

PREDICTION OF SORGHUM BIOMASS USING UAV TIME SERIES DATA AND RECURRENT NEURAL NETWORKS

Ali Masjedi
School of Civil
Engineering
Purdue University
amasjedi@purdue.edu

Neal R. Carpenter
Department of
Agronomy
Purdue University
ncarpen@purdue.edu

Melba M. Crawford
School of Civil
Engineering
Purdue University
mcrawford@purdue.edu

Mitch R. Tuinstra
Department of
Agronomy
Purdue University
drmitch@purdue.edu

Abstract

Phenotyping via Unmanned Aerial Vehicles (UAVs) is of increasing interest for many applications because of their capability to carry advanced sensors and achieve accurate positioning required to collect both high temporal and high spatial resolution data required over relatively limited areas. This paper focuses development of a data analytics based predictive modeling strategy that incorporates multi-sensor data acquisition systems and accommodates environmental inputs. Unsupervised feature learning based on fully connected and convolutional neural networks is investigated. Predictive models based on Recurrent Neural Networks (RNNs) are designed and implemented to accommodate high dimensional, multi-modal, multi-temporal data. Remote sensing data, including Light Detection and Ranging (LiDAR) and hyperspectral inputs, as well as weather data, are incorporated in RNN models. Results from multiple experiments focused on high throughput phenotyping of sorghum for biomass predictions are provided and evaluated for agricultural test fields at the Agronomy Center for Research and Education (ACRE) at Purdue University.

1. Introduction

With the advances in science and technology, it is now possible to sequence the genome of plants rapidly and at low cost [1]. In order to evaluate the performance of the plant varieties with given genetic characteristics, numerous physical and agronomic traits are measured in controlled facilities or in the field during the growing season (phenotyping). Planting the varieties of a crop and measuring dozens of phenotypic traits for each is time consuming, laborious, and expensive. Recently, researchers have begun to explore the use of remotely sensed data to augment and even replace some traditional in-field phenotyping measurements.

Biomass is an important plant characteristic for sorghum-based biofuels, as it is indicative of both the crop condition and the quantity of ethanol that can be produced. Many researchers have studied predictive models for

complex phenotypes such as biomass, based on data derived from Remote Sensing (RS) technologies, via empirical approaches. For example, the importance of sample size, prediction method, and sensor type for biomass predictions was investigated in [2] using support vector machines, random forests, and Gaussian process models. Although traditional empirical methods are powerful tools to model the complex relationships between the RS data and biomass, they are unable to effectively exploit the temporal aspects of the data, incorporate multiple sensing modalities, or include environmental inputs.

In recent years, deep neural networks have been investigated by many researchers in the area of pattern recognition and machine learning (ML). Deep Learning (DL) adds more “depth” (complexity) into the model compared to classical ML and represents the data in a hierarchical way in different layers using various functions to transform the data [3]. If properly trained with adequate labeled data, DL can provide very accurate classification and prediction outputs because of its ability to use more complex models, which also increases flexibility and potential adaptability for various problems.

Extraction of appropriate features, along with using a powerful predictive model, is required for accurate prediction of complex phenotypes such as biomass. Hyperspectral data consist of hundreds of channels that provide detailed spectral information from the image data from which useful predictive features can be extracted. From LiDAR point clouds, multiple features describing the geometric characteristics of the canopy including height, canopy coverage, and vertical distribution of the plant material can be extracted. Although having all these features provides more information for predictive models, it can degrade the accuracy of prediction as many of these features are highly correlated. Generally, this is not a concern for DL-based models, as they have the advantage of being able to learn the features and automatically extract appropriate representations from input data [4] if adequate labeled data are available. For biomass prediction however, the number of data samples is limited, as destructive sampling and processing of biomass is time consuming and expensive. In this situation, it is important to separate feature learning from prediction by first extracting

appropriate features, and then using them in the predictive models.

In this paper, two autoencoder-decoder networks are designed and implemented for unsupervised feature learning from hyperspectral and LiDAR remotely sensed data. The first network consists of a number of fully connected layers, and the second network consists of convolutional, pooling, and de-convolutional layers.

For biomass prediction, having a method that can incorporate the data from multi-year experiments and different sources is important, both for increasing the quantity of data samples and representing the growth scenarios under different weather-related growing conditions. In this study, a Recurrent Neural Network (RNN) is designed to include both hyperspectral and LiDAR data and weather-related data (e.g. solar radiation, temperature, precipitation) to predict crop biomass. The results and analysis provide useful insights on feature importance for prediction and demonstrate the ability to estimate the biomass from remotely sensed data, thereby greatly reducing the required human labor.

2. Feature Learning

Unsupervised feature learning has been investigated by many researchers in different fields [5-7]. In hyperspectral image classification, Chao et al. [8] showed that feature learning based on a stacked sparse auto-encoder model provides robust results. In [9], a network of fully convolutional auto-encoders for image feature learning was developed for image clustering. Inspired by these studies, in this paper, we designed a fully connected auto-encoder network (FAEN) and a convolutional auto-encoders network (CAEN) for unsupervised feature learning of features from both hyperspectral and LiDAR data.

