10.2352/EI.2022.34.15.COLOR-157

Deep Learning Approach for Classifying Contamination Levels with Limited Samples

Min Zhao, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA
Susana Diaz-Amaya, R&D|Sustainability & Outreach—Bayer at Convergence - Bayer Crop Science, West Lafayette, IN, USA
Amanda J. Deering, Department of Food Science, Purdue University, West Lafayette, IN, USA
Lia Stanciu, School of Materials Engineering, Purdue University, West Lafayette, IN, USA
George T.-C. Chiu, School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA
Jan P. Allebach, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA

Abstract

The detection of the contaminants in daily food and drinking water is crucial for global public health. For heavy metals detection of Mercury (Hg) and Arsenic (As), our group has proposed a novel paper-based and microfluidic device integrated with a mobile phone and an image analysis pipeline to capture and analyze the sensor images on-site. Still, the detection of lower contamination levels remains challenging due to the small number of available data samples and large intra-class variance of our application. To overcome this challenge, we explore traditional data augmentation and GAN-based augmentation techniques for synthesizing realistic colorimetric images; and we propose a CNN classifier for five-contamination-levels classification. Our proposed system is trained and evaluated on a limited dataset of 126 phone captured images of five contamination levels. Our system yields 88.1% classification accuracy and 91.92% precision, demonstrating the feasibility of this approach. We believe that this approach of training deep learning models on limited detection images datasets presents a clear path toward phone-based contamination-levels detection.

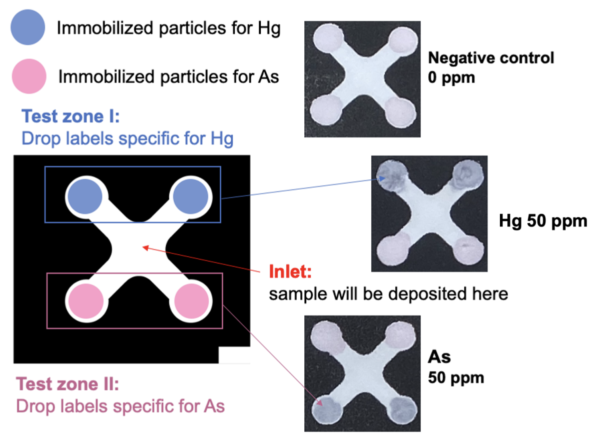


Figure 1. *Detection mechanism of our biosensors and test interpretation. (To illustrate the different particles specific for Hg and As, the particles specific for Hg are labeled blue in the figure; but the actual particles are colored light pink.)*

Introduction

To address the threats of heavy metals As and Hg to food safety, our group has proposed a novel paper-based, microfluidic biosensor. Fig. 1 shows the proposed detection mechanism of our biosensors and the test interpretation. Two kinds of the aptamer-functionalized particles specific to Hg^{2+} and As^{3+} are preloaded on each of the upper, and lower two circular pads, respectively. Test samples are dropped in the inlet of the biosensors. The biosensor shows colorimetric responses in the presence of the target after the test solution interacts with the corresponding particles deposited on the testing areas [1].

To detect and measure heavy metal contaminants in food or liquids, we propose two image analysis methods to obtain a higher prediction accuracy with our developed paper-based devices in our previous work. (1) We convert images of the colorimetric response captured with a mobile phone camera to grayscale images, and then we use ΔE from a white background as our baseline method to correlate the optical properties with the different concentrations of the target, and optimally quantize these responses into five groups to evaluate the prediction accuracy [2], [3]. The prediction accuracy is shown in Table 1. The ΔE from the global background method shows an average prediction performance of 60%. Its effectiveness is restricted by the limited dataset and the insufficient utilization of the spatial information contained in the sensor pad images.

Table 1: Performance of the method based on ΔE from the global background.

