Computers and Electronics in Agriculture 170 (2020) 105197

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Modeling yield response to crop management using convolutional neural networks

Alexandre Barbosa[a], Rodrigo Trevisan[b], Naira Hovakimyan[a], Nicolas F. Martin[b,*]

[a] Mechanical Science and Engineering Department, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[b] Department of Crop Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

A B S T R A C T

Predicting crop yield response to management and environmental variables is a crucial step towards nutrient management optimization. With the increase in the amount of data generated by agricultural machinery, more sophisticated models are necessary to get full advantage of such data. In this work, we propose a Convolutional Neural Network (CNN) to capture relevant spatial structures of different attributes and combine them to model yield response to nutrient and seed rate management. Nine on-farm experiments on corn fields are used to construct a suitable dataset to train and test the CNN model. Four architectures combining input attributes at different stages in the network are evaluated and compared to the most commonly used predictive models. Results show a reduction in the test dataset RMSE up to 68% when compared to multiple linear regression and up to 29% when compared to a random forest. We also demonstrate that higher variability associated with the spatial structure of the data takes the most advantage of this framework.

1. Introduction

Improving crop agronomic management is an essential step towards sustainable cropping systems. Traditional farm management has led to excessive fertilization of crops, generating surplus nutrient flow which ends up polluting the water system (Carpenter et al., 1998). Despite ongoing research on modeling and data generation, the improvement of decision tools is yet to reach its full potential (Antle et al., 2017). So, new methods are needed to take full advantage of the new field technologies and create Decision-Support Systems (DSS) to help farmers increase production while accounting for environmental impacts (Abbott and Murphy, 2007).

In order to create decision tools, descriptive and predictive models of the process are first needed. A large number of models relating environmental and management variables to crop yield have been proposed throughout the last decades. They can be fairly separated into statistical (Kaul et al., 2005), and analytical crop models (Aggarwal, 1995). While analytical crop models are dynamic system simulations based on variables many times not measurable by farmers, empirical models are constrained to how well the dataset represents all possible conditions where the model is to be used and do not have good performance when extrapolated. On-Farm Precision Experimentation (OFPE) is often used to improve empirical models (Rodriguez et al., 2019) by generating site-specific data on yield responses to field management. However, even at a field scale, the spatial structure of environmental (e.g., soil composition and topography) and management variables (e.g., fertilizer application) may affect the yield through phenomena such as nutrient and water transportation (Yang et al., 2009).

Therefore, the spatial structure of environmental and treatment variables plays an important role when trying to create a predictive model for yield. Moreover, the interaction among different explanatory variables may depend on such spatial structures in a nonlinear way. Many spatial econometrics models were developed to account for data's spatial structure on an effort to better describe the relations between explanatory and response variables. In addition, Generalized Least Squares models (Plant, 2012) account for spatial structure using a geostatistical semivariogram while performing linear regressions. This approach, however, defines a fixed range to model the influence of neighboring data in a particular sample given the stationary assumption and considers that such influence is based only on the distance between the points rather than on spatial features of the data. Hence, a way to extract relevant spatial features from the data is needed to overcome this limitation.

Similar spatial feature extraction problem is also present on image recognition, where Convolutional Neural Networks (CNN) have demonstrated significantly higher performance over other methods, as well as in other complex tasks. Neural Networks are for long known to be universal continuous function approximators on a subset of $\mathbb{R}$ (Cybenko, 1989). Convolutional layers can be trained to encode relevant visual (and here we can also use the term "spatial") features of varying complexity when combined with fully connected layers. In

* Corresponding author at: University of Illinois at Urbana-Champaign, 1102 S.Goodwing Ave. Urbana, IL 61801, USA.
E-mail address: nfmartin@illinois.edu (N.F. Martin).

https://doi.org/10.1016/j.compag.2019.105197
Received 21 May 2019; Received in revised form 14 December 2019; Accepted 27 December 2019
Available online 12 February 2020

agriculture, applications of CNNs (Patricio and Rieder, 2018) usually focus on classification problems such as disease and plant identification, and image-based estimations, such as soybean leaf defoliation level. Nevavuori et al. (2019) used a CNN to estimate yield based on multispectral imagery collected during the growing season, not providing actionable pre-season information. To the best knowledge of the authors, the usage of CNNs to learn relevant spatial features from different explanatory variables collected during pre-season, and model the interactions among them as a regression problem has not been explored yet.

This work proposes a CNN to learn relevant spatial structures in different explanatory variables and use them to model the yield response to nutrient and seed rate prescriptions for future predictions. Four different CNN architectures combining the inputs at different stages in the network are compared. Data from nine corn fields across the US were used to test and compare the proposed architectures. Such fields are part of an OFPE with randomized nitrogen and seed rates prescription. The models are trained in a supervised fashion, tested, and compared to a linear model, a fully connected neural network, a random forest regression, and a support vector machine. Further sections in this paper are organized as follows: Section 2 details the construction of the dataset used for testing the CNN models; the CNN architectures are proposed and analyzed in Section 3, while Section 4 describes the experiments and results, and conclusions are made in Section 5.

