Subtle Anomaly Detection of Microscopic Probes using Deep learning based Image Completion

Kosuke Ikeda
Advantest Japan
Applied Research and Technology Japan
Tokyo, Japan
kosuke.ikeda@advantest.com

Keith Schaub
Advantest America, Inc.
Applied Research and Technology
Austin, TX
keith.schaub@advantest.com

Ira Leventhal
Advantest America, Inc.
Applied Research and Technology
San Jose, CA
Ira.leventhal@advantest.com

Yiorgos Makris
University of Texas at Dallas
Computer Science Dept.
Dallas, TX
yiorgos.makris@utdallas.edu

Constantinos Xanthopoulos
University of Texas at Dallas
Computer Science Dept.
Dallas, TX
constantinos.xanthopoulos@utdallas.edu

Deepika Neethirajan
University of Texas at Dallas
Computer Science Dept.
Dallas, TX
deepika.neethirajan@utdallas.edu

I. Problem Statement (*Finding and classifying subtle surface defects in a production manufacturing environment is an important task, in particular, microscopic wafer probe tip manufacturing for a massively parallel wafer probe card.*)

Automated defect inspection in manufacturing of microscopic probes is an important task and often requires machine learning driven solutions. A supervised only approach can be challenging, because production manufacturing process typically have few defects, thus large amounts of labeled training data are generally not available. In this work, we instead employed multiple models in a multi-step process to achieve the end goal of identifying defect and non-defect probe tips.

Thousands to tens-of-thousands of wafer probe tips are individually manufactured and used to build a single-massively parallel production wafer probe card. It is essential that 100%

typically have a manufacturing yield > 99%. Thus, the problem becomes screening out the < 1% probe tip defects in an automated way (the proverbial problem - "finding a needle in a haystack"). If just one defective probe tip slips through the quality inspection stage, it leads to costly re-work and lost time to debug and troubleshoot the probe card to find the one out of a thousand defective probe tips.

II. Proposed Solution

A. *Approach (An unsupervised machine learning approach is proposed)*

Since our application space has a low and insufficient number of defect images due to the extremely high process yield, using a supervised approach was not feasible.

Instead, we utilized an unsupervised approach where the convolutional neural network is trained only with fault-free images to complete specific patches of the image. The reconstructed areas, when compared against the original input image, act as anomaly detectors. Thresholds are then determined to correctly identify PASS/FAIL images.

B. *Highlights of technical challenges and solution*

In cases where yield is exceptionally high, in our case yield > 99%, the defect sample size is extremely small. Using a supervised approach is challenging and impractical, because of lack of FAIL/DEFECT images, high variances of the types of imperfections, and lack of labeled training data. Additionally, some images may be out-of-focus which need to be correctly classified and filtered out in a pre-processing step. Shown in figure 1 is a subset of the thousands of in focus, out-of-focus, defect, and defect-free images. The differences between defect and defect-free images is extremely subtle and difficult to discern even to a human classifier. Additionally, human classification is tedious and time consuming. False positives are most costly and eliminating them is paramount, as false positives will lead to a non-functional probe card requiring expensive engineering and equipment to debug and repair.

978-1-7281-4823-6/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on August 02,2020 at 22:20:25 UTC from IEEE Xplore. Restrictions apply.

Figure 1: Thousands of Unknown Images

III. Experiments and Analysis

A. Experiment Setup

The experimental process is outlined in Figure 2: Process Diagram. We originally applied three processing steps to the input images. The first model was acting as a pre-processing step. Sometimes, due to faulty probe tip processing, the probe tip could be completely missing from the reticle location, or due to camera focusing issues, the image could be Out-of-Focus. In the pre-processing step a YOLO model was applied to correctly identify and locate the probe image. Sobel filtering was also applied to identify an Out-of-Focus image, however, we later removed the Sobel filtering step as redundant after the YOLO model was adequately trained to identify both “missing” and “Out-of-Focus” images. If the probe tip was missing or Out-of-Focus, the YOLO model could filter out those images with greater than 99.9% accuracy. The remaining images were then sent to the primary model. In the primary model, we used a custom DL model to classify defect or defect-free.

In the pre-processing step, using supervised learning, we trained and applied a binary classifier to pre-process the input images and remove the out-of-focus images.

We calculated a sharpness metric on each of the images. This sharpness metric is the variance of the Sobel transformation factor along the horizontal axis across the image. We set a threshold value for the sharpness metric which determines whether an image is In-Focus or Out-of-Focus.

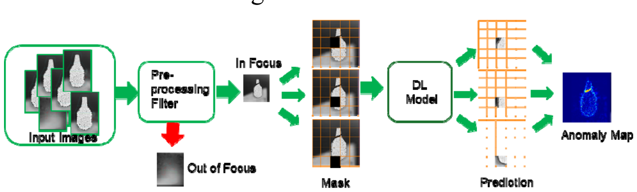

Figure 2: Process Diagram

This threshold value was selected in such a way that it was capable of clearly separating the 2 classes of images. For the Sobel transformation sharpness metric, we set a threshold of 400,000. If the sharpness metric was greater than 400,000 for an image, then the image was binned as ”Clear / In-Focus” and if the sharpness metric was less than 400, 000, the image was binned as ”Blur / Out-of-Focus”. We ran the experiment and the results showed that the images were binned correctly in their respective folders. Figure 3 shows the binning of the images based on the Sobel sharpness metric. The grey bars indicate the images that are In-Focus and the cyan bars indicate the images that are out-of-focus or partially Out-of-Focus. The red dashed line at 400, 000 mark shows the threshold value.