Class	C 1	C 2	C 3	C 4	C 5
Accuracy	91%	60%	43%	20%	86%

(2) To further improve the accuracy, we consider the use of the spectral reflectance of the sensor pad, then develop two different machine learning approaches: k-nearest-neighbor with sequential forward feature selection to determine the best set of features, and random forest with principal component analysis for feature reduction for classifying the level of contamination by As^{3+} into one of five categories. The accuracy of these two models is compared by implementing them with the same training and test datasets. It turns out that the RF model with PCA feature selection performs well in terms of accuracy compared to the k-NN classifier. Table 2 shows the prediction accuracy of the RF model with PCA features. The classification performance yields 86.6%

average accuracy which is higher than the baseline model. The challenging part of the spectral imaging model is that the spectral data must be obtained using an expensive and professional optical component, like a spectroradiometer [4].

Table 2: Performance of the RF model with PCA features.

Class	C 1	C 2	C 3	C 4	C 5
Accuracy	82%	80%	71%	100%	100%

Nowadays, convolutional neural networks (CNN) have gained tremendous popularity in computer vision, especially in the image classification domain for better performance than popular image processing methods [5], [6]. Deep learning algorithms yield high classification accuracy by using large, annotated datasets of images. Therefore, to develop accurate image classifiers for the contamination-levels classification task, we need a large dataset of images of colorimetric responses. However, obtaining large-scale datasets of detection images of contamination levels is challenging because of limited test samples.

One approach to overcome this challenge is to use data augmentation, a standard procedure to obtain good performance by deploying rotation, flip, translation, and scaling techniques. Another emerging deep learning generative model inspired by game theory to synthesize images is the Generative Adversarial Network (GAN) [7]. The GAN model consists of a generator to create fake images and a discriminator to distinguish between the real and fake images. These two parts are trained in an adversarial process. Different variations of the classic GAN models have been proposed. As a representative example, pix2pix is a GAN model addressing image-to-image translation problems [8]. Recent medical and biological imaging applications have shown that the GAN framework can successfully generate images and obtain reasonable performance [9], [10], [11]. Therefore, it is appropriate to apply a GAN model to generate synthetic images for training purposes.

We aim to solve the classification problem posed by a small scale of data samples and large intra-class variance. In this paper, we propose an approach to generate high-quality colorimetric responses from our detection images captured by a phone camera and apply a CNN based on EfficientNet-B0 [12] for the contamination-levels multi-classification. The proposed method is evaluated on five contamination levels, and is compared with our previous work. We hope that the proposed methods can be a strong candidate for phone-based contamination-levels detection. Because the user need only take an image of the test response using their phone camera and feed the captured image into the proposed model, the model can automatically classify the test sample's contamination level.

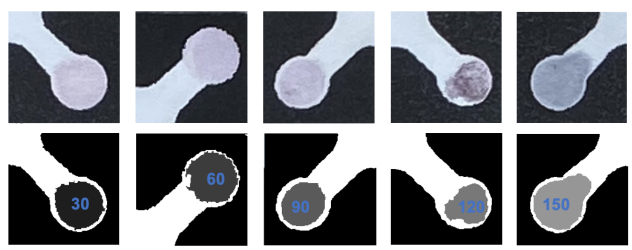
The rest of the paper is organized as follows. In Section 2, we present the dataset, explore two methods for generating realistic synthetic images, and evaluate the classification results achieved by the proposed CNN classifier. Section 3 reports the experimental results for classifying contamination levels. In Section 4, the conclusions are given.

Methodology

The main challenges of our project are the small scale of available data samples and the large intra-class variance. To overcome these challenges, we first use traditional data augmentation techniques to enlarge the training dataset of the colorimetric signals (AUG data), then train the proposed CNN model with these training sets, and test with the real test dataset. Finally, we synthesize realistic images using pix2pix (GAN data), and observe the classification accuracy after adding the GAN data to the AUG training set.

