YIELDNET: A CONVOLUTIONAL NEURAL NETWORK FOR SIMULTANEOUS CORN AND SOYBEAN YIELD PREDICTION BASED ON REMOTE SENSING DATA

A PREPRINT

Saeed Khaki
Industrial Engineering Department
Iowa State University
Ames, Iowa 50011, USA
skhaki@iastate.edu

Hieu Pham
Syngenta
Slater, Iowa 50244, USA
hieu.pham@syngenta.com

Lizhi Wang
Industrial Engineering Department
Iowa State University
Ames, Iowa 50011, USA

December 5, 2020

ABSTRACT

Large scale crop yield estimation is, in part, made possible due to the availability of remote sensing data allowing for the continuous monitoring of crops throughout its growth state. Having this information allows stakeholders the ability to make real-time decisions to maximize yield potential. Although various models exist that predict yield from remote sensing data, there currently does not exist an approach that can estimate yield for multiple crops simultaneously, and thus leads to more accurate predictions. A model that predicts yield of multiple crops and concurrently considers the interaction between multiple crop's yield. We propose a new model called YieldNet which utilizes a novel deep learning framework that uses transfer learning between corn and soybean yield predictions by sharing the weights of the backbone feature extractor. Additionally, to consider the multi-target response variable, we propose a new loss function. Numerical results demonstrate that our proposed method accurately predicts yield from one to four months before the harvest, and is competitive to other state-of-the-art approaches.

***K*eywords** Crop yield · Remote Sensing · Deep learning · Transfer learning

1 Introduction

The use of satellites and other remote sensing mechanisms has proven to be vital in monitoring the growth of various crops both spatially and temporally [1]. Common uses of remote sensing are to extract various descriptive indices such as normalized difference vegetation index (NDVI), temperature condition index, enhanced vegetation index, and leaf area index [2]. This information can be used for drought monitoring, detecting excessive soil wetness, quantifying weather impacts on vegetation, the evaluation of vegetation health and productivity, and for crop yield forecasting [3, 4, 5]. Moreover, large scale collection of environmental descriptors such as surface temperature and precipitation can also be identified from satellite images [6, 7].

Accurate and non-destructive identification of crop yield throughout a crop's growth stage enables farmers, commercial breeding organizations, and government agencies to make decisions to maximize yield output, ultimately for a nation's economic benefit. Early estimations of yield at the field/farm level plays a vital role in crop management in terms of site specific decisions to maximize a crop's potential. Numerous approaches have been applied for crop yield prediction, such as manual surveys, agro-meteorological models, and remote sensing based methods [8]. However, given the size of a farming operations, manual field surveys are neither efficient nor practical. Moreover, modeling approaches have limitations due to the difficulty in determining model parameters, the scalability to large areas, and the computational resources neccessary [9].

Recently, satellite data has become widely available in various spatial, temporal and spectral resolutions and can be used to predict crop yield over different geographical locations and scales [10]. The availability of Earth observation (EO) data has created new ways for efficient, large-scale agricultural mapping [11]. EO enables a unique mechanism to capture crop information over large areas with regular updates to create maps of crop production and yield. Nevertheless, due to the high spatial resolution needed for accurate yield predictions, Unmanned Aerial Vehicles (UAV) have been promoted for data acquisition [12]. Although UAV platforms have demonstrated superior image capturing abilities, without the assistance of a large workforce, accurately measuring yield for large regions is not feasible.

Remote sensing capabilities have been extensively used for estimating crop yield around the world in various scales (field, county, state). Various studies have been performed using various vegetation indices to estimate yield in maize, wheat, grapes, rice, corn, and soybeans using random forests, neural networks, multiple linear regression, partial least squares regression, and crop models [13, 14, 15, 16, 17, 18, 19]. These research papers showcase the potential, power, and accuracy remote sensing has on estimating yield at a large scale.

For the purpose of this paper, we consider two main crops in the Midwest United States (U.S.), corn (Zea mays L.) and soybeans (*Glycine max*). According to the United States Department of Agriculture, in 2019, 89.7 million and 76.1 million acres of corn and soybean were planted, respectively. The combination of these two crops make up approximately 20% of active U.S. farmland [20]. Given the shear size of these farming operations combined with the growing population, actions must be taken to maximize yield. With the help of remote sensing, government officials as well as farmers can monitor crop growth (at various scales) to ensure proper crop management decisions are made to maximize yield potential.

Although approaches exist to estimate yield with various means, current models are limited in that they only estimate yield for a single crop. That is, there does not exist any model making use of remote sensing data to predict the yield of multiple crops simultaneously. Having such a model would result in more accurate predictions by considering the interactions between crops and estimates yields accordingly. Pragmatically, because predictions are simultaneous, the results are less computationally intensive than computing a model for each crop.

We propose a new deep learning framework called YieldNet utilizing a novel deep neural network architecture. This model makes use of a transfer learning methodology that acts as a bridge to share model weights for the backbone feature extractor. Moreover, because of the uniqueness associated to simultaneous yield prediction for two crops, we propose a new loss function that can handle multiple response variables.

Deep learning methods are representation learning methods which can automatically extract relevant features from raw input data. Deep learning models stack multiple non-linear modules which can learn very complex functions [21]. Deep learning methods have extensively been used to solve different problems in the field of agriculture such as crop yield prediction [22, 23, 24, 25, 26, 27, 28], image-based plant phenotyping [29, 30, 31, 32], and drought tolerance classification [33]. These deep learning methods achieved high accuracy and improved the state-of-the-art in the field.

2 Methodology

The goal of this paper is to simultaneously predict the average yield per unit area of two different crops (corn and soybean) both grown in regions of interest (U.S. counties) based on a sequence of images taken by satellite before harvest. Let $I^t_{l,k} \in \mathbb{R}^{H\times W\times d}$ and $Y^c_{l,k} \in \mathbb{R}^+$ denote, respectively, the remotely sensed multi-spectral image taken at time $t \in \{1, ..., T\}$ during the growing season and the ground truth average yield of crop type $c \in \{1, 2\}$ planted in location l at year k, where H and W are the image's height and width and d is the number of bands (channels). Thus, the dataset can be represented as

$$D = \left\{ \left((I^1_{l,k}, I^2_{l,k}, ..., I^T_{l,k}), (Y^1_{l,k}, Y^2_{l,k}) \right) \middle| l \in \{1, ..., L\},\ k \in \{1, ..., K\} \right\},$$

where L and K are the numbers of locations and years, respectively.

