INTERNATIONAL UNION of FOOD SCIENCE AND TECHNOLOGY
IUFoST

IUFoST Scientific Information Bulletin (SIB)
September 2020

Recent Advances of Artificial Intelligence Applications in Food

Summary

Artificial intelligence (AI) is a general heading related to smart systems design that can perform tasks comparable to/superior to human intelligence. The recent success of artificial intelligence is mainly thanks to a specific type of mathematical model named as artificial neural network. With the development of AI algorithms and computational hardware, very complicated artificial neural network composed of billions of neurons and hundreds of layers can be built and trained, which is also the origin of the concept 'deep learning'. As an important branch of artificial intelligence, deep learning has become a well-known term since 2012; it has shown state-of-the-art performance in various data analytical tasks, especially in the image analysis field. Catching the wave of deep learning, some food-related applications have sprung up in recent years. This Scientific Information Bulletin (SIB) will summarize these deep learning application developments in food related fields and will introduce food scientists and technologists to this advanced computational method. It is expected that even a basic understanding of this powerful tool will assist researchers with their current studies while opening opportunities for exploring novel research areas.

Introduction

The term Artificial intelligence (AI) describes the "cognitive" functions of computers and machines. In the 50-year history of AI, this field has made remarkable achievements which play very important roles in both research and industrial applications.

The latest wave of AI started in 2012, and was facilitated by (1) the success of novel computation models like deep learning neural network models (Krizhevsky, Sutskever, & Hinton, 2012); (2) the advent of large open-access datasets (Deng, et al., 2009); (3) the increase of hardware support provided by acceleration devices (such as Graphics Processing Units) and deep learning software libraries (Alom, Hasan, Yakopcic, Taha, & Asari, 2018).

This new AI boom has attracted wide attention from various fields including healthcare, robotics, and basic science, among others, but has only recently been introduced to food science and engineering (Kassim & Palaniappan, 2017). Thus, the *research justification* for this Scientific Information Bulletin (SIB) is to provide a brief introduction to deep learning and a summary of its recent food-related applications. For food scientists and technologists, understanding the basic concept of this powerful tool will potentially change their perspective in how one approaches research problems and explores new research fields.

Deep learning overview

Currently, the term deep artificial intelligence refers to the classic artificial neural network (ANN) with more perceptron layers. ANN is a supervised learning model inspired by the biological neural networks of our nervous system. It is a mathematical model that describes an input/output relationship. This model can be trained based on human experience (training data), and the well-trained model can subsequently be applied into new inputs (test data) for practical applications. Here is a simple example from 2015 Nature review paper (LeCun, Bengio, & Hinton, 2015): “Imagine that we want to build a system that can classify images as containing, say, a house, a car, a person or a pet. We first collect a large data set of images of houses, cars, people and pets, each labelled with its category. During training, the machine is shown an image and produces an output in the form of a vector of scores, one for each category.” In this example, the model input is an image and the output is a class category label (house, car, person, or pet). A typical workflow for training a supervised learning model is summarized in Figure 1.

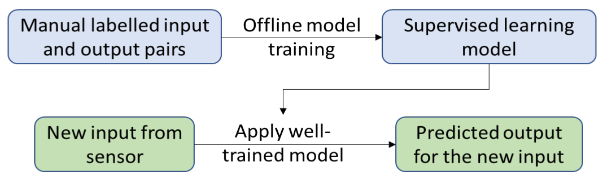

Figure 1. A typical workflow diagram of training a supervised learning model (such as ANN), the training process needs to manually prepare input and output pairs, and the well-trained model can be directly applied to the new data.

The recent success of deep ANNs is largely attributed to a particular type of ANN known as convolutional neural networks (CNN). CNNs can automatically extract local and global data features from raw sensor data without requiring considerable engineering skills or domain expertise (Fu, Xu, Lin, Wong, & Liu, 2016). The effectiveness of CNN has been proven for many common data types, such as images and audio signals.

Recently, the success of deep learning has brought up many new food-related research and applications. These will be introduced in this SIB. However, it is worthwhile to mention that current AI models are considered “narrow AI”, since each model can only focus on a specific targeted task during training. Therefore, researchers still need to choose specific AI models for different food-related applications, and prepare corresponding datasets to train each respective model according to the required task.

