Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# AI-Driven Feature Recognition of SEM Profiles in Deep Reactive Ion Etching Based on Physics-Constrained Variational Autoencoder

**Yi Sun**

sunyi@mail.sim.ac.cn

Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences

**Fang Wang**

Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences
https://orcid.org/0009-0006-5073-8806

**Hao Yu**

Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences

**Yechen Miao**

Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences

**Ke Sun**

Shanghai Institute of Microsystem and Information Technology

**Heng Yang**

Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences
https://orcid.org/0000-0002-8873-9576

**Xinxin Li**

Shanghai Institute of Microsystem and Information Technology,CAS

**Article**

**Keywords:** AI for Manufacturing, Deep Reactive Ion Etching, Scanning Electron Microscopy, Artificial Intelligence, Variational Autoencoder, Image Recognition

**Posted Date:** June 2nd, 2025

**DOI:** https://doi.org/10.21203/rs.3.rs-6758803/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** There is no conflict of interest

# AI-Driven Feature Recognition of SEM Profiles in Deep Reactive Ion Etching Based on Physics-Constrained Variational Autoencoder

*Fang Wang[1,2], Hao Yu[1,2], Yechen Miao[1,2], Ke Sun[1], Yi Sun[1,*], Heng Yang[1,2,*], and Xinxin Li[1,2,*]*
[1]State Key Laboratory of Transducer Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, 200050 Shanghai, China
[2]The School of Graduate Study, University of Chinese Academy of Sciences, 100049 Beijing, China
[*] Corresponding Authors: sunyi@mail.sim.ac.cn, h.yang@mail.sim.ac.cn, xxli@mail.sim.ac.cn

## ABSTRACT

Deep reactive ion etching (DRIE) is critical for fabricating high-aspect-ratio structures in microelectromechanical systems (MEMS), yet its complex, parameter-dependent process poses significant optimization challenges. Artificial intelligence (AI) offers an efficient optimization solution, but its implementation faces the technical challenge of acquiring large-scale data from scanning electron microscopy (SEM) images, the standard for evaluating DRIE etching outcomes. Traditional SEM analysis relies on labor-intensive manual methods, incurring 15-20% errors and hindering high-throughput manufacturing. Existing automated methods, such as CNNs and SVMs, falter with 70-80% accuracy in noisy SEM images, failing to capture the dynamic evolution of etched structures. To address these limitations, we propose a physics-constrained variational level set autoencoder (VLSet-AE) for automated SEM sectional-profile analysis. By integrating physical etching constraints and a three-dimensional framework (time, linewidth, etching depth), VLSet-AE achieves precise contour recognition and nine critical dimensions extraction—scallop depth (2.29%), scallop width (peak-to-peak: 2.05%, valley-to-valley: 6.28%), scallop radius (4.69%), profile angle (0.56%), trench depth (5.46%), bow width (4.35%), mid width (2.43%), and bottom width (4.78%)—with an average error of 3.65% an overall model accuracy of 94.3%, significantly outperforming manual annotation and state-of-the-art alternatives. Compared to seven current models (e.g., CNNs, LSTMs, ResNet), VLSet-AE achieves the shortest training time (20 s), fastest inference time (1.2 s), highest recognition accuracy (96%), and competitive memory usage (50 MB) and parameter count (4.0 million). By enabling efficient, large-scale data acquisition for AI-optimized DRIE processes, VLSet-AE empowers scalable, intelligent manufacturing, unlocking the potential for advanced microfabrication technologies. This approach provides a forward-looking framework for AI-driven MEMS process design and manufacturing, delivering innovative solutions for future AI-assisted microfabrication advancements.

## KEYWORDS

AI for Manufacturing, Deep Reactive Ion Etching, Scanning Electron Microscopy, Artificial Intelligence, Variational Autoencoder, Image Recognition

## INTRODUCTION

Microelectromechanical systems (MEMS) underpin transformative technologies, from smartphones and autonomous vehicles to medical diagnostics, enabling compact, high-performance devices [1]. Deep reactive ion etching (DRIE) plays a pivotal role in MEMS fabrication, serving as a core process for sculpting high-aspect-ratio silicon structures with exceptional precision [2]. DRIE's ability to create intricate features, such as those in sensors and integrated circuits, makes it indispensable for advancing next-generation microsystems.

Although DRIE achieves high precision in the in-plane (lateral) dimensions—accurately replicating mask-defined linewidths—it often exhibits significant variability in the out-of-plane (vertical) direction due to anisotropic etching effects. Phenomena such as scalloping, bowing, and notching commonly arise during deep etching, compromising structural fidelity and introducing uncertainty in depth-dependent features. These challenges stem from the complex interplay between plasma chemistry, ion energy, and cycle timing, making the DRIE process inherently sensitive and difficult to control. Achieving high-aspect-ratio structures in DRIE thus requires precise regulation of etching and passivation cycle times within the Bosch process—a cyclic procedure alternating between etching and passivation steps. Shorter etching cycles, though slower in material removal, enable finer control over etch-passivation alternation, producing smoother sidewalls with minimal scalloping. In contrast, longer cycles increase throughput but result in deeper individual etch steps, generating more pronounced scallop patterns—wavy sidewall irregularities that distort morphology and compromise structural integrity [3]. Compounding these issues, DRIE performance is highly susceptible to operational drift due to factors such as equipment startup/shutdown, chamber conditioning, and tool aging [4]. Consequently, many fabrication lines operate each DRIE system under fixed, recipe-specific configurations tailored to individual MEMS structures, leading to underutilized flexibility and significant resource inefficiencies [5].

Meanwhile, evaluating DRIE outcomes further reflects its complexity, necessitating scanning electron microscopy (SEM) to measure the morphology of etched structures [6]. In a typical MEMS fabrication cleanroom, engineers

meticulously prepare samples—cleaving wafers, mounting specimens, and fine-tuning SEM imaging parameters—to resolve critical features like trench depth and scallop dimensions. Each wafer may yield hundreds of images, with manual analysis requiring hours per image to trace contours and measure dimensions. For large batches, this process extends over weeks, exacerbated by operator fatigue and subjective errors, leading to measurement inaccuracies of 15-20%. This labor-intensive, error-prone workflow underscores the inefficiency of manual SEM analysis and its bottleneck in high-throughput manufacturing [7-8].

Given these challenges, directly optimizing DRIE's nonlinear, multifaceted process poses a formidable challenge due to its intricate parameter interdependencies. Artificial intelligence (AI) provides a transformative solution to address this complexity by leveraging advanced models and large-scale data. How to identify an effective starting point for optimization and acquire large-scale data to support it remains a critical challenge. To address this challenge, we develop AI models to automate feature extraction from SEM images, precisely identifying morphological structures and extracting critical parameters like trench depth and scallop dimensions. These capabilities enable robust correlations between etching outcomes and process parameters, learned through advanced modeling techniques, constructing an AI-driven framework for DRIE optimization, real-time monitoring, and scalable, intelligent manufacturing.

Extracting large-scale feature data from SEM images is a critical technology for developing AI-driven optimization models. However, traditional SEM image analysis methods are limited in accuracy and efficiency. Manual annotation, requiring 1-2 hours per image, incurs 15-20% errors due to operator subjectivity and image variability. Inter-operator discrepancies further erode reproducibility, with critical dimension measurements varying by up to 15% [9]. Existing automated methods, including thresholding (e.g., Canny, Sobel) [10], image recognition, and early machine learning models (e.g., CNNs, SVMs) [18-20], enhance efficiency over manual analysis, achieving accuracies of 70-80% under ideal conditions [11-12], but their limited robustness to noise and complex DRIE morphologies, with performance dropping by 20-30% in noisy, low-contrast SEM images typical of DRIE [13].

Moreover, current recognition methods typically rely on two-dimensional static receptive field scanning—such as pixel-based convolutions over RGB image channels—to identify structural features across entire SEM images [14-15]. These approaches lack spatial adaptivity and are particularly vulnerable to noise and low contrast, leading to frequent misidentification of etched boundaries [16-17]. More critically, they overlook the depth-dependent evolution of DRIE structures, treating profiles as static cross-sections rather than dynamic, layered morphologies [21-23]. In contrast, the proposed level set-based method simulates an adaptive contour evolution process—akin to an expanding bubble that grows from within each structure and halts precisely at material boundaries—thereby capturing depth-varying features with higher accuracy and physical consistency. This evolution-driven mechanism enables robust contour extraction even in noisy or low-contrast SEM images, facilitating large-scale, precise data acquisition essential for AI-driven DRIE process optimization.

To address these gaps, we introduce a physics-constrained variational level set autoencoder (VLSet-AE) for automated contour recognition in DRIE SEM profiles. By leveraging layer-wise scallop segmentation along the etching depth and incorporating physical etching constraints into the decoder, VLSet-AE ensures reconstructed level set functions align with the etching process, enhancing accuracy and consistency. A temporal-scale three-dimensional framework, integrating time, linewidth, and etching depth, enables dynamic spatial and temporal characterization of etching dynamics, with reconstructed scallop segments assembled into complete profiles for precise quantification. VLSet-AE extracts critical dimensions like scallop depth (2.29%), scallop width (peak-to-peak: 2.05%, valley-to-valley: 6.28%), scallop radius (4.69%), profile angle (0.56%), trench depth (5.46%), bow width (4.35%), mid width (2.43%), and bottom width (4.78%). With a 3.65% average error and 94.3% accuracy—outperforming CNN and SVM (70-80% accuracy)—it converges rapidly, enabling real-time process monitoring and advanced etching simulations for scalable, intelligent manufacturing. With a correlation coefficient of 0.998, VLSet-AE enables real-time process monitoring and advanced etching simulations for scalable, intelligent manufacturing. This technology paves a new direction for AI-assisted semiconductor process optimization and intelligent manufacturing, setting the stage for future innovations in high-performance microfabrication and next-generation smart production systems.