2. Dataset construction

This paper uses data from the Data Intense Farm Management (DIFM) project (Bullock, 2016), recorded during the 2017 season. The database contains georeferenced management (generated through OFPE) and environmental data. Each field in the database has on average 40 ha and 200 experimental units, represented by 85 m long and 18 m wide polygons (Bullock et al., in press). Nine fields were selected, from which six are rain-fed fields in Illinois (Fields 2, 4, 7, 8, and 9) and Ohio (Field 5), and three are irrigated fields in Nebraska (Fields 3 and 6) and Kansas (Field 1). Nitrogen and seed rates were randomly assigned from four different levels to each experimental unit in a field, except for field 1, where the nitrogen rate is constant. The explanatory variables chosen for this study are nitrogen and seed rates prescription maps, elevation map and soil's shallow electroconductivity. Soil EC measurements are proxies of chemical and physical properties, including texture, bulk density, soil organic carbon, water content, salinity, and cation exchange capacity (Sudduth et al., 2013). In general, in fields as the ones we are evaluating, it is expected to observe a positive association between EC and soil's capacity to supply water and nutrients to crops. Information about EC is also valuable, given that it is more spatially detailed than that of soil samples (Martín et al., 2005). Additionally, a single cloud-free satellite image was used to characterize the bare soil variability of each field. Only the surface reflectance of the red band from a 3 m spatial resolution Planet Labs (Planet Team, 2017) PSSE4 multispectral image was used. The image was acquired after soil tillage approximately one week before planting when the field was clear form any vegetation and the soil was exposed. The bare soil reflectance has been shown to correlate with soil attributes such as organic matter content and particle size distribution (Khanal et al., 2018) and may work as a proxy to explain yield and nitrogen mineralization potential. Yield data is used as the response variable and was collected by yield monitors during harvesting. Notice that all explanatory variables are available at the beginning of the season, so one can use the model for future optimization of nitrogen and seed rates before applying them. As each field had its data recorded using different equipment and software, a way to put each variable from all fields into the same support is first needed.

The smallest unit of analysis is chosen to be a square with a side of five meters and is represented by a single element in a larger raster spanning the field. Variables stored as polygon data (prescription maps) were sampled to each element in the raster (using the mean method), while the ones stored as point data (elevation map and soil's electroconductivity) were first interpolated using kriging (Plant, 2012) and then rasterized to the desired support (5×5 m cell). The satellite picture was simply resampled to the desired support since it is already a raster.

The key assumption of this work is that the yield at a single unit of analysis depends on the spatial structure of observed explanatory variables around it. It is fair to assume however that the influence of neighbor data over a unit is limited to a certain range. So, by finding such range, models' complexity can be reduced without losing valuable data, which makes the model easier to train and more data-efficient. Therefore, a variogram containing the averaged variability from all explanatory variables is constructed to compute the range that better describes the data. For the selected fields the computed range is approximately 100 m, indicating that no significant higher variability is obtained when considering a greater distance between two points.

Then, a sample used as input for the proposed models described in the next section is defined as a set of 21×21 elements rasters (one for each of the five explanatory variables), spanning a square of 100×100 m around the unit where the yield value is being considered (Fig. 1). Then, to construct a dataset, rasters (with dimension equal to 21×21) are cropped from the original raster variables around each non zero cell in the yield map. Such procedure results in a set of samples containing five rasters each and the value of yield at their respective center cell.

When considering the yield data collected closer than 50 m from the field's boundaries we can see that the rasters centered at such cells cover an area where data is not available. Trying to sample such data would assign null values to these cells, resulting in a non-representative sample. Depending on the size of the field, a considerable number of data points in the yield map lies within 50 m from the border, and simply eliminating such points from the dataset would result in a high percentage loss of information. To solve this problem, a buffer is created around the rasters of each explanatory variable (Fig. 2). The values assigned to cells at the buffer depend on the variable they are representing. The seed and nitrogen maps are buffered with null cells since neither of them is being applied at the buffer. The elevation map and soils' shallow electroconductivity were estimated using kriging to estimate the values at the buffer, while the satellite picture is taken already considering the buffer. Finally, the data is standardized variable-wise to make the model's training process easier, and samples in which yield data is more than three standard deviations from the mean were removed from the dataset.

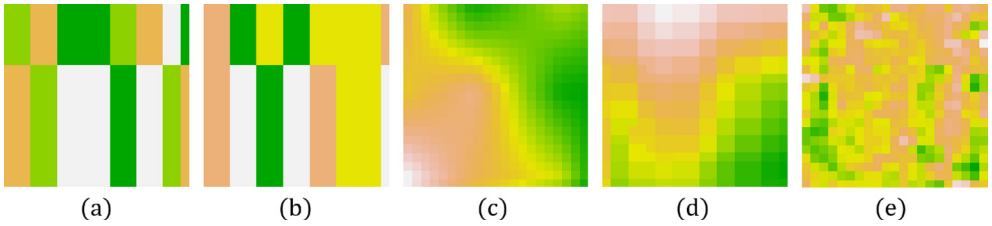

Fig. 1. Rasters representing five different input variables at the same field site: nitrogen rate (a), seed rate (b), elevation map (c), soil's electroconductivity (d), and satellite image (e).

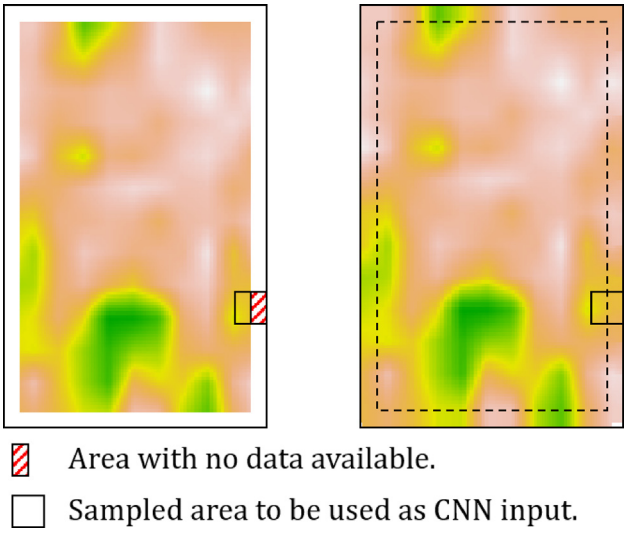

Fig. 2. Example of a buffer created to assign values to areas with no data on an EC map.

