https://doi.org/10.2352/ISSN.2169-4451.2019.35.138

Image Based Quality Assurance of Fabricated Nitrate Sensor*

Qingyu Yang, Yang Yan, Kerry Maize, Xin Jin, Hongjie Jiang, Muhammad Ashraful Alam, Babak Ziaie, George Chiu, Ali Shakouri, Jan P. Allebach
Purdue University, West Lafayette, Indiana 47906, U.S.A.

Abstract

Ground-based nitrate sensors have great potential in agriculture to monitor soil conditions in real time. One path to scalable mass production of inexpensive potentiometric nitrate sensors is reel-to-reel slot-die deposition of ion-selection membranes on screen-printed electrodes. However, this process produces membranes with nonuniform thickness and texture that affects sensor performance. Manually monitoring sensor quality during fabrication costs many hours and human resources. So, we developed a scalable quality assurance method that establishes the relationship between sensor performance and the captured sensor images. The relationship will help us to monitor sensor performance only based on the sensor images. It will reduce the cost of measurement. To accomplish this, we apply both traditional and deep learning techniques for sensor image processing and regression. The traditional approaches are used to detect the useful regions of sensor images. Then we use Convolutional Neural Networks (CNNs) to combine images of the sensor membrane with sensor performance metrics to rapidly predict sensor quality. Successful prediction based on noncontact imaging will help to better control the fabrication process.

Introduction

The scalable Manufacturing of Aware and Responsive Thin Films (SMART) consortium [1] is developing economical printing and manufacturing methods to meet the needs for new low-cost consumer products in health monitoring, agriculture, electronics, architecture, and other fields. In this paper, we focus on inexpensive mass-produced potentiometric nitrate sensors. Its development includes roll-to-roll systems [2], the functional printing, and ion-selective membrane [3] techniques to reduce the cost and accurately measure nitrate level in farm soils in real time.

The electrodes of our nitrate sensor are printed on a polyethylene terephthalate (PET) substrate, which is shown in Fig. 1 (a). Fig. 1 (b) shows the components of the nitrate sensor: a nitrate ion selection membrane layer, a silicone passivation layer, and an underlying electrode layer. Also, the active region within the sensor, defined as the portion of the ion selection membrane that contacts the electrode, is the region of interest in our quality assurance analysis. It is worth mentioning that the variations in the uniformity of our sensor produced by the fabrication process are inevitable and affect the sensor performance. According to measured sensor data and a physics-based model of the sensor, we hypothesize that there exists a relationship between the nonuniformity of the membrane and the sensor performance. Therefore, studying this relationship can help us to better analyze the fabrication process.

To verify our hypothesis, we use traditional methods of image processing to locate the active region in the sensor two-dimensional (2D) image, and analyze influence of ion-selection membrane (ISM) uniformity.

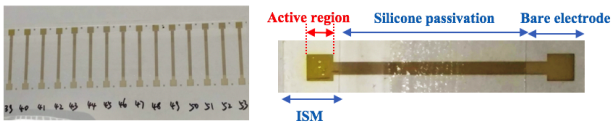

Figure 1: Example of the screen printed electrodes: (a) the screen printed electrodes on the PET substrate; (b) components of the nitrate sensor.

From a traditional perspective, we need to handcraft some features to represent the texturing level of the sensor that may help us establish the relation between the image and the sensor quality. We apply some image processing techniques, such as Normalized Gaussian Pyramid [4], Laplacian of Gaussian (LoG) [5], and Local Binary Pattern (LBP) [6], to quantify the texturing level of the active region. Then, regression is applied between the sensors' texturing level and the sensor performance metrics. However, since "texturing level" has an abstract definition, it is challenging to handcraft valid features to describe the roughness.

In recent years, CNNs, one of the most popular deep learning frameworks, have shown their strong ability in visual feature extraction [7]. Therefore, in this paper, we propose a method that uses the CNNs to extract and select useful and abstract texture features. After quantifying those features, we apply a non-linear regression between the image features and the sensor performance metrics.

Method

The system proposed in this paper consists of both traditional image processing techniques and a deep learning model. As shown in Fig. 2, image processing operations will be applied to the captured image to crop the active-region portion of that image. The active-region image will be fed to CNNs which return a predicted performance metric.