Public Food Image Datasets for Deep Learning of Future Food Service Sectors

Large scale Image recognition is one of the most important applications of deep learning, and relies on a large image dataset for supervised network training. Food-101 (Bossard et al., 2014) is the largest public dataset for food image classification. It consists of 101,000 images of the top 101 food categories from ‘foodspotting.com’ as shown in Figure 2(a). Japanese researchers released two food detection datasets UECFOOD-100 (Matsuda et al., 2012) and UECFOOD-256 (Kawano & Yanai, 2014) for detecting the location of food in images and predicting the categories of foods. UECFOOD-100 contains 9,060 images of the most popular 100 classes of Japanese foods as. UECFOOD-256 further added international foods of various countries with 256 categories. Chinese researchers recently prepared BTBUFood-60 detection dataset to better localize food in real scenarios. It includes 60,000 images of 60 food categories and 78,000 bounding-box labelled sample images which can be used for training the network to accurately detect the position of specific food in the camera field of view. It can further benefit for potential robotic operations for fully-automated warehouses and grocery stores (Cai et al., 2019).

Figure 2. Example images from Food-101 dataset (Bossard, Guillaumin, & Van Gool, 2014).

With these known categories of food, the nutrient content of that food can be estimated based on the food and nutrient database released by [USDA Food and Nutrient Database for Dietary Studies](). Using equipment like 3D cameras and smart phones, the accuracy of nutrient content estimation is expected to rise in the near future. With this goal in mind, Meyers et al. (2015) reported the development of an automated mobile vision food diary system called Im2Calories. This app uses a RGB depth image of food to estimate the volume of food and its respective calories. This helps users monitor and control their dietary behaviors. Amazon Go and Amazon Fresh also try to integrate various AI techniques to push the developments of advanced shopping technologies.

Food Processing Automation Applications

While machine-vision based food processing devices have attracted much attention in the past three decades, the success of deep learning is expected to bring new life to this field. Perhaps, one of the most important applications is the design of autonomous grading and sorting machines. Food grading and category-based sorting are standard image classification tasks with the purpose of classifying food products into different grade levels or categories.

Ponce, Aquino, and Andújar (2019) classified seven different olive varieties based on 2800 images to optimize post-harvesting treatments and manipulation processes). Figure 3 shows their ad-hoc image acquisition system with four sample images for different varieties of olive-fruits. Ni, Wang, Vinson, Holmes, and Tao (2019) designed a machine to classify qualified/defective maize kernels through a dual-stream CNN which inspects both sides of kernels at the same time, achieving 98.2% prediction accuracy for 408 test images. This system is helpful for setting kernels` market prices via quick online kernel inspection. Similarly, Liu et al. (2019) did comprehensive network performance evaluation based on their shrimp classification dataset, which had 1731 images for nine categories of shrimp. The also build their own processing lines with shrimp and imaging chamber for on-line testing. Besides these post-harvesting studies, Rodríguez, García, Pardo, Chávez, and Luque-Baena (2018) utilized 525 images to classify three varieties of plum at the pre-harvest stage to help farmers adapt their harvest to the demand of their market.

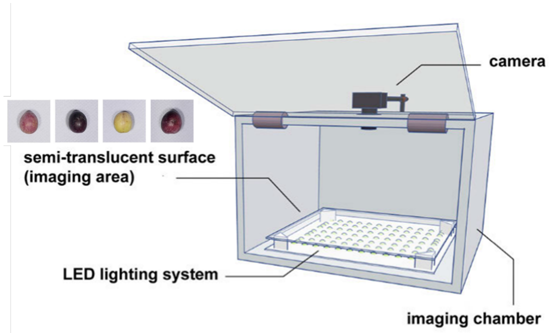

Figure 3. illumination and imaging chamber for olive varieties classification. (Ponce et al., 2019)

Further, AI deep learning methods can improve image analysis tasks that help guide robotic movements for complicated food processing tasks. For example, the Bioimaging and Machine Vision (BMV) Lab at the University of Maryland (UMD) automated the removal of strawberry calyx for processed foods including yogurt, smoothie, and cakes through the integration of a new end-to-end deep network model, AVIDnet, (Wang et al., 2019) into a high-throughput waterjet based strawberry de-calyxing machine (Lin et al., 2017). AVIDnet can automatically recognize the qualified strawberry, and simultaneously generate a 2D decalyx cutline for specific strawberries. These strawberries cutlines were manually labelled beforehand. Figure 4 shows the AVID machine and the architecture of AVIDnet.

