IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 14, 2020, accepted November 21, 2020, date of publication January 20, 2021, date of current version March 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051196

Big Data and Machine Learning With Hyperspectral Information in Agriculture

KENNETH LI-MINN ANG [1], (Senior Member, IEEE), AND JASMINE KAH PHOOI SENG [2], (Member, IEEE)

[1]School of Science and Engineering, University of the Sunshine Coast, Petrie, QLD 4502, Australia
[2]School of Engineering and Information Technology, University of New South Wales, Canberra, ACT 2612, Australia

Corresponding author: Kenneth Li-Minn Ang (lang@usc.edu.au)

ABSTRACT Hyperspectral and multispectral information processing systems and technologies have demonstrated its usefulness for the improvement of agricultural productivity and practices by providing useful information to farmers and crop managers on the factors affecting crop status and growth. These technologies are widely used in a range of agriculture applications such as crop management, crop yield forecasting, crop disease detection, and the monitoring of agriculture land usage, water, and soil conditions. Hyperspectral information sensing can acquire several hundred spectral bands that cover the electromagnetic spectrum of an observational scene in a single acquisition. The resulting hyperspectral data cube contains a large volume of spatial and spectral information. The hyperspectral sequence of images or video further increases the data generation velocity and volume which lead to the Big data challenges particularly in agricultural remote sensing applications. This paper is structured to first give a comprehensive review of representative studies to provide insights into significant research efforts in agriculture using Big data, machine learning and deep learning with the focus on frameworks or architectures, information processing and analytics with hyperspectral and multispectral data. The potential for utilizing Big data, machine learning and deep learning for hyperspectral and multispectral data in agriculture is very promising. The paper then further explores the potential of using ensemble machine learning and scalable parallel discriminant analysis which takes into consideration the spatial and spectral components for Big data in agriculture. To the best of our knowledge, no similar review study on agriculture with Big data, machine learning and deep learning for hyperspectral and multispectral information processing has been reported. Furthermore, the potential of ensemble machine learning and scalable parallel discriminant analysis has not been explored in agriculture information processing. Experiments and data analytics have been performed on hyperspectral data from agriculture for validation. The results have shown the good performance of our approach.

INDEX TERMS Agriculture, big data, machine learning, parallel computing, hyperspectral, multispectral.

I. INTRODUCTION

The authors in [1] project that an increase of approximately 25% to 70% above current production levels may be needed to meet the global crop demand in 2050. This makes it important for farmers and crop growers to utilize emerging technologies to improve productivity to feed the growing global population. The technology and data driven economy and its focus on developing intelligent instrumentation, sensing, robotics, artificial intelligence (AI), machine learning, Big data and data analytics is expected to play a transformative role in agriculture to raise the rate of food production. Big data is increasingly being developed and deployed for many industries, professions, and trade sectors.

For the agriculture sector, Big data provides farmers with useful and actionable information on weather and seasonal patterns, rain and water cycles, fertilizer requirements, and other critical information for harvesting and decision-making. This enables farmers, agricultural suppliers and other stakeholders to make smart decisions such as the cycles for crops planting to increase profitability and the planning of optimal harvesting times leading to improved farm yields. To address the issues of the deployment of Big data in agriculture and Big data which are produced from large-scale networked sensing systems, some authors [2], [3] have presented some reviews for Big data in agriculture. The authors in [2]

The associate editor coordinating the review of this manuscript and approving it for publication was Liang-Bi Chen.

presented a review to develop insights into the usefulness of Big data applications in smart farming and the related socio-economic challenges. The authors in [3] presented a review on some significant research efforts utilizing Big data for crop protection focusing on weed management and control.

A major source of Big data for agriculture comes from hyperspectral and multispectral information processing and remote sensing systems. Remote sensing applications and systems generate a huge amount of earth observation data from many sources (e.g. satellite-based systems, unmanned aerial vehicles (UAVs), ground-based structures) and contribute significantly to the volume of Big data to be processed. Agricultural remote sensing is one of the key enabling technologies to fulfill the potential for precision agriculture. Compared to traditional agriculture approaches, remote sensing approaches for agriculture has the advantages of considering the within-field variability for site-specific management instead of uniform management for the sites [4]. The usefulness of agricultural remote sensing lies in its utilization of global positioning location and geographic information to produce the spatially-varied data for precision agricultural information processing and deployment operations. Agricultural remote sensing is a specialized field to produce the image and spectral data in large volume, variety and complexity to enable decision-making for farmers and crop growers in many areas (e.g. decision support systems for irrigation and fertilization, pest management, crop disease detection, and monitoring of land usage, water and soil properties).

Agricultural remote sensing applications can utilize various data sources including hyperspectral and multispectral data. Hyperspectral and multispectral remote sensing can acquire several hundred spectral bands that cover the electromagnetic spectrum of an observational scene in a single acquisition. The resulting hyperspectral data cube contains a large volume of spatial and spectral information. The hyperspectral sequence of images or video further increases the data generation velocity and volume which lead to the Big data challenges and increase the complexity for information processing and analysis caused by the hyperspectral or multispectral data. The vast amounts of generated data from hyperspectral and multispectral data sources require automated modeling and analysis techniques such as machine learning. The field of machine learning has been defined by [5] as having the goal to program computers to use example data or past experience to solve a given problem. The techniques which have been developed for machine learning is particularly useful to handle the volume and large-scale requirements for Big data applications.

Examples of applications of machine learning in agriculture can be found in [6]. These applications include crop and yield prediction, disease and weed detection, species recognition, soil and water management, animal welfare and livestock management. crop quality for crop management, animal welfare and livestock production for livestock management, water management, soil management, etc. Recent techniques in the field of machine learning have resulted in the development of advanced algorithms termed as deep neural networks (DNN) algorithms and approaches. The authors in [7] defined DNN as computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. DNN methods have significantly improved the state-of-the-art in many fields such as speech recognition, visual object recognition, object detection, drug discovery and genomics.

This paper gives the following contributions. This paper is structured to first give a comprehensive review of representative studies to provide insights into significant research efforts in agriculture using Big data, machine learning and deep learning with the focus on frameworks or architectures, information processing and analytics with hyperspectral and multispectral data. The potential for utilizing Big data, machine learning and deep learning for hyperspectral and multispectral data in agriculture is very promising. The paper then further explores the potential of using ensemble machine learning and scalable parallel discriminant analysis which takes into consideration the spatial and spectral components for Big data in agriculture. To the best of our knowledge, no similar review study on agriculture with Big data, machine learning and deep learning for hyperspectral and multispectral information processing has been reported. Furthermore, the potential of ensemble machine learning and scalable parallel discriminant analysis has not been explored in agriculture information processing. Experiments and data analytics have been performed on hyperspectral data from agriculture for validation. The results have shown the good performance of our approach.

The remainder of the paper is structured as follows: Section II first gives a review of Big data and machine learning for hyperspectral and multispectral data in agriculture. Section III presents the ensemble machine learning and scalable parallel discriminant analysis (EML-SPDA) for agriculture applications and analytics. This section also presents and gives details and discussions of experiments and data analytics. Section IV concludes the paper with some remarks on future works and challenges.

II. REVIEW OF BIG DATA AND MACHINE LEARNING TECHNIQUES FOR HYPERSPECTRAL AND MULTISPECTRAL DATA IN AGRICULTURE

The authors in [8] presented a review on the utilization and deployment of Big data analysis in agriculture. The authors in [3] focused on Big data and machine learning for crop protection. The authors in [9] provided a review of the research focused on the applications of data science and machine learning which are relevant to agricultural systems. The authors in [2] presented a review of Big data in smart farming. These papers presented reviews on Big data or data science related to agriculture, but none of them focused on Big data and machine learning utilizing hyperspectral data for agriculture. There are some authors [4], [10] which have provided a general discussion on Big data in remote sensing. It is noted that these review papers which either focus on

(i) Big data or data science in agriculture or (ii) reviews on machine learning [6] or deep learning [11] for agriculture. Other related works on Big data and sensing systems in smart cities and urban environments can be found in [12] and [13]. The remainder of this section gives an overview of technologies and surveys the potential of Big data, machine learning, AI and deep learning with the focus on spectral, hyperspectral and multispectral data information and processing for agriculture. The works have been summarized into four categories: (1) Big data sources with spectral information; (2) Big data with hyperspectral analytics in agriculture; (3) Machine learning techniques for hyperspectral data analytics in agriculture; and (4) Deep learning techniques for hyperspectral data analytics in agriculture.

A. BIG DATA SOURCES WITH SPECTRAL INFORMATION (BIG SPECTRAL DATA)

Modern hyperspectral sensor technologies have the capabilities of generating very high dimensional imagery with a large number of spectral bands and signatures through the use of sensor optics with a large number of bands and spectral signatures. These technologies make it possible to distinguish materials through spectral information and to provide detailed information about the sensed scene. The sensor technologies from satellite-based hyperspectral imaging systems are also capable of covering vast areas of the earth with high spatial, spectral and temporal resolutions. A hyperspectral image of a single scene can be represented as a large volume three-dimensional (3D) data cube with two spatial dimensions and one spectral dimension.

Sequential scenes are comprised of multiple large volume data cubes and pose significant challenges for Big data. For convenience, we use the term Big spectral data to describe Big data sources with spectral information. There are two main sources for Big spectral data: (1) Big spectral data from satellite imagery; and (2) Big spectral data from unmanned aerial vehicles (UAVs). An example of Big spectral data from satellite imagery is Sentinel-2. Sentinel-2 provides multispectral imaging (MSI) functionalities with spatial, spectral and temporal resolutions, and also has two spectral bands in the red-edge region for distinguishing the different agricultural crops [14]. Table 1 shows a summary of satellites and its hyperspectral/multispectral data capabilities from different countries in the world. These medium-resolution and high-resolution satellites generate huge volumes of hyperspectral or multispectral data which are rapidly increased as Big data or termed as Big spectral data. A second data source for Big spectral data derives from unmanned aerial vehicles (UAVs). As discussed by [15], there are two main classifications of UAV platforms (fixed-wing UAVs and rotary-wing UAVs). Rotary-wing UAVs can be further classified into helicopter UAVs and multi-rotor UAVs. Examples of multi-rotor UAVs are quadcopters, hexacopters and octocopters. These Big spectral data from satellite imagery and UAVs require different approaches for information processing and analytics due to their volume, complexity and characteristics. These lead to many new challenges to be addressed in Big data information processing for agriculture information processing.

TABLE 1. Summary of satellites and its imagery capabilities.

Continent/ Country	Satellite name	Satellite imagery capabilities
Europe	Sentinel-2	Multispectral
	RapidEye	Multispectral
	Spot-6	Panchromatic, Multispectral
	Pleiades-1/2	Panchromatic, Multispectral
United States	WorldView 2	Panchromatic, Multispectral
	WorldView 3	Panchromatic, Multispectral
	WorldView 4	Panchromatic, Multispectral
	IKONOS	Panchromatic, Multispectral
	Quickbird	Panchromatic, Multispectral
	GeoEye-1	Panchromatic, Multispectral
China	CBERS-02B	Panchromatic, Multispectral, WFI Multispectral
	HJ-1A	Multispectral, Hyperspectral
	HJ-1B	Multispectral
	ZY-1 02C	Panchromatic, Multispectral, HR
	ZY-3	Panchromatic, Multispectral
	JL-1	Panchromatic, Multispectral
	GF1	Panchromatic, Multispectral
	GF2	Panchromatic, Multispectral

B. BIG DATA WITH HYPERSPECTRAL ANALYTICS IN AGRICULTURE

This sub-section discusses several representative studies for the application of Big data with hyperspectral analytics in agriculture. A summary of the representative works is shown in Table 2. Agriculture relies on healthy soils to produce quality crops and pastures. One of the real-world Big data challenges initiates from the domain of soil spectroscopy which aims to identify and establish soil spectral libraries (SSLs) and signatures. The authors in [16] proposed an evolutionary fuzzy rule-based system which was applied to real world agricultural Big data. Their work utilized large datasets (GEO-GRADLE and LUCAS SSL libraries) from the area of soil spectroscopy. In this work, the authors proposed a two-stage MapReduce scheme and several adaptations for Big data processing. Their approach adapted an evolutionary fuzzy rule-based algorithm for Big Data termed as DECO3RUM. Their experimental work used real world Big data with hyperspectral information from the area of soil spectroscopy. The data samples were diverse and distributed across a variety of soil and land cover types. The model was evaluated in a Hadoop cluster and simulated on eight virtual servers over a hardware configuration with two Intel Xeon processors and 128GB of RAM.

