Iterative Deep Learning (IDL) for agricultural landscape classification using fine spatial resolution remotely sensed imagery

Huapeng Li [a] *, Ce Zhang [b] *, Shuqing Zhang [a], Xiaohui Ding [c], Peter M. Atkinson [d]

[a] *Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130012, China*

[b] *Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK*

[c] *Guangzhou Institute of Geography, Guangzhou 510070, China*

[d] *Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK*

Abstract

The agricultural landscape can be interpreted at different semantic levels, such as fine low-level crop (LLC) classes (e.g., Wheat, Almond, and Alfalfa) and broad high-level crop (HLC) classes (e.g., Winter crops, Tree crops, and Forage). The LLC and HLC are hierarchically correlated with each other, but such intrinsically hierarchical relationships have been overlooked in previous crop classification studies in remote sensing. In this research, a novel Iterative Deep Learning (IDL) framework was proposed for the classification of complex agricultural landscapes using remotely sensed imagery. The IDL adopts an object-based convolutional neural network (OCNN) as the basic classifier for both the LLC and HLC classifications, which has the advantage of maintaining precise crop parcel boundaries. In IDL, the HLC classification implemented by the OCNN is

* Corresponding author.
E-mail addresses: lihuapeng@iga.ac.cn (H.P. Li), c.zhang9@lancaster.ac.uk (C. Zhang).

conditional upon the LLC classification probabilities, whereas the HLC probabilities combined with the original imagery are, in turn, re-used as inputs to the OCNN to enhance the LLC classification. Such an iterative updating procedure forms a Markov process, where both the LLC and HLC classifications are refined and evolve collaboratively. The effectiveness of the IDL was tested on two heterogeneous agricultural fields using fine spatial resolution (FSR) SAR and optical imagery. The experimental results demonstrate that the iterative process of IDL helps to resolve contradictions within the class hierarchies. The new proposed IDL consistently increased the accuracies of both the LLC and HLC classifications with iteration, and achieved the highest accuracies for each at four iterations. The average overall accuracies were 88.4% for LLC and 91.2% for HLC, for both study sites, far greater than the accuracies of the state-of-the-art benchmarks, including the pixel-wise CNN (81.7% and 85.9%), object-based image analysis (OBIA) (84.0% and 85.8%), and OCNN (84.0% and 88.4%). To the best of our knowledge, the proposed model is the first to identify and use the relationship between the class levels in an ontological hierarchy in a remote sensing classification process. It is applied here to increase progressively the accuracy of classification at two levels for a complex agricultural landscape. As such IDL represents an entirely new paradigm for remote sensing image classification. Moreover, the promising results demonstrate the great potential of the proposed IDL with wide application prospect.

Keywords: Image classification; hierarchical crop classification; iterative deep learning; object-based image analysis (OBIA); convolutional neural network (CNN)

1. Introduction

Food demand is projected to increase by about 50% between 2012 and 2050 in response to global population growth and this poses a great challenge for food production (Alexandratos and Bruinsma, 2012). To cope with such a challenge, a wide range of information on agricultural practices and variables needs to be provided at national-to-global scales, and in a timely manner. Information on crop types, including their spatial distribution, is key to supporting decision-making to reduce local and national food insecurity and to promote agricultural economic development. For example, crop mapping data are required as a base input to support forecasting of agricultural production, which is commonly needed to forecast the potential for and, ultimately avoid, famine (Mkhabela et al., 2011). Data on crop type and their spatial distribution are essential to forecast crop prices and, thus, develop reasonable agricultural subsidy policies (e.g., food aid) (Zhao et al., 2020). In addition, such crop type data are vitally important for a variety of environmental research. For example, crop type is a fundamental input to greenhouse gas (GHG) emission models in view of the great differences in soil carbon flux between crop types (Pena-Barragan et al., 2011).

Remote sensing is an efficient tool for crop classification and mapping due to its synoptic and timely repeat coverage, and cost-effective methodology relative to ground survey. While a number of studies have explored the available medium spatial resolution remote sensing images (such as MODIS, Landsat, CBERS) for crop mapping at a relatively large scale (e.g., Wardlow and Egbert, 2008; Dong et al., 2016; Zhong et al., 2019), parcel-scale detailed maps of crop distributions that are essential for precision agriculture are needed (e.g., fertilization, irrigation, and management) (Zhang et al., 2012). With technological advances, a very large number of fine spatial resolution (FSR) images (e.g., RapidEye, Quickbird, and Gaofen) is now available, providing excellent opportunities to characterise crop type in great detail (Duro et al., 2012; Li et al., 2019a). However, with an increase in spatial resolution, the spectral and spatial variance for a single crop type tends to increase markedly (Li et al., 2019a). The large variances may be further exaggerated by diversified farming practices (Azar et al., 2016), which makes crop type mapping from FSR imagery a very challenging task.

During the last few decades, a number of crop classification methods have been developed for FSR remotely sensed imagery. These approaches can be generalised into two broad categories according to the underlying processing unit: pixel-based and object-based. Pixel-based methods classify crop types based on spectral (or polarimetric) signatures purely without considering the rich spatial information in FSR imagery, and they often achieve limited classification accuracy because of "salt and pepper" noise (Duro et al., 2012; Li et al., 2019b). To overcome these issues, object-based image analysis (OBIA) methods have been developed based on segmented objects (c.f. pixels) (Blaschke, 2010), and are now adopted extensively for crop mapping and classification (Castillejo-Gonzalez et al., 2009; Peña-Barragán et al., 2011; Jiao et al., 2014). These object-based methods utilise not only the within-object information (e.g., spectra and texture), but also contextual information between objects (e.g., the relationship between adjacent objects), thereby achieving increased classification accuracy (Castillejo-Gonzalez et al., 2009; Li et al., 2019b). However, the features employed in OBIA methods are essentially hand-crafted, and their quality depends heavily on individual user expertise and experience (Zhang et al., 2020b). OBIA methods are also challenged by selecting key variables from huge number of object features (Duro et al., 2012).

A major challenge with traditional methods (both pixel-based and object-based) is that they are unable to extract deep-level features from remotely sensed imagery automatically and effectively. Recently, deep learning, which can learn discriminative features in an end-to-end manner, has attracted considerable interest in a variety of research fields (LeCun et al., 2015). Deep convolutional neural networks (CNN), one of the most popular and successful deep learning methods, have demonstrated significant advantages for image processing and analysis (Krizhevsky et al., 2017). Owing to their excellent capability to learn higher-level feature representations, CNNs have achieved impressive results beyond the state-of-the-art in a variety of research fields, such as speech detection (Hinton et al., 2012), image denoising (Zhang et al., 2017) and handwriting recognition (LeCun et al., 2015). Meanwhile, CNNs have also achieved success in remote sensing, such as for object detection (Cheng et al., 2016), panchromatic image sharpening (Scarpa et al., 2018) and remote sensing image classification (Zhang et al., 2018). CNNs have demonstrated huge potential for classifying agricultural landscapes that are spatially and temporally heterogeneous using FSR imagery. Yao et al. (2017) presented a CNN-based approach for crop classification with FSR remote sensing images. Sidike et al. (2019) developed a novel deep progressively expanded network (dPEN) to map crop types and crop residues from FSR WorldView-3 imagery. Li et al. (2020) applied a CNN-transformer approach to perform crop classification using multi-temporal images. Zhang et al. (2020a) recently designed a modified pyramid scene parsing network (MPSPNet) to identify crop areas from FSR images. These pioneering methods, however, only classify the cropland using remotely sensed images, and they overlook the close relationship between crop hierarchies which has proven to be very beneficial to crop classification.

Some previous studies have attempted to incorporate the domain knowledge via a hierarchy of classes into crop mapping. La Rosa et al. (2019) presented the Most Likely Class Sequence (MLCS) post-processing algorithm to incorporate prior knowledge about crop dynamics into crop mapping using a binary transition probability matrix. Martinez et al. (2021) recently adopted the MLCS to enforce prior knowledge about crops' dynamics to the crop classification results of convolutional recurrent networks. Similarly, Giordano et al. (2020) refined crop classification results with crop rotation rules acquired based on previous classification maps. However, these approaches only exploit prior crop rotation knowledge that is local experience-dependent (via temporal hierarchy of classes) for crop mapping, and they are, thus, hard to generalise to other regions. Currently, very few studies have focused on the exploitation of hierarchical ontologies knowledge (via compositional hierarchy of crop classes). In fact, the agricultural landscape can be interpreted at multiple semantic levels (Wardlow and Egbert, 2008). For example, an agricultural landscape might be categorised as summer crops and winter crops at a high-level (i.e., coarse, broad-level), and divided further into corn, sunflower, wheat and oats at a low-level (i.e., fine, detailed-level) (Peña-Barragán et al., 2011). The low-level crop (LLC) and high-level crop (HLC) classes have the same spatial extent and are nested within each other hierarchically. Thus, there is a close, hierarchical relationship between these classes. However, it is still not yet clear whether the relationship between compositional hierarchies can be used to enhance crop classification accuracies.