Dataset description

In this study, the colorimetric responses of 5 contamination levels (As^{3+}) are used as the experimental data. Our optical system first acquires the colorimetric signals of the biosensors. The optical system mainly consists of a photo studio booth (Amzdeal, purchased from Amazon.com) for providing the controlled D65 illumination environment, a mobile phone camera (iPhone 11 Pro Max, CA, USA), and a fixture to hold the mobile phone. Next, we extract the regions of interest and obtain the corresponding segmentation masks, as illustrated in Fig. 2. Through the above-mentioned steps, our dataset consists of 126 phone captured images: 35 in Class 1 (0, 1, 2 ppm), 32 in Class 2 (4, 5 ppm), 22 in Class 3 (10 ppm), 15 in Class 4 (20, 30 ppm), and 22 in Class 5 (50 ppm). All the ROIs are resized to a uniform dimension of 200 × 200 pixels. Finally, we divide the original dataset into a training set, a validation set, and a test set according to the ratio 5: 2: 3, as shown in Table 3.



(a) Class 1 (b) Class 2 (c) Class 3 (d) Class 4 (e) Class 5

Figure 2. *Dataset examples of 5 contamination levels: the ROI images of the colorimetric signal and the corresponding segmentation masks (The numbers in blue are the grayscale values. To distinguish different classes, we use different grayscale values to label different classes' response areas).*

Table 3: Overview of the small-scale dataset showing the division, respectively, into training, validation, and test sets.

Class	Training set	Validation set	Test set	Total
Class 1	17	7	11	35
Class 2	16	6	10	32
Class 3	11	4	7	22
Class 4	7	3	5	15
Class 5	11	4	7	22

Traditional data augmentation

Deep learning algorithms yield high classification accuracy by using large, annotated datasets of images to train a network. This can cause a danger of overfitting when a deep network deals with a limited numbers of training images. One standard method to address this problem uses traditional data augmentation methods. Classic data augmentation techniques include scaling, cropping, flipping, rotation, translation, and other deformations. The color of the image is important to our application, and we aim to train a classifier to predict the unknown test image's contaminant level based on its colorimetric signal. Therefore, we choose the rotation, flipping, and shifting data augmentation methods, and avoid color deformation.

Generative adversarial network

Another promising tool to generate synthetic images is the Generative Adversarial Net (GAN). The GAN model consists of a discriminator D to discriminate between the real and fake images, and a generator G generating fake images to fool the discriminator. These two parts are trained in an adversarial process. Recent studies have shown that the GAN framework can successfully generate images and obtain good performance [13]. Inspired by [10], we explore pix2pix, a variant of conditional GANs for learning the translation from the binary segmentation images to the colorimetric signal images. The loss function is shown in (1).

$$G^* = \arg\min_G \max_D L_{cGAN}(G,D) + \lambda L_{L1}(G) \tag{1}$$

Here the generator G tries to minimize this loss function, whereas the discriminator D tries to maximize it, λ is the hyperparameter that balances the $L1$ loss term, which is used to obtain sharp images. One of the limitations of pix2pix is that it requires paired images to train the network. For our application, the input paired images to train the network are the ROI images of the colorimetric signal and the corresponding segmentation masks, as illustrated in Fig. 2. Then, we only feed the segmentation masks to the trained pix2pix network to generate realistic colorimetric signals.

Proposed CNN architecture

EfficientNet models are based on uniformly scaling the network width, depth, and resolution to yield higher test accuracy and better efficiency with a smaller number of parameters than previous ConvNets, like RestNet-50, and Inception-v2 [14], [15]. EfficientNet consists of a series of models from B0 to B7, and the number of parameters varies from 5×10^6 to 66×10^6 [12]. Considering that we focus on a phone-based contamination detection application, we use EfficentNet-B0 with the least number of parameters for transfer learning and to extract features of the generated detection images. To perform the five-contamination-levels classification task, we add a sequence of two fully connected layers with batch normalization, RELU activation functions, and a dropout layer. Finally, the classification layer contains five output units for 5-class classification based on using the softmax activation function. To test the performance of our proposed CNN classifier, we use accuracy, precision, and F-1 score as the evaluation metrics.