## SYSTEM DESIGN AND METHODOLOGY

### A. Experimental Data Collection and Image Preprocessing

The experimental workflow begins with the design of photolithographic mask patterns, followed by photomask fabrication and lithography on 4-inch single-wafer silicon substrates, as illustrated in **Fig. 1**. The patterned wafers are subsequently etched using DRIE performed on SPTS equipment.

To address the inherent complexity of the DRIE process—where key parameters such as etching cycle time, passivation time, substrate temperature, and chamber pressure significantly influence the resulting etch profiles—we employ a 16-run orthogonal experimental design. This approach enables systematic exploration of the multidimensional parameter space and generates a diverse, representative dataset that reflects the process's intrinsic variability, thereby providing a strong foundation for model training and evaluation. From the DRIE experiments, we obtain 1,000 cross-sectional SEM images,

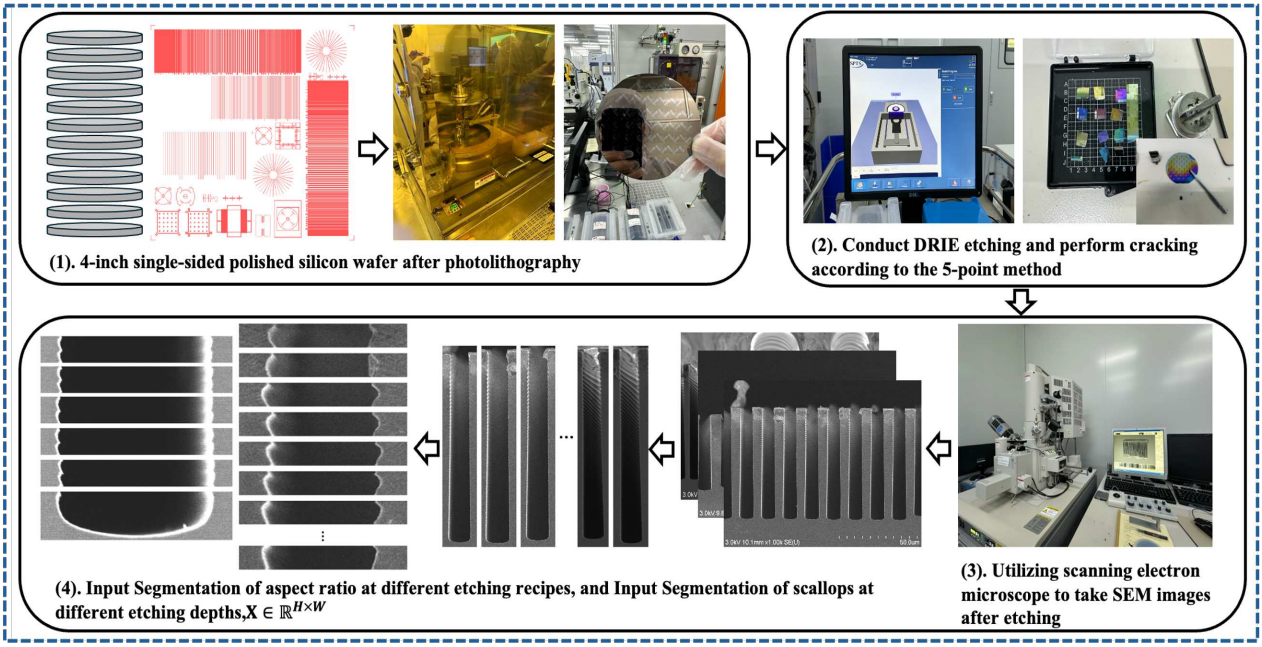



**Fig. 1 Experimental Data Collection and Image Preprocessing.** (1). 4-inch single-sided polished silicon wafer after photolithography. (2). Conducting DRIE etching and perform cracking according to the 5-point method. (3). Utilizing scanning electron microscope to take SEM images after etching. (4). Input Segmentation of aspect ratio at different etching recipes, and Input Segmentation of scallops at different etching depths,$X \in \mathbb{R}^{H\times W}$.

each capturing a unique combination of etching outcomes. To extract meaningful structural features, each image undergoes a tailored preprocessing pipeline. Specifically, we isolate the etched trench region from the background and perform depth-wise segmentation of the scallop structures from top to bottom along the trench axis. These segmented scallop layers—each corresponding to a single etch-passivation cycle—preserve the morphological evolution of the etched profile over time.

This layer-wise scallops not only facilitates fine-grained feature recognition but also establishes a depth-resolved dataset that is ideally suited for training physics-informed models such as VLSet-AE. By capturing both local structural variations and global trench morphology, the resulting dataset enables accurate, scalable, and physically consistent analysis of DRIE processes.

### B. System Architecture of Physics-constrained Variational Level-set Autoencoder

To accurately capture the complex morphological features of deep reactive ion etching (DRIE) profiles from SEM images, we propose a physics-constrained variational level-set autoencoder (VLSet-AE) as illustrated in **Fig. 2a**, the proposed VLSet-AE architecture is built upon a variational autoencoder (VAE), where the encoder transforms high-dimensional SEM images into a compact latent representation z, and the decoder reconstructs the etched structure as a level set function $\phi(x, y)$. Unlike conventional image-based reconstruction that treats contours as pixel boundaries, our method interprets the etched profile as an evolving geometric interface. To ensure that the reconstructed contours reflect the actual physics of DRIE, we introduce the Hamilton-Jacobi equation—a fundamental equation describing interface motion—as a constraint into the decoder's loss function. This enforces that the predicted level set function evolves consistently with how etching interfaces physically propagate (e.g., scallop expansion, sidewall evolution), resulting in contour recognition that is not only more accurate, but also more physically plausible in noisy or irregular SEM images.

Building on this framework, the VLSet-AE architecture consists of a probabilistic encoder–decoder design that allows flexible feature representation and physical constraint embedding. The encoder extracts hierarchical morphological features and maps them to a latent distribution characterized by a mean and variance, enabling stochastic sampling via the reparameterization trick.

The encoder is responsible for compressing high-dimensional SEM image data into a compact latent space representation. Given an input SEM image $\mathcal{X}$, the encoder network $f_{encoder}(\mathcal{X}; \theta_e)$ extracts relevant feature representations and maps them to a lower-dimensional latent space $z \in \mathbb{R}^d$. Instead of directly mapping the input to a deterministic latent vector, the model parameterizes the latent space using a probabilistic distribution. The encoder learns the mean $\mu$ and variance $\sigma^2$ of the latent space representation through a set of neural network layers. Specifically, the latent variables are modeled as a Gaussian distribution where:

$$\mu = W_\mu h + b_\mu, log\sigma^2 = W_\sigma h + b_\sigma$$

Where $h$ is the extracted feature representation from the encoder. To ensure differentiability and allow stochastic sampling, the reparameterization trick is employed:

$$z = \mu + \epsilon \cdot \sigma, \epsilon \sim \mathcal{N}(0,1), \sigma = \sqrt{\exp(log\sigma^2)}$$

This formulation allows backpropagation through the stochastic sampling process, ensuring that the network can be efficiently optimized using gradient-based methods.

Once the SEM image is encoded into the latent space, the decoder reconstructs the level set function $\phi(x, y)$ from the latent variables. The decoder $f_{decoder}(z; \theta_d)$ learns to map the latent representation back to a structured feature space that defines the morphological contours of the etched structure. This reconstruction involves two key steps: first, the decoder generates an intermediate feature map $h'$ as:

$$h' = f_{decoder}(z; \theta_d)$$

Then, the level set function is computed as:

$$\phi(x, y) = W_\phi h' + b_\phi$$

The function $\phi(x, y)$ represents the contour of the SEM image, where the zero level set $\phi(x, y) = 0$ defines the