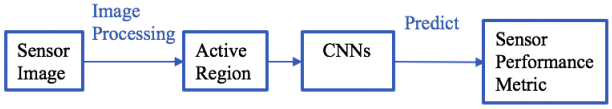

Figure 2: System of prediction with image processing and the CNNs

*Research supported by the SMART Films consortium (https://engineering.purdue.edu/SMART-consortium).

It is worth mentioning the CNNs in Fig. 2 need to be trained. As is the case with many other data-driven models, our networks also need much data for training. Before the training phase, we prepare the training dataset. As shown in Fig. 3, we capture the active region and measure the performance metric for each sensor for training. This dataset was used to update the networks. Once the model has been updated, we use it for prediction.

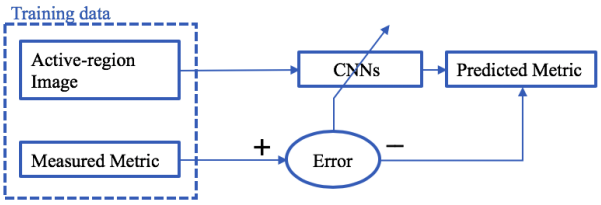

Figure 3: Flowchart of the method of updating the CNNs.

Active Region Capture

To present the image of the sensor, an imaging system is set up before the measurement of sensor performance to obtain a 2D image of the region of interest of the sensor. As shown in Fig. 4, the imaging system consists of a light source, a 0.5 × magnification telecentric lens, and a silicon wafer. We turn on the light source and put the PET substrate with the nitrate sensor on the silicon wafer. Then, the distance between the lens and the substrate is manually adjusted to focus the camera on the active region of the sensor. The camera captures and records each sensor as a 2D color images with 1280 × 1024 pixels. From the captured images, the variations in roughness can be distinguished by visual perception.

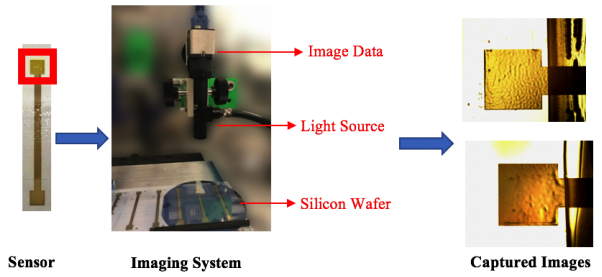

Figure 4: Imaging system for capturing sensor images.

Because we want to consider only the active area of each sensor, we use a traditional method of edge detection to locate and crop the region of interest. The Sobel filter [8] is used to get the gradient of the image intensity. After applying the Sobel filter, the edge will be clearer in the gradient image. Then, we apply Otsu's algorithm [9], which separates image pixels into two classes: foreground and background. After getting the intensity information, Otsu's algorithm and a morphological transformation can be applied to get a binary image with white and black pixels corresponding to the active region and the background, respectively. However, only the large rectangle region is treated as the active region. Therefore, we generate the maximum inscribed rectangle of the active region and crop it as the final image. All input images are cropped to a rectangle as shown in the right image of Fig. 5. The input images do not contain any background information.

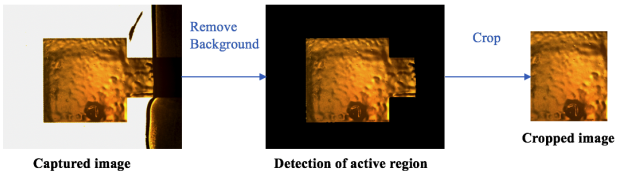

Figure 5: Example of detecting the active region within the captured sensor image.

Sensor Performance Metric Generation

The data of each nitrate sensor performance is generated by measuring the difference between the potential voltages of the membrane and reference sensor in a specific nitrate solution. Fig. 6 (a) shows the measurement process. This measurement records the results at 50 intervals over 24 hours or more. Fig. 6 (b) shows a plot of the measured sensor performance with respect to time. The measured nitrate sensor voltage consists of two characteristic response windows, a transient response and a saturation response. As the measurement time increases, the measured sensor voltage tends to saturate at a constant value after around 5 hours, and becomes more and more stable. So, we select the potential voltage at the end time, which is at 24 hours, as the metric of the ground truth. To prepare a consistent dataset, all sensors were measured in a 0.1 molar nitrate solution in this dataset. Also, it is clear that there are 4 sensors with abnormal performance shown in Fig. 6 (b). We assume that the data with an abnormal behavior is from the measurement process. Thus, the sensors with abnormal performance will not be included in our dataset.