3. CNN model

In this work, five different field's attributes (nitrogen and seed rates, elevation map, soil's electroconductivity, and satellite image) were used to model yield response to nitrogen and seed rate management. Defining how these different sources of information are combined through the network is a crucial step towards obtaining efficient modeling. The problem of finding an appropriate architecture to combine the inputs has been addressed in different research areas. A CNN was proposed in Sun et al. (2017) for flower grading, where three images are necessary to fully describe a flower, and authors concatenate the result of independent convolutions from each image in the first layer before performing subsequent convolutions. Different architectures are also present in the literature related to human action identification in video data. A multi-stream architecture is proposed in (Simonyan and Zisserman, 2014), in which inputs are combined late in the network, demonstrating better performance than architectures using stacked frames as inputs (Karpathy et al., 2014). A 3D CNN was proposed by Tran et al. (2015), which uses 3D filters convolved with adjacent input frames, demonstrating results as good as the state of the art techniques.

Four CNN architectures combining the inputs in different ways are proposed and tested in this work. The first one combines the inputs as in the most commonly used convolutional neural networks, by stacking (we call it ST) them as a multi-channel image (Fig. 3). Sixteen 3×3 filters are used in the first convolutional layer with stride one, followed by a 2×2 max-pooling layer with stride two. Then outputs are flattened and fed to two sequential fully connected ReLU layers with 512 neurons each. Finally, outputs are connected to an output neuron with a linear activation function. The second proposed architecture is a multi-stream Early-Fusion (EF) network. In this network, each input is connected to an independent convolutional layer (configuring a multi-stream network) with eight 3×3 filters each with stride one, followed by a 2×2 max-pooling layer with stride two (Fig. 4). The outputs of the max-pooling layers are then flattened, concatenated, and fed to a fully connected ReLU layer with sixteen neurons, followed by an output neuron with a linear activation function. We name our third architecture as Late Fusion (LF). It is also a multi-stream network similar to the previous one. However, a fully connected ReLu layer with sixteen neurons is added to each stream after the max-pooling layer, followed by a single ReLu neuron (Fig. 5). Then, the five neurons (one from each stream) are concatenated and fed to the last two layers as in the EF architecture. The last proposed architecture is a 3D CNN (named 3D for short) very similar to ST, except the convolutional layer has sixteen $3 \times 3 \times 1$ filters convolved also in the channel dimension with stride one (Fig. 6).

To create a baseline, commonly used models for yield estimation are compared with the proposed CNN architectures. Such models are the multiple linear regression (MLR), a fully connected neural network (FC), a random forest (RF) regression, and a support vector machine (SVM). Fully connected neural networks have received significant attention over the last decade (Dahikar and Rode, 2014; Vani et al., 2015; Betiku and Taiwo, 2015) for being a powerful tool when modeling both linear and nonlinear relations between explanatory and response variables. A broad comparative study is presented by Marko et al. (2017), in which random forest demonstrates higher performance over machine learning competitors when predicting yield. Brdar et al. (2011) explores SVM as a comprehensible methodology for accurate yield prediction. Even so, all these models do not consider the spatial structure of the data, having as inputs the values of each variable at the exact cell where yield is estimated.

4. Experiments and results

Yield response to site-specific management also depends on other environmental factors and management practices with whole field impact (e.g., solar radiation, planting date, seed genetics, among others). These factors vary from one field to another while being constant within the same field. Then, given these whole field conditions, we use site-specific data aiming better recommendations for each farmer individually. However, data within a single field is often autocorrelated and may lead to model over-fitting that might not be generalized depending on the way the test set is chosen. For instance, when randomly partitioning the data, samples from training and test sets are virtually

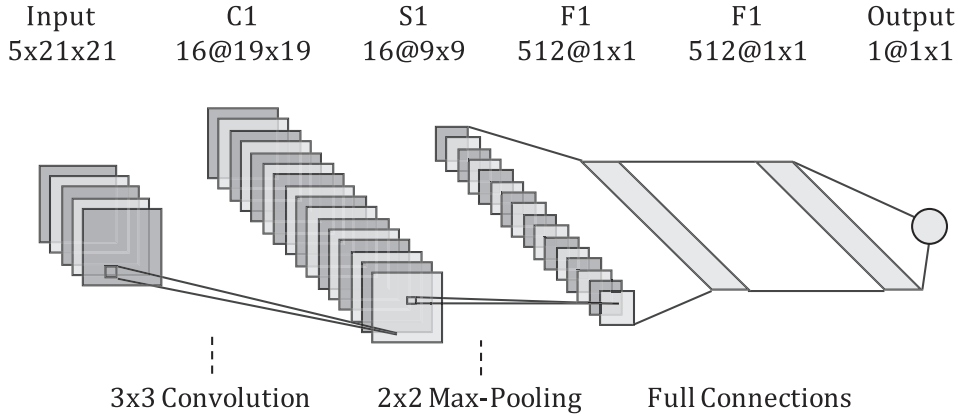

Fig. 3. CNN-ST architecture.

