International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Design and Development of a Deep Learning Model for Electronic Manufacturing Industry Using TeetuNet

Pankaj Kumar Sharma[1], Dr. Sandhya Sharma[2]

Submitted:14/03/2024 **Revised**: 29/04/2024 **Accepted**: 06/05/2024

Abstract: One of the biggest challenges in electronic manufacturing industry is to take care of quality check process of manufacturing semiconductor wafers. Post pandemic the supply chain for electronic chips has disturbed badly and the supply is much less than the demand all over the world. The present systems are not as efficient and smart to reduce the testing time of semiconductor wafer drastically. Many a times it so happens that one or the other kind of defects are present in the manufactured wafer which degrades the quality of wafer and adds up more time to produce same number of wafers. Therefore a much more efficient, reliable quality check system is necessary to tackle the issue. Since the advantage of Artificial Intelligence is in almost every field nowadays, in this paper we propose a much more capable Deep Learning model based on CNN, which can detect a defected and non defected wafer using image. Unlike the present systems where all the testing work is done using sensor data points, this Deep Learning model process on the image of wafer and gives results with greater accuracy.

Keywords: *Artificial Intelligence, Semiconductor wafer, Electronic chip Manufacturing, Deep Learning*

1. Introduction

This template, modified in MS Word 2007 and saved as a "Word 97-2003 Document" for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers.

Semiconductor wafers are the backbone of every electronic item used in multiple products today. Be it Healthcare, Education, Communication, E-commerce, Satellite technology, Space science, Automobile industry, Hospitality, Public Administration, Aerospace, Solar technologies, optics, these wafers are extensively used in every field as part of different multipurpose devices and apparatus. Semiconductor wafer manufacturing is a very complex process which comprises of many steps such as Slicing, Lapping, Etching, Polishing, Doping, Cleaning and Inspection.

Wafer inspection and quality check step comes at the last and plays a significant role in deciding the defected wafer. At present many technical procedures are applied to keep this process quicker and efficient but most of the processes are not much efficient [1]. Current procedures involves multiple sensors to detect various parameters of the wafer and by analyzing the sensor data on a connected system the system further tells about the authenticity of the wafer. Such kind of numeric data processing requires high level of computing resources. At one end such system requires time and demanding resources to be installed at the facility on the other end they also need to be supplied with high amount of processing power in terms of CPU or GPU further huge amount of RAM is also required to process such big volume of data. Even after that the current systems are not that efficient and time consuming. The novel idea in this paper is to use the recent developments in the field of Artificial Intelligence and Deep Learning to enhance the productivity of wafer inspection facility. This paper proposes a state of the art Deep Learning model which can detect a wafer defect by processing its image using Convolution Neural Network (CNN).

1.1. Deep Learning

Artificial Intelligence, Machine Learning and Deep Learning are the most popular terms in the area of automation today. Deep learning is a subset of Machine Learning which further is a part of Artificial Intelligence. AI is a broader term which brings the human intelligence and human behavior to machine and computerized systems. Machine Learning is a method which learns by using already available data values and used in finding patterns and building predictive modeling, automatic analytic model building etc. Deep Learning works on bigger volume of data involving complex computation at a very high level [12]. It uses multi layer neural networks to process and compute the data to build data-driven intelligent models.

2. Literature Review

Arnel C. Fajardo et al [1], proposed a CNN model which was trained on different images of defective and non defective PCB wafers and performed efficiently in

[1] Dept. of ECE, Suresh Gyan Vihar University, Jaipur (Raj) India
ORCID ID : 0009-0003-6411-1308
[2] Dept. of ECE, Suresh Gyan Vihar University, Jaipur (Raj) India
ORCID ID : 0000-0002-5452-8658
** Corresponding Author Email: Sharma.pankaj91@gmail.com*

detecting the kind of defect present in each wafer. The defect classes are spur, scratches, pinholes, mouse bites, excessive conductor, missing conductor, copper track.

Zhijiang Xiong et al [2], proposed a PCB defect detection algorithm based on YOLO v8. This proposed work showed 97% accuracy in defect detection. It has also proposed a defect detection user interface for easy determination.

JunYi Lim et al [3], studied the tradeoff between accuracy and time for the available methods and proposed a new deep learning network which works efficiently for tiny defects on PCB. It is made up of multi scale feature pyramid network and stress upon union loss function. It has achieved mean average precision of 99.17%.

