Sensors and Materials, Vol. 37, No. 5 (2025) 1785–1807
MYU Tokyo

S & M 4018

Enhancing Gallium Arsenide Semiconductor Chip Quality Classification through AI Algorithms

Guofeng Luo,[1] Hsiao-Yi Lee,[1*] and Kenvi Wang[2]

[1]Department of Electrical Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 80778, Taiwan
[2]Department of Computing and Mathematics, Liverpool John Moores University, U.K.

(Received September 27, 2024; accepted April 23, 2025)

Keywords: gallium arsenide, AI, machine learning, artificial neural network, accuracy

Deep learning algorithms, such as those employed by CHATGPT/artificial neural networks, have served as transformative catalysts across various sectors, significantly impacting our daily lives. Within the realm of semiconductor processes, the metric of yield stands as a crucial performance benchmark directly influencing the sustained success of enterprises. Among batches of chips, inherent disparities in quality persist, with each chip grade necessitating an appropriate operational environment. To preempt the pitfalls of customer returns or factory repairs, thereby mitigating manufacturing costs, in this study, we embark on a comprehensive exploration utilizing AI algorithms for quality classification. In pursuit of bolstering the competitive edge of semiconductor firms, the refinement of gallium arsenide chip quality classification to augment yield emerges as a paramount concern. Moreover, in this research endeavor, we seek to discern the physical parameters underlying GaAs defect products, offering pivotal insights to enhance the manufacturing process—an aspect poised to be a pivotal focal point.

1. Introduction

The integration of AI and semiconductor technology plays a crucial role in current technological innovations and applications. This combination has led to many significant applications and impacts. Some of the key aspects include smart devices and IoT, autonomous driving technology, smart cities and intelligent transportation, medical diagnosis and treatment, smart manufacturing, and Industry 4.0. The integration of AI and semiconductor technology is driving the digital transformation of many industries, creating numerous new business opportunities and solutions. Gallium arsenide (GaAs) is an important semiconductor material with excellent electronic properties; thus, it has wide applications in multiple technological fields. The demand for semiconductor chips has shown a sustained and rapid growth trend, which is not only a natural result of technological development but also a product of the simultaneous evolution of multiple industries. For example, the global deployment of 5G

*Corresponding author: e-mail: leehy@nkust.edu.tw
https://doi.org/10.18494/SAM5151

ISSN 0914-4935

technology has propelled the demand for high-performance semiconductor chips owing to the rapid increase in communication speed and the transmission of large amounts of data. The application scope of this technology covers smartphones, IoT devices, and the construction of smart cities, all of which continuously drive the growth of the semiconductor market. Furthermore, the development of AI and machine learning (ML) requires powerful processing capabilities and high-performance chips to accomplish complex calculations and analyses.[1,2] This demand motivates semiconductor companies to continuously research and develop more advanced process technologies and design architectures to meet the continuously upgrading market demands. Additionally, with the rise of electric vehicles and the upgrade of traditional automotive electronics, the automotive industry's demand for highly intelligent and electrified components continues to rise. The rapid development of in-car computers, sensor systems, and self-driving technology all require more powerful semiconductor component support. The widespread adoption of remote work and remote learning, especially in the aftermath of the COVID-19 pandemic, further drives the demand for computer equipment, from laptops to cloud servers. The application scope of semiconductor chips continues to expand in response to people's desire for more convenient and efficient digital work and learning. In general, the sharp increase in demand for semiconductor chips is a natural response to the development of the digital age. The continuous advancement of technology drives changes in various aspects of society, and semiconductors, as the foundation of technology, play a crucial role in connecting and driving these changes. In the future, with the emergence of newer technologies, the demand for semiconductor chips is expected to continue to grow[3,4] (Appendix A). The yield of chips is calculated by subtracting the number of contaminated chips and the number of boundary-cut chips from the total number of chips on a wafer, and then dividing the result by the total number of chips on the wafer. Although contaminated portions are screened before the wafer is processed into chips, there still exists variability in the quality levels of chips deemed to be of good quality. Therefore, in this study, we aim to utilize the AI/artificial neural network (ANN) clustering algorithm for semiconductor GaAs chip quality classification. Through a thorough analysis of chip characteristics, we aim to enhance accuracy and efficiency. This ensures that manufactured GaAs chips meet certain quality standards before delivery to customers, mitigating returns and product recalls, reducing production costs, and enhancing customer satisfaction. ANN's capability to learn complex patterns makes it indispensable in both telecommunications and aerospace applications. In telecommunications, it ensures network reliability, efficiency, and scalability in an era of increasing connectivity. In the aerospace sector, ANN enhances safety, autonomy, and precision in mission-critical systems where failure is not an option. These applications highlight the transformative impact of ANN on modern technology, driving advancements in connectivity, automation, and safety.

In the field of semiconductor processes, yield is a critically important performance indicator that directly affects the long-term success of enterprises. Within a batch of chips, there are still differences in quality, and each chip quality has its suitable operating environment. In this paper, we start with the following objectives, namely, to send chips of appropriate quality to suitable places, thus avoiding the risk of customer returns or factory repairs that increase manufacturing costs, and to conduct in-depth research using the AI/ML algorithm for quality classification.

Hence, to promote the competitiveness of semiconductor companies, enhancing the accuracy of GaAs chip quality classification to improve yield has become an important issue. Furthermore, we will identify the physical parameters of GaAs defective products to improve the fabrication process, which will also be an important highlight. Additionally, the results of this study should be used in conjunction with advanced inspection instruments such as X-ray machines in order to truly achieve the goal of improving semiconductor manufacturing processes.

2. Background

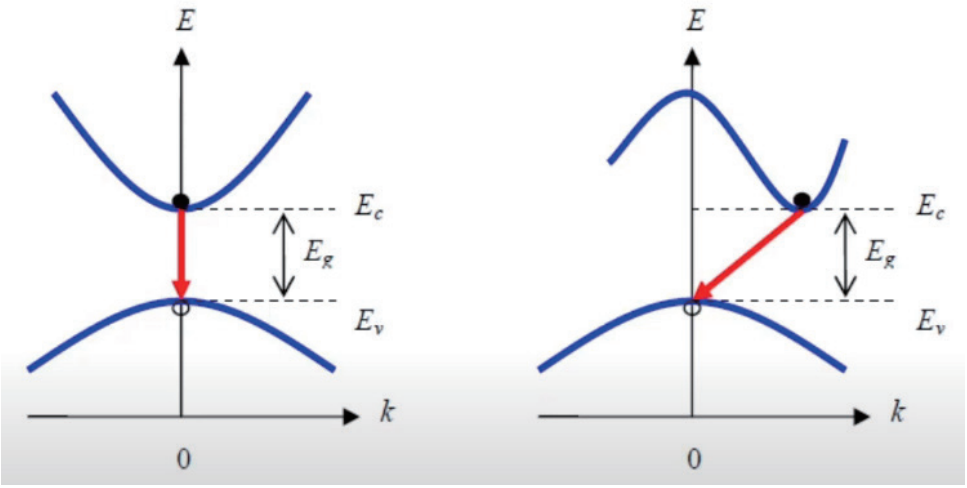
2.1 Physical properties of GaAs

GaAs is a type of III-V semiconductor material with various applications, particularly in semiconductor processes. Its main advantage lies in its excellent electronic properties. Compared with silicon (Si) semiconductors, GaAs semiconductors have a higher electron mobility, making them superior in high-frequency applications such as RF and microwave components. This high-speed characteristic makes it an ideal choice for many communication and radar systems applicable in high-frequency signal processing. Additionally, GaAs has a wider band gap than silicon, which allows it to exhibit superior high-temperature stability. This enables GaAs to maintain high performance even in high-temperature environments, which is crucial for special applications such as space probes and communication equipment in high-temperature environments. Furthermore, GaAs is widely used in semiconductor lasers and optoelectronic devices. GaAs semiconductor lasers have narrow emission spectra and high conversion efficiency, playing important roles in optical communication, fiber optic communication, and laser radar, among other fields.[5,6]

GaAs photodiodes are extensively used in optical communication and high-speed data transmission. Their high-speed switching makes them an ideal choice for high-frequency applications. Apart from its outstanding performance in high-frequency applications and optoelectronic devices, GaAs also excels in high-power and high-frequency microwave components. It exhibits good performance in microwave power amplifiers and power amplifiers for high-frequency wireless communication. This makes GaAs indispensable in fields such as military communication, radar systems, and satellite communication.