For our approach, we do not use end-to-end training due to the main following reasons: (1) the number of labeled training data is limited, and (2) inability to use transfer learning from popular benchmark datasets such as Imagenet due to domain difference and multi-spectral nature of satellite images. Therefore, we reduce the dimensionality of the raw remote sensing images under permutation invariance assumption which states the average yield mostly depends on the number of different pixel types rather than the position of the pixels in images due to infeasibility of end-to-end training. Similar approaches have been used in other studies [25, 34]. As a result, we separately discretize the pixel values for each band of a multi-spectral image $I^t_{l,k}$ into b bins and obtain a histogram representation $h^t_{l,k} \in \mathbb{R}^{b\times d}$. If we obtain histogram representations of the sequence of multi-spectral images denoted as $(h^1_{l,k}, h^2_{l,k}, ..., h^T_{l,k})$ and concatenate them through time dimension, we can produce a compact histogram representations $H_{l,k} \in \mathbb{R}^{T\times b\times d}$ of the sequence of

multi-spectral images taken during growing season. As such, the dataset D can be re-written as the following and we will use this notation in the rest of the paper:

$$D = \left\{ \left(H_{l,k}, (Y^1_{l,k}, Y^2_{l,k}) \right) \middle| l \in \{1, ..., L\},\ k \in \{1, ..., K\} \right\}.$$

Given the above dataset, this paper proposes a deep learning based method, named YieldNet, that learns the desired mapping $H_{l,k} \in \mathbb{R}^{T \times b \times d} \mapsto (Y^1_{l,k}, Y^2_{l,k})$ to predict the yields of two different crops simultaneously.

2.1 Network Architecture

Crop yield is a highly complex trait which is affected by many factors such as environmental conditions and crop genotype which requires a complex model to reveal the functional relationship between these interactive factors and crop yield.

As such, we propose a convolutional neural network model which is a highly non-linear and complex model. Convolutional neural networks belong to the class of representation learning methods which automatically extract necessary features from raw data without the need of any handcrafted features [21, 35]. Figure 1 outlines the architecture of the proposed method. Table 1 shows the detailed architecture of the proposed model. We use 2-D convolution operation which is performed over the 'time' and 'bin' dimensions while considering bands as channels. As such, convolution operation over the time dimension can help capture the temporal effect of satellite images collected over time intervals.

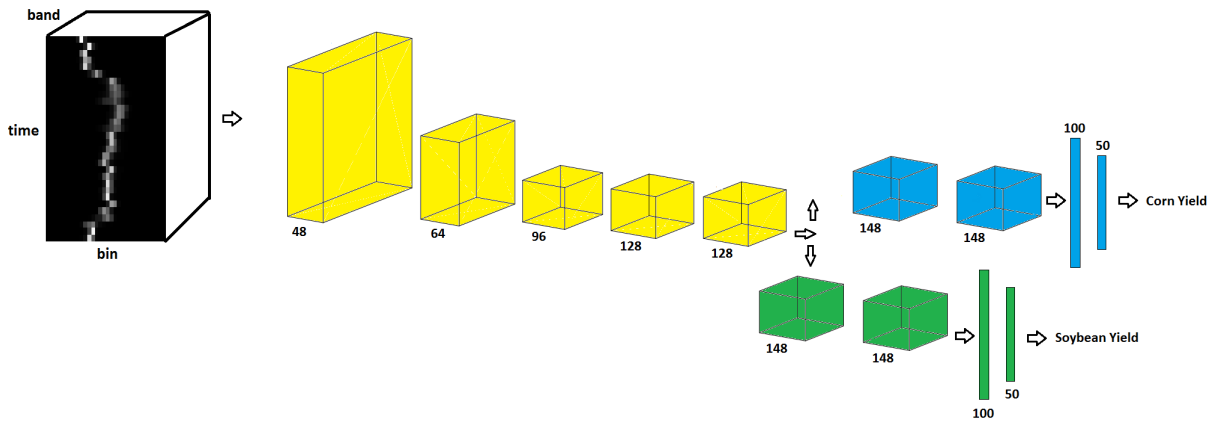

Figure 1: Outline of the YieldNet architecture. The input of the YieldNet is 3-D histograms $H \in \mathbb{R}^{T \times b \times d}$ and the 2-D convolution operation is performed over the 'time' and 'bin' dimensions while considering bands as channels. The number under feature maps indicate the number of channels. Convolutional layers denoted with yellow color work as a backbone feature extractor and share weights for both corn and soybean yield predictions.

Type/Stride	Padding	Filter Size	Number of Filters
Conv/s2	valid	7×7	48
Conv/s2	valid	5×5	64
Conv/s2	same	5×5	96
Conv/s1	same	3×3	128
Conv/s1	same	3×3	128
Corn Head		Soybean Head	
Conv/s1 - same - 3×3 - 148		Conv/s1 - same - 3×3 - 148	
Conv/s1 - same - 3×3 - 148		Conv/s1 - same - 3×3 - 148	
FC-100		FC-100	
FC-50		FC-50	

Table 1: YieldNet architecture. The first 5 convolutional layers work as a backbone feature extractor and share weights for both corn and soybean yield predictions.

In our proposed network, the first 5 convolutional layers share weights for corn and soybean yield predictions which are denoted with yellow color in Figure 1. These layers extract relevant features from input data for both corn and soybean yield predictions. The intuition behind using one common feature extractor is that it significantly decreases the number of parameters of the network, which helps training the model more efficiently given the scarcity of the labeled data. In addition, many low-level features captured by the comment feature extractor reflect general enviromental conditions that are transferable between corn and soybean yields. All convolutional layers are followed by batch normalization [36] and ReLU nonlinearities in our proposed network. Batch normalization is used to accelerate the training process by reducing internal covariate shift and regularizing the network. We use two convolutional layers in both corn and soybean heads which are followed by two fully connected layers.

