AI in Semiconductors

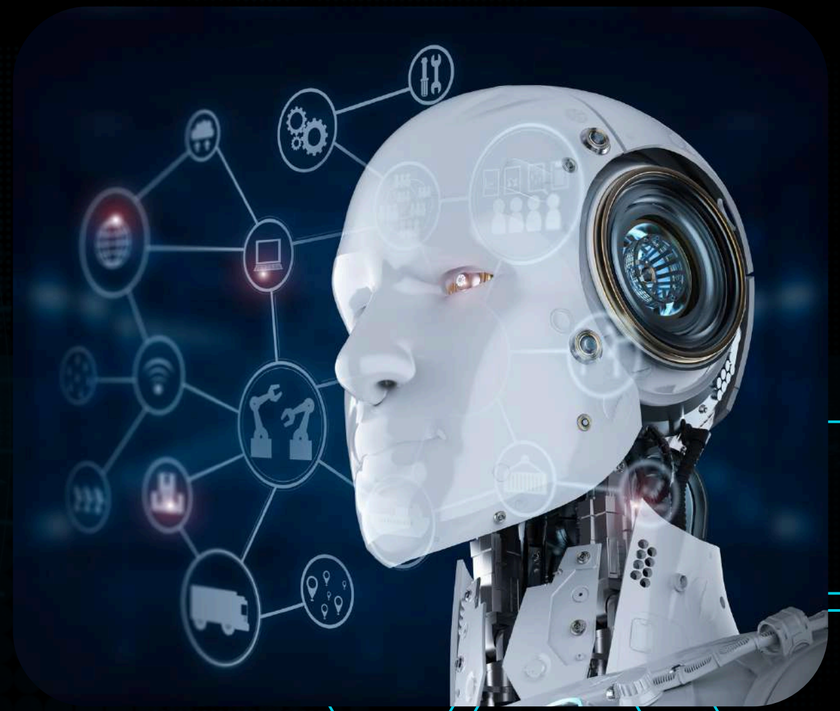
Presented By **Brijesh Sharma (48)**

Introduction

The AI-Semiconductor Revolution

The semiconductor industry stands at an inflection point where artificial intelligence is transforming every aspect of chip development and manufacturing. This revolution is occurring across three fundamental dimensions:

- Design: AI explores design spaces 10,000x larger than human capability
- Manufacturing: Learns subtle process variations invisible to SPC charts
- Testing: Generates adaptive test patterns for emerging failure modes



This transformation matters because:

- Moore's Law is slowing, but architectural innovation through AI provides continued advancement
- AI enables complex co-optimization of power, performance, and area (PPA) that exceeds human capability
- The semiconductor industry faces existential challenges that AI helps solve

Problem Statement:

Why We Need AI

Design Complexity Crisis

- Modern chips contain billions of transistors with 3D architectures
- Human designers can’t manually optimize:
  - Power delivery networks
  - Signal integrity in 2.5D/3D stacks
  - Thermal dissipation paths

Manufacturing Variability

- At 3nm, a single atom’s displacement can cause failures
- Traditional statistical process control misses:
  - Complex multi-variable interactions (e.g., plasma etch + deposition effects)
  - Slow drifts in tool performance

Testing Limitations

Conventional test patterns miss emergent failure modes

Example: AI-found bugs in Apple M1 chips involved timing races human engineers didn’t anticipate

Algorithm Description

Machine Learning in Chip Design

Reinforcement Learning for Physical Design

- Treats chip as a "game board"
- AI "player" moves components and gets rewarded for:
  - Shorter wirelengths
  - Balanced thermal profiles
  - Lower power noise
- Discovers non-intuitive but optimal placements

Graph Neural Networks for Connectivity

- Models chip as mathematical graph
- Predicts congestion and timing issues early
- Understands hierarchical relationships

Adaptive Testing Systems

Anomaly Detection

- Learns normal chip behavior
- Flags statistical outliers
- Catches unexpected failure modes

Intelligent Test Generation

- Focuses testing on high-risk areas
- Adapts based on previous results
- Reduces test time while improving coverage

Anomaly Detection in Semiconductor Testing

Anomaly detection systems learn the "normal" behavior of properly functioning chips and flag statistical outliers that deviate from this baseline. Unlike rule-based testing (which looks for known failure patterns), anomaly detection can catch:

Autoencoders

Autoencoders are the most adopted anomaly detection method in semiconductor manufacturing and testing due to their ability to learn complex patterns without labeled defect data.

1. Autoencoders are the most adopted anomaly detection method in semiconductor manufacturing and testing due to their ability to learn complex patterns without labeled defect data.Unsupervised Learning
    - Doesn’t require labeled defect data (critical because defects are rare and costly to collect).
    - Learns "normal" behavior from good chips/wafer scans.
2. Handles High-Dimensional Data
    - Works with SEM images, parametric test data, power/thermal profiles.
    - Compresses data into a latent space, making anomaly detection efficient.
3. Explainable Anomalies
    - Defects appear as high reconstruction errors (easy to visualize).
4. Adaptable to Different Data Types
    - Works for wafer maps, electrical test results, and even time-series sensor data.