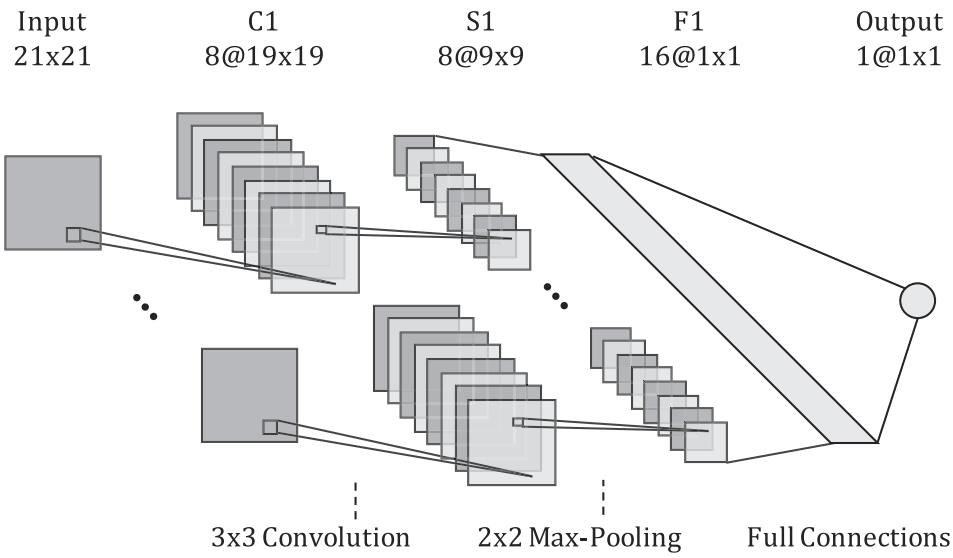

Fig. 4. CNN-EF architecture.

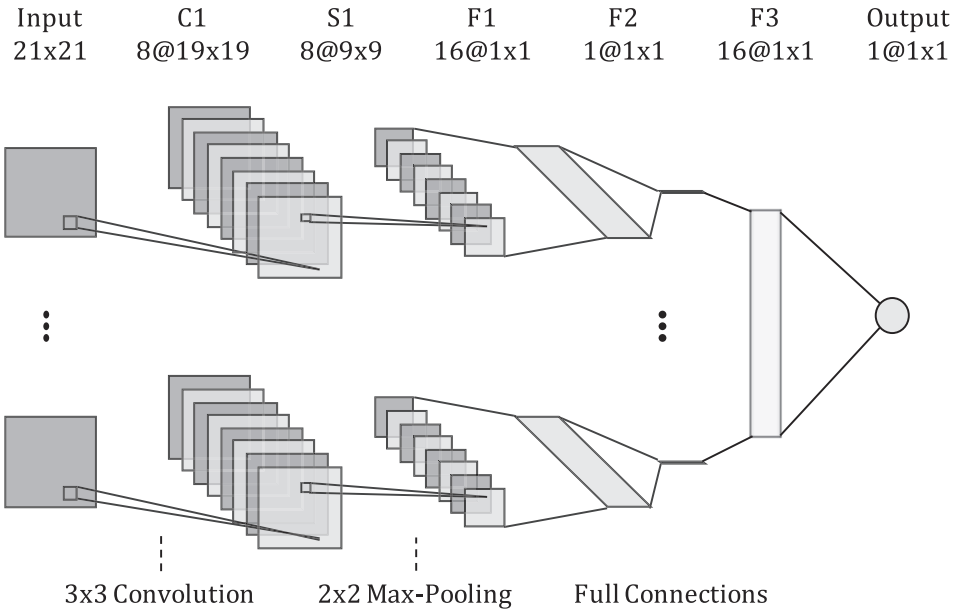

Fig. 5. CNN-LF architecture.

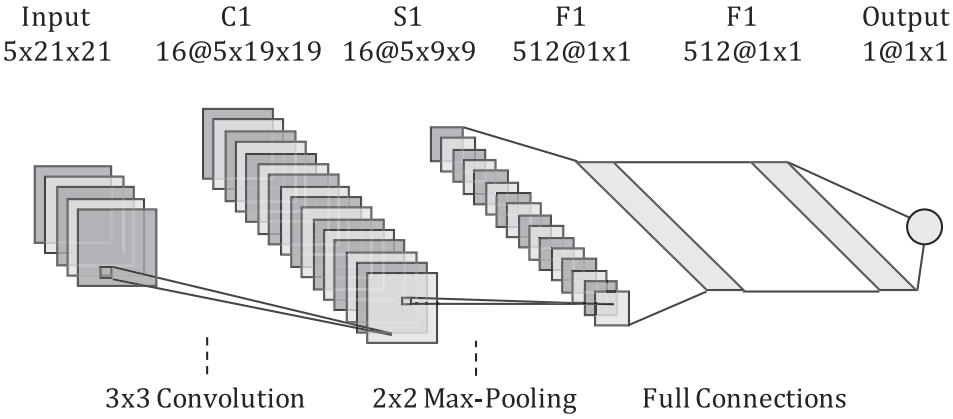

Fig. 6. CNN-3D architecture.

Table 1
Crossvalidation averaged RMSE over test dataset (in terms of yield standard deviation from each field), yield standard deviation, and yield mean [kg/ha].

Field	RMSE								Yield [kg/ha]	
	MLR	FC	LF	EF	3D	ST	SVM	RF	Stdv.	Mean
1	1.53	0.97	**0.66**	0.73	0.75	0.69	0.86	0.92	3240	12500
2	1.29	0.88	**0.83**	0.86	0.87	0.88	0.86	0.90	2290	10700
3	1.90	0.63	**0.58**	0.58	0.59	0.60	0.86	0.59	1230	14400
4	1.03	0.76	**0.75**	0.78	0.77	0.77	0.77	0.76	900	12200
5	0.74	0.72	**0.70**	0.72	0.75	0.73	0.76	0.70	1360	14500
6	1.09	0.51	**0.48**	0.51	0.53	0.56	0.58	0.52	1140	14700
7	0.75	0.72	**0.69**	0.70	0.71	0.73	0.76	0.73	1150	15700
8	1.11	0.94	0.94	0.94	0.94	0.94	**0.89**	0.96	2267	14100
9	1.10	0.69	**0.63**	0.65	0.66	0.66	0.67	0.63	1140	12600
Avg.	1.17	0.76	**0.70**	0.72	0.73	0.73	0.78	0.75	1635	13489

the same, making with the generalization capability of the model is not truly tested. So, we spatially partition the data to account for this problem as proposed by Vucetic et al. (1999).