2.2 Network Loss

To jointly minimize the prediction errors of corn and soybean yield forecasting, we propose the following loss function:

$$L = \max\left(\frac{1}{N_c}\sum_{i=1}^{N_c}\left(\frac{Y_i^c - \hat{Y}_i^c}{\bar{Y}^c}\right)^2, \frac{1}{N_s}\sum_{i=1}^{N_s}\left(\frac{Y_i^s - \hat{Y}_i^s}{\bar{Y}^s}\right)^2\right) \tag{1}$$

Where Y_i^c, $\hat{Y}_i^c$, Y_i^s, $\hat{Y}_i^s$, N_c, N_s, $\bar{Y}^c$, and $\bar{Y}^s$ denote ith average ground truth corn yield, ith predicted corn yield, ith average ground truth soybean yield, ith predicted soybean yield, number of corn samples, number of soybean samples, average corn yield, and average soybean yield, respectively. Our proposed loss function is a normalized Euclidean loss which makes the corn and soybean losses the same scale. We use the maximum function in our proposed loss function to make the training process more stable and ensure that both corn and soybean losses are optimized.

3 Experiments and Results

In this section, we present the dataset used in our study and then report the results of our proposed method along with other competing methods in corn and soybean yield predictions. We conducted all experiments in Tensorflow [37] on a NVIDIA Tesla V100 GPU.

3.1 DATA

The data analyzed in this study included two sets: yield performance and satellite images.

- The yield performance dataset includes the observed county-level average yield for corn and soybean between 2004 and 2018 across 1,132 counties for corn and 1,076 for soybean within 13 states of the U.S. Corn Belt: Indiana, Illinois, Iowa, Minnesota, Missouri, Nebraska, Kansas, North Dakota, South Dakota, Ohio, Kentucky, Michigan, and Wisconsin, where corn and soybean are considered the dominant crops. Figures 2 and 4 depict the US Corn Belt and histogram of yields, respectively. The summary statistics of the corn and soybean yields are shown in Table 2.

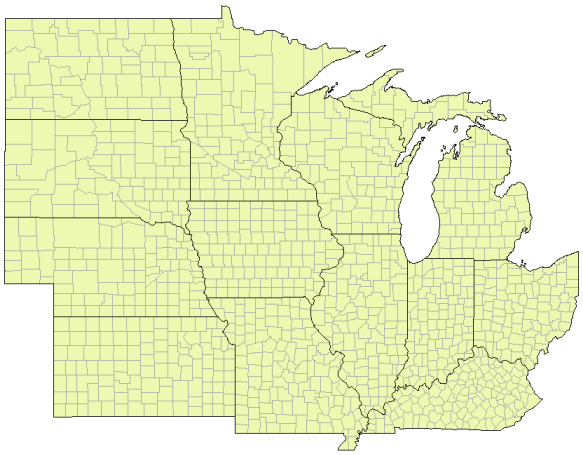

Figure 2: The map of US Corn Belt.

- Satellite data contains MODIS products including MOD09A1 and MYD11A2. MOD09A1 product provides an estimate of the surface spectral reflectance of Terra MODIS bands 1-7 at 500m resolution and corrected for atmospheric conditions [38]. The MYD11A2 product provides an average land surface temperature which has the day and night-time surface temperature bands at 1km resolution [39]. Figure 3 depicts an example multispectral image of land surface temperature and surface spectral reflectance for Adams county in Illinois. These satellite images were captured at 8-days intervals and we only use satellite images captured during growing seasons (March to October). As such, satellite images are collected 30 times a year in our study. We discretize all multispectral images using 32 bins to generate the 3-D histograms $H \in \mathbb{R}^{T \times b \times d}$, where $T = 30$, $b = 32$, and $d = 9$.

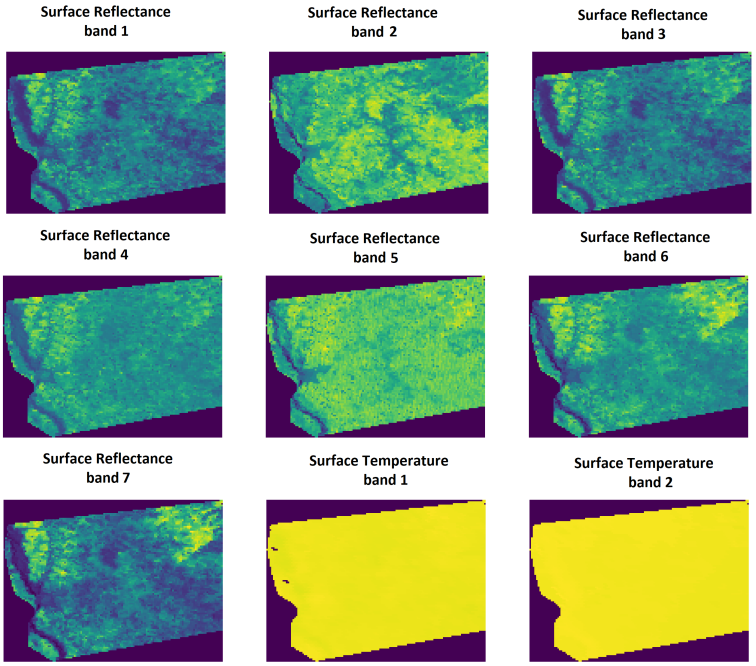

Figure 3: The multispectral images of surface spectral reflectance and land surface temperature for Adams county in Illinois captured on May 8th, 2011.