The authors in [17] proposed a parallel computing approach for hyperspectral identification and classification of oilseed rape waterlogging stress levels. Their work combined hyperspectral imaging and parallel computing to address the challenges of agricultural Big data. In their study, hyperspectral images of these siliques for two oilseed rape varieties (NY 22 and NZ 19) were captured using Resonon Pika XC

TABLE 2. Summary of representative works for big data with hyperspectral analytics in agriculture.

Agriculture targets	References	Techniques or approaches	Features or focus of works
Soil spectroscopy	[16]	Evolutionary fuzzy rule-based algorithm (DECO3RUM)	Evaluation on simulated computational cluster (Hadoop, 8 virtual servers, 2 Intel Xeon processors, 128GB of RAM)
Oilseed rape waterlogging stress	[17]	Artificial neural network (ANN) and support vector machine (SVM)	Implementation on six servers, parallel computing cluster (Spark framework and HDFS (Hadoop Distributed File System))
Remote sensing Big data management	[4]	Agricultural remote sensing Big data	Big data framework for FLTL (four-layer-twelve-level) remote sensing
Big data crop classification (sugar beet, cucumber, maize silage, onion, winter wheat, potatoes).	[18]	Principal component analysis (PCA), Minimum Noise Transform (MNF) and support vector machine (SVM)	Combined spectral and spatial features for improved classification accuracy
Big data geospatial imagery	[19]	Iterative Self-Organizing Data Analysis Technique Algorithm (ISODATA)	Big data classification of 7 land classes (water, shadow, wet, fertile soil, land and forest)

camera, followed by the exposure to three different water logging stress levels (0, 3 and 6 days). Their implementation used six servers, routing and switching devices to form the parallel computing framework using Spark machine learning library and HDFS (Hadoop Distributed File System). The Spark library was used to program and develop two classification algorithms (artificial neural network (ANN) and support vector machine (SVM)). The SVM used the one-against-rest classifier for multiple binary classification. The ANN and SVM were used as classifiers for the hyperspectral data and images using the parallel computing platform. The data from five spectral bands (512, 621, 689, 953 and 961nm) were used as the inputs into the classifiers. For the multiclass classification, the classification accuracy and F1 score of the ANN were higher compared to SVM. For the binary classification, the SVM gave higher accuracy and F1 score. Their results indicated that the ANN was more suitable for multi-class classification on the parallel platform whereas the SVM performed better in binary classification problems.

The authors in [4] proposed a remote sensing data management approach using the four-layer-twelve-level (FLTL) framework as shown in Figure 1. The FLTL is an adaptation of the five-layer-fifteen-level (FLFL) framework proposed by the authors in [20]. The FLTL structure gives a framework for the management of remote sensing and Big data for precision agriculture at regional and farm scales. The production of crop maps is essential for crop classification and the identification of different crops. There are two challenges for crop classification and identification due to the spectral similarity and the huge size of the input data. The authors in [18] proposed crop classification technique which combine various features (spectral, spatial and vegetation index features) to address the spectral similarity challenge for Big data in agriculture. Their technique involves dimensionality reduction using PCA (principal component analysis), MNF (minimum noise transform) in the first stage, followed by the support vector machine (SVM) supervised classification. Their work used six crops to perform the experimental evaluation (sugar beet, cucumber, maize silage, onion, winter wheat, potatoes). Their results showed that combining the vegetation index features with the spectral and spatial features improved the classification accuracy to 98%. The authors in [19] proposed an image classification approach for a study in Florida utilizing unsupervised learning for hyperspectral agricultural images termed as ISODATA (Iterative Self-Organizing Data Analysis Technique Algorithm). Their experimental work used the ENVI (Environment of Visualizing Images) [37] software for geospatial imagery. After performing PCA, the ISODATA algorithm was applied to classify the hyperspectral images for various class types (Water, Shadow, Wet, Fertile soil, Land and Forest). The performance was evaluated and the overall accuracy of the classification process was 75.6%. Another study proposed by the authors in [80] proposed a graph-based learning approach termed as local geometric structure Fisher analysis (LGSFA) for dimensionality reduction. The authors showed that their approach was effective in revealing the manifold structure for high-dimensional hyperspectral data, and their experimental results demonstrated classification results comparable to other state-of-the-art methods. Further information on graph-based learning approaches for hyperspectral information can be found in the survey paper by the authors in [81].

C. MACHINE LEARNING TECHNIQUES FOR HYPERSPECTRAL DATA ANALYTICS IN AGRICULTURE

In the field of agricultural remote sensing, hyperspectral image classification has become an important topic. Hyperspectral data have complex characteristics and a nonlinear relationship amongst the spectral bands and its various component materials. This makes the accurate classification of the sensed scene a challenging task. This subsection presents a review of more recent works on machine learning techniques for multispectral and hyperspectral data analytics in agriculture. A summary of the representative works is shown in Table 3.

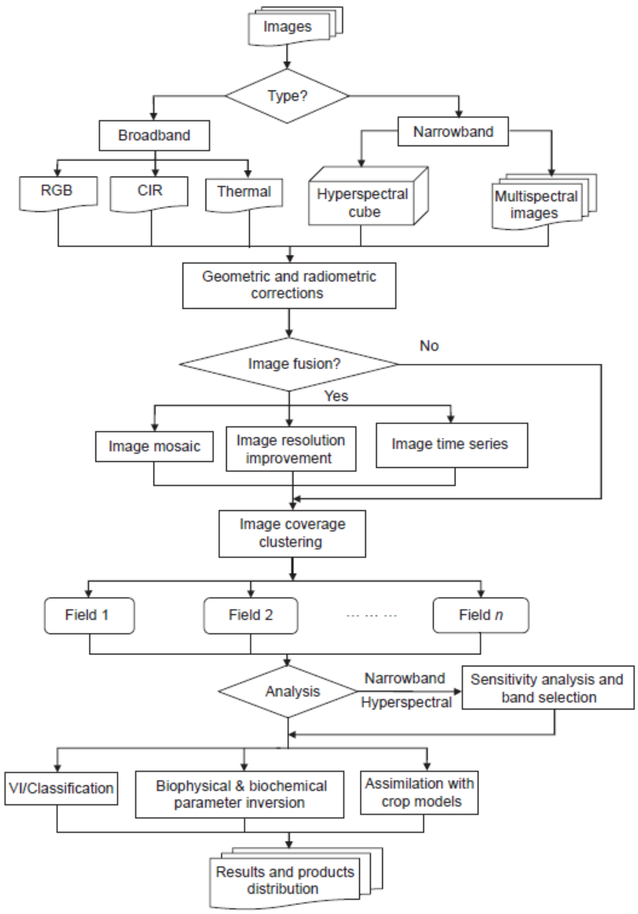

FIGURE 1. **Framework for FLTL remote sensing data management [4].**

The authors in [14] proposed a large-scale crop mapping from multisource remote sensing images in Google Earth Engine. There are three stages in their approach: (1) Harmonic analysis on NDVI data combined with spectral features obtained from satellites (Landsat-8 and Sentinel-2); (2) Utilizing prior constraints of crop distribution and dominance; and (3) Information processing with Google Earth Engine. Their experiments used three crop types (wheat, rapeseed, and corn) to evaluate their approach based on regression tree classification techniques. Their results demonstrated an overall accuracy of 84.25%. Their work also showed that the distribution of the crops in the region of their study was related to agricultural climate, topography and cultivation practices. The authors in [21] proposed an approach to analyze crop fields evolution by utilizing spatial, spectral and temporal S2-SITS data. Their approach consisted of three major stages: (1) Building a vegetation map by combining the spatial and spectral data with temporal NDVI data; (2) Constructing a NDVI time series for a crop field and defining an adaptive regression model with a multilayer perceptron neural network (MLP-NN); and (3) Extracting and analysing the spatial-temporal information from the NDVI time series. The performance of their approach was validated by experiments carried out on S2-SITS data acquired over an area located in Barrax, Spain.

The authors in [22] proposed a spatial-spectral classification framework for Sentinel-2 time series data for land cover mapping. Their approach utilized mathematical morphology and image processing techniques to extract the spatial trends from satellite image time series (SITS) data. These data were then combined with the available spectral and temporal information to improve the discrimination ability among different land cover classes. The obtained spatial–spectral representation was classified with a random forest (RF) classifier. Experiments were conducted on two study sites characterized by different heterogeneous land covers. The sites were Reunion Island study site located in the Indian Ocean and another site in the southwest of France. Their experimental and analysis results have demonstrated the significance of the proposed approach and the validity to combine the spatial and spectral information for land cover classification.

The authors in [23] proposed a sparse kernel logistic regression approach and an incremental learning technique for import vector machines (IVM) for sequential classification of hyperspectral data. Their approach included the addition of new training samples and the deletion of non-informative training samples to improve the classification accuracy while maintaining memory and run-time efficiencies. The incremental learning strategy enables an efficient update of the classifier model without a full re-training from scratch to allow it to handle large data sets. Remote sensing datasets were used to validate the performance of the incremental IVM. The experiments aimed to classify 16 classes. The performance of the IVM was also compared to the SVM for classification accuracy. Their experimental results demonstrated that the IVM and SVM performed comparably in terms of classification performance. However, the number of import vectors was lower when compared to the number of support vectors and remains constant or only slightly increases with an increasing number of training samples.

The authors in [24] proposed machine learning techniques for crop classification using temporal multispectral satellite images. In their approach, several machine learning models were investigated and applied to crop classification of Sentinel-2 satellite image data. The selected study area was the region of Andhra Pradesh in India. The machine models in their study included SVM, random forest, RNN with LSTM and RNN with GRU.

Their results showed that the SVM produced the highest classification performance of 95.9% with the ground surveyed crop areas. The authors in [25] proposed a system for the classification of rice seed varieties using RGB and hyperspectral images. The spatial and spectral features were extracted from the RGB images and hyperspectral image data cubes. The high dimensional spectral feature sets were further reduced using LDA [72]. Their work compared four combinations of the spatial and spectral features: (1) Spatial only; (2) Spectral only; (3) Combination of spatial and spectral features; and (4) Combination of LDA features from spectral data and spatial features. The random forest classifier with the four schemes were used to perform the classification.

TABLE 3. Summary of representative works for machine learning techniques for hyperspectral data analytics in agriculture.

Agriculture targets	References	Techniques or approaches	Features or focus of works
Crop mapping (wheat, rapeseed and corn)	[14]	Harmonic analysis on NDVI time-series data combined with spectral features (Landsat-8 and Sentinel-2)	Large-scale mapping from Google Earth Engine remote sensing images
Spatial-temporal evolution of crop fields	[21]	Multitemporal map built by fusing spectral, spatial and temporal NDVI information	Exploit spectral, spatial and temporal information of S2-SITS data
Land cover mapping	[22]	Mathematical morphology to extract spatial characteristics from SITS, combined with spatial-spectral representation and random forest classifier	Spatial-spectral classification framework for Sentinel-2 time series data
Hyperspectral image classification	[23]	Sparse kernel logistic regression and incremental import vector machines (IVM)	Incremental learning enables efficient update of classifier model without full re-training to handle large data sets
Crop classification	[24]	SVM, random forest, RNN with LSTM and RNN with Gated Recurrent Unit (GRU)	Crop classification of Sentinel-2 satellite temporal remote sensing image data
Classification of rice seed species	[25]	Four combinations of spatial and spectral features, random forest classifier	Combination of spatial and spectral features give good classification performance and discrimination ability
Classification of glycyrrhiza seeds (Glycyrrhiza uralensis Fish, Glycyrrhiza inflata Bat and Glycyrrhiza glabra L)	[26]	SVM and Partial Least Squares Discriminant Analysis (PLS-DA) model	Near infrared hyperspectral imaging with model discriminant analysis could be used to identify glycyrrhiza varieties, origins and planting patterns
Early stage banana disease detection	[27]	SVM classifier with radial basis function (RBF) kernels	Time-series hyperspectral images trained by samples from late infected stage could predict disease in earlier stage
Canopy chlorophyll measurements	[28]	Spectral-temporal response surface (STRS) and Bayesian theory	STRS approach outperformed direct interpolation and direct interpolation with spectral dimension imputation
Mapping agricultural tillage practices	[29]	Kernel extreme learning machine (KELM)	KELM outperformed traditional methods like SVM and random forest
Forecast powdery mildew on barley leaves	[30]	Cycle-consistent adversarial networks (CycleGAN)	Model-based predictor could provide daily forecast one week earlier for better planning of plant protection
Sorghum biomass prediction	[31]	Support vector regression (SVR) and multilayer perceptron (MLP)	Predict biomass from LiDAR point clouds and hyperspectral data
Hyperspectral remote sensing image classification	[32]	Support vector machine (SVM)	Self-training method with spatial majority filtering to find unlabeled samples for SVM classifier training
Vine water status prediction	[33]	Multilayer perceptron (MLP) to predict relation between spectral bands and vine water status	Plant stresses could be predicted with an accuracy of 0.68 to 0.87
Soybean classification	[34]	Single Hidden Layer Feedforward Networks (SLFN) trained with ELM or Optimally Pruned ELM	Best results obtained with 70 spectral bands, significant improvement over previous works
Classification of agricultural landscapes	[35]	Three supervised algorithms (decision tree, random forest and SVM)	RF and SVM classifiers gave better depiction of riparian, wetland and crop land cover types compared to DT
Modelling alpine grassland forage phosphorus	[36]	Three classifier models (artificial neural network, SVM and random forest)	FD-IBs + SVM model gave optimum forage model, account for 88% of forage phosphorus variation

The performances of the proposed approaches were evaluated on a large dataset of 90 rice seed varieties with 96 seeds per variety. The experimental results showed that the combination of spatial features and spectral features could give good classification performance and improve discrimination ability to eliminate the impure species from rice seed samples.