Rey-chue hwang et al [4], used AOI i.e. area of interest based calculations in model where the authors have pressed on automatic optical inspection which has proved to be the best for defect detection. The dataset was taken from industry and the overall model deals with the problem of poor imaging because if poor lighting and shadows. A very few parameters are used in making of model and the accuracy given is 95%, recall is 94%, detection time is 0.027 seconds.

Marcos Antonio Andrade et al [5], proposed a system for PCB defect detection using visual computing and deep learning for production optimization. In this work they have merged traditional computing with newer deep learning technique and the proposed model has achieved greater than 90% accuracy.

Bixian Feng et al [6], worked on variable defects where they have used global information while emphasizing on very small area and diverse style. Defect Detection Transformer is used which is designed with the base of residual swin transformer. They have focused on multi head self attention mechanism for feature extraction. Experiments were conducted on PKU-market-PCB and DeepPCB dataset

Zhewei Liu et al [7], proposed a cost sensitive DL model based on siamese network by using optimization algorithm NSGA-II to check for cost sensitive threshould value. This has produced exciting results in terms of accuracy of prediction. The official predicting accuracy claimed is 97.6% and the shoving average is given 33.32% in training.

Ye Chan Kim et al [8], developed an automated system for wafer defect detection and localize them with accuracy. The images are first scanned with line space patterns and then a deep learning neural network is designed and used to localize the defects. Model is optimized for design variables before training, after that data augmentation is used to enhance the data sample.

After a detailed survey of literature pertaining to already available methods and deep learning models for wafer defect detection it is important to point out that many models are performing efficiently with good accuracy of prediction. But there are still some significant issues and research gaps present which must be addressed.

Some instances of research gaps are as follows:

Image De-noising – it is important to remove all noise form images before processing them in to the model. The image data with noisy images cannot give better results in terms of prediction. De-noising the images helps the model to generalize effectively.

Accuracy – achieved accuracy is good but can be improved for better detection.

Model size- most of the studied models are very heavy in size. They go in GBs a significant improvement is needed to reduce the size of model.

Computation cost- the number of parameters in the studied models is way too high. It requires high computational cost to process the dataset on certain number of parameters. A light weight model with lesser parameters is required to bring down the computation cost in terms of GPU, RAM and processor.

Problem Statement and Objective

We are provided with semiconductor wafer images from industry and by using these images we have to make a deep learning model which can detect defective and non defective wafers with greatest accuracy. The outcome of this will be a self sufficient deep learning model which can predict defected or non defected wafer within microseconds.

3. Solving Approach- Implementation of Base Model

Use either SI (MKS) or CGS as primary units. (SI units are strongly encouraged.) English units may be used as secondary units (in parentheses). This applies to papers in data storage**.**

We already have pre-trained network with us which has been trained on thousand of images and posses the weights and bias data within. Efficient Net Algorithm is one such Algorithm which has given good results in image recognition tasks. It is easy to implement with open source availability of resources.

Efficient Net

Efficient Net was created by Google AI research team [13]. It uses many layers of convolution in order to compute the weights and biases. It is based on compound scaling and Neural Architectural Search (NAS) these two things works together in fusion to achieve best performance and resource efficiency. Compound scaling is achieved by bringing down the parameters and

Flops (Floating point Operations per Second) on the other hand it uses NAS to create the baseline Architecture [13].

Table 1. Accuracy and Loss Results

Model	*Validation Loss*	*Validation Accuracy (%)*	*Test Loss*	*Test Accuracy(%)*
EfficientNetB0	1.05	73.68	1.03	73.68
EfficientNetB1	0.56	73.68	0.48	81.57
EfficientNetB2	1.08	73.68	0.88	73.68
EfficientNetB3	2.18	26.32	2.42	26.31
EfficientNetB4	1.06	73.68	0.46	76.31
EfficientNetB5	1.63	73.68	1.01	73.68
EfficientNetB6	0.36	81.58	0.23	92.10
EfficientNetB7	1.87	63.16	0.67	78.94
EfficientNetV2B0	1.39	73.68	1.15	73.68
EfficientNetV2B1	1.40	73.68	1.24	73.68
EfficientNetV2B2	0.53	76.32	0.13	94.73
EfficientNetV2B3	0.48	76.32	0.13	94.73
EfficientNetV2S	2.15	89.47	0.53	94.73