Experiments with the dataset constructed in Section 2 were conducted in RStudio environment using the Keras and Tensorflow packages. The models proposed in the previous sections were trained and evaluated in each field independently to get full advantage of the site-specific data available in the dataset. Data from each field was spatially partitioned in five stripes perpendicularly to the longest dimension of the field to maximize the distance between samples. Three stripes (60% of the data) were used for training the model, one (20% of the data) for validation, and one (20% of the data) for tests. A grid search was performed to define hyperparameters such as the number of filters in the convolutional layers, the number of layers and neurons in the fully connected layers, and the dropout regularization probability for all models. Each architecture was trained over the same database using Adam optimizer (Kingma and Ba, 2014) to update the network weights, and the validation set was used online as an early stop criterion (allowing at maximum eight consecutive iterations of increase in its RMSE value) to avoid overfitting the data.

Cross-validation was used to evaluate the model using five folds according to the spatial partitioning. Table 1 shows the averaged RMSE over the five test sets for each model in each field. Notice that since yield data is standardized, the RMSE value represents a fraction of the standard deviation from the original yield data (shown in the most right column) in each field. The best results are formatted in bold.

The CNN-LF model is the one with the lowest RMSE value for eight of the nine tested fields, being field eight the only in which the SVM had a lower value. We show further in this section that this field is the one where the variability is the least explained by the spatial structure of the data. The random forest regression shows the second best results in four of the fields.

The better performance of CNN-LF over other CNNs may be explained by the way each architecture combines the input variables in the network. The CNN-ST architecture (usually applied to image classification problems) is limited to modeling element-wise linear combinations between different variables (at the first convolutional layer), even that the interaction between the result of this combination with the response variable can be modeled with no restrictions. To make this statement clear, suppose a n × n filter convolved with stacked input with L different variables (each one is a channel in the stacked frame). Now let $w_{i,j}$ be the filter's weights, $y_{u,v}$ the convolution output at position (u, v), and $x_{k,i,j}$ be the value at variable k aligned with position (i, j) in the filter. Then we have that y_{uv} is given by Eq. (1):

$$y_{uv} = \sum_{k=1}^{L} \sum_{i=1}^{n} \sum_{j=1}^{n} (w_{ij} x_{kij}) \tag{1}$$

Fig. 7 shows an example where two stacked inputs are convolved with a single 2D filter with stride equal to one. It is evident in this example that the output of the convolutional layer only contains a linear combination of the values from each input, and that the information from each separate input is not preserved. As a consequence, any nonlinear interaction between the input variables will not be modeled at the fully connected layers (that work as function approximators) (see Fig. 8).

One possible solution to overcome this limitation is to convolve the filter input-wise so the information from each input is carried separately through the convolutional layer. In fact, the proposed CNN-3D architecture considers the stacked different inputs as a third dimension to convolve the filter. However, even that this architecture solves the problem of carrying the information from each input to be combined in the fully connected layers, it introduces a new undesired feature. By using this architecture, the output of the convolutional layer increases with the number of inputs since the filters are convolved with all inputs independently. This may lead to an unnecessary increase in the number of parameters in the network if we assume that the same filter may not be relevant (as feature extractors) for all input variables, resulting in a model with a lower data efficiency.

The multi-stream architecture addresses both problems discussed above. It performs independent convolutions with the inputs while using a different set of filters for each one. When comparing the LF and EF models it is reasonable to say that LF focuses on better feature extraction from inputs while EF can model more complex interaction between variables. In addition, by reducing the dimension of each input before combining them, the LF model becomes easier to train, leading to higher data efficiency.

When looking to Table 1, one can notice that the CNN models have a modest advantage over their competitors in some fields, while in others this advantage is much higher. To investigate the possible cause of this difference, the variability of the response variables is observed. For that, spherical models were fitted to the variogram. Let's define c_0 as the nugget variance and $(c_0 + c_1)$ as the sill variance. Then, c_0 represents the variance observed at a lag distance of 0 m, and is a function of stochastic effects and measurement error (Taylor et al., 2007), while c_1 accounts for the additional variance that comes from the spatial structure of the data for a given range. Since the CNN framework focuses on modeling unknown spatial structures of the data, we propose to estimate how much of the total variability observed in each field comes from such structure. This resembles the Cambardella metric (Cambardella et al., 1994) used to calculate spatial dependence, except here we use a ratio of c_1 to total variance, given by Eq. (2):

$$C_{\text{index}} = 1 - \frac{c_0}{c_0 + c_1} \tag{2}$$

Table 2 shows the C_{index} calculated for the yield data of all nine fields and the RMSE reduction of CNN-LF architecture when compared to MLR, FC, SVM and RF. It is clear and also expected that in fields with high C_{index} (i.e. yield variability highly dependent on the spatial structure of the data) the proposed model presents a higher reduction on the RMSE value when compared to FC and MLR. On the other hand, in fields with low C_{index} a simpler model as FC or even MLR also has good performance, getting closer to CNN model, what results on a lower improvement when comparing them.

Then, yield data or even an explanatory variable (which is assumed to have some impact on yield) can be used to assess the expected improvement on yield prediction before using the proposed framework, which is, in fact, more computationally expensive than commonly used methods.