The authors in [26] presented the research work for the classification of glycyrrhiza by utilizing NIR hyperspectral imaging. The study used seed samples from three glycyrrhiza varieties which were collected from four origins and two planting patterns. The authors used spectral information collected from 288 bands (948 nm to 2512 nm). The classifier was developed using the SVM and PLS-DA (Partial Least Squares Discriminant Analysis) models. Their experiments showed that the SVM model gave classification accuracies of 93%. Their work demonstrated that NIR hyperspectral imaging with model discriminant analysis could be used for the identification of different glycyrrhiza varieties, origins and planting patterns. The authors in [27] utilized machine learning methods for banana disease detection. The authors used hyperspectral images with spectral wavelengths ranging from 364 nm to 1031 nm with a spectral resolution of 4.55 nm. Three classes were considered for disease classification: (1) Dead; (2) Dying; and (3) Healthy. Their approach utilized morphological techniques from image processing to extract the spatial and spectral features from the banana leaf samples at both early and late stages. The SVM was used for the classification task. Their experimental results demonstrated that the hyperspectral images analysis classifier which was trained by using the samples from banana leaves at late infected stages could be better used to predict the disease in the earlier infected banana leaves compared to utilizing the raw spectral information.

The authors in [28] presented a novel spectral–temporal response surface (STRS) approach by utilizing Bayesian theory to interpolate spectral information into multispectral imagery. They also compared their approach with two earlier methods (direct interpolation and direct interpolation with spectral dimension imputation) for constructing the STRS. Their experimental results showed that the proposed Bayesian STRS approach outperformed the two earlier approaches. The Bayesian STRS gave correlations of 0.83 with leaf area index (LAI) and 0.77 with canopy chlorophyll measurements compared to correlation values of 0.27 for LAI and 0.09 for canopy chlorophyll measurements for the direct interpolated STRS. The authors in [29] proposed an extreme learning machine (ELM) classifier for mapping agricultural tillage practices from hyperspectral remote sensing imagery. The ELM is a single hidden layer feed forward neural network. The authors implemented the kernel version of the ELM termed as the kernel ELM (KELM). A spatial convolution filter was adopted to generate the spatial and spectral features by incorporating information from surrounding pixels, which were used as the inputs into the KELM. The authors conducted the experiments on airborne hyperspectral images and their experimental results showed that the KELM could outperform other traditional approaches like SVM and random forest.

The authors in [30] proposed an approach to predict the spread of powdery mildew on barley leaves by utilizing hyperspectral image data. The authors used the cycle-consistent adversarial networks (CycleGAN) which is a special type of a generative adversarial network (GAN). The GAN consists of two neural networks termed as the generator G and the discriminator D. The CycleGAN consists of two generators G and F. In their experiments, they analyzed healthy barley leaves and leaves which were inoculated by powdery mildew. Their experiments showed that their predictive model was able to forecast the disease spread from the image time-series. The authors in [31] focused on the prediction of sorghum biomass prediction utilizing remote sensing data with high spatial and temporal resolutions. The authors proposed two approaches to perform the biomass prediction: (1) Nonlinear regression models to predict biomass directly from remote sensing data based on features from LiDAR point clouds and hyperspectral data. Two nonlinear regression models support vector regression (SVR) and multilayer perceptron (MLP) were developed. The authors used the parameter settings for SVR and MLP as described in [38]; and (2) Agricultural Production Systems Simulator using remote sensing data to parametrize the crop model, and then simulate the biomass. Evaluations were performed for both approaches to demonstrate the usefulness of the approaches.

The authors in [32] proposed a self-training method and utilized a spatial majority filtering technique to locate the unlabeled samples that could assist in the SVM classifier training. The approach utilizes the assumption that the class labels of neighboring pixels are reliable and the authors proposed a majority voting-based algorithm. The performance of the algorithm is improved by considering the spectral similarity between a center and its surrounding pixels. The authors performed experimental results with agricultural datasets (including Indian Pines and Salinas) and confirmed the effectiveness of the approach for improving the classification accuracy in cases when the number of labelled samples is limited. The authors in [33] demonstrated that spectral images of crops could be used to for nutrient deficiencies detection. Their approach used multispectral cameras mounted on UAV to predict the vine water status using neural network models. In their investigation, they computed the Normalized Difference Vegetation Index (NDVI) from the spectral image data for soil and plant classification. They utilized the multilayer perceptron (MLP) to different spectral bands to predict the relation between the information contained in the spectral bands and the vine water status. Their experimental results showed that plant stresses such as nutrient components could be predicted with an accuracy of 0.68 to 0.87.

The authors in [34] proposed an approach using the extreme learning machine (ELM) for soybean classification from remote sensing hyperspectral images. In their approach, the spectral data is transformed into a hyper spherical representation and an image gradient is computed. The classification was performed by feedforward networks trained with two methods: (1) ELM; and (2) Optimally Pruned ELM (OP-ELM). In the ELM approach, the training consisted of random generation of the hidden layer weights followed by solving a linear system of equations by least squares for the estimation of the output layer weights. The authors used several classes (Perdiz, Monsoy 8544, Monsoy 9010, Kaiabi and Tabarana) in their evaluation of datasets.

Their experimental results showed that the best results were obtained with 70 bands which gave significant improvement over previous results reported in the literature. Furthermore, the OP-ELM gave improved results over other state-of-the-art methods using only the information from one spectral band. The authors in [35] provided a study of pixel-based and object-based image analysis with machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. The authors performed comparisons using three supervised machine learning algorithms (decision tree (DT), random forest (RF), and support vector machine (SVM)). Their experiments showed that all the three classifiers were able to depict the broad land cover types with acceptable accuracies. One finding was that the RF and SVM classifiers were able to give better predictions of riparian, wetland and crop land cover types compared to the DT classifier which had more errors for these classes. Another finding was that the object-based analysis required more computational time compared to the pixel-based analysis.

The authors in [36] proposed a machine learning approach based on hyperspectral remote sensing and agricultural factors (topography, soil, vegetation and meteorology) for modelling alpine grassland forage phosphorus. Their approach utilized the correlation factors (CFs) and correlation bands (CBs) based on fifteen variables and four types of spectral transformations (original spectral (OR), log spectral (1/R), first derivative (FD) and continuum removal spectral (CR)). The authors used three classifier models (artificial neural network (ANN), support vector machine (SVM) and random forest (RF)) in their approach for their experimental evaluation. Their results showed that the FD and CR spectral models could retrieve more feature bands located in the NIR and SWIR regions than the Log (1/R) and OR spectral models for the forage phosphorus estimation. Their work also showed that the combination of IBs and other factors (longitude and monthly mean temperature) increased the accuracy of the forage estimation when compared with the models that used IBs alone. The FD-IBs + SVM model gave the optimum forage model and could account for 88% of the variation of forage phosphorus in alpine grassland.

This sub-section has demonstrated the potential of deploying machine learning techniques for hyperspectral data analytics in agriculture. The representative works which have been discussed show a wide variety of agriculture applications (e.g. crop mapping, prediction of plant diseases and stresses, classification of species, canopy measurements, etc.) which would benefit by the combination of machine learning techniques with hyperspectral data analytics. Some popular machine learning approaches which have demonstrated potential for agriculture applications include the SVM, IVM, MLP, ELM, discriminant analysis, random forest, etc.

D. DEEP LEARNING TECHNIQUES FOR HYPERSPECTRAL DATA ANALYTICS IN AGRICULTURE

In recent years, deep learning approaches have demonstrated significant improvements in the area of advanced machine learning. Several deep learning approaches have been proposed for solving problems including image classification in agriculture. This subsection presents a review of some recent representative studies on deep learning techniques for multispectral and hyperspectral data analytics in agriculture. A summary of the representative works is shown in Table 4. The authors in [39] presented a technical tutorial on the state of the art of deep learning approaches for remote sensing data.

There are different approaches that have been proposed for deep learning networks such as CNNs (convolutional neural networks), DBNs (deep belief networks), AEs (autoencoders) and SCs (sparse coders). The CNN [40] is a multilayer network architecture composed of several stages for hierarchical representation and feature extraction. Each stage consists of three layers: (1) convolutional layer; (2) nonlinearity layer; and (3) pooling layer. The deep structure of CNNs allows the network model to function as highly abstract feature detectors and to map the input features into representations that can improve the performance of the subsequent classification. The DBN [41] is a generative model that contain many layers of hidden variables. The DBN is trained one layer at a time in an unsupervised manner by restricted Boltzmann machines (RBMs). The AE [42] is a symmetrical neural network that is used to learn the features from a data set in an unsupervised manner by minimizing the reconstruction error between the input data at the encoding layer and its reconstruction at the decoding layer. The SC [43] is an unsupervised approach for learning sets of overcomplete bases to represent data efficiently to find a set of basis vectors which can be used to represent an input vector as a linear combination of these basis vectors.

The authors in [67] presented an overview on spatial and spectral information fusion approaches and techniques for hyperspectral image classification. In their work, the authors grouped spatial-spectral information fusion approaches into three categories: (1) segmentation-based approaches where objects are used for classification; (2) feature fusion approaches; and (3) decision fusion approaches where information from several classifiers are combined to achieve the final classification strategy. The authors reviewed different techniques in these categories. The performances of various fusion methods were evaluated for classification accuracy and running time on popular hyperspectral datasets including Indian Pines and Salinas. The results showed that the feature fusion methods could provide superior classification accuracy compared to other methods at the cost of requiring more computational and processing time.

The authors in [44] proposed a deep learning approach for semantic segmentation termed as DeepLab to extract the spatial features of hyperspectral images. The first principal components were used as the label image for the DeepLab training. Normalization was performed using the z-score on the original spectral bands and the extracted spatial features. The spectral and spatial information were combined using a weighted fusion rule and passed into a SVM for classification. The proposed approach had two significant advantages

TABLE 4. Summary of representative works for deep learning techniques for hyperspectral data analytics in agriculture.