Compound Scaling

There are three types of scaling – Depth Scaling, Width Scaling, Resolution Scaling

Depth $D = \alpha$

Width $W = \beta \wedge \Phi$

Resolution $R = \gamma \wedge \Phi$

$F = DWR = \alpha \ \ \beta \wedge \Phi \ \gamma \wedge \Phi$, where $\alpha = 1.20$, $\beta = 1.10$, $\gamma = 1.15$

Value of Φ comes from Grid Search, F is scaling factor which is used to upscale the network

Because of its automatic scaling feature this Algorithm is performing nicely in terms of Accuracy and Loss. Here we are looking for maximum Accuracy and Minimum Loss for our deep learning model. There are total of 15 Algorithms pertaining to Efficient Net wherein there are 8 belongs to base version and 4 belongs to 2nd version rest to later versions. Each Algorithm has its own unique architecture and different number of parameters, size, Top 1% Accuracy, Top 5% Accuracy as mentioned in the official documents.

Implementation

We have the dataset of images of both kinds defected and Non defected so these images are used to pass through Algorithms. All images are divided into three parts for training set, validation set and test set, each set is further divided into 2 parts as defected and Non defected. These parts are treated as two classes and problem is treated as "**Binary classification Problem**" where Defected = class '0', Non defected = class '1'

3.1. Types of Graphics

The following list outlines the different types of graphics published in IEEE journals. They are categorized based on their construction, and use of color / shades of gray:

3.1.1. Color/Grayscale figures

Figures that are meant to appear in color, or shades of black/gray. Such figures may include photographs, illustrations, multicolor graphs, and flowcharts.

3.1.2. Line Art figures

Figures that are composed of only black lines and shapes. These figures should have no shades or half-tones of gray, only black and white.

3.1.3. Tables

Data charts which are typically black and white, but sometimes include color.

3.2. Multipart figures

Figures compiled of more than one sub-figure presented side-by-side, or stacked. If a multipart figure is made up of multiple figure types (one part is lineart, and another is grayscale or color) the figure should meet the stricter guidelines.

3.3. File Formats For Graphics

Format and save your graphics using a suitable graphics processing program that will allow you to create the images as PostScript (PS), Encapsulated PostScript (.EPS), Tagged Image File Format (.TIFF), Portable Document Format (.PDF), or Portable Network Graphics (.PNG) sizes them, and adjusts the

As shown in Results 15 different Deep Learning models were built using 15 Algorithms of EfficientNet series but none of the above models gave desired results in terms of prediction Accuracy and loss. We are looking for maximum Accuracy and minimum loss, among all the above models only **EfficientNetV2B3** gave best result as Test Accuracy of **94.73** and Test Loss of **0.13**. That is the best result we are able to achieve by this Algorithm.

resolution settings. If you created your source files in one of the following programs you will be able to submit the graphics without converting to a PS, EPS, TIFF, PDF, or PNG file: Microsoft Word, Microsoft PowerPoint, or Microsoft Excel. Though it is not required, it is strongly recommended that these files be saved in PDF format rather than DOC, XLS, or PPT. Doing so will protect your figures from common font and arrow stroke issues that occur when working on the files across multiple platforms. When submitting your final paper, your graphics should all be submitted individually in one of these formats along with the manuscript.

3.4. Vector Art

In order to preserve the figures' integrity across multiple computer platforms, we accept files in the following formats: .EPS/.PDF/.PS. All fonts must be embedded or text converted to outlines in order to achieve the best-quality results.

3.5. Color Space

The term color space refers to the entire sum of colors that can be represented within the said medium. For our purposes, the three main color spaces are Grayscale, RGB (red/green/blue) and CMYK (cyan/magenta/yellow/black). RGB is generally used with on-screen graphics, whereas CMYK is used for printing purposes.

All color figures should be generated in RGB or CMYK color space. Grayscale images should be submitted in Grayscale color space. Line art may be provided in grayscale OR bitmap colorspace. Note that "bitmap colorspace" and "bitmap file format" are not the same thing. When bitmap color space is selected, .TIF/.TIFF/.PNG are the recommended file formats.