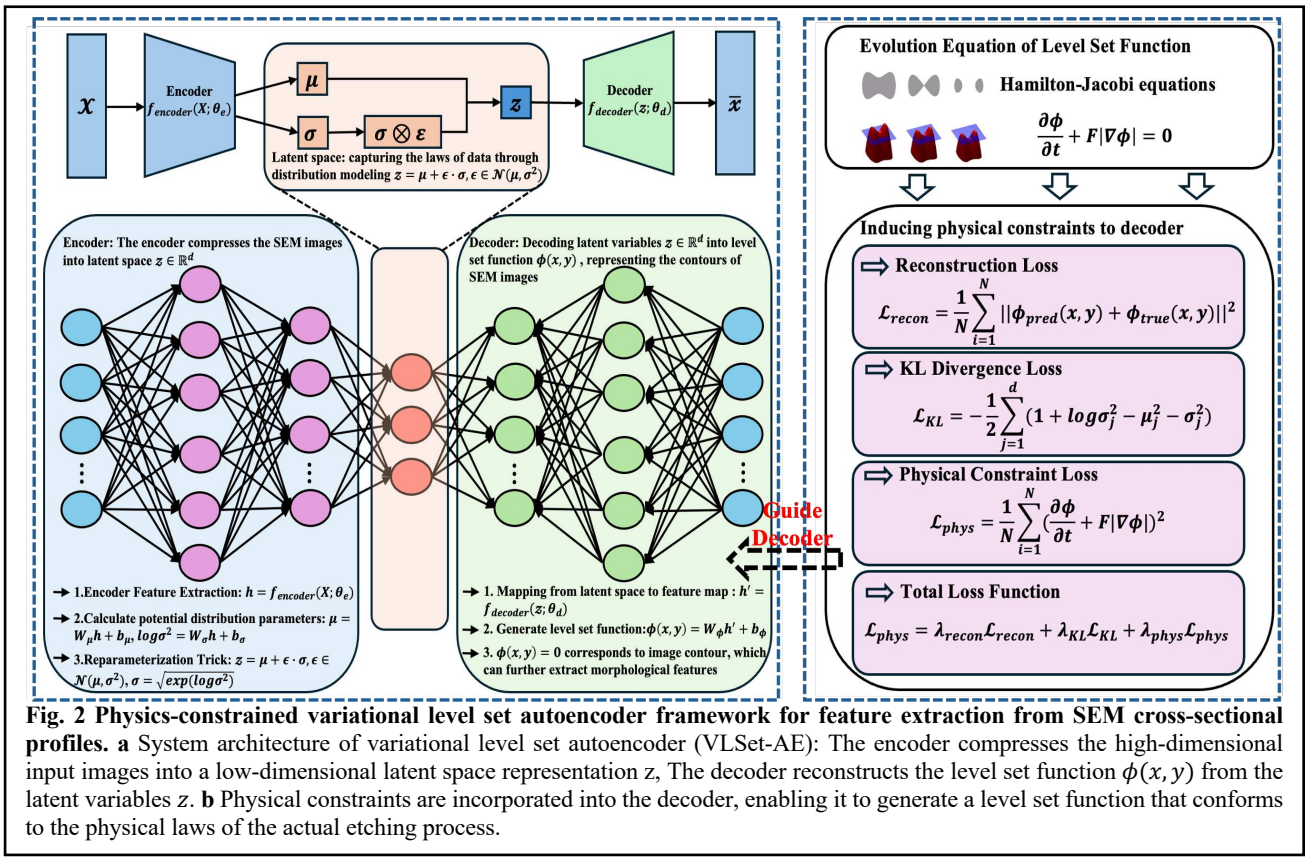



**Fig. 2 Physics-constrained variational level set autoencoder framework for feature extraction from SEM cross-sectional profiles. a** System architecture of variational level set autoencoder (VLSet-AE): The encoder compresses the high-dimensional input images into a low-dimensional latent space representation z, The decoder reconstructs the level set function $\phi(x, y)$ from the latent variables $z$. **b** Physical constraints are incorporated into the decoder, enabling it to generate a level set function that conforms to the physical laws of the actual etching process.

boundaries of the etched structure. This approach allows the model to accurately capture the morphology of the scallops and trenches that arise from the DRIE process.

A unique aspect of VLSet-AE is the incorporation of physical constraints into the decoder's loss function, as shown in **Fig. 2b**. By embedding the Hamilton-Jacobi equation, which governs the evolution of level set functions, the model ensures that the generated contours remain physically meaningful. The Hamilton-Jacobi equation provides a dynamic evolution framework for level set functions, ensuring that the extracted features correspond to real physical phenomena observed in the etching process. This constraint enforces smoothness and consistency in the extracted contours, reducing artifacts that may arise due to image noise or improper segmentation.

### C. Design of Encoder and Decoder in VLSet-AE

In the VLSet-AE model, the core of both the encoder and decoder is based on the Variational Autoencoder (VAE), integrated with the theory of level set functions. This approach maps SEM images into a low-dimensional latent space representation $z$, from which the level set function $\phi(x, y)$ is decoded and generated.

The objective of the encoder is to map the input SEM image $X \in \mathbb{R}^{H \times W}$ into the distribution parameters $(\mu, \sigma^2)$ of a low-dimensional latent space representation $z \in \mathbb{R}^d$, where $\mu$ represents the mean vector, indicating the center of the latent variables, and $\sigma^2$ represents the variance vector, reflecting the uncertainty of the latent variables. The encoding process begins with feature extraction from the input image X using a series of convolutional operations, denoted as $h =$

$f_{encoder}(X;\theta_e)$, where $f_{encoder}$ is a nonlinear mapping function (such as CNN or multi-layer perceptron), and $\theta_e$ represents the corresponding parameters. Subsequently, the latent distribution parameters are computed as $\mu = W_\mu h + b_\mu$, and $log\sigma^2 = W_\sigma h + b_\sigma$, where $W_\mu$, $W_\sigma$, $b_\mu$, $b_\sigma$ are trainable weights and biases, and $log\sigma^2$ is used to avoid directly optimizing the variance, thus enhancing numerical stability. To sample the latent variable $z \sim \mathcal{N}(\mu, \sigma^2)$, the reparameterization trick is applied, yielding $z = \mu + \epsilon \cdot \sigma$, where $\epsilon \sim \mathcal{N}(0,1)$ is noise sampled from a standard normal distribution, and $\sigma = \sqrt{\exp(log\sigma^2)}$.

The objective of the decoder is to transform the latent variable $z \in \mathbb{R}^d$, into a level set function $\phi(x,y)$, which implicitly represents the contour of the SEM image. The decoder begins by mapping the latent variable $z$ into a two-dimensional feature map $h' \in \mathbb{R}^{H' \times W'}$ through a nonlinear transformation $h' = f_{decoder}(z;\theta_d)$, where $f_{decoder}$ represents the decoder's mapping function (such as a deconvolutional network or multi-layer perceptron), and $\theta_d$ denotes the associated parameters. The decoder then generates the level set function $\phi(x,y)$, typically represented as a two-dimensional tensor, via the equation $\phi(x,y) = W_\phi h' + b_\phi$, where $W_\phi$ and $b_\phi$ are the weights and biases of the final layer of the decoder. The function $\phi(x,y)$ is used to implicitly encode the contour, with $\phi(x,y) = 0$ indicating the boundary of the shape. The decoder thus reconstructs the level set function $\phi(x,y)$ from the latent variable $z$, facilitating the generation of high-dimensional images from low-dimensional representations. The zero level set $\phi(x,y) = 0$ corresponds to the image contour, enabling further extraction of morphological features.

### D. Physical Constraints in Level-set Function of VLSet-AE

The physical constraints in the level set function are a key innovation of VLSet-AE model. By introducing these constraints, the model is able to generate level set functions $\phi(x,y)$ that adhere to the physical laws governing the actual etching process, ensuring the physical consistency of the contours. These constraints are based on the principles of the Hamilton-Jacobi equation and the level set method, integrated within a deep learning framework for optimization.

The level set method represents the position of the interface using an implicit function $\phi(x,y,t)$, where $\phi(x,y,t) = 0$ defines the exact interface location (i.e., the contour), $\phi(x,y,t) > 0$ represents the exterior, and $\phi(x,y,t) < 0$ represents the interior of the interface. In the etching process, the interface evolves over time, primarily influenced by factors such as the normal velocity $F$, which depends on etching parameters like power, gas flow, and ion energy, as well as the directional nature of the etching process, which may be isotropic (uniform etching speed) or anisotropic (speed dependent on direction). This dynamic behavior is described by the Hamilton-Jacobi equation:

$$\frac{\partial \phi}{\partial t} + F|\nabla \phi| = 0$$

where $\frac{\partial \phi}{\partial t}$ represents the rate of change of the level set function over time, indicating the movement of the interface, and $|\nabla \phi|$ represents the gradient magnitude of the level set function, describing the direction and magnitude of the interface change.

The goal of introducing the physical constraints is to ensure that the level set function $\phi(x,y)$ generated by the decoder is consistent with the actual physical laws of the etching process, preventing the generation of contours that do not align with the real etching behaviour, such as unrealistic waviness or improper contour shapes. These physical constraints are incorporated into the model's loss function through the Hamilton-Jacobi equation, as shown by the physical loss term:

$$\mathcal{L}_{phys} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\partial \phi}{\partial t} + F|\nabla \phi|\right)^2$$

where $\mathcal{L}_{phys}$ evaluates whether the generated level set function adheres to the physical principles, and $F$ represents the etching normal velocity, which can be a constant, variable, or a function of spatial coordinates. The gradient magnitude $|\nabla \phi|$ describes the interface change direction.

To compute the necessary derivatives in the deep learning framework, automatic differentiation is employed to calculate the time derivative $\frac{\partial \phi}{\partial t}$ and spatial gradients $|\nabla \phi|$, where $|\nabla \phi|$ is derived from the spatial coordinates $(x,y)$ as:

$$|\nabla \phi| = \sqrt{\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2}$$

The physical constraints are incorporated into the total loss function, which is minimized to ensure that the decoder generates level set functions that follow etching laws.

### E. Loss function Design

To train the encoder-decoder architecture, a composite loss function is formulated that integrates the reconstruction loss, KL divergence, and physical consistency loss, each weighted by corresponding coefficients, as shown in **Fig. 2b**.