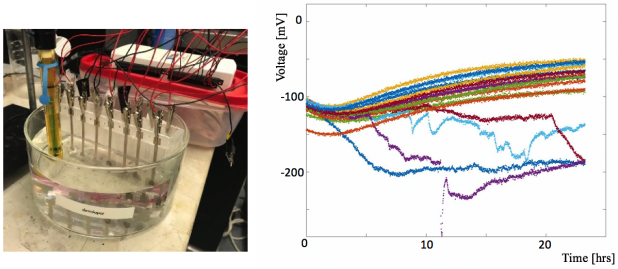

Figure 6: Measurement of sensor performance for 16 sensors in 24 hours: (a) measurement process; (b) recorded results.

Convolutional Neural Networks

CNNs have proven their abilities in high-level feature extraction and regression tasks. As we mentioned above, CNNs are data-driven models that need a large amount of data and ground truth values to tune the parameters in the model [10]. However, the number of sensors in our dataset is limited. To deal this problem, we use the pre-trained VGG-16 model [11]. As the pre-trained model already has powerful ability to extract image features, the initial CNNs' parameters with pre-trained weights not only requires less training data, but also can save a lot of training time. In our case, the VGG-16 was trained on ImageNet [12]. It already can extract visual features from 1000 kinds of objects. Fig. 7 shows the architecture of VGG-16. VGG-16 model consists of 13 convolutional layers and 3 regression layers. The first 13 convolutional layers are used to extract the high-level features; and the remaining 3 regression layers are doing non-linear re-

gression. The convolutional layers use convolution and maximum pooling to map each 2D sensor image to a 25,088-dimension feature space. To complete the regression task, three layers perform a non-linear regression that regresses 25,088 features, which is $7 \times 7 \times 512$ features from the convolutional layers, into one number: the voltage measured by each sensor. Each layer includes a linear fully connected layer, an ReLU function [13], and a "Dropout" process [14]. Fig. 8 shows how two fully connected layers are working. The fully connected layer means each node of the output depends on all nodes of the input. From a mathematical perspective, as shown in Eq. 1, each element of the fully connected layer's output, which represented as $\hat{x}_j$, is a linear combination of all the input elements x_i. The fully connected layer is doing a linear transformation where ω_{mn} are the weights and b_n represent the bias. Following this linear transformation, the ReLU, which not explicitly shown in Fig. 8, can be expressed as Eq. 2. It only passes the positive output elements. Since ReLU is a non-linear function, the combination of ReLU and linear transformation is also a non-linear transformation. In addition, Dropout is an random process that drops half the elements from each layer randomly. As shown in Fig. 8, the dash line means the connection was been dropped. In other words, the networks randomly selected half the input nodes to fully connect to all the output nodes. This process reduces the number of parameters and can be used as regularization to prevent the overfitting.

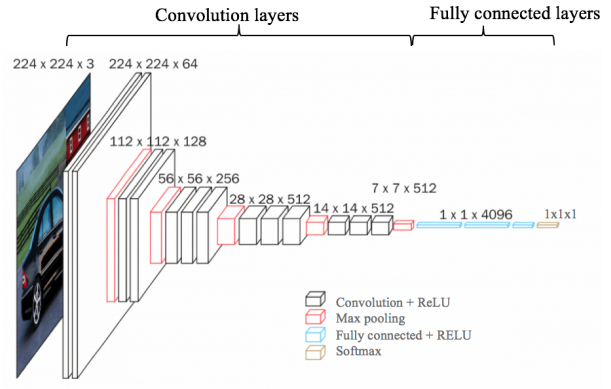

Figure 7: Architecture of the CNNs with VGG-16.