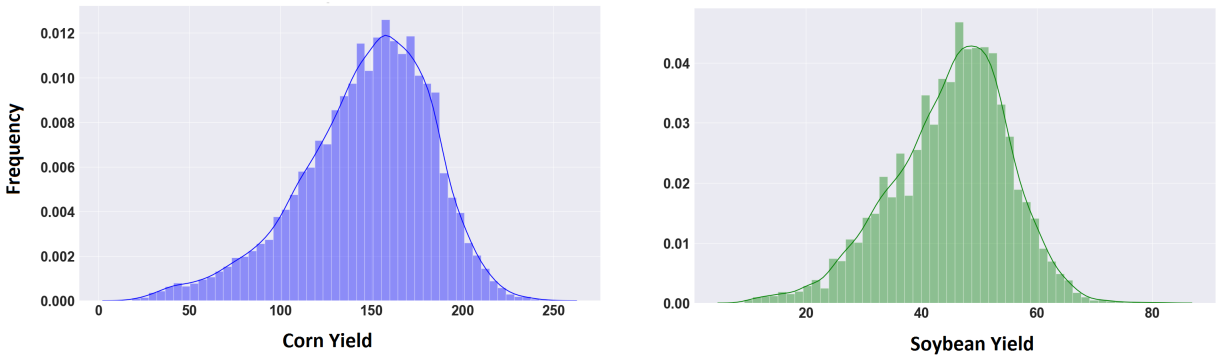

Figure 4: The histograms of the corn and soybean yields. The unit of yield is bushels per acre.

Summary statistics	Corn	Soybean
Year range	2004-2018	2004-2018
Average yield	146.68	45.02
Standard deviation of yield	36.03	10.08
Minimum yield	18.3	9.3
Maximum yield	246.7	82.3
Number of locations	1,132	1,076
Number of observations	13,992	12,502

Table 2: The summary statistics of yield data across all years. The unit of yield is bushels per acre.

3.2 Design of Experiments

We compare our proposed method with the following models to evaluate the efficiency of our method:

Random forest (RF) [40]: RF is a non-parametric ensemble learning method which is robust against overfitting. We set the number and the maximum depth of trees in RF to be 150 and 20, respectively. We tried different number and the maximum depth of trees and found that these hyperparameters resulted in the most accurate predictions.

Deep feed forward neural network (DFNN): DFNN is a highly nonlinear model which stacks multiple fully connected layers to learn the underlying functional relationship between inputs and outputs. The DFNN model has 9 hidden layers, each having 50 neurons. We used batch normalization for all layers. The ReLU activation function was used in all layers except the output layer. The model was trained for 120,000 iterations. Adam optimizer was used to minimize the Euclidean loss function.

Regression tree (RT) [41]: RT is a nonlinear model which does not make any assumption regarding the mapping function. We set the maximum depth of the regression tree to be 12 to decrease the overfitting.

Lasso [42]: Lasso adopts L_1 norm to reduce the effect of unimportant features. Lasso model also can be used as a baseline model to compare the linear and nonlinear effect of input features. We set the L_1 coefficient of Lasso model to be 0.05.

Ridge [43]: Ridge regression is similar to the Lasso model except it uses L_2 norm as a penalty to reduce the effect of unimportant features. We set the L_2 coefficient of Ridge model to be 0.05.

3.3 Training Details

YieldNet network was trained in an end-to-end manner. We initialized the network parameters with Xavier initialization [44]. To minimize the loss function defined in Equation 1, we used Adam optimizer [45] with learning rate of 0.05% and a mini-batch size of 32. The network was trained 4,000 iterations to convergence. We do not use dropout [46] because batch normalization also has a regularization effect on the network.

3.4 Final Results

After having trained all models, we evaluated the performances of our proposed method along with other competing models to predict corn and soybean yields. To completely evaluate all models, we took three years 2016, 2017, and 2018 as test years and predicted corn and soybean yields 4 times a year during the growing season on 23rd day of July, August, September, and October. Tables 3 and 4 present the corn and soybean yield prediction results, respectively, and compare the performances of models with respect to the root-mean-square error (RMSE) evaluation metric.

Test Date		Models					
Year	Month	Ridge	Lasso	RF	DFNN	RT	YieldNet
2016	July	29.40	27.35	26.68	28.27	34.23	**20.82**
	August	24.50	24.72	23.55	25.05	28.75	**19.27**
	September	25.42	24.24	24.02	24.43	29.36	**18.59**
	October	26.96	23.60	24.05	25.32	29.82	**18.51**
2017	July	31.37	33.60	29.41	32.01	33.68	**22.57**
	August	22.35	24.52	23.70	24.77	29.79	**20.25**
	September	21.97	22.29	23.39	22.70	29.95	**19.06**
	October	22.86	22.37	23.06	23.24	29.10	**17.89**
2018	July	21.14	23.81	27.67	25.64	32.96	**20.98**
	August	20.49	22.31	24.76	22.82	29.62	**18.46**
	September	19.18	20.55	24.87	21.75	29.28	**18.27**
	October	23.81	21.18	25.75	23.24	29.27	**18.08**
Average		24.12	24.21	25.08	24.94	30.48	**19.39**

Table 3: The RMSE of corn yield prediction performance of models. The average $\pm$ standard deviation for corn yield in years 2016, 2017, and 2018 are, respectively, 165.72 ± 30.35, 168.50 ± 32.88, and 170.77 ± 34.95. The numbers of test samples in years 2016, 2017, 2018 for corn yield prediction are 885, 882, and 784, respectively. The unit of RMSE is bushels per acre.

Test Date		Models					
Year	Month	Ridge	Lasso	RF	DFNN	RT	YieldNet
2016	July	11.73	11.68	10.42	11.38	11.36	**6.69**
	August	8.82	10.05	9.45	8.97	10.37	**5.54**
	September	8.66	5.87	9.44	8.57	10.17	**5.41**
	October	10.74	9.63	9.42	9.39	10.26	**5.47**
2017	July	7.58	8.88	7.60	7.70	10.11	**6.42**
	August	5.56	7.16	6.08	5.68	8.18	**5.50**
	September	5.81	7.64	5.97	6.0	7.80	**5.11**
	October	6.12	5.89	5.93	6.0	7.68	**5.15**
2018	July	8.41	11.17	9.34	9.22	11.83	**6.81**
	August	6.63	9.20	8.15	7.35	10.17	**6.24**
	September	6.50	8.94	7.98	7.40	9.51	**5.54**
	October	6.86	9.44	8.26	7.87	9.87	**5.48**
Average		7.79	8.80	8.17	7.96	9.78	**5.78**

Table 4: The RMSE of soybean yield prediction performance of models. The average $\pm$ standard deviation for corn yield in years 2016, 2017, and 2018 are, respectively, 53.94 ± 7.23, 50.24 ± 8.72, and 53.17 ± 9.72. The numbers of test samples in years 2016, 2017, and 2018 for soybean yield prediction are 773, 776, and 663, respectively. The unit of RMSE is bushels per acre.