Agriculture targets	References	Techniques or approaches	Features or focus of works
Semantic image segmentation	[44]	Deep convolutional networks	DeepLab - deep learning approach for semantic segmentation
Hyperspectral image classification	[45]	Deep learning feature extraction and classification of spectral-spatial HSI using cross domain CNN model	Approach gave good classification accuracy and simple implementation, making use of the available spatial features
Crop classification (tomatoes, corn, rice, grapes, alfalfa, sunflower, clover, almonds, walnuts)	[46]	Hybrid CNN and transformer architecture	CNN-transformer gave significant performance improvement over traditional methods (random forest, SVM, multitemporal CNN, CNN-LSTM)
Hyperspectral image classification	[47]	Active transfer learning with Hierarchical Stacked Sparse Autoencoder (SSAE) networks	Proposed method gave promising performance compared with state-of-the-art approaches
Hyperspectral image classification	[48]	Spectral-spatial deep learning framework for hyperspectral image classification	Proposed framework outperformed other deep learning methods, specifically for small scale classes
Drug crops identification	[49]	Ensemble of CNN classifiers	Promising machine learning approach for drug crops detection
Land cover (urban, water, bare land) and crop (wheat, clover, sugar beet)	[50]	Deep CNN (DCNN) for multi-temporal pixel-based land cover and crop classification	DCNN gave 89% classification accuracy for crops and land cover classes
Hyperspectral image classification	[51]	Deep learning framework (CNN) with markov random fields (MRF) for spatial-spectral classification	Approach gave comparable results with some state-of-the-art methods
Rice variety distribution maps generation (Reiziq, Sherpa, Topaz, YRM, Langi)	[52]	Deep CNN for spectral and time domains	Deep CNN gave 92.87% accuracy compared to 57.49% with SVM
Prediction of cadmium residue in lettuce leaves	[53]	Deep learning with stacked auto-encoders (SAE) and partial least squares SVM regression (LSSVR)	Deep learning approach showed good potential for detecting heavy metal content in lettuce leaves
Classification of corn seedling (W22, BxM, B73, PH207, Mo17) cold damage	[54]	Ten-layer CNN model	Spectral analysis and CNN model could provide useful reference for detecting cold damage in corn seedlings
Detection of aflatoxin in peanuts	[55]	Five-layer CNN architecture	Approach gave recognition rates of 96% and 90% on pixel and kernel levels
Classification of agriculture and urban subclasses	[56]	CNN model and two modalities (hyperspectral and LiDAR)	Classified map with an overall accuracy of 96 percent for the fused modalities
Prediction of wheat fungal outbreaks	[57]	Various deep learning models (DNN, CNN, RNN, LSTM)	Demonstrated CNN and LSTM outperformed traditional classifiers
Winter wheat yield estimation	[58]	CNN model	CNN could provide a useful reference for estimating crop yield, TensorFlow
Hyperspectral image classification	[59]	Subspace-based feature extraction and CNN model	Approach led to performance improvement compared to conventional feature extraction
Crop identification and discrimination	[60]	Parallel Convolutional Neural Network architecture	PCNN gave higher performance than ANN on dataset after 5000 iterations
Grapevine variety identification	[61]	SVM and CNN classifier models	CNN gave best classification accuracies of 91.63% and 93.82%
Agricultural and non-agricultural land detection	[62]	CNN model	Additional training data that are unfamiliar decreased performance of CNN
Crop classification	[63]	Convolutional, recurrent and hybrid neural networks	Hybrid approach gave best performance, 90% of parameters are allocated to modelling temporal data
Prediction of late blight in potato crops	[64]	Spectral band differences to create additional datasets with different band combinations for training	Random forest and CNN models outperformed other models
Hyperspectral image classification	[65]	Deep learning method for spectral-spatial classification based on single gate recurrent unit (GRU)	Outperformed traditional and deep learning methods, extracted more homogeneous feature representations
Hyperspectral image classification	[66]	Deep metric learning (DML) neural network	Approach gave satisfactory classification performance compared to other metric or deep learning models

when compared with other deep learning approaches: (1) The spectral features are extracted at multiple scales; and (2) The approach avoids reduction of the spatial resolution. The work was validated and demonstrated the superiority of the DeepLab feature extraction method particularly for small scale classes which contains limited number of pixels. Other examples of studies for using deep learning for hyperspectral data analytics in agriculture can be found in [73]–[75] and [77].

The authors in [45] proposed a deep learning feature extraction and classification of spectral-spatial HSI using a cross domain CNN model for classification. Their approach used a guided filter to compute the filter output. The authors used three principal components from the HSI as the guided image. The resultant spatial feature maps at different scales were combined to generate the hyperspectral data cube containing the spatial features. The spatial feature vectors of each pixel were reshaped to form a two-dimensional image which was passed into the CNN for classification. The experimental results showed that the approach gave good classification accuracy and had a simple implementation while making full use of the available spatial features.

The authors in [46] proposed a hybrid CNN and transformer architecture for crop classification on multitemporal and multispectral data. In their research, a dataset with 65 acquiring dates were collected from Sentinel-2 A/B and Landsat-8 for a region in central California. Their approach used two steps. The first step obtained scale-consistent feature and position features from the multitemporal sequence. In the second step, the encoder module was used to express the correlation of the sequence to obtain the depth characteristics of the sequence. The proposed CNN-transformer approach was evaluated on a dataset with a crop matrix that included several crops (tomatoes, corn, rice, grapes, alfalfa, sunflower, clover, almonds, walnuts and specialty crops (watermelons, carrots, onions, peas). The classification results showed that the proposed CNN-transformer architecture resulted in a significant performance improvement compared with other traditional methods such as random forest, SVM, and other deep learning (multitemporal CNN and CNN-LSTM) models.

The authors in [47] proposed an approach for hyperspectral image classification using Hierarchical Stacked Sparse Autoencoder (SSAE) networks to learn sparse feature representations. The SSAE networks were applied to extract the spatial and spectral features. The ATL (active transfer learning) sampling method was used to select a subset of the unlabeled samples for labelling and to add them to the training set at each iteration. The authors performed a comprehensive evaluation on three popular hyperspectral data sets including the Salinas Valley dataset which contains 204 bands. Experimental results demonstrated that the proposed method gave promising performance compared with many state-of-the-art approaches.

The authors in [48] proposed a deep learning framework based on DeepLab for hyperspectral image classification (HSIC). There are two stages in their approach for spectral–spatial HSIC. The first stage extracts the spatial features of HSI pixel-to-pixel at multiple scales and avoids the reduction of spatial resolution. This is followed by the weighted fusion of the spatial and spectral features. In the second stage, these fused features are input into the SVM for the final classification. The performance of their framework was tested on two well-known public HSI datasets including the Indian Pines dataset which lies in a predominantly agricultural region and the University of Pavia dataset and compared with some conventional deep learning techniques. Their results revealed good classification performance and that the proposed framework outperformed other deep learning methods, especially for small scale classes.

The authors in [49] proposed a fusion approach for the identification of drug crops from remote sensing images. Their data-driven approach to characterize these drug crops takes into account the complementary information from the NIR channel and false-colour image representations. The different CNN architectures were applied to distinct image representations, which were able to represent complementary characterizations of such crops. These representations were then input to an ensemble of CNN classifiers using multiple architectures. The approach was validated using a dataset containing Cannabis Sativa crops in a Brazilian region called the Marijuana Polygon. Their proposed approach gave high mean F-measure, accuracy and low false detections, and demonstrated a promising approach for machine-learning approaches for drug crops detection in remote sensing images. The authors in [50] proposed a seasonal land cover and crop classification approach using the Deep CNN (DCNN) architecture. Their work investigated the pixel-based crops and land cover classification on several dates for the same agricultural season from the Sentinel satellite. The experiments were performed for some major crops and land cover classification in Egypt. The architecture used 10 spectral bands from the Sentinel-2 satellite imagery during the winter season of 2016. The proposed architecture was also compared with other techniques such as support vector machines (SVMs), random forests (RFs) and k-nearest neighbours (k-NNs). The results revealed that the DCNN achieved about 89% average accuracy for major crops and land cover classes.

The authors in [51] proposed a deep learning framework with CNN and markov random fields (MRF) for spatial-spectral classification of hyperspectral images (HSI). Their approach can be summarised into two stages: (1) A CNN model was built to learn the deep spectral features and the classification of HSI and the class posterior probability distribution was estimated. The input into the CNN was the pixel vectors, thus the CNN is a pixel-classifier in the spectral domain; and (2) The MRF-based multilevel logistic (MLL) prior encoded the spatial information to regularize the classification result from CNN. The MRF-based loopy belief propagation (LBP) was used to learn the marginal probability distribution in HSI to derive the correlation for both the spectral and spatial features. Their experiments used three public datasets including University of Pavia dataset and two agriculture related datasets (Indian Pines dataset and Salinas dataset). Their approach was compared with some state-of-the-art methods, and results revealed the good performance of their approach. The authors in [52] proposed an approach for generating rice variety distribution maps using deep CNN learning in spectral and temporal domains for Sentinel-2 data. In their work, the deep CNN network was applied towards separating rice varieties at the Coleambally Irrigation Area, NSW, Australia, during the 2016-17 rice growing season. Five rice varieties (Reiziq, Sherpa, Topaz, YRM 70 and Langi) were investigated. Their experiments investigated the separability of the rice varieties based on the spectral and temporal patterns. The temporal curves for two spectral indices NDVI and LSWI were charted over the growing period. The performance of CNN was also compared with SVM. Their results showed that the deep CNN gave a classification accuracy of 92.87% compared to 57.49% with the SVM. Amongst the varieties, Sherpa gave the highest producer accuracy of 98%.

The authors in [53] proposed a deep learning-based regression approach to utilize hyperspectral data for the prediction of cadmium residue in lettuce leaves. Their deep learning approach consisted of stacked auto-encoders (SAE) and partial least squares support vector machine regression (LSSVR). Their approach was applied together with Vis-NIR HSI technique to obtain depth features for cadmium prediction in lettuce leaf. In their approach, the Vis-NIR hyperspectral images of 1120 lettuce leaf samples were collected from the region of lettuce leaf and pre-processed with spectral pre-treatment methods. The authors used several algorithms (Successive Projections Algorithm (SPA), Partial Least Squares Regression (PLSR) and SAE) to locate the optimum wavelengths. The LSSVR model was built based on characteristic wavelengths. The results showed that the deep learning approach showed good potential for detecting heavy metal content in lettuce leaves. The authors in [54] proposed a CNN model for classification of five varieties of corn seedling cold damage recognition. Their approach aimed to extract spectral features in the Vis-NIR range to estimate the cold damage of corn seedlings. The pre-processing of spectral data was performed using application of Gaussian low-pass filter and Savitzky-Golay smoothing method combined with its first-order derivative. The CNN modelling using 3600 pixels were sampled from the region of interests. The CNN used a ten-layer model for classification accuracy and computational efficiency. Their results showed that the proposed approach gave high correlation for different types of corn seedlings given by the traditional chemical method (W22 (41.8%), BxM (35%), B73 (25.6%), PH207 (20%) and Mo17 (14%)), and demonstrated that spectral analysis based on CNN modelling could provide a useful technique for detecting cold damage in corn seedlings.

The authors in [55] developed a hyperspectral imagery system using CNN to detect aflatoxin in peanuts using a grating module, SCOMS camera, and electric displacement platform. The authors used 146 hyperspectral images cubes of 73 peanut samples before and after contamination by aflatoxin. Their CNN architecture consisted of five hidden layers: (1) Input layer; (2) Convolution layer; (3) Sub-sampling layer; (4) Convolution layer; and (5) Sub-sampling layer. The output layer was a fully connected layer. Their approach gave recognition rates of 96% and 90% on pixel and kernel levels respectively, and gave better results compared with traditional classifiers such as KNN, SVM and BP-ANN. The authors in [56] applied the deep learning algorithm based on CNN to classify agriculture and urban subclasses. The authors considered two modalities, hyperspectral data and LiDAR data in their work. The hyperspectral data had the advantages of being able to identify the surface objects based on their material composition. However, it has the disadvantages of failing the identification when two or more objects composed of the same materials have different heights. On the other hand, the LiDAR data had the advantages of being able to discriminate the objects of different heights. The complementary nature of both the data modalities are fused to increase the classification accuracy. Their work used the dataset from National Ecological Observatory Network (NEON) [68]. Using the proposed methodology, a classified map was obtained with an overall accuracy of 96% for the fused modalities.

The authors in [57] proposed a framework for predicting Ethiopian wheat fungal outbreaks using hyperspectral satellite imagery and deep feature learning. The authors compared various deep learning models including Deep Neural Networks (DNNs), Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs) to automatically learn the spectral features. They evaluated all models with the following parameters (20-fold nested cross validation, minibatches of 16, dropout rate of 0.5, 40 histogram buckets, 16 filters of size 3×3, 1 unidirectional LSTM layer with 512 hidden cells and 64-unit fully connected layer). Their experimental results demonstrated that the CNN and LSTM approach significantly outperformed that of traditional classifiers.

The authors in [58] proposed an approach for winter wheat yield estimation from multitemporal remote images using CNN. In their approach, they applied histogram dimensionality reduction and time series fusion to generate the input layer for the CNN. The CNN was built to extract the features of winter wheat growth from multitemporal MODIS images for yield estimation in North China. It consisted of the input layer, seven convolution layers, seven activation layers, seven batch normalization layers, three dropout layers, two full connection layers, and an output layer. Their work was implemented by TensorFlow and the results showed good performance and that the estimated yield of winter wheat based on time-series remote sensing images was highly correlated with statistical data (Pearson r value of 0.82), and demonstrated that the CNN could provide a useful reference for estimating crop yield. The authors in [59] proposed a deep learning approach by combining subspace feature extraction and CNNs for hyperspectral image classification. There were two major steps in their approach: (1) Subspace-based feature extraction to reduce the dimensionality of the hyperspectral images by calculating the orthonormal basis of correlation matrix for each class; and (2) CNN hyperspectral image classification using majority voting strategy applied to the output of CNNs for each feature of certain classes. Experiments were conducted on two real hyperspectral data sets including the Indian Pines dataset covering the agricultural Indian Pines test site in Northwestern Indiana. Their results showed that the proposed strategy gave a performance improvement compared to conventional feature extraction strategies. An overall classification accuracy of 98.1% was obtained for the Indian Pines dataset.