The reconstruction loss $\mathcal{L}_{recon}$ is designed to quantify the discrepancy between the predicted level set function and the ground truth contour. Mathematically, it is defined as

$$\mathcal{L}_{recon} = \frac{1}{N}\sum_{i=1}^{N}\left\|\phi_{pred}(x,y) - \phi_{true}(x,y)\right\|^2$$

Where $N$ is the number of samples. This loss function penalizes differences between the generated contour $\phi_{pred}(x,y)$ and the true contour $\phi_{true}(x,y)$, thereby encouraging the decoder to produce accurate and detailed reconstructions that preserve the salient morphological features of the original SEM images.

The KL divergence $\mathcal{L}_{KL}$ term is introduced to regularize the latent space. Specifically, it measures the divergence between the posterior distribution of the latent variables and a pre-defined standard normal prior. It is given by

$$\mathcal{L}_{KL} = -\frac{1}{2}\sum_{j=1}^{d}(1 + log\sigma_j^2 - \mu_j^2 - \sigma_j^2)$$

Where $d$ is the dimensionality of the latent space, and $\mu_j$ and $\sigma_j^2$ represent the mean and variance of the latent variables, respectively. This loss term serves to constrain the latent space, ensuring that the encoded representations are both smooth and continuous, which is critical for the generative performance of the model.

The physical consistency loss $\mathcal{L}_{phys}$ enforces that the generated level set functions adhere to the physical laws underlying the etching process. Based on the Hamilton-Jacobi equation, the loss is formulated as

$$\mathcal{L}_{phys} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\partial\phi}{\partial t} + F|\nabla\phi|\right|^2$$

Where $\frac{\partial\phi}{\partial t}$ denotes the temporal derivative of the level set function, $F$ is the normal etching velocity, and $|\nabla\phi|$ is the magnitude of the spatial gradient of $\phi$. This term penalizes deviations from the expected physical behavior dictated by the Hamilton-Jacobi equation, thereby ensuring that the model-generated contours are not only mathematically consistent but also physically plausible.

The total loss function used to train the model is a weighted sum of the three aforementioned components:

$$\mathcal{L}_{total} = \lambda_{recon}\mathcal{L}_{recon} + \lambda_{KL}\mathcal{L}_{KL} + \lambda_{phys}\mathcal{L}_{phys}$$

Here, $\lambda_{recon}$, $\lambda_{KL}$, and $\lambda_{phys}$ are hyperparameters that control the relative importance of the reconstruction, KL divergence, and physical consistency losses, respectively. By optimizing this composite loss function, the model is encouraged to generate high-fidelity reconstructions of SEM image contours.

### F. Adaptive Contour Recognition for Segmented Scallop Layers

The segmented scallop layers obtained from SEM images are subsequently analyzed using the VLSet-AE model for fine-grained feature extraction and morphological characterization. For each scallop segment, the level set function $\phi(x,y)$ is initialized at the geometric center of the layer and undergoes a dynamic outward evolution, analogous to the expansion of an inflating balloon, as shown in **Fig. 3**. This propagation continues until the evolving interface encounters the sidewalls of the etched structure, at which point the expansion automatically halts. This self-regulating and

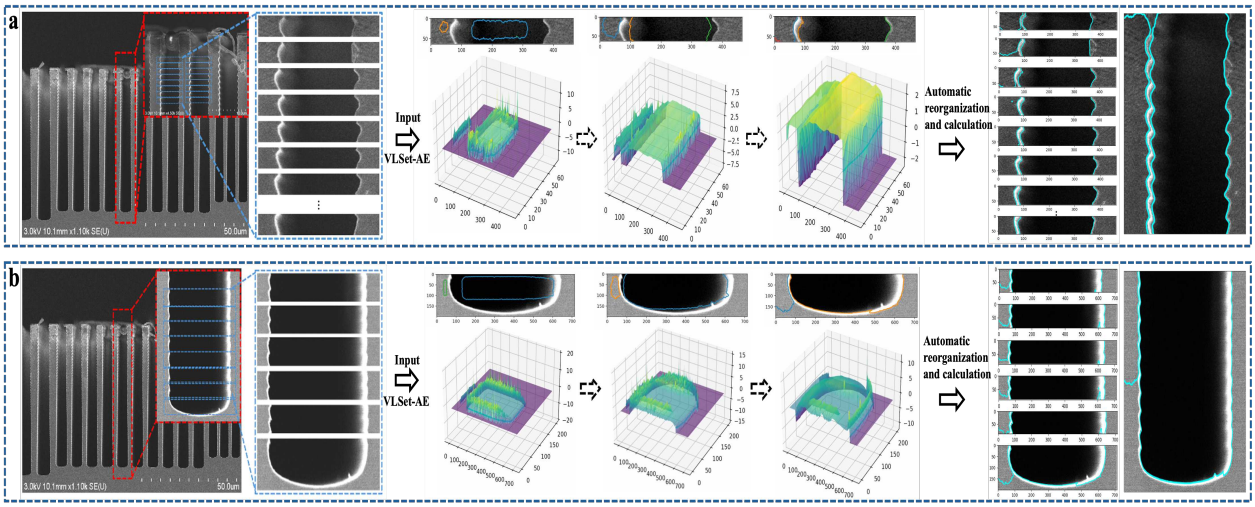



**Fig. 3 Contour Recognition and Evolution Visualization: The SEM images are segmented along the etching depth into individual scallop layers, enabling the extraction of multi-layer scallop features and the automated calculation of contour edges and feature dimensions using the VLSet-AE model. a** Input segmented scallop images from the top and middle sections of SEM along etching depth into the VLSet-AE, which automatically recognizes and extracts features of cross-sectional profiles. **b** Input segmented scallop images from the bottom sections of SEM along etching depth into the VLSet-AE, which automatically recognizes, and extracts features of cross-sectional profiles.

adaptive evolution mechanism enables the model to delineate structural boundaries with high precision in real time. By initiating contour growth from within the feature and terminating it upon boundary contact, the method naturally suppresses over-segmentation and enhances robustness against image noise and artifacts. As a result, it effectively captures the intricate morphological details of scallop formations, including subtle curvature variations and edge transitions. Each scallop layer is processed independently using this mechanism, and the extracted morphological descriptors—such as trench depth, scallop depth, and sidewall roughness—are subsequently assembled to reconstruct a complete etching profile. This layer-wise, contour-driven reconstruction approach provides a high-resolution, physically consistent understanding of the DRIE process, offering valuable insights into etch uniformity, structural integrity, and process stability across depth.

**Figure 3** presents a visual demonstration of the VLSet-AE in contour extraction and morphological analysis of DRIE profiles from scanning electron microscope (SEM) images, which consists of two cases, labeled a and b, each illustrating the application of VLSet-AE to different etching profiles, corresponding to distinct etching structural variations.

In both cases, the leftmost SEM images depict etched trenches with distinct scallop formations, characteristic of the Bosch process in DRIE. A magnified region of interest (ROI) is highlighted in red, where the structure undergoes segmentation into individual scallop layers, visualized as a set of horizontal slices extracted along the trench depth. These slices serve as input for VLSet-AE, where each scallop layer is independently analyzed and reconstructed through the model's adaptive level set evolution. The middle section of the **Fig. 2** illustrates the level set function evolution in a 3D representation. The extracted contours are transformed into 3D surface plots, where the expansion dynamics of the level set function can be observed. As described earlier, the level set function expands outward like an inflating balloon, adjusting adaptively as it interacts with the sidewalls of the trench. The collision of the evolving level set with the trench walls is evident, ensuring an accurate delineation of the etching boundaries. On the right side of the figure, the results of the automatic contour extraction and feature calculation are displayed. The processed scallop layers are realigned and reconstructed, forming a complete etching profile with well-defined contours. The extracted profiles, overlaid on the original SEM images, confirm the precise identification of scallop boundaries, with cyan-colored contours clearly marking the structural edges. This automatic organization enables the direct calculation of key morphological parameters, including trench depth, scallop depth, and so on, essential for evaluating etching uniformity and process stability.

By leveraging the VLSet-AE model, this approach establishes a solid foundation for the automated extraction of SEM-based feature parameters. Specifically, our proposed VLSet-AE model enables precise contour recognition of multi-segment scallops and reconstructs the identified segments to obtain a complete etched profile, laying the groundwork for subsequent automated feature detection. This facilitates advanced morphological analysis and intelligent process optimization in semiconductor manufacturing.