As shown in Tables 3 and 4, our proposed method outperforms other methods to varying extents. The Ridge and Lasso had comparable performances for both corn and soybean yield prediction, though, Ridge performed slightly better compared to the Lasso. DFNN showed a better performance than RF and RT for both corn and soybean yield predictions. Despite the linear modeling structure, Ridge regression performed better than RF, and DFNN, which indicates that DFNN and RF cannot successfully capture the nonlinearity of remote sensing data, resulting in a poor performance compared to Ridge and Lasso. RT had a weak performance compared to other methods due to being prone to overfitting. RF performs considerably better than RT because of using ensemble learning, which makes it robust against overfitting. Our proposed method significantly outperformed the other methods due to multiple factors: (1) the convolution operation in the YieldNet model captures both the temporal effect of remote sensing data collected over growing season and the spatial information of bins in histograms, (2) the YieldNet network uses transfer learning between corn and soybean yield predictions by sharing the weights of the backbone feature extractor, and (3) using a shared backbone feature extractor in the YieldNet model substantially decreases the number of model's parameters and subsequently helps training process despite the limited labeled data.

The prediction accuracy decreases as we try to make prediction earlier during the growing season (e.g. July and August) due to loss of information. All models except our proposed model do not show a clear decreasing pattern in accuracy as

we go from October to July, which indicates they cannot fully learn the functional mapping from satellite images to the yield.

We also report the yield prediction performance of our proposed model with respect to another evaluation metric, mean absolute error (MAE), in Table 5. As shown in Table 5, our proposed method accurately predicted corn yield one month, two months, three months, and four months before the harvest with MAE being 9.92%, 8.88%, 8.36%, and 7.8% of the average corn yield. The proposed model also accurately predicted soybean yield one month, two months, three months, and four months before harvest with MAE being 10.05%, 9.06%, 8.01%, and 7.67% of the average corn yield. The proposed model is slightly more accurate in soybean yield forecasting than corn yield forecasting, which is due to the higher variation in the corn yield compared to the soybean yield.

Crop	2016				2017				2018			
	Jul.	Aug.	Sep.	Oct.	Jul.	Aug.	Sep.	Oct.	Jul.	Aug.	Sep.	Oct.
Corn	17.5	15.46	14.03	14.16	17.25	15.78	13.76	12.72	16.05	14.23	15.06	13.06
Soybean	5.18	4.98	4.26	3.92	5.35	4.44	4.34	4.28	5.5	5.03	4.18	4.04

Table 5: The MAE of the corn and soybean yield prediction performances of our proposed model.

We visualized the error percentage maps for the corn and soybean yield predictions for the year 2018. As shown in Figures 5 and 6, the error percentage is below 5% for most counties, which indicates that our proposed model provides a robust and accurate yield prediction across US Corn Belt.

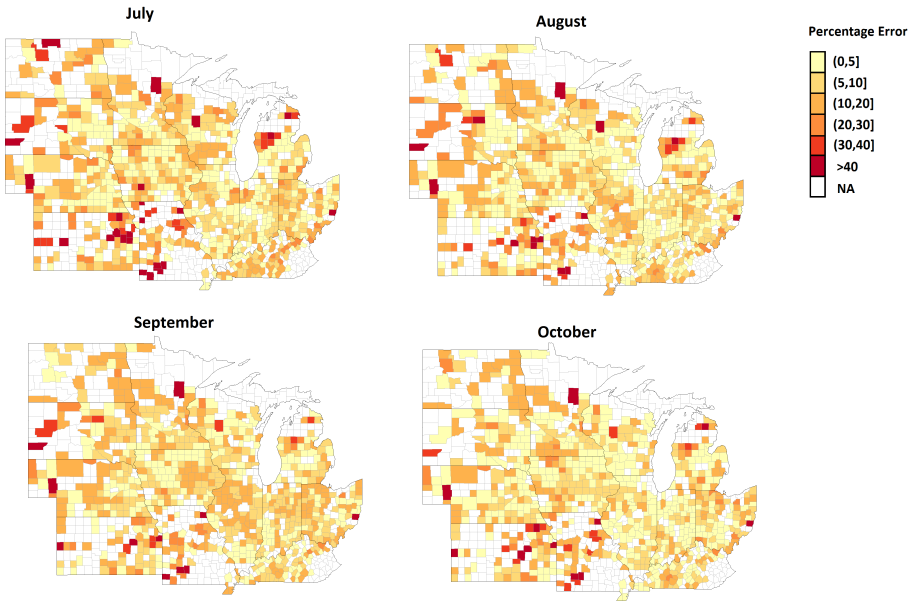

Figure 5: The error percentage maps for the 2018 corn yield prediction which is done during growing season in July, August, September, and October. The counties with white color indicate the ground truth yield were not available for those counties in 2018.