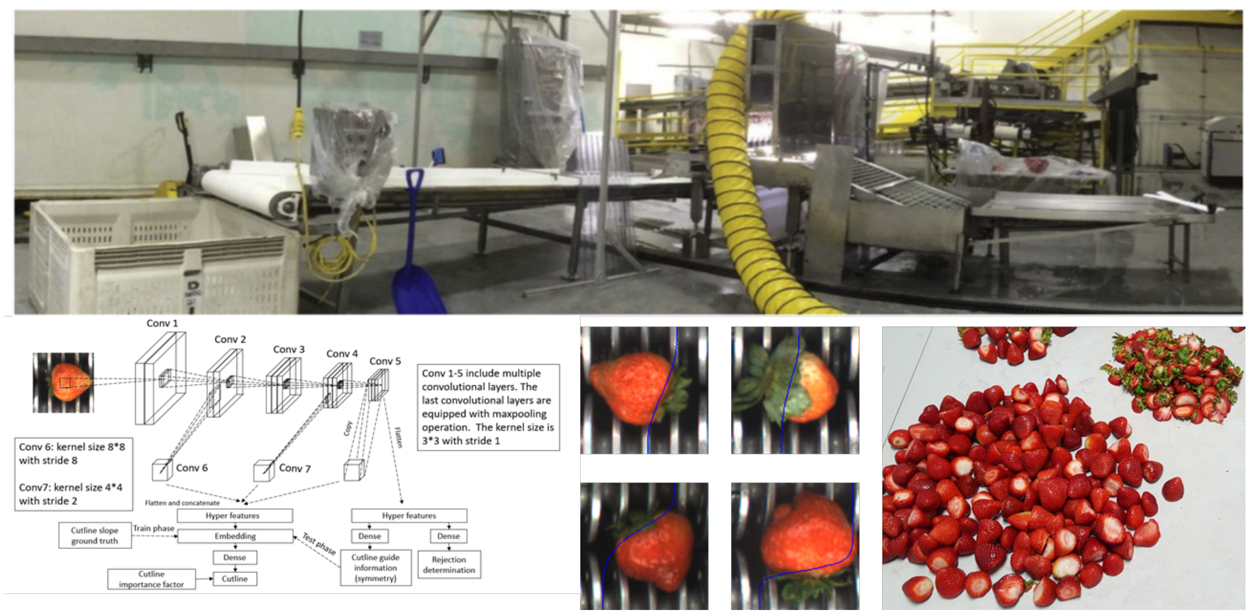

Figure 4. Automated strawberry decalyx machine integrated with end-to-end cutline generation network (Wang, Vinson, Seibel, Cheng, & Tao, 2019)

Instead of directly outputting the robotic movement trajectories, the deep learning model can also be designed to identify specific image features, which can further generate the robotic movement trajectories. In the crab processing machine designed by the BMV Lab at UMD, we firstly utilized the deep neural networks to locate the crab back-fin knuckles, and then, to remove blue crab legs, the robotic movement trajectories were generated by traditional template matching methods (Wang et al., 2018). This machine is expected to alleviate the labor shortage problem in crab industry, with the idea

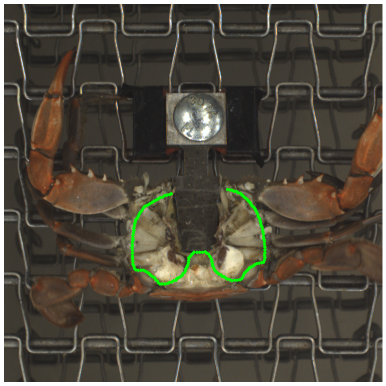

Figure 5. An example crab leg removal trajectory generated based on deep

that combining deep learning methods and traditional methods will increase the robustness of the machine. An example crab core extraction and leg removal trajectory is shown in Figure 5.