The authors in [60] proposed a novel Parallel Convolutional Neural Network (PCNN) architecture for the pixel-wise identification and discrimination of crop types using AVIRIS-NG hyperspectral images. For band selection, two techniques PCA and back traversal of pre-trained ANN were used to identify an optimal set of bands having higher inter-class separability and lower intra-class variability. To discriminate different crop stages for the same crop type, two different CNN models were trained separately using two sets of crops. During the prediction phase, the results of both models were combined in parallel to decide the final class label based on the highest probability. Their experimental results showed that the PCNN achieved slightly higher performance than ANN on augmented test dataset consistently after 5000 iterations with almost identical training parameters. The PCNN achieved the best test accuracy of 99.1%,

The authors in [61] aimed to investigate the possibility to separate one grapevine variety from an enlarged group of other varieties when the number of samples was significantly increased. Their work was used to separate samples of one variety from 63 other varieties. The SVM and CNN classifiers were applied to separate two varieties (Touriga Franca (TFvar) and Touriga Nacional (TNvar)) from all the remaining varieties. The built classifiers used the one-vs-all binary type to indicate if a spectrum belonged to a certain variety or not. Their work showed that it is possible to separate the leaf spectra of TNvar or TFvar from the spectra of 62 other varieties. In the case of TNvar, the SVM gave better classification performance compared to the CNN. The SVM could classify 63% of the non-TNvar spectra and 81% of the TNvar spectra. For TFvar, the CNN gave the best performance with the non-TFvar and the TFvar spectra with correct classification percentages of 91% and 93% respectively.

The authors in [62] utilized deep learning approaches to detecting agricultural and non-agricultural land. Their methodology was based on classification with CNNs and transfer learning using AlexNet. The area of study consisted of the Ionian islands in Greece. The study used two datasets (EuroSAT and Demokritos) which were partitioned into two categories (agricultural and non-agricultural). The agricultural category included four class categories (Annual Crop, Permanent Crop, Herbaceous Vegetation, and Pasture) whereas the non-agricultural included another four class categories (Residential, Sea-Lake, Highway, and Forest). The experimental results showed that the extra information used for the training data that were unfamiliar to the Greek data decreased the performance of the CNN. The authors in [63] investigated approaches utilizing deep learning models for classification of crop types from multi-spectral time series data. In this work, the authors proposed approaches using convolutional, recurrent and hybrid neural networks for evaluating the importance of spatial and temporal structures in the data. Their experiments were conducted on imagery from Sentinel-2. Their results showed that the hybrid configurations which allocated most of the parameters (up to 90%) for modelling the temporal structure of the multi-spectral data gave the best performance.

The authors in [64] applied deep learning methods for the prediction of the severity of late blight in potato crops caused by Phytophthora infestans. Their work used a UAV to capture images of different phenotypes of potato crops with a multispectral sensor. The authors performed comparisons with other machine learning algorithms including random forests, MLP and support vector regression. Their results showed that the random forest and the CNN models gave the best performance for the identification of infested potato crops. The authors in [65] proposed a deep learning method for spatial-spectral classification for hyperspectral images based on the single gate recurrent unit (GRU). The authors conducted experiments on the different input modes in GRU of spectral information and investigated different ways of fusing the spatial information. By comparing the different utilization patterns with several spatial information fusion methods, their approach demonstrated a higher performance for accuracy and efficiency. Their experimental results on datasets revealed that their approach outperformed other traditional and deep learning methods, and also had the advantages of extracting homogeneous discriminative feature representations. The authors in [66] proposed a deep metric learning (DML) neural network for the classification of hyperspectral images. Their work aimed to decrease the distances between same classes and increase the distances between different classes by multilayers nonlinear projection. Their approach was different from other conventional metric learning methods where the proposed DML method had the capability to exploit the non-linear information between samples with multi-layers nonlinear transformation. The experiments used three datasets (Indian Pines, Pavia University, and Salinas) to validate the proposed spatial-spectral DML method. Their experimental results showed that the proposed approach could achieve classification performance which were comparable with other metric learning or deep models.

This sub-section has demonstrated the potential of deploying deep learning techniques for hyperspectral data analytics in agriculture. Several representative works which have been discussed show that deep learning approaches significantly outperformed that of traditional machine learning classifiers for agriculture applications. The representative works which

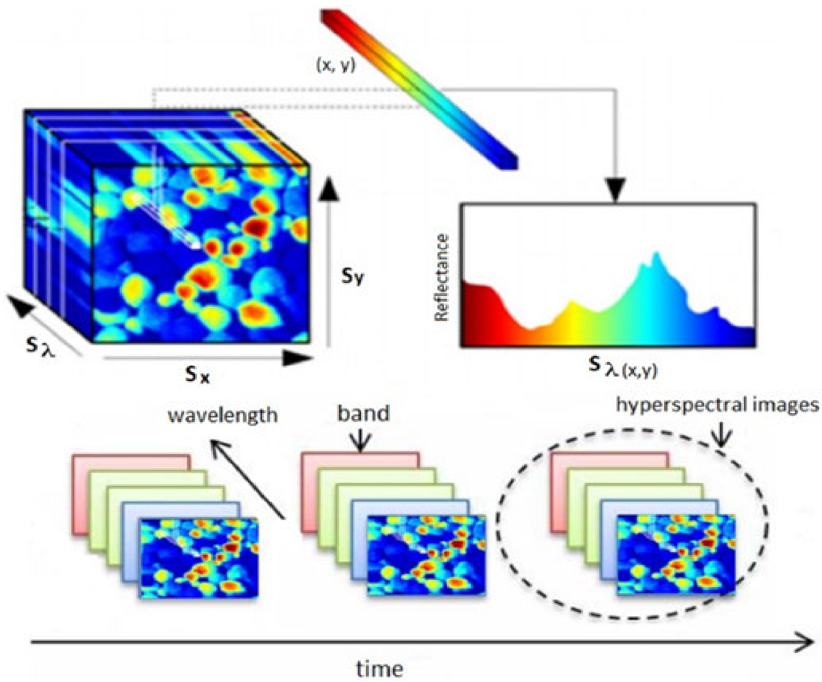

FIGURE 2. 3D cube representation for Big hyperspectral data.

have been discussed show a wide variety of agriculture applications (e.g. semantic crop segmentation and classification, land cover classification, drug crops identification, agricultural and non-agricultural land detection, grapevine identification, prediction of crop diseases, etc.) which would benefit by the combination of deep learning techniques with hyperspectral data analytics. Many studies employ the CNN deep learning model. Other deep learning approaches which have demonstrated potential for agriculture applications include RNN, LSTM, DNN, DML, etc.

III. ENSEMBLE MACHINE LEARNING AND SCALABLE PARALLEL DISCRIMINANT ANALYSIS FOR HYPERSPECTRAL IMAGE CLASSIFICATION

The previous section (Section II) has given a comprehensive overview of agriculture with Big data, machine learning and deep learning for hyperspectral and multispectral information processing. There are several challenges which need to be further addressed to achieve the potential of Big data and hyperspectral information processing in agriculture: (1) The need for efficient machine learning algorithms and classifiers, and also to overcome the shortage of high-quality and labeled training images (e.g. semi-supervised or weakly supervised approaches); (2) The need for efficient and scalable computational architectures for efficient information processing; (3) The need for standardization and ease of use for different remote sensing formats and sensor resolutions particularly for non-expert users; and (4) The need for data management systems to support the efficient storing and indexing of geographical metadata.

As discussed in Section II and illustrated in Tables 3 and 4, hyperspectral image classification is a popular and important application for agriculture. This section gives brief discussions and explores the potential of ensemble machine learning and scalable parallel discriminant analysis (SPDA) for agriculture information processing towards the application of hyperspectral image classification. A similar approach to the proposed SPDA has been previously reported for human emotion and sentiment classification from unstructured Big data [69]. However, the potential of ensemble machine learning and scalable parallel discriminant analysis (EML-SPDA) has not been explored in agriculture information processing. The approach utilizes a tree-based conquer and divide mechanism with an ensemble of classifiers. This part of the paper discusses the EML-SPDA to address Challenges (1) and (2) for Big hyperspectral data for agricultural systems. A difference between the previous work and the proposed approach is that the work in [69] was targeted towards two-dimensional facial image data, whereas the proposed approach is targeted towards large volume three-dimensional (3-D) hyperspectral spatial-spectral data cubes (i.e. Big hyperspectral data). The 3-D hyperspectral data cube structure requires a careful arrangement of the data information processing to preserve the spatial-spectral relationships and for the tree-based conquer and divide mechanism and parallel information processing. The section first gives some discussions on the proposed EML-SPDA approach and is then followed by details and discussions on experiments and data analytics to validate the approach.

A. DISCUSSIONS ON PROPOSED APPROACH

Figure 2 shows the 3-D cube representation for Big hyperspectral data. The hyperspectral cube comprises of two spatial dimensions and one spectral dimension. The data in the cube

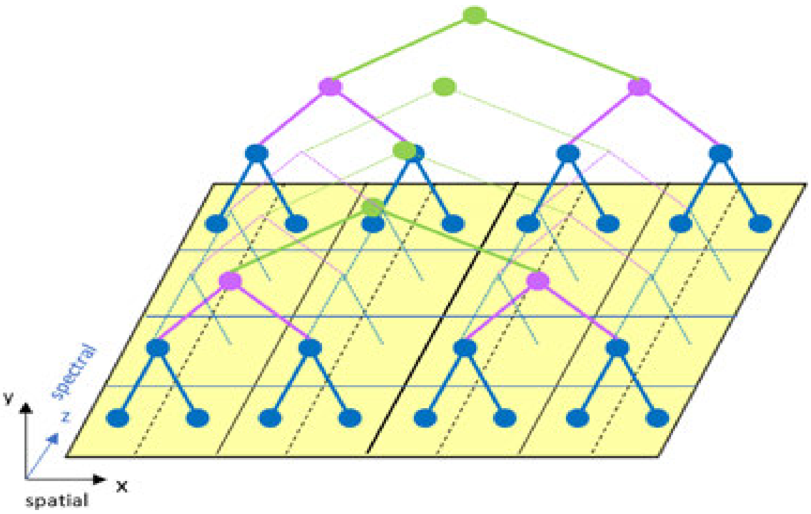

FIGURE 3. Tree-based Conquer and Divide Mechanism.

is re-arranged (split) using a tree-based organization for the conquer and divide mechanism for the parallel information processing as shown in Figure 3. The mechanism first divides the hyperspectral cube system into spatial-spectral localized computational cells. The hyperspectral cube is first divided into the horizonal planes called spatial-spectral planes and each spatial-spectral plane is linearly separated into spatial-spectral bands. The tree-based conquer and divide mechanism is then performed on these spatial-spectral bands. The mechanism breaks the bands into smaller bands based on multiple trees branched recursion. There are different algorithms and techniques which can be applied to perform the information processing using the proposed EML-SPDA framework. For the hyperspectral image classification task, we illustrate the conquer and divide approach using the linear discriminant analysis (LDA) supervised machine learning technique [72], [76], [78]. To perform the LDA using the EML-SPDA approach, the 3-D hyperspectral cube is first mapped into a two-dimensional array structure. Let $X \in R^{d\times n} = [X_1, X_2, \ldots, X_k]$ denote the data matrix partitioned into k classes in which $X_i \in R^{d\times n_i}$ denotes samples from the i^{th} class, i = 1,2,..., k, and $n = \sum_{i=1}^{k} n_i$. Using the notations of S_w, S_b, and S_t to denote the within-class scatter matrix, between-class scatter matrix, and total scatter matrix respectively, the LDA class separability criterion can be formulated as

$$G = \underset{G}{argmax}\, \frac{Tr\left(G^T S_b G\right)}{Tr\left(G^T S_w G\right)}. \tag{1}$$

Table 5 shows a summary of some notations used for the EML-SPDA scheme.

Figure 4 shows the algorithm to perform the conquer and divide mechanism for the EML-SPDA LDA implementation using the RQ decomposition following a binary tree splitting and re-merging mechanism. The RQ decomposition is a counterpart to the well-known QR decomposition. The output of a RQ decomposition for a $m \times n$ matrix is a diagonal

Input: Data set $X = [x_1, x_2, \ldots, x_k] \in R^{d\times n}$, Class label $C = [c_1, c_2, \ldots, c_k] \in R^{1\times n}$.

1. *Binary tree row-based splitting of data matrix*: Split data matrix X into two sub-matrices X_{EVEN} and X_{ODD} containing even rows and odd rows of X respectively. Further split X_{EVEN} and X_{ODD} into four sub-matrices (X_1, X_2, X_3, X_4), and do the parallel decompositions.