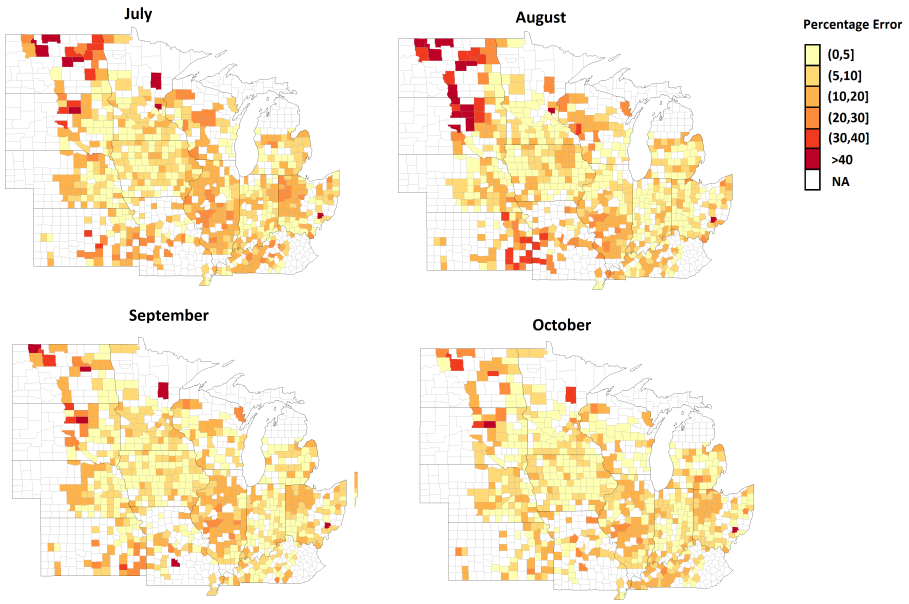

Figure 6: The error percentage maps for the 2018 soybean yield prediction which is done during growing season in July, August, September, and October. The counties with white color indicate the ground truth yield were not available for those counties in 2018.

To further evaluate the prediction results of our proposed model, we created the scatter plots of ground truth yield against predicted yield for the year 2018. Figure 7 depicts the scatter plots for the corn yield prediction during the growing season in months July, August, September, and October. The corn scatter plots indicate that the YieldNet model can successfully forecast yield months prior to the harvest.

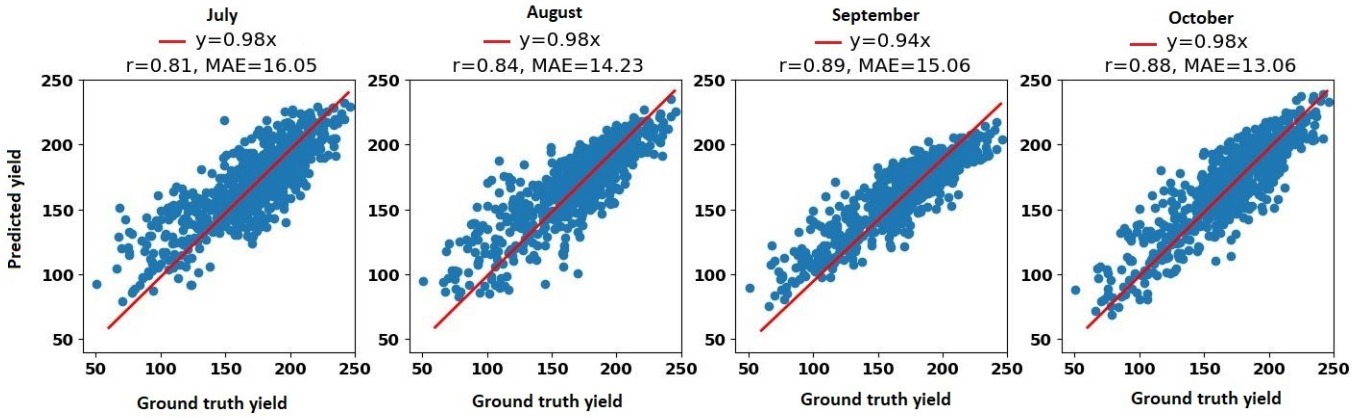

Figure 7: The scatter plots for the 2018 corn yield prediction during the growing season in months July, August, September, and October. MAE and r stand for mean absolute error and correlation coefficient, respectively. The unit of yield is bushels per acre.

Figure 8 depicts the scatter plots for the soybean yield prediction during the growing season in months July, August, September, and October. The corn scatter plots indicate that the YieldNet model provides reliable and accurate yield months prior to the harvest.

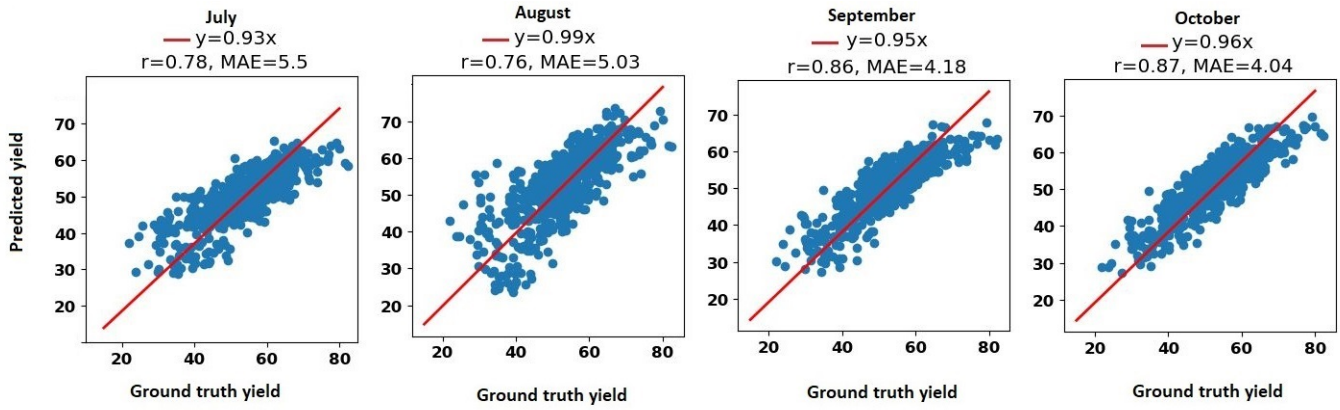

Figure 8: The scatter plots for the 2018 soybean yield prediction during the growing season in months July, August, September, and October. MAE and r stand for mean absolute error and correlation coefficient, respectively. The unit of yield is bushels per acre.