Anomaly Detection in Semiconductor Testing

Step-by-Step Process

1. Training Phase (Normal Data Only)
   - Feed thousands of defect-free wafer images or chip test results into the autoencoder.
   - The model learns to compress (encode) and reconstruct (decode) normal patterns.
2. Inference Phase (Detecting Defects)
   - Pass new wafer/chip data through the trained autoencoder.
   - Defective chips/wafers will have high reconstruction errors because their patterns differ from trained "normal" data.
   - A threshold is set (e.g., 99th percentile of error) to flag anomalies.
3. Visualization & Root Cause Analysis
   - Heatmaps highlight where reconstruction fails (pinpoints defect locations).
   - Engineers review flagged cases to update models or adjust processes.

Real-World Case Study

TSMC’s Autoencoder-Based Defect Detection

Problem

- Traditional inspection missed subtle, random defects in 3nm wafer scans.
- Human inspectors had fatigue issues, leading to escapes (missed defects).

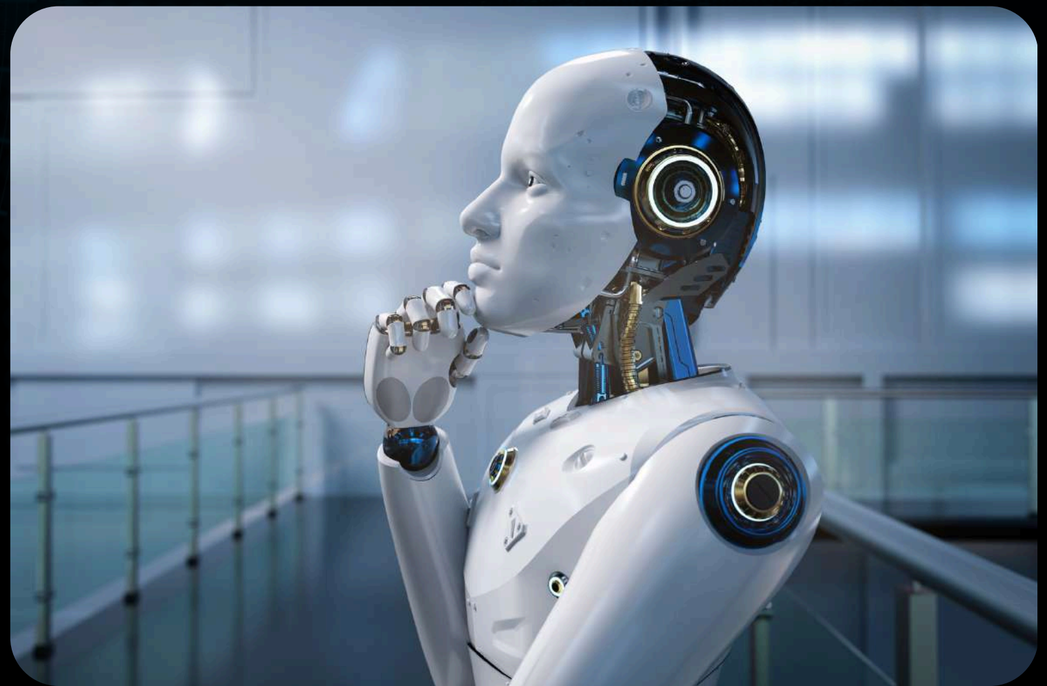
Solution

- Trained a Convolutional Autoencoder (CAE) on defect-free SEM images.
- Used pixel-wise reconstruction error to flag anomalies.

Results

Metric	Autoencoder	Human Inspectors
Defect Detection Rate	**99.1%**	95.3%
False Positives	**0.05%**	0.2%
Inspection Speed	**10x faster**	Manual review

- Cost Savings: Reduced wafer scrap by $3M/month at a single fab.
- Scalability: Deployed across all advanced nodes (5nm, 3nm, etc.).



Comparison with Other Methods

Algorithm	Pros	Cons	Best Used For
Autoencoder	No labeled data needed, handles images/sensor data	Requires large "normal" dataset	Wafer defect detection, parametric test outliers
One-Class SVM	Good for small datasets	Struggles with high-dimensional data	Early-stage process monitoring
Isolation Forest	Fast for big data	Less accurate for subtle defects	Gross failure screening
GMM	Captures multi-modal data	Sensitive to noise	Process variation monitoring

Results

Quantifiable Impact:

Area	Improvement	Business Impact
Design Time	50-70% faster	$50M+/project savings
Defect Rate	30-40% reduction	$10M/wafer lot saved
Power Efficiency	12-25% better	Extends battery life 3+ hours

Applications

A. Next-Gen Devices
- Smartphones: AI-optimized SoCs (e.g., Apple Neural Engine)
- Autonomous Vehicles: Fail-safe AI chips with self-monitoring

