

Article

Research on Reverse Decoupling and Optimization of White-Light Interference Signals Based on Deep Learning

Yanzhong Ma ^{1,†}, Chi Chen ^{1,†}, Lu Chen ^{1,*}, Ji Zhang ¹ , Xiaojun Tian ¹, Guangrui Wen ² and Zihao Lei ^{2,*} 

¹ Skyverse Technology Co., Ltd., Shenzhen 518110, China; mayanzhong@skyverse.cn (Y.M.); chenchi@skyverse.cn (C.C.); zhangji1@skyverse.cn (J.Z.); tianxiaojun@skyverse.cn (X.T.)

² School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China; grwen@xjtu.edu.cn

* Correspondence: chenlu@skyverse.cn (L.C.); leizihao@xjtu.edu.cn (Z.L.)

† These authors contributed equally to this work.

Abstract

In the intelligent operation and maintenance of core process equipment in semiconductor manufacturing, the thickness and morphological parameters of sub-micron multilayer transparent films on wafer surfaces are critical quality indicators that govern device performance and yield. White Light Interferometry (WLI) enables the high-precision measurement of thin-film thickness and topography parameters, and is widely deployed in high-precision manufacturing fields such as semiconductors. However, when measuring sub-micron multilayer transparent films, WLI faces challenges including low computational efficiency, severe parameter coupling, and non-unique solutions, making it difficult to meet the demands of high-throughput online inspection. To address these issues, this paper proposes a deep learning-based method for decoupling and optimizing WLI signals through inverse modeling. The approach establishes an innovative hybrid intelligent framework: first, a training dataset is generated based on interferometric system modeling and simulation; then, a classification model is employed to intelligently categorize the acquired interference signals, decomposing the complex multimodal inversion problem into several sub-problems with simpler patterns. Next, for each signal category, a specialized closed-loop deep network model is designed and trained. This network integrates an inverse prediction network with a forward reconstruction network in series. During training, both prediction error and reconstruction error are jointly used as the loss function, ensuring solution uniqueness and physical consistency, thereby enhancing inversion reliability and robustness. This research provides an effective solution for high-precision online optical measurement of complex thin-film structures, offering significant theoretical value and broad industrial application prospects.

Keywords: intelligent fault perception; white light interferometry; deep learning; inverse problems; cascaded networks; thin film measurement



Academic Editor: Ignacio Bosch Roig

Received: 7 August 2026

Revised: 20 September 2026

Accepted: 28 September 2026

Published: 30 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The long-term operational reliability of critical equipment such as deposition and etching tools directly determines production efficiency and product yield in semiconductor manufacturing. Issues including chamber temperature drift, airflow disturbances, and precision degradation of mechanical components often manifest themselves sensitively through structural anomalies of multilayer transparent films on wafer surfaces—such as thickness deviations or morphological distortions. Consequently, online high-precision

perception of film parameters is not only a prerequisite for process quality control, but also a critical enabler for early degradation awareness and intelligent operation and maintenance of core process equipment.

White light interferometry (WLI) is a high-precision, non-contact 3D surface topography measurement technique [1]. It offers nanoscale vertical resolution and rapid full-field imaging, making it widely used for quality control and process monitoring in advanced fields such as semiconductor manufacturing [2], micro-electromechanical systems (MEMS) [3], and precision optical components [4]. However, as semiconductor processes advance to smaller nodes and device structures become more complex, achieving rapid, large-area, and high-precision online measurement of multilayer transparent thin films with submicron thickness presents a critical technical bottleneck. When measuring such samples using traditional WLI, the interference signals become extremely complex due to the superposition of reflections from the film's top surface, bottom interface, and internal layers. Conventional peak extraction or phase analysis algorithms struggle to accurately separate surface topography from film structural information. Furthermore, the diversity of film structures across wafer surfaces makes a single model inadequate for area measurements, severely limiting WLI's application in advanced semiconductor process control.

To extract film parameters from complex interference signals, various methods have been developed, yet each has inherent limitations. Template matching (Library Matching) is a common approach [5]. It involves pre-building a large library of simulated signals and comparing the measured signal against each template to find the best match. However, its computational and storage costs grow exponentially with parameter dimensions and precision, hindering real-time application. More critically, the discrete nature of the template library caps the achievable measurement accuracy. Severe parameter coupling is another issue, as parameters like layer thickness and refractive index jointly determine the optical path difference (OPD) and influence fringe features. This many-to-one mapping makes parameter inversion difficult. Another category is iterative fitting based on physical models [6], which optimizes parameters by minimizing residuals between measured and model-predicted signals [7]. While physically interpretable, this method is computationally intensive, prone to local minima, and unreliable for multilayer structures. These traditional methods share a fundamental challenge inherent to the optical inverse problem: due to the periodic nature of spectral responses, entirely different parameter combinations can produce nearly indistinguishable interference signals within noise tolerance. This one-to-many mapping often causes traditional inversion algorithms to fail or yield ambiguous solutions.

Recently, deep learning, particularly Deep Neural Networks (DNNs), has emerged as a new paradigm for solving complex optical inverse problems, leveraging its strong nonlinear mapping and automatic feature extraction from vast datasets. However, directly training an inverse network to map spectra to structural parameters faces convergence issues due to the aforementioned one-to-many problem [8]. To overcome this core challenge, this paper proposes a deep learning-based method for decoupling and optimizing WLI signals via inverse modeling. We designed an innovative hybrid intelligent framework following a "classify first, then specialize" strategy. First, a training dataset is generated via interferometric system modeling and simulation. A classification model then intelligently categorizes acquired interference signals, decomposing the complex multimodal inversion problem into several simpler, well-defined sub-problems. Subsequently, for each signal category, a specialized closed-loop deep network model is designed and trained. This model cascades an inverse prediction network with a forward reconstruction network. During training, a combined loss function incorporating both parameter prediction error and signal reconstruction error is used. This approach circumvents the one-to-many

training dilemma, ensuring solution uniqueness and physical consistency while enhancing inversion robustness.

