Artificial Neural Networks based Approaches: Simulation Toys or Real Solutions?

Kurosh Madani

Images, Signals, and Intelligent System Laboratory (LISSI / EA 3956)
Paris-XII University, Senart Institute of Technology
Avenue Pierre Point, Lieusaint, 77127, France
madani@univ-paris12.fr

Abstract. If over the past decades understanding, modeling and improvement of learning and generalization capabilities of Artificial Neural networks have been subject to a particular attention, nowadays, these connectionist models have to face up to new challenges dealing with industrial requirements and real-world dilemmas. The target is to explore this pertinent point through a set of ANN based solutions developed in order to defy a number of applicative challenges dealing with industrial and real world requirements.

1 Introduction

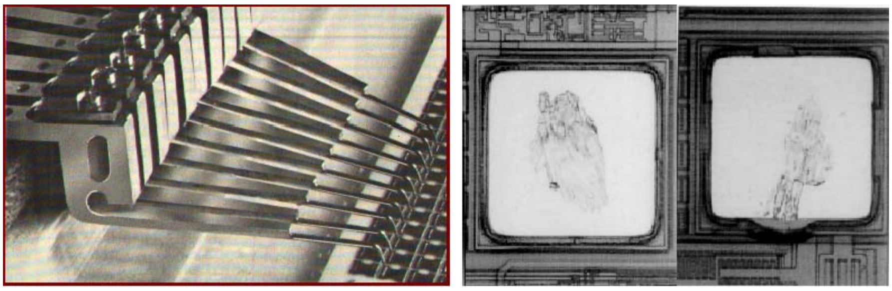
Overcoming limitations of conventional approaches thank to their learning and generalization capabilities, Artificial Neural Networks (ANN) made appear a number of expectations to design "intelligent" information processing systems. If over the past decades understanding, modeling and improvement of learning and generalization capabilities of these bio-inspired models have been subject to a particular attention, nowadays, these connectionist models have to face up to new challenges dealing with industrial requirements and real-world dilemmas. In other words, now the time is to fulfill the following question: do ANN based approaches remain simulation toys or they reveal real solutions for reaching farther technological borders?

The present paper aims to contribute in discussion around the possible answer to the above-formulated question. The target is to discuss and explore this pertinent question through a set of ANN based solutions developed in order to defy a number of applicative challenges dealing with industrial and real world requirements.

The present paper is organized in following way: the next section will introduce the first application of ANN based approaches dealing with "automated visual inspection" in industrial production of VLSI devices. Section 3, will present another industrial application of ANN based concepts focusing fault detection and defects' classification in high-tech optical devices production. Finally the last section of this paper will conclude the present article and discuss a number of perspectives.

2 ANN based Probe Mark Inspection in VLSI Chips Production

One of the main steps in VLSI circuit production is the testing step. This step verifies if the final product (VLSI circuit) operates correctly or not. The verification is performed thank to a set of characteristic input signals (stimulus) and associated responses obtained from the circuit under test. A set of such stimulus signals and associated circuit's responses are called test vectors. Test vectors are delivered to the circuit and the circuit's responses to those inputs are catch through standard or test dedicated Input-Output pads (I/O pads) called also vias. As in the testing step, the circuit is not yet packaged, the test task is performed by units, which are called probers including a set of probes performing the communication with the circuit. The problem is related to the fact that the probes of the prober may damage the circuit under test. So, an additional step consists of inspecting the circuit's area to verify vias (I/O pads) status after circuit's testing: this operation is called developed Probe Mark Inspection (PMI). Fig.1-a shows a picture of probes relative to such probers. Fig.1-b gives examples of faulty and correct vias.



