A Large-Area Image Sensing and Detection System Based on Embedded Thin-Film Classifiers

Warren Rieutort-Louis, *Member, IEEE*, Tiffany Moy, *Student Member, IEEE*, Zhuo Wang, *Student Member, IEEE*, Sigurd Wagner, *Fellow, IEEE*, James C. Sturm, *Fellow, IEEE*, and Naveen Verma, *Member, IEEE*

***Abstract*—This paper presents a large-area image sensing and detection system that integrates, on glass, sensors and thin-film transistor (TFT) circuits for classifying images from sensor data. Large-area electronics (LAE) enables the formation of millions of sensors spanning physically large areas; however, to perform processing functions, thousands of sensor signals must be interfaced to CMOS ICs, posing a critical limitation to system scalability. This work presents an approach whereby image detection of shapes is performed using simple circuits in the LAE domain based on amorphous silicon (a-Si) TFTs. This reduces the interfaces to the CMOS domain. The limited computational capability of TFT circuits as well as high variability and high density of process defects affecting TFTs and sensors is overcome using a machine-learning algorithm known as error-adaptive classifier boosting (EACB) to form embedded weak classifiers. Through EACB, we show that high-dimensional sensor data from a-Si photoconductors can be reduced to a small number of weak-classifier decisions, which can then be combined in CMOS to achieve strong-classifier performance. For demonstration, a system classifying five shapes achieves performance of >85%/>95% [true-positive (tp)/true-negative (tn) rates] [near the level of an ideal software-implemented support vector machine (SVM) classifier], while the total number of signals from 36 sensors in the LAE domain is reduced by 3.5-9×.**

***Index Terms*—Amorphous silicon (a-Si), boosting, classification, image detection, machine learning, sensing, thin film, thin-film transistor (TFT), variability.**

I. Introduction

LARGE-AREA electronics (LAE) is a technology that allows us to deposit thin films of semiconductors and insulators at low temperature. This makes its processing compatible with a diverse set of materials, enabling the formation of a wide range of sensors. Just a few demonstrated examples of these are shown in Fig. 1, including gas sensors [1], light sensors, strain sensors [2], and pressure sensors [3]. In addition to this, low-temperature processing makes it possible to use substrates such as glass or plastic which can be large and flexible, spanning areas approaching square meters. A representative example is X-ray imagers, also shown in Fig. 1, which today integrate over 10 million amorphous silicon (a-Si) pixels over planes that span 0.5 m × 0.5 m.

The ability to form diverse types of sensors and the ability to integrate many of these, spatially distributed over large areas, enable systems that can interface much more extensively with the physical world. However, a key challenge in such systems is that there now exists a large amount of distributed sensor data that require processing and analysis.

Two possible options for processing the sensor data exist. First, in addition to sensors, LAE also enables the formation of thin-film transistors (TFTs), which could be used to implement processing functions. Second, all of the sensor data can be sent to CMOS ICs for processing; this results in hybrid systems based on, e.g., a-Si LAE and CMOS [4]. In practice, the first option is challenging because low-temperature processing, from which LAE derives benefits in terms of sensing, also results in low electrical performance of TFTs. Typical mobilities (μ), unity current gain frequencies (f_T), and required supply voltages are orders of magnitude worse than those of silicon CMOS transistors.

N-channel a-Si TFTs, which are the workhorse transistors used in flat-panel display applications today, have electron mobilities in the area of 1 $\mathrm{cm}^2/\mathrm{Vs}$, unity current gain frequencies in the area of 1 MHz, and threshold voltages around 1 V (compared to approximately $500\,\mathrm{cm}^2/\mathrm{Vs}$, 300 GHz, 0.3 V, for typical CMOS transistors). In addition, the TFTs exhibit high variability in parameters such as threshold voltage and mobility. This can be seen in Fig. 2, where these quantities are measured over an array of 100 devices that we have fabricated on a single glass substrate.

Given the disparity in electrical performance and reliability between TFTs and CMOS transistors, the hybrid system option is appealing. However, this now requires a potentially large number of physical interfaces between the two different technologies.

As an example, X-ray imagers commonly available today, address this by employing an active matrix, whereby pixels are read out row/column-by-row/column. This can substantially reduce the number of interfaces (by roughly a square-root factor). However, for megapixel imagers, this still implies thousands of interfaces, and further, active-matrix architectures are applicable to highly regular sensor arrangements, not the irregular arrangements over large areas being envisioned for many LAE systems [4], [5]. Thus, alternate approaches that

Manuscript received April 20, 2015; revised September 01, 2015; accepted September 29, 2015. Date of publication November 05, 2015; date of current version December 30, 2015. This paper was approved by Guest Editor Jan Genoe. This work was supported in part by a Princeton Harold W. Dodds Honorific Fellowship, NSF under Grant ECCS-1202168 and Grant CCF-1218206, and Systems on Nanoscale Information fabriCs (SONIC), one of six SRC STARnet Centers, sponsored by MARCO and DARPA.

The authors are with Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: w@rieutortlouis.com; tmoy@princeton.edu; zhuow@princeton.edu; wagner@princeton.edu; sturm@princeton.edu; nverma@princeton.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSSC.2015.2489842

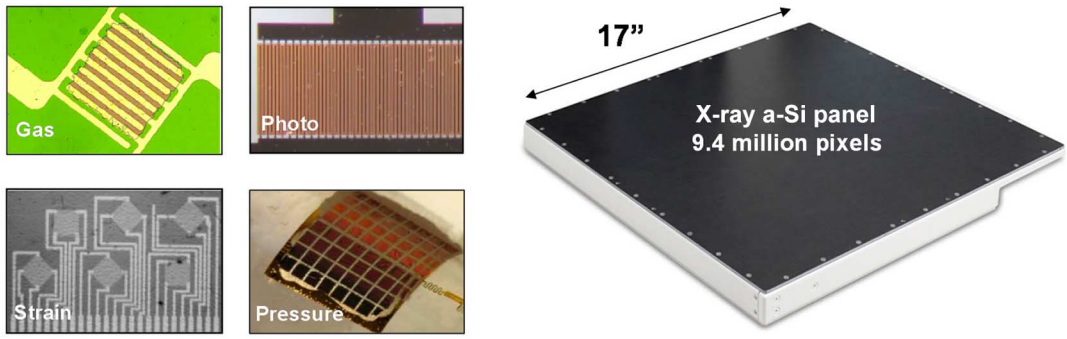

Fig. 1. Examples of thin-film sensors [1]–[3] and a large-area X-ray imager [17].

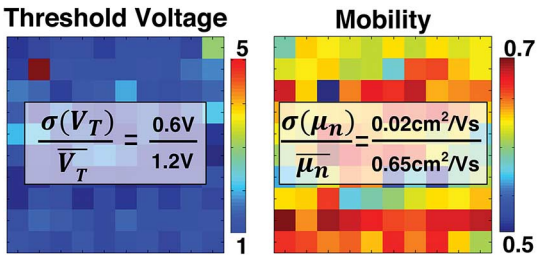

Fig. 2. Variation in mobility and threshold voltage across 100 TFTs on a single glass sample.