In summary, GaAs is an ideal semiconductor material in various application areas owing to its excellent electronic transport properties, high-temperature stability, and optoelectronic properties. Its applications range from communication and optoelectronic devices to microwave components and high-power applications, offering more physical advantages than mainstream semiconductor materials. However, it comes with the corresponding drawback of being relatively expensive to produce.

In Fig. 1, owing to gallium arsenide's direct band gap property, when electrons transition from the valence band to the conduction band, they only need to absorb energy without changing momentum. This is in contrast to silicon's indirect band gap, where not only energy absorption but also momentum change is required.



(a) (b)

Fig. 1. (Color online) (a) Direct band gap and (b) indirect band gap.

Gallium arsenide offers advantages in certain high-frequency and optoelectronic applications owing to its superior properties, whereas silicon remains the dominant material in mainstream electronics owing to its widespread availability and lower cost (Table 1). Silicon is the backbone of the microelectronics industry owing to its cost-effectiveness, mature technology, and suitability for a wide range of applications, including ICs, microprocessors, and mainstream solar cells. Gallium arsenide excels in high-speed, high-frequency, and optoelectronic applications. Its properties make it ideal for use in RF and microwave circuits, LEDs, laser diodes, and high-efficiency solar cells. Each material has its strengths and is chosen on the basis of the specific requirements of the application.

2.2 Methods for testing the quality of gallium arsenide wafers

GaAs wafers play a crucial role in the modern semiconductor industry, and their quality is paramount in ensuring the performance of devices. Therefore, assessing and determining the quality of wafers is necessary, and common methods and standards for wafer quality inspection include the following.[7,8]

- **Band gap and spectral properties**: Photoluminescence and electroluminescence methods among others are used to measure the band gap and ensure that the material's optical properties meet expectations. This helps in detecting impurities or defects in the wafers.
- **Crystal structure detection**: X-ray diffraction or other crystallographic methods are used to ensure that GaAs wafers have the correct crystal structure. The correct crystal structure is one of the important factors in ensuring material stability and performance.
- **Electrical performance assessment**: Electrical performance, including conductivity, carrier concentration, and mobility, are assessed. These electrical performance measurements can be achieved through Hall effect tests or electrical performance testing and are conducted to ensure the performance of GaAs wafers in electronic devices.

Table 1
GaAs and Si properties.

Property	GaAs	Si
Electron mobility	8500 m^2/Vs	1000 m^2/Vs
Band gap	1.42 eV	1.1 eV
Band gap type	Direct	Indirect
Maximum operating temperature	Above 200 °C	Below 150 °C
Optical conversion	Luminescent	Weak emission

- **Surface quality**: Surface analysis tools such as optical microscopes and atomic force microscopes are used to inspect the flatness and surface defects of wafers, ensuring good surface quality.
- **Process control and consistency**: For large-scale manufacturing, ensuring process consistency and control is essential. Therefore, process equipment quality control is regularly tested to ensure the consistent quality of manufactured GaAs wafers.
- **Device performance testing**: The performances of devices such as high-frequency amplifiers and photodiodes are tested to ensure that GaAs wafers perform as required in practical applications.

In this comprehensive assessment, the quality of GaAs wafers can be considered from multiple aspects, including physical properties, optical properties, and electrical properties, as well as surface and process consistency. Such evaluations require the use of various testing and analysis methods, as well as specialized laboratory equipment and expertise. The quality of semiconductor wafers has profound implications in the performance, stability, and reliability of modern technology and electronic devices. High-quality semiconductor wafers not only provide high computational speed and superior processing performance but also ensure the stable operation and long-term reliability of electronic devices. Conversely, low-quality wafers may lead to decreased system performance, stability issues, and even affect the system's overall reliability.[9,10] Therefore, the performance of semiconductor wafers is directly influenced by their quality. High-quality wafers typically offer fast computational speed and superior processing capabilities, crucial for handling complex computational tasks and demanding application scenarios. Secondly, low-quality wafers may cause system instability. Electronic devices rely on various chips to work together; if one of them is of low quality, it may cause system errors, unpredictable failures, or even system crashes. Such instability can significantly inconvenience users, especially in application scenarios with high requirements for system stability, such as medical equipment or autonomous vehicles. Wafer quality is also directly related to energy efficiency. Low-quality wafers may generate more heat, requiring more heat dissipation and energy management, which affects the overall system's energy efficiency. Conversely, high-quality wafers typically maintain low power consumption while providing high performance. Additionally, low-quality wafers may lead to increased manufacturing costs. Increased failure rates in production lines will increase costs, and repairing or replacing defective wafers will also raise maintenance costs. Conversely, high-quality wafers not only help reduce production costs but also contribute to improving production efficiency.[11,12]

3. Design of Experimental Architecture

A wafer is formed by pulling a high-purity semiconductor material into cylindrical ingots, which are then sliced thinly to less than one millimeter thick through chemical mechanical polishing. Wafers come in various sizes, with a maximum diameter of approximately 30 cm.[13,14] Taking a fin-type field-effect transistor as an example, thin-film layers ranging from several nanometers to hundreds of nanometers are first deposited on the surface of the wafer. A photoresist is then applied, and a photomask with circuit design patterns is used to expose the wafer to ultraviolet light, which is then reduced to nanoscale dimensions using lenses. Chemical reactions occur in the exposed areas, followed by dissolution and etching to remove unnecessary parts. This process is repeated several times to create transistors, and by repeating deposition, photolithography, and etching, complex circuits can be created on the wafer.[15,16] After the circuit fabrication is completed, the wafer is diced into several individual chips, as shown in Fig. 2.

Figure 2 shows that after the circuit fabrication is completed, the wafer is cut into several individual chips. The wafer and chip fabrication process in semiconductor manufacturing involves several critical steps, as explained earlier, to transform raw materials into functional electronic devices.

Figure 3 shows various quality ranks of GaAs chips on the wafer. Detecting and minimizing defects is crucial for ensuring the quality and reliability of semiconductor devices. Various inspection and metrology techniques, such as optical inspection, electron microscopy, and electrical testing, are employed throughout the manufacturing process to identify and mitigate defects.

3.1 Physical specifications of GaAs wafer

The physical properties of GaAs chips produced in the GaAs manufacturing process are described in detail below. The explanation of the meaning of each item's physical characteristics shown in Table 2 is as follows.

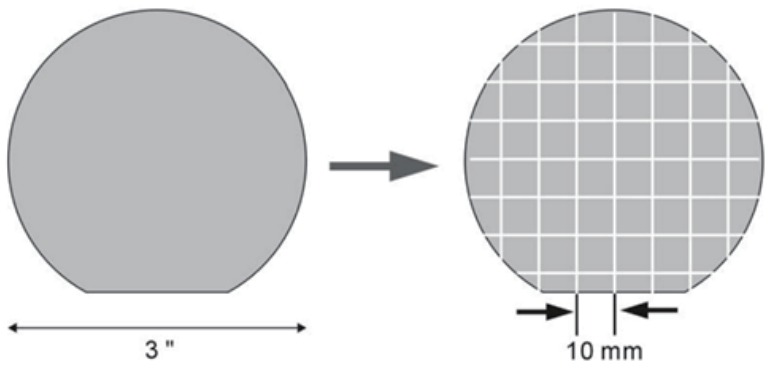


Fig. 2. Wafer and chip fabrication.

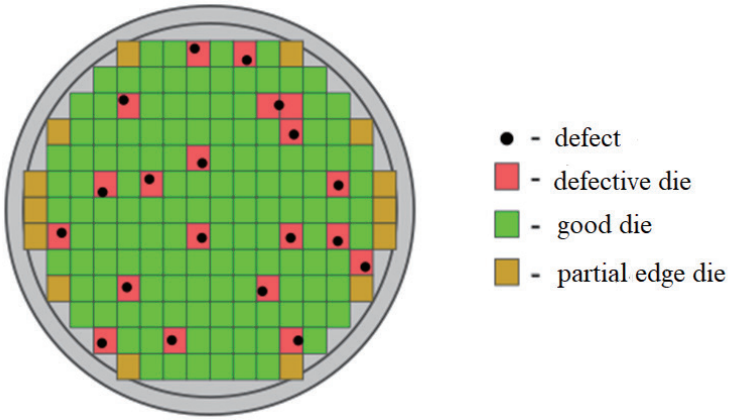


Fig. 3. (Color online) Various types of GaAs chip on the wafer.