(a) (b)

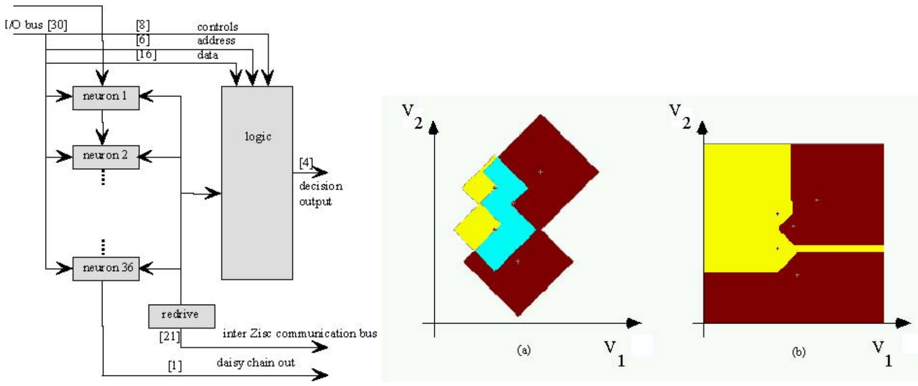
Fig. 1. Photograph giving an example of probes in industrial prober (a). Example of probe impact: correct and faulty (b).

Many prober constructors had already developed PMI software based on conventional pattern recognition algorithms with little success]. The difficulty is related to the compromise between real time execution (production constraints) and methods reliability. In fact, even sophisticated hardware using DSPs and ASICs specialized in image processing are not able to perform sufficiently well to convince industrials to switch from human operator (expert) defects recognition to electronically automatic PMI. That's why a neural network based solution has been developed and implemented on ZISC-036 neuro-processor, for the IBM Essonnes plant. The main advantages of developed solutions are real-time control and high reliability in fault detection and classification tasks. Our automatic intelligent PMI application, detailed in [1] and [2], consists of software and a PC equipped with this neural board, a video acquisition board connected to a camera and a GPIB control board connected to a wafer prober system. Its goal is image analysis and prober control.

The IBM ZISC-036 (see [3] and [4]) is a parallel neural processor based on the RCE and KNN algorithms. Each chip is capable of performing up to 250 000 recognitions per second. Thanks to the integration of an incremental learning algorithm, this circuit is very easy to program in order to develop applications; a very few number of functions (about ten functions) are necessary to control it. Each ZISC-036 like neuron implements two kinds of distance metrics called L1 and LSUP respectively. Relations (1) and (2) define the above-mentioned distance metrics were Pi represents the memorized prototype and Vi is the input pattern. The first one (L1) corresponds to a polyhedral volume influence field and the second (LSUP) to a hyper-cubical influence field.

$$L1 : dist = {}_{i=0}^{n}| V_i - P_i | \quad (1)$$

$$LSUP : dist = \max_{i=0 \ldots n} | V_i - P_i | \quad (2)$$



(a) (b)

Fig. 2. IBM ZISC-036 chip's bloc diagram (a) and an example of input feature space mapping in a 2-D space using ROI and 1-NN modes, using norm L1 (b).

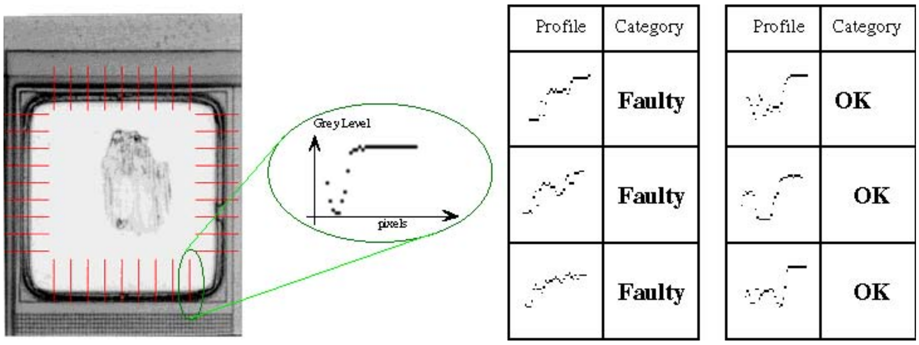
Figure 2 gives the ZISC-036 chip's bloc diagram and an example of input feature space mapping in a 2-D space. A 16 bit data bus handles input vectors as well as other data transfers (such as category and distance), and chip controls. Within the chip, controlled access to various data in the network is performed through a 6-bit address bus. ZISC-036 is composed of 36 neurons. This chip is fully cascadable which allows the use of as many neurons as the user needs (a PCI board is available with a 684 neurons). A neuron is an element, which is able to:

- memorize a prototype (64 components coded on 8 bits), the associated category (14 bits), an influence field (14 bits) and a context (7 bits),
- compute the distance, based on the selected norm (norm L1 given by relation or LSUP) between its memorized prototype and the input vector (the distance is coded on fourteen bits),
- compare the computed distance with the influence fields,
- communicate with other neurons (in order to find the minimum distance, category, etc.),
- adjust its influence field (during learning phase).