Fig. 9 shows a qualitative comparison between predicted yield maps using the CNN-LF and FC model in field four. It is possible to observe that the spatial structure of the data is better captured by the CNN-LF model, while FC results in a "blurrier" map. The difference between the residual maps from both models shows a spatial structure that resembles the spatial distribution observed on the true yield map,

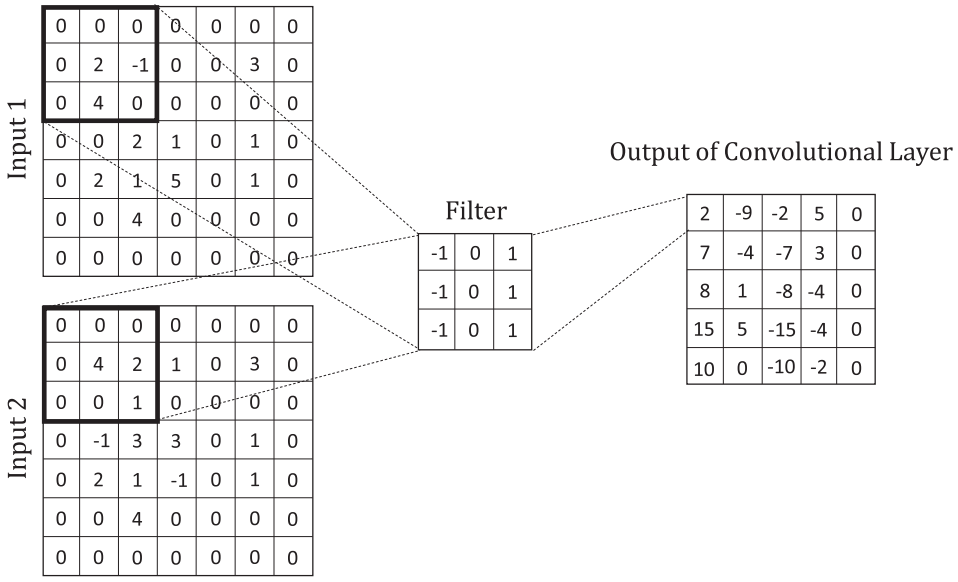

Fig. 7. Two stacked inputs convolved with a 2D filter.

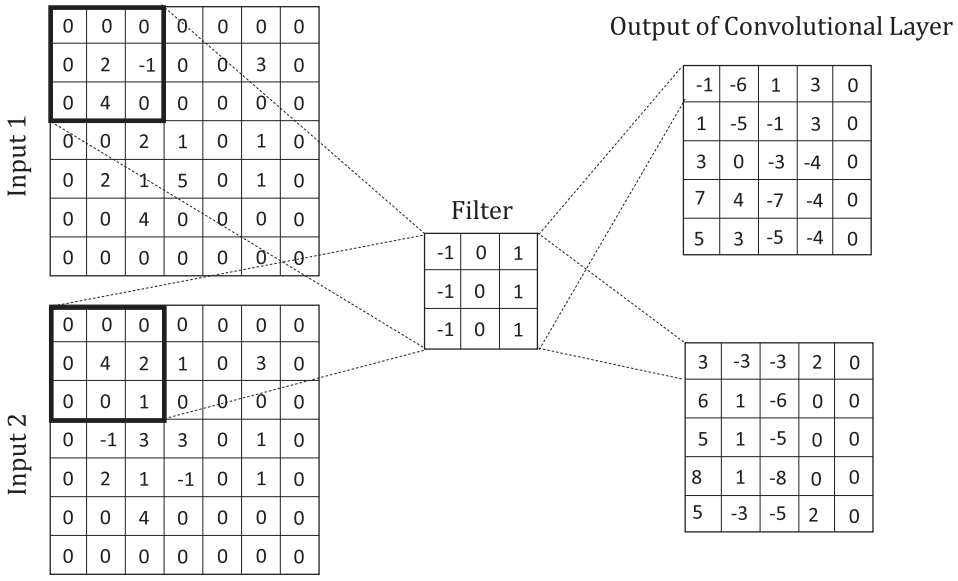

Fig. 8. Two stacked inputs convolved with a 3D filter.

Table 2
Percent decrease in RMSE value using CNN-LF when compared to FC and MLR for each field, and their respective C index for the yield map.

Field	Decrease in RMSE [%]				C_{index}
	MLR	FC	SVM	RF	
1	57	32	24	29	0.47
2	36	5	3	8	0.36
3	70	8	33	3	0.18
4	27	1	3	1	0.10
5	5	3	8	1	0.12
6	56	6	17	8	0.50
7	7	3	8	4	0.29
8	16	0	−5	2	0.04
9	43	8	6	0	0.33
Avg.	35	7	11	6	0.27

indicating that the main difference between both predictions lies indeed on the spatial structure of the data. Moreover, one can observe that most of the higher values in the residuals maps are located closer to field's borders. The data in such regions are expected to have more noise due to yield monitoring system errors caused by the change in harvester speed and direction, and cannot be explained by the selected explanatory variables.

Since this framework aims to model the yield response to crop management, we propose a sensitivity analysis of the network for each of the inputs. We want to ensure the manageable variables (nitrogen and seed rates) are really considered in the model after the training process and not vanished by possibly higher importance of environmental variables. A sensitivity index is then obtained by Eq. (3):

$$\zeta_i = \frac{1}{S}\sqrt{\sum_{k=1}^{S}\left(\frac{\partial y}{\partial i_k}(I_k)\right)^2} \tag{3}$$

where S is the number of elements in the input raster, i is the input label, y represent the predicted yield, and I_k is the input vector of a

