trained models have strict spatiotemporal limitations. Many studies focus on small subsets of satellite scenes and their models have a large chance to fail when the spatiotemporal extent changes. The features extracted from one scene are inapplicable to other regions or dates, due to the seasonal and regional phenology of various crop types. On the other hand, although the automatic level of classification routine is improving, it still requires human intervention to manually tune and supervise the processes. To overcome the first difficulty, our study uses solid pixels in CDL to enrich the insufficient ground truth data. Solid pixels denote those pixels which are trusted to be accurate. Several criteria are used to amid the judgment, including whether the eight-directional neighbor pixels are the same value, whether the pixel is on the boundary, or whether the respective quality flags are on. Only the pixels which meet all the criteria are considered as solid/trusted pixels.

Manuscript received December 27, 2019; revised March 16, 2020; accepted April 16, 2020. Date of publication May 14, 2020; date of current version May 29, 2020. This work was supported in part by the U.S. National Science Foundation under Grant AGS 1740693 and Grant CNS 1739705, and in part by ESIPLab Geoweaver project. *(Corresponding author: Liping Di.)*

Ziheng Sun, Liping Di, and Hui Fang are with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, Virginia 22030 USA (e-mail: zsun@gmu.edu; ldi@gmu.edu; hfang1288@gmail.com).

Annie Burgess is with Earth Science Information Partners, Boulder, Colorado 80304 USA (e-mail: annieburgess@esipfed.org).

Digital Object Identifier 10.1109/JSTARS.2020.2990104

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

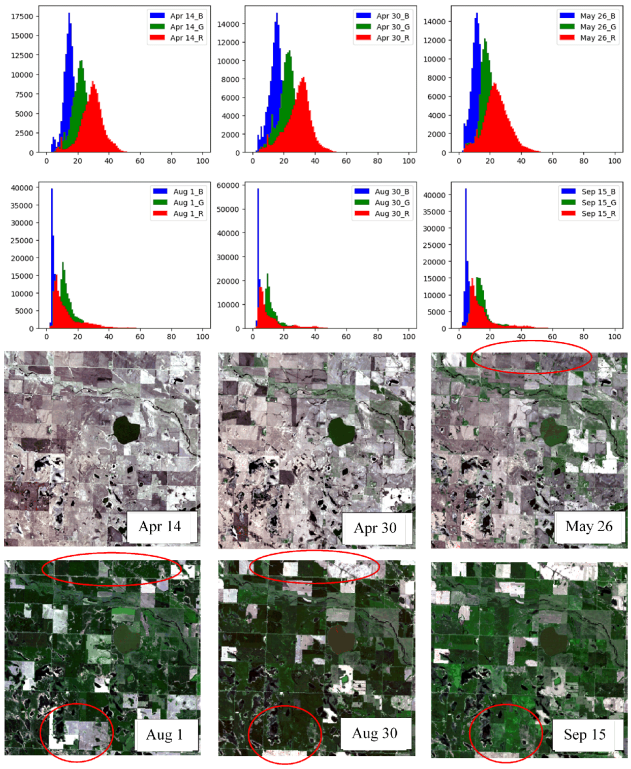

Fig. 1. Landsat spectral pattern of crops and adjacent land cover in eastern North Dakota (*X*-axis: reflectance value; *Y*-axis: appearance frequency); the bottom images show the greening up and harvesting of crops (showing that the neighbor fields have very different farming calendar).

To address the second difficulty, this study also uses an advanced workflow system [11]–[15] to compose an automatic workflow for automating the tedious small processing steps, manipulating computing facilities, and coping with large-volume Landsat images. The workflow is used to yield in-season crop maps from satellite images of North Dakota collected in the green-up and prematurity stages of major crops such as corn and soybean. The core function of the workflow is some machine learning algorithms for classifying the pixels into semantic types with actual meanings.

Machine learning is a family of data analysis methods that automate analytical model building based on the idea that algorithms can learn from data, identify patterns, and make decisions with the minimal human intervention [16]–[19]. Researchers have experimented with many machine learning methods to derive crop maps directly from satellite observations. Among all the machine learning methods, neural network has been a popular one since the 1980s [20]. Many types of neural networks have been created since the efficient gradient descent method, back-propagation, was invented in the 1980s [21]. In recent years, the augmented version of neural networks, deep neural network (DNN), gains a lot of attention as they outperform many other existing solutions in many image analysis tasks. Deep learning (DL) is a general designation of DNNs to learn very abstract features from data and use them to make predictions on new data. It is initially inspired by the biological nervous system and now broadly used in computer vision, speech recognition, and natural language processing. DNN has been tested and proven effective in the computer vision domain. Comparing to conventional task-wise algorithms such as K-nearest neighbor, random forest (RF), and support vector machines, a well-trained DNN model seems to be able to extract more useful and common features and is expected to significantly reduce costs by being directly reused across multiple tasks. However, using a new technique in a new domain, e.g., agriculture, will always meet some bottleneck challenges due to the differences in data, requirements, and rules.

In crop mapping [22], semantic segmentation is the key step that uses spectral reflectance, localized properties, and spatial context information to determine the land cover type. One example of the state-of-the-art DNNs for semantic segmentation is fully convolutional network (FCN). FCN replaces the fully connected (dense) output layers in conventional convolutional neural networks (CNN) with layers that can output 2-D images. But FCNs cannot take global context on various scales into account. To solve the disadvantages, many complex neural networks were created, such as AlexNet, VGG, GoogLeNet, ResNet, R-CNN, DeepLabv3, SegNet, U-Net, and so forth. Although the design goals of these networks vary widely according to their initial use cases, most of them can support semantic segmentation task.

Besides, due to the high similarity in crop variety (barley and wheat, red beans and black beans, grass, and alfalfa, etc.), crop recognition is more difficult than many other vegetation types. To accurately distinguish the crops, this article takes advantage of SegNet [23] that has a stack of encoders and decoders and a soft-max output layer. The encoders help extract very high-level features and the decoders and soft-max layer together translate the extracted features back to full-input-image-size maps [23]. To better fit in crop mapping tasks, the SegNet model is tuned to be sensitive to minor spectral differences while maintaining accurate extraction of high-dimension large-scale contextual features.

An advanced workflow management system is used to automate the tedious large dataset processing, to enhance time efficiency and reduce scientific reproducibility issues [13], [16], [24]. We implemented the DNN crop mapping workflow in Geoweaver—an open-source web-based workflow system [25]–[27], to help us integrate distributed resources, automate the pre/postprocessing workflows and track the provenance of the final crop maps [25]. Training DNN models is a back-and-forth process and requires a lot of skills and expertise. Geoweaver is an effective tool for recording the training history and locating the roots of problems. The DNN workflow runs regularly on near-real-time observed satellite images and quickly produces in-season crop maps for new Landsat scenes that are observed.

This remainder of the article is organized as follows. The related work is investigated in Section II. Section III introduces the used DNNs. Section IV introduces the study area and data source. Section V proposes the framework integrating distributed DNN workflows on advanced cyberinfrastructure. Section VI describes the experimental results. The discussion is in Section VII. Section VIII concludes the article and gives the future work.

II. Literature Review

Machine learning has drawn a lot of attention in recent years. The recent research literature is reviewed. Fritz *et al.* [28] produced a 1-km global cropland percentage map of 2005 by integrating satellite images and individual cropland maps of various scales and got an overall accuracy (OA) of 82.4%. Estel *et al.* [29] used MODIS normalized difference vegetation index (NDVI) times series to map active and fallow farmland and provide the first European-wide map of abandoned farmland (cropland and grassland) and cultivation. They used the RF classifier with independent observations from the field and satellite images. Gao *et al.* [7] studied mapping crop progress at field scales by fusion of Landsat and MODIS imagery. Roy and Yan [30] evaluated the 5- and 7-parameter linear and nonlinear harmonic models in simulating crop growth. Muller *et al.* [31] studied the use of spectral-temporal variability metrics to separate cropland, pasture, natural savanna vegetation, and other relevant land cover classes for a savanna landscape in the Brazilian Cerrado using Landsat. Xiong *et al.* [32] used Google Earth Engine and MODIS NDVI to calculate the extent of croplands in Africa. Teluguntla *et al.* [33] also used RFs on Google Earth Engine to generate Landsat-derived cropland extent products of Australia and China.

Kussul *et al.* [34] described a multilevel DL architecture to classify crop maps from multitemporal multisource satellite imagery including Sentinel-1A and Landsat-8. They first used an unsupervised neural network to process optical segmentation and missing data restoration, and then used supervised fully connected multilayer perceptron to generate crop maps for a test site in Ukraine. Their result accuracy can reach higher than 85% for all the major crops (wheat, maize, sunflower, soybeans, and sugar beet). Sun *et al.* [35] have proposed an end-to-end framework using long short-term memory recurrent neural network to take the temporal pattern of crops across image time series to improve the accuracy of predicted cropland maps. Waldner *et al.* [36] proposed a solution for automated annual cropland mapping using knowledge-based temporal features. They tried to build five knowledge-based temporal features that remain stable over time and a classifier to deliver cropland maps based on those features. They tested the system in Argentina, Belgium, China, and Ukraine, and found that their method increases the stability of the classifier allowing its reuse from year to year without recalibration. Were *et al.* [37] compared the results of support vector regression, RFs, and neural networks in mapping soil organic carbon stocks in the Eastern Mau Forest Reserve.

These studies laid the foundation for producing in-season cropland maps with high resolution and detailed information on crop coverage and categories more automatically and more time wisely.