Besides in plane applications, deep learning models also integrate with some in-field machines for field monitoring. Many attempts apply deep-learning based detection models to assist apple harvesting (Kang & Chen, 2020; Onishi et al., 2019), as shown in Figure 6, and yield estimations (Koirala, Walsh, Wang, & McCarthy, 2019). Similar research has been extended to different species including strawberries (Yu, Zhang, Yang, & Zhang, 2019) and tomatoes (Zhang et al., 2018). With further combinations of automation and AI, more food-related real-world intelligence machines are expected to crop up in following years.

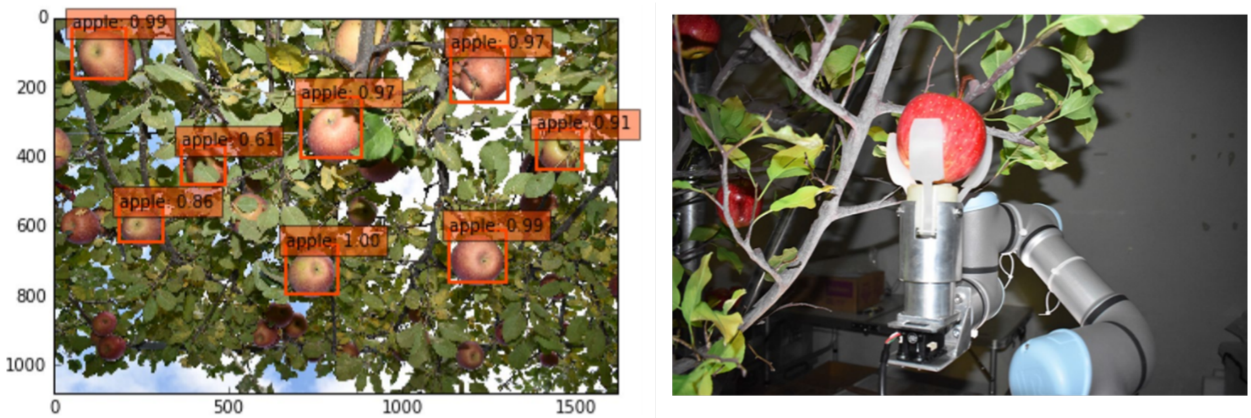

Figure 6. Left: example apple detection results based on deep learning model SSD. Right: apple harvesting based on UR3 robots. (Onishi et al., 2019)

Spectroscopy and hyperspectral imaging analysis

Advanced spectroscopy and hyperspectral imaging techniques offer more invisible information compared to regular RGB image sensor, which can better inspect the food product qualities (Sun, 2010). Conventional spectral signal analysis methods focus on one or more signal peaks independently. Instead, deep-learning based spectral signal analysis can better consider the intra-wavelength relationships of signal peaks for specific classification and regression tasks. This concept was applied by researchers from USA (Liu et al., 2018), who combined CNN and stacked sparse auto-encoder (SSAE) to recognize defect cucumbers, as shown in Figure 7. In their design, CNN was firstly used to identify detecting regions. Within the detected region, SSAE can extract hidden features from the mean spectral signal to judge various cucumber qualities.

Furthermore, deep learning is a potential and powerful chemometrics tool for qualitative spectroscopic analysis. Spectroscopy, especially near or short infrared spectroscopy (with wavelength range 780-2500nm), is a powerful tool in food analysis which can reflect chemical bonds vibrations of food constitutes. As a consequence of the physics of diffuse transmittance and reflectance and the complexity of the spectra, spectroscopy analysis is normally carried out using multivariate mathematics models (Osborne, 2006). Researchers explored the feasibility of CNN to predict the content value using some public and self-collected spectroscopy datasets. Compared to large image datasets in other fields, these spectroscopy datasets are relatively small (<1000 data samples). One of the most popular datasets is [a corn kernel spectroscopy dataset,](#) released by Eigenvector Research, Inc., which offers the moisture, oil, protein and starch values of corn, as shown in Figure 8. Researchers have also tried to use CNN or auto-encoder models to study different food products including wheat (Cui & Fearn, 2018), berries (Zhang et al., 2020), orange juice, and wine (Malek, Melgani, & Bazi, 2018). One of the main drawbacks of current CNN based spectroscopic analysis is that the model performance is very sensitive to the network hyperparameter selections (Acquarelli et al., 2017), which is mainly explained by the limited size of training data. Also, as with other ANN models, the running mechanisms inside CNN models are 'black-boxes' which make the model hard to explain from a scientific point of view. Although some researches have started to explain the wavelength importance inside the network (Ni, Wang, & Tao, 2019), the design of reliable deep learning models still has a long way to go without the availability of a large spectroscopic dataset.