The FAEN network consists of three fully connected hidden layers as well as input and output layers as shown in Figure 1. The first hidden layer consists of more neurons than the input layer, allowing the network to expand the information of the input features to a higher dimensional space. The second layer then has a few neurons (four in this paper) that are considered as the learned features. This part of the network is referred to as the encoder, which abstracts (encodes) the information of the input features into a few features. Then the second part of the network (decoder), which is symmetric with the first part, reconstructs the input features from the learned features.

The second network, CAEN, is similar to the first, but uses convolutional layers instead of fully connected layers, as shown in Figure 2. Using convolutional layers for feature learning allows the network to take advantage of the potential relation between the input features. Also, there are fewer parameters to optimize during the training of the network.

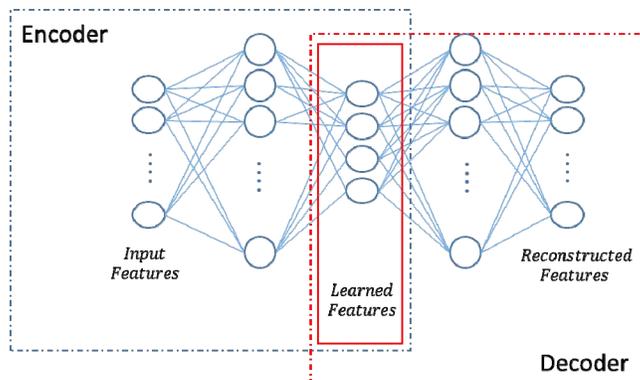


Figure 1. Unsupervised feature learning using fully connected auto-encoder network (FAEN).

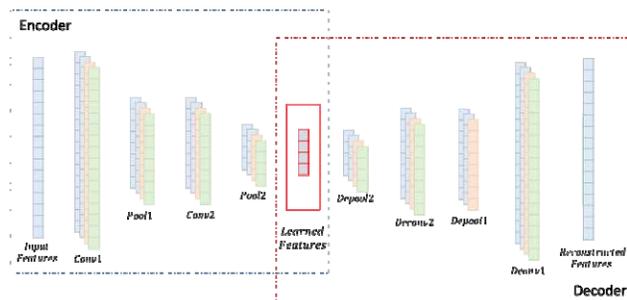


Figure 2. Unsupervised feature learning using convolutional auto-encoder network (CAEN).

The size of output layer in both networks is the same as the input layer as they reconstruct the input features. The cost function is the average sum-of-squared error between input and reconstructed features over the entire sample data.

3. Predictive Models

In this paper, two models are developed for biomass prediction, one based on Support Vector Regression (SVR) and the other based on Recurrent Neural Networks (RNN).

3.1. Support Vector Regression

SVR is a supervised non-parametric regression technique, and therefore, no assumptions regarding the underlying data model are required. SVR transforms the original feature space into a higher dimensional space [10], with the goal of finding a hyperplane to predict the training data set. More details for SVR are provided in [11]. The optimal values of the kernel function parameters are found by a general k-fold cross-validation in a grid search.

3.2. Recurrent Neural Network

RNNs are powerful models for representing sequential data and are particularly effective in learning long-term dependencies. Mou et al. [12] proposed an RNN model to classify hyperspectral pixels by considering the

sequence-based data structure. In [13], a novel RNN network was proposed for online disturbance detection from a satellite image time series. RS-based prediction was also investigated by You et al. [14] for soybean yield prediction using an RNN.

Long Short Term Memory (LSTM) and Gated recurrent units (GRU) have proven to be successful for accommodating the difficulty of training simple recurrent networks [15]. An LSTM unit is capable of learning long-term dependencies due to its structure, which incorporates gates that regulate the learning process. The operations are accomplished by the input, forget, and output gates. Based on the information provided by the input and forget gates, the cell state, which represents the memory of the unit, is updated. The GRU architecture is similar to LSTM, but simpler to compute and implement as it consists of two gates: reset and update gates [16].

In this study, GRU and LSTM models are developed and compared for multi-temporal prediction of sorghum biomass. The proposed network is shown in Figure 3.

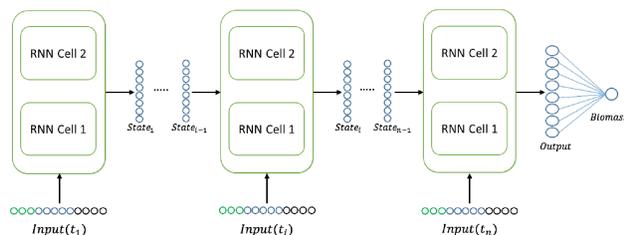


Figure 3. Recurrent neural network developed in this study for biomass prediction.

The input vector for the proposed RNN includes hyperspectral features, LiDAR features, and weather related features as shown in Figure 4.