III. Deep Neural Networks

Crop mapping is an image semantic segmentation task and many DNNs can be utilized. We use SegNet as our DNN model because of its demonstrated capability in identifying similar classes in street view recognition [23], [38]–[40] and we expect it reproduces a similar result in crop mapping. The original version

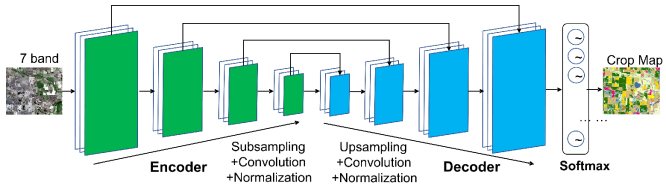

Fig. 2. SegNet.

of SegNet has four pairs of encoder and decoder layers [23] (see Fig. 2). Each encoder has a subsampling layer, a dense convolution layer, and a batch normalization layer [41]. The subsampling layer uses a nonoverlapping 2×2 or 3×3 pooling window. Rectified linear unit (ReLU) is the activation in the convolution layer [42]. The batch normalization layer is used to normalize the activations of the previous layer at each batch to accelerate network training by maintaining the mean activation close to zero and the standard deviation close to one [43]. Each decoder has an upsampling layer, a dense convolution layer, and also a batch normalization layer. The upsampling layer magnifies its input using the memorized pooled indices to a larger tile with higher resolution, which is the reverse of the subsampling layer [41]. The output layer uses soft-max activation to classify each pixel independently [44]. The output is an N-band image where N is the number of classes in the target hierarchy. In this case, N is the number of crop categories. The chosen loss function is categorical cross-entropy [45], and the optimizer function is Adadelta [46]. Categorical cross-entropy is the cross-entropy loss function for multiclass classification. The loss increases as the predicted probability diverge from the actual label value. A perfect model would have zero loss. Adadelta optimizer can adapt the learning rate according to the moving window of gradient updates and delivers a more stable learning curve.

Besides SegNet, U-Net is another popular DNN for such a task. It consists of a contracting path to capture context and a symmetric expanding path that enables precise localization [38], [47]. The initial experiment on the biomedical image database of the IEEE international symposium on biomedical imaging (ISBI) challenge [47] shows that U-Net can be trained end-to-end from very few images and outperforms other convolutional networks. The learning and testing speed is very fast. The original U-net structure is divided into two halves. The first-half of U-net includes repeated 3×3 convolution layers (unpadded convolutions) [48], each convolution neuron followed by a ReLU and a 2×2 max pooling operation [49] with stride 2 for down-sampling. Every step in the second half has an upsampling of the feature map followed by a 2×2 deconvolution layer [50] that reduces the number of feature channels, a concatenation with the fused feature maps from the first half, and two 3×3 convolutions, each followed by an ReLU. Padding after each step is necessary due to the loss of border pixels in every convolution. At the final layer, a convolution layer is used to map each feature vector to the number of classes. To make the tiling able to be merged back to a seamless map, it is important to set the tile size to ensure all max-pooling operations are applied to a layer with an even x- and y-size [51]. Same as SegNet, the output layer adopts the soft-max activation function

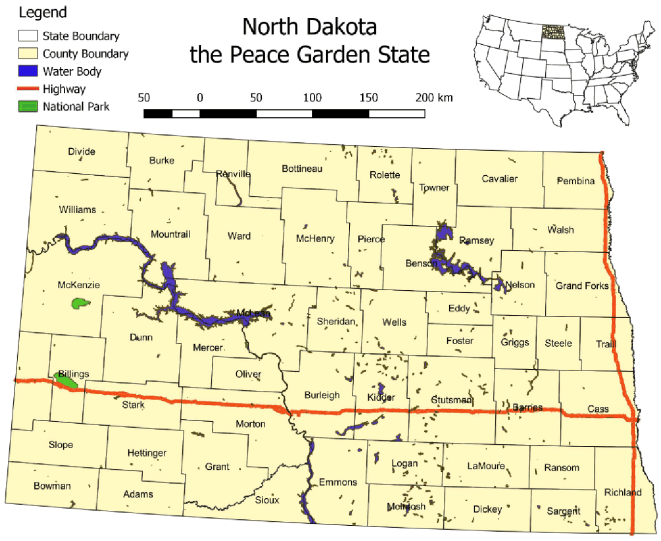

Fig. 3. North Dakota.

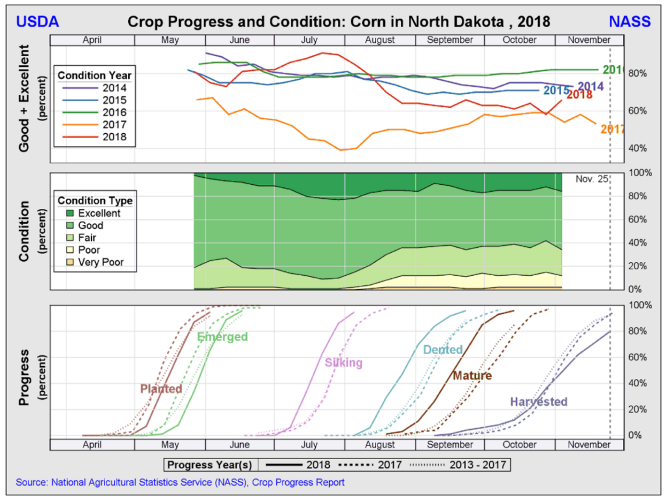

Fig. 4. Growing stage time chart of corn in North Dakota, in 2018 (cited from USDA NASS website https://bit.ly/36ddGyN).

and the loss function uses cross-entropy [38]. Considering U-Net has more parameters and is relatively slower to train than SegNet when the training dataset is big, we chose SegNet instead of U-Net out of totally financial efficiency reasons.

IV. Study Area and Data

A. Study Area

U.S. agriculture has distinct regional patterns in the Central Valley, Midwest, and Great Plains. For example, California's Central Valley is the most concentrated region for valuable cash crops, including fruits, nuts, berries, vegetables, etc. The Midwest region has dominated U.S. corn production for many years and is widely known as the "corn belt" [52]. The broad version of "corn belt" states includes North Dakota, South Dakota, Ohio, Michigan, Iowa, Illinois, Indiana, Nebraska, Kansa, Minnesota, Missouri, Wisconsin, and Kentucky. These states are located in the Great Plains with the great latitudinal and climatic span stretching from North to South. Our ultimate goal is to train a model that can apply to this entire region. In this study, we choose North Dakota—the Peace Garden state (see Fig. 3), as our study area to experiment and validate our approach. Once the model accuracy becomes reliable, we will expand the coverage to the entire corn-belt region.

B. Dataset

Landsat 8 imagery is used as model inputs and CDL, NASS reports (see Figs. 4 and 5), regional crop maps and ground-collected datasets are used as training labels. Most of the ground truth data were obtained via field surveys and roadside photo samples. Some crop field boundaries are got by digitalizing the high-resolution images from the National Agriculture Imagery Program [53]. The data sources include the Common Land Unit of USDA Farm Service Agency, the data portal of the state government of North Dakota [54], and the data archive of North Dakota State University Agriculture Experiment Station [55],

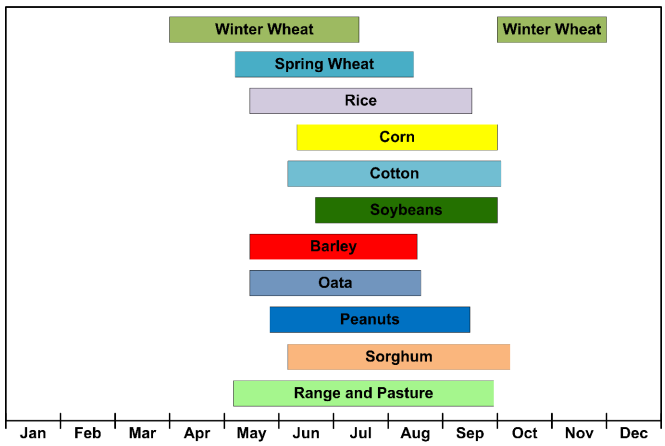

Fig. 5. Growing season chart of U.S. major crops (source: USDA NASS).

[56]. These data programs have been carried out by experienced data collectors following a series of protocols and managed by government departments or universities. The quality of the ground data is very reliable.

The crop phenology is complicated and the crops in two neighbor fields are in very different stages by a very high chance. Spatial scale, observation date, and greenness are the three important factors used in choosing Landsat scenes and collecting ground truth. Landsat resolution is suitable for measuring dynamics at the field level. In 2018, the corn in North Dakota is planted in April and May, silking in late June and entire July, become dented and mature in September, and harvested in October and November. It shows that every year the period and duration of each growing stage shifts back and forth. We need to choose Landsat 8 collected in the silking, dented, and mature stages when the crops have the largest leaf area and the strongest spectral reflectance of crop plants. To make it easier, the selection of Landsat images is automated by calculating NDVI from band 4 and band 5 and counting the pixels whose NDVI values are larger than 0.4 in the scene. In most cases,

moderate vegetation, e.g., crops at an early stage, tends to vary between 0.4 and 0.6. Anything above 0.6 indicates the highest possible density of green leaves [57]. If the number exceeds a certain threshold (thresholds differ among regions), the image would be considered as suitable for crop mapping. We will use the scenes from 2013, 2014, and 2015 as training datasets and test the trained model on the scenes, in 2016 and 2017.