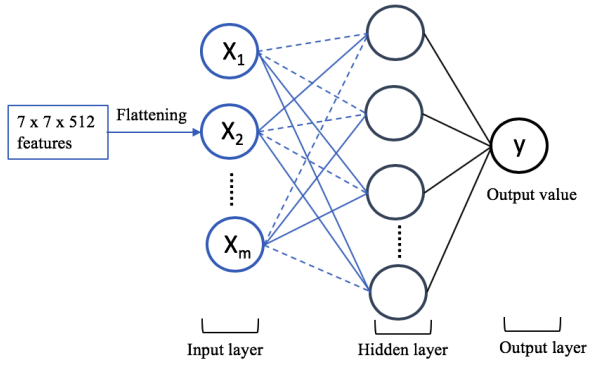

Figure 8: Linear connection algorithm in fully connected layers.

$$\begin{aligned}
\hat{x}_1 &= x_1 \times \omega_{11} + x_2 \times \omega_{21} + \ldots + x_m \times \omega_{m1} + b_1 \\
\hat{x}_2 &= x_1 \times \omega_{12} + x_2 \times \omega_{22} + \ldots + x_m \times \omega_{m2} + b_2 \\
&\vdots \\
\hat{x}_n &= x_1 \times \omega_{1n} + x_2 \times \omega_{2n} + \ldots + x_m \times \omega_{mn} + b_n
\end{aligned} \tag{1}$$

$$y = f(\hat{x}) = \hat{x}^+ = max(0, \hat{x}) \tag{2}$$

We define the loss to be the absolute difference between regression result and the ground truth in the prepared dataset. The loss value will be backpropogate to each parameter in the CNNs [15]. The optimizer will tune each parameter to reduce the loss. Once the training loss converges, that means our networks can extract the useful high-level features and find the non-linear regression function relating the captured sensor image features and measured sensor voltage.

Experiments and Results

Following the above methods of dataset preparation, we collect 71 active-region images with the corresponding metric of the sensor performance. Also, we randomly select 51 images for training, 10 images for validation, and 10 images for testing. For a fair comparison, we also do an experiment that creates the prediction model by a traditional machine learning approach on the same training, validation, and testing datasets. In contrast to the CNNs, the traditional machine learning approach only uses the hand-crafted features of the membrane roughness to correlate with the sensor performance metric. The experiments will use the training dataset to train the prediction models. Then, the validation set will be fed to models so that we can estimate the performances of these models on testing set. Especially for CNNs, we need to use the performance from validation set to tune the parameters so that the overfitting problem can be prevented. The final comparison will be operated on testing set.

Traditional Machine Learning Approach for Comparison: Linear Regression Model

The traditional machine learning approach used in this paper is a combination of Local Binary Pattern (LBP) and the Least Square (LS) regression model [16]. The LBP, which is a descriptor that is invariant to rotation and illumination of the grayscale image, is one of the most widely used texturing descriptor in the computer vision area. As shown in Fig. 9, we first convert the 2D active-region images to grayscale. Then, we apply a Gaussian filter to reduce the noise and smooth the grayscale image. The LBP descriptor can be applied to the smoothed images. Each pixel in the smoothed image will be set to a decimal value based on a comparison between the anchor pixel and the surrounding pixels. In our case, eight surrounding pixels are chose within a circle centered at the anchor pixel with radius, three pixels. After conversion, the 2D image, which stores the converted decimal values, contains the texturing information of the original image. Fig. 9 (c) shows an example visualizing the 2D LBP array. Then, the one-dimensional (1D) LBP array can be constructed from the 2D LBP array by its histogram. The number of bins in the histogram depends on the number of prototypes of patterns, which is

decided by the number of surrounding pixels. Due to the nonuniform patterns in the active-region images, we select the number of prototypes to be twelve. The counts of the twelve prototypes is constructed to the elements of LBP array, which is the input to the LS regression model.

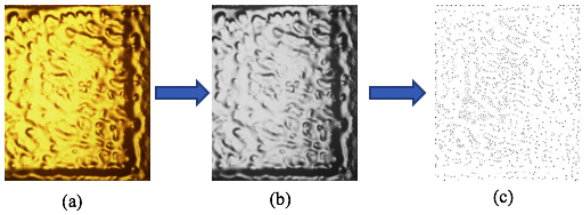

Figure 9: Example showing the visualization of the 2D LBP array of an active-region image: (a) original RGB active-region image; (b) grayscale rendering of (a); (c) LBP array of (b), where the decimal values have been stretched to the range 0 to 255, and inverted, so white corresponds to decimal value 0 and black corresponds to the maximum decimal value of 255.