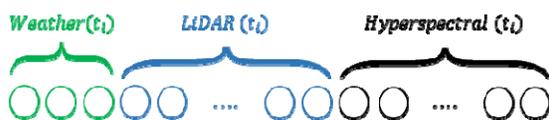


Figure 4. The input vector at time t_i for the proposed RNN.

4. Data and Experimental Setup

4.1. Field Ground Reference Data

The field experiments were conducted over approximately 2.8 ha sorghum breeding trials in different fields of the Purdue Agronomy Center for Research and Education (ACRE) at Purdue University in Indiana. The field experiments consisted of three distinct trials in four panels; the hybrid calibration (HyCal) panel, the inbred calibration (InCal) panel, the Sorghum Biodiversity (SbDiv) panel, and the Sorghum Biodiversity test cross

(SbDiv-tc) panel. The HyCal panel consisted of 18 commercial hybrids from distinct market classes including both forage and grain sorghum replicated four times in twelve row plots in both the 2017 and 2018 growing seasons. The InCal panel included 60 and 54 inbred lines in 2017 and 2018, respectively, a subset of the SbDiv panel that encompasses the genetic diversity of the SbDiv panel, which were replicated twice in four row plots in both years. The SbDiv panel consisted of 840 inbred lines replicated twice in four row plots in the 2017 growing season. The SbDiv-tc included 630 hybrids derived from the SbDiv panel with a common maternal parent replicated twice in four row plots in the 2018 growing season. Details of the experiments are provided in Table 1.

Table 1. Experimental design for the 2017 and 2018 growing seasons.

Trial	Year	Genotype	# of plots	# of varieties	Sowing Date	Harvest Date
HyCal	2017	hybrid	72	18	May 15	Sep 31
	2018	hybrid	72	18	May 1	Aug 9
InCal	2017	inbred	120	60	May 16	Oct 15
	2018	inbred	108	54	May 1	Aug 9
SbDiv	2017	inbred	1800	840	May 17	Oct 15
SbDiv-tc	2018	hybrid	1600	630	May 8	Aug 14

4.2. Remote Sensing Data

Hyperspectral and LiDAR remote sensing data were collected from UAVs. All remote sensing data acquisition platforms were flown with Global Navigation Satellite System/Inertial Navigation System (GNSS/INS) units for direct georeferencing. Visible Near-Infrared (VNIR) hyperspectral data were collected with a Headwall Photonics Nano-Hyperspec pushbroom scanner, which acquires data in 272 spectral bands at 2.2 nm/band from 400 nm to 1000 nm. LiDAR data were collected with a Velodyne VLP-16 3D LiDAR sensor. The dates of the remote sensing data sets are listed in Table 2.

Table 2. Remote Sensing Data Sets Details

Year	Data Type	Dates
2017	LiDAR	7/07, 7/12, 7/25, 8/08, 8/23
	Hyperspectral	7/04, 7/14, 7/25, 8/08, 8/23
2018	LiDAR	7/02, 7/18, 7/23, 8/01, 8/06
	Hyperspectral	7/03, 7/18, 7/25, 7/30, 8/09

From each LiDAR point cloud and for each plot, 21 features including 30th, 50th, 75th, 90th, 95th, 99th, and 100th percentiles of height, canopy cover, and multiple statistical features were extracted. From hyperspectral data, the average of the respective values associated with the vegetation pixels extracted via pixel-based classification within a row in a given plot were computed and used as input features for further analysis. More information about pre-processing of the data is available at [17].

4.3. Experimental Setup

Results of two experiments are presented: the proposed method for feature learning is investigated initially. For each experiment, the data from all other experiments are used to train the proposed the FAEN and CAEN models, then the learned features are evaluated as the input features in SVR models to predict end of season biomass for that experiment. The robustness of performance of the SVR models and the proposed RNN is evaluated by training these models on one experiment and using the models to predict biomass of all other experiments.

5. Results

5.1. Feature Learning Results

The results of feature learning include both the proposed FAEN and CAEN architectures. For training, all the features from all the multi-variety plots and all five available dates are stacked in an input vector. For example, for the HyCal 2017, there are 72 plots (from Table 1) and five dates, so $5 \times 72 = 360$ sample data points are available for training. Once training is completed, the trained network

can be used to extract features for other data sets (any experiment at any date).

Figure 5 (a) shows the original, reconstructed, and learned features from hyperspectral data for one of the varieties in the HyCal 2017 experiment using FAEN trained on the same data sets. Figure 5 (b) shows the same features using the same network, but training was performed on the SbDiv 2017 data sets. The Root Mean Square Error (RMSE) of the reconstruction is also provided in Figure 5. The reconstructed features are very similar to the original features in both training scenarios based on the small RMSE values observed for each reconstruction. This indicates that it is possible to train the network on some experiments, and then use the trained network for other experiments. This is important for two reasons: 1) if there are experiments where the number of samples for training is not adequate, pre-trained networks can be used; 2) this provides the opportunity to use a trained network to extract features for all the experiments in both 2017 and 2018, which represent a broader range of environmental conditions. This can potentially facilitate transfer learning required for multi-year predictions.