## G. Temporal-scale Three-dimensional Etching Trench Feature Recognition, Extraction and Calculation

Building upon the contour recognition results obtained through the VLSet-AE model, as illustrated in **Fig. 4**, we introduce a three-dimensional contour feature recognition model that characterizes etching trench features within a temporal-scale framework. This model constructs a three-dimensional representation where time sequence serves as the x-

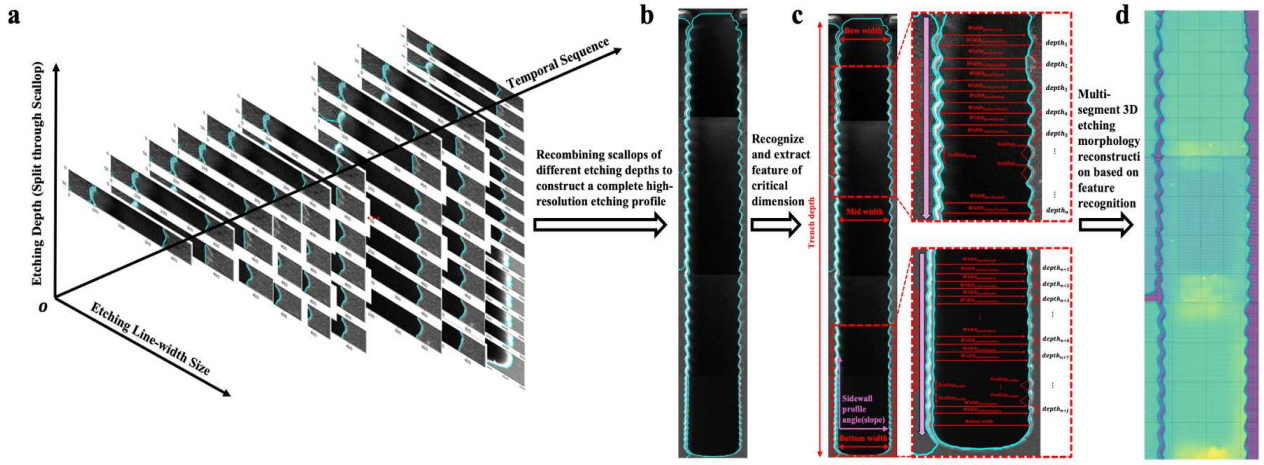



**Fig. 4 A temporal-scale three-dimensional etching trench feature recognition, extraction and calculation framework based on VLSet-AE model. a** Temporal-scale three-dimensional etching feature recognition framework based on VLSet-AE is constructed, with time as the x-axis, linewidth as the y-axis, and etching depth as the z-axis. **b** The identified scallop segments are reconstructed into a complete etching profile. **c** Recognition and extraction of critical dimensions. Various critical dimensions of the etched profile are automatically recognized and extracted by the model. **d** Multi-segment 3D etching morphology reconstruction. By integrating the features from all depths, the model produces a full 3D morphology of the trench structure, allowing for detailed analysis of the etched profile.

axis, etching linewidth as the y-axis, and etching depth as the z-axis, enabling a comprehensive understanding of the dynamic evolution of the etching process.

To achieve a complete representation of the etched structure, the identified scallop contours are reassembled along the etching depth direction, effectively reconstructing the full trench morphology, as shown in Fig. 3b. During this process, the model not only extracts and reconfigures individual scallop segments but also simultaneously computes critical dimension parameters at both the scallop and trench levels. For each scallop, the model quantifies key features, including scallop depth, scallop radius, peak-to-peak width, and valley-to-valley width. At the trench scale, it further calculates nine essential structural parameters, such as trench opening width, mid-depth width, bottom width, overall trench depth, and sidewall angle, providing a detailed characterization of the etching profile, as shown in Fig. 3c. As the model automates the calculation of key feature parameters, it simultaneously simulates the 3D depth information of the trench structure. Through the evolution of the level set function within the model, the scallop segments at different depths are progressively integrated, allowing for a comprehensive view of both vertical and lateral features, as shown in Fig. 3d. By integrating features from all depths, the VLSet-AE model generates a complete 3D morphology of the trench structure, facilitating a detailed analysis of the etched profile. This dynamic process ensures that the model captures the intricate etching dynamics and accurately reconstructs the structure's final form.

This systematic feature extraction framework ensures a comprehensive, high-fidelity reconstruction of the trench morphology, facilitating an accurate, automated, and quantitative analysis of the etching process. By integrating dynamic contour recognition with multi-scale feature analysis, this framework establishes a robust foundation for data-driven etching process optimization and predictive modeling, enabling precise control over semiconductor microfabrication processes.

## RESULTS AND DISCUSSION

### A. Visual Comparison of Contour Recognition: Traditional CNN vs. Proposed VLSet-AE Model

To further highlight the advantages of the proposed physics-constrained variational level set autoencoder (VLSet-AE), we present a visual comparison of its contour recognition performance against the conventional convolution neural network (CNN) receptive field convolution method, as illustrated in **Fig. 5**. The CNN approach (Fig. 5a) relies on a static two-dimensional receptive field scanning mechanism (Fig. 5a1), where a convolution kernel extracts local features through multi-layer operations. However, when applied to SEM images of DRIE etched profiles (Fig. 5a2–a3), the CNN method struggles to capture intricate scallop structures, resulting in discontinuous and blurry contours (green lines), particularly in noisy regions. This limitation stems from the lack of physical constraints and depth-aware feature extraction, leading to suboptimal recognition accuracy in complex morphologies.

Specifically, Fig. 5a2 illustrates the application of the traditional CNN method, revealing significant issues in contour recognition within SEM images of etched profiles. Notable problems include repeated contour identification and incomplete edge detection, such as the failure to recognize the contour structure at the etched opening, where identification gaps are evident. Furthermore, in the recognition of individual scallop segments, the CNN method produces multiple overlapping boundary contours. This issue arises because the traditional receptive field-based convolution operates as a "sequential" scanning method, which is prone to repeated identification errors at complex boundaries, leading to misinterpretations of the etched structure. Similarly, Fig. 5a3 demonstrates the overall structural contour recognition using the CNN method, further exposing its limitations. In Fig. 5a3(2), there are evident issues with missing contours in the overall structure, exacerbated by constraining factors such as the SEM image's shooting angle, light intensity, and clarity, which degrade the CNN's ability to accurately capture the etched profile. Additionally, in Fig. 5a3(3) and Fig. 5a3(4), boundary recognition errors and automatic feature calculation inaccuracies are apparent, where the extracted contours fail to align with the true edges of the scallop structures, leading to erroneous morphological parameter estimation. Finally, Fig. 5a3(5) employs a CNN-based semantic segmentation model, which struggles to distinguish between different etched boundaries, resulting in unclear boundary feature identification and semantic segmentation errors. These issues highlight the CNN's challenges in handling the complex and variable nature of DRIE SEM images, particularly in distinguishing subtle morphological differences.

In contrast, the proposed VLSet-AE (Fig. 5b) leverages a physics-constrained level set framework to achieve superior contour recognition. As shown in Fig. 5b1, the level set function initializes at the center of the scallop layer and evolves dynamically through multiple stages, adaptively expanding until it reaches the contour edges, where it automatically stops. This process, governed by the Hamilton-Jacobi equation, ensures that the generated contours align with the physical dynamics of the etching process. The evolution of contour area, perimeter, and speed norm (Fig. 5b1, right) demonstrates the stability and controllability of the level set function, with the area and perimeter converging smoothly and the speed norm decreasing steadily over iterations. Unlike the traditional "sequential" convolution scanning method, our adaptive level set function evolution approach is better suited for precise contour recognition in irregular images. In various SEM images of etched profiles, we allow the level set function to adaptively identify an appropriate initial evolution position before iteratively evolving. This evolution is regulated through a velocity field, enabling controlled adaptation. As observed