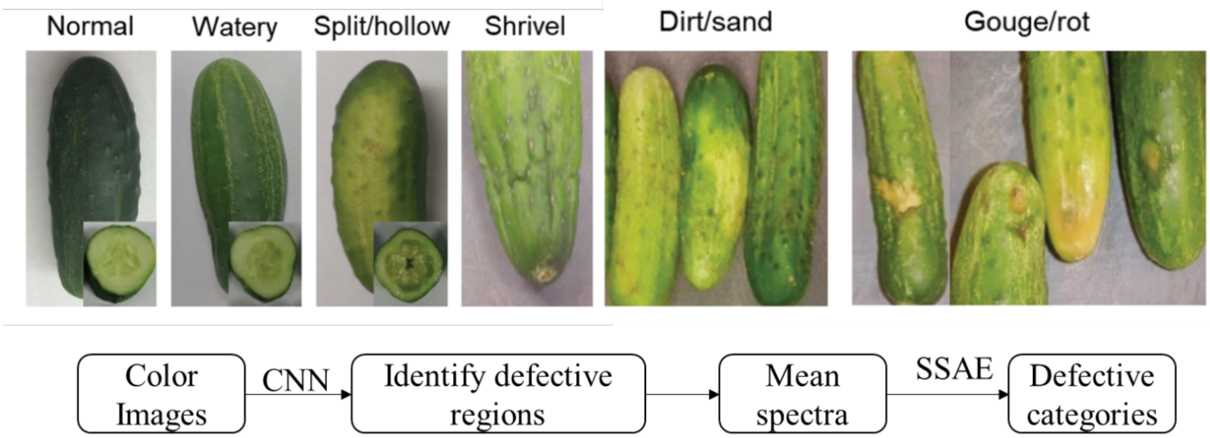

Figure 7. Deep learning based spectral analysis for cucumber defection detections. (Liu, He, Cen, & Lu, 2018)

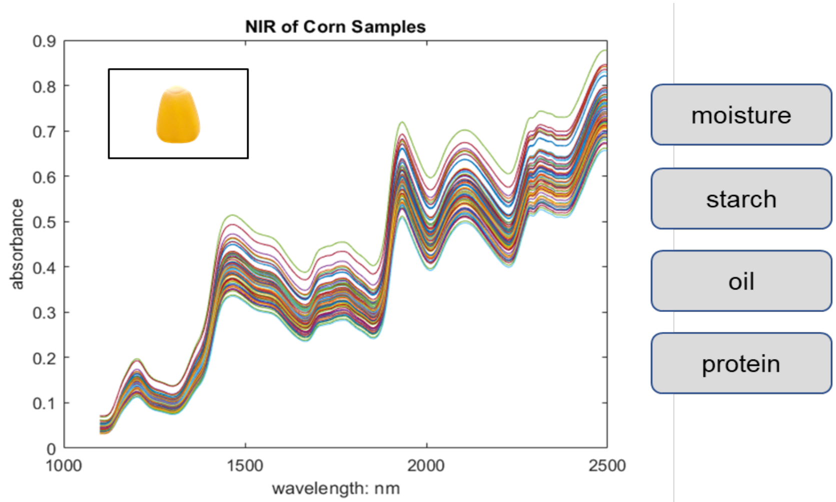

Figure 8. Corn kernel spectroscopy dataset released by Eigenvector Research, Inc (https://eigenvector.com/resources/data-sets/)