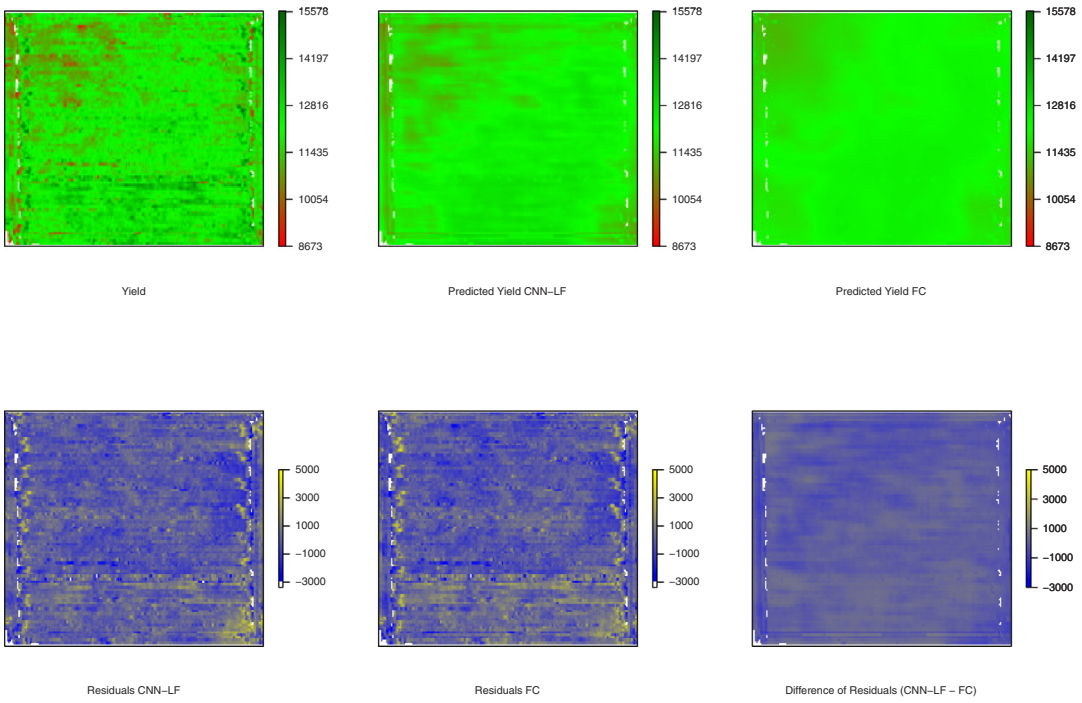

Fig. 9. True yield, predicted yield using CNN-LF, predicted yield using FC, residuals for both models, and difference of residuals for field 4 [kg/ha].

sample k. This equation is based on the sensitivity analysis through partial derivatives method (Montavon et al., 2018). The idea is to evaluate how the predicted yield changes with small perturbations of each input assuming the other input variables as constants. The value of the partial derivative depends on the point of the input space it is being evaluated and can be misleading interpreted if such a point does not represent the entire input space. To address this problem, as our CNN model is constructed from a dataset we believe represents the input space, we calculate the partials for every sample in our dataset. Finally, we square the values (we are interested in the magnitude) to add them, take the square root, and divide it by the number of samples to obtain the desired index.

Notice that since the inputs are standardized we can directly compare them to check for relative sensitivity among variables. Table 3 shows the index ζ for each input for the nine studied fields, revealing that the model is, in fact, sensitive to the selected manageable variables. Field 1 is an exception, showing an index of zero for nitrogen rate, which is expected since a constant rate was used for this field.

5. Conclusions

A CNN model for yield prediction based on pre-season treatments and environmental variables was proposed. Four different architectures based on convolutional neural networks were tested on nine corn fields and compared to a multiple linear regression model, a fully connected neural network (NN), a support vector machine (SVM), and a random forest (RF) regression. Tests were conducted using data from the same field in which each model was trained, and after computing the RMSE values for the test sets the CNN-LF architecture demonstrated the best performance among all. This model has the advantage of allowing nonlinear combinations among input variables while keeping the model with a relatively low number of parameters. The multi-stream approach used in this model focuses on variable-wise feature extraction. Moreover, the proposed late fusion architecture (CNN-LF) keeps each stream separated for more layers, resulting in better spatial feature extraction from the data. Thus, it is expected that filters trained with data from a field can learn features also relevant to different ones, allowing the application of transfer learning. In this way, this model can be trained in one field and used on a different one where OFPE were not conducted before (with fewer data available), remaining to only retrain the fully connected with a much smaller number of samples.

Fields with yield variability highly associated with the spatial structure of the data get the most benefit from this framework, and this can be estimated beforehand using information from a variogram. The demonstrated generalization power and sensitivity to manageable variables make this model suitable for nitrogen and seed rates optimization for the following season in order to maximize expected yield or even profit.

The proposed CNN framework brings the power of neural networks to geospatial problems. It allows creating models with no need of handcrafting features or making spatial distribution assumptions, which are incorporated into the learning process.

Table 3
Sensitivity index ζ for each input.

Field	NR	SR	Elev.	EC.	Soil
1	0.00	2.86	0.84	0.14	0.45
2	1.98	1.54	1.22	0.65	0.71
3	1.22	5.23	0.11	0.25	0.56
4	0.86	1.11	1.48	0.74	0.31
5	0.76	0.22	0.83	0.47	0.38
6	2.81	3.51	0.17	0.29	0.29
7	1.45	0.92	0.38	0.22	0.46
8	0.63	0.55	2.32	0.02	0.17
9	1.16	0.50	0.04	0.32	0.06
Avg.	1.21	1.83	0.82	0.34	0.38

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.compag.2019.105197.

References

Abbott, L., Murphy, D., 2007. What is Soil Biological Fertility? pp. 1–15. doi:https://doi.org/10.1007/978-1-4020-6619-1_1.

Aggarwal, P., 1995. Uncertainties in crop, soil and weather inputs used in growth models: Implications for simulated outputs and their applications. Agric. Syst. 48, 361–384. https://doi.org/10.1016/0308-521X(94)00018-M.