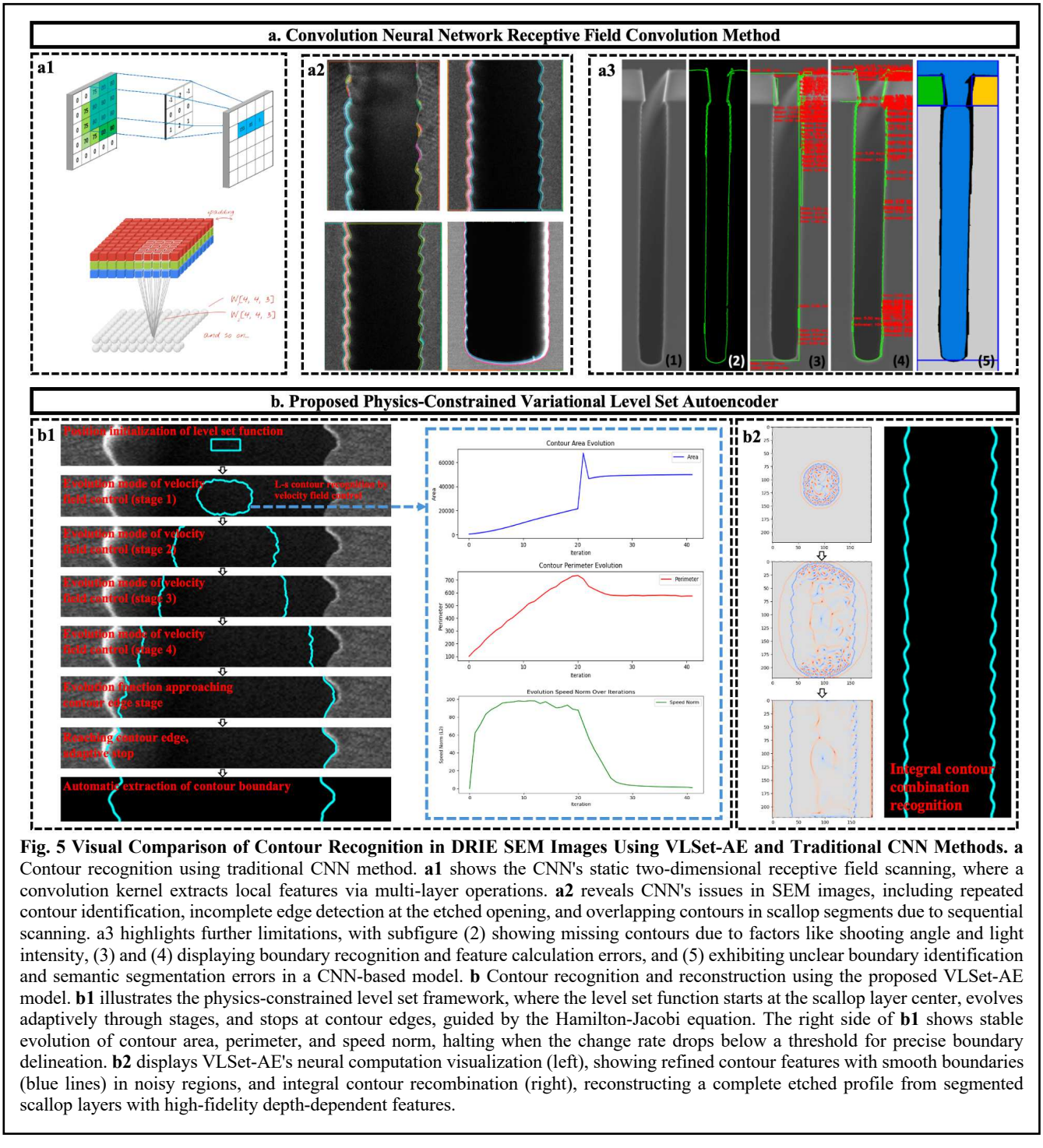



**Fig. 5 Visual Comparison of Contour Recognition in DRIE SEM Images Using VLSet-AE and Traditional CNN Methods. a** Contour recognition using traditional CNN method. **a1** shows the CNN's static two-dimensional receptive field scanning, where a convolution kernel extracts local features via multi-layer operations. **a2** reveals CNN's issues in SEM images, including repeated contour identification, incomplete edge detection at the etched opening, and overlapping contours in scallop segments due to sequential scanning. a3 highlights further limitations, with subfigure (2) showing missing contours due to factors like shooting angle and light intensity, (3) and (4) displaying boundary recognition and feature calculation errors, and (5) exhibiting unclear boundary identification and semantic segmentation errors in a CNN-based model. **b** Contour recognition and reconstruction using the proposed VLSet-AE model. **b1** illustrates the physics-constrained level set framework, where the level set function starts at the scallop layer center, evolves adaptively through stages, and stops at contour edges, guided by the Hamilton-Jacobi equation. The right side of **b1** shows stable evolution of contour area, perimeter, and speed norm, halting when the change rate drops below a threshold for precise boundary delineation. **b2** displays VLSet-AE's neural computation visualization (left), showing refined contour features with smooth boundaries (blue lines) in noisy regions, and integral contour recombination (right), reconstructing a complete etched profile from segmented scallop layers with high-fidelity depth-dependent features.

in Fig. 5b1 (right), when the evolution function's speed or the recognized area falls below a predefined change rate threshold, the process adaptively halts, effectively reaching the contour boundary. This mechanism ensures robust and accurate delineation of complex etched structures, even in the presence of noise and morphological irregularities. Fig. 5b2 provides a visualization of the neural computation process and the final contour recognition results. On the left side of Fig. 5b2, we present the neural learning visualization of the algorithm model applied to the overall structural contours of the SEM image. This visualization reveals how the model iteratively learns and refines the contour features, accurately capturing the intricate scallop patterns with continuous and smooth boundaries (blue lines), even in regions with high noise or complex morphology. This demonstrates the model's capability to generalize across diverse etched profiles while maintaining high precision. On the right side, the integral contour recombination (Fig. 5b2, right) enables the reconstruction of a complete etched profile by assembling segmented scallop layers, capturing depth-dependent features with high fidelity. By integrating the segmented contours along the etching depth, VLSet-AE ensures a comprehensive representation of the complex morphology, providing a solid foundation for subsequent quantitative analysis and process optimization in DRIE applications.

## B. Critical Dimension Performances Evaluation of VLSet-AE for DRIE Etched Profile Analysis

In this section, we present a comprehensive performance evaluation of the VLSet-AE model by comparing its predictions with manual measurements, focusing on error distribution, model comparison, training stability, and prediction accuracy across various dimensions.

To ensure the reliability and reproducibility of our experiments, all algorithm models, including the proposed VLSet-AE and comparative models (CNN, LSTM, SVM, Random Forest, ResNet, GoogleNet, and AttentionNet), were evaluated on a high-performance computing setup. Specifically, the experiments were conducted on a workstation equipped with an NVIDIA RTX4060Ti GPU, featuring 16 GB of GDDR6 memory, 4352 CUDA cores, and a base clock speed of 2.31 GHz (boost up to 2.54 GHz). The system was paired with an AMD Ryzen 7 5800X CPU (8 cores, 16 threads, 3.8 GHz base clock, boost up to 4.7 GHz), 32 GB of DDR4 RAM (3200 MHz), and ran on Ubuntu 22.04 LTS with CUDA 12.2 and PyTorch 2.0 for model implementation. This hardware configuration provided sufficient computational power to handle the intensive neural computations and large-scale SEM image processing required for our study.

The evaluation begins with the analysis of nine critical dimensions extracted by VLSet-AE, such as scallop depth, scallop width (peak-to-peak and valley-to-valley), scallop radius, total profile angle, trench depth, bow width, mid width, and bottom width. As shown in **Fig. 6a**, a schematic diagram of the etched profile visually annotates these dimensions on a cross-sectional SEM image, clarifying their locations and definitions within the etched structure. This serves as a foundational illustration for understanding the feature parameters extracted by the model. We then compute the numerical values of these feature parameters using VLSet-AE and compare them against manual measurements to analyze error metrics.

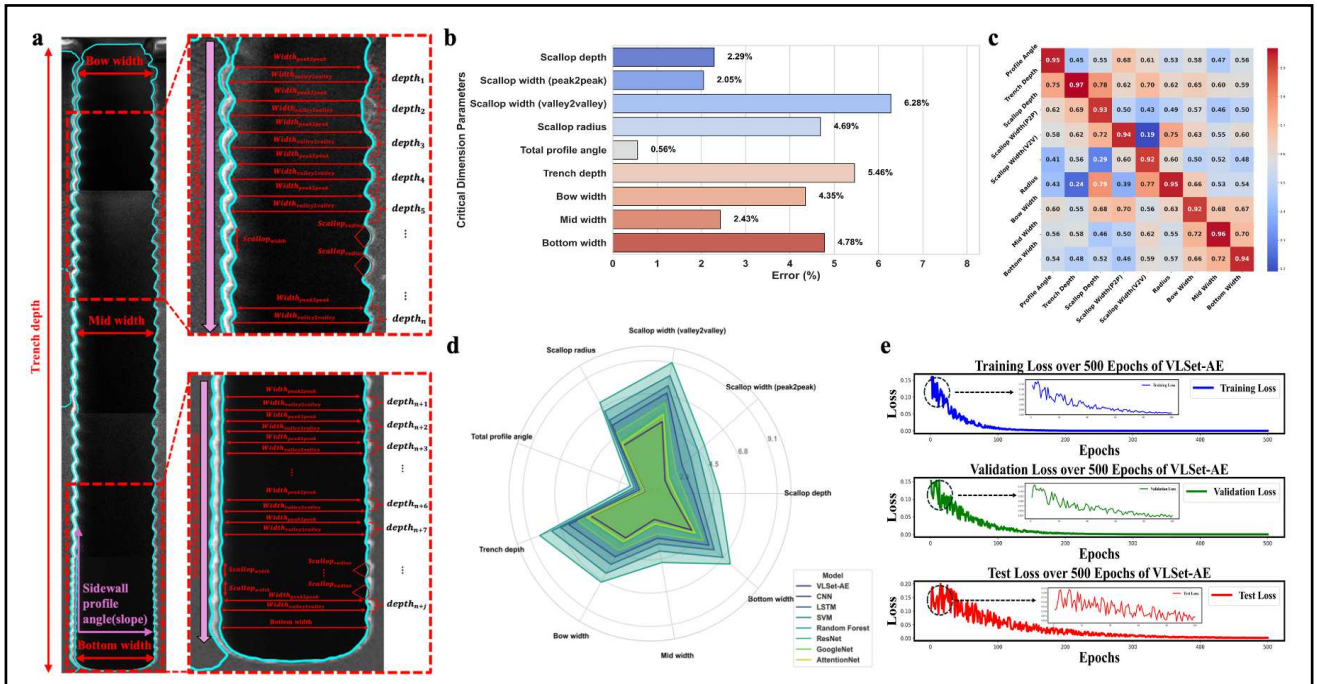



**Fig. 6 Critical Dimension Performances Evaluation of VLSet-AE for DRIE Etched Profile Analysis. a** Recognition and error analysis of nine critical dimensions using VLSet-AE, including scallop depth, scallop width (peak-to-peak and valley-to-valley), scallop radius, total profile angle, trench depth, bow width, mid width, and bottom width. **b** The VLSet-AE model's computed values are compared directly against manual measurement data. **c** Training, validation, and test loss trajectories over 500 epochs, demonstrating VLSet-AE's stable convergence. **d** Radar chart comparing error rates of VLSet-AE and other models across nine dimensions. **e** Confusion matrix of VLSet-AE predictions vs. manual measurements for nine dimensions.