To fill this knowledge gap, a novel Iterative Deep Learning (IDL) approach that is capable of learning discriminative features and utilising the relationship between different crop class levels, was proposed in this paper to solve progressively the problem of classifying complex agricultural landscapes. In IDL, the agricultural landscape is interpreted at two semantic levels, namely fine low-level crop (LLC) classes and broad high-level crop (HLC) classes. The LLC and HLC are classified using an object-based CNN (OCNN) to maintain the boundary of the crop parcels. A Markov process is formulated in the IDL to progressively and iteratively model the joint distribution between the predicted LLC and HLC variables. During the iterative progress, the LLC and HLC classifications interact with and complement each other, thus, increasing their accuracies. To the best of our knowledge, this is the first attempt to classify automatically a complex agricultural landscape using deep learning by considering hierarchical ontologies in relation to the crop system. The proposed IDL method was tested over two heterogeneous agricultural fields, respectively, using FSR Synthetic Aperture Radar (SAR) and optical imagery.

2. Methods

2.1 Convolutional neural network (CNN)

A CNN is intrinsically a deep neural network consisting of several pairs of convolutional and pooling layers (i.e., hidden layers). The convolutional layer is adopted to extract multi-level feature representations through convolutional filters, followed by an activation function to enhance non-linearity. The max-pooling layer is employed to strengthen the generalisation ability of the network. The parameters of the CNN network (i.e., weights and biases) are learnt using a stochastic gradient descent algorithm. Finally, one or more fully connected layers is employed on top of the last max-pooling layer, with a softmax function being included to predict the final classification results.

2.2 Object-based convolutional neural network (OCNN)

The OCNN was developed by Zhang et al. (2018) to allow application of the CNN to FSR imagery for land use classification, while maintaining the geometric integrity of ground *objects* and enhancing computational efficiency. The OCNN places an image patch at the centroid of each object to extract multi-level feature representations for prediction (Li et al., 2019b). While employing the same training process as the standard pixel-wise CNN using labelled image patches, the prediction of the OCNN model is assigned to each segmented object acquired from remotely sensed imagery.

2.3 Iterative Deep Learning (IDL) model

An agricultural landscape can be interpreted as comprising low-level crop (LLC) and high-level crop (HLC) classes arranged in a hierarchical ontological structure, as mentioned above. The basic assumption of the proposed IDL is that the LLC and HLC classifications are intrinsically correlated and complementary to each other. The general workflow of the proposed classification model is illustrated by Fig. 1, where LLC and HLC classifications are achieved jointly at each iteration, and they refine each other iteratively.

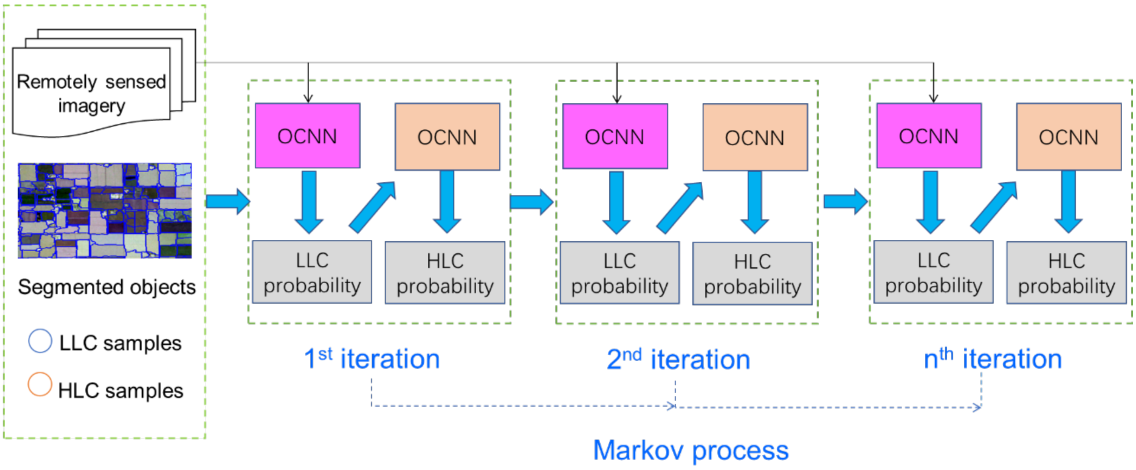

Fig. 1. General workflow of the proposed Iterative Deep Learning method for LLC and HLC classifications.

In the IDL model, the HLC classification probabilities are conditional upon the LLC classification probabilities within each iteration, and the joint probability distribution between LLC and HLC of the current iteration (i) is impacted by the probability distribution of the previous iteration. Such a hierarchical classification framework can be formulated as a Markov process as follows:

$$P\left(\mathrm{LLC}^{i}, \mathrm{HLC}^{i}\right) = P\left(\mathrm{LLC}^{i}, \mathrm{HLC}^{i} \middle| \mathrm{LLC}^{i-1}, \mathrm{HLC}^{i-1}\right) \tag{1}$$

where i represents the number of iterations within the Markov process, and LLC^{i} and HLC^{i} denote the LLC and HLC classifications at the i-th iteration, respectively. The LLC and HLC classifications were achieved by using two submodels of IDL (denoted as LLC-submodel and HLC-submodel) with the OCNN classifier.

Let **M** represent a scene of remote sensing imagery, with m and n denoting the number of classes for LLC and HLC, respectively. Let $\mathbf{O}$=$(o_1, o_2, \dots, o_j, \dots, o_u)$ represent the set of segmented objects from **M**, where o_j and u are the j-th object and the total number of objects, respectively. Let $\mathbf{T}_{\mathrm{LLC}}$ = ($t_{\mathrm{LLC}1}, t_{\mathrm{LLC}2}, \dots, t_{\mathrm{LLC}k}, \dots, t_{\mathrm{LLC}v}$) and $\mathbf{T}_{\mathrm{HLC}}$ =

$(t_{\mathrm{HLC}1}, t_{\mathrm{HLC}2}, \dots, t_{\mathrm{HLC}k}, \dots, t_{\mathrm{HLC}v})$ represent the set of training samples of LLC and HLC, respectively, where $t_{\mathrm{LLC}k}$ and $t_{\mathrm{HLC}k}$ are the k-th samples of the LLC and HLC, respectively, and v is the total number of samples. The $\mathbf{T}_{\mathrm{LLC}}$ and $\mathbf{T}_{\mathrm{HLC}}$ were employed to train the OCNN models to achieve the LLC and HLC classifications, respectively. Note that the samples contained in $\mathbf{T}_{\mathrm{LLC}}$ and $\mathbf{T}_{\mathrm{HLC}}$ are the same and the samples of a specific HLC class are constituted by samples of one or more LLC classes (e.g., HLC Forage samples may consist of LLC Alfalfa and Hay samples).

Suppose the hierarchical relationship between LLC and HLC can be expressed via a function f, and the classification probabilities of the LLC and HLC classifications can be represented as:

$$P\left(\mathrm{LLC}^{i}, \mathrm{HLC}^{i}\right) = f(\mathrm{LLC}^{i-1}, \mathrm{HLC}^{i-1}, \mathbf{M}, \mathbf{O}, \mathbf{T}_{\mathrm{LLC}}, \mathbf{T}_{\mathrm{HLC}}) \quad (2)$$

where LLC^{i-1} and HLC^{i-1} denote the LLC and HLC classification outputs of the previous (i.e., (i-1)-th) iteration, respectively; $\mathbf{M}$ and $\mathbf{O}$ are the original remotely sensed image and the set of object-based segmentations, respectively; $\mathbf{T}_{\mathrm{LLC}}$ and $\mathbf{T}_{\mathrm{HLC}}$ are the LLC and HLC samples in which the locations in the image and the corresponding class labels are recorded. These elements serve as the inputs of the IDL model, with the joint probability distribution between LLC and HLC as the output of the model.

The input to the LLC-submodel is remotely sensed imagery combined with the probabilities of the HLC classification from the previous iteration, whereas the HLC-submodel takes only the probabilities of LLC classification as the input evidence. The LLC and HLC classification probabilities and their output maps are elaborated in detail as follows:

(1) LLC classification probabilities

The original imagery $\mathbf{M}$ and the HLC classification probabilities output from the previous iteration $P\left(\mathrm{HLC}^{i-1}\right)$ are combined for LLC classification as:

$$\mathbf{M}_{\mathrm{LLC}}{}^{\boldsymbol{i}} = \mathrm{Concate}(\mathbf{M}, P\left(\mathrm{HLC}^{i-1}\right)) \quad (3)$$

where Concate denotes a function to concatenate the imagery **M** with the HLC classification probabilities $P\left(\mathrm{HLC}^{i-1}\right)$. In other words, the function combines spatially the bands contained in $P(\mathbf{X})^{i-1}$ with those in **M** as the input for the next iteration. For the case of $i = 1$, the $P\left(\mathrm{HLC}^{i-1}\right)$ are empty (NULL) and $\mathbf{M}_{\mathrm{LLC}}{}^{\boldsymbol{i}}$ is, thus, equivalent to the original imagery $\mathbf{M}$.