Two kinds of registers hold information in ZISC-O36 architecture: global registers and neuron registers. Global registers hold information for the device or for the full network (when several devices are cascaded). There are four global registers implemented in ZISC-036: a 16-bits Control & Status Register (CSR), a 8-bits Global Context Register (GCR), a 14-bits Min. Influence Field register (MIF) and a 14-bits Max. Influence Field register (MAF). Neuron registers hold local data for each neuron. Each neuron includes five neuron registers: Neuron Weight Register (NWR), which is a 64-by-8 bytes register, a 8-bits Neuron Context Register (NCR), Category register (CAT), Distance register (DIST) and Neuron Actual Influence Field register (NAIF). The last three registers are both 14-bites registers. Association of a context to neurons is an interesting concept, which allows the network to be divided in several subsets of neurons. Global Context Register (GCR) and Neuron Context Register (NCR) hold information relative to such subdivision at network and neuron levels respectively. Up to 127 contexts can be defined.

The process of analyzing a probe mark can be described as following: the PC controls the prober to move the chuck so that the via to inspect is precisely located under the camera; an image of the via is taken through the video acquisition board, then, the ZISC-036 based PMI:

- finds the via on the image,
- checks the integrity of the border (for damage) of via,
- locates the impact in the via and estimates its surface for statistics.

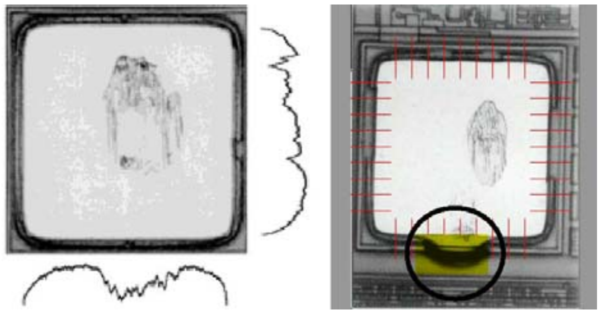


(a) (b)

Fig. 3. Example of profiles extraction after via centring process (a). Example of profiles to category association during the learning phase (b).

All vias of a tested wafer are inspected and analyzed. At the end of the process, the system shows a wafer map which presents the results and statistics on the probe quality and its alignment with the wafer. All the defects are memorized in a log file. In summary, the detection and classification tasks of our PMI application are done in three steps: via localization in the acquired image, mark size estimation and probe impact classification (good, bad or none).

The method, which was retained, is based on profiles analysis using kennel functions based ANN. Each extracted profile of the image (using a square shape, figures 3 and 4) is compared to a reference learned database in which each profile is associated with its appropriated category. Different categories, related to different needed features (as: size, functional signature, etc).



(a) (b)

Fig. 4. Profiles extraction for size and localization of the probe mark (a). Experimental result showing a fault detection and its localization in the via (b).

Experiments on different kinds of chips and on various probe defects have proven the efficiency of the neural approach to this kind of perception problem. The developed intelligent PMI system outperformed the best solutions offered by competitors by 30%: the best response time per via obtained using other wafer probers was about 600 ms and our neural based system analyzes one via every 400 ms, 300 of which were taken for the mechanical movements. Measures showed that the defect recognition neural module's execution time was negligible compared to the time spent for mechanical movements, as well as for the image acquisition (a ratio of 12 to 1 on any via). This application is presently inserted on a high throughput production line.