**Figure 6b** illustrates the prediction errors for nine critical dimension parameters—scallop depth, scallop width (peak to peak), scallop width (valley to valley), scallop radius, total profile angle, trench depth, bow width, mid width, and bottom width—under one etching recipe. The VLSet-AE model's computed values are compared directly against manual measurement data. Notably, the total profile angle achieves the lowest error (approximately 0.56%), suggesting that the model excels at capturing angular features. By contrast, the scallop width (valley–valley) exhibits the highest deviation (around 6.28%), likely reflecting increased geometric complexity and measurement sensitivity in that dimension. The remaining parameters show intermediate error values (2–5%), underscoring the model's overall consistency and reliability in handling diverse critical dimensions within this specific etching context.

**Figure 6c** displays the confusion matrix for nine different critical dimension parameters, comparing the VLSet-AE model's automatic predictions with manual measurements. Each cell in the matrix represents the correspondence between a predicted parameter (horizontal axis) and the actual parameter (vertical axis), with darker shades of red indicating higher accuracy and lighter or blue shades reflecting greater misclassification. Notably, the diagonal entries show relatively strong

agreement for most dimensions, such as bow width and profile angle, underscoring the model's effectiveness in accurately identifying and quantifying these parameters. By contrast, certain off-diagonal cells reveal moderate confusion among geometrically similar features—particularly those with subtle morphological differences (e.g., scallop width in its peak–peak and valley–valley definitions). These instances of cross-parameter misclassification suggest that while VLSet-AE robustly captures most topographical nuances, additional refinement or more extensive training data may further mitigate overlaps in closely related dimensions. Such the confusion matrix underscores the VLSet-AE model's capacity to discern distinct etching parameters.

Model comparison is conducted using a radar chart to evaluate the average error rates of VLSet-AE against other models (CNN, LSTM, SVM, Random Forest, ResNet, GoogleNet, AttentionNet) across nine dimensions under 16 orthogonal etching recipes as in **Fig. 6d**. VLSet-AE exhibits the smallest error polygon, outperforming others, which show higher errors and reduced robustness for complex etching predictions. A complementary radar chart details the prediction errors for each model-dimension pair, confirming VLSet-AE's lowest errors in dimensions like scallop depth and trench depth, with competitive performance in challenging parameters like scallop width (valley-to-valley).

The training stability of VLSet-AE is assessed through its training, validation, and test loss trajectories over 500 epochs for the same etching recipe as in **Fig. 6e**. The training loss (blue) decreases rapidly and stabilizes, showing efficient learning with minimal overfitting, while the validation (green) and test (red) losses converge to stable plateaus, demonstrating robust generalization across datasets and the model's ability to capture complex topographical features.

## C. Performance Testing and Evaluation of Different Algorithm Models

This section evaluates the performance of the VLSet-AE model against seven advanced algorithms—CNN, LSTM, SVM, Random Forest, ResNet, GoogleNet, and AttentionNet—across multiple metrics to demonstrate its effectiveness in SEM image analysis for DRIE etched profiles. The evaluation encompasses training stability, prediction accuracy, error correlation, residual analysis, and computational efficiency.

To assess the training stability of VLSet-AE and competing models, we analyse the training, validation, and test loss trajectories over 500 epochs under the same etching recipe. **Fig. 7a** compares the training, validation, and test loss curves over 500 epochs for the eight models under the same etching recipe experimental condition. While all methods show decreasing loss trajectories, VLSet-AE converges more quickly and attains lower final loss values, suggesting stronger feature extraction and better generalization. CNN and ResNet also achieve relatively stable results but remain above VLSet-AE's loss plateau. In contrast, methods like SVM and LSTM exhibit more pronounced fluctuations or slower convergence, highlighting their limitations when computing intricate etching critical dimensions.

**Figure 7b** displays a correlation analysis of automatic calculation values versus actual measurements for eight models. Each data point represents an individual measurement, with the horizontal axis indicating the actual value and the vertical axis showing the model's predicted value. The ideal line ($y = x$) signifies perfect agreement between prediction and measurement. Overall, all models exhibit high linear correlation, as evidenced by data points clustering around the diagonal. Among them, VLSet-AE achieves the highest correlation coefficient ($R = 0.998$), indicating that its predictions align most closely with the actual measurements. Models such as CNN and LSTM also display relatively strong correlations ($R = 0.995$ and $R = 0.991$, respectively), though they still fall slightly short of VLSet-AE's performance. Figure 4h shows the corresponding residual analysis, where each point's vertical position represents the difference between the automatic calculation value and the actual measurement (Automatic calculation value – Actual), and the horizontal axis indicates the actual measurement. The horizontal line at zero denotes perfect alignment between the automatic calculation value and the actual value. Points scattered above this line suggest overestimation, whereas points below it reflect underestimation. As the plot illustrates, VLSet-AE's residuals remain tightly clustered around zero across the entire measurement range, indicating more precise and consistent automatic calculations. While other models also exhibit reasonable distributions of residuals, they generally show broader dispersion or occasional bias, especially at higher or lower measurement values. Taken together, **Fig. 7g** and **Fig. 7h** reinforce that VLSet-AE not only correlates strongly with actual data but also maintains minimal calculation error, thereby demonstrating robust and reliable performance across diverse measurement ranges.

We compare the computational complexity and performance of eight models across multiple metrics—namely training time, inference time, memory usage, parameter count, and accuracy—using a radar chart for visualization. For the four metrics where lower values indicate better performance (training time, inference time, memory usage, and parameter count), we compute an efficiency score according to

$$efficiency = \frac{\min\left(Index_Metric\right)}{Index_Metric}$$

This normalization method transforms these metrics into efficiency scores, where higher values denote superior performance. For instance, if VLSet-AE achieves the shortest raw training time (e.g., 20 s), then its normalized value is 1, denoting the best training efficiency among all models. On the radar chart, a score of 1 appears closest to the outer boundary—visually the "longest" radius—yet in practical terms signifies the shortest training duration and thus superior performance.

By contrast, for the "accuracy" metric (where higher is better), we employ min–max normalization:

$$Normalized\ Accuracy = \frac{Accuracy - \min(Accuracy)}{\max(Accuracy) - \min(Accuracy)}$$

Through this approach, all metrics are rescaled to lie between 0 and 1, with higher values corresponding to better overall performance. After normalization, each metric is plotted on the radar chart, providing a direct comparison of each model's comprehensive performance; larger areas on the chart thus indicate higher efficiency and stronger predictive capabilities.