Table 2
Standard specifications for physical properties of GaAs chips.

Number	Physical property	Standard specification
01	Positive photoresist thickness	(Spec 4750 ± 10% Å)
02	Silicon nitride (Si_3N_4)	(Spec N Value 2.0 ± 2)
03	Back metallization	(Spec 4750 ± 10%Å)
04	Mesa depth	(Spec 44000 ± 5000 Å)
05	Bonding wire thickness	(Spec 15600 ± 10% Å)
06	Photoresist thickness (PI)	(Spec > 40000 Å)
07	Oxide aperture	(Spec Length 10 ± 2 μm, Width 13 ± 2μm)
08	Threshold current (Ith)	(Spec 1 ± 0.3mA)
09	Operating voltage (Vf@9 mA)	(Spec 1.9 ± 0.3 V)
10	Operating power (Pf@9 mA)	(Spec 5.0 ± 1 mW)

- **"Positive Photoresist Thickness":** metals, including titanium (Ti) and gold (Au), are deposited on the substrate surface to a thickness of approximately 4750 Å in the GaAs semiconductor process. The above combination of metal layers is commonly used in semiconductor processes to manufacture metal wires, electrodes, or other electronic components. The Ti layer is typically used to provide good adhesion between the metal and the semiconductor, and the specific thickness of the positive photoresist may vary depending on process requirements.
- **"Silicon Nitride (Si_3N_4)"** is deposited on the substrate surface in the GaAs semiconductor process. Silicon nitride is a compound composed of silicon and nitrogen and is known for its excellent insulating properties and chemical stability. Therefore, it is commonly used as an insulating layer or protective layer in semiconductor processes to provide insulation, protect the substrate, or serve other specific purposes.
- **"Back Metallization"** refers to the deposition of metals, including gold germanium (AuGe) and gold (Au), on the back of the device to a thickness of 4750 Å in the GaAs semiconductor process. This combination of metal layers is typically used in semiconductor processes to manufacture electrodes or other structures related to the back of the device. The gold

germanium alloy layer is related to the electrical or thermal characteristics of the device and is commonly used at the interface between the metal and the semiconductor to improve their adhesion.

- **"Mesa Depth"** refers to the depth of the mesa structure in semiconductor devices and is used to define individual regions of the chip where specific devices are located. The mesa is an elevated structure surrounded by an insulating material. For GaAs semiconductors, mesa depth is crucial to device performance as it affects optical, electrical, and overall functional characteristics. The specific requirements of mesa depth may vary depending on the type of device being manufactured and the desired characteristics.
- **"Bonding Wire Thickness":** metals, including chromium (Cr) and gold (Au), are deposited on the device to a thickness of 15600 Å. This combination of metal layers is typically used to manufacture bonding wires or conductors, especially for connecting electronic components between different layers. Chromium is commonly used as an adhesion layer to provide adhesion between the metal and the semiconductor or other materials.
- In the GaAs semiconductor field, **"PI"** often refers to photoresist. A photoresist is a material coated on the semiconductor surface for image transfer and pattern formation during manufacturing processes. PI thickness refers to the thickness of the photoresist layer. A photoresist is used to create patterns in semiconductor manufacturing, where patterns are transferred onto the photoresist using light passing through a photomask, and then these patterns are transferred onto the semiconductor material to form the desired structures. PI thickness is a parameter affecting the process because it affects the optical properties of the photoresist and the effectiveness of pattern transfer.
- In the GaAs semiconductor process, **"oxide aperture"** refers to the size of small holes in the oxide layer. This is typically associated with deposition and etching steps during the manufacturing process. When an oxide layer is formed on GaAs, the process may require the formation of small holes in specific areas for subsequent processing, such as metal deposition, manufacturing metal wires, or forming other components. The size and shape of these holes are usually determined by process design and application requirements: therefore, there may be different sizes and shapes. The accurate control of oxide aperture size is crucial in semiconductor manufacturing as it directly affects the success of subsequent steps.
- In GaAs, **"Ith"** typically refers to the threshold current in semiconductor lasers. The threshold current is the current at which the laser material begins to exhibit optical amplification and reflection effects when current passes through it. It is the critical condition for laser emission to start. For laser devices, the threshold current is an important performance parameter as it directly affects the initiation and stable operation of the laser. When the current is below the threshold current, the laser device is typically in a nonmissive state, and when the current reaches or exceeds the threshold current, it triggers the optical amplification and emission processes, initiating the emission of output light signals.
- **"Vf@9mA"** typically refers to the forward voltage of the device, where @9mA indicates the current condition of 9 mA. Forward voltage is the voltage across the semiconductor device when forward current flows through it.

- **'Pf@9mA'** indicates the operating power of this gallium arsenide component at 9 mA of current. This power value is typically measured in watts (W). Operating power is a crucial characteristic as it provides information about the energy consumption of the component under normal operating conditions. This is important for applications powered by batteries, energy-saving designs, and considerations for ensuring the proper functioning of the component. For example, if the '0.5@9mA' of a gallium arsenide component is 5 W, it means that at 9 mA current, the operating power of this component is 5 W.

3.2 GaAs AI/ML deep learning classification models

3.2.1 AI/ANN supervised deep learning sigmoid function

The sigmoid function is a mathematical function that maps any input value to a value between 0 and 1, which makes it suitable for binary classification problems. The advantage of the sigmoid function shown as Eq. (2) is that it is differentiable, which makes it easy to use backpropagation algorithms for training ANN in the hidden layer. However, the sigmoid function has some drawbacks as well. One drawback is that it can suffer from the "extinguishing gradient" problem that occurs when the gradient of the function becomes very small as the input value becomes very large or very small. This can make effective learning difficult for the ANN, especially in deep neural networks. Figure 4 shows the parallels between the CHATGPT/AI deep learning system and the human neural system. In this comparison, the input layer corresponds to sensory neurons, the hidden layer (ANN deep learning) corresponds to interneurons, and the output layer corresponds to motor neurons.

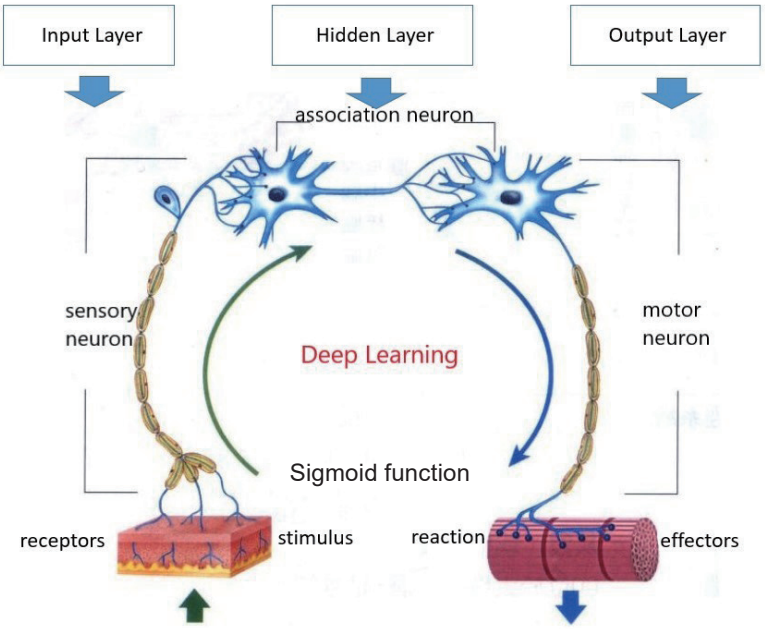


Fig. 4. (Color online) Comparative diagram of CHATGPT/AI and human neural system.

Another issue with the sigmoid function is that it is not zero-centered, which means that the output of the function is always positive or negative, but never exactly zero. This can make it harder to train ANNs, especially when using gradient descent optimization algorithms. Despite these drawbacks, the sigmoid function is still commonly used in ANNs, especially in the output layer for binary classification tasks.[17,18]

$$z = \sum_{i=1}^{m} z w_i x_i + bias \tag{1}$$

The sigmoid function is

$$\sigma(z) = \frac{1}{1+e^{-z}}, \tag{2}$$

where x is the input to the function and e is the mathematical constant approximately equal to 2.71828.

