Computers and Electronics in Agriculture 147 (2018) 70–90

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Review

Deep learning in agriculture: A survey

Andreas Kamilaris*, Francesc X. Prenafeta-Boldú

Institute for Food and Agricultural Research and Technology (IRTA), Spain

ARTICLE INFO

Keywords:
Deep learning
Agriculture
Survey
Convolutional Neural Networks
Recurrent Neural Networks
Smart farming
Food systems

ABSTRACT

Deep learning constitutes a recent, modern technique for image processing and data analysis, with promising results and large potential. As deep learning has been successfully applied in various domains, it has recently entered also the domain of agriculture. In this paper, we perform a survey of 40 research efforts that employ deep learning techniques, applied to various agricultural and food production challenges. We examine the particular agricultural problems under study, the specific models and frameworks employed, the sources, nature and pre-processing of data used, and the overall performance achieved according to the metrics used at each work under study. Moreover, we study comparisons of deep learning with other existing popular techniques, in respect to differences in classification or regression performance. Our findings indicate that deep learning provides high accuracy, outperforming existing commonly used image processing techniques.

1. Introduction

Smart farming (Tyagi, 2016) is important for tackling the challenges of agricultural production in terms of productivity, environmental impact, food security and sustainability (Gebbers and Adamchuk, 2010). As the global population has been continuously increasing (Kitzes et al., 2008), a large increase on food production must be achieved (FAO, 2009), maintaining at the same time availability and high nutritional quality across the globe, protecting the natural ecosystems by using sustainable farming procedures.

To address these challenges, the complex, multivariate and unpredictable agricultural ecosystems need to be better understood by monitoring, measuring and analyzing continuously various physical aspects and phenomena. This implies analysis of big agricultural data (Kamilaris et al., 2017b), and the use of new information and communication technologies (ICT) (Kamilaris et al., 2016), both for short-scale crop/farm management as well as for larger-scale ecosystems' observation, enhancing the existing tasks of management and decision/policy making by context, situation and location awareness. Larger-scale observation is facilitated by remote sensing (Bastiaanssen et al., 2000), performed by means of satellites, airplanes and unmanned aerial vehicles (UAV) (i.e. drones), providing wide-view snapshots of the agricultural environments. It has several advantages when applied to agriculture, being a well-known, non-destructive method to collect information about earth features while data may be obtained systematically over large geographical areas.

A large subset of the volume of data collected through remote sensing involve images. Images constitute, in many cases, a complete picture of the agricultural environments and could address a variety of challenges (Liaghat and Balasundram, 2010; Ozdogan et al., 2010). Hence, imaging analysis is an important research area in the agricultural domain and intelligent data analysis techniques are being used for image identification/classification, anomaly detection etc., in various agricultural applications (Teke et al., 2013; Saxena and Armstrong, 2014; Singh et al., 2016). The most popular techniques and applications are presented in Appendix A, together with the sensing methods employed to acquire the images. From existing sensing methods, the most common one is satellite-based, using multi-spectral and hyperspectral imaging. Synthetic aperture radar (SAR), thermal and near infrared (NIR) cameras are being used in a lesser but increasing extent (Ishimwe et al., 2014), while optical and X-ray imaging are being applied in fruit and packaged food grading. The most popular techniques used for analyzing images include machine learning (ML) (K-means, support vector machines (SVM), artificial neural networks (ANN) amongst others), linear polarizations, wavelet-based filtering, vegetation indices (NDVI) and regression analysis (Saxena and Armstrong, 2014; Singh et al., 2016).

Besides the aforementioned techniques, a new one which is recently gaining momentum is deep learning (DL) (LeCun et al., 2015; LeCun and Bengio, 1995). DL belongs to the machine learning computational field and is similar to ANN. However, DL is about "deeper" neural networks that provide a hierarchical representation of the data by means of various convolutions. This allows larger learning capabilities and thus higher performance and precision. A brief description of DL is

* Corresponding author.
E-mail address: andreas.kamilaris@irta.cat (A. Kamilaris).

https://doi.org/10.1016/j.compag.2018.02.016
Received 14 July 2017; Received in revised form 7 February 2018; Accepted 10 February 2018
Available online 22 February 2018
0168-1699/