As shown in Fig. 10, once the hand-crafted features, in this case, the LBP array values, are extracted, we can feed them to the linear regression model to get the predicted performance metric. To generate the model, we use least squares regression to set up a linear relationship between the hand-crafted features to and the predicted sensor performance. In a 1D space, least square regression finds a line that best fits the input points. In our case, the training process finds a hyperplane that best fits the training dataset. Here, the loss function is defined as the absolute difference between the prediction and the ground truth.

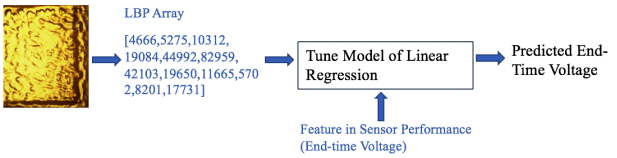

Figure 10: System of prediction with image processing and the traditional machine learning approach.

Deep Learning Approach: CNNs

In contrast to the training process of traditional machine learning, the deep learning approach uses the 2D active-region images to be the inputs directly. All input images are resized to 224 $\times$ 224 pixels and fed to the CNNs. Also, training networks requires a graphics processing unit (GPU) for fast computation. As the development of deep learning, many useful deep learning platforms and libraries are released so that people can manipulate operations to networks in the convenient way. In this paper, we use Pytorch [17], which contains many useful high-level application programming interfaces, as our platform.

In addition, in the training process, the hyper-parameters need to be specified to best fit our dataset to achieve better performance, and reduce the time cost in the training process. Since the initial convolutional layers use a pre-trained model and the dataset is not large, a small learning rate 0.00001 is used in the training process. Also, the number of epochs in the training process is 5000 to guarantee that the training process has converged and that the validation loss remains stable. After 5000 epoches, we test our trained model on the test dataset.

Results and Comparisons

Our experiments create two different prediction models: the linear regression model and the CNNs. These two models are based on the 51 training images with the corresponding metric of sensor performance. We use the 10 testing images in our dataset to test the performance of the two prediction models. The value of the end-time potential voltages in the sensor performance is predicted through the testing process based on the testing images. To evaluate the performance of each model, we obtain the values of loss in the training, validation, and testing, and calculate the accuracy of prediction. The loss function is defined by Eq. 3. N represents the number of images in the target set. In addition, the accuracy of prediction will be calculated by Eq. 4 to show the performance of our models. $\hat{V}_f$ is the predicted end-time potential voltage and V_f is the measured value, which is the ground truth. Here, N equals 10, which is the number of the testing images.

$$Loss = \frac{\sum |\hat{V}_f - V_f|}{N} \quad [mV] \tag{3}$$

$$Accuracy = 1 - \frac{\sum \frac{|\hat{V}_f - V_f|}{|V_f|}}{N} \tag{4}$$

As shown in Fig. 11, in the experiment results of the CNNs, the loss values in the training and validation sets are both converging. The training loss starts from 119.1 mV and becomes stable around 7.8 mV after 2000 epoches. The validation loss also converges from 113.5 mV to 16.2 mV. Although the validation loss is larger than the training loss, the difference is still in a reasonable range so that the model fits both the training and the validation sets. These results verify our assumption that there is a relationship between the features of the captured sensor image and the measured sensor performance. Also, this relationship can be approximated by our non-linear CNNs. Fig. 12 shows the results of our predicted end-time potential voltages and the measured ground truth. The predicted results track well with the measured end-time potential voltage. The accuracy of our prediction is 85.8%.

In addition, the traditional machine learning approach with a linear regression also creates a prediction model. Table 1 compares performance between the CNNs and the linear regression model. It can be seen that the CNNs achieve higher accuracy. This means that the networks can extract more useful high-level features than the hand-crafted features. In the other words, the LBP array in the traditional machine learning only captures limited characteristics of the sensor. Also, the table shows that the training loss of the networks is much smaller than the LS model. Thus, the non-linear regression of the CNNs better fits our dataset than the linear regression model.