The OCNN model for LLC classification is trained using the LLC training samples ($\mathbf{T}_{\mathrm{LLC}}$) as follows:

$$\mathrm{OCNN}_{\mathrm{LLC}}^{i} = \mathrm{OCNN.Train}(\mathbf{M}_{\mathrm{LLC}}{}^{\boldsymbol{i}}, \mathbf{T}_{\mathrm{LLC}}) \quad (4)$$

The LLC classification probabilities $P(\mathrm{LLC}^{i})$ at the i-th iteration can be predicted using the trained OCNN model as follows:

$$P(\mathrm{LLC}^{i}) = \mathrm{OCNN}_{\mathrm{LLC}}^{i}.\mathrm{Predict}(\mathbf{M}_{\mathrm{LLC}}{}^{\boldsymbol{i}}, \mathbf{O}) \quad (5)$$

Note that the $P(\mathrm{LLC}^{i})$ has the same spatial size as the imagery **M**, and the dimensions of $P(\mathrm{LLC}^{i})$ are equal to the number of LLC classes, with each band of the $P(\mathrm{LLC}^{i})$ corresponding to probabilities of a specific LLC class.

(2) HLC classification probabilities

Different from the LLC-IDL, the HLC-submodel uses only the LLC classification probabilities as the inputs. The training of the HLC classifier is represented as follows:

$$\mathrm{OCNN}_{\mathrm{HLC}}^{i} = \mathrm{OCNN.Train}(P(\mathrm{LLC}^{i}), \mathbf{T}_{\mathrm{HLC}}) \quad (6)$$

The HLC classification probabilities are predicted using the trained OCNN model as follows:

$$P(\mathrm{HLC}^i) = \mathrm{OCNN}^i_{\mathrm{HLC}}.\mathrm{Predict}(P(\mathrm{LLC}^i), \mathbf{O}) \quad (7)$$

By using Eq. (5), the probability of being assigned to each HLC class for each segmented object is achieved within each iteration. Like the $P(\mathrm{LLC}^i)$, the spatial size of $P(\mathrm{HLC}^i)$ is the same as the extent of the original imagery $\mathbf{M}$. The dimension of $P(\mathrm{HLC}^i)$ is equal to the number of HLC classes, and each dimension corresponds to the probabilities of a specific HLC class.

The probabilities of LLC ($P(\mathrm{LLC}^i)$) and HLC ($P(\mathrm{HLC}^i)$) are updated at each iteration. The final LLC ($\mathbf{M}_{\mathrm{LLCresult}}$) and HLC ($\mathbf{M}_{\mathrm{HLCresult}}$) classification maps are achieved based on the probabilities output at the last iteration as follows:

$$\mathbf{M}_{\mathrm{LLCresult}} = \arg\max(P(\mathrm{LLC}^N)) \quad (8)$$

$$\mathbf{M}_{\mathrm{HLCresult}} = \arg\max(P(\mathrm{HLC}^N)) \quad (9)$$

where arg max is a function classifying each object of the imagery as the class with the maximum membership, and N is the maximum number of iterations for the IDL model.

The proposed Iterative Deep Learning model has three major advantages as follows:

1. Hierarchical classifications of LLC and HLC are achieved in an automatic way.
2. Both the LLC and HLC classifiers evolve collaboratively and classification accuracy is increased progressively.
3. The training samples applied for both of the submodels of IDL are essentially the same, without extra substantial sampling workload.

3. Experimental results and analysis

3.1 Study area and materials

In this research, two agricultural regions (S1 and S2) located in the centre of the Sacramento Valley, California were chosen as the study areas (Fig. 2). The agricultural systems of the Sacramento Valley are highly complex and heterogeneous in crop composition and, thereby, are ideal for evaluating the effectiveness of the proposed IDL method. The first study site (S1) is in Solano county, with ten dominant low-level (detailed-level) crop categories identified, namely Almond, Walnut, Alfalfa, Hay, Clover, Winter wheat (denoted as Wheat hereafter), Corn, Sunflower, Tomato and Pepper. The second study site (S2) is situated in Yolo county, consisting of nine low-level crop categories, including Almond, Walnut, Grass, Alfalfa, Wheat, Corn, Sunflower, Tomato and Cucumber. These low-level categories for both S1 and S2 can be aggregated into five high-level (broad-level) categories, namely Tree crops, Forage, Winter crops, Summer crops, and Vegetables and Fruits (denoted as Vegetables hereafter), as illustrated by Table 1.

Table 1

The high-level crop (HLC) classes with descriptions and the corresponding low-level crop (LLC) components.

HLC	Study site	Description	LLC
Tree crops	S1, S2	Permanent crops, woody structures, growing season: spring to fall.	Walnut, Almond
Forage	S1, S2	Permanent crops, herbaceous structures, growing season: spring to fall with several rounds of cuttings.	Alfalfa, Hay, Clover, Grass
Winter crops	S1, S2	Non-permanent crops, herbaceous structures, growing season: mid-fall to late-spring of the next year.	Winter wheat
Summer crops	S1, S2	Non-permanent crops, herbaceous structures, growing season: mid-spring to early-autumn.	Corn, Soybean, Sunflower

Vegetables	S1, S2	Non-permanent crops, herbaceous structures, growing season: mid-spring to late-summer.	Tomato, Pepper, Cucumber

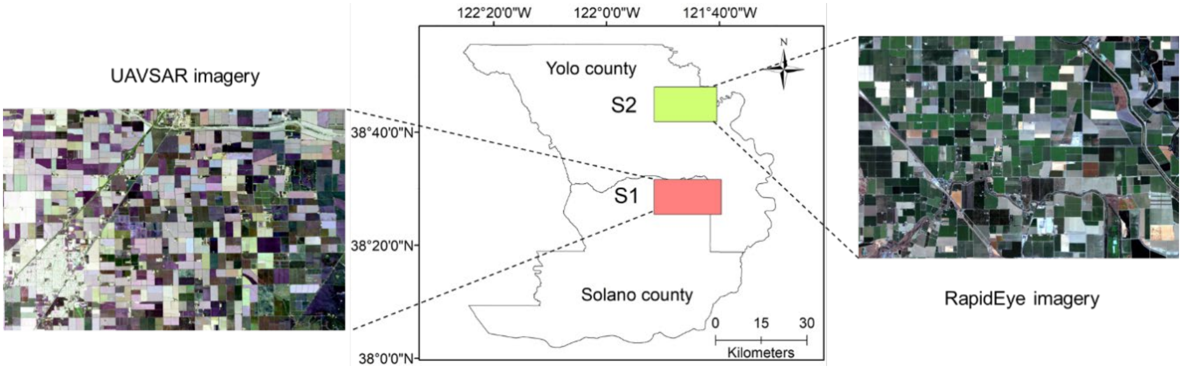

Fig. 2. Geographical locations of the two study areas with the corresponding remotely sensed images.

In S1, a scene of an Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR) image, an airborne L-band polarimetric radar system (Fore et al., 2015), was acquired on 29 Aug, 2011. The UAVSAR datasets employed were in the GRD format (calibrated and ground range projected) with a fine spatial resolution of 5 m and a spatial extent of 3474×2250 pixels. Three linear polarizations (i.e., HH, HV, and VV) as well as three polarimetric parameters (entropy, anisotropy and alpha angle) generated by the Cloude-Pottier decomposition were used as input variables to the classifiers.

In S2, a cloud-free RapidEye image with five spectral bands (blue, green, red, red edge, and near infrared), was captured on 10 July, 2016. The image employed in this research is a Level 3A Ortho product (i.e., sensor, radiometric and geometric correction already implemented) with a fine spatial resolution of 5 m and a spatial extent of 3222×2230 pixels. The atmospheric and topographic correction method was applied to the image to acquire surface reflectance for input to the image classifiers.

The Cropland Data Layer (CDL) of the United States Department of Agriculture (USDA) was employed as the ground sampling reference. The CDL is generated annually using medium spatial resolution images and a large number of ground samples (Boryan et al., 2011). Due to its very high quality, the CDL has been used widely as the ground reference in a variety of applications (e.g., Whelen and Siqueira, 2017; Cai et al., 2018; Li et al., 2019a). The crop parcels of both study sites were identified and delineated manually according to the CDL datasets. To acquire representative samples, crop parcels with an area below 5 ha were not considered (Li et al., 2019a). Training and validation sample points were collected within the separated training and validation polygons, respectively, using a stratified random sampling scheme to ensure they come from different crop polygons. A stratified random sampling procedure was adopted for sampling (training and validation), with the number of samples for a specific crop type being proportional to its total area. A total of 1415 and 1262 sample points were collected within S1 and S2, respectively, with an average of about 120 samples for each class. To evaluate comprehensively the classifications, wall-to-wall assessment was adopted for both sites. That is, all pixels within the testing polygons were used for accuracy assessment.