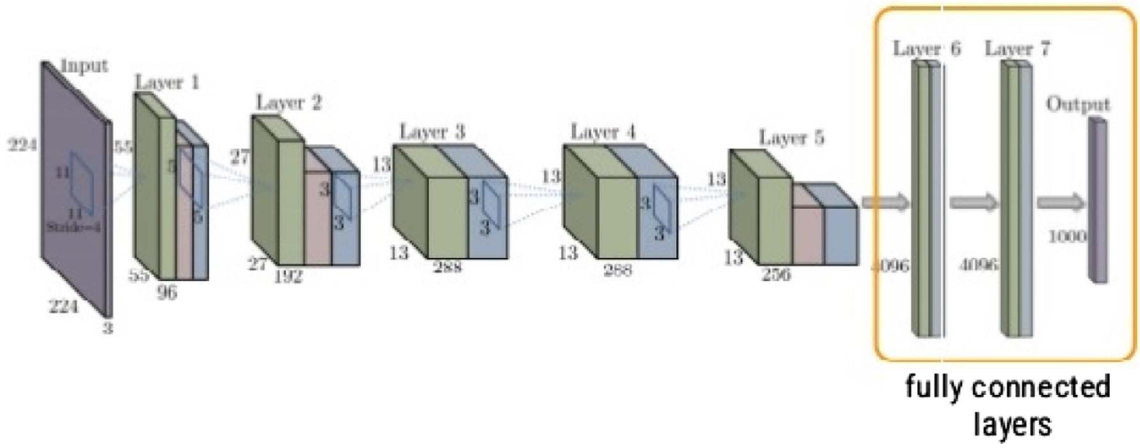

Fig. 1. CaffeNet, an example CNN architecture.
Source: Sladojevic et al. (2016).

attempted in Section 3.

The motivation for preparing this survey stems from the fact that DL in agriculture is a recent, modern and promising technique with growing popularity, while advancements and applications of DL in other domains indicate its large potential. The fact that today there exists at least 40 research efforts employing DL to address various agricultural problems with very good results, encouraged the authors to prepare this survey. To the authors' knowledge, this is the first such survey in the agricultural domain, while a small number of more general surveys do exist (Deng and Yu, 2014; Wan et al., 2014; Najafabadi et al., 2015), covering related work in DL in other domains.

2. Methodology

The bibliographic analysis in the domain under study involved two steps: (a) collection of related work and (b) detailed review and analysis of this work. In the first step, a keyword-based search for conference papers or journal articles was performed from the scientific databases IEEE Xplore and ScienceDirect, and from the web scientific indexing services Web of Science and Google Scholar. As search keywords, we used the following query:

["deep learning"] AND ["agriculture" OR "farming"]

In this way, we filtered out papers referring to DL but not applied to the agricultural domain. From this effort, 47 papers had been initially identified. Restricting the search for papers with appropriate application of the DL technique and meaningful findings[1], the initial number of papers reduced to 40.

In the second step, the 40 papers selected from the previous step were analyzed one-by-one, considering the following research questions:

1. Which was the agricultural- or food-related problem they addressed?
2. Which was the general approach and type of DL-based models employed?
3. Which sources and types of data had been used?
4. Which were the classes and labels as modeled by the authors? Were there any variations among them, observed by the authors?
5. Any pre-processing of the data or data augmentation techniques used?
6. Which has been the overall performance depending on the metric adopted?
7. Did the authors test the performance of their models on different datasets?
8. Did the authors compare their approach with other techniques and, if yes, which was the difference in performance?

Our main findings are presented in Section 4 and the detailed information per paper is listed in Appendix B.

3. Deep learning

DL extends classical ML by adding more "depth" (complexity) into the model as well as transforming the data using various functions that allow data representation in a hierarchical way, through several levels of abstraction (Schmidhuber, 2015; LeCun and Bengio, 1995). A strong advantage of DL is feature learning, i.e. the automatic feature extraction from raw data, with features from higher levels of the hierarchy being formed by the composition of lower level features (LeCun et al., 2015). DL can solve more complex problems particularly well and fast, because of more complex models used, which allow massive parallelization (Pan and Yang, 2010). These complex models employed in DL can increase classification accuracy or reduce error in regression problems, provided there are adequately large datasets available describing the problem. DL consists of various different components (e.g. convolutions, pooling layers, fully connected layers, gates, memory cells, activation functions, encode/decode schemes etc.), depending on the network architecture used (i.e. Unsupervised Pre-trained Networks, Convolutional Neural Networks, Recurrent Neural Networks, Recursive Neural Networks).

The highly hierarchical structure and large learning capacity of DL models allow them to perform classification and predictions particularly well, being flexible and adaptable for a wide variety of highly complex (from a data analysis perspective) challenges (Pan and Yang, 2010). Although DL has met popularity in numerous applications dealing with raster-based data (e.g. video, images), it can be applied to any form of data, such as audio, speech, and natural language, or more generally to continuous or point data such as weather data (Sehgal et al., 2017), soil chemistry (Song et al., 2016) and population data (Demmers et al., 2012). An example DL architecture is displayed in Fig. 1, illustrating CaffeNet (Jia et al., 2014), an example of a convolutional neural network, combining convolutional and fully connected (dense) layers.

Convolutional Neural Networks (CNN) constitute a class of deep, feed-forward ANN, and they appear in numerous of the surveyed papers as the technique used (17 papers, 42%). As the figure shows, various convolutions are performed at some layers of the network, creating different representations of the learning dataset, starting from more

[1] A small number of papers claimed of using DL in some agricultural-related application, but they did not show any results nor provided performance metrics that could indicate the success of the technique used.