Conclusion and future prospective

In this bulletin, some basic ideas of artificial intelligence, and specifically deep learning were introduced along with some recent deep learning achievements in food engineering fields. Though deep learning has achieved remarkable performance in many tasks, manually-labelled data is still indispensable component based on current deep AI algorithms. Therefore, preparing and publishing well-labelled food-related dataset will accelerate the successful application of deep learning in food related fields and cater to the trend of mobile computing. Besides existing calorie estimation applications, estimating food freshness and qualities via cell phones are also expected to yield huge market benefits. Meanwhile, advanced machine vision technologies can accelerate the development of high throughput intelligence machines for pre-harvesting and post-harvesting during food processing. Fundamental food scientists can use the powerful tools offered by advanced deep learning models to explain trends in research data. While CNN has achieved extraordinary performances in the analysis of RGB color images, its applications must be studied further for explaining advanced sensory information such as one-dimensional spectroscopic data, three-dimension hyperspectral data, or information from other sensors.

Websites

Definition of artificial intelligence. https://en.wikipedia.org/wiki/Artificial_intelligence#History_of_AI

Definition of ANN. https://en.wikipedia.org/wiki/Artificial_neural_network

Definition of object detection. https://en.wikipedia.org/wiki/Object_detection

Definition of image segmentation.
https://en.wikipedia.org/wiki/Image_segmentation#Segmentation_benchmarking

References

Acquarelli, J., van Laarhoven, T., Gerretzen, J., Tran, T. N., Buydens, L. M., & Marchiori, E. (2017). Convolutional neural networks for vibrational spectroscopic data analysis. *Analytica chimica acta, 954*, 22-31.

Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., & Asari, V. K. (2018). Recurrent residual convolutional neural network based on u-net (R2U-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*.

Bossard, L., Guillaumin, M., & Van Gool, L. (2014). Food-101–mining discriminative components with random forests. In *European Conference on Computer Vision* (pp. 446-461): Springer.

Cai, Q., Li, J., Li, H., & Weng, Y. (2019). BTBUFood-60: Dataset for Object Detection in Food Field. In *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 1-4): IEEE.

Cui, C., & Fearn, T. (2018). Modern practical convolutional neural networks for multivariate regression: Applications to NIR calibration. *Chemometrics and Intelligent Laboratory Systems, 182*, 9-20.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255): Ieee.

Fu, H., Xu, Y., Lin, S., Wong, D. W. K., & Liu, J. (2016). Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. In *International conference on medical image computing and computer-assisted intervention* (pp. 132-139): Springer.

Kang, H., & Chen, C. (2020). Fast implementation of real-time fruit detection in apple orchards using deep learning. *Computers and electronics in agriculture, 168*, 105108.

Kassim, Y. M., & Palaniappan, K. (2017). Extracting retinal vascular networks using deep learning architecture. In *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1170-1174): IEEE.

Kawano, Y., & Yanai, K. (2014). Automatic expansion of a food image dataset leveraging existing categories with domain adaptation. In *European Conference on Computer Vision* (pp. 3-17): Springer.

Koirala, A., Walsh, K. B., Wang, Z., & McCarthy, C. (2019). Deep learning–Method overview and review of use for fruit detection and yield estimation. *Computers and electronics in agriculture, 162*, 219-234.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature, 521*(7553), 436.

Lin, J., Holmes, M., Vinson, R., Ge, C., Pogoda, F. C., Mahon, L., Gentry, R., Seibel, G. E., Chen, X., & Tao, Y. (2017). Design and testing of an automated high-throughput computer vision guided waterjet knife strawberry calyx removal machine. *Journal of Food Engineering, 211*, 30-38.

Liu, Z., He, Y., Cen, H., & Lu, R. (2018). Liu, Z., He, Y., Cen, H., & Lu, R. (2018). Deep feature representation with stacked sparse auto-encoder and convolutional neural network for hyperspectral imaging-based detection of cucumber defects. *Transactions of the ASABE*, 61(2), 425-436.

Liu, Z., Jia, X., & Xu, X. (2019). Study of shrimp recognition methods using smart networks. *Computers and Electronics in Agriculture, 165*, 104926.

Malek, S., Melgani, F., & Bazi, Y. (2018). One-dimensional convolutional neural networks for spectroscopic signal regression. *Journal of Chemometrics, 32*(5), e2977.