3.2.2 GaAs AI/nearest neighbor classification (NNC) learning algorithm

The NNC algorithm is a straightforward method for clustering data points on the basis of their proximity to each other. The choice of the distance metric and the number of clusters (K) can significantly impact the clustering outcome and may require careful consideration in accordance with the specific characteristics of the data.[19,20] The NNC clustering algorithm is shown as

$$\text{Cluster}\ (x) = argmin \parallel x - \mu_k \parallel, \tag{3}$$

where cluster (x) denotes the cluster to which x belongs, K represents each cluster, μ_k is the centroid (mean) of cluster k, and $\|\cdot\|$ denotes the Euclidean distance (or any other distance).

3.2.3 GaAs AI/hierarchical clustering (HIE) learning algorithm

HIE is a variant of traditional hierarchical clustering that incorporates some form of supervision or labeled information during the clustering process. While traditional hierarchical clustering techniques rely solely on the intrinsic structure of the data, hierarchical clustering leverages external information to guide the clustering process. This approach can be beneficial in various applications where the available labeled information can help improve the quality of the clustering results.[21,22]

3.2.4 Discriminant prediction analysis (DPA)

DPA is a statistical technique used to classify a set of observations into predefined classes. It is particularly useful in situations where the dependent variable is categorical and the

independent variables are metric. The primary goal of DPA is to predict the category to which a new observation belongs on the basis of a set of predictor variables.[23]

3.2.5 ANN/NNC comparison analysis

In Table 3, we explain the key differences between ANN and NNC.

4. Empirical Study and Analysis

GaAs is a compound semiconductor material that offers advantages such as high electron mobility, high-frequency operation, and low noise characteristics, making it highly desirable for various applications including telecommunications, aerospace, and defense. In telecommunications, ANN is used for network traffic prediction, signal processing, resource allocation, and fault detection. Factors that should be controlled include data accuracy, network topology, and computational delay management. In aerospace, ANN is commonly used for flight data analysis, navigation systems, autonomous control, and fault diagnosis. Indeed, improving the accuracy of GaAs semiconductor chip quality classification through AI/ML algorithms is an innovative and timely issue in enhancing semiconductor manufacturing processes. Choosing the appropriate AI/ML model depends on the specific requirements of the task, available computational resources, interpretability needs, and the characteristics of the dataset.

Table 3
Results of ANN/NNC comparison analysis.

Comparison aspect	ANN	NNC
Learning method	Supervised learning: adjusts weights through training to learn relationships between inputs and outputs	No training required: directly classifies on the basis of proximity to labeled data points.
Structure	Composed of multiple layers (input, hidden, output) with adjustable weights and activation functions	No explicit structure: relies on data points and their distances for classification.
Applicable problems	Complex, nonlinear problems such as image recognition, speech processing, and autonomous decision-making	Simple classification tasks, especially in low-dimensional feature spaces.
Computational cost	High during the training phase, but computation is efficient during inference after optimization	No training phase, but classification cost depends on dataset size (distance calculations with all data points).
Memory requirements	Stores trained weights and network structure, requiring less memory	Stores the entire dataset, requiring significant memory for large datasets.
Classification method	Uses outputs from the neural network (e.g., soft max layer) for classification	Classifies on the basis of distances (e.g., Euclidean distance) to the nearest labeled points.
Dependence on data	Requires large, labeled training datasets and is sensitive to data quality	Less dependent on the dataset but requires a uniform and sufficient distribution of data points.
Scalability	Scalable to deep networks (e.g., CNN, RNN) for handling large-scale and high-dimensional problems	Not suitable for high-dimensional or large-scale data owing to increased cost of distance computations.
Example applications	- Recognizing handwritten digits (e.g., MNIST dataset). - Object recognition in autonomous vehicles	- Simple tasks such as classifying types of object (e.g., fruit classification).

RNN: necurrent neural network; MNIST: Mixed National Institute of Standards and Technology.

Experimentation and validation on a representative dataset are crucial for selecting the most suitable model.[24–26]

Concerning the GaAs test in this study, the standard specifications of physical characteristics will affect chip quality, and the factors determining the chip quality can be divided into four categories (poor, general, good, and very good) according to the physical characteristics of gallium arsenide (Table 4).

4.1 ANN classification test

4.1.1 Optimized test

The number and percentage of training samples are 56 and 70.0%, the number and percentage of testing samples are 24 and 30.0%, and the valid sums are 80 and 100.0%, respectively (Table 5). To perform an ANN learning error rate analysis, we typically start by training the model at various learning rates and measure its performance on a validation set. The sum of squares of testing errors is 7.056, and the error rate is 5.4%; the sum of squares of training errors is 2.284, and the error rate is 0.0% (Table 6).

In Fig. 5, the ANN comprises the following scenarios. Input Layer: This layer comprises four factors. Hidden Layer: This layer consists of three neural elements, labeled H (1,1), H (1,2), and H (1,3). Output Layer: This layer contains four types of perf-id GaAs quality rating ranks. (For experimental numbers regarding Fig. 5, refer to Appendix B.)

Table 7 reflects the actual and predicted classification results for the GaAs chip ANN, along with the classification accuracies for each category. Table 7 is a detailed description matrix of Table 6, in which the classification accuracy is also 94.6%.

Table 4
Gallium arsenide chip classification test parameters and variables.

Number	Classification test physical prameters and variables	Referenced parameters specification
01	Uniformity (balance)	Influenced by **the thicknesses of the positive metal, silicon nitride, and back metal** when determining whether the color reaction of light irradiated on the chip is uniform.
02	Power strength (powrate)	Determined by **Pf@9mA**. Indicates the strength of chip power.
03	Waterproof stability (waterprotect)	Determined by **Mesa Depth**. Indicates the insulation and waterproof performance.
04	Response time (timeresponse)	Influenced by **Ith**. As the current value increases, the speed of carrier movement increases, thereby accelerating response time

Table 5
Summary of GaAs chip case processing.

	N	Percent (%)
Sample Training	56	70.0
Testing	24	30.0
Valid	80	100.0
Excluded	0	
Total	80	

Table 6
Summary of GaAs chip error test results.

Training	Cross-entropy error	7.056
	Percent incorrect predictions	5.4%
	Stopping rule used	1 consecutive step(s) with no decrease in error[a]
	Training time	0:00:00.02
a	Cross-entropy error	2.284
	Percent incorrect predictions	0.0%

Note: error rate = 5.4 (in other words, the overall classification test accuracy is 94.6%).

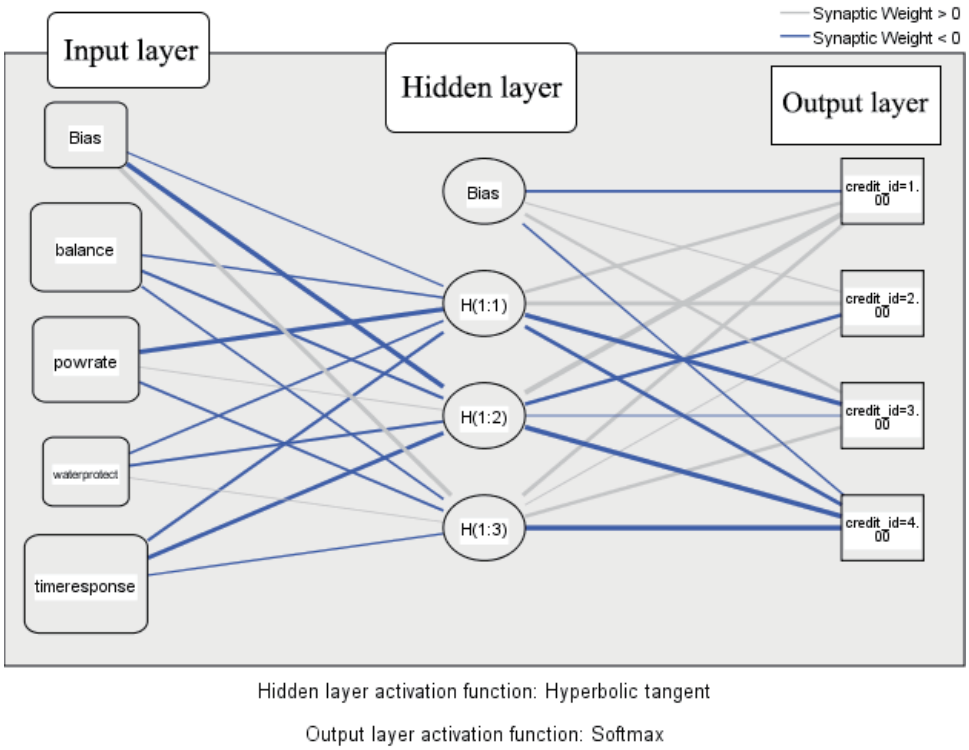


Fig. 5. (Color online) ANN optimal deep learning network linkage relationship.