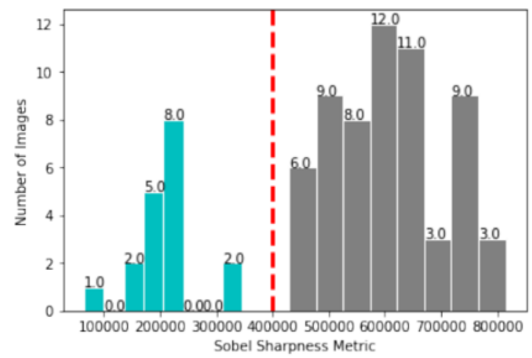

Figure 3: Histogram of Number of Images binned In/Out of Focus

The remaining images were then fed to a DL model that was trained only with known good images using unsupervised learning. This method is known as Imprinting. The unsupervised imprinting approach takes advantage of the high frequency of defect-free images where the model learns what defect-free images look like. Once trained, the input image is fed 48 times to the DL model, where each iteration is purposefully missing a small section of the image. The DL model then completes the missing image by replacing the missing block with its best estimation of the missing patch. The L2norm between the estimated patch and actual patch is calculated for each of the 48 blocks and the maximum result is retained and used as a threshold to identify defect vs non-defect.

B. Results

Figure 4 shows the confusion matrix. The model correctly predicted a defect free probe 99.68% and a defective probe 98.35% of the time. Importantly, FP’s 1.65% result indicates the DL model’s ability to prevent false positives (costliest element), however, more work is needed to achieve FP = 0 %.

		Prediction	
		Good	Bad
Actual	Good	TP=99.68% (29905)	FN=0.31% (95)
	Bad	FP=1.65% (3)	TN=98.35% (179)

Figure 4: Confusion Matrix of DL

Figure 5 is a sample indicating good (defect free) images correctly predicted. The leftmost image is the input image. The center image is the 48 compilations from the DL model. This center result indicates the difference between the input image and the models estimation of each of the 48 blocks. Ideally,

978-1-7281-4823-6/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on August 02,2020 at 22:20:25 UTC from IEEE Xplore. Restrictions apply.

this area should be zero indicating no difference between the input image and the estimated completed image. The rightmost image is a heat map of the final output image where values higher than the threshold valueare highlighted on the image. For these defect images, we do not see any defects highlighted.

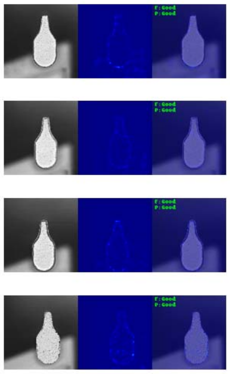

Figure 5: Good Images correctly predicted

Figure 6 is a sample showing bad (defective) images correctly predicted. Similarly, the leftmost image is the original input. The center image indicates and highlights the differences between the input image and the model's 48 estimated completions. You can see various highlighted areas indicating that these areas are different from ideal. The rightmost highlights those differences if they surpass the empirically determined threshold value.

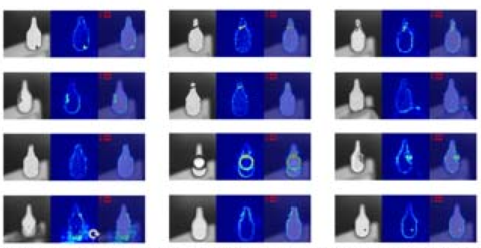

Figure 6: Defective Images correctly predicted

Results from analyzing Advantest America, Inc. production parametric probe tip images demonstrate that this approach is suitable for the subtle detections necessary of visible anomalies and is a viable solution for identifying subtle defects in a production probe manufacturing environment.

IV. Comparison to State-of-the-Art

Our DL method achieved similar results to reference [1] Anomaly detection using deep learning-based image completion"

In the future, additional models can be developed that utilize the output data of this model to classify the type of defect (i.e. "crack", "large nodule", "debris near tip", "center hold"), which the current rules-based method is unable to do.

References

[1] M. Haselmann, D. P. Gruber, P. Tabatabai, "Anomaly detection using deep learning based image completion," IEEE 17th International Conference on Machine Learning and Applications (ICMLA), 2018.

[2] Thomas Schlegl, Philipp Seeboeck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, Georg Langs, Christian Doppler Laboratory for Ophthalmic image analysis, "Unsupervised anomaly detection with generative adversarialnetworks to guide marker discovery", IPMI 2017

978-1-7281-4823-6/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on August 02,2020 at 22:20:25 UTC from IEEE Xplore. Restrictions apply.