Experiments and results

To solve the multi-class classification problem posed by the small scale of our dataset and its large intra-class variance, we explore two kinds of data augmentation techniques, and compare their effectiveness for classifying the five contamination levels. The experiments are set up as follows:

(1) We enlarge our training dataset (AUG training data) by using traditional data augmentation, then calculate the test accuracy of our proposed CNN model trained with the different numbers of the AUG data.

(2) To compare the classification effects between the traditional data augmentation method and pix2pix, we use the AUG training dataset that yields the highest test accuracy to train pix2pix. Then we input the specific segmentation masks to pix2pix to generate the corresponding realistic colorimetric signal images.

Traditional data augmentation evaluation

According to the proposed approach, we first apply rotation, flipping, and shifting to produce a large number of images for the training and validation datasets. Here, $N_{rotation} = 70$, $N_{flip} = 2$, and $N_{trans} = 24$. Inspired by [16], we randomly sample the augmented images to additively form the different training dataset groups $D_{train1} \subset D_{train2} \subset ... \subset D_{train8}$ such that D_{train1} only consists of the original training dataset, each class of D_{train2} includes 1,000 samples, ..., and each class of D_{train8} includes 7,000 images. Then, we randomly select 1,000 images per class for the validation dataset.

We train the proposed CNN classifier separately for each set of the training groups and evaluate the test results on the same original test dataset. The accuracy results for five contamination levels with the increasing training datasets are illustrated in Fig. 3. It shows that the classification results improve from 61.9% with no AUG data to 88.1% (D_{train4}, the optimal AUG data group). We also notice that after D_{train4}, the classification results drop down slightly and continue to fluctuate around 80%. Table 4 presents the confusion matrix for the optimal AUG training data group D_{train4}. The classification performance using only classic data augmentation (D_{train4}) yields 91.92% average precision, 86.6% average recall, and 86.7% average F-1 score.

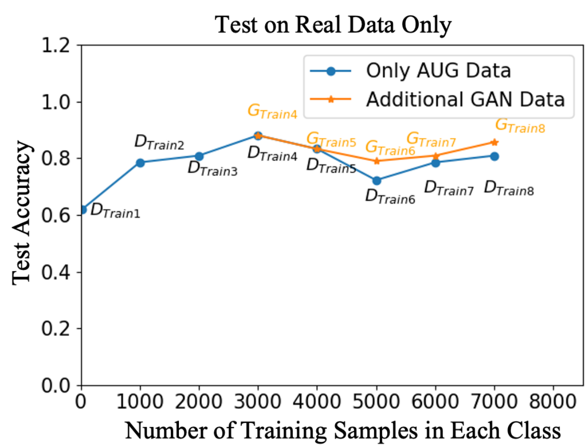


Figure 3. Classification results for the five-classes test data as a function of the training set size.

Pix2pix data augmentation evaluation

Researchers have reported that the augmented images produced by the traditional data augmentation approach are highly correlated; and GANs are a promising approach to generate a large, diversified dataset of images for training purposes [10], [16]. So, we use the optimal AUG data group D_{train4} to train pix2pix and input the segmentation masks of the well-trained

Table 4: Confusion matrix for the CNN model trained with the D_{train4} group.

Ground	Predicted Class						
Truth	C1	C2	C3	C4	C5	Precision	Recall
C1	11	0	0	0	0	84.6%	100%
C2	1	9	0	0	0	75%	90%
C3	1	3	3	0	0	100 %	42.9%
C4	0	0	0	5	0	100%	100%
C5	0	0	0	0	7	100%	100%

pix2pix to generate a realistic colorimetric images dataset. Here, we also additively form the synthetic training group datasets (GAN data) $D_{train4} \subset G_{train5} \subset G_{train6} \subset G_{train7} \subset G_{train8}$. To avoid the influence of the edges in the training group on the classification, the segmentation masks of G_{train5} and D_{train5} are the same, and this requirement applies to the rest of the training groups, i.e. G_{train6} and D_{train6} are the same, and so on.