Fig 2. Images showing the visual symptoms cause by fungal disease

3.6. Accepted Fonts Within Figures

When preparing your graphics IEEE suggests that you use of one of the following Open Type fonts: Times New Roman, Helvetica, Arial, Cambria, and Symbol. If you are supplying EPS, PS, or PDF files all fonts must be embedded. Some fonts may only be native to your operating system; without the fonts embedded, parts of the graphic may be distorted or missing.

A safe option when finalizing your figures is to strip out the fonts before you save the files, creating "outline" type. This converts fonts to artwork what will appear uniformly on any screen.

3.7. Using Labels Within Figures

3.7.1. Figure Axis labels

Figure axis labels are often a source of confusion. Use words rather than symbols. As an example, write the quantity "Magnetization," or "Magnetization *M*," not just "*M*." Put units in parentheses. Do not label axes only with units. As in Fig. 1, for example, write "Magnetization (A/m)" or "Magnetization ($A \cdot m^{-1}$)," not just "A/m." Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)," not "Temperature/K."

Multipliers can be especially confusing. Write "Magnetization (kA/m)" or "Magnetization (10^3 A/m)." Do not write "Magnetization (A/m) × 1000" because the reader would not know whether the top axis label in Fig. 1 meant 16000 A/m or 0.016 A/m. Figure labels should be legible, approximately 8 to 10 point type.

3.7.2. Subfigure Labels in Multipart Figures and Tables

Multipart figures should be combined and labeled before final submission. Labels should appear centered below each subfigure in 8 point Times New Roman font in the format of (a) (b) (c).

3.7.3. Referencing a Figure or Table Within Your Paper

When referencing your figures and tables within your paper, use the abbreviation "Fig." even at the beginning of a sentence. Do not abbreviate "Table." Tables should be numbered with Roman Numerals.

4. Conclusion

A conclusion might elaborate on the importance of the work or suggest applications and extensions.

4.1. Appendix

Appendixes, if needed, appear before the acknowledgment.

5. References and Footnotes

5.1. References

References need not be cited in text. When they are, they appear on the line, in square brackets, inside the punctuation. Multiple references are each numbered with separate brackets. When citing a section in a book, please give the relevant page numbers. In text, refer simply to the reference number. Do not use "Ref." or "reference" except at the beginning of a sentence: "Reference [3] shows ... ." Please do not use automatic endnotes in Word, rather, type the reference list at the end of the paper using the "References" style.

Reference numbers are set flush left and form a column of their own, hanging out beyond the body of the reference. The reference numbers are on the line, enclosed in square brackets. In all references, the given name of the author or

editor is abbreviated to the initial only and precedes the last name. Use them all; use *et al*. only if names are not given. Use commas around Jr., Sr., and III in names. Abbreviate conference titles. When citing IEEE transactions, provide the issue number, page range, volume number, year, and/or month if available. When referencing a patent, provide the day and the month of issue, or application. References may not include all information; please obtain and include relevant information. Do not combine references. There must be only one reference with each number. If there is a URL included with the print reference, it can be included at the end of the reference.

Other than books, capitalize only the first word in a paper title, except for proper nouns and element symbols. For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation See the end of this document for formats and examples of common references.

Acknowledgements

This research was supported/partially supported by [Name of Foundation, Grant maker, Donor]. We thank our colleagues from [Name of the supporting institution] who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper. We thank [Name Surname, title] for assistance with [particular technique, methodology], and [Name Surname, position, institution name] for comments that greatly improved the manuscript.

Author contributions

Mr. Pankaj Kumar Sharma: Data collection,Conceptualization, Methodology, Field study,Writing-Original draft preparation

Dr. Sandhya Sharma: Data curation, Software, Investigation,Validation & Testing., Reviewing ,Editing and finalizing the manuscript

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] 2nd ed., vol. 3, J . Peters, E d . New Y o r k , NY, USA: McGraw-Hill,1964,pp.15–64.

[2] W.-K. Chen, Linear Networks and Systems. Belmont, CA, USA: Wadsworth, 1993, pp. 123–135.

[3] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," *IEEE Trans. Electron Devices*, vol. ED-11, no. 1, pp. 34–39, Jan. 1959, 10.1109/TED.2016.2628402.

[4] E. P. Wigner, "Theory of traveling-wave optical laser," *Phys. Rev*., vol. 134, pp. A635–A646, Dec. 1965.

[5] E. H. Miller, "A note on reflector arrays," *IEEE Trans. Antennas Propagat*., to be published.

