

Early Discovery of Tomato Foliage Diseases Based on Data Provenance and Pattern Recognition

Diogo Nunes¹, Carlos Werly¹, Gizelle Kupac Vianna¹,
and Sérgio Manuel Serra da Cruz^{1,2,3(✉)}

¹ UFRRJ – Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, Brazil
{diogo_c_nunes, carlos_werly, kupac, serra}@ufrjr.br

² PPGMMC/UFRRJ – Programa de Pós Graduação Modelagem Matemática e
Computacional, Seropédica, RJ, Brazil

³ PET-SI/UFRRJ – Programa de Educação Tutorial-Sistemas de Informação,
Seropédica, RJ, Brazil

Abstract. This work presents an approach focused in enhancing the quality of tomato crops. We are developing and using low cost computational strategies to support early detection of the late blight. Our approach consorts tomatoes cultivars in an experimental field with inexpensive computer-aided resources based on Web and Android mobile tools in which workers collect scouting data and annotations and take images about the state of the crop, and in image filtering techniques and pattern recognition to detect foliage diseases on tomatoes images. In this study, we use provenance metadata about field observations, images and farmers' annotations as well, to improve the efficiency and accuracy of the patterns recognition algorithms. Our identification method achieved a hit rate of 94.12 %, using a reduced set of digital images of the tomato crops.

Keywords: Provenance · Pattern recognition · Neural networks

1 Introduction

Side by side with a very competitive agribusiness, Brazil have a dynamic and diversified family agriculture, composed by 4.3 million small agricultural establishments, responsible for the production of very important products. This diversity at production gives a huge economic value for the Brazilian agriculture, which nowadays experiments a strong rhythm of growth at its productivity [1]. In 2013, it contributed approximately with 6 % of Brazil's Gross National Product [2]. Thus there are real key benefits in ensuring the quality of data used by farmers, smallholder and agronomists to support activities such as monitoring permanent or temporary crops, and planning for sustainable development.

Among the temporary crops with very expressive value of production are the tomatoes. Tomatoes are climacteric soft skin fruits, highly susceptible to diseases and contamination, mainly through injured skin or damaged tissues during the plantation stage [3]. The indiscriminate use of pesticides in tomato crops brings serious problems

to human health and to environment. Last but not least, smallholders may lack the resources to comply the increasingly strict standards of food safety, as customer traceability needs, safety verification and inventory control, mainly because they do not have full access to computer-based systems that alert, detect or predict the occurrence of diseases. Such facilities may aid smallholders and farmers to reduce the chemical management of tomato diseases, to grow their income levels and to deliver healthier products. This research presents an approach focused in enhancing the quality of tomatoes crops based on provenance annotations and neural networks.

2 Materials and Methods

We are developing a loosely coupled distributed system that uses low cost computational strategies to support early detection of the late blight [4], the commonest tomato disease in Brazil. Late blight is severe and results in losses to tomato crops, especially in colder and wet periods of the year. The disease is visually recognized by the appearance of dark spots on tomato leaves, whose blotches vary from brown or gray to pale green, often located at the edges of the tomato leaves [3, 4].

Our approach consorts real field data from 66 different genotypes of tomatoes cultivars (plant varieties produced by breeding, grown under the principles of organic agriculture in a controlled experimental field) with inexpensive computer-aided resources based on:

(i) *Android-based mobile tools* in which smallholders may collect scouting data attributes for pests such as the occurrence of weeds, insects and diseases. They could also collect common annotations about the crop status and other environmental information, take georeferenced photos about the infected (or suspected) tomato stems and leaves, with a built-in camera. The tools would also manage the collected data and provenance annotations;

(ii) *Image filtering techniques and pattern recognition based on Multilayer Perceptron neural networks* to detect late blight on tomatoes images. Once acquired, these scouting data, annotations and images are transferred from the mobile devices to a base station that hosts the Web application and an Internet connection. To ensure that the data collected in the field have a good level of quality, the dataset is pre-processed with the objective to produce higher quality data. The processor checks the missing or erroneous data and annotations, the outliers and the quality of the images before performing pattern recognition algorithms. The Web application allows users to investigate rectified faulty data and annotate datasets with provenance to reduce error propagation on long-term evaluations. Besides, it is designed to allow the creation of analytical reports to empower smallholder and farmers, allowing them to: visualize the spreading of late blight on the field; reduce chemical sprayings; send or receive reports and image samples to agronomist professional services.

3 Foliage Disease Pattern Recognition

In this study, we use provenance annotations about field observations, to improve the efficiency and accuracy of patterns recognition algorithms, previously developed by our

research group [5, 6]. The pattern recognition algorithm performs the following steps: first, it reduces the definition of the colored images about 70 % in order to speed the performance of further procedures; second, the reduced colored image is converted to black and white images; third, the reduced color images are used to generate novel images, containing only red, green and black pixels. Over the images created by the previous steps, we conduct a counting of pixels from different color bands. These counters are then stored in a repository that contains the absolute number of pixels of each color. Next, data are normalized, generating a new repository of those variables. Finally, the variables are evaluated to recognize the pattern of late blight disease, more specifically; the recognition used techniques of MLP neural networks [7].

4 Conclusion

Our foliage diseases identification approach achieved a hit rate of 94.12 %, using a reduced set of digital images and annotations about the tomato crops. The main role of provenance in our research is to ensure food safety, consumer protection and to offer opportunities to smallholders to run their operations more productively. The software tools were designed to use low cost apparatus and to guarantee the ease of use, because smallholders have little room for error and little expertise in digital processing with sophisticated programs.

Acknowledgements. We are grateful by the financial support provided by FAPERJ (E-26/112.588/2012 and E-26/110.928/2013 and FNDE-MEC-SeSU).

References

1. IBGE - Brazil in Figures (2013). http://biblioteca.ibge.gov.br/visualizacao/periodicos/2/bn_2013_v21.pdf
2. IBGE - Contas Nacionais Trimestrais Indicadores de Volume e Valores Correntes. (2013) http://ftpibge.gov.br/Contas_Nacionais/Contas_Nacionais_Trimetrais/Fasciculo_Indicadores_IBGE/pib-vol-val_201304caderno.pdf
3. Nakano, O.: As pragas das hortaliças: seu controle e o selo verde. Horticultura Brasileira, Vol. 17, n.1 UnB (1999)
4. Correa, F.M., Bueno Filho, J.S.S., Carmo, M.G.F.: Comparison of three diagrammatic keys for the quantification of late blight in tomato leaves. *Plant Pathol.* **58**, 1128–1133 (2009)
5. Vianna, G.K., Cruz, S.M.S.: Análise Inteligente de Imagens Digitais no Monitoramento da Requeima em Tomateiros. Anais do IX Congresso Brasileiro de Agroinformática. Cuiabá, MT (2013)
6. Cruz, S.M.S., Campos, M.L.M., Mattoso, M.: Towards a taxonomy of provenance in scientific workflow management systems. In: Proceedings of the SERVICES 2009 Congress, pp. 259–266, Los Angeles (2009)
7. Bishop, C.M.: *Neural Networks for Pattern Recognition*. Oxford University Press, New York (1995)