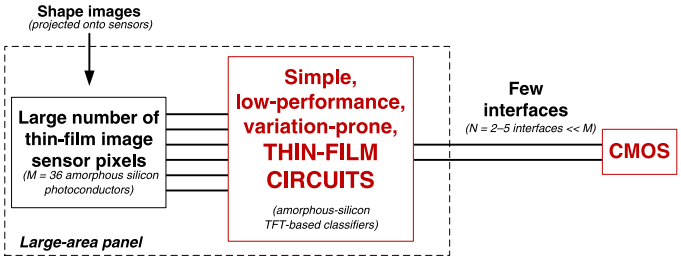

Fig. 3. Reduction in interfaces from the large-area panel is achieved using simple TFT circuits.

have the potential to greatly reduce the number of interfaces are of high interest.

In this paper, we explore an approach that addresses this interface challenge by taking data from many image-sensor pixels and processing them using simple, low-performance, variation-prone thin-film circuits [6]. This results in a reduced number of signals, which can then be transmitted from the large-area panel for image detection and classification. This approach is illustrated in Fig. 3. The driving insight is that a reduction in interfaces is achieved by transmitting only higher level information, more closely tied to the application of interest, as opposed to the raw pixel data. Systems where the raw pixel data may not be of interest include application-specific custom imagers whose role is to identify the presence of a specific sensed attribute (an object on a sensing plane, a heat zone sensed in a specific location, or a particular pressure configuration on a surface). In Section II, we identify such information, which corresponds to elemental classification decisions derived using low-performance devices. These decisions are leveraged in a machine-learning algorithm to derive a high-accuracy classification decision. In Section III, we describe the TFT circuits

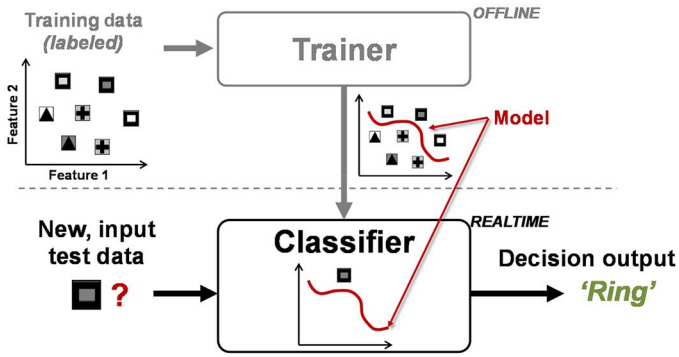

Fig. 4. Architecture of a machine-learning classification system.

that implement the simple classifier blocks required. Finally, in Section IV, we demonstrate the complete image-detection system, classifying five shape classes with a performance close to that of a support vector machine (SVM), a widely used strong classifier whose complexity precludes implementation using TFT circuits.

II. Machine Learning to Enable TFT-Based Classifiers

A number of algorithms have emerged from the domain of machine learning that enable data-driven methods for modeling and analyzing application signals. This enables the creation of models for inference from data that may be too complex to otherwise model analytically. A number of these algorithms have been previously implemented using analog circuits [7], [8]. The particular inference we focus on in this work is the classification of shapes from image-sensor signals. Using machine-learning algorithms for classification, we can employ previously observed instances of data to create a model by which classification can be performed on future instances of data.

The basic operation of such a classifier is shown in Fig. 4, consisting of two key components. The first component is a trainer, which is used to learn the classification model from previous data; this is typically performed off line. The second is the classifier itself, which uses the model to classify incoming data continuously and in real time. Fig. 4 illustrates the example of shape classification, detecting rings from all other shapes. A training set provides feature vectors along with labels of the shapes to the trainer. The trainer uses data mapped in a

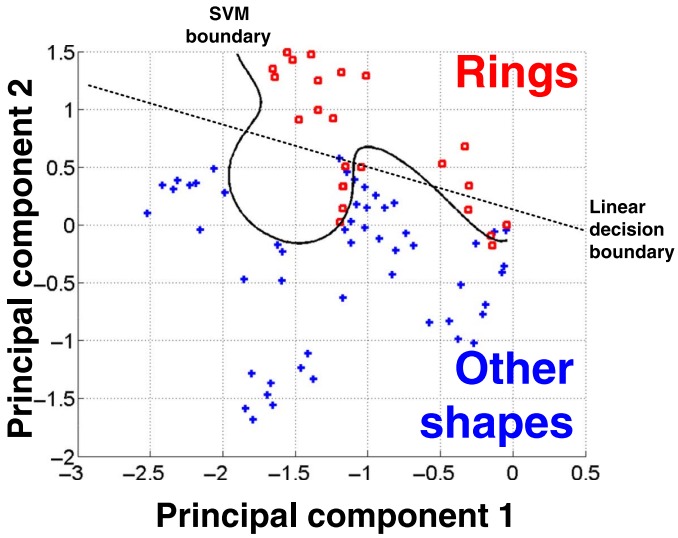

Fig. 5. Measured image data illustrates the inadequacy of linear decision boundaries, compared to more flexible boundaries such as that obtained with an SVM classifier.

feature space, along with associated labels, to form a decision boundary. The representation of this boundary, referred to as a model, is then fed to a classifier, which distinguishes rings from all other shapes from the incoming sensor data. Depending on how the data from the two classes (in this case rings vs. other shape) is distributed in the feature space, a flexible and potentially complex decision boundary may be required. This is discussed in greater detail next, using actual data from the proposed system.

A. Strong and Weak Classifiers

In machine learning, there exist the concepts of a *strong classifier* and a *weak classifier*. A strong classifier is one that can be trained to fit arbitrary data distributions, while a weak classifier is one that cannot be. Weak classifiers typically result in a high rate of errors, dependent on the precise distribution of data in the application. For this reason, weak classifiers are often inadequate, and our ultimate objective is to create a strong classifier.

To motivate this more concretely, we present first a linear classifier, which is a weak classifier. A linear classifier is implemented by applying a weight, derived from training, to each input feature and then taking a linear combination over the weighted features; mathematically this amounts to a dot-product operation between the feature vector $\vec{x}$ and a model vector $\vec{c}$, derived from training. The result of the dot product is then compared to a threshold b, in order to make classification decisions

$$y = \begin{cases} \text{rings}, & \vec{c}\cdot\vec{x} > b \\ \text{other shapes}, & \vec{c}\cdot\vec{x} \leq b. \end{cases} \quad (1)$$

Fig. 5 shows data corresponding to the actual shapes from the image-detection system demonstrated (specifically, rings vs. all other shapes). The feature vector is the 36 raw outputs from image-sensing pixels (photoconductors), onto which shape images have been illuminated. While the actual feature

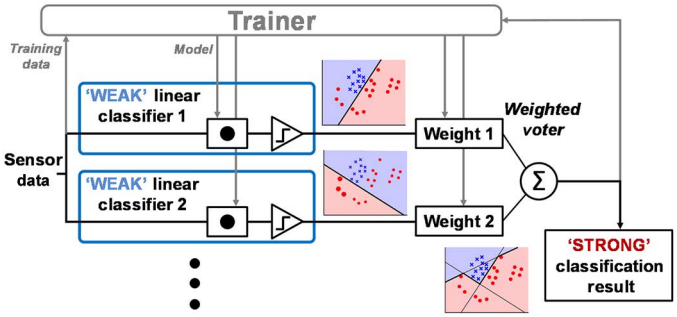

Fig. 6. Architecture of the AdaBoost algorithm.

vector dimensionality is 36, to help visualize the data, principal component analysis is performed, projecting the data vectors onto two principal components. A linear-classifier decision boundary, obtained from training, is also shown. As seen, such a boundary is inadequate, resulting in many of data points being misclassified. A more flexible decision boundary from a strong classifier is required, as obtained from training an SVM classifier with radial-basis function (RBF) kernel, also shown in Fig. 5. Unfortunately, such a classifier requires the ability to reliably perform complex computations, and thus its direct implementation is not feasible using TFT circuits.