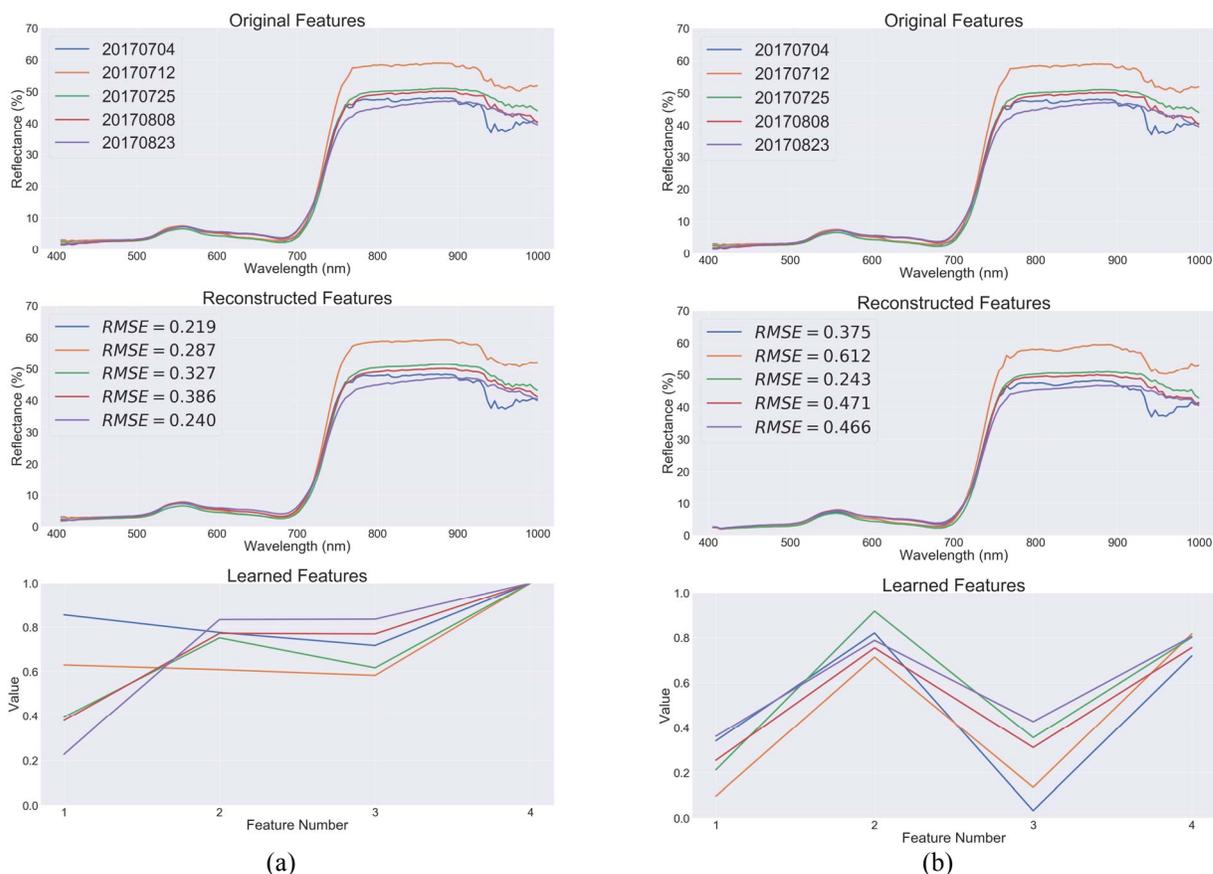


Figure 5. Original, reconstructed, and learned features for one variety of the HyCal 2017 experiment using FAEN trained on the (a) HyCal 2017 and (b) SbDiv 2017 hyperspectral data sets.

To further investigate these issues, the FAEN and CAEN are trained on each experimental data set first, and the trained networks are then used to learn the features for all the other experiments. The learned features for each experiment were then used in an SVR model to predict the end-of-season biomass for that experiment using 3-fold cross validation. The R^2 s of the predictions compared to ground reference data are provided in Tables 3 and 4. In most cases, the highest accuracies were obtained when the training and testing were both associated with the same experiment (panel). However, for all experiments using the networks trained on data sets from other experiments provided comparable results. In some cases, training using the data from the combined experiments (last column of Tables 3 and 4) provided the most accurate prediction. These results indicate that it may not be necessary to train the proposed networks on the data sets of the same experiment for biomass prediction. It should be noted however, that it is critical to use high quality data for training and evaluating networks over multiple panels and time periods.

Table 3. R^2 values of SVR-based biomass prediction (of field based results) using latent features extracted by the proposed FAEN trained on different experimental data sets and tested on all experiments.

