Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

An Image Analysis Environment for Species Identification of Food Contaminating Beetles

Daniel Martin[1], Hongjian Ding[2], Leihong Wu[3], Howard Semey[2], Amy Barnes[2], Darryl Langley[2], Su Inn Park[4], Zhichao Liu[3], Weida Tong[3], Joshua Xu[3,*]

[1]School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85281, USA
[2]ARL Chemistry Lab1, Arkansas Regional Laboratory, Office of Regulatory Affairs, US Food and Drug Administration, AR 72079, USA
[3]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR 72079, USA
[4]Data Science Team, System Technology, Samsung Austin Semiconductor LLC, Austin, TX 78754, USA
*Correspondence: Joshua Xu (Zhihua.xu@fda.hhs.gov)

Abstract

Food safety is vital to the well-being of society; therefore, it is important to inspect food products to ensure minimal health risks are present. The presence of certain species of insects, especially storage beetles, is a reliable indicator of possible contamination during storage and food processing. However, the current approach of identifying species by visual examination of insect fragments is rather subjective and time-consuming. To aid this inspection process, we have developed in collaboration with FDA food analysts some image analysis-based machine intelligence to achieve species identification with up to 90% accuracy. The current project is a continuation of this development effort. Here we present an image analysis environment that allows practical deployment of the machine intelligence on computers with limited processing power and memory. Using this environment, users can prepare input sets by selecting images for analysis, and inspect these images through the integrated panning and zooming capabilities. After species analysis, the results panel allows the user to compare the analyzed images with reference images of the proposed species. Further additions to this environment should include a log of previously analyzed images, and eventually extend to interaction with a central cloud repository of images through a web-based interface.

Background

Food sourcing has become increasingly global in recent years. This greater distance between source and endpoint has increased transportation and storage times, leading to a higher risk of contamination before food products reach consumers. Thus, food inspection has grown in importance for enforcing food safety regulations. A common indicator of contamination acquired during processing and storage is the presence of storage insects (Olsen et al. 2001).

The current inspection process used throughout the industry involves manual examination by analysts (Olsen et al. 1996), and is subjective and time-consuming. In order to standardize, streamline, and automate the contaminant identification process, the authors have proposed the use of image analysis (Gonzalez et al. 2009) and machine learning (Mohri et al. 2012) to mirror the current identification process. This paper is a follow-up to previous research by Park et al. (preprint), which can be accessed in the references.

This paper provides an overview of an image analysis environment for identifying beetle elytra, or hardened forewings, as well as an evaluation of the implementation technology.

Implementation

The image analysis environment was implemented in MATLAB. MATLAB was chosen as it is widely used for image processing and machine learning due to its function toolboxes for those areas. It also includes a Graphical User Interface Design Environment (GUIDE) tool that simplifies the process of constructing basic layouts of interface components. Although MATLAB is a commercial application, the MATLAB Compiler Runtime (MCR) library needed to run compiled MATLAB applications is free, making MATLAB an attractive solution for creating a deployable application.

The main screen of the environment is the Image Manager, which features an image preview area, a list of possible inputs in the current folder, and information about the selected image. A settings dialog is also present for the user to change the number of subimages taken as well as the current neural network to use.

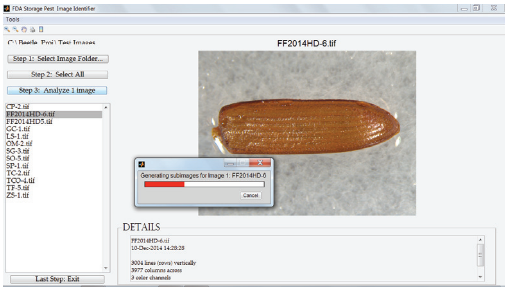

Figure 1: Choosing Images for Analysis.

Once the analysis is finished, the Analysis Report window appears. This window contains preview areas for the input and reference images, as well as the resulting species name and confidence level. As in the Image Manager window, the user can zoom and pan both images for better visual comparison and confirmation. Currently the program only returns the resulting species with the highest confidence level out of the 15 reference storage beetle species. Both limitations, i.e., the number of resulting candidates and the number of reference species included in the original model by Park et al., will be addressed in the future. Finally, the results can be printed out to an HTML log for later reference.

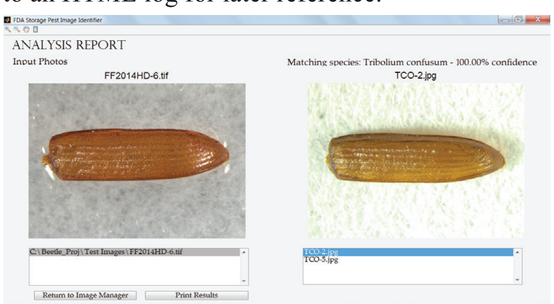

Figure 2: Results of an Analysis.

During implementation, the authors have overcome several performance challenges with MATLAB in its functionality and design. For example, compiled MATLAB applications can run only with the MCR of the same version as the compiler. Compiling MATLAB code with a different version may affect the program, as the implementation of functions might change between versions. The task of compiling MATLAB is further complicated by the absence of a toolbox dependency checker. Finally, the official documentation provides mostly rudimentary instructions for using GUIDE and MATLAB Compiler, forcing the developer to the unofficial forums for most common problems.

While MATLAB presents a few challenges, especially with its user interface and compiler tools, its strengths in image processing and machine learning make it competitive at this stage.

Conclusion

This paper presents an image analysis environment to streamline routine food contaminant identification. It is easy to use, standardized, and reproducible. By reducing the manpower and time costs of inspection, this program aids in increasing processing capacity. Its analysis is faster than current manual inspection.

In the future, this program can be enhanced in several ways, such as allowing multiple result windows to be open at once; being able to identify hairs, bird feathers, and other contaminants; and finally, porting the system to a web platform.

Acknowledgements

Daniel Martin and Su Inn Park are grateful to the National Center for Toxicological Research (NCTR) of U.S. FDA for the Summer Student Research Program administered by the Oak Ridge Institute for Science and Education. The authors appreciate the critical review of the manuscript by Dr. Vikrant Vijay and Mr. Joseph Meehan of FDA/NCTR.

Disclaimer

The views presented in this article do not necessarily reflect the views of the US Food and Drug Administration.

References

Gonzalez, R. C.; Woods, R. E.; and Eddins, S. L. 2009. *Digital Image Processing Using MATLAB*, 2nd ed. Knoxville, TN: Gatesmark Publishing.

Mohri, M.; Rostamizadeh, A.; and Talwalkar, A. 2012. *Foundations of Machine Learning.* Cambridge, MA: The MIT Press.

Olsen, A. R.; Gecan, J. S.; Ziobro, G. C.; and Bryce, J. R. 2001. Regulatory Action Criteria for Filth and Other Extraneous Materials V. Strategy for Evaluating Hazardous and Nonhazardous Filth. *Regulatory Toxicology and Pharmacology* 33(3):363-392.

Olsen, A. R.; Sidebottom, T. H.; and Knight, S. A. 1996. *Fundamentals of Microanalytical Entomology: A Practical Guide to Detecting and Identifying Filth in Foods.* Boca Raton, FL: CRC Press.

Park, S.; Bisgin, H.; Ding, H.; Semey, H.; Langley, D.; Tong, W.; and Xu, J. Species Identification of Food Contaminating Beetles by Recognizing Patterns in Microscopic Images of Elytra Fragments. Preprint, submitted September 9, 2015.