4 Ablation Study

In order to examine the usefulness of using a single deep learning model for simultaneously predicting the yield of two crops, we perform the following analysis. We train two separate models one for corn yield prediction and another for soybean yield prediction which are as follows:

YieldNet^{corn}: This model has exactly the same network architecture as the YieldNet model except we removed the the soybean head from the original YieldNet network. As a result, *YieldNet*^{corn} can only predict the corn yield.

YieldNet^{soy}: This model has exactly the same network architecture as the YieldNet model except we removed the the corn head from the original YieldNet network. As a result, *YieldNet*^{soy} can only predict the soybean yield.

Tables 6 and 7 compare the yield prediction performances of the above-mentioned models with the original YieldNet model.

Model	2016				2017				2018			
	Jul.	Aug.	Sep.	Oct.	Jul.	Aug.	Sep.	Oct.	Jul.	Aug.	Sep.	Oct.
YieldNet^{corn}	21.48	19.48	19.34	19.05	23.36	21.84	18.76	18.04	21.84	19.51	18.74	18.48
YieldNet	20.82	19.27	18.59	18.51	22.57	20.25	19.06	17.89	20.98	18.46	18.27	18.08

Table 6: The RMSE of the corn yield prediction performances of the YieldNet and *YieldNet*^{corn} models.

Model	2016				2017				2018			
	Jul.	Aug.	Sep.	Oct.	Jul.	Aug.	Sep.	Oct.	Jul.	Aug.	Sep.	Oct.
YieldNet^{soy}	7.81	7.03	6.22	5.94	6.84	6.17	5.26	5.23	7.95	7.4	5.92	5.86
YieldNet	6.69	5.54	5.41	5.47	6.42	5.50	5.11	5.15	6.81	6.24	5.54	5.48

Table 7: The RMSE of the soybean yield prediction performances of the YieldNet and *YieldNet*^{soy} models.

As shown in Tables 6 and 7, the YieldNet model which simultaneously predicts corn and soybean yields outperforms individual *YieldNet*^{corn} and *YieldNet*^{soy} models. The YieldNet provides more robust and accurate yield predictions compared to the other two individual models, which indicates that transfer learning between corn and soybean yield prediction improves the yield prediction accuracy for both crops.

5 Discussion and Conclusion

Our numerical results illustrate that our approach to simultaneously predicting yield for both corn and soybeans is possible and can achieve higher accuracy than individual models. By utilizing transfer learning between corn and soybean yield to share the weights of the backbone feature extractor, YieldNet was able to substantially decrease the number of learning parameters. Our transfer learning approach enabled us to save on computation resources while

also maximizing our prediction accuracy. Moreover, the accuracy achieve using a four month look-ahead has a lot of important implications for crop management decisions. With accurate yield predictions at various time points, decision makers now have the ability to change crop management practices to ensure yield is being maximized throughout its growth stage.

Although our approach highlighted corn and soybean in the U.S. market, this approach is applicable to any number of crops in any region. Due to the strength of our deep learning framework in combination with a generalized loss function, our approach is ready for scale. To improve the accuracy of our methodology, more data can be gathered and this can be left as a future extension to this work, along side more crops and more regions. It is the hope of this paper that our approach and results showcase the power deep learning for simultaneous yield prediction can have on the remote sensing community and the larger agricultural community as a whole.

Funding

This work was partially supported by the National Science Foundation under the LEAP HI and GOALI programs (grant number 1830478) and under the EAGER program (grant number 1842097).

References

[1] Evangelos Anastasiou, Athanasios Balafoutis, Nikoleta Darra, Vasileios Psiroukis, Aikaterini Biniari, George Xanthopoulos, and Spyros Fountas. Satellite and proximal sensing to estimate the yield and quality of table grapes. *Agriculture*, 8(7):94, 2018.

[2] Carlos Antonio da Silva, Marcos Rafael Nanni, Paulo Eduardo Teodoro, and Guilherme Fernando Capristo Silva. Vegetation indices for discrimination of soybean areas: A new approach. *Agronomy journal*, 109(4):1331–1343, 2017.

[3] NA Quarmby, Mi Milnes, TL Hindle, and N Silleos. The use of multi-temporal ndvi measurements from avhrr data for crop yield estimation and prediction. *International Journal of Remote Sensing*, 14(2):199–210, 1993.

[4] Felix Kogan, Anatoly Gitelson, Edige Zakarin, Lev Spivak, and Lubov Lebed. Avhrr-based spectral vegetation index for quantitative assessment of vegetation state and productivity. *Photogrammetric Engineering & Remote Sensing*, 69(8):899–906, 2003.

[5] Ramesh P Singh, Sudipa Roy, and F Kogan. Vegetation and temperature condition indices from noaa avhrr data for drought monitoring over india. *International journal of remote sensing*, 24(22):4393–4402, 2003.

[6] Yuei-An Liou and Sanjib Kumar Kar. Evapotranspiration estimation with remote sensing and various surface energy balance algorithms—a review. *Energies*, 7(5):2821–2849, 2014.

[7] Lisheng Song, Shaomin Liu, William P Kustas, Ji Zhou, and Yanfei Ma. Using the surface temperature-albedo space to separate regional soil and vegetation temperatures from aster data. *Remote Sensing*, 7(5):5828–5848, 2015.

[8] Jakob Geipel, Johanna Link, and Wilhelm Claupein. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. *Remote Sensing*, 6(11):10335–10355, 2014.

[9] Justin Van Wart, K Christian Kersebaum, Shaobing Peng, Maribeth Milner, and Kenneth G Cassman. Estimating crop yield potential at regional to national scales. *Field Crops Research*, 143:34–43, 2013.

[10] Marjorie Battude, Ahmad Al Bitar, David Morin, Jérôme Cros, Mireille Huc, Claire Marais Sicre, Valérie Le Dantec, and Valérie Demarez. Estimating maize biomass and yield over large areas using high spatial and temporal resolution sentinel-2 like remote sensing data. *Remote Sensing of Environment*, 184:668–681, 2016.