Train \ Test		Data Type	2017			2018			All
			HyCal	InCal	SbDiv	HyCal	InCal	SbDiv- tc	
2017	HyCal	H	0.63	0.63	0.47	0.56	0.52	0.57	0.53
		L	0.65	0.68	0.65	0.67	0.66	0.61	0.68
		HL	0.78	0.78	0.78	0.72	0.75	0.70	0.75
	InCal	H	0.58	0.63	0.63	0.59	0.64	0.62	0.64
		L	0.49	0.56	0.54	0.54	0.56	0.51	0.51
		HL	0.64	0.70	0.57	0.56	0.64	0.62	0.60
	SbDiv	H	0.62	0.64	0.65	0.62	0.64	0.65	0.67
		L	0.62	0.61	0.62	0.61	0.60	0.60	0.62
		HL	0.71	0.72	0.73	0.70	0.72	0.72	0.74
2018	HyCal	H	0.71	0.69	0.55	0.56	0.61	0.52	0.60
		L	0.76	0.69	0.62	0.73	0.72	0.65	0.70
		HL	0.77	0.75	0.60	0.75	0.73	0.69	0.79
	InCal	H	0.59	0.53	0.56	0.46	0.50	0.57	0.50
		L	0.48	0.51	0.54	0.57	0.58	0.41	0.58
		HL	0.69	0.61	0.63	0.59	0.66	0.62	0.66
	SbDiv-tc	H	0.44	0.44	0.43	0.42	0.42	0.43	0.43
		L	0.36	0.35	0.35	0.36	0.36	0.34	0.35
		HL	0.45	0.45	0.44	0.43	0.43	0.43	0.43

** H refers to hyperspectral, L refers to LiDAR, and HL is combined hyperspectral and LiDAR data sets

For most experiments in Tables 3 and 4, the combined set of learned features from LiDAR and hyperspectral data provided more accurate results than using features from only hyperspectral or LiDAR. This indicates the importance of using both geometric and spectral related features for biomass predictions. For HyCal experiments in both 2017 and 2018, LiDAR-based features provided more accurate results than hyperspectral-based features, while for all other experiments, hyperspectral-based features provided more accurate results. This can be explained by considering the differences in the genotypes planted in different experiments. The HyCal experiments consisted of 18 commercial genotypes with a wide range of structural characteristics that can be reflected in LiDAR-based geometric features, while the InCal and SbDiv experiments had more similar genotypes, and thereby more similar plant structure.

Table 4. R^2 values of SVR-based biomass prediction (of field based results) using latent features extracted by the proposed CAEN trained on different experimental data sets and tested on all experiments.

Train \ Test		Data Type	2017			2018			All
			HyCal	InCal	SbDiv	HyCal	InCal	SbDiv- tc	
2017	HyCal	H	0.60	0.57	0.52	0.58	0.56	0.59	0.47
		L	0.70	0.67	0.61	0.66	0.53	0.72	0.67
		HL	0.67	0.71	0.65	0.64	0.56	0.54	0.71
	InCal	H	0.65	0.67	0.69	0.64	0.65	0.64	0.69
		L	0.55	0.53	0.55	0.53	0.54	0.58	0.55
		HL	0.59	0.51	0.58	0.49	0.52	0.55	0.58
	SbDiv	H	0.68	0.71	0.70	0.67	0.69	0.68	0.69
		L	0.61	0.58	0.62	0.60	0.58	0.63	0.62
		HL	0.74	0.74	0.75	0.74	0.74	0.76	0.75
2018	HyCal	H	0.65	0.64	0.63	0.64	0.65	0.64	0.69
		L	0.72	0.59	0.50	0.74	0.72	0.63	0.64
		HL	0.73	0.68	0.58	0.72	0.72	0.66	0.62
	InCal	H	0.58	0.60	0.59	0.59	0.54	0.50	0.58
		L	0.59	0.63	0.63	0.60	0.60	0.61	0.53
		HL	0.57	0.62	0.64	0.65	0.65	0.60	0.61
	SbDiv-tc	H	0.46	0.47	0.47	0.45	0.47	0.47	0.46
		L	0.37	0.36	0.37	0.36	0.37	0.35	0.36
		HL	0.48	0.49	0.48	0.48	0.49	0.49	0.48