Conclusion

Our results show that both the CNNs and the linear regression model can successfully predict the single feature end-time voltage of nitrate sensor performance based on active region images of sensors. It verifies the assumption that a relationship exists

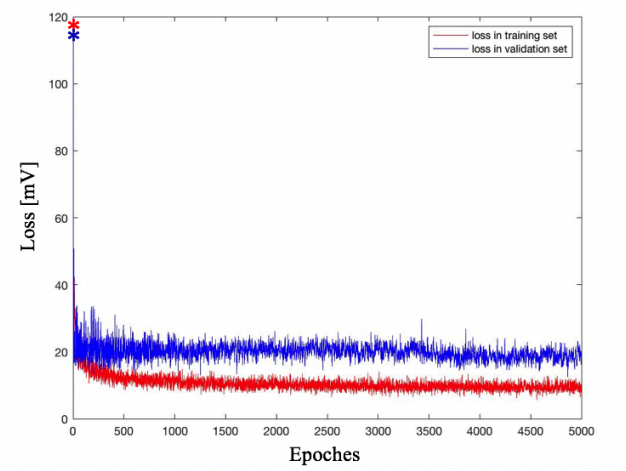

Figure 11: Loss values in each epoch in the experiment results of the CNNs (red line represents the loss values in training set; blue line represents the loss values in validation set; "*" labels the starting points of these two lines).

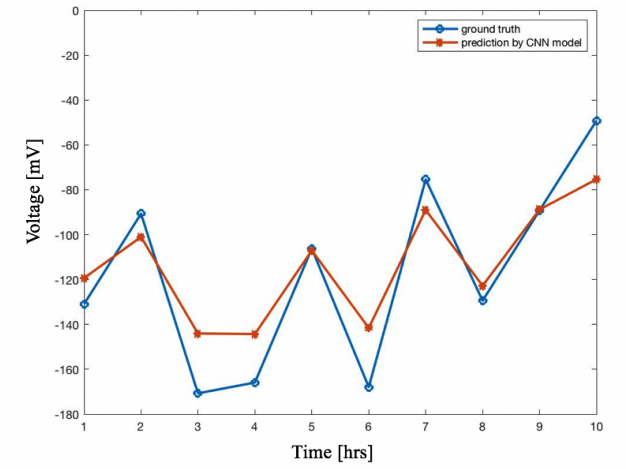

Figure 12: Prediction results for 10 testing sensors, and ground truth for comparison (red line represents the predicted values; blue line represents the ground truth values).

between nitrate sensor images and sensor performance. From the comparison, the networks achieves a higher accuracy of 85.8% than the linear regression model. The convolutional layers can extract more comprehensive features than the hand-crafted method. We conjecture that this is because the nonuniform patterns on the sensor membrane make it difficult for the local hand-crafted features to represent as much information about the physical characteristics of the sensor.

Although training the CNNs with the pre-trained VGG-16 obtains the higher accuracy of prediction, the prediction still needs to be improved. Currently, the number of the training images is insufficient to create the model. Thus, we need to increase the size of the dataset by including images from more nitrate sensors and more measurements. Also, data augmentation methods, such as changing the illumination of the images can be used to enlarge the dataset. On the other hand, the networks designed in this paper are based on the VGG-16 architecture. We also consider that it may be desirable to reduce the model complexity by reducing the number of convolutional layers to make the model better fit our dataset.

Table 1: Performance of the prediction models.

Model	Training Loss	Validation Loss	Testing Loss	Accuracy
CNNs	7.80	16.15	23.28	0.858
LS Model	28.24	31.71	26.52	0.747

References

[1] "SMART Films Consortium," Brick Nanotechnology Center, Purdue University, West Lafayette, IN. [Online]. Available: https://engineering.purdue.edu/SMART-consortium

[2] R. R. Søndergaard, M. Hösel, and F. C. Krebs, "Roll-to-roll fabrication of large area functional organic materials," *Journal of Polymer Science Part B: Polymer Physics*, vol. 51, no. 1, pp. 16–34, 2013.

[3] J. Hu, A. Stein, and P. Bühlmann, "Rational design of all-solid-state ion-selective electrodes and reference electrodes," *TrAC Trends in Analytical Chemistry*, vol. 76, pp. 102–114, 2016.

