WAFERMAP PATTERNS CLUSTERING VIA VARIATIONAL AUTOENCODERS

Elahe Ghalebi, Radu Grosu

E191 - Institute of Computer Engineering

INTRODUCTION

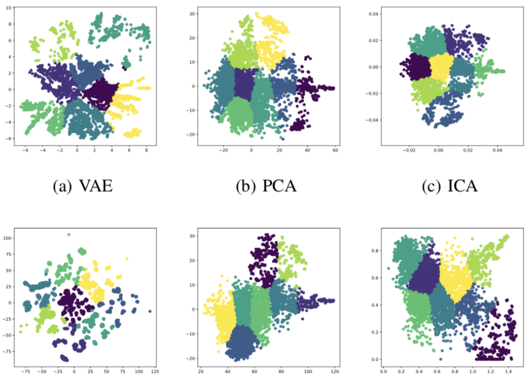
"Industry 4.0" has confronted with new challenges to be able to compete in the increasing global markets in the past couple of years. The industries decided to adapt their manufacturing processes to new paradigms such as automated root cause analysis and decision making in the semiconductor manufacturing. Defect detection from given sensory data is one major goal of their contribution. Defects are exhibited in typical shapes such as rings, spots, repetitive patters or scratches. Several methods has been proposed to recognize these patterns based on traditional image processing [1,2]. There also have been proposed methods on supervised training of mixture models [3], neural networks [4], or support-vector machines [5]. All these methods require a human expert to manually label the data of training dataset. By using unsupervised approaches the intervention of human experts is eliminated and the hidden dependencies between various kinds of wafer defects are detected automatically which enables detection of patterns that were unknown or overlooked before. Our goal is to provide an unsupervised method for clustering the wafermap patterns using variational autoencoder. Our algorithm is proceed in three steps: (1) pre-processing of the raw wafer dataset using computer vision techniques, (2) extracting features by training variational autoencoder on the cleaned data, and (3) clustering latent features and group the wafers based on a distance measure using k-means clustering.

PRE-PROCESSING PHASE

The wafer dataset contains some irregular shapes of wafermaps for missing values (or holes) within the wafer area. We binarize the wafer by replacing all present values with 1 and all missing values with 0. We then close the small holes by mathematical morphology approach. All binary morphological operations are the result of combining two morphological operators–erosion and dilation. The former is sufficient for closing the holes in the binary clipping mask. Although random measurement errors caused by physical limitations of manufacturing device lower the quality of recorded data, they are usually within certain range and do not cause major problems when building a predictive model. However, occasional large inaccuracies or malfunctions in measurement process can introduce errors called outliers. The mean and standard deviation, used to to characterize normal distribution, are especially susceptible to perturbations caused by outliers. On the other hand, median is resistant to gross errors in up to 50% of the samples. Our method uses modified Z-score to detect the outliers.

FEATURE EXTRACTION PHASE

The dataset obtained by pre-processing can be seen as a random discrete variable X consisting of d individual samples $x_1, \ldots, x_d$. In order to overcome the course of dimensionality for the classification task, we extract only the most salient features $z_1, \ldots, z_d$ represented by normally distributed variable Z. The approach used for this task is based on auto- encoding variational Bayes.



(a) VAE (b) PCA (c) ICA

(d) t-SNE (e) truncated SVD (f) NMF

Figure 1: Wafer dataset projected into two dimensional latent feature space and clustered with k-means into 8 clusters. Different feature extraction methods have been used: (a) Variational Autoencoder; (b) Principal Component Analysis; (c) Independent Component Analysis; (d) t-Distributed Stochastic Neighbor Embedding; (e) Truncated Singular-value Decomposition; (f) Non-Negative Matrix Factorization.

CLUSTERING PHASE

Clustering groups set of objects together based on a similarity between them. We typically recognize two types of clustering -hierarchical clustering which groups the set of objects into a hierarchical tree and partitioning clustering which groups the set of objects into a disjoint subsets, s.t. each object is exactly in one subset. Given a set of n latent features $(x_1, x_2, \ldots, x_n)$, we cluster them with k-means into a disjoint clusters $C_1, C_2, \ldots, C_k$. The objective of this algorithm is to optimize squared Euclidean distance between the latent feature x and the centroid μ_i of the cluster C_i till convergence.

EXPERIMENTS

In order to evaluate the performance of the algorithm presented in this paper, we have compared the variational autoencoder approach with other commonly used decomposition methods. The same pre-processing and clustering algorithm have been used in with all tested methods. Detected clusters for two dimensional latent space can be seen in Figure 1.

REFERENCES

[1] L. Breaux and B. Singh. Automatic defect classification system for patterned semiconductor wafers. In Proceedings of International Symposium on Semiconductor Manufacturing, pages 68-73, Sep 1995.

[2] Frederic Duvivier. Automatic detection of spatial sig- nature on wafermaps in a high volume production. In Proceedings of the 14th International Symposium on Defect and Fault-Tolerance in VLSI Systems, DFT 99, pages 61-, Washington, DC, USA, 1999. IEEE Computer Society.

[3] Fourmun Lee, A. Chatterjee, and D. Croley. Advanced yield enhancement: computer-based spatial pattern analysis. part 1. In IEEE/SEMI 1996 Advanced Semiconductor Manufacturing Conference and Workshop. Theme-Innovative Approaches to Growth in the Semiconductor Industry. ASMC 96 Proceedings, pages 409-415, Nov 1996.

[4] F.L. Chen, Sheng-Che Lin, K. Yih-Yuh Doong, and K.L. Young. Logic product yield analysis by wafer bin map pattern recognition supervised neural network. 2003 5th International Conference on ASIC. Proceedings (IEEE Cat. No.03TH8690).

[5] Li-Chang Chao and Lee-Ing Tong. Wafer defect pattern recognition by multi-class support vector machines by using a novel defect cluster index. Expert Systems with Applications, 36(6):10158 - 10167, 2009.