C. Landsat 8 Surface Reflectance (SR) Product

SR is the amount of light reflected by the earth's surface [58]. The Landsat SR products account for atmospheric effects such as aerosol scattering and thin clouds to enhance image brightness and can help the detection and characterization of earth's surface change. Landsat 8 OLI SR products, which are used in this study, are generated using the Landsat Surface Reflectance Code (LaSRC), which makes use of the coastal aerosol band to perform aerosol inversion tests, uses auxiliary climate data from MODIS, and a unique radiative transfer model [59]. LaSRC hardcodes the view zenith angle to "0," and the solar zenith and view zenith angles are used for calculations as part of the atmospheric correction. Most day-lit (descending) Landsat 8 scenes can be processed to SR. Newly collected Landsat 8 scenes become suitable for processing within a few days of data acquisition [60]. However, the efficacy of land SR correction is likely to be reduced in hyperarid or snow-covered regions, areas with low sun angle conditions, coastal regions where land area is small, relative to adjacent water, or areas with extensive cloud contamination [61]. Corrections may not be accurate to data acquired over high latitudes ($> 65°$ North or South). People can refer to the quality assessment band for pixel-level condition and validity flags [62], [63]. Besides, Landsat atmospheric correction and SR retrieval algorithms are not ideal for water bodies due to the inherently low level of water-leaving radiance, and the consequential very low signal to noise ratio. Similarly, SR values greater than 1.0 can be encountered over bright targets such as snow and playas. These are known computational artifacts in the Landsat SR products. Within 65° North or South, Landsat SR products are very reliable for measuring the actual reflectance of the crop plants and release scientists from the job of atmospheric correction [58].

V. Approach

Most DL researches in agriculture directly reuse the neural networks designed by computer scientists, which have been proven workable in many crop classification tasks [34], [64], [65]. The current major challenges in applying DL in agriculture are preprocessing the existing observation into the learning-ready format, and postprocessing the neural network results into crop maps. Although there are many DL tools, web systems, libraries, it is still hard to put together all the heterogeneous pieces into a consistent pipeline [66], [67]. The proposed framework integrates DNN and cyberinfrastructure to build the workflow (see Fig. 6) and produce crop maps from remote sensing images automatically.

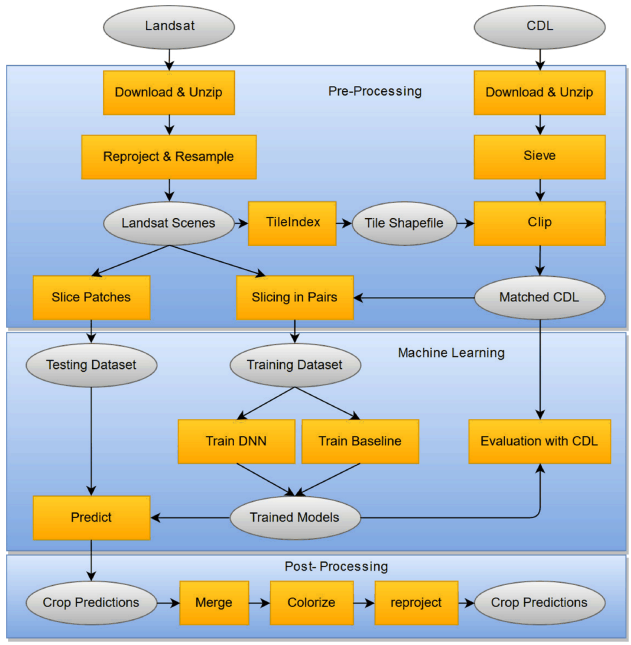

Fig. 6. Workflow.

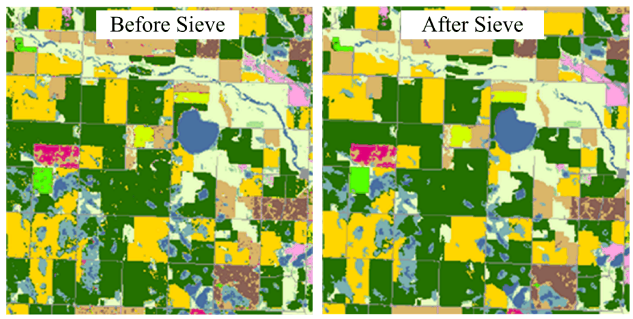

Fig. 7. Sieve process (threshold: 15).

A. Preprocessing

To make the training more efficient, the preprocessing of DNN needs to take extra steps compared to the conventional classification scheme. Besides the normal preprocessing operation like calibration, atmospheric correction, and spatial enhancement, Landsat and cropland map products must be prepared into input batches and output masks respectively. Landsat images (downloaded via USGS EarthExplorer) will be first unzipped and reprojected to NAD83/Conus Albers (EPSG:5070) projection and resample to exact 30-m spatial resolution (reprojecting will change the resolution).

CDL for North Dakota is downloaded and unzipped into ERDAS IMG files. We use geospatial data abstraction library (GDAL) [68] to read and sieve CDL to remove the isolated pixels and reduce data uncertainties and complexity. As shown in Fig. 7, the sieving process removes raster polygons smaller than a provided threshold size (in pixels) and replaces them with the pixel value of the largest neighbor polygon. It can greatly eliminate the inconsistent small-area crops within the other

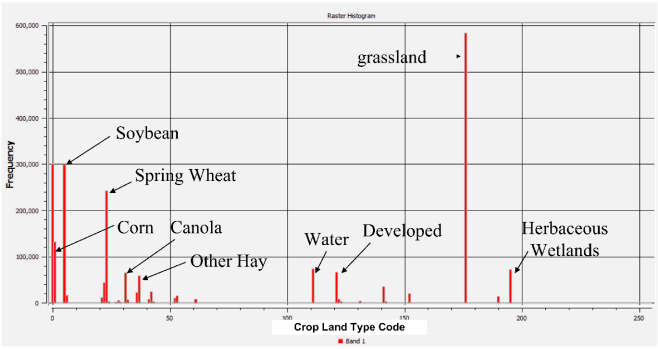

Fig. 8. Naturally biased crop distribution in North Dakota (2017).

large-scale fields, e.g., the unexpected soybean fields within the cornfields. The downside is that some good pixels, such as the river sections, are removed too. It could be complemented by applying an additional water/road data layer to force update the sieve results with ground truth. A pixel-to-pixel mapping between Landsat and CDL will be established based on the resampled Landsat and sieved CDL. Any location mismatch in the process will result in huge failures. To make them exactly match, we use GDAL [68] to extract the boundary polygon of every scene and use the polygon to clip the CDL. The resampled Landsat images (1–7 bands) and the matched CDL images are sliced into two folders: inputs (Landsat bands) and targets (CDL). We filtered all those out-of-season batches (before emerging or after harvesting), bad pixels, cloud pixels, snow/ice pixels, edge-effect pixels, untrusted pixels (confidence coefficient is low), etc., to ensure training dataset quality.

CDL image values range from 0 to 255. However, not all of the 256 values are used to label the crop types and there are some empty values. In particular, only 132 values correspond to real land cover types in the hierarchy. The one-hot encoding method is used to map the crop classes (0–255) into a matrix of numbers ranging from 0 to 131. One pixel is corresponding to one row in the matrix. In the row, the item respective to the correct class has a value of 1, and all the others are 0. The hierarchy without empty classes will make the texture of the results clean and unified.

B. Machine Learning on Imbalanced Dataset

This work reuses the USDA NASS crop classification system which has 132 crop classes. Unfortunately, the training dataset is naturally seriously biased (imbalanced) for many of the minority crops. As shown in Fig. 8, in 2017 most area of North Dakota is covered by grassland and the most grown crops are soybean, spring wheat, corn, and canola. Other crops have a much smaller growing area. The crops are not equally represented in either the reality or the training dataset. To better represent the minority crops, we could use resampling techniques to increase the ratio of the minority crops in the training dataset.

A DNN model with seven input channels and 132 output channels is created based on the original SegNet model. The seven bands of Landsat are fed into the seven channels. The data will be digested by the subsequent encoders to do the convolutional computation to extract high-level features from the bands. The information from different bands may be combined or exaggerated in the process in many ways to discover the most common and effective features. Later, the features will be translated by the decoders to get a same-size crop map. The model needs a lot of training before working properly.

The ultimate goals of model training include not only getting higher classification accuracy but also achieving generalization. A well-trained neural network allows people to make predictions on data the model never met. DL has two core procedures—training and validation. The commonly used algorithm for training neural networks is backpropagation which fed the errors from the classification of the previous record back into the neural network to modify the weights on the links among neurons for the next run and repeat for many iterations. Usually, the training dataset is split at a certain ratio, e.g., 8:2, in which 80% percent of samples will be used in training and the rest 20% will be for testing. In every iteration, the model will calculate its errors on both portions. The accuracy got from the 80% portion is called training accuracy, while the accuracy from the 20% portion is called testing accuracy. Via the setting, we can determine overfitting and underfitting in the learning process. Overfitting happens when the neural network learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on the new data. It will prevent the trained network from applying to new data and generalization. Generalization means how well the concepts learned by the model apply to strange datasets that the model is never trained on before. Underfitting means the model is not trained enough and is not suitable to give good performance on either training dataset or testing dataset. Overfitting and underfitting, two typical problems caused by poor generalization, permanently coexist with neural network training. That is one of the major reasons why neural network training is always back and forth, and very hard to balance between the concrete prior knowledge and abstract representation features. The dividing line between overfitting and underfitting is vague and floating per case. It requires a lot of experience and expertise from DL practitioners. For example, in Fig. 9 we show the learning curves of accuracy and loss function by training DNN on one Landsat scene and CDL. After training for three hours with 1000 iterations on the training batches, the model is still underfitting (the training and validation accuracy curves are not going apart) and has space to improve. The training and validation accuracy are very close and their mean values both reach 80%. There are many techniques and tricks to help prevent overfitting, including Dropout, Bias, L1, and L2 regulation. But none of them can substitute human supervision at present.