** H refers to hyperspectral, L refers to LiDAR, and HL is combined hyperspectral and LiDAR data sets

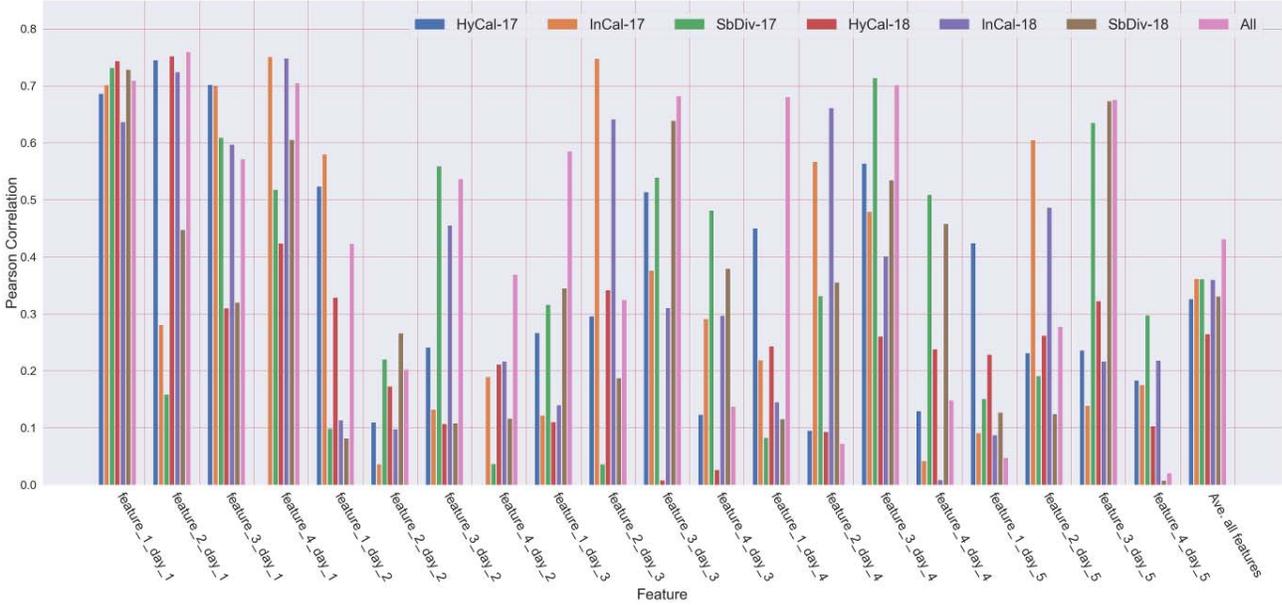


Figure 6. Pearson correlation coefficient between the biomass data and hyperspectral features learned by FAEN trained on each experiment individually and all experiments. The correlations are provided for each feature for each date, as well as the average of the correlation.

Figure 6 shows the Pearson correlation coefficients between the end of season biomass and hyperspectral features learned by FAEN trained on data from different experiments individually and tested on all experiments. The figure shows that 1) the correlation between input features 1-4 from each day of the season and biomass can be very different, 2) the features from the first day of the season were more correlated with the final biomass than later season features, 3) the average of the correlation of all the features with biomass shows that when training, using all the trials provided the features most related to the biomass which is consistent with R^2 results provided in Table 3.

The highest R^2 value for each experiment using FAEN and CAEN is provided in Table 5, which shows the FAEN provided slightly more accurate results for the HyCal and InCal in both years while CAEN resulted higher R^2 for the SbDiv in 2017 and SbDiv-tc in 2018. A possible explanation for these differences, although small, could be that for the SbDiv experiments, more sample data are available than the other experiments (as more plots are planted in these experiments as noted on Table 1) implying that the CAEN might require more samples to train than FAEN.

Table 5. The highest R^2 of the SVR predictions using the features learned using the proposed FAEN vs. CAEN from Tables 3 and 4.

Experiment		FAEN	CAEN
2017	HyCal	0.78	0.72
	InCal	0.70	0.69
	SbDiv	0.74	0.76
2018	HyCal	0.79	0.74
	InCal	0.69	0.65
	SbDiv-tc	0.45	0.49

5.2. Multi-year Biomass Prediction

In the previous section, the training and validation of the SVR models was performed on sample data of the same experiment. In this section, the SVR and RNN models were trained for each experiment, and the models are then used for biomass prediction of all the experiments. For all predictions in this section, the features learned from the proposed FAEN (trained on all the data sets of all experiments) were used as the input features for the predictive models.

Table 6 shows the results of biomass prediction using SVR models training on each experiment and validating on all other experiments. The highest R^2 for each experiment was achieved when the training and validation were performed on the same experiment (splitting the sample data into training and validation categories and using 3-fold cross validation). Figure 7 shows the prediction results of an SVR model that was trained with the HyCal 2017 data and tested on the InCal 2017 and HyCal 2018 data. In Figure 7-(b), although the R^2 is zero, a relationship between the predicted and ground reference biomass data is obvious.

Tables 7 and 8 show the results of similar predictions with the proposed RNN in Figure 3 with GRUs (RNN-GRU) and LSTMs (RNN-LSTM) as the RNN cells, respectively. The number of iterations was set to 100,000, but the model with the highest R^2 was identified after training. Figure 8 shows the training loss, training and test accuracies for RNN-GRU when training and testing were performed on the InCal 2018 and SbDiv 2017 experiments, respectively.