To further investigate the applicability of the presented IDL method, another scene of RapidEye image (Level 3A Ortho product) covering the first study site (denoted as S1') was collected on 14 Aug, 2014 for image classification. In S1', the county-level land use survey data in year 2014 by California Department of Water Resources (CDWR) were employed as ground reference to collect samples. Each land parcel within S1' was visited by staff in regional offices of CDWR, and the land use attributes (including specific crop type) were recorded during the visits (Zhong et al., 2019). The crop categories identified in S1' were exactly the same as S1, with a total of 1223 sample being collected for model training.

3.2 IDL model architecture and parameters

3.2.1 Image segmentation

The image segmentation procedure is the basis for the IDL since the LLC and HLC classifications are implemented on the segmented objects. A multi-resolution segmentation (MRS) algorithm (Baatz and Schaepe, 2000) was applied using the eCognition 9.0 software to acquire the segmented objects. Followed by the suggestions of Duro et al. (2012), the "scale" parameter was optimised first, and then the other two parameters (shape and compactness) were tuned successively, until the segmented objects matched well with crop boundaries based on visual inspection. The scale parameter of the MRS was tuned through cross-validation as 30 and 180 for S1 and S2, respectively, with slightly over-segmented results being achieved (i.e., the segmented objects are homogeneous). The Shape and Compactness parameters were optimised as 0.2 and 0.7 for S1, and 0.3 and 0.6 for S2. In total, 3040 and 3867 objects were generated for S1 and S2, respectively.

3.2.2 Model structure and parameters

In the proposed IDL model, a standard CNN classifier is applied to classify each segmented object (OCNN) at both the LLC and HLC classification levels, with the centroid of each object taken as the convolutional point (i.e., the centre of image patch) of the CNN (Zhang et al., 2018; Li et al., 2019b). The CNN within the IDL method needs to predefine hyperparameters to achieve the optimal classification results. Herein, the CNN was parameterised in S1 and directly generalized in S2, as detailed below.

The structure of the CNN employed in the IDL (denoted as CNN-IDL, hereafter) was similar to AlexNet with six hidden layers and small convolutional filter sizes (5×5 for the first convolutional layer and 3×3 for the remaining layers) (Fig. 3). The number of filters was tuned as 64 to extract multi-level feature representations for each segmented object. The input window size was optimised to 32×32. To alleviate the possibility for over-fitting problem, dropout regularization was applied with an optimised dropout value of 0.25. The maximum number of epochs was set to 500 to allow the network to converge through backpropagation. As the predicted CNN scores (i.e., probabilities) are often over-confident (Guo et al., 2017), the CNN-IDL model was calibrated during model training process with a label smoothing factor of 0.05 on validation set (Muller et al., 2019).

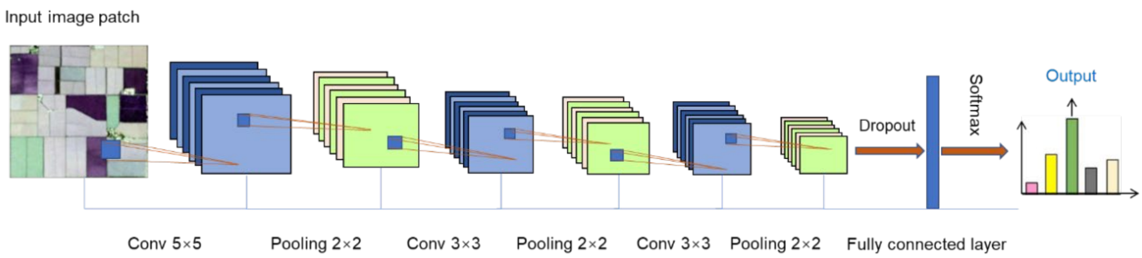

Fig. 3. Model architecture of the CNN network employed in the IDL model.

3.3 Benchmarks and parameter settings

To test comprehensively the effectiveness of the proposed IDL model, traditional object-based image analysis (OBIA), standard pixel-wise CNN (PCNN), and object-based CNN (OCNN) were applied as benchmarks. To provide a fair comparison, the structure of the two CNN-based benchmarks (i.e., PCNN and OCNN) was the same as that in the CNN-IDL network (i.e., three pairs of convolutional and max-pooling layers). Parameters including filter size, dropout value and epoch were also identical to those of the CNN-IDL. The three benchmarks are described briefly as follows:

OBIA: The OBIA was implemented based on the segmentation results achieved in Section 3.2.1. A range of hand-coded features were obtained from each segmented object, including spectral features, texture, and geometry. These hand-crafted feature representations were used as the input variables of a parameterised SVM classifier.

PCNN: The standard pixel-wise CNN classifies all pixels of the imagery using densely overlapping patches. The input window size of the PCNN was tuned as 24×24 through cross-validation for both study sites. The number of filters for each hidden layer was 32. The other control parameters were the same as for the CNN-IDL.

OCNN: Unlike the PCNN, the OCNN takes the segmented objects (Section 3.2.1) as the functional unit (Zhang et al., 2018b; Li et al., 2019b). A standard CNN was trained in the OCNN to predict the label of each object. Settings of the parameters were identical to those of the CNN-IDL.

3.4 Classification analysis and results

3.4.1 IDL classification accuracies

The presented IDL method was implemented 10 times (with 10 iterations in each implementation) for each study site to evaluate its accuracy and robustness. Fig. 4 plots the average overall accuracy (OA) of the IDL against iteration from iteration 1 to 10. It can be observed that the OAs of the LLC and HLC classifications in S1 started from 82.25% and 90.05%, respectively, then increased rapidly from iteration 2 to 3, and reached the greatest OAs of 87.94% and 91.83% at iteration 4 (Fig. 4(a)). The accuracies of both LLC and HLC tend to be stable (around 88% and 92%) after iteration 4 (i.e., from iteration 5 to 10), with the OA of HLC being higher than that of the LLC by about 4%. A similar trend of increasing accuracy with iteration was found for the second study site (S2) (Fig. 4(b)). Specifically, the OAs of the LLC and HLC classifications (from 84.90% and 88.66%, respectively) increased gradually with iteration until iteration 4, where the greatest OAs of 88.46% and 90.37% were achieved for LLC and HLC, respectively. The OAs of both LLC and HLC stabilised from iteration 5 to 10. The difference in accuracy between the LLC and HLC classifications in S2 was about 2%.

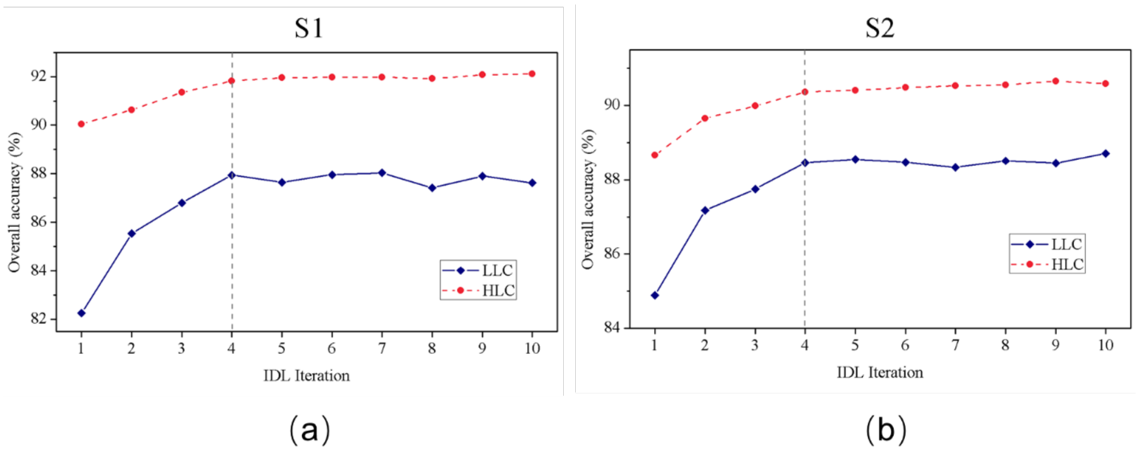

Fig. 4. Plots of overall accuracy achieved by the proposed IDL against iteration for both S1 and S2. The optimal accuracies of both LLC and HLC classifications are obtained by iteration 4 as indicated by the gray dashed line.

3.4.2 IDL classification results

Fig. 5. Two typical image subsets of the LLC and HLC classifications in S1 achieved using the LLC-submodel and HLC-submodel, respectively. Note that the red and yellow circles highlight incorrect and correct classifications, respectively.