C. Postprocessing

Translating the machine learning results back to cropland maps requires a series of post processes. The output of the neural network is processed by a soft-max function which turns numbers into a probability array indicating the probability of the inputted objects belonging to each class (one-hot encoded matrix). The classes with the highest probability will be assigned to the objects. The DNN results would be in the shape of (batch size, batch width $*$ batch height, 132). They will be mapped back

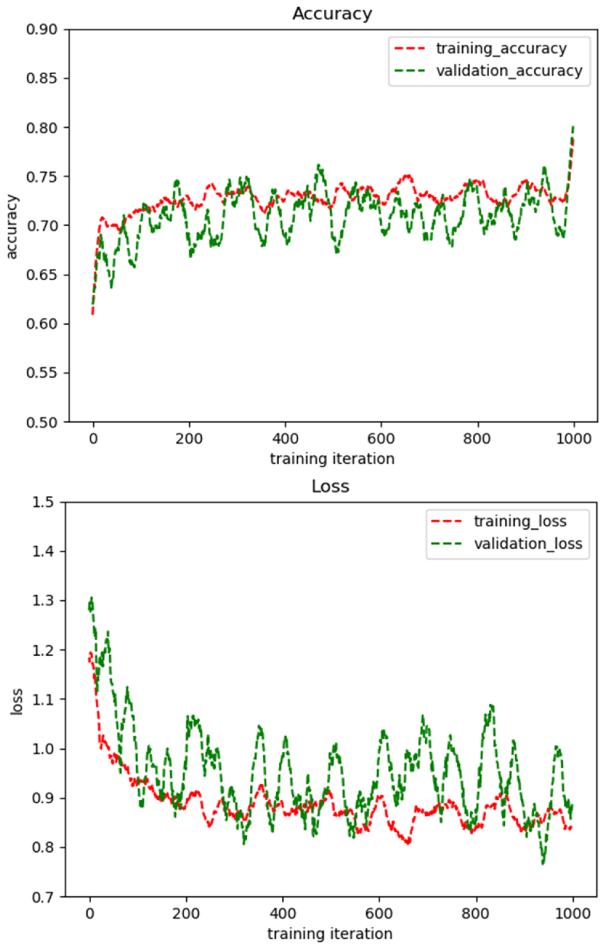

Fig. 9. Curves of model validation accuracy and loss function value in training processes.

to the 0–255 range, the same as CDL by adding empty values and rendered using the CDL color style. The result batches will be merged into a big map, reprojected and rendered using the same color styles as CDL. Due to uncertainties in classification, the predicted results must integrate some ground truth files to force correct some parts of the areas such as the residential area, high way, waterbody, cloud, snow, cloud shadow, etc., which are not the emphasis of this research. We will directly use the data from other sources such as NASA NLCD [69] and OpenStreetMap [70] to fill in those regions. The results will be evaluated by several metrics. The common metrics such as OA, producer accuracy, consumer accuracy, Kappa coefficient, and F-1 score, will be all calculated based on the confusion matrix [71], [72]. Section VI will give more details about the evaluation results.

D. Setup on Advanced Workflow System

Considering the large data volume, tedious pre/post processing steps, hyperparameterization in the neural network, most DL tasks cannot be carried out by a single machine but need multiple machines with high-performance parallel computing devices such as GPU to accelerate the process into a reasonable time frame. Cyberinfrastructure such as Google Earth Engine [73] and Amazon Web Service (AWS) [74], [75] have been

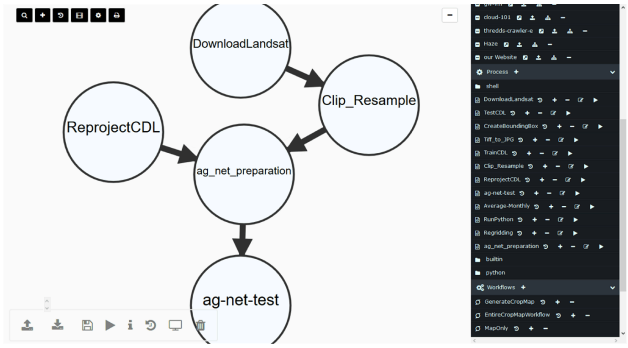

Fig. 10. Production workflow in Geoweaver.

proven greatly useful in big data storage, processing, analysis, and dissemination. Many online software resources such as Google Cloud Datalab, Google AI, FloydHub, OpenML, BigML, etc., are also available to help. Besides public resources, many institutes have their own datasets, legacy machine learning code, computational workstations, rack servers, private cloud, cluster, and cloud instances. Most of the cyberinfrastructures are not directly compatible with each other regarding the current big data processing framework, and the transfer of data and application is very painful. To successfully set up a DL workflow on the hybrid of private and public cyberinfrastructures, advanced tools are required to connect and make them collaborate with each other in a seamless way [24], [76]–[79].

The DNN workflow is complicated and time consuming and involves a series of tools, software, and scripts in various languages. For example, the scripts for re-projecting and resampling are based on GDAL [68] and written in C language. The scripts for plotting the statistics results use R and matplotlib. The DNN definition and training/testing programs are built on Keras in Python. The overall volume of the original Landsat scenes from USGS is around 5 TB after decompression. However, the intermediate products generated by the preprocessing and postprocessing steps take more than 30-TB disk space. The DL code is deployed on a DELL server with two NVIDIA Tesla K-80 GPUs. All the preprocessed training samples must be transferred to that GPU server for training and testing. Managing this compound workflow is a daunting challenge for DL practitioners. Instead of manual management, an efficient management tool will largely decrease the complexity, make the workflow reusable, and allow the tracking of provenance. In this study, we use Geoweaver to set up our production environment (see Fig. 10). Geoweaver is a graphical workflow tool for managing DL workflows in a distributed environment. It can substitute Secure Shell client, File Transfer Protocol client, virtual private network (VPN), and workflow software. It supports interacting with multiple GPU servers and uses shell scripts to automate the run of preprocessing, training, testing, and postprocessing in DL workflows. We can upload and browse the result files in the GPU server in a one-stop manner.

Additionally, provenance information is automatically captured and recorded, which can enable the inspection of each step so a user could quickly locate misclassification and investigate the cause of problems [80]. The atom processes in

TABLE I
SETUP CONFIGURATION

	Scene	Year	Month
One Year	All	2017	Apr, May, Jun, July, Aug
One Scene	Row 032, Path 027	2013-2017	Apr, May, Jun, July, Aug
One Month	All	2013-2017	July
Multiple Everything	All	2013-2017	Apr, May, Jun, July, Aug

Geoweaver can be heterogeneous so that scientists can compose their legacy code without rewriting the C++ code in Python or vice versa. Geoweaver is a decentralized cross-platform system running in web browsers, and every DL group could install a private instance on their machines (including their laptops) to manage their resources. The workflow and scripts recorded in a Geoweaver are reusable in any other Geoweaver instances regardless of the underlying hardware environment, which makes it very convenient to reproduce the crop maps or generate new maps from new Landsat scenes by various user groups. We have successfully used Geoweaver in carrying out the experiments by many times which saved us at least 30% of time on interacting with the servers, scripts, code, version control, result files, data transferring, etc. Geoweaver is the cyberinfrastructure tool for students, researchers, engineers, or even citizen scientists who are devastated in managing various servers, heterogeneous language programs, code version management, or a tremendous number of files.

VI. EXPERIMENT AND ANALYSIS

We have successfully set up DNN workflows on Geoweaver and run it with various parameters (see Table I). We wrote our SegNet using Keras and Tensorflow and tested on two NVIDIA Tesla K80 GPUs, and 10-core Intel Xeon 4116 CPU. The Landsat 8 scenes of North Dakota from 2013 to 2017 are retrieved. Most of them are observed between June and September. Many crops in the crop classification hierarchy are not grown in North Dakota. Small-frequency crops will have fewer samples which would lead to a lower recall. In this study, we mainly focus on the major crops in North Dakota and will improve the accuracy of minority crops in the future by extending the study area to increase the size of minority crop samples.

A. Comparison With Baseline Models

The changes of study regions will alter model performance. We need first test DNN to compare with traditional baseline models to prove that DNN is usable in this region with the provided satellite images. We choose RF as our representative of baselines models and use one Landsat scene 032027 captured in July 2016 as our test image. We initially trained both DNN and RF on the above half of the image and tested the model on the other half. Fig. 11 displays some result comparison. Comparing to the ground truth polygons, DNN and RF agree

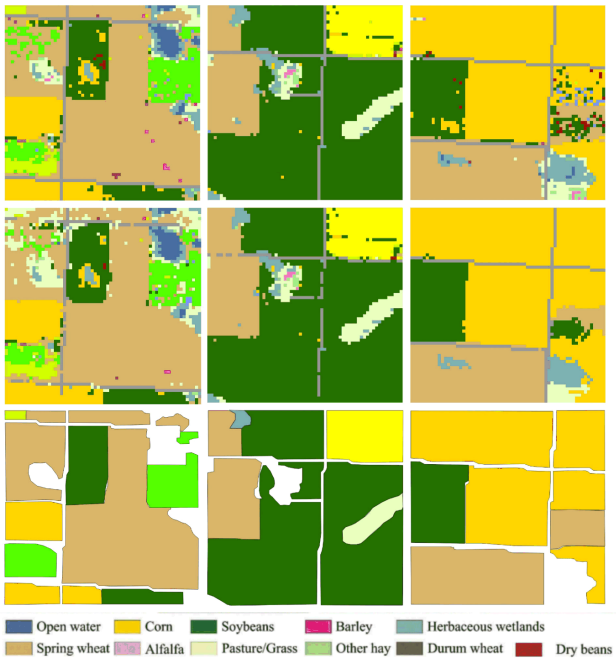

Fig. 11. Comparison of DNN results with RF results in July 2016 (the first row are RF results; the second row are DNN results; the third row are ground truth.).

very well in the distribution of corn and soybeans. In the DNN results, corn and soybean fields have less isolated pixels than RF results which probably benefits from the noise-reducing ability of convolutional layers. They also have very close results in spring wheat, hay. Wetlands and water bodies are correctly recognized in both models. The roads are more broken in the DNN results. Both models classified some barley and dry beans spots in the middle of spring wheat fields and generate some herbaceous wetlands in the soybean fields (the first column) while the ground truth shows that it is a pure soybean field. That might be caused by a small pond formed by rainfall or irrigation. Considering the DNN is still underfitting and has space for improvement, this test validates that DNN is applicable in cropland classification of North Dakota.