In the following sections, the proposed approach is presented for enabling TFT implementations of the computations required for strong classification. Two ideas are exploited. The first idea is *boosting*. This addresses the problem of inadequate fitting of weak classifiers to the data distributions, by employing an ensemble of weak-classifier decisions to form a strong classifier. Specifically, we employ the boosting algorithm known as adaptive boosting (AdaBoost) [9]. The second idea is *error-adaptive classifier boosting (EACB)* [10]. This addresses the problem of high variability and nonideality in the weak-classifier implementations, by training the weak classifiers in a manner that is aware and responsive to the decision errors that these cause.

B. Adaptive Boosting

AdaBoost uses a fundamentally weak-classifier building block, such as a simple linear classifier, which cannot fit to arbitrary training data. The architecture of the algorithm is shown in Fig. 6. AdaBoost trains the weak classifiers iteratively, by emphasizing the training data instances that are incorrectly fit by previous weak classifiers. As a result, subsequent weak classifiers adaptively improve overall fitting to the training set. The final classification decision is then made by performing weighted voting over the individual classifier outputs (weighted voting is simply implemented using a signed adder). The overall result is a strong classifier, yet constructed from a set of weak classifiers, which may be substantially simpler to implement.

An important attribute of AdaBoost noted by theoretical work is that fitting to arbitrary distributions can be achieved even with extremely weak classifiers, namely, which perform only marginally better then 50/50 guessing [9]. This opens the option to use weak classifiers that have a topology viably implemented using simple TFTs circuits.

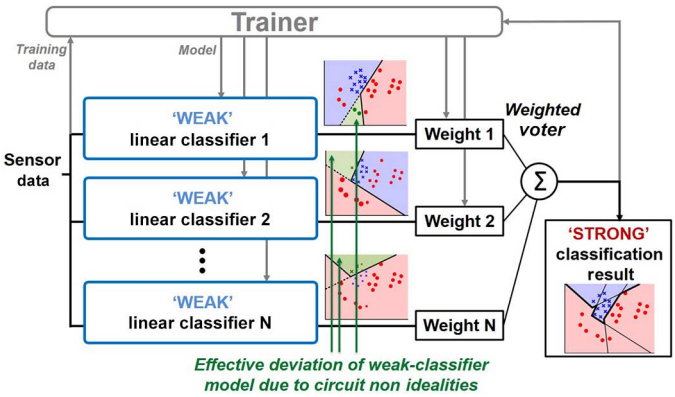

Fig. 7. Error-adaptive classifier boosting allows for the addressing of errors in preceding weak-classifier iterations that arise due to circuit nonidealities.

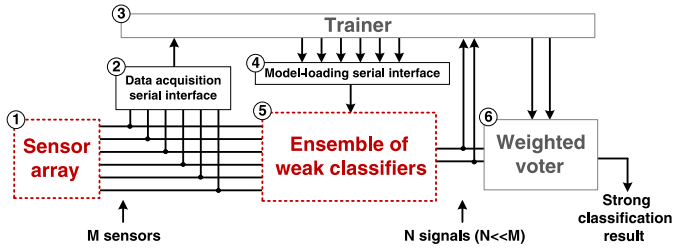

Fig. 8. Components required for the trainer and classifier.

However, TFT implementation of even simple weak-classifier circuits faces practical challenges. In addition to the fitting errors incurred with weak classifiers, implementation using TFTs will suffer from large static variations and high circuit fault rates. To overcome this, we employ EACB.

C. Error-Adaptive Classifier Boosting

Previous work has resulted in an algorithm known as EACB, whereby weak-classifier implementations that substantially deviate from their nominal behavior can be employed within the framework of AdaBoost [10]. EACB exploits the fact that, in AdaBoost, weak classifiers are trained iteratively. This means that knowledge of the errors due to circuit nonidealities in a particular instance of weak-classifier hardware can be used to adaptively train subsequent iterations of weak classifiers. By exploiting data-driven training, complex and severe errors manifesting from the nonideal implementation of previous weak-classifier iterations can thus be effectively overcome. This is shown schematically in Fig. 7, where the nominal linear decision boundaries of the weak classifiers are perturbed due to circuit nonidealities. Subsequent weak-classifier models are adapted to address the resulting misclassified points.

Previous work on EACB has resulted in two key outcomes [10]. First, it has shown that the performance achieved through this approach corresponds to the fundamental level of information that is preserved in the presence of the nonidealities; namely, the performance achieved corresponds to the mutual information between the error-affected data and the true class membership, regardless of the magnitude or rate of the errors. Second, it has shown that EACB training can be achieved efficiently using low-complexity embedded hardware (i.e., microcontroller with low MIPS and modest embedded memory [10]). Further, utilizing the error-affected data from the system itself in order to perform training, various methods can also be employed for estimating training labels, thereby achieving training entirely within the system itself.

We show in this work that EACB applied to AdaBoost enables the implementation of a high-performance classifier (achieving performance near the level of an ideal SVM) using simple, variation-prone TFT circuits for the weak classifiers. This leaves only the weighted voter requiring error-free computation, as previous work has shown that final classification performance is sensitive to voter errors [10]. Section III describes the implementation and analysis for an image-detection system, which thus employs TFT-based weak classifiers to substantially reduce the number of interfaces from the LAE domain, with the weighted voter (signed adder) presumed to be implemented in the CMOS domain.

III. Thin-Film Classifier Implementation

Having outlined the algorithmic approach used to create a strong classifier from simple, variation-prone TFT circuits, this section describes the implementation. Fig. 8 illustrates the following components required in a complete system:

1) a sensor array (image-sensing pixels);
2) a serial interface for acquiring training data;
3) a trainer for constructing the classification model from training data;
4) a serial interface for loading the models from training into the weak classifiers;
5) the weak classifiers themselves for applying the models in order to derive image-detection decisions on the sensor data in real time;
6) a weighted voter for deriving the final image-detection decision.

Among the components, the sensor array [labeled (1)] and the classifier [labeled (5)] operate continuously, and are thus particularly critical. The trainer [labeled (3)] operates infrequently (one time in our demonstration) and is presumed to be implemented on a CMOS IC, using an algorithm similar to that previously demonstrated [10]. Both the training-set acquisition interface and the model-loading interface [labeled (2) and (4)] require thin-film scanning circuits to control switches for multiplexing data onto a serial interface. The scanning circuits can be simple digital circuits and can operate at low speed, since training occurs infrequently. Similar circuits are routinely incorporated for row-scanning in active-matrix systems. As an example, scanning circuits previously presented can be used for this purpose [11]. For the current system demonstration, these circuits are not included. Finally, having reduced the interfaces from the sensor-array signals to a small number of weak-classifier decisions, the weighted voter [labeled (6)] is presumed to be implemented on a CMOS IC,

using a signed adder (since its errors impact overall performance [10]). Thus, the focus of the system demonstration is the sensor array and ensemble of weak classifiers, which operate continuously.