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} R_1 Q_1 \\ R_2 Q_2 \\ R_3 Q_3 \\ R_4 Q_4 \end{bmatrix}$$

2. *Parallel decompositions on intermediate sub-matrices*: Compute RQ decompositions of $X_{z1}$and X_{z2}.

$$X_{z1} = [R_{z1} Q_{z1}] = [R_1 Q_1 - R_1 Q_1 Q_2^{\mathrm{T}} Q_2]$$
$$X_{z2} = [R_{z2} Q_{z2}] = [R_3 Q_3 - R_3 Q_3 Q_4^{\mathrm{T}} Q_4]$$

3. *Re-merging of intermediate (int) sub-matrices to compute class separability criterion.*

$$X_{\mathrm{int1}} = [R_{int1} Q_{int1}] = \begin{bmatrix} R_{z1} & R_1 Q_1 Q_2^{\mathrm{T}} \\ 0 & R_2 \end{bmatrix} \begin{bmatrix} Q_{z1} \\ Q_2 \end{bmatrix}$$
$$X_{\mathrm{int2}} = [R_{int2} Q_{int2}] = \begin{bmatrix} R_{z2} & R_3 Q_3 Q_4^{\mathrm{T}} \\ 0 & R_4 \end{bmatrix} \begin{bmatrix} Q_{z2} \\ Q_4 \end{bmatrix}$$
$$X_z = [R_z Q_z] = \left[R_{int1} Q_{int1} - R_{int1} Q_{int1} Q_{int2}^{\mathrm{T}} Q_{int2}\right]$$
$$G = \begin{bmatrix} R_z & R_{int1} Q_{int1} Q_{int2}^{\mathrm{T}} \\ 0 & R_{int2} \end{bmatrix}^{-\mathrm{T}} \begin{bmatrix} Q_z \\ Q_{int2} \end{bmatrix} C$$

Output: $Y = G^T X \in R^{k\times n}$

FIGURE 4. Algorithm for EML-SPDA LDA conquer and divide mechanism.

matrix R of size $n \times n$ and an orthogonal matrix Q of size $m \times n$. The first split stage divides the $d \times n$ data matrix into even rows and odd rows containing two $d/2\times n$ sub-matrices. The second split stage further sub-divides into four sub-matrices containing $d/4\times n$ elements. The RQ decomposition is then performed on each of the sub-matrices to complete the splitting stage. For this EML-SPDA approach for LDA, on a

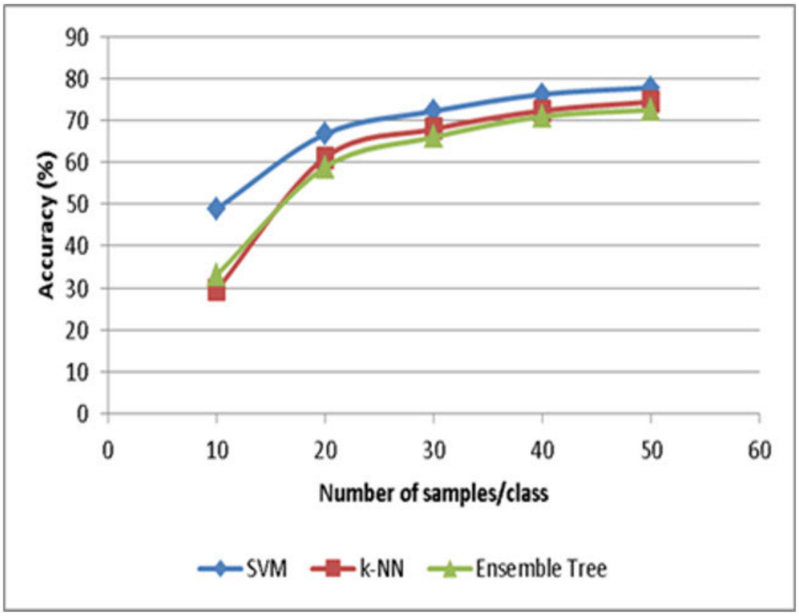

FIGURE 5. Performance accuracy on Indian Pines dataset for different classifiers.

TABLE 5. Summary of notations for EML-SPDA.

Notation	Description
X	Data matrix
X_i	Data matrix of the i^{th} class
d	Number of features/dimensions
n	Number of data samples
n_i	Number of data samples of the i^{th} class
C	Class vector of labels
c_i	Class label of the i^{th} class
S_w	Within-class scatter matrix
S_b	Between-class scatter matrix
S_t	Total scatter matrix

multiprocessor computing platform, each RQ decomposition can be allocated to be performed on a separate processing unit to be computed in parallel. Note that Figure 4 only shows the splitting suitable for four computational processing units. Further stages of splitting can be performed to accommodate a computing hardware platform with a higher number of processors. A significant advantage is that the number of decompositions to be performed can be tailored to suit the computational capability (e.g. number of processors or cores) to achieve the meta-scalability information processing required for the architecture and platform. The re-merging mechanism takes the separate RQ local outputs from the RQ splitting stages and together with the label of class vectors, C combines the local outputs into a global output to obtain the transformation matrix, G for the LDA.

B. DISCUSSIONS ON EXPERIMENTS AND DATA ANALYTICS

This sub-section gives discussions on the experimental implementation and testing for the EML-SPDA and elaborates on the datasets used, the computational setup and the results and discussions.

Experiments: The first set of experiments demonstrates the performance efficacy and the second set of experiments demonstrates the speedup in computational times for EML-SPDA which can be obtained with implementation on parallel processing (in our case multicore) architectures. The experiments aim to demonstrate the efficacy of the conquer-and-divide mechanism for EML-SPDA on parallel architectures using the binary tree row-based re-merging mechanisms.

Data: These set of experiments used the AVIRIS Indian Pines dataset [70]. The Indian Pines dataset covers the agricultural Indian Pines test site in Northwestern Indiana and was collected by the AVIRIS sensor. This dataset contains 16 classes or categories and is a cube size of $145 \times 145 \times 220$ with a spatial resolution of 20 m and a spectral range from 0.2 to 2.4 μm. Table 6 shows the class categories for the AVIRIS Indian Pines dataset.

Computational setup: These set of experiments used an Intel i7 workstation with a 2.2-GHz CPU (4 cores) and 16 GB of RAM.

Results & Discussion: Figure 5 shows the performance accuracy of EML-SPDA for the binary tree row-based conquer and re-merging mechanisms using three different classifiers (SVM, k-NN and ensemble trees) for the Indian Pines dataset. These classifiers were chosen to be representative of the different classification approaches which are available. Other classifiers (e.g. random forest classifiers, Bayesian classifiers, logistic regression, etc.) could be used to perform the classification task. The random forest classifier is an example of an ensemble machine learning (EML) classifier. Other examples of EML approaches are bagging, boosting and stacking. The ensemble tree approach used in the experiments employed adaptive boosted trees [82]. The SVM used the Gaussian kernel, and the k-NN used a value of $k = 10$.

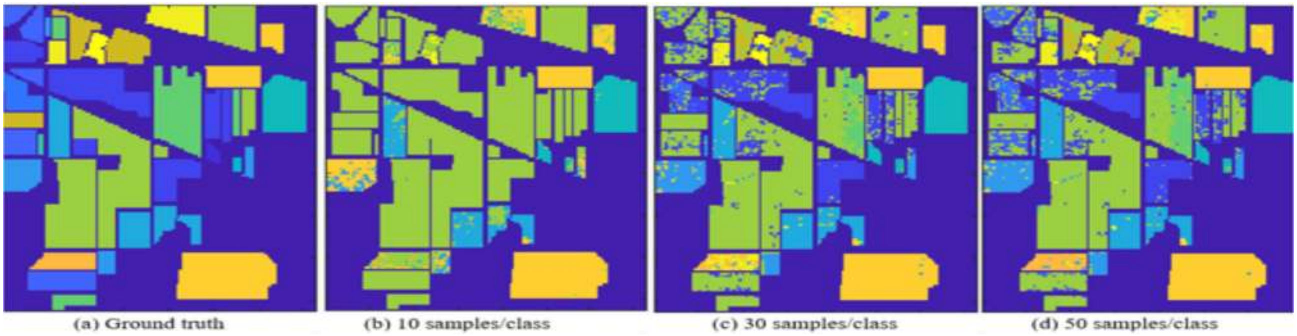

FIGURE 6. Visual classification results for different samples/class.

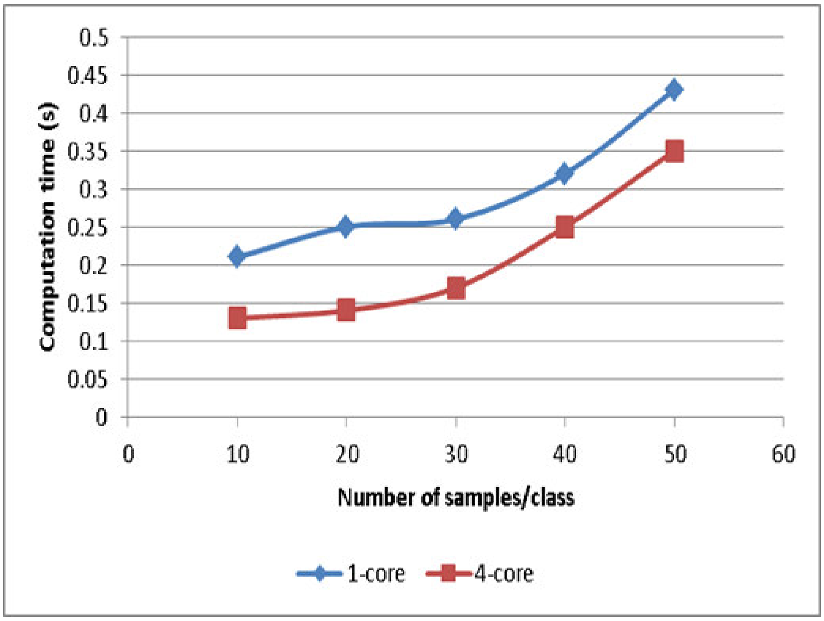

FIGURE 7. Computational time on multicore architectures.

TABLE 6. AVIRIS Indian pines hyperspectral dataset and its class categories for agriculture.

Class	Category
1	Alfalfa
2	Corn-notill
3	Corn-mintill
4	Corn
5	Grass-pasture
6	Grass-trees
7	Grass-pasture-mowed
8	Hay-windrowed
9	Oats
10	Soybean-notill
11	Soybean-mintill
12	Soybean-clean
13	Wheat
14	Woods
15	Buildings-Grass-Trees-Drives
16	Stone-Steel-Towers

The classifiers were trained using a range of samples from 10 to 50 for each class. Amongst the classifiers, the highest accuracy was obtained using the SVM. Note that the focus of the paper is more on the dimensionality reduction using the conquer-and-divide EML-SPDA LDA mechanism, and less on experimenting with improved classifiers to improve the recognition performance. However, we note that the EML-SPDA LDA performed comparably in terms of classification accuracy with the methods and techniques discussed in [71]. Furthermore, the results showed improved accuracy as the number of samples used for training was increased with a classification accuracy of 77.8% for SVM. The results also showed that for the classifiers trained using 20 samples/class or higher, the k-NN classifiers performed comparably with the SVM. Using the lower complexity k-NN classifiers compared with the more complex SVM classifiers can give advantages trade-offs to reduce the implementation complexity at a slight reduction in performance accuracy. Figure 6 shows some visual classification results for the Indian Pines dataset using the SVM classifier with a Gaussian kernel. Only the visual classification results for the SVM classifier are shown because it was the best performing amongst the various classifiers. The leftmost columns show

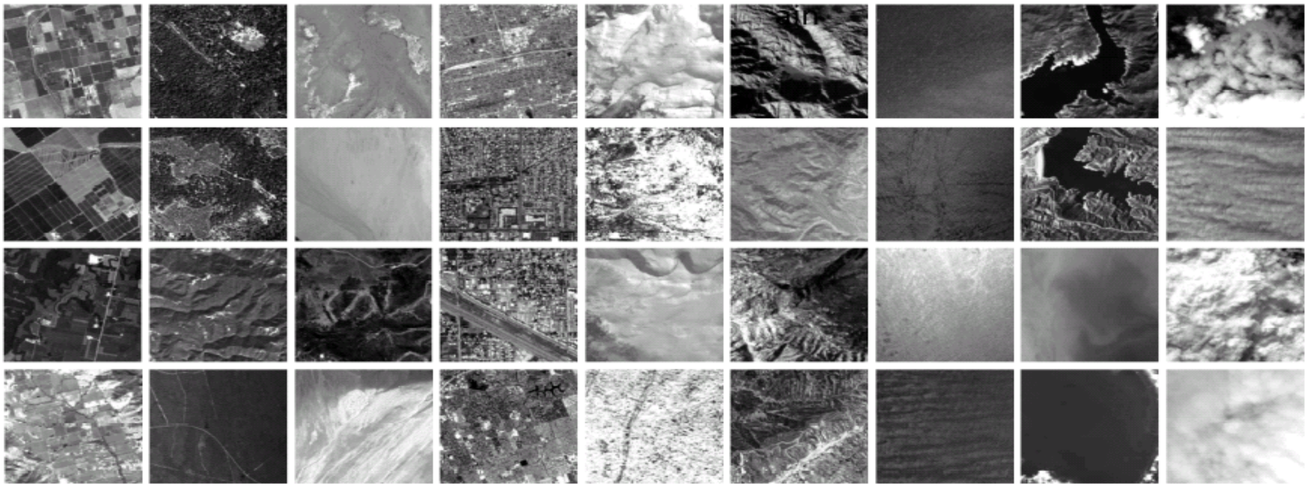

FIGURE 8. Samples for ICONES hyperspectral dataset.

the ground truth results, and the columns moving towards the right show the classification results for increasing number of training samples/class.