B. Comparison of Different Training Datasets

The selection of training samples has a direct impact on the success of the DNN model. High-quality training dataset will greatly accelerate learning and enhance accuracy. Four training datasets are created to find out which configuration could maximize learning efficiency. As displayed in Table I, the training datasets are customized according to the used Landsat scenes, observed year, and observed month. The first dataset uses the Landsat images from all the scenes and months in one year (2017). The second dataset uses only one scene at Row 032 Path 027 and collected all the images of that scene from 2013 to 2017. The third dataset uses the images observed in July of all the years at all the scenes in North Dakota. July is selected because it is when many crops have the largest leaf areas while fewer crops are harvested. The fourth dataset contains all the images. Each dataset is split into training and validation subsets and trained

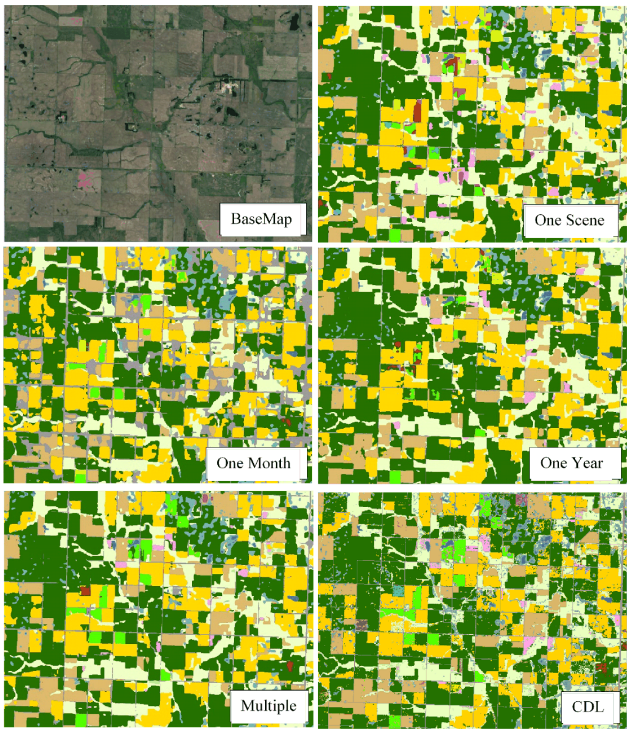

Fig. 12. Comparison of corn and soybean on different training datasets.

DNN on all the four datasets by 10 epochs. The four trained models are tested on the same validation image and the results are displayed in Fig. 12. All the results are similar in corn and soybean fields. The one-scene result misclassified some large soybean fields into cornfields, some pasture into alfalfa, some spring wheat fields into barley fields. The month result has more pixels absorbed by the highway and residential area. The one-month and one-scene results are less accurate than the one-year and multiple results. Furthermore, the multiple-everything result has better performance in peas than the one-year result. Overall, multiple-everything training dataset yields the best OA.

C. Compare With CDL

We fully trained DNN on the multiple training datasets and use new test images to predict crop maps and compare them with the NASS CDL. Fig. 13 shows the side-by-side comparison of several typical croplands in North Dakota including sugarbeets, corn, soybean, dry beans, canola, sunflower, peas, durum wheat, barley, lentils, flaxseed, etc. The left maps are the result of DNN from Landsat 8 image on August 3, 2016; the right maps are CDL. In all the chosen croplands, soybean and corn are very close and the differences are mostly in the field edges. DNN results remove many isolated pixels within the corn/soybean fields and make the crop maps more consistent and cleaner. In herbaceous fields, wetland, pasture, and hay are mostly identified correctly in DNN results with less isolated water body pixels than CDL [see Fig. 13(a)]. Durum wheat, a unique category in U.S. crops which is mainly distributed in North Dakota, is separated from spring wheat very well by DNN [see Fig. 13(c)]. Spring wheat and barley are correctly recognized at most times, but DNN still confused the two sometimes due to their high similarity in spectral characteristics. Also, DNN feels hard to tell the differences between alfalfa and pasture/grassland and misclassifies some grassland into alfalfa. The spectral-unique minority crops such as sugarbeets [see Fig. 13(a), (e)], sunflower [see Fig. 13(b)], canola [see Fig. 13(d)], and peas [see Fig. 13(d)] are generally accurate in DNN results, although some of them are classified into spring wheat or corn. The DNN results have relatively irregular field shapes than CDL as a result of its encoder–decoder fusion. CDL fields have better field boundaries which could be reversely used to improve the DNN maps. Besides, DNN can successfully recognize the large developed space (towns, cities, highways, etc.), while missing some small roads between fields [see Fig. 13(a) and (e)]. That is because small roads are extremely underrepresented in the training dataset. As the field roads are barely changed, we can fix it by directly integrating the road pixels from CDL and NLCD. Overall, the DNN maps give a reasonable prediction on crop distribution purely based on SRs and show high agreement with CDL in major crops in big farmlands. For the minority crops, such as sunflower, canola, sugarbeets, peas, lentils, flaxseed, DNN still did a good job in distinguishing them despite their seriously underrepresented samples. It could be foreseen that the accuracy of minority crops would increase if more samples are involved to further balance the biases.

D. Quantitative Assessment

To quantitatively evaluate the accuracy, we chose three regions in North Dakota (see Fig. 14) and calculated their confusion matrix based on solid/trusted pixels in CDL (see Fig. 15). A solid pixel means all the eight neighbors have the same value with itself, which will eliminate the edge pixels with a high risk of errors in CDL. The three chosen regions are southeast of Darling Lake, Missouri River basin, and west of Devils Lake. The confusion matrix is normalized to highlight the underrepresented crops. In the Darling Lake region, spring wheat (23) and grassland (176) have >90% OA [see Fig. 15(a)]. Soybean and corn's accuracies are 89% and 74%, respectively. The accuracy of soybean is pulled back by dry beans which are misclassified into soybeans in that region. Sunflower is 82% and peas have 73% accuracy. Dry beans are mostly misclassified into soybeans in this region. Alfalfa has 61% OA while the other 28% of alfalfa are classified into grassland in DNN results. Barley is 72% accurate and 14% is classified as spring wheat and the other 10% is classified as grassland. In this herbaceous wet region, DNN did very well in recognizing soybean, sunflower, spring wheat, and grassland. In the Missouri river basin, which is a suburban area with occasional croplands, DNN correctly recognized 98% pasture/grassland and did fairly good on corn (81%), soybean (75%), sunflower (70%), spring wheat (80%), peas (81%), and alfalfa (73%) [see Fig. 15(b)]. It has poor results in peas (46%) many of which were classified into spring wheat (31%). It might be caused by the fact that in this region, the portion of the pea field is very small (barely seen) and the boundary pixels are occupied by the neighbor spring wheat fields. It reflects that the less area a crop has, the less accurate the

