biosystems engineering 172 (2018) 84–91

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Reassessment

Factors influencing the use of deep learning for plant disease recognition

Jayme G.A. Barbedo

Embrapa Agricultural Informatics, Av. André Tosello, 209, C.P. 6041, Campinas, SP 13083-886, Brazil

ARTICLE INFO

Article history:
Received 22 March 2018
Received in revised form
25 April 2018
Accepted 30 May 2018

Keywords:
Image processing
Deep neural nets
Transfer learning
Image database
Disease classification

Deep learning is quickly becoming one of the most important tools for image classification. This technology is now beginning to be applied to the tasks of plant disease classification and recognition. The positive results that are being obtained using this approach hide some issues that are seldom taken into account in the respective experiments. This article presents an investigation into the main factors that affect the design and effectiveness of deep neural nets applied to plant pathology. An in-depth analysis of the subject, in which advantages and shortcomings are highlighted, should lead to more realistic conclusions on the subject. The arguments used throughout the text are built upon both studies found in the literature and experiments carried out using an image database carefully built to reflect and reproduce many of the conditions expected to be found in practice. This database, which contains almost 50,000 images, is being made freely available for academic purposes.

1. Introduction

Since the dawn of agriculture, plant diseases cause substantial economic, social and environmental losses. Prophylactic actions are not always enough to prevent apidemics, thus careful monitoring is essential for early detection and consequent application of control measures. Traditionally, crop inspection has been carried out visually by people with some training or experience detecting plant disorders. As for any activity carried out by humans, this approach is subject to psychological and cognitive phenomena that may lead to bias, optical illusions and, ultimately, to error (Bock, Poole, Parker, & Gottwald, 2010). More importantly, trained plant pathologists are not always available, especially in poor and isolated areas. It is also worth noting that many agricultural areas are too expansive to be properly monitored throughout (Barbedo, 2013). Image-based tools can thus play an important role in detecting and recognising plant diseases when human assessment is unsuitable, unreliable or unavailable.

At present, the potential of automated tools has yet to be realised. In previous works, we have investigated the characteristics of the proposals found in the literature (Barbedo, 2013) and the main challenges that still prevent this kind of technology from being adopted in practice (Barbedo, 2016). Methods based on conventional machine learning techniques have been relatively successful under limited and constrained setups, but many of the difficulties associated with the intrinsic characteristics of the problem could not be properly handled. With the inception of deep learning concepts, the answer to those limitations seemed to be close. Indeed, since 2015 research on plant disease detection has strongly veered towards using deep learning.

According to Ferentinos (2018), deep learning refers to “the use of artificial neural network architectures that contain a quite large number of processing layers, as opposed to

E-mail address: jayme.barbedo@embrapa.br.
https://doi.org/10.1016/j.biosystemseng.2018.05.013

shallower architectures of more traditional neural network methodologies". Among deep learning tools, arguably the most commonly used are the Convolutional Neural Networks (CNN) (Krizhevsky, Sutskever, & Hinton, 2012). This kind of neural network requires fewer artificial neurons than conventional feedforward neural networks, being particularly suitable for image recognition. CNNs usually require a very large number of samples to be trained; however, in many real-world applications, it is expensive or unfeasible to collect the training data needed by the models (Pan & Yang, 2010). Thus, many authors are applying the concept of transfer learning to reuse pretrained networks (e.g. GoogLeNet and AlexNet), in which case predictions are done on examples that are not from the same distribution as the training data (Bengio, 2012). The conjunction of deep learning and transfer learning, together with the development of Graphics Processing Units (GPU), has provided a powerful tool for classification and recognition of diseases in plants (Ferentinos, 2018).