[6] E. E. Reber, R. L. Michell, and C. J. Carter, "Oxygen absorption in the earth's atmosphere," Aerospace Corp., Los Angeles, CA, USA, Tech. Rep. TR-0200 (4230-46)-3, Nov. 1988.

[7] J. H. Davis and J. R. Cogdell, "Calibration program for the 16-foot antenna," Elect. Eng. Res. Lab., Univ. Texas, Austin, TX, USA, Tech. Memo. NGL-006-69-3, Nov. 15, 1987.

[8] *Transmission Systems for Communications*, 3rd ed., Western Electric Co., Winston-Salem, NC, USA, 1985, pp. 44–60.

[9] *Motorola Semiconductor Data Manual*, Motorola Semiconductor Products Inc., Phoenix, AZ, USA, 1989.

[10] G. O. Young, "Synthetic structure of industrial plastics," in Plastics, vol. 3, Polymers of Hexadromicon, J. Peters, Ed., 2nd ed. New York, NY, USA: McGraw-Hill, 1964, pp. 15-64. [Online]. Available: http://www.bookref.com.

[11] *The Founders' Constitution*, Philip B. Kurland and Ralph Lerner, eds., Chicago, IL, USA: Univ. Chicago Press, 1987. [Online]. Available: http://press-pubs.uchicago.edu/founders/

[12] The Terahertz Wave eBook. ZOmega Terahertz Corp., 2014. [Online]. Available: http://dl.z-thz.com/eBook/zomega_ebook_pdf_1206_sr.pdf. Accessed on: May 19, 2014.

[13] Philip B. Kurland and Ralph Lerner, eds., *The Founders' Constitution.* Chicago, IL, USA: Univ. of Chicago Press, 1987, Accessed on: Feb. 28, 2010, [Online] Available: http://press-pubs.uchicago.edu/founders/

[14] J. S. Turner, "New directions in communications," *IEEE J. Sel. Areas Commun*., vol. 13, no. 1, pp. 11-23, Jan. 1995.

[15] W. P. Risk, G. S. Kino, and H. J. Shaw, "Fiber-optic frequency shifter using a surface acoustic wave incident at an oblique angle," *Opt. Lett.*, vol. 11, no. 2, pp. 115–117, Feb. 1986.

[16] P. Kopyt*et al., "*Electric properties of graphene-based conductive layers from DC up to terahertz range," *IEEE THz Sci. Technol.,* to be published. DOI: 10.1109/TTHZ.2016.2544142.

[17] PROCESS Corporation, Boston, MA, USA. Intranets: Internet technologies deployed behind the firewall for corporate productivity. Presented at INET96AnnualMeeting.[Online]. Available:

http://home.process.com/Intranets/wp2.htp

[18] R. J. Hijmans and J. van Etten, “Raster: Geographic analysis and modeling with raster data,” R Package Version 2.0-12, Jan. 12, 2012. [Online]. Available: http://CRAN.R-project.org/package=raster

[19] Teralyzer. Lytera UG, Kirchhain, Germany [Online]. Available: http://www.lytera.de/Terahertz_THz_Spectroscopy.php?id=home, Accessed on: Jun. 5, 2014

[20] U.S. House. 102nd Congress, 1st Session. (1991, Jan. 11). *H. Con. Res. 1, Sense of the Congress on Approval of Military Action*. [Online]. Available: LEXIS Library: GENFED File: BILLS

[21] Musical toothbrush with mirror, by L.M.R. Brooks. (1992, May 19). Patent D 326 189 [Online]. Available: NEXIS Library: LEXPAT File: DES

[22] D. B. Payne and J. R. Stern, “Wavelength-switched passively coupled single-mode optical network,”in *Proc. IOOC-ECOC,* Boston, MA, USA,1985,
pp.585–590.

[23] D.Ebehardand E. Voges,“Digital single side band detection for interferometric sensors,”presented at the*2ndInt. Conf.Optical Fiber Sensors,* Stuttgart, Germany,Jan.2-5, 1984.

[24] G. Brandli and M. Dick, “Alternatingcurrent fed power supply,”U.S.Patent 4 084 217,Nov.4,1978.

[25] J. O. Williams, “Narrow-band analyzer,” Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, USA, 1993.

[26] N. Kawasaki, “Parametric study of thermal and chemical nonequilibrium nozzle flow,” M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.