An advantage of the EML-SPDA is the conquer-and-divide mechanism for implementation speed-up on parallel computational units. A further investigation was performed to look at the computational time for the EML-SPDA algorithm on multicore architectures for the different datasets. The experiments were conducted on an Intel i7 workstation with a 2.2-GHz CPU (4 cores) and 16 GB of RAM. The comparison in Figure 7 shows the computational times for different number of samples/class for the Indian Pines dataset for running on one-core and four-core architectures. For the dataset, the four-core splitting and re-merging architecture gave a speedup of 1.22 times for the Indian Pines dataset and demonstrating the usefulness of the proposed techniques. It is expected that a higher speedup can be obtained on computational platforms with larger number of computational units (e.g. GPU and massively parallel processors).

For a final investigation, we used a recently developed and published large dataset termed as the ICONES Hyperspectral Satellite Images Dataset (ICONES- HSI) [79]. To the best of our knowledge, the ICONES-HSI dataset is the largest hyperspectral (approximately 36GB) and most recent (published in 2019) dataset available for researchers. This dataset contains 486 remote sensing patches of dimensions 300×300 hyperspectral pixels which were generated from the NASA JPL AVIRIS. The spectral radiance measurement data is sampled in 224 contiguous spectral channels/bands between 365 and 2497 nm. The patches in the dataset are classified into nine categories (Agriculture, Forest, Desert, Urban, Snow, Mountain, Ocean, Wetland and Cloud). Figure 8 shows some representative samples for the nine categories. The spatial-spectral feature for a patch contains $300 \times 300 \times 224$ pixel measurements. In our experiments, we did not use the last six patches for the Cloud category

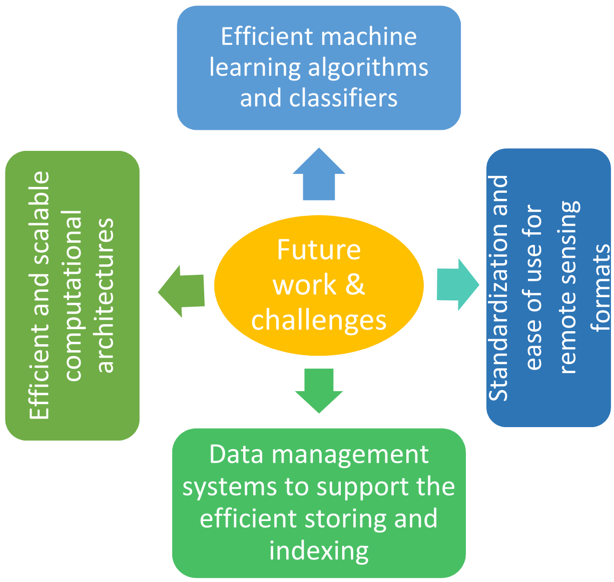

FIGURE 9. Future work and challenges for Big data and hyperspectral information processing in agriculture.

resulting in a data matrix of $20{,}160{,}000 \times 480$. The dimensionality reduced data matrix was passed to two different classifiers (SVM and ensemble tree) to perform the classification tasks which returned 98.8% and 94.4% recognition rates respectively. Figure 9 shows a summary of future work and challenges for Big data and hyperspectral information processing in agriculture.

IV. CONCLUSION AND FUTURE WORK

Big data and machine learning in remote sensing for agriculture is very promising. This paper has provided a comprehensive review of the research efforts in remote sensing in agriculture using Big data and machine learning. There are several challenges which need to be further addressed to achieve the potential of Big data and hyperspectral information processing in agriculture: (1) The need for efficient machine learning algorithms and classifiers, and also to overcome the shortage of high-quality and labeled training images (e.g. semi-supervised or weakly supervised approaches); (2) The need for efficient and scalable computational architectures for rapid information processing; (3) The need for standardization and ease of use for remote sensing formats and sensor resolutions particularly for non-expert users; and (4) The need for data management systems to support the efficient storing and indexing of geographical metadata. The latter part of the paper has proposed the EML-SPDA to address Challenges (1) and (2) for Big hyperspectral data in agricultural information processing. For Challenge (1), the LDA EML-SPDA can perform comparably with other state-of-the-art methods although these methods are not designed for scalability and parallel processing for hyperspectral data. The experimental results have validated the performance of the approach. For Challenge (2), the EML-SPDA has addressed the challenge of traditional conquer-and-divide mechanism which breaks and recursively solves the subproblems of the original, and finally combines the solutions to the subproblems but does not guarantee the optimal solutions for discriminative analytics. The ensemble parallelism machine learning which can be used with many existing machine learning techniques has also been proposed for applications involving Big hyperspectral classification or prediction. In the future, we plan to extend our work by incorporating and re-designing other data analytics into our proposed framework to further address the above challenges.

REFERENCES

[1] M. C. Hunter, R. G. Smith, M. E. Schipanski, L. W. Atwood, and D. A. Mortensen, "Agriculture in 2050: Recalibrating targets for sustainable intensification," *BioScience*, vol. 67, no. 4, pp. 386–391, Apr. 2017.

[2] S. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, "Big data in smart farming–a review," *Agricult. Syst.*, vol. 153, pp. 69–80, May 2017.

[3] R. H. L. Ip, L. M. Ang, K. P. Seng, J. C. Broster, and J. E. Pratley, "Big data and machine learning for crop protection," *Comput. Electron. Agricult.*, vol. 151, pp. 376–383, Aug. 2018.

[4] Y. Huang, Z.-X. Chen, T. Yu, X.-Z. Huang, and X.-F. Gu, "Agricultural remote sensing big data: Management and applications," *J. Integrative Agricult.*, vol. 17, no. 9, pp. 1915–1931, Sep. 2018.

[5] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA, USA: MIT Press, 2020.

[6] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, Aug. 2018.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[8] A. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, "A review on the practice of big data analysis in agriculture," *Comput. Electron. Agricult.*, vol. 143, pp. 23–37, Dec. 2017.

[9] N. Tantalaki, S. Souravlas, and M. Roumeliotis, "Data-driven decision making in precision agriculture: The rise of big data in agricultural systems," *J. Agricult. Food Inf.*, vol. 20, no. 4, pp. 344–380, Oct. 2019.

[10] Y. Ma, H. Wu, L. Wang, B. Huang, R. Ranjan, A. Zomaya, and W. Jie, "Remote sensing big data computing: Challenges and opportunities," *Future Gener. Comput. Syst.*, vol. 51, pp. 47–60, Oct. 2015.

[11] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agricult.*, vol. 147, pp. 70–90, Apr. 2018.

[12] L.-M. Ang, K. P. Seng, A. M. Zungeru, and G. K. Ijemaru, "Big sensor data systems for smart cities," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1259–1271, Oct. 2017.

[13] L.-M. Ang and K. P. Seng, "Big sensor data applications in urban environments," *Big Data Res.*, vol. 4, pp. 1–12, Jun. 2016.

[14] X. Liu, H. Zhai, Y. Shen, B. Lou, C. Jiang, T. Li, S. B. Hussain, and G. Shen, "Large-scale crop mapping from multisource remote sensing images in Google Earth engine," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 13, pp. 414–427, 2020.

[15] J. Kim, S. Kim, C. Ju, and H. Il Son, "Unmanned aerial vehicles in agriculture: A review of perspective of platform, control, and applications," *IEEE Access*, vol. 7, pp. 105100–105115, 2019.

[16] N. L. Tsakiridis, J. B. Theocharis, and G. C. Zalidis, "An evolutionary fuzzy rule-based system applied to real-world Big Data-the GEO-CRADLE and LUCAS soil spectral libraries," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2018, pp. 1–8.

[17] J. Xia, B. Huang, Y. W. Yang, H. X. Cao, W. Zhang, L. Xu, Q. Wan, Y. Ke, W. Zhang, and D. Ge, "Hyperspectral identification and classification of oilseed rape waterlogging stress levels using parallel computing," *IEEE Access*, vol. 6, pp. 57663–57675, 2018.

[18] S. Reshma, S. Veni, and J. E. George, "Hyperspectral crop classification using fusion of spectral, spatial features and vegetation indices: Approach to the big data challenge," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 380–386.

[19] S. A. El_Rahman, "Big data analysis: Hyperspectral image processing for agriculture applications," *Int. J. Comput. Digit. Syst.*, vol. 5, no. 4, pp. 225–234, Jul. 2016.

[20] D. Wang, F. Zheng, J. Lai, T. Yu, J. Li, and S. Guo, "A new parallel algorithm based on five-layer fifteen-level remote sensing data organization," *Microcomput. Inf.*, vol. 1, pp. 1–5, Mar. 2012.

[21] Y. T. Solano-Correa, F. Bovolo, L. Bruzzone, and D. Fernandez-Prieto, "A method for the analysis of small crop fields in Sentinel-2 dense time series," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 2150–2164, Mar. 2020.

[22] Y. J. E. Gbodjo, D. Ienco, and L. Leroux, "Toward spatio–spectral analysis of Sentinel-2 time series data for land cover mapping," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 307–311, Feb. 2020.

[23] R. Roscher, B. Waske, and W. Forstner, "Incremental import vector machines for classifying hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3463–3473, Sep. 2012.

[24] R. Koppaka and T.-S. Moh, "Machine learning in indian crop classification of temporal multi-spectral satellite image," in *Proc. 14th Int. Conf. Ubiquitous Inf. Manage. Commun. (IMCOM)*, Jan. 2020, pp. 1–8.

[25] S. D. Fabiyi, H. Vu, C. Tachtatzis, P. Murray, D. Harle, T. K. Dao, I. Andonovic, J. Ren, and S. Marshall, "Varietal classification of rice seeds using RGB and hyperspectral images," *IEEE Access*, vol. 8, pp. 22493–22505, 2020.

[26] Q. Han, Y. Li, and L. Yu, "Classification of glycyrrhiza seeds by near infrared hyperspectral imaging technology," in *Proc. Int. Conf. High Perform. Big Data Intell. Syst. (HPBD&IS)*, May 2019, pp. 141–145.

[27] W. Liao, D. Ochoa, L. Gao, B. Zhang, and W. Philips, "Morphological analysis for banana disease detection in close range hyperspectral remote sensing images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 3697–3700.

[28] C. M. Gevaert, J. Suomalainen, J. Tang, and L. Kooistra, "Generation of spectral–temporal response surfaces by combining multispectral satellite and hyperspectral UAV imagery for precision agriculture applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3140–3146, Jun. 2015.

[29] D. Lee, "Mapping agricultural tillage practices using extreme learning machine," in *Proc. 8th Int. Conf. Agro-Geoinformatics (Agro-Geoinformatics)*, Jul. 2019, pp. 1–4.

[30] A. Forster, J. Behley, J. Behmann, and R. Roscher, "Hyperspectral plant disease forecasting using generative adversarial networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 1793–1796.

[31] A. Masjedi, J. Zhao, A. M. Thompson, K.-W. Yang, J. E. Flatt, M. M. Crawford, D. S. Ebert, M. R. Tuinstra, G. Hammer, and S. Chapman, "Sorghum biomass prediction using uav-based remote sensing data and crop model simulation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 7719–7722.

[32] D. Han, Q. Du, and N. H. Younan, "Semisupervised classification of hyperspectral remote sensing images with spatial majority voting," in *Proc. 9th IAPR Workshop Pattern Recogniton Remote Sens. (PRRS)*, Dec. 2016, pp. 1–4.