3 Automated Classification of High-Tech Optical Devices' Defects in Industrial Production

Fault diagnosis in industrial environment is a challenging but crucial task, since it ensures products' nominal specification and manufacturing control. Concerning High-Tech optical industry, a major step for high-quality optical devices' faults diagnosis concerns scratches and digs defects detection and characterization in such products. These kinds of aesthetic flaws, shaped during different manufacturing steps, could provoke harmful effects on optical devices' functional specificities, as well as on their optical performances by generating undesirable scatter light, which could seriously damage the expected optical features. A reliable diagnosis of these defects becomes therefore a crucial task to ensure products' nominal specification. Moreover, such diagnosis is strongly motivated by manufacturing process correction requirements in order to guarantee mass production quality with the aim of maintaining acceptable production yield.

Unfortunately, detecting and measuring such defects is still a challenging problem in production conditions and the few available automatic control solutions remain ineffective. That's why, in most of cases, the diagnosis is performed on the basis of a human expert based visual inspection of the whole production. However, this conventionally used solution suffers from several acute restrictions related to human operator’s intrinsic limitations (reduced sensitivity for very small defects, detection exhaustiveness alteration due to attentiveness shrinkage, operator’s tiredness and weariness due to repetitive nature of fault detection and fault diagnosis tasks). Figure 5 gives an example of High-Tech optical products, showing four optical filters. The same figure shows an example of visual inspection process of the aforementioned defects, requiring expert knowledge and a consequent delay [5].

Fig. 5. Example of High-Tech optical devices performing optical filtering (a) and the visual fault detection, performed by an expert (b).

To construct an automatic diagnosis system, we propose an approach based on three main operations: detection, classification and decision. Our motivation to adopt the approach dissociating detection and diagnosis tasks is based on requirement relative to the frame of industrial production. In fact, two complementary options could de required in industrial production environment. The first is inherent to mass production where it is not always necessary to diagnose whole manufactured products during the production, but it is crucial to detect the presence of defects in order to state if the number of defects is conform to the process’ intrinsic limitations. However, at the same time, diagnosis ability could help to state (offline) if detected defects are due to intrinsic limitations of the used manufacturing process or a number of them correspond to different derivations. The second situation is specific to High-Tech products manufacturing requirements, where additionally to systematic defect detection it is crucial to state on nature of detected defects in order to reach high-quality specifications.

Fig. 6. Block diagram of the proposed optical devices diagnosis system.

To perform this challenging task, we choose to use neural network based techniques, which have shown many attractive features in complex patterns recognition and classifications. The outline of the process, we propose to use in order to carry out the defect classification is shown in the diagram of figure 6. As one could remark, the whole system includes four main stages (tasks): defect detection stage, data extraction module, dimensionality reduction stage and classification module [5]. The detection approach is based on Nomarski’s microscopy issued imaging (NMI) [6], [7], [8], [9]. This method provides robust detection and reliable measurement of outward defects (essentially scratches and digs defects), making plausible a fully automatic inspection of optical products.

The aim of the second stage is to extract defects images from Nomarski detector issued digital image. A new method has been proposed including four phases:

- Pre-processing: Nomarski issued digital image transformation in order to reduce lighting heterogeneity influence and to enhance the aimed defects' visibility,
- Adaptive matching: adaptive process to match defects,
- Filtering and segmentation: noise removal and defects' outlines characterization.
- Defect image extraction: correct defect representation construction.

(a) (b) (c) (d)

Fig. 7. Examples of images of different characteristic items obtained after "data extraction" stage: (a) scratch; (b) dig; (c) dust; (d) cleaning marks.

The dimensionality reduction task consists on constructing a set of lower dimension homogenous invariant (regarding translation and rotation) representation of characteristic items obtained from "data extraction" stage. In fact, processing high-dimensional data may induce several problems among which, the exponentially increase of the number of samples, required to reach a predefined level of precision in approximation tasks, with the dimension of considered space [10]. Concerning a ANN based classification, taking into account the above-mentioned problem, one can intuitively understand that the number of samples needed to properly learning high-dimensional data becomes quickly too large to be collected by a real systems.