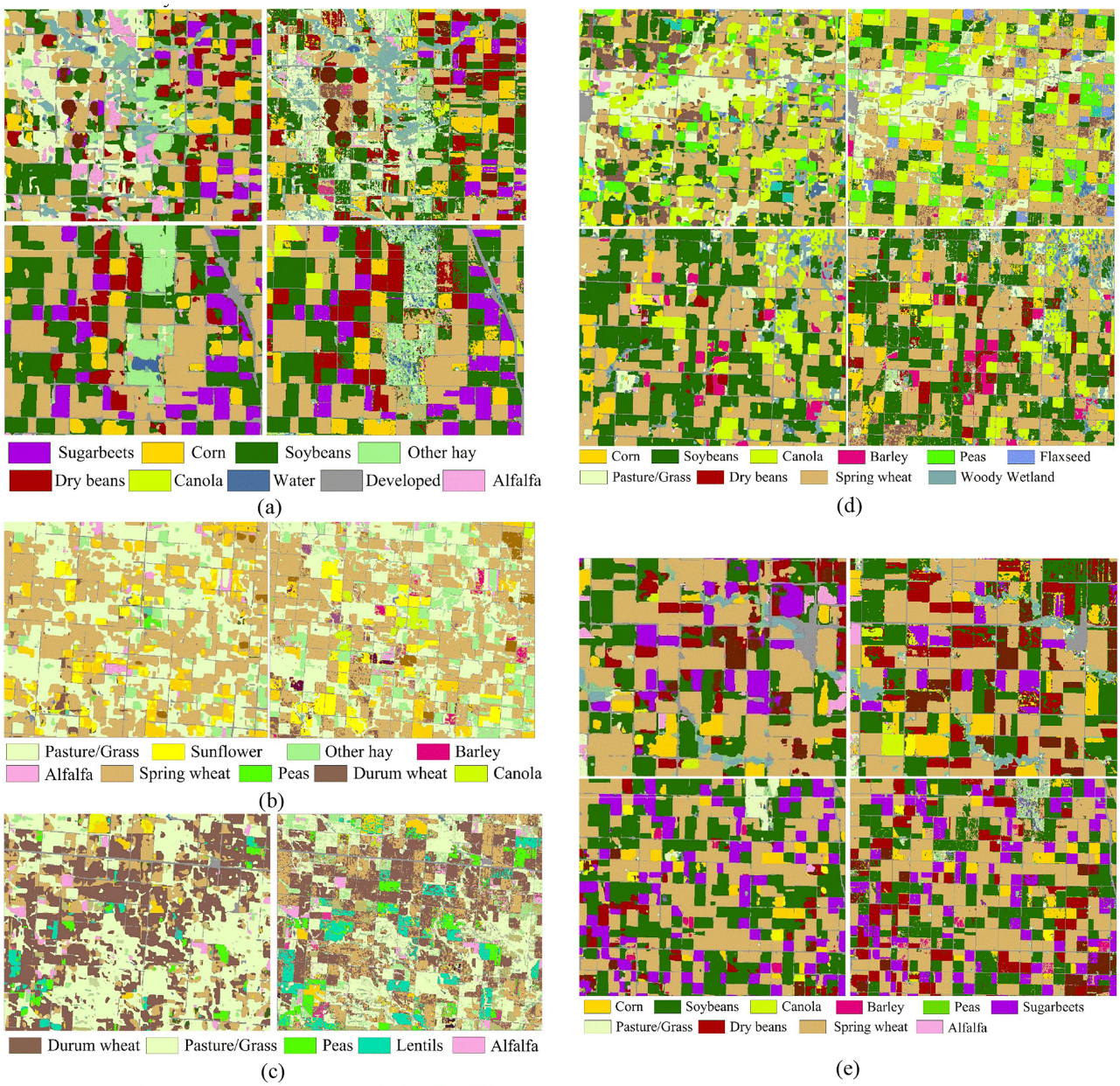

Fig. 13. Comparison between the DNN results (left) and CDL (right). (a) Herbaceous area, (b) sunflower, (c) durum wheat, (d) canola, (e) sugarbeets.

results will be. Reasonably speaking, the results should be more accurate if the corresponding crop has more area. This hypothesis is validated in the west region of Devils Lake, a traditional agricultural county where the majority of the land is covered by major crops. In Fig. 15(c), the major crops have very high accuracy, e.g., corn (85%), soybean (87%), sunflower (82%), barley (79%), spring wheat (93%), dry beans (92%), peas (86%), and grassland/pasture (90%). In this region, DNN still struggles in distinguishing alfalfa (53%) from grassland/pasture (32%). The fallow fields are not recognized very well in all three regions. Fallow lands are covered by wild vegetation such as alfalfa and weeds in the growing season and are acceptable being classified as alfalfa or grassland. We will study how to identify fallow farm fields from artificial alfalfa crop fields or real grassland in the future. Open water is misclassified into herbaceous wetland and we are comfortable with that disagreement as the water coverage is changing over the year and the source data used in CDL might be observed at a different date from the Landsat scene. We will study DL by taking advantage of these biological and optical features to improve accuracy. The herbaceous wetland has some misclassification to fallow land. Overall, DNN has high accuracy in major crops in big farms and may misclassify minority crops in the wetlands and suburban regions where the crop fields are scattered and have dramatic seasonal changes.

The final accuracy metrics are listed in Table II. The precision of the Missouri River basin is 0.8214, Devils Lake is 0.8463, and Lake Darling is 0.8241. These metrics are very impressive considering the complicated crop hierarchy and the huge number

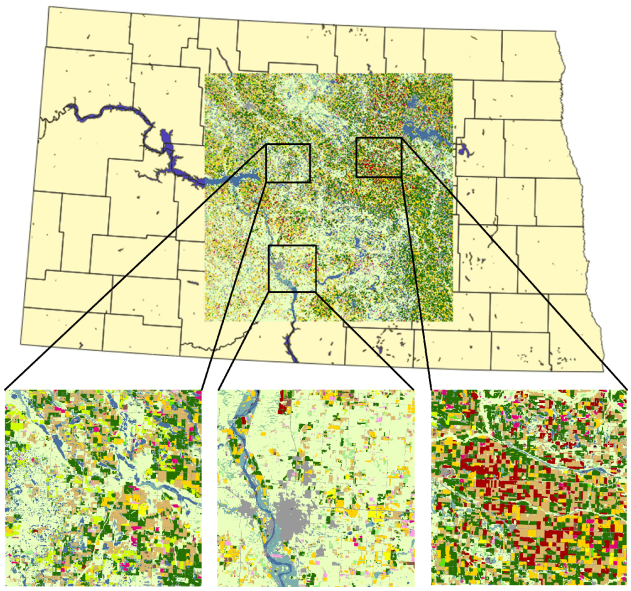

Fig. 14. Validation region (Lake Darling, Missouri River Basin, and Devils Lake).

TABLE II
ACCURACY METRICS OF DNN

Metrics	Missouri River Basin	Devils Lake	Lake Darling
Accuracy	0.8025	0.8265	0.7885
Precision	0.8214	0.8463	0.8241
Recall	0.828	0.8536	0.8691
Kappa	0.8166	0.8424	0.8056
F1	0.8232	0.8476	0.8439

of pixels in Landsat images. The high kappa and recall (>80%) indicate that the number of correctly classified crops by DNN covers most crops including the minority crops. The DNN result of Devils Lake has the highest value in the four metrics (except recall) which proves our conclusion (DNN accuracy is higher in big farmlands). The equations for the metrics are listed as follows (explanation can be found in [81]):

$$\text{Precision} = \frac{n_{\text{correct}}}{n_{\text{all predicted}}}$$

$$\text{Recall} = \frac{n_{\text{correct}}}{n_{\text{all real}}}$$

$$\text{kappa} = \frac{p_0 - p_e}{1 - p_e}$$

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

VII. DISCUSSION

A. Training Dataset Selection

NDVI and trusted pixel methods are used to get more samples from CDL to enrich the training dataset. NDVI is used

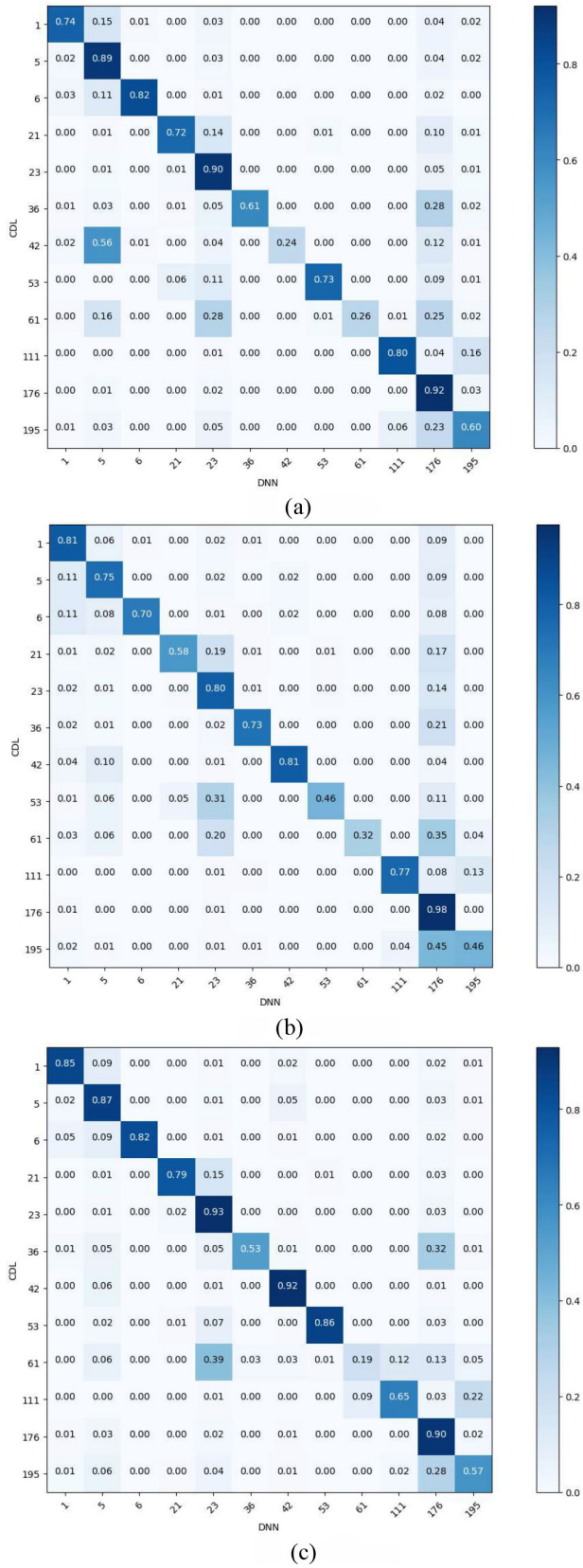

Fig. 15. Normalized confusion matrix between refined CDL and DNN 2017 result (corn-1, soybean-5, sunflower-6, barley-21, spring wheat-23, alfalfa-36, dry beans-42, peas-53, fallow-61, open water-111, grassland-176, and herbaceous wetland-195. The complete crop categories are in Appendix A). (a) Darling Lake. (b) Missouri River. (c) Devils Lake.

to automatically identify pixels with crop/vegetation-covered at the observation time to automatically select the appropriate Landsat scenes for classification. We extract the pixels from CDL and make them go through a series of filtering processes to remove the ambiguous pixels which have a high chance to be incorrect. The natural quality of Landsat pixels should be seen as a critical criterion to determine whether adding them into the training dataset. The edge pixels, boundary pixels, low-confident pixels, cloud/shadow pixels, should be eliminated to ensure all the training pixels are highly accurate. The rest pixels which have higher accuracy will be merged into the training dataset to permit the training of DNN. The underlying crop features will be more easily learned from SR if the crop samples are sufficiently available in the training dataset. According to the results, the multiple-scene-multiple-month-multiple-year training dataset is better than only using images observed in one scene, one year, or one month. In other words, "more is better."