[33] T. Poblete, S. Ortega-Farías, M. A. Moreno, and M. Bardeen, ''Artificial neural network to predict vine water status spatial variability using multispectral information obtained from an unmanned aerial vehicle (UAV),'' *Sensors*, vol. 17, no. 11, p. 2488, 2017.
[34] R. Moreno, F. Corona, A. Lendasse, M. Graña, and L. S. Galvão, ''Extreme learning machines for soybean classification in remote sensing hyperspectral images,'' *Neurocomputing*, vol. 128, pp. 207–216, Mar. 2014.
[35] D. C. Duro, S. E. Franklin, and M. G. Dubé, ''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.*, vol. 118, pp. 259–272, Mar. 2012.
[36] J. Gao, B. Meng, T. Liang, Q. Feng, J. Ge, J. Yin, C. Wu, X. Cui, M. Hou, J. Liu, and H. Xie, ''Modeling alpine grassland forage phosphorus based on hyperspectral remote sensing and a multi-factor machine learning algorithm in the east of tibetan plateau, China,'' *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 104–117, Jan. 2019.
[37] ENVI. (2009). *ENVI Reference Guide: ENVI Version 4.7.* [Online]. Available: https://www.l3harrisgeospatial.com/docs/using_envi_Home.html
[38] Z. Zhang, A. Masjedi, J. Zhao, and M. M. Crawford, ''Prediction of sorghum biomass based on image based features derived from time series of UAV images,'' in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 6154–6157.
[39] L. Zhang, L. Zhang, and B. Du, ''Deep learning for remote sensing data: A technical tutorial on the state of the art,'' *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.
[40] 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.
[41] G. E. Hinton, S. Osindero, and Y.-W. Teh, ''A fast learning algorithm for deep belief nets,'' *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
[42] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, ''Extracting and composing robust features with denoising autoencoders,'' in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, 2008, pp. 1096–1103.
[43] H. Lee, A. Battle, R. Raina, and A. Y. Ng, ''Efficient sparse coding algorithms,'' in *Proc. Adv. Neural Inf. Process. Syst.*, 2007, pp. 801–808.
[44] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, ''Semantic image segmentation with deep convolutional nets and fully connected CRFs,'' 2014, *arXiv:1412.7062*. [Online]. Available: http://arxiv.org/abs/1412.7062
[45] Y. Guo, H. Cao, J. Bai, and Y. Bai, ''High efficient deep feature extraction and classification of spectral-spatial hyperspectral image using cross domain convolutional neural networks,'' *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 345–356, Jan. 2019.
[46] Z. Li, G. Chen, and T. Zhang, ''A CNN-transformer hybrid approach for crop classification using multitemporal multisensor images,'' *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 13, pp. 847–858, 2020.
[47] C. Deng, Y. Xue, X. Liu, C. Li, and D. Tao, ''Active transfer learning network: A unified deep joint spectral–spatial feature learning model for hyperspectral image classification,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1741–1754, Mar. 2019.
[48] Z. Niu, W. Liu, J. Zhao, and G. Jiang, ''DeepLab-based spatial feature extraction for hyperspectral image classification,'' *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 251–255, Feb. 2019.
[49] A. Ferreira, S. C. Felipussi, R. Pires, S. Avila, G. Santos, J. Lambert, J. Huang, and A. Rocha, ''Eyes in the skies: A data-driven fusion approach to identifying drug crops from remote sensing images,'' *IEEE J. Sel. Topics Appl. Earth Observ.Remote Sens.*, vol. 12, no. 12, pp. 4773–4786, Dec. 2019.
[50] N. Laban, B. Abdellatif, H. M. Ebeid, H. A. Shedeed, and M. F. Tolba, ''Seasonal multi-temporal pixel based crop types and land cover classification for satellite images using convolutional neural networks,'' in *Proc. 13th Int. Conf. Comput. Eng. Syst. (ICCES)*, Dec. 2018, pp. 21–26.
[51] C. Qing, J. Ruan, X. Xu, J. Ren, and J. Zabalza, ''Spatial-spectral classification of hyperspectral images: A deep learning framework with Markov random fields based modelling,'' *IET Image Process.*, vol. 13, no. 2, pp. 235–245, Feb. 2019.
[52] Y. Guo, X. Jia, and D. Paull, ''Mapping of rice varieties with Sentinel-2 data via deep CNN learning in spectral and time domains,'' in *Proc. Digit. Image Comput., Techn. Appl. (DICTA)*, Dec. 2018, pp. 1–7.
[53] Z. Xin, S. Jun, T. Yan, C. Quansheng, W. Xiaohong, and H. Yingying, ''A deep learning based regression method on hyperspectral data for rapid prediction of cadmium residue in lettuce leaves,'' *Chemometric Intell. Lab. Syst.*, vol. 200, May 2020, Art. no. 103996.
[54] W. Yang, C. Yang, Z. Hao, C. Xie, and M. Li, ''Diagnosis of plant cold damage based on hyperspectral imaging and convolutional neural network,'' *IEEE Access*, vol. 7, pp. 118239–118248, 2019.
[55] Z. Han and J. Gao, ''Pixel-level aflatoxin detecting based on deep learning and hyperspectral imaging,'' *Comput. Electron. Agricult.*, vol. 164, Sep. 2019, Art. no. 104888.
[56] S. N. Chaudhri, N. S. Rajput, K. P. Singh, and D. Singh, ''Different modality based remote sensing data fusion approach for efficient classification of agriculture and urban subclasses,'' in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 5710–5713.
[57] R. Pryzant, S. Ermon, and D. Lobell, ''Monitoring ethiopian wheat fungus with satellite imagery and deep feature learning,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 39–47.
[58] H. Mu, L. Zhou, X. Dang, and B. Yuan, ''Winter wheat yield estimation from multitemporal remote sensing images based on convolutional neural networks,'' in *Proc. 10th Int. Workshop Anal. Multitemporal Remote Sens. Images (MultiTemp)*, Aug. 2019, pp. 1–4.
[59] T. Alipourfard, H. Arefi, and S. Mahmoudi, ''A novel deep learning framework by combination of subspace-based feature extraction and convolutional neural networks for hyperspectral images classification,'' in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 4780–4783.
[60] H. Patel, N. Bhagia, T. Vyas, B. Bhattacharya, and K. Dave, ''Crop identification and discrimination using AVIRIS-NG hyperspectral data based on deep learning techniques,'' in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 3728–3731.
[61] A. M. Fernandes, A. B. Utkin, J. Eiras-Dias, J. Cunha, J. Silvestre, and P. Melo-Pinto, ''Grapevine variety identification using 'Big Data' collected with miniaturized spectrometer combined with support vector machines and convolutional neural networks,'' *Comput. Electron. Agricult.*, vol. 163, Aug. 2019, Art. no. 104855.
[62] E. Charou, G. Felekis, D. B. Stavroulopoulou, M. Koutsoukou, A. Panagiotopoulou, Y. Voutos, E. Bratsolis, P. Mylonas, and L. Likforman-Sulem, ''Deep learning for agricultural land detection in insular areas,'' in *Proc. 10th Int. Conf. Inf., Intell., Syst. Appl. (IISA)*, Jul. 2019, pp. 1–4.
[63] V. S. F. Garnot, L. Landrieu, S. Giordano, and N. Chehata, ''Time-space tradeoff in deep learning models for crop classification on satellite multi-spectral image time series,'' in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 6247–6250.
[64] J. Duarte-Carvajalino, D. Alzate, A. Ramirez, J. Santa-Sepulveda, A. Fajardo-Rojas, and M. Soto-Suárez, ''Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms,'' *Remote Sens.*, vol. 10, no. 10, p. 1513, Sep. 2018.
[65] E. Pan, X. Mei, Q. Wang, Y. Ma, and J. Ma, ''Spectral-spatial classification for hyperspectral image based on a single GRU,'' *Neurocomputing*, vol. 387, pp. 150–160, Apr. 2020.
[66] X. Cao, Y. Ge, R. Li, J. Zhao, and L. Jiao, ''Hyperspectral imagery classification with deep metric learning,'' *Neurocomputing*, vol. 356, pp. 217–227, Sep. 2019.
[67] M. Imani and H. Ghassemian, ''An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges,'' *Inf. Fusion*, vol. 59, pp. 59–83, Jul. 2020.
[68] National Ecological Observatory Network (NEON). *Data Products NEON.DP1.30010.001, NEON.DP3.30006.001, NEON.DP1.30003.001, NEON.DP3.30024.001*. Battelle, Boulder, CO, USA. Accessed: Oct. 20, 2018. [Online]. Available: http://data.neonscience.org
[69] J. Kah Phooi Seng and K. Li-Minn Ang, ''Multimodal emotion and sentiment modeling from unstructured big data: Challenges, architecture, & techniques,'' *IEEE Access*, vol. 7, pp. 90982–90998, 2019.
[70] M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe, ''Band AVIRIS hyperspectral image data set,'' in *Purdue University Research Repository*, 2015, p. 220. [Online]. Available: https://purr.purdue.edu/publications/1947/about?v=1
[71] F. Leyuan, N. He, S. Li, J. Anotonio Plaza, and P. Javier, ''A new spatial–spectral feature extraction method for hyperspectral images using local covariance matrix representation,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3534–3546, Jun. 2018.
[72] Z. Fan, Y. Xu, and D. Zhang, ''Local linear discriminant analysis framework using sample neighbors,'' *IEEE Trans. Neural Netw.*, vol. 22, no. 7, pp. 1119–1132, Jul. 2011.

[73] J. Feng, J. Chen, L. Liu, X. Cao, X. Zhang, L. Jiao, and T. Yu, ''CNN-based multilayer spatial–spectral feature fusion and sample augmentation with local and nonlocal constraints for hyperspectral image classification,'' *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 4, pp. 1299–1313, Apr. 2019.

[74] N. He, L. Fang, S. Li, A. Plaza, and J. Plaza, ''Remote sensing scene classification using multilayer stacked covariance pooling,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 12, pp. 6899–6910, Dec. 2018.

[75] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, ''Multisource remote sensing data classification based on convolutional neural network,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 937–949, Feb. 2018.

[76] W. Li, F. Feng, H. Li, and Q. Du, ''Discriminant analysis-based dimension reduction for hyperspectral image classification: A survey of the most recent advances and an experimental comparison of different techniques,'' *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 1, pp. 15–34, Mar. 2018.

[77] J. Yang, Y. Q. Zhao, and J. C. W. Chan, ''Learning and transferring deep joint spectral–spatial features for hyperspectral classification,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4729–4742, Aug. 2017.

[78] J. K. P. Seng and K. L.-M. Ang, ''Big feature data analytics: Split and combine linear discriminant analysis (SC-LDA) for integration towards decision making analytics,'' *IEEE Access*, vol. 5, pp. 14056–14065, 2017.

[79] O. Ben-Ahmed, T. Urruty, N. Richard, and C. Fernandez-Maloigne, ''Toward content-based hyperspectral remote sensing image retrieval (CB-HRSIR): A preliminary study based on spectral sensitivity functions,'' *Remote Sens.*, vol. 11, no. 5, p. 600, Mar. 2019.

[80] F. Luo, H. Huang, Y. Duan, J. Liu, and Y. Liao, ''Local geometric structure feature for dimensionality reduction of hyperspectral imagery,'' *Remote Sens.*, vol. 9, no. 8, p. 790, Aug. 2017.

[81] L. Zhang and F. Luo, ''Review on graph learning for dimensionality reduction of hyperspectral image,'' *Geo-Spatial Inf. Sci.*, vol. 23, no. 1, pp. 98–106, Jan. 2020.

[82] P. Viola and M. Jones, ''Fast and robust classification using asymmetric adaboost and a detector cascade,'' in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, pp. 1311–1318.

KENNETH LI-MINN ANG (Senior Member, IEEE) received the B.Eng. and Ph.D. degrees from Edith Cowan University, Australia. He was an Associate Professor of Networked and Computer Systems with the School of Information and Communication Technology (ICT), Griffith University. He is currently a Professor with the School of Science and Engineering, University of Sunshine Coast. He has published over 180 articles in journals and international refereed conferences. His research interests include big data analytics, multimedia Internet-of-Things, embedded systems, wireless multimedia sensor systems, reconfigurable computing, development of real-world computer systems, and machine learning. He is a Fellow of the Higher Education Academy, U.K.

JASMINE KAH PHOOI SENG (Member, IEEE) received the B.Eng. and Ph.D. degrees from the University of Tasmania, Australia. She is currently an Adjunct Professor with the School of Engineering and Information Technology, UNSW. Before returning to Australia, she was a Professor and the Department Head of Computer Science and Networked System with Sunway University. Before joining Sunway University, she was an Associate Professor with the School of Electrical and Electronic Engineering, Nottingham University. She has published over 230 papers in journals and international refereed conferences. She is the lead author of the book *Multimodal Analytics for Next-Generation Big Data Technologies and Applications*. Her research interests include data analytics, Big data, machine learning, Artificial Intelligence (AI) and intelligent systems, Internet of Things (IoT), multimodal signal processing, pervasive computing and sensor networks, HCI and affective computing, and mobile software development.