A. Weak-Classifier Implementation

The weak classifiers are implemented using TFT circuits. In particular, each weak classifier implements an approximation to a linear classifier, represented by the function previously shown in (1). The dot-product operation required is implemented using a TFT circuit, which also is capable of storing the weights $\vec{c}$ obtained from model training. The weak classifiers thus reduce M sensor signals to N dot-product outputs. The reduced signals are then presumed to feed a CMOS IC, which implements thresholding and weighted voting. Next, we will look closer at the implementation of the model-programmable weak classifiers.

1) Weak "Linear" Classifier: The dot product within the linear classifiers is implemented as shown in Fig. 9. Approximation to multiplication between each sensor and each weight from model training is achieved through a series combination of TFTs. The TFTs are biased with M1 in saturation (having transconductance g_{m1}) and M2 in the linear region (giving effective resistance R_{B2}). For M2, we thus have

$$\frac{1}{R_{B2}} \approx \frac{dI_1}{dV_{ds,M2}} = k\frac{W}{L}\left[(V_{B1} - V_T) - V_{ds,M2}\right] \quad (2)$$

and for M1 we have

$$I_1 = k\frac{W}{2L}(V_{S1} - I_1 R_{B2} - V_T)^2 = \frac{g_{m1}}{2}(V_{S1} - I_1 R_{B2} - V_T). \quad (3)$$

Rearranging this results in

$$I_1 = \frac{g_{m1}(V_{S1} - V_T)}{2 + g_{m1} R_{B2}}. \quad (4)$$

If $g_{m1}R_{B2} \gg 2$ is sufficiently satisfied, and M2 is in deep triode such that $(V_{B1} - V_T) \gg V_{ds,M2}$ is sufficiently satisfied, the output current of the series connected TFTs implements an approximation to multiplication

$$I_1 \approx \frac{(V_{S1} - V_T)}{R_{B2}} \approx k\frac{W}{L}(V_{B1} - V_T)(V_{S1} - V_T). \quad (5)$$

Following this, the summation required within the dot-product operation is then realized by combining the output currents from these branches through the resistor R.

However, in practice, this circuit only approximates multiplication. This is because of nonlinearity (arising due to the use of the g_{m1} term), approximations (arising as described above), and variations (arising in all the TFTs). From measurements taken at representative bias voltages, the actual transfer functions achieved for each input (gate of M1 and gate of M2) to the output current are illustrated in Fig. 10. We see that the transfer function from M1, over the designated voltage range (output of the sensor), is fairly linear, and thus implements a good approximation to multiplication. On the other hand, the transfer function from M2 is substantially nonlinear, and thus implements a somewhat worse approximation to multiplication (the ideal multiplication transfer function is overlayed in dotted curves). This motivates feeding the sensor signal to M1 and applying the model weight to M2. The reason for this is that the model weight does not change continuously; rather, once derived from training, the proper biasing level (V_B) can be determined computationally, by applying the inverse of the nominal transfer function from M2, thus mitigating the impact of its nonlinearity. Such a transfer function can be obtained either from device measurements or from circuit simulations (the former is used for the results presented). In either case, the transfer function should represent the behavior after averaging out the effects of random variations, since random variations are overcome by the EACB approach. On the other hand, the sensor signal changes continuously, and is thus preferable to apply directly to the circuit (i.e., without any transform to mitigate the impact of nonlinearity, which would necessitate additional circuitry).

In addition to nonideal deviations in the nominal transfer functions, the TFT circuits also suffer from substantial variation. Shown as error bars in Fig. 10, the resulting variations in the transfer functions (measured from 10 circuits) are severe. Specifically, the device-level variations are measured as reported in Table I. To overcome this, the EACB algorithm is leveraged, and analysis from simulation is presented below.

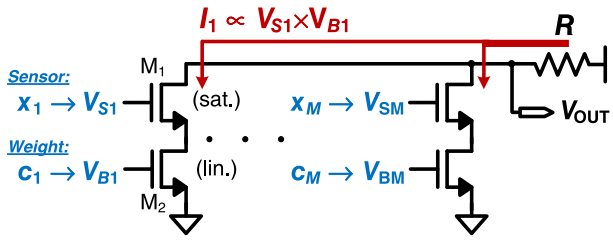

Fig. 9. TFT implementation of the linear-classifier dot-product approximation.

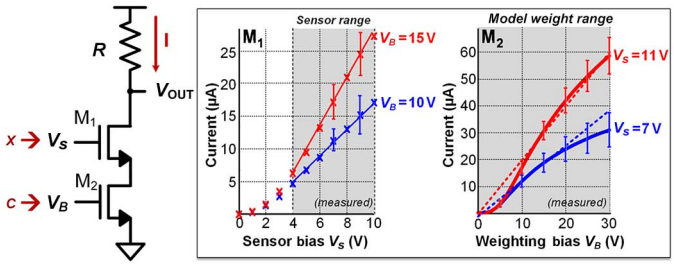

Fig. 10. Measured nonideal deviation and variation in the dot-product approximation circuit (over 10 circuits).

TABLE I
MEASURED VARIATION IN TFT DEVICE PARAMETERS
(OVER 100 DEVICES)

Measured levels of variation					
Output Resistance	$\overline{r_0}$ = 10 MΩ σ = 0.1 MΩ	Threshold Voltage	$\overline{V_T}$ = 2 V σ = 0.3 V	Mobility	$\overline{\mu}$ = 0.7 cm²/Vs σ = 0.02 cm²/Vs

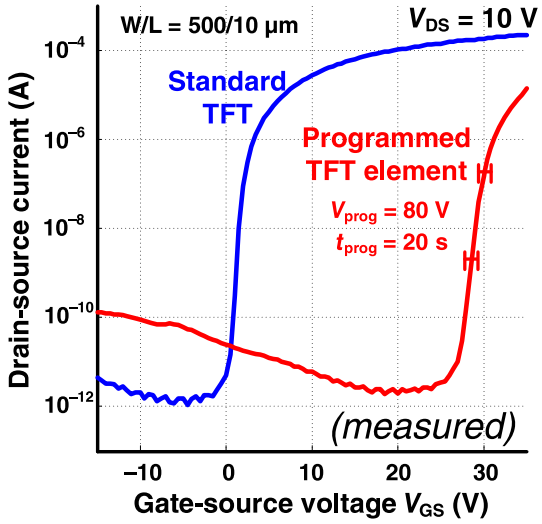

Fig. 11. TFT I–V curves showing threshold-voltage shift for model-programming in TFT classifier.