B. Usability of DNN Crop Maps

DNN maps can reflect the distribution of major crops such as corn, soybean, spring wheat, sugar beets, barley, sunflower, alfalfa, etc. The major crop distribution can be directly used for agricultural activity advisory. For minority crops, DNN maps are still incapable of producing a reliable judgment, especially for wetland and suburban regions. The results of minority crops such as peas, canola, etc. are less accurate and can only be used as prediction results. The quality of DNN maps can be further enhanced by a series of postprocesses involving other data sources to force correct those misclassified fields, which we will continue work in the future.

C. Reusability of Trained DNN Model

One major advantage of DL models over other machine learning models is high reusability. However, reusing a trained DNN model is still subject to spatial and temporal limitations. The spatial extent is limited to the coverage region of training samples. It means the DNN model trained in North Dakota may not be applicable in other agriculture states such as Ohio. Observation dates of the input Landsat images also matter for DNN. If the training images are majorly observed in July, the trained models will not perform well on the images captured in May. To eliminate these limitations, expanding the spatiotemporal coverage of the training dataset might be the only solution.

D. Integration of DNN and Cyberinfrastructure

The learning process is always back and forth, and the overall duration of training is long. Each training epoch takes about two hours on two K80 GPUs. We normally run 30 epochs which take about two and a half days. Each prediction process takes around 30 min. Also, the preprocessing will cost half an hour. Merging all the results into one map will take another half an hour. We used Geoweaver to chain all these processes into a highly automated workflow. The sequential process on all 21 scenes flying over North Dakota will take no more than one day. The parallel process will take no more than two hours. In other words, with the help of Geoweaver, this DNN solution is capable of making in-season maps available only less than two days after the retrieval of Landsat scenes. The workflow setup can also scale when the used dataset outpaces the initial experimental database. DL can usually go "deeper" into the datasets by consuming more computational resources. Making DL work in a distributed high-performance computing environment is a major requirement of scalable DL. This research uses Geoweaver to manage distributed resources and allow users to work across multiple nodes. Integration with advanced cyberinfrastructure is a big step forward for realizing large-scale crop mapping.

VIII. Conclusion and Future Work

There are constantly unexpected problems during applying a new technique in a new domain, especially when there are more than one hundred classes and the study region contains multiple scenes. Besides the algorithm itself, the inputs and the training/testing methods have significant impacts on the algorithm performance. This article attempts to explicitly describe those problems and address them using the proposed workflow to ensure the DL model could output what is desired. A novel full-stack workflow using deep CNNs and cyberinfrastructure is proposed to learn the underlying features of various crops and use the trained model to produce in-season crop maps from Landsat images. Geoweaver is used to manage the learning workflow in a distributed environment. We tested the trained DNN model in North Dakota and got satisfying results on major crops such as corn, soybean, barley, spring wheat, alfalfa, etc. The accuracy in all three test regions is higher than 82%. Both precision and recall of DNN are very competitive comparing to CDL. The DNN model is reusable in new Landsat images of North Dakota and could easily scale to bigger regions (e.g., the other corn belt states) via cyberinfrastructure. This study discovers that the DNN model will be more accurate if the training dataset includes more Landsat images from multiple scenes, months, and years. The trained DNN model can better recognize major crops in big farms but will struggle in differentiating minority crops in wetland and suburban regions.

In the future, we will continue to leverage the existing big data facilities to expand the spatiotemporal coverage of the training dataset on both major crops and minority crops to produce large-scale in-season crop maps. The current model is still underfitting and we will integrate more high-performance computational platforms to collaborate on training to further improve its performance. We will study how to identify fallow farm fields from artificial alfalfa crop fields or real grassland. Improving the accuracy of minority crops will be another important direction. The quality of DNN maps can be further enhanced by a series of post-processes involving other data sources to force correct those misclassified fields, which we will continue to study in our next stage of work.

Acknowledgment

The authors would like to thank J. Gaigalas for proofreading the manuscript.

REFERENCES

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

[2] S. K. Seelan, S. Laguette, G. M. Casady, and G. A. Seielstad, "Remote sensing applications for precision agriculture: A learning community approach," *Remote Sens. Environ.*, vol. 88, pp. 157–169, 2003.

[3] P. J. Pinter Jr. *et al.*, "Remote sensing for crop management," *Photogrammetric Eng. Remote Sens.*, vol. 69, pp. 647–664, 2003.

[4] L. Di, G. Y. Eugene, L. Kang, R. Shrestha, and Y.-Q. Bai, "RF-CLASS: A remote-sensing-based flood crop loss assessment cyber-service system for supporting crop statistics and insurance decision-making," *J. Integrative Agriculture*, vol. 16, pp. 408–423, 2017.

[5] M. A. Friedl *et al.*, "Global land cover mapping from MODIS: algorithms and early results," *Remote Sen. Environ.*, vol. 83, pp. 287–302, 2002.

[6] X. Zhang *et al.*, "Monitoring vegetation phenology using MODIS," *Remote Sen. Environ.*, vol. 84, pp. 471–475, 2003.

[7] F. Gao *et al.*, "Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery," *Remote Sen. Environ.*, vol. 188, pp. 9–25, 2017.

[8] C. Boryan, Z. Yang, R. Mueller, and M. Craig, "Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program," *Geocarto Int.*, vol. 26, pp. 341–358, 2011.

[9] W. Han, Z. Yang, L. Di, and R. Mueller, "CropScape: A web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support," *Comput. Electron. Agriculture*, vol. 84, pp. 111–123, 2012.

[10] H. Xiang and L. Tian, "An automated stand-alone in-field remote sensing system (SIRSS) for in-season crop monitoring," *Comput. Electron. Agriculture*, vol. 78, pp. 1–8, 2011.

[11] Z. Sun, P. Yue, and L. Di, "GeoPWTManager: A task-oriented web geoprocessing system," *Comput. Geosci.*, vol. 47, pp. 34–45, 2012.

[12] Z. Sun *et al.*, "Automation of customized and near-real-time vegetation condition index generation through cyberinfrastructure-based geoprocessing workflows," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 11, pp. 4512–4522, Nov. 2014.

[13] Z. Sun, L. Di, A. Chen, P. Yue, and J. Gong, "The use of geospatial workflows to support automatic detection of complex geospatial features from high resolution images," in *Proc. 2nd Int. Conf. Agro-Geoinform. (Agro-Geoinform.)*, 2013, pp. 159–162.

[14] Z. Sun and P. Yue, "The use of Web 2.0 and geoprocessing services to support geoscientific workflows," in *Proc. 18th Int. Conf. Geoinform.*, 2010, pp. 1–5.

[15] P. Yue *et al.*, "GeoPW: Laying blocks for the geospatial processing web," *Trans. GIS*, vol. 14, pp. 755–772, 2010.

[16] Z. Sun, H. Fang, L. Di, P. Yue, X. Tan, and Y. Bai, "Developing a web-based system for supervised classification of remote sensing images," *GeoInformatica*, vol. 20, pp. 1–21, 2016.

[17] Z. Sun, H. Fang, L. Di, and P. Yue, "Realizing parameterless automatic classification of remote sensing imagery using ontology engineering and cyberinfrastructure techniques," *Comput. Geosci.*, vol. 94, pp. 56–67, 2016.

[18] Z. Sun, H. Fang, M. Deng, A. Chen, P. Yue, and L. Di, "Regular shape similarity index: A novel index for accurate extraction of regular objects from remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3737–3748, Jul. 2015.

[19] M. C. You, Z. Sun, L. Di, and Z. Guo, "A web-based semi-automated method for semantic annotation of high schools in remote sensing images," in *Proc. 3rd Int. Conf. Agro-Geoinform.*, 2014, pp. 1–4.

[20] A. J. Surkan and L. Di, "Fast trainable pattern classification by a modification of Kanerva's SDM model," in *Proc. Int. Joint Conf. Neural Netw.*, 1989, pp. 347–349.

[21] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.

[22] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez, "A review on deep learning techniques applied to semantic segmentation," 2017, *arXiv:1704.06857.*

[23] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[24] Z. Sun *et al.*, "CyberConnector: a service-oriented system for automatically tailoring multisource Earth observation data to feed Earth science models," *Earth Sci. Inform.*, vol. 11, pp. 1–17, 2017.

[25] Z. Sun and L. Di, "Geoweaver: a web-based prototype system for managing compound geospatial workflows of large-scale distributed deep networks," *ESIP*, 30 Jan. 2019, doi: 10.6084/m9.figshare.7629491.v1.

[26] Z. Sun, "Some Basics of Deep Learning in Agriculture," *ESIP*, 30 Jan. 2019, doi: 10.6084/m9.figshare.7631615.v1.

[27] Z. Sun, L. Di, A. Burgess, J. A. Tullis, and A. B. Magill, "Geoweaver: Advanced cyberinfrastructure for managing hybrid geoscientific AI workflows," *ISPRS Int. J. Geo-Inf.*, vol. 9, 2020, Art. no. 119.

[28] S. Fritz *et al.*, "Mapping global cropland and field size," *Global Change Biol.*, vol. 21, pp. 1980–1992, 2015.

[29] S. Estel, T. Kuemmerle, C. Alcántara, C. Levers, A. Prishchepov, and P. Hostert, "Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series," *Remote Sens. Environ.*, vol. 163, pp. 312–325, 2015.

[30] D. Roy and L. Yan, "Robust Landsat-based crop time series modelling," *Remote Sens. Environ.*, vol. 238, 2018, Art. no. 110810.

[31] H. Müller, P. Rufin, P. Griffiths, A. J. B. Siqueira, and P. Hostert, "Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape," *Remote Sens. Environ.*, vol. 156, pp. 490–499, 2015.