Table 6. The R^2 value (relative to field based measurements) of biomass predictions using SVR

Train \ Test		2017			2018		
		HyCal	InCal	SbDiv	HyCal	InCal	SbDiv-tc
2017	HyCal	0.69	0.27	-6.89	-3.09	-1.35	0.64
	InCal	0.20	0.64	-14.86	-8.98	-4.90	-0.98
	SbDiv	0.52	0.56	0.74	-46.83	-12.73	-13.19
2018	HyCal	-0.50	-2.52	-4.90	0.78	-0.01	-1.01
	InCal	0.34	0.06	-0.80	-0.67	0.70	-0.86
	SbDiv-tc	0.26	-0.15	-0.95	-5.38	-0.58	0.43

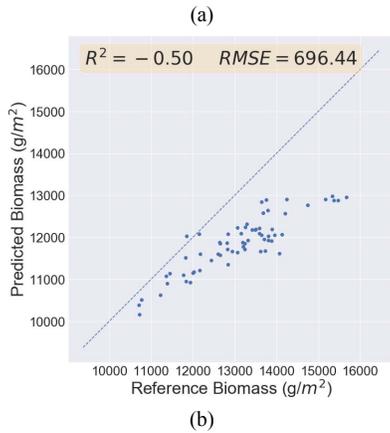
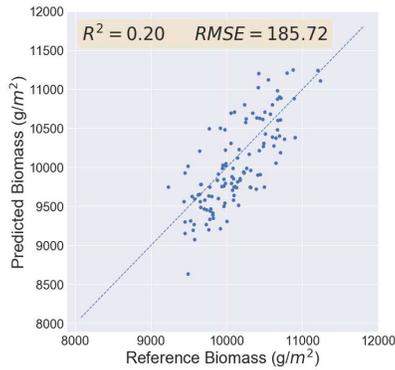


Figure 7. Biomass prediction results of SVR model trained on the HyCal 2017 experiment and tested on (a) InCal 2017 and (b) HyCal 2018.

From Tables 6, 7, and 8, when the training and validation are performed with different experiments of the same year, predictions are more accurate compared to training and validation with different experiments of different years. Also, training on the 2017 experiments and validating on 2018 experiments yielded more accurate predictions than training on the 2018 experiments and validating on the 2017 experiments. In general, the R^2 of predictions with RNNs are higher than those with SVRs. The possible explanations for this are that RNNs are more powerful than SVRs in modeling time series data, and that with RNN, the training

can be controlled to avoid the overfitting by early termination of the training. Comparing Tables 7 and 8 shows that RNN-GRU provided marginally better results than RNN-LSTM which could be because LSTM is more complex than GRU and requires learning of more parameters.

From Tables 6 to 8, the predictions for the HyCal experiment in both 2017 and 2018 and the SbDiv experiment in 2017 are more accurate than other experiments. The reason could be that genotypes in the HyCal experiments are more varied in terms of physical structural characteristics, and using LiDAR-based features in the models resulted in more accurate predictions for this experiment (similar to the results in Tables 3 and 4). For the SbDiv in 2017, although the genotypes are not as varied as HyCal in terms of physical structural characteristics, the predictions are comparable to those from the HyCal experiments which could be attributed to the greater number of plots (and therefore number of samples) than other experiments, providing a better training for the regression models. The reason(s) for poor predictions for SbDiv-tc is unclear to the authors.

Table 7. The R^2 of biomass predictions using RNN-GRU

Train \ Test		2017			2018		
		HyCal	InCal	SbDiv	HyCal	InCal	SbDiv-tc
2017	HyCal	0.72	0.46	0.79	0.00	0.25	0.62
	InCal	0.58	0.64	0.60	0.00	0.00	0.35
	SbDiv	0.52	0.55	0.72	0.00	0.55	0.33
2018	HyCal	0.46	0.00	0.00	0.84	0.11	0.01
	InCal	0.33	0.55	0.00	0.40	0.60	0.41
	SbDiv-tc	0.25	0.35	0.00	0.05	0.33	0.45

Table 8. The R^2 of biomass predictions using RNN-LSTM

Train \ Test		2017			2018		
		HyCal	InCal	SbDiv	HyCal	InCal	SbDiv-tc
2017	HyCal	0.63	0.29	0.71	0.00	0.26	0.58
	InCal	0.57	0.65	0.59	0.00	0.00	0.42
	SbDiv	0.28	0.60	0.74	0.00	0.50	0.00
2018	HyCal	0.56	0.00	0.00	0.75	0.00	0.00
	InCal	0.37	0.55	0.00	0.36	0.49	0.33
	SbDiv-tc	0.22	0.34	0.00	0.00	0.25	0.43

6. Conclusions and Future Work

Two deep learning based networks were investigated in this study for unsupervised feature learning from hyperspectral and LiDAR data. A recurrent neural network