To provide a visualization of how the two submodels of the IDL complement each other iteratively, typical subsets of the LLC and HLC classifications produced by the LLC-submodel (IDL-LLC) and HLC-submodel (IDL-HLC) are presented from iteration 1 to 4 for S1 and S2 in Figs. 5 and 6, respectively. Two typical subsets are illustrated for each of the study sites. For the first subset of S1, two adjacent parcels of Sunflower were misclassified as Pepper at iterations 1 to 3 by the IDL-LLC, as illustrated by the red circles in Fig. 5 (a), but they were correctly classified as Summer crops by the IDL-HLC (see the yellow circles in Fig. 5 (b)). With the valuable information provided by the IDL-HLC at iteration 3, Sunflower were accurately classified from Pepper at iteration 4. Besides this, the misclassifications between Sunflower and Tomato were rectified progressively with the help of IDL-HLC, and they were completely discriminated from each other at iteration 4 (Fig. 5 (a)). In turn, the IDL-LLC modified the classification errors of IDL-HLC during the iterative process. For example, a misclassified parcel of Winter crops produced by the IDL-HLC at iteration 1 was rectified at iteration 2 (Fig. 5 (b)) with the correct information about crop class (i.e., Winter wheat) provided by the IDL-LLC at iteration 2 (Fig. 5 (a)). Similar to subset 1, the IDL-LLC and IDL-HLC rectified each other iteratively in the second subset (Fig. 5 (c and d)). Clearly, Tomato and Pepper were misclassified as each other by the IDL-LLC at iterations 1 and 2 (Fig. 5 (c)). Fortunately, they were correctly labelled as Vegetable by the IDL-HLC at iteration 2 (Fig. 5 (d)), which helped the IDL-LLC discriminate Tomato from Pepper accurately at iteration 3 (Fig. 5 (c)).

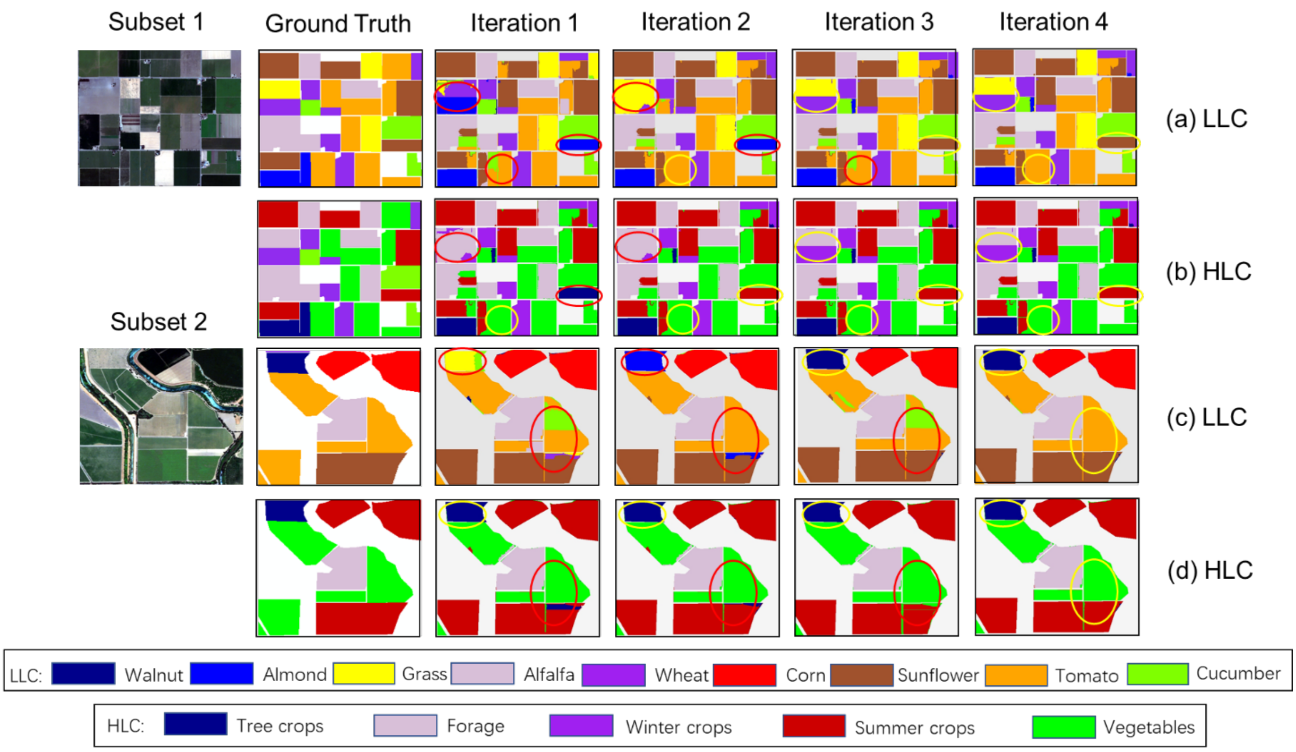

Fig. 6. Two typical image subsets of the LLC and HLC classifications in S2 achieved using the LLC-submodel and HLC-submodel, respectively. Note that the red and yellow circles highlight incorrect and correct classifications, respectively.

Regarding S2, a Sunflower parcel was erroneously mapped as Almond by the IDL-LLC initially (i.e., iterations 1 and 2) in the first subset, as shown by the red circle in Fig. 6 (a). The parcel was correctly identified by the IDL-HLC at iteration 2 (Fig. 6 (b)), which helped IDL-LLC classify the parcel at iteration 3. In turn, the IDL-LLC helped IDL-HLC differentiate Forage and Winter crops at iteration 3, as shown in Fig. 6 (a and b). Like the first subset of S2, the LLC and HLC classification accuracies were increased progressively with iteration in the second subset. For example, a Walnut parcel falsely identified by the IDL-LLC at iterations 1 and 2 (Fig. 6 (c)) was distinguished at iteration 3 with the support of IDL-HLC, in which the high-level class of the parcel was labelled correctly (i.e., Tree crops, Fig. 6 (d)). At the same time, a Tomato parcel mislabelled by

IDL-LLC at iteration 1-3 was correctly identified at iteration 4 (Fig. 6 (c)), thanks to the correct classification information (i.e., Vegetables) achieved by the IDL-HLC (Fig. 6 (c)).

3.4.4 Benchmark comparison for the LLC and HLC classifications

Classification results: To further test the effectiveness of the IDL, a range of benchmarks, including pixel-wise CNN (PCNN), object-based image analysis (OBIA), and object-based CNN (OCNN), were compared with the IDL for both the LLC and HLC classifications in S1 and S2, respectively.

The low-level crop classification maps of S1 and S2 are presented in Fig. 7 (a-b) and (c-d), respectively. As illustrated by the figures, classifications of the proposed IDL-LLC were consistently more accurate compared to those of the benchmarks over both study sites. For the PCNN classification, severe salt-and-pepper noise and linear artifacts were observed in Fig. 7 (a-d); Sunflower, Tomato and Pepper were frequently confused with each other (Fig. 7 (a and b)). For the OBIA and OCNN, smooth LLC classification results were obtained while keeping the precise boundaries of the crop parcels; the classification accuracies of Tomato and Pepper were increased. However, both OBIA and OCNN failed to differentiate Sunflower and Pepper, as well as Walnut and Grass (Fig. 7 (a and d)). Besides, parts of Grass and Cucumber were misclassified as other LLC classes, as shown in Fig. 7 (c). The above issues were resolved by the proposed IDL (i.e., IDL-LLC), which produced clearly the smoothest and most accurate results.