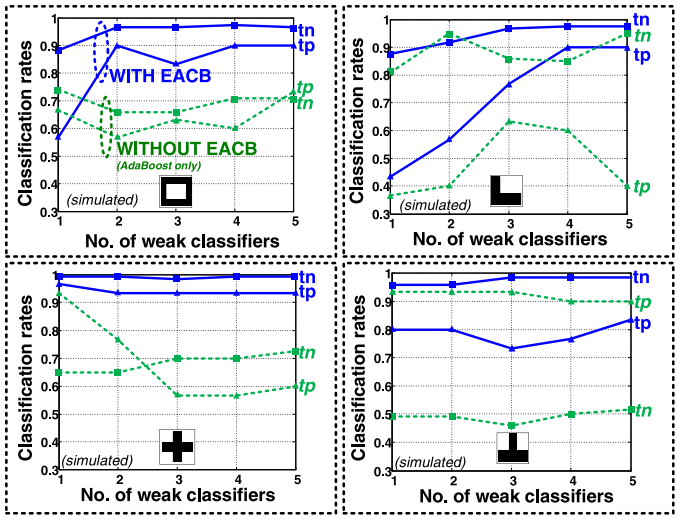

Fig. 14. Simulated classification rates with and without the implementation of EACB (shown for one vs. all classification of four shapes).

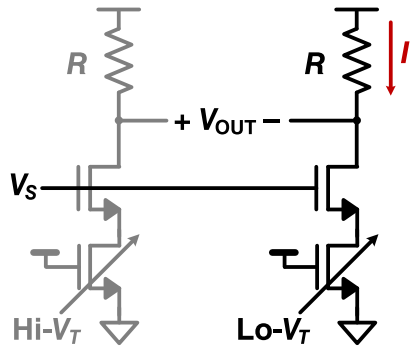

Fig. 12. Complete pseudodifferential multiplication circuit for the application of positive and negative model weights.

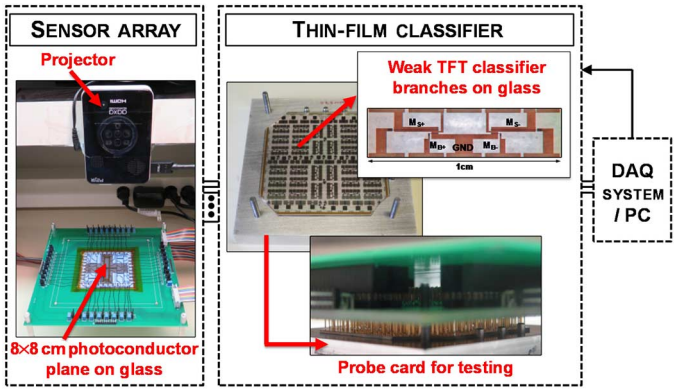

Fig. 15. Thin-film prototype.

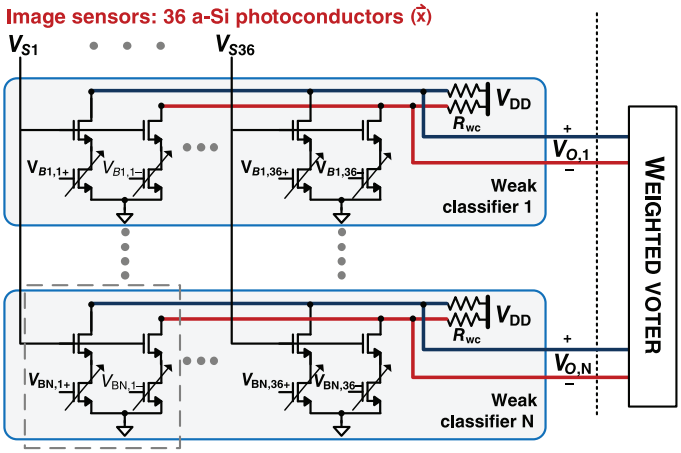

Fig. 13. Complete thin-film sensing and classification system.

2) Classifier Model Programmability: As mentioned, in the multiplication circuit, the gate of M2 is used to apply the trainer-derived model weight. This implies that the circuit must support storage of a programmable analog voltage at the gate of M2. To achieve this, nonvolatile charge trapping in the gate dielectric of a TFT is used [12], [13]. Application of a large electric field on the TFT gate effectively results in a large, controllable threshold-voltage shift. This is shown in Fig. 11 which plots the measured drain current versus gate voltage for a standard TFT, and for a TFT programmed to have a threshold-voltage shift of 30 V (with error bars showing results over six devices). For large model weights, a small positive threshold-voltage shift is applied, while for small model weights, a large positive threshold-voltage shift is applied. Though training typically only occurs once in the system, this form of programming is reversible through the application of a large negative voltage on the TFT gate.

One complication that arises in practice is that the model weights derived from training can be negative. To address this, pseudodifferential multiplication branches are used to create the complete multiplication circuit, as shown in Fig. 12. For example, to apply a negative weight, the left branch is programmed with a large threshold voltage, essentially turning it OFF, while the right branch is programmed with a low threshold voltage, appropriately set.

3) Overall Implementation: Fig. 13 shows the overall implementation with the weak classifiers comprising the

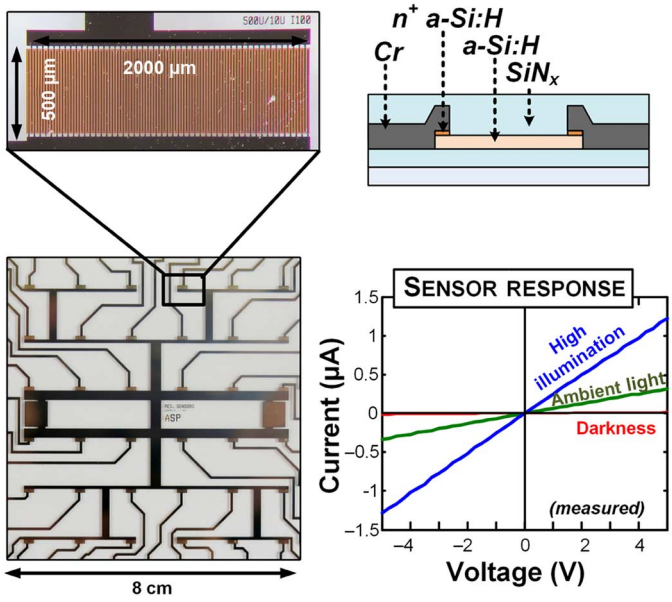

Fig. 16. Thin-film photoconductors.

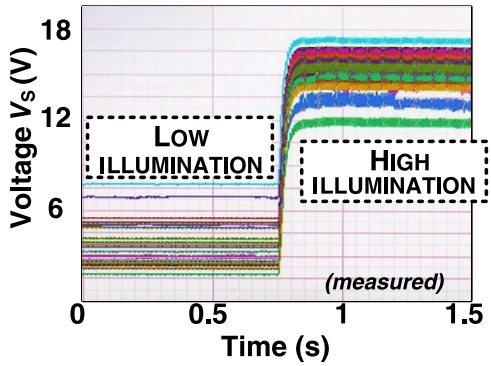

Fig. 17. Variability in thin-film photoconductors.