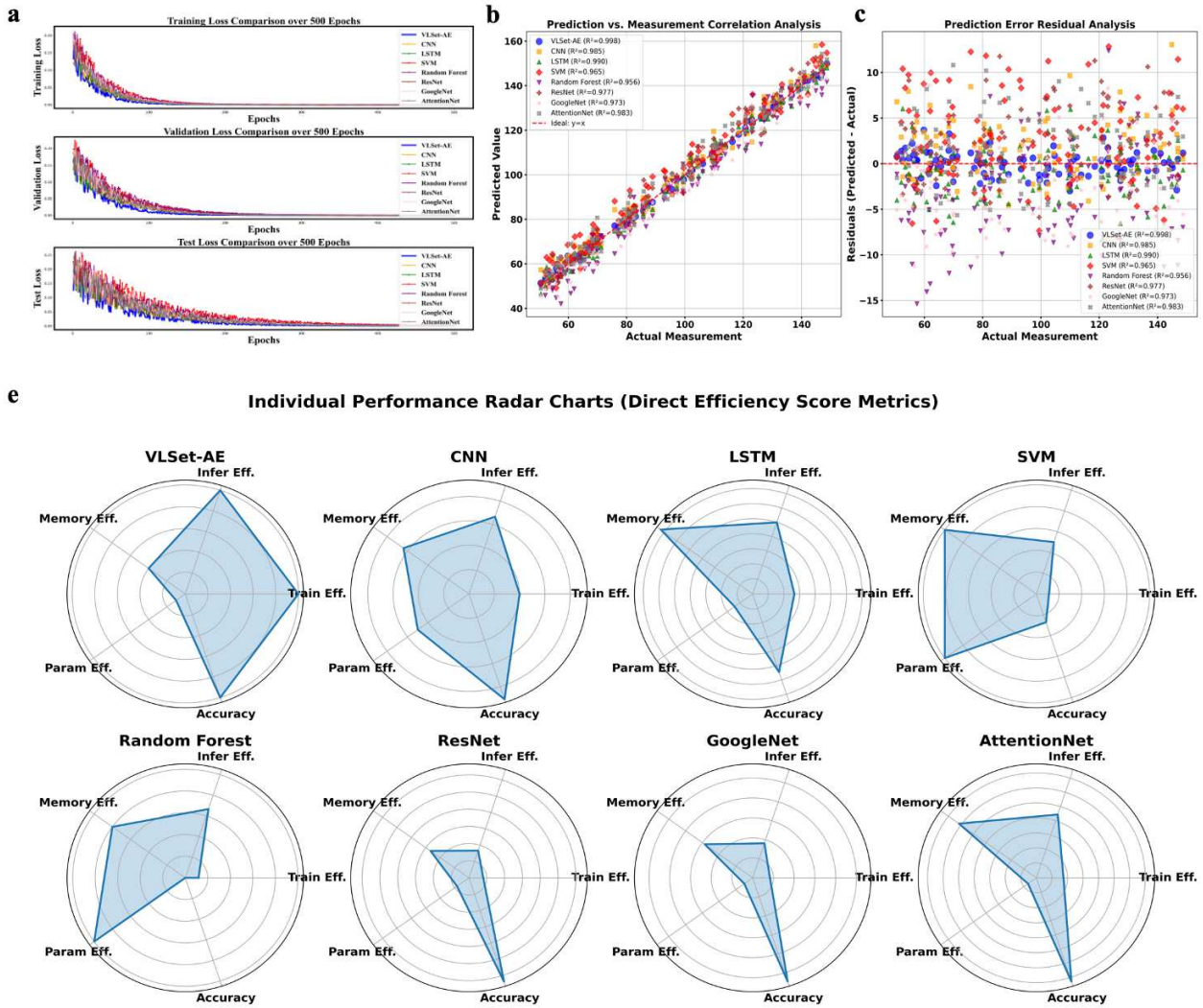



**Fig. 7 Performance Testing and Evaluation of Different Algorithm Models. a** Training, validation, and test loss comparison among multiple models. **b** Correlation analysis of automatic calculation vs. actual measurements. **c** Residual analysis of automatic calculation vs. actual measurements. **d** Individual performance radar charts of direct efficiency scores (training time, inference time, memory usage, parameter count, accuracy).

As illustrated in **Fig. 7d**, the proposed VLSet-AE model consistently exhibits superior performance across the majority of evaluated dimensions, including training efficiency, inference efficiency, parameter efficiency, and predictive accuracy. These results suggest that VLSet-AE not only accelerates the overall training and inference processes but also attains high predictive fidelity, rendering it highly suitable for scenarios where rapid model deployment and accurate predictions are paramount. It is noteworthy that while VLSet-AE does not achieve the single best memory-efficiency score among all evaluated models, its memory footprint remains within a competitive range. The slight increase in memory usage is arguably offset by the model's substantial gains in other critical metrics, implying that the overall trade-off is decisively favorable. In particular, the synergy between low parameter counts, rapid convergence, and high accuracy underscores the model's practical applicability in resource-limited or latency-sensitive environments.

Consequently, despite this minor limitation in memory consumption, VLSet-AE demonstrates a balanced and robust profile, outperforming conventional approaches such as CNN, LSTM, or more complex architectures like ResNet and AttentionNet in most respects.

## CONCLUSION

In conclusion, we propose a physics-constrained variational level set autoencoder (VLSet-AE) that revolutionizes automated contour recognition and feature extraction in scanning electron microscopy (SEM) cross-sectional profiles of deep reactive ion etching (DRIE). By integrating layer-wise scallop segmentation along the etching depth and embedding physical etching constraints within the decoder, VLSet-AE ensures that reconstructed level set functions faithfully represent the etching dynamics. A novel three-dimensional framework, incorporating time, linewidth, and etching depth, captures the spatiotemporal evolution of etched profiles with unprecedented precision. Assembled from reconstructed scallop segments, complete profiles enable accurate quantification of nine critical dimensions: scallop depth (2.29% error), scallop width (peak-to-peak: 2.05%; valley-to-valley: 6.28%), scallop radius (4.69%), profile angle (0.56%), trench depth (5.46%), bow width (4.35%), mid width (2.43%), and bottom width (4.78%). Rigorous experimental validation, encompassing error distributions, loss trajectories, correlation analyses, residual assessments, and confusion matrices (Fig. 4a–i), confirms VLSet-AE's exceptional performance. Achieving an average feature recognition error of 3.65%, an overall accuracy of 96%, and a correlation coefficient of 0.998 with ground truth measurements, VLSet-AE markedly surpasses manual annotation (15–20% error) and conventional models, such as CNNs and SVMs (70–80% accuracy in noisy conditions). Benchmarking against seven state-of-the-art models (including CNNs, LSTMs, and ResNet), VLSet-AE exhibits the shortest training duration (20 s), fastest inference time (1.2 s per image), minimal memory footprint (50 MB), and a competitive parameter count (4.0 million). These attributes underscore its computational efficiency and robustness, facilitating real-time process monitoring, advanced three-dimensional morphology simulations, and scalable data acquisition for AI-driven DRIE optimization. By overcoming the limitations of traditional SEM analysis, VLSet-AE establishes a transformative paradigm for intelligent, high-precision microfabrication, heralding a new era of AI-assisted manufacturing for next-generation microelectromechanical systems.

## ACKNOWLEDGEMENTS

This work is supported by the National Key R&D Plan Project (2023YFB3207900) and Guangdong Provincial Key Field R&D Plan Project (2020B1111120004).

## CONFLICT OF INTERESTS

The authors declare no competing interests.

## REFERENCES


1. Laermer F, Franssila S, Sainiemi L, et al. Deep reactive ion etching[M]//Handbook of silicon based MEMS materials and technologies. Elsevier, 2020: 417-446.
2. Clerc P A, Dellmann L, Gretillat F, et al. Advanced deep reactive ion etching: a versatile tool for microelectromechanical systems[J]. Journal of Micromechanics and Microengineering, 1998, 8(4): 272.
3. Yeom J, Wu Y, Selby J C, et al. Maximum achievable aspect ratio in deep reactive ion etching of silicon due to aspect ratio dependent transport and the microloading effect[J]. Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena, 2005, 23(6): 2319-2329.
4. Abhulimen I U, Polamreddy S, Burkett S, et al. Effect of process parameters on via formation in Si using deep reactive ion etching[J]. Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena, 2007, 25(6): 1762-1770.
5. Hajare R, Reddy V, Srikanth R. MEMS based sensors–A comprehensive review of commonly used fabrication techniques[J]. Materials Today: Proceedings, 2022, 49: 720-730.
6. Xu T, Tao Z, Li H, et al. Effects of deep reactive ion etching parameters on etching rate and surface morphology in extremely deep silicon etch process with high aspect ratio[J]. Advances in Mechanical Engineering, 2017, 9(12): 1687814017738152.
7. Li Y, Zhang H, Yang R, et al. Fabrication of sharp silicon hollow microneedles by deep-reactive ion etching towards minimally invasive diagnostics[J]. Microsystems & nanoengineering, 2019, 5(1): 41.
8. Morse E. Deep Reactive-Ion Etching Process Development and Mask Selection[D], 2020.
9. Postek M T, Vladár A E. Does your SEM really tell the truth? —How would you know? Part 1[J]. Scanning: The Journal of Scanning Microscopies, 2013, 35(6): 355-361.
10. Canny J. A computational approach to edge detection[J]. IEEE Transactions on pattern analysis and machine intelligence, 1986 (6): 679-698.
11. Gesho M, Chaisoontornyotin W, Elkhatib O, et al. Auto-segmentation technique for SEM images using machine learning: Asphaltene deposition case study[J]. Ultramicroscopy, 2020, 217: 113074.
12. Yao L, Chen Q. Machine learning in nanomaterial electron microscopy data analysis[M]//Intelligent Nanotechnology. Elsevier, 2023: 279-305.
13. Maraghechi S, Hoefnagels J P M, Peerlings R H J, et al. Correction of scanning electron microscope imaging artifacts in a novel digital image correlation framework[J]. Experimental mechanics, 2019, 59: 489-516.
14. Sun W, Duan F, Zhu J, et al. An Edge Detection Algorithm for SEM Images of Multilayer Thin Films[J]. Coatings, 2024, 14(3): 313.
15. Rani G E, Murugeswari R, Rajini N. Edge detection in scanning electron microscope (SEM) images using various algorithms[C]//2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2020: 401-405.
16. Selvakumar P, Hariganesh S. The performance analysis of edge detection algorithms for image processing[C]//2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16). IEEE, 2016: 1-5.
17. Kim H, Han J, Han T Y J. Machine vision-driven automatic recognition of particle size and morphology in SEM images[J]. Nanoscale, 2020, 12(37): 19461-19469.
18. Qiao Y, Mei Z, Luo Y, et al. DeepSEM-Net: Enhancing SEM defect analysis in semiconductor manufacturing with a dual-branch CNN-Transformer architecture[J]. Computers & Industrial Engineering, 2024, 193: 110301.
19. Giannatou E, Papavieros G, Constantoudis V, et al. Deep learning denoising of SEM images towards noise-reduced LER measurements[J]. Microelectronic Engineering, 2019, 216: 111051.
20. LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436-444.
21. Modarres M H, Aversa R, Cozzini S, et al. Neural network for nanoscience scanning electron microscope image recognition[J]. Scientific reports, 2017, 7(1): 13282.
22. Iwata H, Hayashi Y, Hasegawa A, et al. Classification of scanning electron microscope images of pharmaceutical excipients using deep convolutional neural networks with transfer learning[J]. International Journal of Pharmaceutics: X, 2022, 4: 100135.
23. Li C, Su J, Yu L, et al. A variational level set method image segmentation model with application to intensity inhomogene magnetic resonance imaging[J]. Digital Medicine, 2018, 4(1): 5-15.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalMaterial.zip