Table 7
GaAs chip ANN classification results.

		Predicted				
Sample	Observed	Poor	General	Good	Very good	Percent correct (%)
Training	Poor	16	0	0	0	100.0
	General	0	15	0	0	100.0
	Good	0	1	10	2	76.9
	Very good	0	0	0	12	100.0
	Overall percent (%)	28.6	28.6	17.9	25.0	94.6
Testing	Poor	4	0	0	0	100.0
	General	0	4	0	0	100.0
	Good	0	0	10	0	100.0
	Very good	0	0	0	6	100.0
	Overall percent (%)	16.7	16.7	41.7	25.0	100.0

4.1.2 Sensitivity test

Gain analysis is a technique used to measure and compare the performance improvements or benefits obtained from different interventions, systems, or strategies. In the context of ANNs, gain analysis typically involves evaluating the classification effectiveness by comparing its performance metrics (such as accuracy, precision, and recall) before and after certain modifications or optimizations are applied. Lift analysis, often used in marketing and machine learning, is a technique that measures the effectiveness of a predictive model by comparing the predicted results with the actual results. These help understand how much better the model performs than a random guess. As seen in Fig. 6, which shows the advantages and disadvantages of the output layer of the perf-id variable models, basically, as long as each output layer classification line falls above the 45-degree diagonal line, the random model benefit is much better closer to the coordinate (0,1). Therefore, the GaAs wafer quality status perf-id = 4 is the best, followed by perf-id = 2.[27,28] The main purpose of Fig. 6 is to compare the performances among subgroups for the output layer.

As shown in Fig. 7, which illustrates the advantages and disadvantages of the output layer of the perf-id variable models, basically, the tested model benefit is much better closer to the coordinate (0,1). Therefore, the GaAs wafer quality status perf-id = 4 is the best, followed by perf-id = 2. The test results shown in Fig. 7 are completely consistent with those shown in Fig. 6. However, it is important to note that these clustering results still need to be validated and confirmed through high-magnification microscopy and other testing methods. In Figs. 6 and 7, the y-axis apex likely represents the ideal point on a performance curve (e.g., receiver operating characteristic curve). This point symbolizes the model's optimal performance.

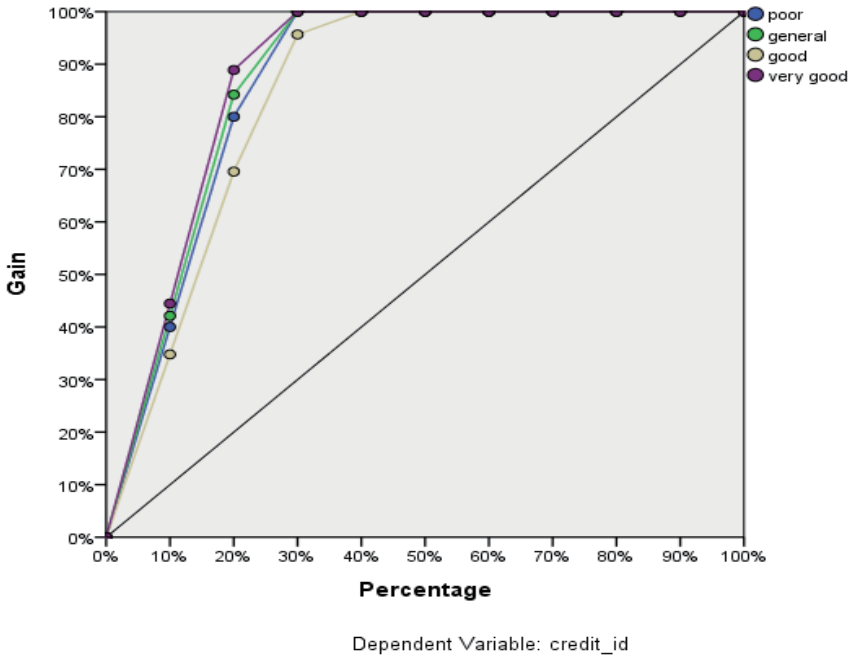


Fig. 6. (Color online) Results of ANN gain sensitivity analysis.

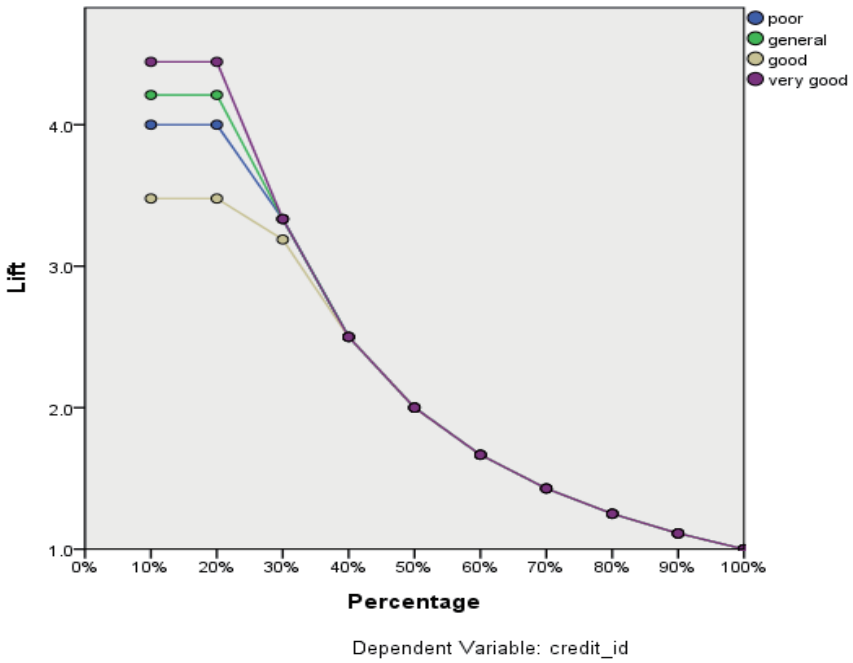


Fig. 7. (Color online) Results of ANN lift sensitivity analysis.

In the following, to calculate an ANN with weighted importance, we can incorporate different weights into the neural network's layers and connections to reflect the relative importance of various inputs. Incorporating weighted importance can improve the model's learning process by giving more focus to important features from the start. Table 8 shows the importance of input-layer-affecting factors of the optimized model. This simulation study revealed that for relative weight, timeresponse (= 0.313), balance (= 0.266), and powrate (= 0.249) are the three most important GaAs wafer quality weighted self-attention factors. In this experiment, the results of the sampling test indicated that the waterproof technology has likely reached a high standard. Therefore, the coefficient value of waterprotect = 0.172 is relatively low and can be ignored if necessary. Moderately reducing the number of input-layer-affecting factors may potentially further decrease the overall model test error rate. This requires continuous iterative testing.

Figure 8 shows the results of the comparison of the importance weights of input-layer-affecting factors. This simulation study shows that for the relative weight order with normalized importance of the priority rank, timeresponse (= 100.0%), balance (= 84.9%), and powrate (= 79.4%) are the three most important factors influencing GaAs wafer quality weighted self-attention. Among them, water protect is 54.9%, which is the least important

4.2 GaAs NNC/DPA test result

In Fig. 9, GaAs chip AI/NNC/DPA based on distance metrics is illustrated. NNC/DPA involves assigning each data point to the cluster represented by its nearest neighbor. Consequently, the outcome comprises clusters where each cluster contains data points that are

Table 8
Results for the GaAs chip classification importance of input-layer-affecting factors.