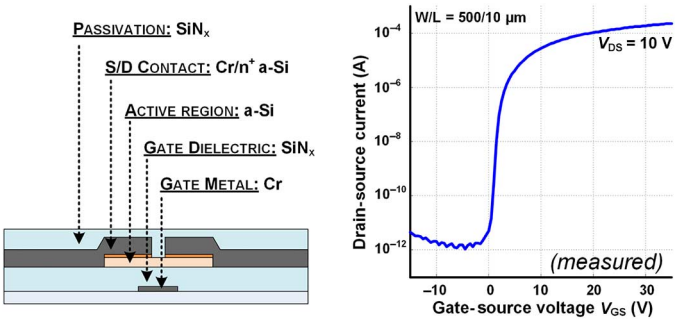

Fig. 18. Amorphous silicon TFT.

summed outputs from the pseudodifferential multiplication units. Thirty-six sensor signals, corresponding to the outputs from photoconductor pixels for image sensing, feed the ensemble of weak classifiers, all appropriately V_T-shifted based on the model weights derived from training. The outputs, thus corresponding to the dot-product results, are subsequently thresholded to derive the weak-classifier decisions. The ultimate goal of this implementation is to reduce the large number of sensor

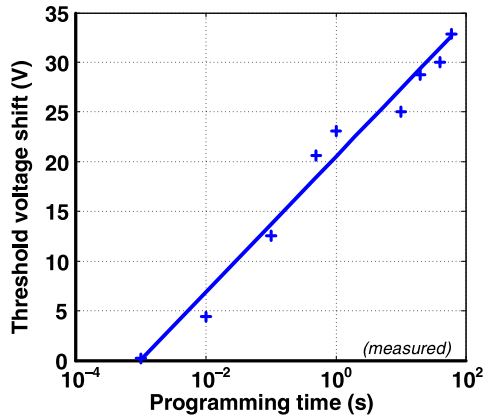

Fig. 19. Effect of programming time on threshold-voltage shift of amorphous silicon TFT.

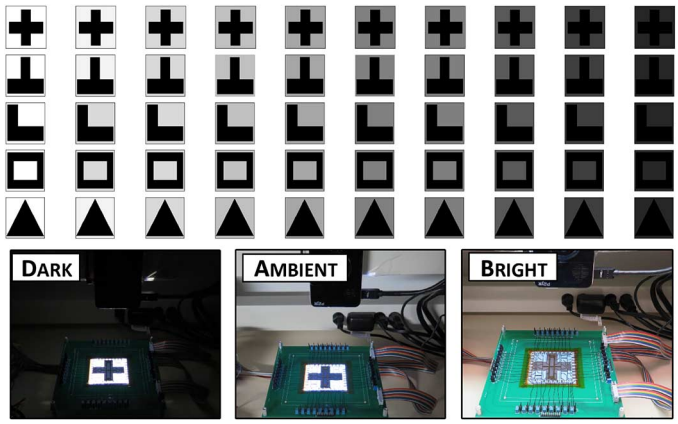

Fig. 20. Subset of dataset used for classification.

signals to just a small number interfaces for the weak-classifier decisions. Although here inputs from all 36 sensors are fed to the weak classifiers, the implementation of an offline feature selection algorithm could also be considered. Such algorithms, which select the sensors providing the greatest information for a given classification application, can substantially reduce the number of inputs [14].

To analyze the system, simulations are performed by creating statistical models of the TFT-based weak classifiers. First, we extract a Level 61 SPICE model of the TFTs from device measurements. Transistor-level simulations are then performed in Spectre to characterize the transfer functions of the multiplication unit, which is used in the weak classifiers. The transfer functions are represented numerically and augmented with measured device-level variation data to create a statistical model of the units usable in MATLAB. Monte Carlo simulations are then performed in MATLAB of the entire system, consisting of multiple weak classifiers composed of variation-affected multiplier units. Using the shape dataset, the weak classifiers are trained, employing both the AdaBoost and EACB algorithms. Fig. 14 shows the simulation results for classifiers trained to four shapes. The metrics used for characterizing classification performance are true-positive (tp) rate and true-negative (tn) rate. As an example, considering the classification of rings versus all the other shapes in the dataset, true positive

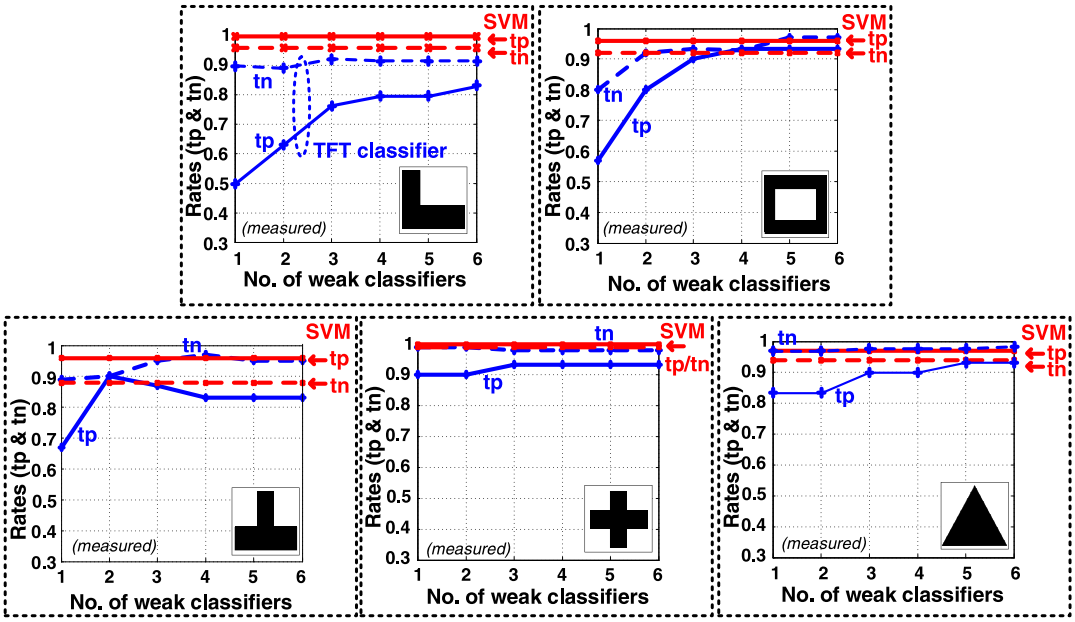

Fig. 21. Measured classification results.

rate (6) and true negative rate (7) are shown at the bottom of the page. As can be seen in Fig. 14, while AdaBoost alone is inadequate, EACB enables high classification performance to be achieved.

IV. System Demonstration

In this section, the thin-film system prototype is presented. As shown in Fig. 15, it consists of two glass substrates onto which both thin-film a-Si sensors and transistors are fabricated in-house using a maximum processing temperature of 180°C.

The first sample is an 8×8 cm array of 36 image sensors onto which images are projected using a microprojector. The outputs from the sensors are passed to the second sample, consisting of the TFT-based weak classifiers. To aid testability, interfacing between the two samples is achieved using a probe card, which permits the acquisition of both weak-classifier and sensor outputs, for processing and analysis by a PC.

A. Image Sensors

The image-sensing array is constructed using sensor pixels based on thin-film photoconductors. The photoconductors are pictured in Fig. 16. These are formed by interdigitated undoped a-Si with a thickness of 150 nm. As shown, these exhibit strong change in I–V response under different illumination conditions.