The application of deep learning to plant pathology problems started to gain momentum after 2015 (Table 1). As promising as the results seem to be, they must be interpreted with some observations in mind: a) most of the studies cited above use transfer learning in their experiments (Brahimi, Boukhalfa, & Moussaoui, 2017; Cruz, Luvisi, Bellis, & Ampatzidis, 2017; Ferentinos, 2018; Liu, Zhang, He, & Li, 2018; Mohanty, Hughes, & Salathé, 2016), and even those that do not apply this technique use CNN architectures that are similar to existing ones (Amara, Bouaziz, & Algergawy, 2017; DeChant et al., 2017; Lu, Yi, Zeng, Liu, & Zhang, 2017; Oppenheim & Shani, 2017); b) many studies used images contained in several versions of the PlantVillage dataset (Amara et al., 2017; Brahimi et al., 2017; Cruz et al., 2017; Ferentinos, 2018; Mohanty et al., 2016). As a consequence, most studies are applying similar tools to similar datasets. So, it is no surprise that there is not much variation in the results reported in the literature. In addition, there are many factors that affect deep learning-based tools when they are used under real field conditions, but in most cases these factors are only briefly discussed (or not considered at all). This causes the practical use of tools for automatic disease recognition to be still very limited. Although initiatives such as Plantix (PEAT, Berlin, Germany) are trying to change this scenario, there is still much to be done in order to effectively introduce this kind of technology to the daily routine of farms.

This article provides an in-depth analysis of the main factors that affect the performance of deep learning-based tools for plant disease recognition under realistic conditions. The goal was to provide some guidelines to make the investigation of deep learning-based methods for disease recognition more thorough and realistic.

The relevant factors mentioned above and discussed in detail in Section 4 were derived from experiments using CNNs. This was done by carefully analysing each misclassification produced by the model, and then associating them with a specific causal factor. This analysis provided a wealth of information which was used to draw out the remarks associated with each one of those factors. When appropriate, causes of misclassification were linked to error sources reported in the literature, thus providing further evidence of their generality. Thus, in the context of this work, the absolute accuracies yielded by the model are not nearly as important as the underlying causes for the errors. It is also important to emphasise that the image database used in the experiments was carefully built to reflect and reproduce many of the conditions expected to be found in practice. This database, which contains almost 50,000 images, is being made freely available for academic purposes at the address https://www.digipathos-rep.cnptia.embrapa.br/.

2. Materials and methods

The database used in the experiments is freely available and contains almost 50,000 images of 171 diseases affecting 21 plant species. However, only images of corn diseases were used in the context of this work. This subset was chosen because it contains the widest variety of conditions and a reasonable number of images for each of the nine diseases (Table 2), all of which are caused by fungi.

One of the main advantages of the deep learning approach is that, in general, the symptoms do not have to be explicitly identified in the image. However, relevant information is mostly concentrated in the symptoms themselves and their surrounds. Thus, in order to increase the size of the database and to test how the CNN would perform with more localised information, the original samples were divided into smaller images containing individual lesions or localised symptom regions (Fig. 1).

Some rules were applied for consistency in this division: a) images were manually blacked out prior to the subdivision; b) healthy tissue occupied at least 20% of the cropping area; c) isolated symptoms were taken individually; d) clustered

Table 1 – Studies employing deep learning for plant disease recognition.

Reference	Network	Dataset	Accuracy
Amara et al. (2017)	CNN (LeNet architecture)	PlantVillage	92%–99%
Brahimi et al. (2017)	CNN (AlexNet, GoogLeNet)	PlantVillage	99%
Cruz et al. (2017)	CNN (Modified LeNet)	Olive tree images (own)	99%
DeChant et al. (2017)	CNN (Pipeline)	Corn images (own)	97%
Ferentinos (2018)	CNN (Several)	PlantVillage	99%
Fuentes, Yoon, Kim, and Park (2017)	CNN (Several)	Tomato images (own)	83%
Liu et al. (2018)	CNN (AlexNet)	Apple images (own)	98%
Lu et al. (2017)	CNN (AlexNet inspired)	Rice images (own)	95%
Mohanty et al. (2016)	CNN (AlexNet, GoogLeNet)	PlantVillage	99%
Oppenheim and Shani (2017)	CNN (VGG)	Potato images (own)	96%