[4] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, "Pyramid methods in image processing," *RCA Engineer*, vol. 29, no. 6, pp. 33–41, 1984.

[5] A. Huertas and G. Medioni, "Detection of intensity changes with subpixel accuracy using Laplacian-Gaussian masks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 5, pp. 651–664, 1986.

[6] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 7, pp. 971–987, 2002.

[7] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1733–1740.

[8] N. Kanopoulos, N. Vasanthavada, and R. L. Baker, "Design of an image edge detection filter using the Sobel operator," *IEEE Journal of Solid-State Circuits*, vol. 23, no. 2, pp. 358–367, 1988.

[9] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.

[11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[12] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[13] L. Wu, C. Shen, and A. Van Den Hengel, "Deep linear discriminant analysis on fisher networks: A hybrid architecture for person re-identification," *Pattern Recognition*, vol. 65,

pp. 238–250, 2017.
[14] S. J. Rennie, V. Goel, and S. Thomas, "Annealed dropout training of deep networks," in *IEEE Spoken Language Technology Workshop (SLT)*, 2014, pp. 159–164.
[15] C.-X. Feng and X.-F. Wang, "Surface roughness predictive modeling: neural networks versus regression," *IIE Transactions*, vol. 35, no. 1, pp. 11–27, 2003.
[16] J. A. Nelder and R. W. Wedderburn, "Generalized linear models," *Journal of the Royal Statistical Society: Series A (General)*, vol. 135, no. 3, pp. 370–384, 1972.
[17] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," *in NIPS Autodiff Workshop: The Future of Gradient-based Machine Learning Software and Techniques*, 2017.

Author Biographies

Qingyu Yang received her B.S.(2017) in Electrical Engineering from Purdue University and currently is Ph.D. student from Purdue ECE. Her research focuses on image quality analysis, computer vision, object detection, and deep learning applications.

Yang Yan received her B.S. (2016) and M.S. (2019) in Electrical Engineering from Purdue University. Her primary area of research has been image processing, image segmentation, and image quality evaluation.

Kerry Maize received his B.S. in Electrical Engineering and Computer Science from UC Berkeley in 2002. He came to UCSC in 2005 to pursue graduate study in the areas of quantum electronics and nanoscience. His current work is focused on optical coherence tomograohy, and thermal device characterization. Prior to engineering Kerry worked in journalism.

Xin Jin is currently a Ph.D. candidate from Purdue ECE. His research interests include design and analysis of next generation electrochemical bio-sensors for IoT applications.

Hongjie Jiang joined Professor Babak Ziaie group at August 2013 and received his Doctoral degree at October 2018 from ECE Purdue. Currently he works as a postdoctoral Researcher also at ECE Purdue. His research interests include sensor and actuator, microfluidics, printed electronics, bioMEMS, etc.

Muhammad Ashraful Alam holds the Jai N. Gupta Endowed Chair professorship at Purdue University, where his research focuses on the physics and technology of semiconductor devices. From 1995 to 2003, he was with Bell Laboratories, Murray Hill, NJ working on optoelectronic integrated circuits. Since joining Purdue in 2004, Dr. Alam has published over 300 papers on biosensors, transistors, solar cells, and flexible electronics. He is a fellow of IEEE, APS, and AAAS.

Babak Ziaie is currently a professor in School of Electrical and Computer Engineering at Purdue University since Jan 2005. His research interests are related to the biomedical applications of MEMS (BioMEMS), mobile health (eHealth), and low cost flexible transducer for medical applications. He has more than 200 publications and 8 issued patents in these areas.

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 IST.

Ali Shakouri is a professor of electrical and computer engineering and director of Birck Nanotechnology Center at Purdue University. He received his PhD from Caltech. His current research is on nanoscale heat transport and electrothermal energy conversion. He has developed lock-in imaging for temperature measurement and for real-time monitoring of functional film manufacturing. He is also leading a team to manufacture low-cost smart internet of thing (IoT) devices and sensor network for applications in advanced manufacturing and agriculture.

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 IST and SPIE, and was named Honorary Member of IST, the highest award that IST bestows. He has received the IEEE Daniel E. Noble Award, the IST/OSA Edwin Land Medal, is a Fellow of the National Academy of Inventors, and is a member of the National Academy of Engineering.