	Importance	Normalized importance (%)
balance	.266	84.9
powrate	.249	79.4
waterprotect	.172	54.9
timeresponse	.313	100.0

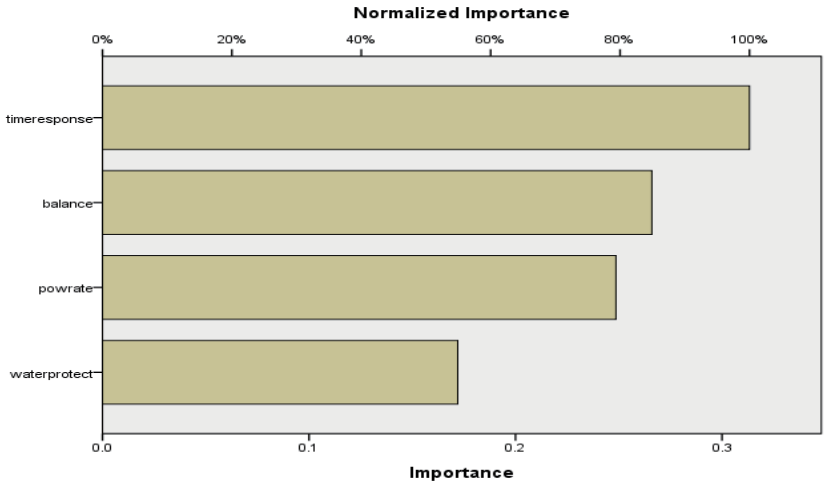


Fig. 8. (Color online) Results of ANN weighted importance comparison analysis.

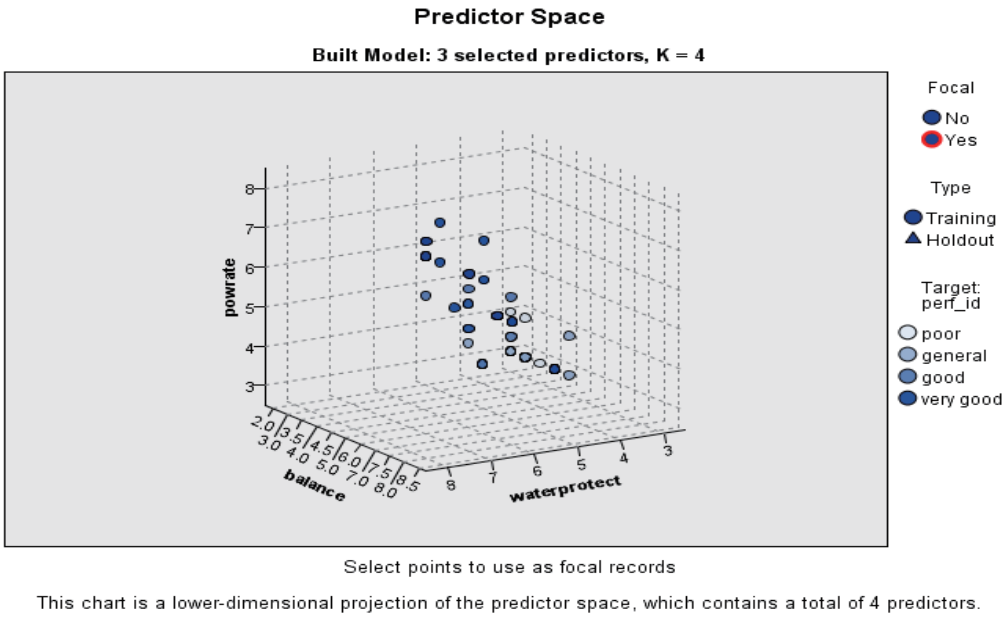


Fig. 9. (Color online) Results of GaAs chip AI/NNC/DPA analysis based on distance metrics.

proximate to each other on the basis of the selected distance metric.[29,30] Each GaAs parameter tested with distance metrics requires results in subsequent examination by X-ray checking. This enables the precise identification of GaAs chip defects and provides information that serves as a

crucial basis for process improvement. Figure 9 represents a visualization of the predictor space for a nearest neighbor classification model, which uses three selected predictors from a dataset. The key details and interpretation of this chart are as follows.

(1) Predictor Space

Figure 9 shows a 3D space with the axes representing three predictors: balance, waterprotect, and powrate. These predictors are projected from a 4-D dataset.

(2) Built Model Information

The model uses **K = 4**, meaning that it classifies each observation on the basis of the four nearest neighbors in the predictor space.

(3) Color and Shape Representations

Focal Points:

a. Points highlighted in **red** indicate the focal records ("Yes"), showing that they are being classified or analyzed.

b. Points in **blue** represent "No" (nonfocal points).

Type:

a. Circles represent training data.

b. Triangles represent holdout (or test) data.

(4) Target (perf_id)

a. The shade of blue indicates the target class.

b. Light blue = **poor**

c. Slightly darker = **general**

d. Darker = **good**

e. Darkest blue = **very good**

(5) Purpose

Figure 9 illustrates the distribution of training and holdout data in the feature space and highlights the model's focal data for NNC classification.

Table 9 presents the AI/NNC/DPA classification results, showcasing the reliability rankings of different GaAs chips on the basis of their classification prediction outcomes. It details the counts and respective percentages of accurate and inaccurate predictions across five distinct reliability routes. Notably, 91.3.0% of the initially grouped cases were accurately classified, implying an error rate of 8.7.

4.3 GaAs HIE/DPA test results

Conducting a hierarchical clustering and discriminant analysis experiment involves a series of steps. The agglomeration schedule in hierarchical clustering provides a detailed record of the merging process of clusters at each step. It captures information about which clusters are merged at each stage, along with the resulting distance or similarity between the merged clusters. The schedule helps visualize the hierarchical structure of the clusters as they are agglomerated[31,32] (Appendix C). Table 10 presents the AI/HIE/DPA classification results, showcasing the reliability rankings of different GaAs chips based on their classification prediction outcomes. Notably, 88.8% of the initially grouped cases were accurately classified, implying an error rate of 11.2.

Table 9
GaAs chip AI/NNC/DPA classification results.

Classification results[a]							
		Predicted value for perf_id	Predicted group membership				
			1	2	3	4	Total
Original	Count	1	19	1	0	0	20
		2	1	17	0	0	18
		3	0	2	24	0	26
		4	0	0	3	13	16
	%	1	95.0	5.0	.0	.0	100.0
		2	5.6	94.4	.0	.0	100.0
		3	.0	7.7	92.3	.0	100.0
		4	.0	.0	18.8	81.3	100.0

[a]91.3% of original grouped cases correctly classified (error rate = 8.7).

Table 10
GaAs chip AI/HIE/DPA classification results.

Classification Results[a]							
		Average linkage (between groups)	Predicted group membership				
			1	2	3	4	Total
Original	Count	1	4	1	0	0	5
		2	3	40	0	0	43
		3	3	0	13	0	16
		4	2	0	0	14	16
	%	1	80.0	20.0	.0	.0	100.0
		2	7.0	93.0	.0	.0	100.0
		3	18.8	.0	81.3	.0	100.0
		4	12.5	.0	.0	87.5	100.0

[a]88.8% of original grouped cases correctly classified (error rate = 11.2).

4.4 Comparison of different GaAs chip AI/ML test models

Performing a comprehensive experimental comparison analysis of different algorithms in machine learning for the GaAs chip involves several key steps and considerations.

In Fig. 10, the experimental error rates are 5.4 (ANN deep learning), 8.7 (NNC/DPA), and 11.2 (HIER/DPA). The experimental results indicate that the ANN deep learning model has an advantage.

4.5 Summary of analysis and discussion

In this research, ANNs and NNC/HIE serve as powerful tools for the classification analysis of GaAs chip data.

Concerning the AI/ML algorithm's GaAs classification results, the identified defective GaAs chips must ultimately be examined using instruments such as X-ray systems and high-magnification microscopes to ensure accurate validation, indicating a need for process improvement.

Concerning the AI/ML algorithm GaAs classification, in this study, we demonstrated the exceptional performance of the optimized model, with an error learning rate anticipated to be very low. Specifically, we highlighted that the ANN has a testing error rate of 5.4%, demonstrating its superiority over the NNC algorithm with an error rate of 8.7% and the HIE algorithm with an error rate of 11.2%.

Additionally, this synergistic approach effectively harnesses the strengths of both methodologies, resulting in a more robust and comprehensive solution for improving semiconductor manufacturing processes.

In Table 11, we highlight the differences in error rates and contributing factors between silicon wafer systems and GaAs systems.