To implement the complete sensor pixel, this photoconductor is configured as one leg of a voltage divider, where the other leg is a fixed thin-film resistor (750 kΩ). The output voltage from 36 sensors over time is shown in Fig. 17 under two illumination conditions. Although a strong photoresponse can be observed, a substantial amount of variation across the sensors can also be seen. The use of a data-driven machine-learning algorithm allows for modeling of the acquired data even in the presence of such large variations across the sensors.

B. Thin-Film Transistors

The TFTs used in the weak classifiers are fabricated using our standard a-Si process technology [15]. The TFT layers are pictured in Fig. 18 alongside a typical TFT transfer curve.

As described previously, programmability of the weak classifier is achieved via gate-dielectric charge trapping in TFTs with the same structure, simply by applying a large programming voltage on the gate. Typical programming (erase) voltages are +80 V (−80 V), with a drain and source voltage of 0 V in both cases. A typical threshold-voltage shift versus programming time characteristic is shown in Fig. 19 (as described earlier, this threshold-voltage shift can be mapped appropriately to a model weight obtained from training by applying the inverse transfer function of the multiplication circuit). It is also worth noting that the variation in programmed and erase states is small ($\sigma(V_T) \leq 1V$), so the desired shifts can be achieved repeatably. This level of variation in V_T is easily tolerated as a result of using the EACB algorithm.

A potential future focus for the implementation is the programming voltages and programming times required, both of

$$\text{True positive rate} = \frac{\text{number of correctly classified } ring\ instances \text{ in the dataset datapoints}}{\text{total number of } ring\ instances \text{ in the dataset}} \tag{6}$$

$$\text{True negative rate} = \frac{\text{number of correctly classified } other\ shapes}{\text{total number of instances of } other\ shapes \text{ in the dataset}} \tag{7}$$

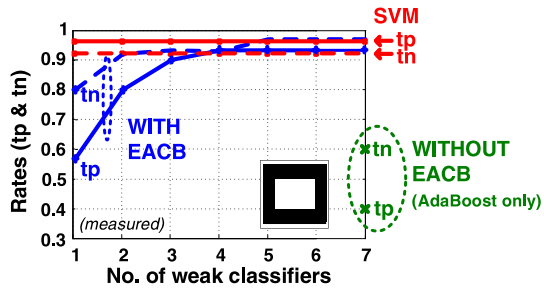

Fig. 22. Effect of error-adaptive classifier boosting (EACB).

which are somewhat large. Substantial room exists to optimize these. For instance, previous work has demonstrated that these can be reduced by using a slightly modified TFT structure, consisting of the addition of an explicit a-Si floating gate [16]. This previous work also highlights that retention time and biasing stresses are strongly affected by the chosen memory structure. For the programmable elements described in this work, a charge retention of 80% of programmed state is typical after 1 h; however, structures achieving retention of 80% over a duration of greater than 3 months have been realized (e.g., [16]).

C. Image Classification System

Some of the images from the dataset used for testing the thin-film shape classifier are shown in Fig. 20. In total, 150 data instances are used, comprising 5 shapes with 10 background intensities, under three illumination conditions.

Fig. 21 shows plots of tp and tn rates for one-versus-all classification of each shape with respect to the number of weak classifiers used. As can be seen for all shapes, boosted performance is observed, as TFT weak-classifier iterations are added. Indeed, high overall classification performance is achieved with very few weak classifiers in all cases.

For reference, horizontal lines show the performance achieved when performing classification using an ideal MATLAB-implemented SVM classifier, a widely used strong classifier. Between just two and five weak classifiers (iterations) are required to achieve performance at the level near that of the SVM, implying a substantial reduction in the number of interfaces, compared to the 36 raw sensor inputs.

Fig. 22 highlights the importance of using the EACB algorithm as opposed to simply using the AdaBoost algorithm, illustrated for the case of classification of ring versus other shapes. Without EACB, even with seven weak-classifier iterations, convergence is not achieved and low classification performance is observed, whereas EACB leads to substantially boosted performance.

V. Conclusion

This paper presents the implementation of a thin-film sensing and classification system based on LAE. While LAE enables the formation of diverse and expansive arrays of sensors, the resulting TFTs have low performance and high variability, precluding the implementation of complex functions for processing sensor signals. The presented system uses a machine-learning algorithm, known as EACB, which enables the creation of a strong classifier using an ensemble of simple classifier circuits having low electrical performance and high variability. Implemented using TFTs, the weak-classifier circuits reduce the sensor outputs to a small number of signals, which can be provided to a CMOS IC to make the final classification decisions.

High classification performance is demonstrated for a shape classification system, at the level of a strong complex classifier (SVM), even in the presence of substantial sensor and TFT variability. Trained to detect five different shapes, a reduction in signals of 3.5–9$\times$ is achieved with respect to the 36 raw sensor outputs. Though this represents a small-scale proof of concept, the approach of reducing data to weak-classifier decisions is promising as the size of sensing array scales. This is because algorithms such as AdaBoost (employed within EACB) perform powerful optimizations during weak-classifier training. In addition, architectural techniques, such as an active matrix, used in conjunctions with such algorithms, can enable substantial reduction in the total number of interfaces, and present a compelling direction for future work.

Acknowledgment

The authors would like to thank Systems on Nanoscale Information fabriCs (SONIC), one of six SRC STARnet Centers, sponsored by MARCO and DARPA.

References

[1] H. Wang, L. Chen, J. Wang, Q. Sun, and Y. Zhao, "A micro oxygen sensor based on a nano sol–gel TiO_2 thin film," *Sensors*, vol. 14, no. 9, pp. 16423–16433, Sep. 2014.

[2] L. Zhou, S. Jung, E. Brandon, and T. N. Jackson, "Flexible substrate micro-crystalline silicon and gated amorphous silicon strain sensors," *IEEE Trans. Electron Devices*, vol. 53, no. 2, pp. 380–385, Feb. 2006.

[3] C. Dagdeviren *et al.*, "Conformable amplified lead zirconate titanate sensors with enhanced piezoelectric response for cutaneous pressure monitoring," *Nat. Commun.*, vol. 5, no. 4496, pp. 1–10, Aug. 2014.

[4] Y. Hu *et al.*, "A self-powered system for large-scale strain sensing by combining CMOS ICs with large-area electronics," *IEEE J. Solid-State Circuits*, vol. 49, no. 4, pp. 838–850, Apr. 2014.

[5] L. Huang *et al.*, "Reconstruction of multiple-user voice commands using a hybrid system based on thin-film electronics and CMOS," in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2015, pp. C196–C197.

[6] W. Rieutort-Louis, T. Moy, Z. Wang, S. Wagner, J. Sturm, and N. Verma, "A large-area image sensing and detection system based on embedded thin-film classifiers," *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, no. 16.2, pp. 1–3, 2015.

[7] J. Zhang, Z. Wang, and N. Verma, "A matrix-multiplying ADC implementing a machine-learning classifier directly with data conversion," *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, no. 18.4, pp. 1–3, 2015.

[8] R. Genov and G. Cauwenberghs, "Kerneltron: Support vector machine in silicon," *IEEE Trans. Neural Netw.*, vol. 14, no. 5, pp. 1426–1434, 2003.

[9] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997.