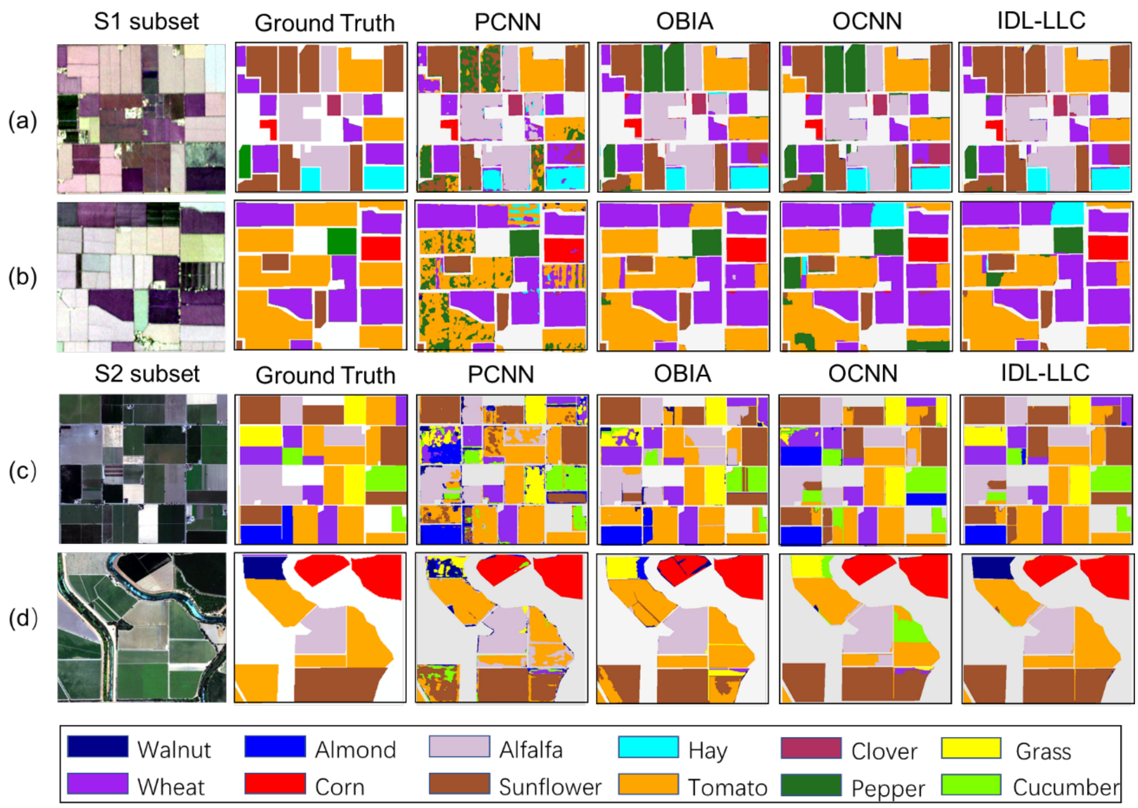

Fig. 7. Image subset comparison amongst PCNN, OBIA, OCNN, and IDL-LLC in both S1 and S2.

In terms of high-level crop classification, the most accurate results were achieved by the proposed method (IDL-HLC) in S1 (Fig. 8 (a-b)) and S2 (Fig. 8 (c-d)). In contrast, the PCNN classification maps produced much undesirable salt-and-pepper noise, especially in the Vegetables and Winter crops parcels (Fig. 8 (a and d)). A large number of pixels near the boundary of crop parcels were classified incorrectly (Fig. 8 (a and d)). By using the segmented objects, the OBIA and OCNN reduced significantly the salt-and-pepper noise, and increased the classification accuracy, accordingly. However, they did not perform well in discriminating HLC classes with similar spectral characteristics. For example, the OBIA often misclassified Summer crops and Vegetables, as well as Tree crops and Forage with each other (Fig. 8 (a and c)), and the OCNN was unable to distinguish between Winter crops, Summer crops, and Vegetables (Fig. 8 (a and d)). These issues were resolved by the proposed IDL-HLC.

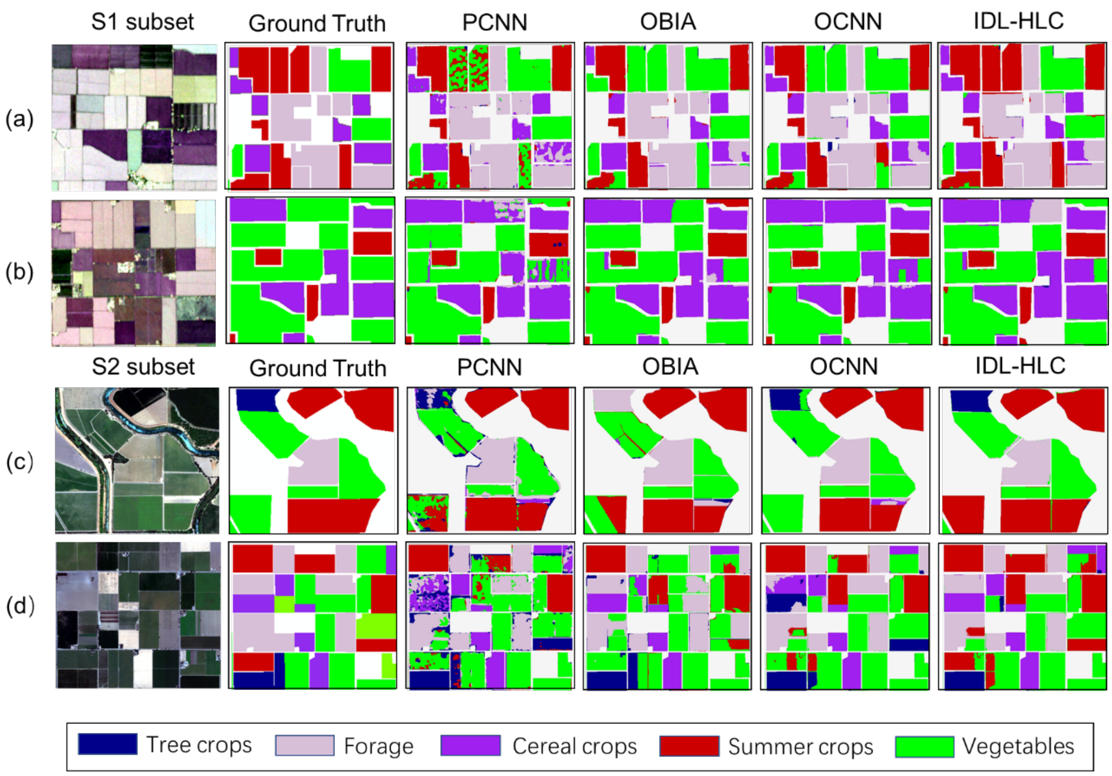

Fig. 8. Image subset comparison amongst PCNN, OBIA, OCNN, and IDL-HLC in both S1 and S2.

Accuracy assessment: To provide a quantitative assessment of classification accuracy, the proposed IDL method was compared with benchmarks using the overall accuracy (OA), Kappa coefficient (κ) and per-class mapping accuracy. The accuracy of LLC classification is summarised in Tables 2 and 3 for S1 and S2, respectively. The IDL-LLC consistently obtained the greatest overall accuracy of 87.89% and 88.94% (κ=0.86 and 0.87) for S1 and S2, respectively, better than for the OCNN at 82.97% and 84.95% (κ=0.80 and 0.82), the OBIA at 85.95% and 82.01% (κ=0.84 and 0.78), and the PCNN at 81.00% and 82.04% (κ=0.78 and 0.79). For the HLC, the accuracy assessment is presented in Tables 4 and 5 for S1 and S2, respectively. The tables show that the IDL-HLC is consistently more accurate (OA=91.74% and 90.72%, and κ=0.89 and 0.88 for S1 and S2) than the benchmarks.

The class-wise mapping accuracy assessment results for the LLC (Tables 2 and 3) and HLC (Tables 4 and 5) classifications in S1 and S2 also demonstrate the superiority of the proposed IDL method. For the LLC classification, the IDL-LLC obtained the greatest accuracy for most of the LLC classes in S1 and nearly all LLC classes (except Walnut) in S2. The largest increases in accuracy were seen for the most challenging Clover class in S1 and Grass class in S2, with accuracies of 83.20% and 80.34%, respectively, for IDL-LLC; markedly greater than for the OCNN (73.89% and 58.68%), OBIA (78.25% and 48.58%), and PCNN (63.38% and 50.94%). The IDL-LLC also produced a large increase in accuracy for Hay, Wheat, and Sunflower in S1 (58.40%, 84.41% and 93.74%), and Almond, Alfalfa, and Cucumber in S2 (86.39%, 82.18% and 82.18%), increasing by around 5%-10% compared to the benchmarks. Moderate increases in accuracy were obtained for Alfalfa in S1 and Wheat and Sunflower in S2, with an average increase of about 3%-5%. For the other LLC classes, only a slight average increase in accuracy were achieved in comparison with the benchmarks.

For the HLC classification (Tables 4 and 5), the IDL-HLC consistently produced the greatest accuracy for nearly all crop classes in S1 and S2, as shown by the bold font in the tables. The most remarkable accuracy increase achieved by the IDL-HLC was achieved for the Winter crops in S1 and Forage in S2 (84.67% and 90.20), much higher than for the OCNN (76.29% and 84.01%), OBIA (78.55% and 75.05%), and PCNN (76.12% and 77.12%). Moderately increased accuracies were produced for Forage and Summer crops in both sites and Tree crops and Vegetables in S2, with an average increase of around 3%-6% compared with the benchmark methods. The IDL-HLC resulted in no significant increase in accuracy for Tree crops in S1, with a slight increase in accuracy in comparison with the benchmarks.

The effectiveness of the proposed IDL method was further demonstrated in comparison with benchmarks using an additional Rapideye satellite imagery. The OA and κ are in accordance with the classification results of S1 and S2. As shown in Table 6, the IDL approach achieved the highest OA of 78.49% for LLC classification and 83.76% for HLC classification, consistently higher than the OCNN (74.23% and 77.62%), the OBIA (72.96% and 73.59%), and the PCNN (71.26% and 77.00%). Such coherency of classification accuracy further confirms the wide applicability of the proposed IDL method.