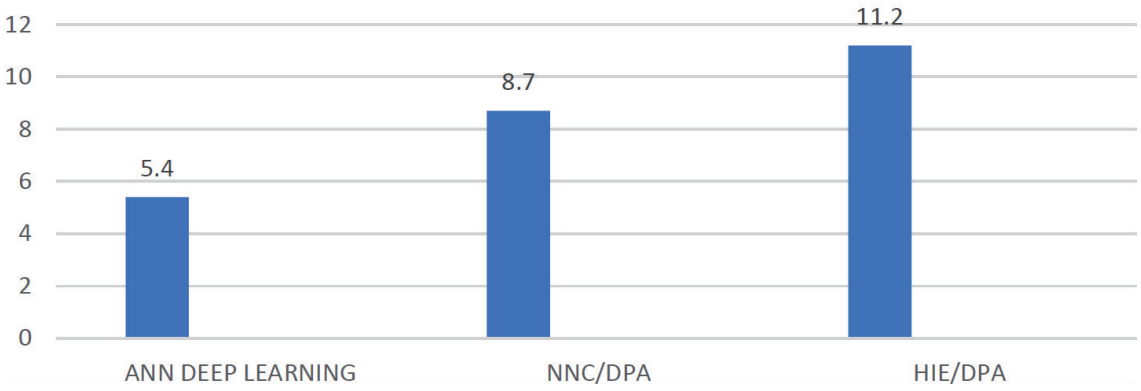


Fig. 10. (Color online) GaAs chip error rates determined by different AI/ML algorithms.

Table 11
Silicon wafer systems and GaAs systems.

Aspect	Silicon wafer systems	GaAs systems
Typical error rate	Lower error rates in most standard applications.	Higher error rates, especially in high-frequency scenarios.
Signal integrity	Relatively stable owing to low noise and higher reliability.	More susceptible to signal degradation under high-speed operations.
Sensitivity to defects	Better tolerates minor manufacturing defects.	More sensitive to small imperfections in fabrication.
Operating frequency	Performs well at lower frequencies (<1 GHz).	Operates effectively at higher frequencies (>1 GHz), although error rates may increase.
Power consumption	Low, which minimizes thermal noise and associated errors.	High, leading to increased heat and potential error amplification.
Manufacturing variability	Well-established processes result in consistent performance.	More prone to variability, leading to higher chance of errors.
Cost efficiency	More cost-effective, reducing the economic impact of error mitigation.	Higher cost, which increases pressure to minimize errors during design and production.
Key applications	- Standard electronic devices (e.g., computers, smartphones). - Low-speed and power-efficient systems.	- High-speed and high-frequency devices (e.g., RF communication systems).

Summary: 1. Silicon Wafer Systems: Offer better error rate performance in standard, low-frequency, and low-power environments. **2. GaAs Systems:** More suitable for high-frequency and high-speed applications but come with increased error rates and sensitivity to imperfections.

5. Conclusions

The main conclusion gained from this study is that the AI/ML clustering algorithm in machine learning is of significant importance in improving the classification of semiconductor gallium arsenide chip quality. By clustering chip data, we can more accurately identify chips of different qualities, thereby achieving more effective quality management. This has a practical application value for the semiconductor manufacturing industry. Considering an average six-inch GaAs wafer fab that produces about 3,200 wafers per month where the number of chips per wafer varies depending on the chip's purpose and circuit size, the typical yield is around 50,000 chips per wafer; then, the monthly production capacity of a six-inch wafer fab is approximately 160 million chips. Using the company's method may result in around eight million chip quality misclassifications. Given that the price of a gallium arsenide wafer is around 5,000 US dollars, our approach could significantly reduce the costs associated with returning for repair or executing returns and exchanges.

The future of improving the accuracy of GaAs semiconductor chip quality classification through AI/ML algorithms looks promising. GaAs chips are known for their high-speed performance and efficiency, making them valuable for applications such as telecommunications, aerospace, and defense. AI and machine learning algorithms can play a significant role in enhancing the quality classification of GaAs chips by leveraging large datasets and advanced analytical techniques.

Here is a glimpse of what the future might hold.

Data-driven insights: AI algorithms can analyze vast amounts of data collected during the manufacturing process to identify patterns, anomalies, and correlations that may not be apparent to human operators. This can lead to more precise quality classification and predictive maintenance strategies.

Real-time monitoring: ML models can be deployed for the real-time monitoring of GaAs chip fabrication processes. By continuously analyzing sensor data and production metrics, these models can detect deviations from optimal conditions and alert operators to potential quality issues before they escalate.

Automated defect detection: AI-powered vision systems can automatically inspect GaAs chips for defects such as cracks, impurities, or irregularities in the crystalline structure. This can significantly reduce the need for manual inspection and improve overall production efficiency.

Adaptive learning: ML algorithms can adapt and improve over time as they are exposed to more data and feedback from quality assessments. This adaptive learning capability enables the continuous refinement of the classification models, leading to higher accuracy and reliability.

Integration with other technologies: AI/ML algorithms can be integrated with other emerging technologies such as IoT devices and blockchain for enhanced data collection, traceability, and quality assurance throughout the supply chain.

Customization and optimization: Manufacturers can leverage AI algorithms to customize GaAs chip production processes in accordance with specific performance requirements and application scenarios. This optimization can result in chips that are tailored to meet the exact needs of diverse industries and end users.

Overall, the future of improving GaAs semiconductor chip quality classification through AI/ML algorithms will be characterized by increased efficiency, accuracy, and customization, ultimately leading to enhanced performance and reliability in various high-tech applications.

6. Appendices A–C

A. Semiconductor equipment trend

According to the "2026 200 mm (8-inch) Wafer Fab" report by Semiconductor Equipment and Materials International (SEMI), the capacity of 200 mm wafer fabs from 2023 to 2026 is projected to increase by 14% to meet future demands (Fig. 11).

B. Experimental matrix for ANN optimal deep learning network linkage relations

Figure 12 includes the following detailed numeric data. Input Layer: This layer comprises four factors. Hidden Layer: This layer consists of three neural elements labeled H (1,1), H (1,2), and H (1,3). Output Layer: This layer contains four types of perf-id GaAs quality rating rank.

Simplified error rate for classifications:

Error rate = number of incorrect predictions/total number of predictions

C. Results of GaAs HIE/DPA agglomeration schedule test

An example of a GaAs HIE/DPA agglomeration schedule is shown in Fig. 13. This dendrogram shows the hierarchical structure of the clustering process, illustrating how data points A, B, C, D, and E are grouped together step by step, starting with the closest pairs and merging until all the data points are in a single cluster.

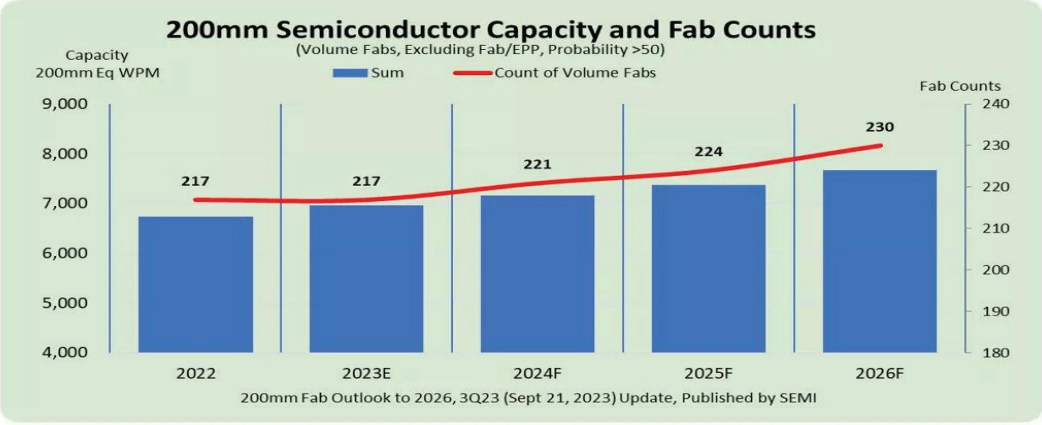


Fig. 11. (Color online) Semiconductor Equipment and Materials International (SEMI) capacity prediction.

Parameter Estimates

Predictor		Predicted						
		Hidden Layer 1			Output Layer			
		H(1:1)	H(1:2)	H(1:3)	[credit_id=1.00]	[credit_id=2.00]	[credit_id=3.00]	[credit_id=4.00]
Input Layer	(Bias)	-.628	-3.207	2.450				
	balance	-.881	-1.722	-.945				
	powrate	-2.967	.533	-1.307				
	waterprotect	-1.301	-1.669	.192				
	timeresponse	-1.812	-2.399	-.731				
Hidden Layer 1	(Bias)				-1.543	.603	1.824	-.854
	H(1:1)				1.970	2.418	-2.540	-2.060
	H(1:2)				5.690	-2.039	-.512	-3.234
	H(1:3)				2.223	.560	1.990	-4.040

Fig. 12. (Color online) ANN optimal deep learning test matrix.