Antle, J.M., Jones, J.W., Rosenzweig, C.E., 2017. Next generation agricultural system data, models and knowledge products: introduction. Agric. Syst. 155, 186–190. https://doi.org/10.1016/j.agsy.2016.09.003.

Betiku, E., Taiwo, A.E., 2015. Modeling and optimization of bioethanol production from breadfruit starch hydrolyzate vis-à-vis response surface methodology and artificial neural network. Renewable Energy 74, 87–94.

Brdar, S., Culibrk, D., Marinkovic, B., Crnobarac, J., Crnojevic, V., 2011. Support vector machines with features contribution analysis for agricultural yield prediction. In: Second International Workshop on Sensing Technologies in Agriculture, Forestry and Environment (EcoSense 2011). Belgrade, Serbia, pp. 43–47.

Bullock, D., 2016. Using precision technology in on-farm field trials to enable data-intensive fertilizer management. USDA-NIFA Food Security Grant 2016-68004-24769.

Bullock, D.S., Boerngen, M., Tao, H., Maxwell, B., Luck, J.D., Shiratsuchi, L., Puntel, L., Martin, N.F., 2019. The data-intensive farm management project: changing agronomic research through on-farm precision experimentation. Agron. J. 111, 2736–2746. https://doi.org/10.2134/agronj2019.03.0165.

Cambardella, C., Moorman, T., Novak, J., Parkin, T., Karlen, D., Turco, R., Konopka, A., 1994. Field-scale variability of soil properties in central iowa soils. Soil Sci. Soc. Am. J. 58. https://doi.org/10.2136/sssaj1994.03615995005800050033x.

Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol. Appl. 8, 559–568. https://doi.org/10.1890/1051-0761(1998) 008[0559:NPOSWW]2.0.CO;2.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. MCSS 2, 303–314.

Dahikar, S., Rode, V., 2014. Agricultural crop yield prediction using artificial neural network approach. Int. J. Innov. Res. Electric. Electron. Instrum. Control Eng. 2, 683–686.

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L., 2014. Large-scale video classification with convolutional neural networks. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1725–1732. doi:https://doi.org/10.1109/CVPR.2014.223.

Kaul, M., Hill, R.L., Walthall, C., 2005. Artificial neural network for corn and soybean yield prediction. Agric. Syst. 85, 1–18. https://doi.org/10.1016/j.agsy.2004.07.009.

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., Shearer, S., 2018. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. Comput. Electron. Agric. 153, 213–225.

Kingma, D., Ba, J., 2014. Adam: A method for stochastic optimization. Int. Conf. Learn. Represent.

Marko, O., Brdar, S., Panić, M., Šašić, I., Despotović, D., Knežević, M., Crnojević, V., 2017. Portfolio optimization for seed selection in diverse weather scenarios. PloS One 12, e0184198.

Martín, N.F., Bollero, G., Bullock, D.G., 2005. Associations between field characteristics and soybean plant performance using canonical correlation analysis. Plant Soil 273, 39–55.

Montavon, G., Samek, W., Müller, K.R., 2018. Methods for interpreting and understanding deep neural networks. Digit. Signal Process. 73, 1–15. https://doi.org/10.1016/j.dsp.2017.10.011.

Nevavuori, P., Narra, N., Lipping, T., 2019. Crop yield prediction with deep convolutional neural networks. Comput. Electron. Agric. 163, 104859.

Patricio, D.I., Rieder, R., 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. Comput. Electron. Agric. 153, 69–81. http://www.sciencedirect.com/science/article/pii/S0168169918305829, doi:https://doi.org/10.1016/j.compag.2018.08.001.

Planet Team, San Francisco, C., 2017. Planet application program interface: In space for life on earth. https://api.planet.com.

Plant, R., 2012. Spatial Data Analysis in Ecology and Agriculture Using R. CRC Press, Boca Raton.

Rodriguez, D.G., S. Bullock, D., Boerngen, M., 2019. The origins, implications, and consequences of yield-based nitrogen fertilizer management. Agronomy J. 111. https://doi.org/10.2134/agronj2018.07.0479.

Simonyan, K., Zisserman, A., 2014. Two-stream convolutional networks for action recognition in videos. CoRR abs/1406.2199. arXiv:1406.2199.

Sudduth, K.A., Myers, D.B., Kitchen, N.R., Drummond, S.T., 2013. Modeling soil electrical conductivity–depth relationships with data from proximal and penetrating eca sensors. Geoderma 199, 12–21.

Sun, Y., Zhu, L., Wang, G., Zhao, F., 2017. Multi-input convolutional neural network for flower grading. J. Electric. Comput. Eng. 2017, 1–8. https://doi.org/10.1155/2017/9240407.

Taylor, J.A., Praat, J.P., Bollen, A.F., 2007. Spatial variability of kiwifruit quality in orchards and its implications for sampling and mapping. HortScience horts 42, 246–250.

Tran, D., Bourdev, L., Fergus, R., Torresani, L., Paluri, M., 2015. Learning spatiotemporal features with 3D convolutional networks. In: 2015 IEEE International Conference on Computer Vision, pp. 4489–4497.

Vani, S., Sukumaran, R., Savithri, S., 2015. Prediction of sugar yields during hydrolysis of lignocellulosic biomass using artificial neural network modeling. Bioresource Technol. 188. https://doi.org/10.1016/j.biortech.2015.01.083.

Vucetic, S., Fiez, T., Obradovic, Z., 1999. A data partitioning scheme for spatial regression. In: IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339), IEEE. pp. 2474–2479.

Yang, Y., Zhu, J., Tong, X., Wang, D., 2009. The spatial pattern characteristics of soil nutrients at the field scale. In: In: Li, D., Zhao, C. (Eds.), Computer and Computing Technologies in Agriculture II, vol. 1. Springer, US, Boston, MA, pp. 125–134.