Table 2 LLC classification accuracy comparison amongst PCNN, OBIA, OCNN and the proposed IDL applied to the first study area (S1). The largest accuracies are highlighted in bold font.

Low-level class (S1)	PCNN	OBIA	OCNN	IDL-LLC
Walnut	89.17	**97.11**	94.14	93.32
Almond	94.16	89.60	**96.05**	92.44
Alfalfa	82.71	88.47	91.20	**91.23**
Hay	47.76	48.11	54.99	**58.40**
Clover	63.38	78.25	73.89	**83.20**
Wheat	76.35	83.78	78.56	**84.41**
Corn	92.44	86.47	91.80	**93.10**
Sunflower	84.94	82.28	83.39	**93.74**
Tomato	88.63	**92.83**	85.44	90.75
Pepper	58.29	**79.79**	44.84	69.19
OA (%)	81.00	85.95	82.97	**87.89**
Kappa	0.78	0.84	0.80	**0.86**

Table 3 LLC classification accuracy comparison amongst PCNN, OBIA, OCNN and the presented IDL applied to the second study area (S2). The largest accuracies are highlighted in bold font.

Low-level class (S2)	PCNN	OBIA	OCNN	IDL-LLC
Walnut	83.27	79.49	**88.31**	87.28
Almond	76.24	74.79	79.97	**86.39**
Grass	50.94	48.58	58.68	**80.34**
Alfalfa	78.49	77.88	85.99	**89.98**
Wheat	85.16	88.56	92.31	**94.64**
Corn	97.18	96.74	96.48	**98.99**
Sunflower	85.51	84.38	82.94	**89.00**
Tomato	86.95	86.63	86.42	**88.50**
Cucumber	71.20	71.73	78.30	**82.18**
OA (%)	82.40	82.01	84.95	**88.94**
Kappa	0.79	0.78	0.82	**0.87**

Table 4 HLC classification accuracy comparison amongst PCNN, OBIA, OCNN and the presented IDL applied to the first study area (S1). The largest accuracies are highlighted in bold font.

High-level class (S1)	PCNN	OBIA	OCNN	IDL-HLC
Tree crops	94.93	93.59	**95.19**	94.85
Forage	88.61	90.27	89.63	**92.65**
Winter crops	76.12	78.55	76.29	**84.67**
Summer crops	90.34	88.42	90.81	**94.96**
Vegetables	91.23	91.43	91.59	**93.08**
OA (%)	87.91	88.48	88.49	**91.74**

Kappa	0.84	0.85	0.85	**0.89**

Table 5 HLC classification accuracy comparison amongst PCNN, OBIA, OCNN and the presented IDL applied to the second study area (S2). The largest accuracies are highlighted in bold font.

High-level class (S2)	PCNN	OBIA	OCNN	IDL-HLC
Tree crops	82.75	85.80	87.29	**90.58**
Forage	77.12	75.05	84.01	**90.20**
Winter crops	86.37	88.55	90.23	**94.01**
Summer crops	86.23	85.79	87.73	**90.41**
Vegetables	84.81	81.59	**90.45**	90.03
OA (%)	83.80	83.20	88.29	**90.72**
Kappa	0.79	0.78	0.85	**0.88**

Table 6 Classification accuracy comparison amongst PCNN, OBIA, OCNN and the presented IDL for the S1' from the Rapideye satellite image. The largest accuracies are highlighted in bold font.

	Accuracy	PCNN	OBIA	OCNN	IDL
LLC classification	OA(%)	71.26	72.96	74.23	**78.49**
	Kappa	0.67	0.69	0.70	**0.75**
HLC classification	OA(%)	77.00	73.59	77.62	**83.76**
	Kappa	0.70	0.65	0.71	**0.79**

4. Discussion

Agro-ecosystems can be considered as highly complex and heterogeneous dynamical systems influenced by both human-related and natural environmental conditions (e.g., climate and soil conditions). Due to their highly complex and dynamic nature, identifying crop types from FSR images remains a great challenge, even for deep learning-based algorithms (e.g., Sidike et al., 2019; Li et al., 2019b). Solving such a difficult task normally requires a very deep network with a large number of samples and huge computing resource (Sidike et al., 2019), which may not be achievable or affordable. Seeking conceptually sound solutions to resolve such complex tasks is, therefore, of great value. An agricultural landscape can be conceptualized as comprising a set of crop types represented at different ontological levels in a hierarchical structure. For example, the high-level crop class (HLC) Forage consists of the low-level crops (LLC) Alfalfa, Hay and Clover in S1. The paper proposes to exploit the hierarchical relationship between the LLC and HLC classes to increase the accuracies of classifying both levels to address the challenging problem of classifying complex agricultural landscapes using FSR remotely sensed imagery.

A novel Iterative Deep Learning (IDL) framework was proposed which progressively models the relationship between the LLC and HLC levels through a Markov Process. The two sub-models (LLC-submodel and HLC-submodel) complement each other through information transformation and interaction. Spectral similarities exist amongst LLCs from different HLCs (Li et al., 2019), such that the CNNs often misclassify one LLC as the other. For example, Sunflower and Walnut were misclassified as other LLC classes at the beginning of the iterative process (i.e., without HLC classification information) (Fig. 5(a) and Fig. 6(c)). Fortunately, the corresponding crop parcels were classified accurately (i.e., Summer crops and Tree crops) at the HLC level (Fig. 5(b) and Fig. 6(d)); this may be due to the unique structural characteristics of Summer crops and Tree crops (Li et al., 2019). Similarly, the differentiation of LLCs within a given HLC can also be enhanced with more accurately identified HLC classes, as shown in Fig. 5 (c), where the complex classification issue between Tomato and Pepper was solved from iteration 3. In short, the more accurate HLC classification can feedback unique and valuable information to increase the accuracy of LLC classification. In turn, with the improved prediction of the LLCs, the HLCs can be distinguished more accurately since HLCs are essentially constituted perfectly by averaging the LLCs. For example, the classification of LLC at iteration 3 helped to identify Winter crops for the HLC classification (Fig. 6 (a and b)). The positive feedback process in the IDL between the LLC and HLC levels refines, updates and reinforces the two classifications in a complementary way through iteration.

It should be noted that the CNN predicted scores (i.e., CNN predicted probabilities) are usually poorly calibrated, often tend to be over-confident (Guo et al., 2017). For example, a prediction score of 0.9 for a crop parcel does not necessarily mean it can be correctly identified with 90% probability. As such, it is very essential and useful to calibrate deep learning. In the proposed IDL model, CNN prediction scores were calibrated to classification probabilities via Label Smoothing (Müller et al., 2019). Such calibration not only improves the iteration efficiency for IDL, reaching the highest accuracy with only four iterations, but also increases the accuracy of predictions, rising from ~84% to ~88%.

As mentioned above, previous studies improved crop classifications with prior crop rotation knowledge acquired via temporal hierarchy of classes. Their central idea is to explicitly define a transition probability matrix of which classes can follow others in a crop rotation use-case (La Rosa et al., 2019; Giordano et al., 2020). The major shortcoming of such methods is that they rely on a huge amount of knowledge (past datasets or experts' experience) about local practices on crop rotations to generate a transition probability matrix, which makes them more rigid and brittle for other use-cases. Being subject to scattered human knowledge, these methods are, thus, hard to generalise to other regions. In contrast, no prior crop rotation knowledge is required for the Iterative Deep Learning approach proposed in this work. Through an integration of compositional hierarchies (well-accepted knowledge) in an end-to-end manner, the proposed approach is more generalisable and applicable in practice, as demonstrated by the promising results over both study sites.

In this research, the HLC was defined according to our semantic knowledge serving as extra input information to the OCNN classifier. Care and attention should be taken in defining the HLC classes so that the LLCs within the same HLC share similar characteristics (such as spectra, structure and texture). In the proposed IDL framework, we designed and classified two crop hierarchies (i.e., LLC and HLC) which can be further extended to many more hierarchies based on demand in practice. For example, according to the time of reproductive development (e.g., early, mid, and late), certain Summer crops (e.g., Corn and Soybean) may consist of several sub-classes (Sidik et al., 2019), leading to the possibility to formulate a new deeper crop hierarchy. In addition, the proposed IDL was implemented at a single "optimal" scale (i.e., input window size of OCNN). To address the challenges of the diversity and complexity of cropland parcels in terms of size and shape a Scale Sequence OCNN (SS-OCNN), which integrates continuously increasing spatial scales into the classification process, can be employed by the IDL to further improve the classification of crop type.

Along with the development of remote sensing applications, FSR remote sensing image classification is increasingly demanded. Given its great potential to change the paradigm of remote sensing classification, the proposed IDL, thus, has a wide application prospect.