Fig. 4 shows the synthesized high-quality colorimetric images of five-classes with the well-trained pix2pix. Fig. 3 shows

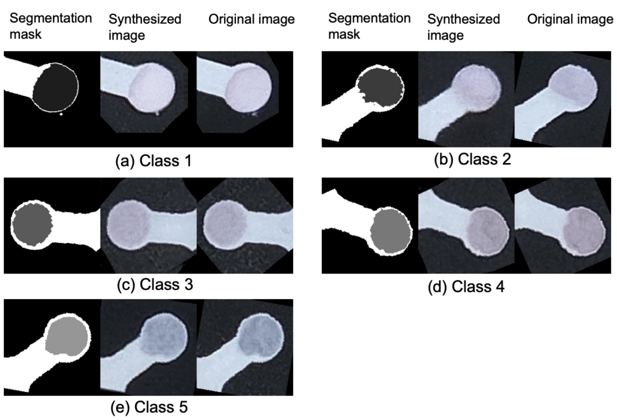


Figure 4. *Synthesized examples of five-classes with pix2pix. For each class, the figures from left to right are a segmentation mask, a synthesized image generated by a well-trained pix2pix, and the colorimetric signal's original image.*

the test accuracy of the GAN-based synthetic augmentation experiments. Even though the highest classification result is still obtained with D_{train4}, adding the synthetic training data does improve the accuracy for the training sets G_{train6}, G_{train7}, and G_{train8}, compared with the same training set number in the AUG training group. It also reduces the fluctuation in accuracy.

Conclusion

In this paper, we focus on solving the multi-class classification problem posed by a small scale dataset and large intra-class variance. We propose a CNN classifier and explore two kinds of data augmentation techniques to compare their effectiveness for a classification task. Moreover, we conclude that this proposed approach demonstrates promising results for a contamination-levels classification task with limited data.

Acknowledgments

This manuscript is based upon work supported by the U.S. Department of Agriculture, Agricultural Research Service, under Agreement No. 59-8072-6-001. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.

References

[1] S. Diaz Amaya, M. Zhao, J. Allebach, G.T.C Chiu, and L. Stanciu, "Ionic strength influences on biofunctional Au-decorated microparticles for enhanced performance in multiplexed colorimetric sensors", *ACS Applied Materials & Interfaces*, p. 32397, July 2020.

[2] M. Zhao, S. Diaz Amaya, L.K. Lin, S.A. Jin, L.K. Lin, A.J. Deering, L. Stanciu, G.T.C Chiu, and J. Allebach, "Image Analytics for Food Safety", *Proceedings of Imaging and Multimedia Analytics in a Web and Mobile World 2020 (part of Electronic Imaging 2020)*, p. 302, January 2020.

[3] M. Zhao, R. Zhang, S. Diaz Amaya, S.A. Jin, L.K. Lin, A.J. Deering, L. Stanciu, G.T.C Chiu, and J. Allebach, "Detection, imaging, and quantification of DNA-based pathogen based on inkjet-printed test strips", *NIP & Digital Fabrication Conference*, p. 177, Sepetember 2019.

[4] M. Zhao, S. Diaz Amaya, S.A. Jin, L.K. Lin, A.J. Deering, L. Stanciu, G.T.C Chiu, and J. Allebach, "Spectral Imaging to Differentiate Contaminant Levels", *Proceedings of Displaying, Processing, Hardcopy, and Applications 2021 (part of Electronic Imaging 2021)*, p. 342, January 2021.

[5] S.P. Mohanty, D.P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection", *Frontiers in Plant Science*, p. 1419, September 2016.