B. Advanced Packaging
- AI-driven 3D IC integration:
  - Optimizes through-silicon via (TSV) placement
  - Manages thermal stresses

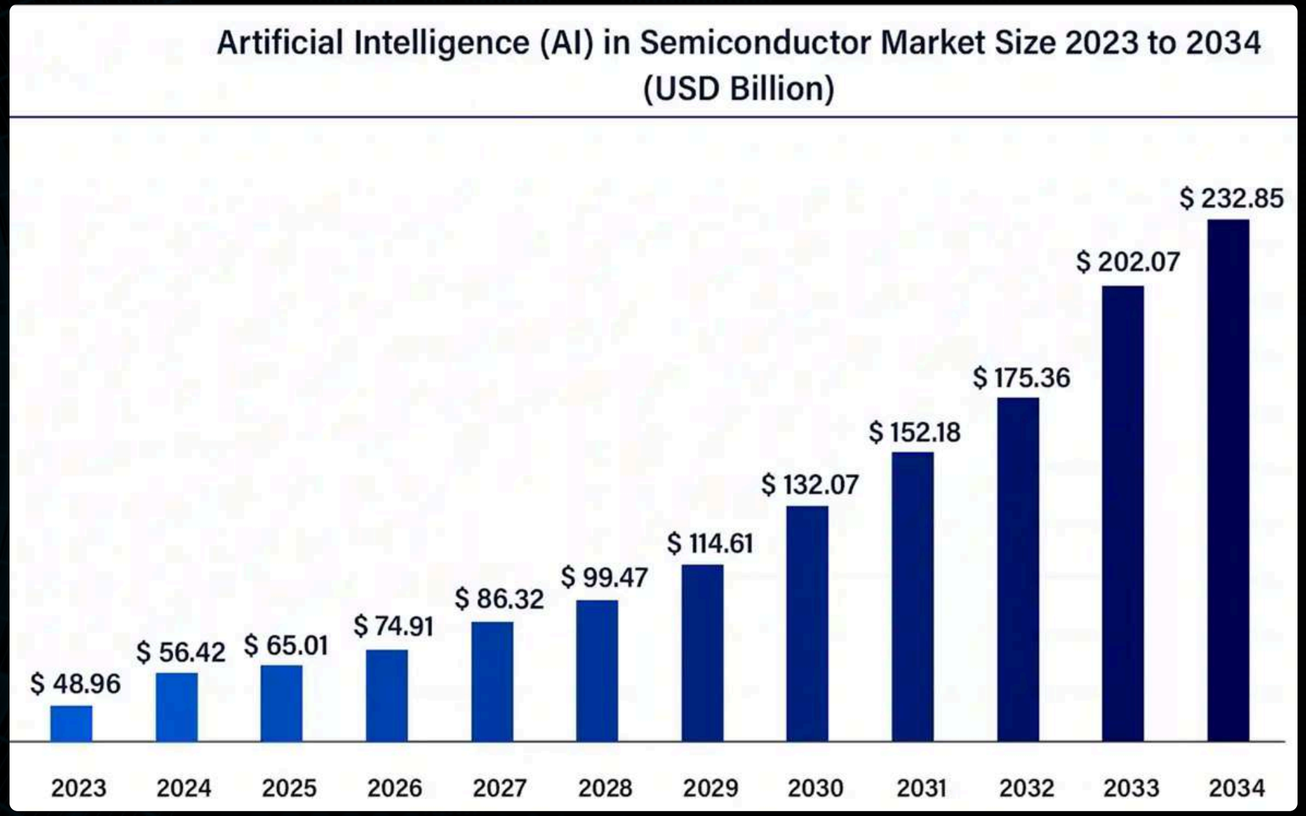
C. Quantum Computing
- AI co-designs qubit layouts for:
  - Minimal crosstalk
  - Maximum coherence time

Conclusion

Shaping Tomorrow Together

- AI isn’t just improving semiconductors—it’s redefining what’s possible:
  - Chips that optimize themselves during operation
  - Foundries that self-calibrate tools in real-time

We’re entering an era where the most advanced chips may be designed not by humans, but by systems trained on the collective wisdom of decades of semiconductor physics.



References

Shaping Tomorrow Together

Technical Papers:

- S. Dey, S. Nandi and G. Trivedi, "Machine Learning for VLSI CAD: A Case Study in On-Chip Power Grid Design," 2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI), Tampa, FL, USA, 2021, pp. 378-383, doi: 10.1109/ISVLSI51109.2021.00075. keywords: {Electromigration;Machine learning;Very large scale integration;Power grids;System-on-chip;Transistors;Indexes;IR Drop;Electromigration;Machine Learning;On-Chip Power Grid;VLSI CAD},

- "Reinforcement Learning for Physical Design" (Nature, 2021)
  https://arxiv.org/abs/2004.10746

Industry Reports:

- McKinsey "AI in Semiconductors 2025"
- SEMI Roadmap for AI/ML Adoption

Thank You