A first data dimensionality reduction is performed using "Fourier-Mellin" transformation as it provides invariant descriptors, which are considered to have good coding capacity in classification tasks [11], [12], [13] and [14]. In fact, because of different sizes of items' images and their relative positions (due to translation and rotation) it is necessary to have a "normalize" representation for classification stage's input patterns. After "Fourier-Mellin" transformation each item image obtained from the second stage is represented as a 13-D vector. Then, such 13-D translation-rotation invariant vector is normalized thanks to a centring-reducing transformation, modify ing each feature F_i conformably to the relation (3), where M is the mean value of the feature F_i over the database and σ its standard deviation.

$$F_i = \frac{F_i - M}{\sigma} \tag{3}$$

A second data dimensionality reduction is then performed using a projection based approach from a high-dimensional space to a lower-dimensional space. Several techniques have been studied, among which, Kohonen Self-Organizing Map (SOM), Curvilinear Component Analysis (CCA) and Curvilinear Distance Analysis (CDA).

These aspects are presented and discussed in [15]. SOM is often used in industrial engineering [16], [17] to characterize high-dimensional data or to carry out classification tasks. Its main advantage is to offer an additional "pre-classification" ability. However, it suffers from a number of drawbacks: first the SOM's topology is static and should be fixed a priori; moreover the method defines only a discrete nonlinear subspace; finally algorithm is computationally too expensive to be practically applied for projection space dimension higher than 3 [15], [18]. Concerning CCA [19], if it is able to reproduce the topology of an n-dimension original space in a new p-dimension space (where p<n) without any a priori configuration of the topology (constituting its advantage regarding SOM), it remains essentially a linear projection (its main drawback regarding complex classification tasks) [20]. CDA involves curvilinear distances allowing dealing with non-linear manifolds (its main advantage) [21]. Finally, the classification stage, including an unsupervised learning based "pre-classification" stage and a supervised learning based classifier, performs the defects' classification operation.

Table 1. Validation results relative to the MLP classification performances.

Training database dimensionality	Correct classification	Standard Deviation
13	76 %	1.33 %
2	94 %	0.87 %

The validation of proposed approach has been done using data issued from real industrial production process [15]. Two Multi-Layer Perceptron (MLP) based classifiers, one classifying 13-D data and the other classifying 2-D data have been implemented. The first MLP structure includes13 input neurons, 35 neurons in hidden layer, and 2 output neurons (13-35-2 MLP) and the second one engages 2 input neurons, 35 neurons in hidden layer, and 2 output neurons (2-25-2 MLP). For both of above-indicated structures, the training was achieved 20. Results are presented in Table 4. These results clearly prove that the considered classification problem is simplified, when properly reformulated in an appropriated lower dimensional space.

4 Conclusion and Perspectives

The main goal of this paper was to show how ANN models could be sources of inspiration in emergence of real industrial solutions. The presented applications and issued results show the significant potentiality of connectionist architectures for designing real world applications dealing with complex industrial dilemmas.

Acknowledgements

I would like to express my gratitude to Dr. Veronique Amarger, Dr. Abdennasser Chebira, Dr. Amine Chohra and Dr. Christophe Sabourin, working as staff members of my research team (and members of LISSI Lab.) involved in presented works. Also, I would like to express my gratitude as well to Dr. Ghislain De Tremiolles, who have worked as Ph.D. student on ZISC-036 neuro-processor Based Probe Mark Inspection as to IBM-France that supported this project. In the same way, I would like express gratitude to Mr. Mathieu Voiry my Ph.D. student working with SAGEM on Automated Classification of High-Tech Optical Devices' Defects for his valuable work allowing accomplished progresses in this project. Finally, I would acknowledge SAGEM industrial group for supporting Mr. Voiry and this challenging project.