5. Conclusion

In this research, a novel Iterative Deep Learning (IDL) method was proposed for complex agricultural landscape classification through iterative interaction between low-level crop (LLC) and high-level crop (HLC) classifications. The hierarchical relationship between LLC and HLC was specified using a Markov process, which allows the LLC and HLC predictions to refine each other gradually. Experiments in two heterogeneous crop areas using two types of FSR remotely sensed imagery illustrated that the IDL was consistently more accurate than state-of-the-art benchmarks for both LLC and HLC classification. In particular, small biomass crop classes with indistinct remote-sensing spectra (e.g., Clover and Grass), which were very difficult to discriminate, were classified accurately. We, therefore, conclude that the proposed IDL is an effective method for crop classification using FSR remotely sensed imagery. Meanwhile, the IDL is readily generalisable to other ecosystems (or landscapes) with hierarchical relationships. It, thus, represents a potentially useful tool for a wide range of classification tasks in remote sensing.

Acknowledgements

The OCNN approach was developed during a PhD studentship “Deep Learning in massive area, multi-scale resolution remotely sensed imagery” (NO. EAA7369), sponsored by Lancaster University and Ordnance Survey (the national mapping agency of Great Britain). Ordnance Survey owns the intellectual property arising from the project, together with a US patent pending: 'Object Based Convolutional Neural Network' (US application number 16/156044). Lancaster University wishes to thank Ordnance Survey for permission to publish this paper and for the supply of aerial imagery and the supporting geospatial data which facilitated the PhD (which is protected as Crown copyright).

Funding

This research was co-funded by the National Natural Science Foundation of China (41301465), the Jilin Province Science and Technology Development Program (20170204025SF), and the Open Fund of State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (Grant No. 20R04).

References

Alexandratos, N., Bruinsma, J., 2012. World Agriculture towards 2030/2050: The 2012 Revision, ESA Working Paper No. 12-03. Food and Agriculture Organization, Rome.

Azar, R., Villa, P., Stroppiana, D., Crema, A., Boschetti, M., Brivio, P.A., 2016. Assessing in-season crop classification performance using satellite data: a test case in Northern Italy. Eur J Remote Sens 49, 361-380.

Baatz, M.; Schaepe, A. Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. Herbert Wichmann Verlag: Heidelberg, Germany, 2000.

Blaschke, T., 2010. Object based image analysis for remote sensing. Isprs J Photogramm 65 (1), 2-16.

Boryan, C., Yang, Z.W., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. Geocarto Int. 26 (5), 341–358.

Cai, Y.P., Guan, K.Y., Peng, J., Wang, S.W., Seifert, C., Wardlow, B., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sens Environ 210, 35-47.

Castillejo-Gonzalez, I.L., Lopez-Granados, F., Garcia-Ferrer, A., Pena-Barragan, J.M., Jurado-Exposito, M., de la Orden, M.S., Gonzalez-Audicana, M., 2009. Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. Comput Electron Agr 68 (2), 207-215.

Cheng, G., Zhou, P.C., Han, J.W., 2016. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. Ieee T Geosci Remote 54 (12), 7405-7415.

Dong, J., Xiao, X., Menarguez, M.A., Zhang, G., Qin, Y., Thau, D., Biradar, C., Moore, B., 2016. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. Remote Sens. Environ. 185, 142–154.

Duro, D.C., Franklin, S.E., Dube, M.G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens Environ 118, 259-272.

Giordano, S., Bailly, S., Landrieu, L., Chehata, N., 2020. Improved Crop Classification with Rotation Knowledge using Sentinel-1. and-2 Time Series. Photogramm Eng Rem S 86 (7), 431-441.

Guo, C., Pleiss, G., Sun, Y., Weinberger, K.Q., 2017. On calibration of modern neural networks. In Proc. 34th International Conference on Machine Learning pp. 1321-1330.

Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, A.R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T.N., Kingsbury, B., 2012. Deep Neural Networks for Acoustic Modeling in Speech Recognition. Ieee Signal Proc Mag 29 (6), 82-97.

Jiao, X.F., Kovacs, J.M., Shang, J.L., McNairn, H., Walters, D., Ma, B.L., Geng, X.Y., 2014. Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. Isprs J Photogramm 96, 38-46.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet Classification with Deep Convolutional Neural Networks. Commun Acm 60 (6), 84-90.

La Rosa, L.E.C., Feitosa, R.Q., Happ, P.N., Sanches, I.D., da Costa, G.A.O.P., 2019. Combining Deep Learning and Prior Knowledge for Crop Mapping in Tropical Regions from Multitemporal SAR Image Sequences. Remote Sens-Basel 11 (17), 2029.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436-444.

Li, H.P., Zhang, C., Zhang, S.Q., Atkinson, P.M., 2019a. Full year crop monitoring and separability assessment with fully-polarimetric L-band UAVSAR: A case study in the Sacramento Valley, California. Int J Appl Earth Obs 74, 45-56.

Li, H.P., Zhang, C., Zhang, S.Q., Atkinson, P.M., 2019b. A hybrid OSVM-OCNN Method for Crop Classification from Fine Spatial Resolution Remotely Sensed Imagery. Remote Sens-Basel 11 (20)

Li, H.P., Zhang, C., Zhang, S.Q., Atkinson, P.M., 2020a. Crop classification from full-year fully-polarimetric L-band UAVSAR time-series using the Random Forest algorithm. Int J Appl Earth Obs 87

Li, Z.T., Chen, G.K., Zhang, T.X., 2020b. A CNN-Transformer Hybrid Approach for Crop Classification Using Multitemporal Multisensor Images. Ieee J-Stars 13, 847-858.

Martinez, J.A.C., La Rosa, L.E.C., Feitosa, R.Q., Sanches, I.D., Happ, P.N., 2021. Fully convolutional recurrent networks for multidate crop recognition from multitemporal image sequences. Isprs J Photogramm 171, 188-201.

Mkhabela, M.S., Bullock, P., Raj, S., Wang, S., Yang, Y., 2011. Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. Agr Forest Meteorol 151 (3), 385-393.

Muller, R., Kornblith, S., Hinton, G., 2019. When Does Label Smoothing Help? NIPS (2019), pp. 4696-4705.

Pena-Barragan, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens Environ 115 (6), 1301-1316.

Scarpa, G., Vitale, S., Cozzolino, D., 2018. Target-Adaptive CNN-Based Pansharpening. Ieee T Geosci Remote 56 (9), 5443-5457.

Sidike, P., Sagan, V., Maimaitijiang, M., Maimaitiyiming, M., Shakoor, N., Burken, J., Mockler, T., Fritschi, F.B., 2019. dPEN: deep Progressively Expanded Network for mapping heterogeneous agricultural landscape using WorldView-3 satellite imagery. Remote Sens Environ 221, 756-772.

Stuckens, J., Coppin, P.R., Bauer, M.E., 2000. Integrating contextual information with per-pixel classification for improved land cover classification. Remote Sens Environ 71 (3), 282-296.

Wardlow, B.D., Egbert, S.L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the US Central Great Plains. Remote Sens Environ 112 (3), 1096-1116.

Whelen, T., Siqueira, P., 2017. Use of time-series L-band UAVSAR data for the classification of agricultural fields in the San Joaquin Valley. Remote Sens Environ 193, 216-224.

Yao, C.; Zhang, Y.; Zhang, Y.; Liu, H. Application of Convolutional Neural Network in Classification of High Resolution Agricultural Remote Sensing Images. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W7, 2017, pp.989-992.

Zhang, C., Harrison, P.A., Pan, X., Li, H.P., Sargent, I., Atkinson, P.M., 2020a. Scale Sequence Joint Deep Learning (SS-JDL) for land use and land cover classification. Remote Sens Environ, 237.

Zhang, C., Kovacs, J.M., 2012. The application of small unmanned aerial systems for precision agriculture: a review. Prec. Agric., 13, 693-712.

Zhang, C., Sargent, I., Pan, X., Li, H.P., Gardiner, A., Hare, J., Atkinson, P.M., 2018. An object-based convolutional neural network (OCNN) for urban land use classification. Remote Sens Environ 216, 57-70.

Zhang, C., Sargent, I., Pan, X., Li, H.P., Gardiner, A., Hare, J., Atkinson, P.M., 2019. Joint Deep Learning for land cover and land use classification. Remote Sens Environ 221, 173-187.

Zhang, D.J., Pan, Y.Z., Zhang, J.S., Hu, T.G., Zhao, J.H., Li, N., Chen, Q., 2020b. A generalized approach based on convolutional neural networks for large area cropland mapping at very high resolution. Remote Sens Environ 247

Zhang, K., Zuo, W.M., Chen, Y.J., Meng, D.Y., Zhang, L., 2017. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. Ieee T Image Process 26 (7), 3142-3155.

Zhao, J., Zhong, Y.F., Hu, X., Wei, L.F., Zhang, L.P., 2020. A robust spectral-spatial approach to identifying heterogeneous crops using remote sensing imagery with high spectral and spatial resolutions. Remote Sens Environ, 239.

Zhong, L.H., Hu, L.N., Zhou, H., 2019. Deep learning based multi-temporal crop classification. Remote Sens Environ 221, 430-443.