[10] Z. Wang, R. Schapire, and N. Verma, "Error-adaptive classifier boosting (EACB): Leveraging data-driven training towards hardware resilience for signal inference," *IEEE Trans. Circuits Syst. I (TCAS-I)*, vol. 62, no. 4, pp. 1136–1145, Apr. 2015.

[11] T. Moy *et al.*, "Thin-film circuits for scalable interfacing between large-area electronics and CMOS ICs," in *Proc. Device Res. Conf. Dig. Tech. Papers*, Jun. 2014, pp. 271–272.

[12] M. Powell, C. van Berkel, A. Franklin, S. Deane, and W. Milne, "Defect pool in amorphous-silicon thin-film transistors," *Phys. Rev. B*, vol. 45, p. 4160, 1992.
[13] Y. Huang, S. Wagner, and J. Sturm, "Nonvolatile amorphous-silicon thin-film-transistor memory structure for drain-voltage independent saturation current," *IEEE Trans. Electron Devices*, vol. 58, no. 9, pp. 2924–2927, Sep. 2011.
[14] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.
[15] B. Hekmatshoar, K. Cherenack, A. Kattamis, K. Long, S. Wagner, and J. C. Sturm, "Highly stable amorphous-silicon thin-film transistors on clear plastic," *Appl. Phys. Lett.*, vol. 93, no. 03213, pp. 1–3, 2008.
[16] Y. Huang, "Novel approaches to amorphous silicon thin film transistors for large area electronics," Ph.D. dissertation, Department of Electrical Engineering, Princeton Univ., Princeton, NJ, USA, 2011.
[17] Varian. (2015). *Amorphous Silicon X-Ray Imagers* [Online]. Available: https://www.varian.com/

Warren Rieutort-Louis (S'12–M'15) received the B.A. (Hons.) and M.Eng. degrees in electrical and information engineering from Trinity College, Cambridge University, Cambridge, U.K., in 2009, the M.A. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2012, and the Ph.D. degree in Electrical Engineering at Princeton University in 2015.

He was a Graduate Teaching Fellow with Princeton McGraw Center for Teaching and Learning. His research interests include thin-film materials, processes, devices, and circuits for large-area electronic systems.

Dr. Rieutort-Louis was the recipient of the IBM Ph.D. Fellowship, the Andlinger Center Maeder Fellowship in Energy and the Environment, the Princeton Harold W. Dodds Honorific Fellowship.

Tiffany Moy (S'14) received the B.S.E. (*magna cum laude*) and M.A. degrees in electrical engineering from Princeton University, Princeton, NJ, USA, in 2012 and 2014, respectively, where she is currently pursuing the Ph.D. degree in Electrical Engineering.

Her research interests include thin-film devices and circuits for hybrid large-area electronics/CMOS system design.

Zhuo Wang (S'13) received the B.S. degree in microelectronics from Peking University, Beijing, China, in 2011, the M.A. degree in Electrical Engineering, from Princeton University, Princeton, NJ, USA, in 2013. He is currently pursuing the Ph.D. degree in electrical engineering at Princeton University, Princeton, NJ, USA.

His research interests include leveraging statistical approaches such as machine learning, for achieving hardware relaxation in an algorithmic and architectural level in resource-constrained platforms such as embedded sensing systems.

Mr. Wang was the recipient of the 2011 Peking University Best Undergraduate Dissertation Award, the 2014 ICASSP Conference Best Paper Award Nomination, and the 2015 Princeton University Honorific Fellowship.

Sigurd Wagner (SM'78–F'00) received the Ph.D. degree in chemistry from the University of Vienna, Vienna, Austria, in 1968.

Following a Postdoctoral Fellowship at Ohio State University, Columbus, OH, USA, he worked from 1970 to 1978 with the Bell Telephone Laboratories, Murray Hill, Holmdel, NJ, USA, on semiconductor memories and heterojunction solar cells. He then joined the Solar Energy Research Institute (now NREL), Golden, CO, USA, as the Founding Chief of the Photovoltaic Research Branch. Since 1980, he has been a Professor of Electrical Engineering with Princeton University, Princeton, NJ, USA; in 2015, he became Professor Emeritus and Senior Scholar.

Dr. Wagner is a member of Princeton's Large-Area Systems Group whose goal is to demonstrate complete large-area applications based on hybrid thin film/CMOS architectures. He is a fellow of the American Physical Society and a Member of the Austrian Academy of Science. He was the recipient of the Nevill Mott Prize for his groundbreaking research, both fundamental and applied, on amorphous semiconductors and chalcopyrites, and the ITC Anniversary Prize for pioneering research on flexible and stretchable large-area electronics, and the comprehensive study of its mechanical behavior.

James C. Sturm (S'81–M'85–SM'95–F'01) was born in Berkeley Heights, NJ, USA, in 1957. He received the B.S.E. degree in electrical engineering and engineering physics from Princeton University, Princeton, NJ, USA, in 1979, and the M.S.E.E. in electrical engineering and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1981 and 1985, respectively.

In 1979, he joined Intel Corporation, Santa Clara, CA, USA, as a Microprocessor Design Engineer, and in 1981, he was a Visiting Engineer at Siemens, Munich, Germany. In 1986, he joined the Faculty of Princeton University, where he is currently the Stephen R. Forrest Professor in Electrical Engineering. Since 2003, has been the Founding Director of the Princeton Institute for the Science and Technology of Materials (PRISM), Princeton, NJ, USA, and from 1994 to 1995, he was a von Humboldt Fellow with the Institut fuer Halbleitertechnik, University of Stuttgart, Stuttgart, Germany. His research interests include silicon-based heterojunctions, flexible and large-area electronics, photovoltaics, nano-bio interface, and three-dimensional (3-D) integration.

Dr. Sturm is a member of the American Physical Society and the Materials Research Society and was a National Science Foundation Presidential Young Investigator. He served on the organizing committee of IEDM (1988–1992 and 1998–1999), having chaired both the solid-state device and detectors/sensors/displays committees. He has served on the Boards of Directors of the Materials Research Society and the Device Research Conference, and was a Co-Founder of Aegis Lightwave. From 1996 to 1997, he was the Technical Program Chair and General Chair of the IEEE Device Research Conference, respectively. He was the recipient of over 10 awards for teaching excellence.

Naveen Verma (S'03–M'09) received the B.A.Sc. degree in electrical and computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2003, and the M.S. and Ph.D. degrees in electrical engineering from Massachusetts Institute of Technology, Cambridge, MA, USA, in 2005 and 2009, respectively.

Since July 2009, he has been with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, where he is currently an Associate Professor. His research interests include advanced sensing systems such as low-voltage digital logic and SRAMs, low-noise analog instrumentation and data-conversion, large-area sensing systems based on flexible electronics, and low-energy algorithms for embedded inference, especially for medical applications.

Dr. Verma was the recipient or corecipient of the 2006 DAC/ISSCC Student Design Contest Award, the 2008 ISSCC Jack Kilby Paper Award, the 2012 Alfred Rheinstein Junior Faculty Award, the 2013 NSF CAREER Award, the 2013 Intel Early Career Award, the 2013 Walter C. Johnson Prize for Teaching Excellence, the 2013 VLSI Symp. Best Student Paper Award, and the 2014 AFOSR Young Investigator Award.