References

1. G. Trémiolles (de), P. Tannhof, B. Plougonven, C. Demarigny, K. Madani, Visual Probe Mark Inspection, using Hardware Implementation of Artificial Neural Networks, in VLSI Production, Lecture Notes in Computer Science - Biological and Artificial Computation : From Neuroscience to Technology, Edited by : Jose Mira, Roberto M. Diaz and Joan Cabestany - Springer Verlag Berlin Heidelberg, pp 1374-1383, 1997, (1997)
2. K. Madani, G. De Tremiolles, P. Tanhoff, Image processing using RBF like neural networks : A ZISC-036 based fully parallel implementation solving real world and real complexity industrial problems, Journal of Applied Intelligence N°18, 2003, Kluwer Academic Publishers, pp. 195-231, (2004)
3. G. Trémiolles (de), K. Madani, P. Tannhof, A New Approach to Radial Basis Function's like Artificial Neural Networks , NeuroFuzzy'96, IEEE European Workshop, Vol. 6 N° 2, pp 735-745, April 16 to 18, Prague, Czech Republic, (1996)
4. G. De Tremiolles, Contribution à l'étude théorique des modèles neuromimétiques et à leur validation expérimentale: mise en œuvre d'applications industrielles ("Contribution to the theoretical study of neuromimetic models and to their experimental validation: use in industrial applications"), Ph.D. Report, University of PARIS XII – Val de Marne, (1998)
5. K. Madani, M. Voiry, V. Amarger, N. Kanaoui, A. Chohra, F. Houbre "Computer Aided Diagnosis Using Soft-Computing Techniques and Image's Issued Representation : Application to Biomedical and Industrial Fields", COMPUTING, Special Issue on Intelligent Computing, Vol.5, Issue 3, ISSN 1727-6209, pp. 43-53, (2007)
6. P. Bouchareine, Métrologie des Surfaces. Techniques de l'Ingénieur, vol. R1390, 1999 (in French).
7. S. Chatterjee, Design Considerations and Fabrication Techniques of Nomarski Reflection Microscope. Optical Engineering, vol. 42, no. 8, pp. 2202-2212, (2003).
8. M. Voiry, F. Houbre, V. Amarger, and K. Madani, Toward Surface Imperfections Diagnosis Using Optical Microscopy Imaging in Industrial Environment, Workshop on Advanced Control and Diagnosis, Mulhouse, France, pp. 139-144, (2005).
9. M. Voiry, K. Madani, V. Amarger, F. Houbre, Toward Automatic Defects' Clustering in Industrial Production Process Combining Optical detection and Unsupervised Artificial Neural Network techniques, Artificial Neural Networks and Intelligent Information Processing, INSTICC Press, N°ISBN: 978-972-8865-68-9, pp. 25-34, (2006).
10. M. Verleysen: Learning high-dimensional data. LFTNC'2001 - NATO Advanced Research Workshop on Limitations and Future Trends in Neural Computing (2001).

11. A. Choksuriwong, H. Laurent, and B. Emile, Comparison of invariant descriptors for object recognition. IEEE International Conference on Image Processing (ICIP) pp. 377-380, (2005)
12. S. Derrode, Représentation de Formes Planes à Niveaux de Gris par Différentes Approximations de Fourier-Mellin Analytique en vue d'Indexation de Bases d'Images. Phd Thesis Rennes I Univ., (in French) (2005)
13. F. Ghorbel, A Complete Invariant Description for Gray Level Images by the Harmonic Analysis Approach. Pattern Recognition, vol. 15, pp. 1043-1051, (1994)
14. G. Ravichandran and M. Trivedi, Circular-Mellin features for texture segmentation. IEEE Trans. Image Processing, vol. 4, pp. 1629-1640, (1995)
15. M. Voiry, K. Madani, V. AmargerJ. Bernier, Impact of Data Dimensionality Reduction on Neural Based Classification: Application to Industrial Defects, ANNIIP 2007 (to be published in May, 2007).
16. T. Kohonen, E. Oja, O. Simula, A. Visa, and J. Kangas, Engineering Applications of the Self-Organizing Maps. Proceedings of the IEEE, vol. 84, no. 10, pp. 1358-1384, (1996)
17. J. Heikkonen and J. Lampinen: Building Industrial Applications with Neural Networks., Proc.European Symposium on Intelligent Techniques, ESIT'99 (1999).
18. T. Kohonen, Self Organizing Maps, 3rd edition, Berlin: Springer, (2001)
19. P. Demartines and J. Hérault: Vector Quantization and Projection Neural Network, Lecture Notes in Computer Science, vol. 686, International Workshop on Artificial Neural Networks, p.328-333 (1993).
20. P. Demartines and J. Hérault: CCA : "Curvilinear Component Analysis", Proceedings of 15th workshop GRETSI (1995).
21. J. A. Lee, A. Lendasse, N. Donckers, and M. Verleysen: A Robust Nonlinear Projection Method, European Symposium on Artificial Neural Networks - ESANN'2000 (2000).