[6] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks", *Communications of the ACM*, p. 84, May 2017.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", *Advances in Neural Information Processing Systems*, p. 2672, December 2014.

[8] P. Isola, J.Y Zhu, T. Zhou, and A.A. Efros, "Image-to-image translation with conditional adversarial networks", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 1125, July 2017.

[9] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification", *Neurocomputing*, p. 321, December 2018.

[10] H. Hu, Y. Lei, D. Xin, V. Shkolnikov, S. Barcelo, J.P. Allebach, and E.J. Delp, "2D label free microscopy imaging analysis using machine learning", *Proceedings of Computational Imaging XVIII (part of Electronic Imaging 2020)*, p. 341, January, 2020.

[11] Y. Liu, Y. Zhou, X. Liu, F. Dong, C. Wang, and Z. Wang, "Wasserstein gan-based small-sample augmentation for new-generation artificial intelligence: a case study of cancer-staging data in biology", *Engineering*, p. 156, February 2019.

[12] M. Tan and Q.V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks", *arXiv preprint arXiv:1905.11946*, May 2019.

[13] P. Costa, A. Galdran, M.I. Meyer, M.D. Abràmoff, M. Niemeijer, A.M. Mendonça, and A. Campilho, "Towards adversarial retinal im-

age synthesis", *arXiv preprint arXiv:1701.08974*, January 2017.

[14] K. He, X. Zhang, S. Ren, and J. Sun, 2016. "Deep residual learning for image recognition", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 770, June 2016.

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 2818, June 2016.

[16] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification", *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, p. 289, April 2018.

Author Biography

Min Zhao *is a graduate student in Electrical and Computer Engineering Department at Purdue University under the supervision of Professor Jan P. Allebach. She received her B.S. degree in Mechanical Engineering in 2010 from Qingdao Technology University, Shandong, China. Her current research interests include image analysis, computer vision, machine learning, and deep learning.*

Susana Diaz-Amaya*, Ph.D. in materials engineering from Purdue University with a solid background in microbiology from the Pontifical Javeriana University, Colombia. Susana started her industrial career in 2011, as leader of R&D and independent consultant for business acceleration. Currently, she serves Bayer Crop Science as the head of the innovation site "Bayer at Convergence".*

Amanda Deering *is a clinical associate professor of Food Science at Purdue University. Her research focuses on examining the internalization of human pathogenic bacteria in plants, and routes of contamination that can contribute to plants harboring pathogenic bacteria. Amanda also works closely with industry to develop and test novel sanitization treatments that can be used for fresh produce, and works directly with Indiana fruit and vegetable growers to address food safety issues on the farm.*

Lia Stanciu *is a professor of Materials Engineering at Purdue University. She earned her Ph.D. in Materials Science from U.C. Davis in 2003. In 2005, she joined Purdue's School of Materials Engineering as a faculty member. Her main research areas include advanced biomaterials for implantation, materials for biosensing devices and nanotechnology.*

George T. Chiu *is a Professor of Mechanical Engineering with courtesy appointments in Electrical and Computer Engineering and Psychological Sciences at Purdue University. He received the B.S. degree from National Taiwan University and the M.S. and Ph.D. degrees from University of California at Berkeley. His research interests are mechatronics and control with applications to digital printing and imaging systems, digital fabrications and functional printing. He is a Fellow of ASME and IS&T.*

Jan P. Allebach *is Hewlett-Packard Distinguished Professor of Electrical and Computer Engineering at Purdue University. Allebach was named Electronic Imaging Scientist of the Year by IS&T and SPIE, and was named Honorary Member of IS&T, the highest award that IS&T bestows. He has received the IEEE Daniel E. Noble Award, the IS&T/OSA Edwin Land Medal, the IS&T Johann Gutenberg Prize, is a Fellow of the National Academy of Inventors, and is a member of the National Academy of Engineering.*