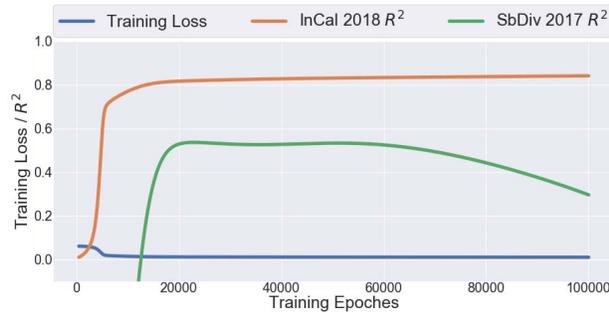


Figure 8. Training the proposed RNN-GRU on the InCal 2018 and validating on the SbDiv 2017.

was then developed with hyperspectral, LiDAR, and weather data inputs to predict sorghum biomass. The results of using the proposed network for training on one experiment and predicting biomass for other experiments with different types of sorghum varieties illustrate the potential of the network for biomass prediction, and the challenges relative to small sample sizes, including weather and sensitivity to the associated ground reference information. In ongoing studies, the feature learning networks with other possible variations, such as increasing the number of hidden layers and changing the hyper-parameters as well as using RNNs for hyperspectral data is being tested. Also, RNN architecture is being tested on prediction of biomass by simultaneously training on multiple years of data.

7. Acknowledgment

This work was supported by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy under Grant DE-AR0000593. The authors thank the Purdue TERRA team, especially Evan Flatt for his work on system integration, flying, and data processing; Meghdad Hasheminasab, Tian Zhao, Magdy Elbahnasawy, Tamer Shamseldin, Radhika Ravi, Yun-Jou Lin, and Yi Chun Lin for their work on collecting and processing the RGB and LiDAR data; Karoll Quijano, Ruya Xu, Taojun Wang, Behrokh Nazari, and Zhou Zhang for their work on hyperspectral data collection and processing; professor Ayman Habib for his valuable input throughout this work.

References

- [1] J. W. Davey, P. A. Hohenlohe, P. D. Etter, J. Q., Boone, J. M. Catchen, Genome-wide genetic marker discovery and genotyping using next-generation sequencing, *Nature Reviews Genetics*, vol. 12, no. 7, pp. 499-510, 2011.
- [2] F. E. Fassnacht, F. Hartig, H. Lati, C. Berger, J. Hernández, P. Corvalán, and B. Koch, Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass, *Remote Sensing of Environment*, vol. 154, pp. 102-114, 2014.

- [3] J. Schmidhuber, Deep Learning in neural networks: An overview, *Neural Networks*, vol. 61, pp. 85-117, 2015.
- [4] Y. Lecun, Y. Bengio, and G. Hinton, Deep learning, *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [5] S. Hoo-Chang, M. R. Orton, D. J. Collins, S. J. Doran, M. O. Leach, Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1930-1943, Aug. 2013.
- [6] Y. C. Tang, C. Eliasmith, Deep networks for robust visual recognition, *Proc. ICML*, pp. 1055-1062, 2010.
- [7] C. Farabet, C. Couprie, L. Najman, Y. LeCun, Learning hierarchical features for scene labeling, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1915-1929, Aug. 2013.
- [8] T. Chao, P. Hongbo, and L. Yansheng, Unsupervised Spectral-Spatial Feature Learning with Stacked Sparse Autoencoder for Hyrspectral Imagery Classification, *IEEE Geoscience and Remote Sensing Letters*, vol.12, pp. 2438-2442, 2015.
- [9] F. Li, H. Qiao, B. Zhang, and X. Xi, Discriminatively boosted image clustering with fully convolutional auto-encoders, *Pattern Recognition*, vol. 83, pp. 161-173, 2018.
- [10] N. Cristianini and J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods. Cambridge university press, 2000.
- [11] A. J. Smola and B. Schlkopf, A tutorial on support vector regression, *Statistics and Computing*, vol. 14, pp. 199-222, 2004.
- [12] L. Mou, S. Member, P. Ghamisi, X. Xiang Zhu, and S. Member, Deep Recurrent Neural Networks for Hyperspectral Image Classification, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, p. 3639, 2017.
- [13] Y. L. Kong, Q. Huang, C. Wang, J. Chen, J. Chen, and D. He, Long short-term memory neural networks for online disturbance detection in satellite image time series, *Remote Sensing*, vol. 10, no. 3, pp. 1-13, 2018.
- [14] J. You, X. Li, M. Low, D. Lobell, and S. Ermon, Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data, in *Proceedings of the Thirty-First AAAI conference on Artificial Intelligence (AAAI-17)*. Association for the Advancement of Artificial Intelligence, 2017, pp. 4559-4566.
- [15] A. Karpathy, J. Johnson, and L. Fei-Fei. Visualizing and understanding recurrent networks, Cornell Univ. Lab. 2015, arXiv:1506.02078.
- [16] R. Fu, Z. Zhang, and L. Li, Using LSTM and GRU neural network methods for traffic flow prediction, *31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2016, pp. 324-328.
- [17] 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 *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. Valencia, Spain, 2018, pp. 7719-7722.