[11] Rémy Fieuzal, C Marais Sicre, and Frédéric Baup. Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks. *International journal of applied earth observation and geoinformation*, 57:14–23, 2017.

[12] V Sagan, M Maimaitijiang, P Sidike, M Maimaitiyiming, H Erkbol, S Hartling, KT Peterson, J Peterson, J Burken, and F Fritschi. Uav/satellite multiscale data fusion for crop monitoring and early stress detection. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2019.

[13] Louise Leroux, Mathieu Castets, Christian Baron, Maria-Jose Escorihuela, Agnès Bégué, and Danny Lo Seen. Maize yield estimation in west africa from crop process-induced combinations of multi-domain remote sensing indices. *European Journal of Agronomy*, 108:11–26, 2019.

[14] Rocío Ballesteros, Diego S Intrigliolo, José F Ortega, Juan M Ramírez-Cuesta, Ignacio Buesa, and Miguel A Moreno. Vineyard yield estimation by combining remote sensing, computer vision and artificial neural network techniques. *PRECISION AGRICULTURE*, 2020.

[15] Wen Zhuo, Jianxi Huang, Li Li, Xiaodong Zhang, Hongyuan Ma, Xinran Gao, Hai Huang, Baodong Xu, and Xiangming Xiao. Assimilating soil moisture retrieved from sentinel-1 and sentinel-2 data into wofost model to improve winter wheat yield estimation. *Remote Sensing*, 11(13):1618, 2019.

[16] Mohamad M Awad. Toward precision in crop yield estimation using remote sensing and optimization techniques. *Agriculture*, 9(3):54, 2019.

[17] Maitiniyazi Maimaitijiang, Vasit Sagan, Paheding Sidike, Sean Hartling, Flavio Esposito, and Felix B Fritschi. Soybean yield prediction from uav using multimodal data fusion and deep learning. *Remote Sensing of Environment*, 237:111599, 2020.

[18] Pablo Rischbeck, Salah Elsayed, Bodo Mistele, Gero Barmeier, Kurt Heil, and Urs Schmidhalter. Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley. *European Journal of Agronomy*, 78:44–59, 2016.

[19] K Kuwata and R Shibasaki. Estimating corn yield in the united states with modis evi and machine learning methods. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci*, 3(8):131–136, 2016.

[20] USDA. USDA long-term agricultural projections. `https://www.usda.gov/oce/commodity/projections/`, 2019.

[21] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[22] Saeed Khaki and Lizhi Wang. Crop yield prediction using deep neural networks. *Frontiers in plant science*, 10:621, 2019.

[23] Jie Sun, Liping Di, Ziheng Sun, Yonglin Shen, and Zulong Lai. County-level soybean yield prediction using deep cnn-lstm model. *Sensors*, 19(20):4363, 2019.

[24] Saeed Khaki, Hieu Pham, Ye Han, Andy Kuhl, Wade Kent, and Lizhi Wang. Deepcorn: A semi-supervised deep learning method for high-throughput image-based corn kernel counting and yield estimation. *arXiv preprint arXiv:2007.10521*, 2020.

[25] Jiaxuan You, Xiaocheng Li, Melvin Low, David Lobell, and Stefano Ermon. Deep gaussian process for crop yield prediction based on remote sensing data. In *Thirty-First AAAI conference on artificial intelligence*, 2017.

[26] Saeed Khaki, Lizhi Wang, and Sotirios V Archontoulis. A cnn-rnn framework for crop yield prediction. *Frontiers in Plant Science*, 10:1750, 2020.

[27] Saeed Khaki, Zahra Khalilzadeh, and Lizhi Wang. Predicting yield performance of parents in plant breeding: A neural collaborative filtering approach. *Plos one*, 15(5):e0233382, 2020.

[28] Mohsen Shahhosseini, Rafael A Martinez-Feria, Guiping Hu, and Sotirios V Archontoulis. Maize yield and nitrate loss prediction with machine learning algorithms. *Environmental Research Letters*, 14(12):124026, 2019.

[29] Michael P Pound, Jonathan A Atkinson, Alexandra J Townsend, Michael H Wilson, Marcus Griffiths, Aaron S Jackson, Adrian Bulat, Georgios Tzimiropoulos, Darren M Wells, Erik H Murchie, et al. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *Gigascience*, 6(10):gix083, 2017.

[30] Lawrence Mosley, Hieu Pham, Yogesh Bansal, and Eric Hare. Image-based sorghum head counting when you only look once. *arXiv preprint arXiv:2009.11929*, 2020.

[31] Saeed Khaki, Hieu Pham, Ye Han, Wade Kent, and Lizhi Wang. High-throughput image-based plant stand count estimation using convolutional neural networks. *arXiv preprint arXiv:2010.12552*, 2020.

[32] Saeed Khaki, Hieu Pham, Ye Han, Andy Kuhl, Wade Kent, and Lizhi Wang. Convolutional neural networks for image-based corn kernel detection and counting. *Sensors*, 20(9):2721, 2020.

[33] Saeed Khaki, Zahra Khalilzadeh, and Lizhi Wang. Classification of crop tolerance to heat and drought—a deep convolutional neural networks approach. *Agronomy*, 9(12):833, 2019.

[34] Jie Sun, Zulong Lai, Liping Di, Ziheng Sun, Jianbin Tao, and Yonglin Shen. Multilevel deep learning network for county-level corn yield estimation in the us corn belt. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:5048–5060, 2020.

[35] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[36] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[37] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, pages 265–283, 2016.

[38] E Vermote. MOD09A1 MODIS/terra surface reflectance 8-day l3 global 500m sin grid v006. *NASA EOSDIS Land Processes DAAC*, 2015.

[39] Z Wan, S Hook, and G Hulley. MOD11A2 MODIS/Terra land surface temperature/emissivity 8-day l3 global 1km sin grid v006. *NASA EOSDIS Land Processes DAAC*, 2015.

[40] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[41] Leo Breiman, Jerome Friedman, Charles J Stone, and Richard A Olshen. *Classification and regression trees*. CRC press, 1984.

[42] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[43] Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

[44] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.

[45] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

[46] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.