Matsuda, Y., Hoashi, H., & Yanai, K. (2012). Recognition of multiple-food images by detecting candidate regions. In *2012 IEEE International Conference on Multimedia and Expo* (pp. 25-30): IEEE.

Meyers, A., Johnston, N., Rathod, V., Korattikara, A., Gorban, A., Silberman, N., Guadarrama, S., Papandreou, G., Huang, J., & Murphy, K. P. (2015). Im2Calories: towards an automated mobile vision food diary. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1233-1241).

Ni, C., Wang, D., & Tao, Y. (2019). Variable weighted convolutional neural network for the nitrogen content quantization of Masson pine seedling leaves with near-infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 209*, 32-39.

Ni, C., Wang, D., Vinson, R., Holmes, M., & Tao, Y. (2019). Automatic inspection machine for maize kernels based on deep convolutional neural networks. *Biosystems engineering, 178*, 131-144.

Onishi, Y., Yoshida, T., Kurita, H., Fukao, T., Arihara, H., & Iwai, A. (2019). An automated fruit harvesting robot by using deep learning. *ROBOMECH Journal, 6*(1), 13.

Osborne, B. G. (2006). Near-infrared spectroscopy in food analysis. *Encyclopedia of analytical chemistry: applications, theory and instrumentation*.

Ponce, J. M., Aquino, A., & Andújar, J. M. (2019). Olive-Fruit Variety Classification by Means of Image Processing and Convolutional Neural Networks. *IEEE Access, 7*, 147629-147641.

Rodríguez, F. J., García, A., Pardo, P. J., Chávez, F., & Luque-Baena, R. M. (2018). Study and classification of plum varieties using image analysis and deep learning techniques. *Progress in Artificial Intelligence, 7*(2), 119-127.

Sun, D.-W. (2010). *Hyperspectral imaging for food quality analysis and control*: Elsevier.

Wang, D., Holmes, M., Vinson, R., Seibel, G., & Tao, Y. (2018). Machine Vision Guided Robotics for Blue Crab Disassembly---Deep Learning Based Crab Morphology Segmentation. In *2018 ASABE Annual International Meeting* (pp. 1): American Society of Agricultural and Biological Engineers.

Wang, D., Vinson, R., Seibel, G., Cheng, X., & Tao, Y. (2019). Vision Intelligence guided strawberry decalyx machine--automatic vision-guided intelligent decalyxing network (AVIDnet). In *2019 ASABE Annual International Meeting* (pp. 1): American Society of Agricultural and Biological Engineers.

Yu, Y., Zhang, K., Yang, L., & Zhang, D. (2019). Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. *Computers and electronics in agriculture, 163*, 104846.

Zhang, C., Wu, W., Zhou, L., Cheng, H., Ye, X., & He, Y. (2020). Developing deep learning based regression approaches for determination of chemical compositions in dry black goji berries (Lycium ruthenicum Murr.) using near-infrared hyperspectral imaging. *Food Chemistry*, 126536.

Zhang, L., Jia, J., Gui, G., Hao, X., Gao, W., & Wang, M. (2018). Deep learning based improved classification system for designing tomato harvesting robot. *IEEE Access, 6*, 67940-67950.

This Scientific Information Bulletin (SIB) was prepared by **Yang Tao, PE, PhD, Director of Bioimaging and Machine Vision Lab, Professor and Dongyi Wang, PhD candidate in the Fischell Department of Bioengineering, University of Maryland, College Park, USA.** This SIB was reviewed and approved by the IUFoST Scientific Council.

ABOUT IUFoST

The International Union of Food Science and Technology (IUFoST) is the global scientific organization working with more than 300,000 food scientists and technologists in over 100 countries around the world. IUFoST is a full scientific member of ISC (International Science Council) and the only elected representative of Food Science and Technology in the ISC. IUFoST represents food science and technology to international organizations such as WHO, FAO, UNDP, UNIDO, The World Bank, and others. IUFoST organizes world food congresses, among many other activities, to stimulate the ongoing exchange of knowledge and to develop strategies in those scientific disciplines and technologies relating to the expansion, improvement, distribution and conservation of the world's food supply.

General Secretariat, IUFoST: secretariat@iufost.org, www.iufost.org