[32] J. Xiong *et al.*, "Automated cropland mapping of continental Africa using Google Earth Engine cloud computing," *ISPRS J. Photogrammetry Remote Sens.*, vol. 126, pp. 225–244, 2017.

[33] P. Teluguntla *et al.*, "A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform," *ISPRS J. Photogrammetry Remote Sens.*, vol. 144, pp. 325–340, 2018.

[34] 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. 14, no. 5, pp. 778–782, May 2017.

[35] Z. Sun, L. Di, and H. Fang, "Using long short-term memory recurrent neural network in land cover classification on Landsat and Cropland data layer time series," *Int. J. Remote Sens.*, vol. 40, pp. 593–614, 2018.

[36] F. Waldner, G. S. Canto, and P. Defourny, "Automated annual cropland mapping using knowledge-based temporal features," *ISPRS J. Photogrammetry Remote Sens.*, vol. 110, pp. 1–13, 2015.

[37] K. Were, D. T. Bui, Ø. B. Dick, and B. R. Singh, "A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape," *Ecological Indicators*, vol. 52, pp. 394–403, 2015.

[38] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234–241.

[39] V. Badrinarayanan, A. Handa, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling," 2015, *arXiv:1505.07293.*

[40] V. Iglovikov and A. Shvets, "TernausNet: U-net with VGG11 encoder pretrained on imagenet for image segmentation," 2018, *arXiv:1801.05746.*

[41] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2015, pp. 3431–3440.

[42] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 807–814.

[43] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167.*

[44] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[45] J. T. Springenberg, "Unsupervised and semi-supervised learning with categorical generative adversarial networks," 2015, *arXiv:1511.06390.*

[46] M. D. Zeiler, "ADADELTA: an adaptive learning rate method," 2012, *arXiv:1212.5701.*

[47] O. Goksel *et al.*, "Overview of the VISCERAL challenge at ISBI," in *Proc. VISCERAL Challenge@ ISBI*, 2015, pp. 6–11.

[48] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," 2016, *arXiv:1603.07285.*

[49] J. Nagi *et al.*, "Max-pooling convolutional neural networks for vision-based hand gesture recognition," in *Proc. IEEE Int. Conf. Signal Image Process. Appl.*, 2011, pp. 342–347.

[50] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in *Proc. IEEE Int. Conf. Comput. Vision*, 2015, pp. 1520–1528.

[51] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2017.

[52] C. B. Flora, "Social capital and sustainability: Agriculture and communities in the Great Plains and Corn Belt," *Res. Rural Sociology Develop.*, vol. 6, pp. 227–246, 1995.
[53] USDA. National agriculture imagery program. (2003). [Online]. Available: https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/
[54] NorthDakotaGIS. North Dakota GIS hub data portal. (2012). [Online]. Available: https://gishubdata.nd.gov/
[55] C. W. Lee, "NDSU PLSC 210 - horticulture science - plant classification," 2010. [Online]. Available: https://www.ndsu.edu/pubweb/chiwonlee/plsc210/topics/chap2-classification/classification.html
[56] NDSU. *Archive - Ag News From NDSU*. (2020). [Online]. Available: https://www.ag.ndsu.edu/news/topics/
[57] EOS. NDVI FAQ: All you need to know about NDVI. (2019). [Online]. Available: https://eos.com/blog/ndvi-faq-all-you-need-to-know-about-ndvi/
[58] J. G. Masek *et al.*, "A landsat surface reflectance dataset for North America, 1990–2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 68–72, Jan. 2006.
[59] E. Vermote, C. Justice, M. Claverie, and B. Franch, "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product," *Remote Sens. Environ.*, vol. 185, pp. 46–56, 2016.
[60] D. P. Roy *et al.*, "Landsat-8: Science and product vision for terrestrial global change research," *Remote Sens. Environ.*, vol. 145, pp. 154–172, 2014.
[61] N. Flood, "Continuity of reflectance data between Landsat-7 ETM+ and Landsat-8 OLI, for both top-of-atmosphere and surface reflectance: A study in the Australian landscape," *Remote Sens.*, vol. 6, pp. 7952–7970, 2014.
[62] Z. Zhu, C. E. Woodcock, C. Holden, and Z. Yang, "Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time," *Remote Sens. Environ.*, vol. 162, pp. 67–83, 2015.
[63] M. Feng, C. Huang, S. Channan, E. F. Vermote, J. G. Masek, and J. R. Townshend, "Quality assessment of Landsat surface reflectance products using MODIS data," *Comput. Geosci.*, vol. 38, pp. 9–22, 2012.
[64] A. Kamilaris and F. X. Prenafeta-Boldu, "Deep learning in agriculture: A survey," *Comput. Electron. Agriculture*, vol. 147, pp. 70–90, 2018.
[65] M. Dyrmann, H. Karstoft, and H. S. Midtiby, "Plant species classification using deep convolutional neural network," *Biosyst. Eng.*, vol. 151, pp. 72–80, 2016.
[66] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, 2015, Art. no. 1.
[67] Y. Bengio, "Deep learning of representations: Looking forward," in *Proc. Int. Conf. Statist. Lang. Speech Process.*, 2013, pp. 1–37.
[68] OSGEO. GDAL-OGR: Geospatial Data Abstraction Library/Simple Features Library Software. (2008). [Online]. Available: http://www.gdal.org
[69] C. Homer *et al.*, "Completion of the 2011 national land cover database for the conterminous United States–representing a decade of land cover change information," *Photogrammetric Eng. Remote Sens.*, vol. 81, pp. 345–354, 2015.
[70] M. Haklay and P. Weber, "Openstreetmap: User-generated street maps," *IEEE Pervasive Comput.*, vol. 7, no. 4, pp. 12–18, Oct./Dec. 2008.
[71] M. Story and R. G. Congalton, "Accuracy assessment: A user's perspective," *Photogrammetric Eng. Remote Sens.*, vol. 52, pp. 397–399, 1986.
[72] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sens. Environ.*, vol. 80, pp. 185–201, 2002.
[73] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017.
[74] E. Amazon, *Amazon Elastic Block Store (EBS)*. Seattle, WA, USA: Amazon Web Services Inc., 2013.
[75] J. Varia and S. Mathew,"Overview of Amazon web services," 2014. [Online]. Available: http://cabibbo.dia.uniroma3.it/asw-2014-2015/altrui/AWS_Overview.pdf
[76] Z. Sun, L. Di, and J. Gaigalas, "SUIS: Simplify the use of geospatial web services in environmental modelling," *Environ. Model. Softw.*, vol. 119, pp. 228–241, 2019.
[77] J. Gaigalas, L. Di, and Z. Sun, "Advanced cyberinfrastructure to enable search of big climate datasets in THREDDS," *ISPRS Int. J. Geo-Inf.*, vol. 8, 2019, Art. no. 494.
[78] Z. Sun, L. Di, B. Cash, and J. Gaigalas, "Advanced cyberinfrastructure for intercomparison and validation of climate models," *Environ. Model. Softw.*, vol. 123, 2019, Art. no. 104559.
[79] Z. Sun *et al.*, "Advanced cyberinfrastructure for agricultural drought monitoring," in *Proc. 8th Int. Conf. Agro-Geoinform. (Agro-Geoinform.)*, 2019, pp. 1–5.
[80] Z. Sun, P. Yue, L. Hu, J. Gong, L. Zhang, and X. Lu, "GeoPWProv: Interleaving map and faceted metadata for provenance visualization and navigation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 11, pp. 5131–5136, Nov. 2013.
[81] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, pp. 35–46, 1991.

Ziheng Sun received the Ph.D. degree from Wuhan University, in 2015.

He is a Research Assistant Professor with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA, USA. He is a practitioner of using the cutting-edge technologies including artificial intelligence and high-performance computing to answer critical scientific questions in geosciences. He has published 12 first-author papers on international journals. He invented RSSI, a novel index for artificial object recognition from high-resolution aerial images, and proposed parameterless automatic classification solution for reducing the parameter-tuning burden on scientists. He used long short-term memory neural network to learn the seasonal changes to better recognize crops. He has successfully built a number of geospatial cyberinfrastructure systems for better disseminating, processing, visualizing, and understanding big spatial data.

Liping Di (Senior Member, IEEE) received the Ph.D. degree from University of Nebraska–Lincoln, in 1991.

He is a Professor and the Director with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA, USA. He is internationally known for his extraordinary contributions to the geospatial information science/geoinformatics, especially to the development of geospatial interoperability technology and the federal, national, and international geographic information and remote sensing standards. He is the pioneer in agro-geoinformatics and in the development of web-based, advanced, distributed geospatial systems and tools. He has engaged in the geoinformatics and earth system research for more than 30 years and has authored or coauthored over 450 publications. He has received a total of over $55 million research grants from federal agencies, international organizations, and private companies.

Hui Fang received the M.S. degree in geographic information system from Wuhan University, Wuhan, China, in 2011.

She used to work in the industry to develop remote sensing image processing module in commercial geographical information systems. Her research interests include remote sensing, image classification, computer vision, and forestry science.

Annie Burgess received the Ph.D. degree from University of Utah, in 2013.

She is the Director of ESIP Lab, Boulder, CO, USA. Her career has bridged Earth Science and informatics. During her graduate work with the University of Utah, Salt Lake City, UT, USA, she managed, analyzed, and distributed an immense amount of data related to her research in snow hydrology. The primary data product created during her Ph.D. work is currently distributed through NASA/JPL. As a Postdoc with the University of Southern California, Los Angeles, CA, USA, she developed software for the unique needs of the Polar Science Community. She knows the power of connecting earth scientists with technical and collaborative infrastructure. At ESIP, she utilizes her technical savvy and networking skills to run their innovation program, also known as the ESIP Lab.