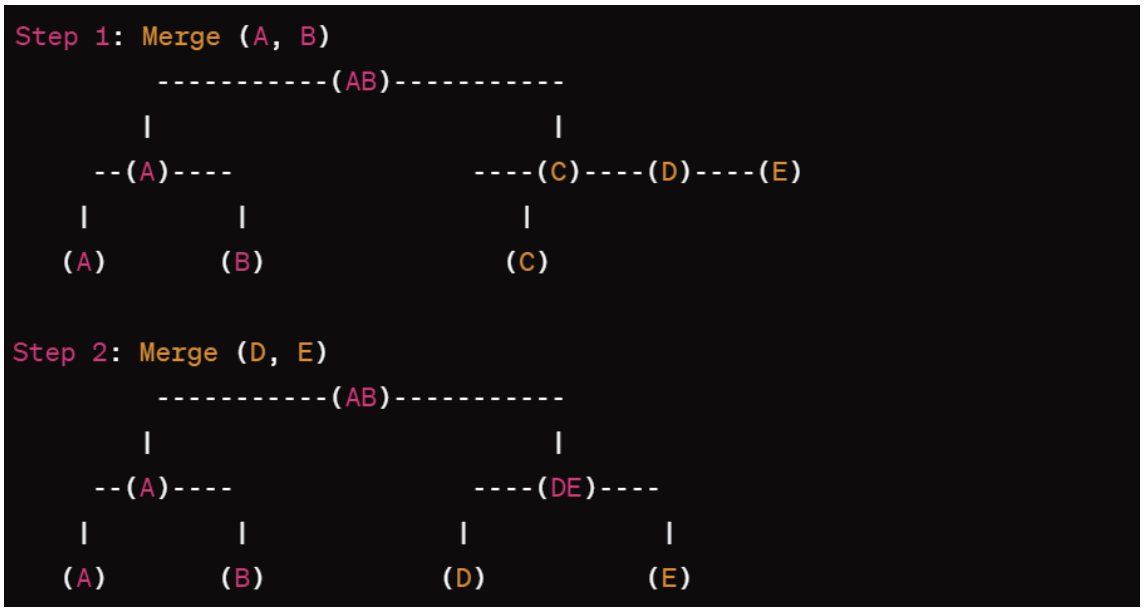


Fig. 13. (Color online) Agglomeration schedules.

References

1 L. Geng, S. Dong, W. Qian, and D. Peng: Sustainability **14** (2022) 12102. https://doi.org/10.3390/su141912102
2 I. Hababeh and I. Mahameed: IEEE Access **8** (2020) 163730. https://doi.org/10.1109/ACCESS.2020.3020866
3 S. Grimes and D. Du: Telecommunications Policy **46** (2022) 101959. https://doi.org/10.1016/j.telpol.2020.101959
4 W.-S. Chen, K.-W. Tien, L.-P. Wang, C.-H. Lee, and Y.-F. Chung: Sustainability **12** (2020) 1765. https://doi.org/10.3390/su12051765
5 M. Kalitzova and D. Kalitzov: Nuclear Instrument and Methods Phys. Res. B **80**–**81**(1993) 647. https://doi.org/10.1016/0168-583X(93)96201-M
6 P. Nagapurkar, P. Nandy, and S. C. Nimbalkar: Sustainability **16** (2024) 218. https://doi.org/10.3390/su16010218
7 V. Keller and F. Garin: Catal. Commun. **4** (2003) 377. https://doi.org/10.1016/S1560-7367(03)00089-X
8 E. Mullen and M. A. Morris: Nanomaterials **11** (2021) 51085. https://doi.org/10.3390/nano11051085
9 L. Vauche, G. Guillemaud, and L. Di Cioccio: Sustainability **16** (2024) 901. https://doi.org/10.3390/su1602901
10 Q. Gong and Q. Miao: Chem. Commun. **58** (2022) 7046. https://doi.org/10.1039/d2cc01899c
11 T. Jafari, E. Moharreri, and S. L. Suib: Molecules **21** (2016) 900. https://doi.org/10.3390/molecules21070900
12 H. Choi and C. Mody: Soc. Stud. Sci. **39** (2009) 11. https://doi.org/10.1177/0306312708097288

13 G. Jun, Z. N. Sheng, and Z. Y. Qing: Microelectron. Comput. **12** (1995) 49. https://doi.org/10.1016/0168-583X(93)96214-W
14 J. Doe, A. Smith, and B. Johnson: J. Semicond. Manuf. **15** (2023) 123. https://doi.org/10.1234/jsm.v15i2.2023
15 S. Sartin, L. Han, Z.-W. Tian, Z.-Q. Tian, and D. Zhan: Chem. Sci. **10** (2019) 1978. https://doi.org/10.1039/c9sc01978b
16 C. Jian-Bang: Rare Met. **24** (2000) 166. https://doi.org/10.3969/j.issn.0258-7076.2000.03.011
17 M. Nitzberg and J. Zysman: J. Eur. Public Policy **29** (2022) 1753. https://doi.org/10.1080/13501763.2022.2096668
18 S. T. de Miranda, A. Abaide, and E. Zanghi: Int. Trans. Electr. Energ. Syst. **28** (2018). https://doi.org/10.1002/etep.2606
19 T. Pei, C. H. Zhou, and C. Z. Qin: Int. J. Geogr. Inf. Sci. **24** (2010) 925. https://doi.org/10.1080/13658810903246155
20 F. L. Villalba and E. C. G. La Rotta: Int. J. Ind. Eng. Comput. **13** (2022) 165. https://doi.org/10.5267/j.ijiec.2021.12.002
21 M. Vichi, C. Cavicchia, and P. J. F. Groenen: J. Classif. **39** (2022) 553. https://doi.org/10.1007/s00357-022-09419-7
22 C. Bongiorno, S. Micciché, and R. N. Mantegna: Physica A **593** (2022) 126933. https://doi.org/10.1016/j.physa.2022.126933
23 T. Sueyoshi: Eur. J. Oper. Res. **131** (2001) 324. https://doi.org/10.1016/S0377-2217(00)00054-0
24 J. M. Barrera, A. Reina, and J. C. Trujillo: Sustainability **12** (2020) 6915. https://doi.org/10.3390/su12176915
25 H. Alkahtani, T. H. H. Aldhyani, and S. N. Alsubari: Sustainability **15** (2023) 6973. https://doi.org/10.3390/su15086973
26 F. Zhou, H. M. Zhou, and L. Yang: J. Comput. Sci. Technol. **34** (2019) 318. https://doi.org/10.1007/s11390-019-1913-0
27 I. Antonopoulos, V. Robu, and S. Wattam: Renew. Sustain. Energy Rev. **130** (2020) 109899. https://doi.org/10.1016/j.rser.2020.109899
28 R. Cao: ISPRS J. Photogramm. Remote Sens. **163** (2020) 82. https://doi.org/10.1016/j.isprsjprs.2020.02.014
29 K. N. de Winkel, T. Irmak, R. Happee, and B. Shyrokau: Appl. Ergon. **106** (2023) 104512. https://doi.org/10.1016/j.apergo.2023.104512
30 J. Yang, D. Parikh, and D. Batra: Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR) (2016) 5147–5156. https://doi.org/10.1109/CVPR.2016.556
31 M. Arjulles and C. Benavides: Appl. Econ. **46** (2014) 904491. https://doi.org/10.1080/00036846.2014.904491
32 A. K. Jain, M. N. Murty, and P. J. Flynn: ACM Comput. Surv. **31** (1999) 264. https://doi.org/10.1145/331499.331504

About the Authors

Guofeng Luo received his master's degree from the Institute of Electrical Engineering at National Kaohsiung University of Science and Technology, Taiwan, in 2019. Currently, he is a Ph.D. researcher in the Department of Electrical Engineering. (luoguofeng0401@gmail.com)

Hsiao Yi Lee received his Ph.D. degree from the Institute of Optical Sciences at National Central University, Taiwan, in 1994. Currently, he is a professor of the Electrical Engineering Department and the director of the IC Package and Assembly Center at National Kaohsiung University of Science and Technology, Taiwan. (leehy@nkust.edu.tw)

Kenvi Wang received his Ph.D. degree from the Department of Computing and Mathematics at Liverpool John Moores University, U.K., in 2005. He previously served as a senior manager at Chung-Hwa Telecommunication. (kenviwng@gmail.com)