This study aims to fundamentally address the triple challenge of computational efficiency, parameter coupling, and solution non-uniqueness faced by traditional techniques when processing multilayer film interference signals. It provides a viable solution for achieving high-precision online measurement with a surface step height accuracy of 3 nm.

2. Fundamental Theories and Methods

2.1. Physical Model of White Light Interferometry

White light interferometry operates on the principle of low-coherence interference. As shown in Figure 1a, light from a broadband source (e.g., xenon lamp or white LED) is directed into the microscope as a coaxial illuminator by a beam splitter. The beam then travels through an aperture stop and enters the interferometric objective. Inside this objective, a second beam splitter divides the light. One beam is directed toward a reference mirror (reference arm), while the other is focused onto the surface of the test sample (measurement arm). The two reflected beams recombine at the detector, generating interference. By vertically scanning the sample along the Z-axis using a precise piezoelectric transducer (PZT) or translation stage, the optical path difference between the two arms is varied. The detector records a series of intensity values $I(h)$ corresponding to different scan positions h , forming the white-light interference signal for that point (also called the interferogram, as shown in Figure 1b).

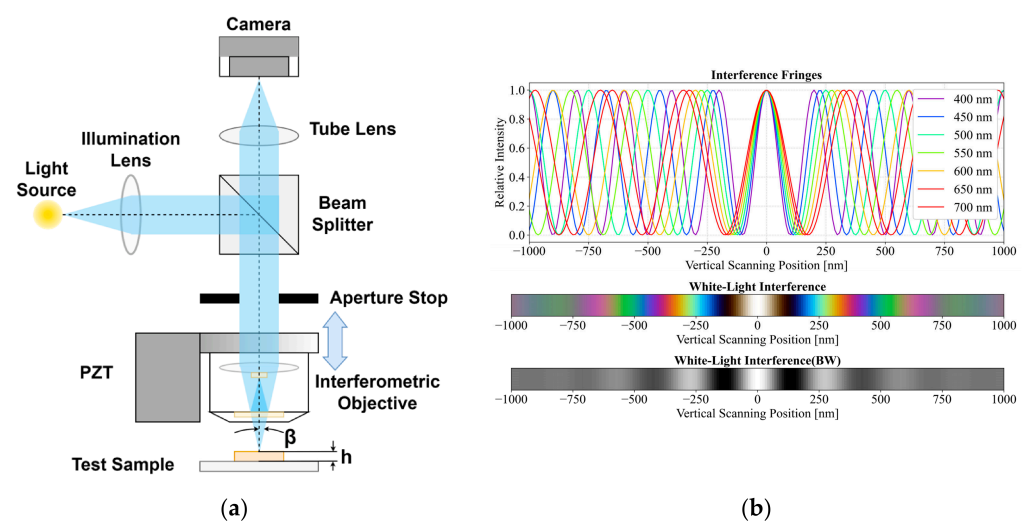


Figure 1. Principle and Signal Model of white light interferometry. (a) Schematic Diagram of a White Light Interferometer System; (b) Synthesis of Broad-Spectrum Interference Intensity.

In white light interferometry, the interference signal $I(h)$ can be described by the following integral model, taking into account factors such as the source spectral distribution $S(k)$ and the numerical aperture effect $U(\beta)$ of the optical system [9]:

$$I(h) = \int \left[I_{ref} + I_{obj} + 2\sqrt{I_{ref}I_{obj}}\cos(2\beta hk + \theta(k)) \right] U(\beta)S(k)\beta d\beta dk \quad (1)$$

In Equation (1), k denotes the wavenumber ($k = 2\pi/\lambda$), β represents the incidence angle, I_{ref} and I_{obj} are the intensities of the reference and object beams, respectively, and $\theta(k)$ is the phase introduced by sample reflection. Within this model, the phase $\theta(k)$ serves as the sole link connecting the measurable optical signal to the sample's underlying physical microstructure—specifically, the thickness d , refractive index n , and extinction coefficient k of each film layer. Consequently, the ultimate objective of all inversion algorithms is to

accurately retrieve these physical structural parameters from the measured interference signal $I(h)$.

2.2. Transfer Matrix Method for Multilayer Film Reflection Spectrum (TMM)

To establish an accurate forward model that maps physical structural parameters to spectral reflectivity $r(k)$, we employ the Transfer Matrix Method (TMM). As a classical and precise electromagnetic theory based on Maxwell's equations, TMM is used for calculating the optical response of planar-parallel multilayer film structures.

As illustrated in Figure 2, this method models an n -layer thin-film system as a sequence of interfaces and layers. The system consists of the incident medium L_0 , a series of Film Layers $L_1, L_2, \dots, L_n$, and the exit substrate L_{n+1} . The thickness of layer L_0 is defined as the air gap between the measurement surface and the reference plane. The propagation, reflection, and transmission behavior of light waves within this system can be described using two types of matrices: the boundary matrix P and the propagation matrix Q .

$$P_n = \begin{bmatrix} 1 & 1 \\ k_n & -k_n \end{bmatrix} \quad (2)$$

$$Q_n = \begin{bmatrix} e^{ik_n d_n} & 0 \\ 0 & e^{-ik_n d_n} \end{bmatrix} \quad (3)$$

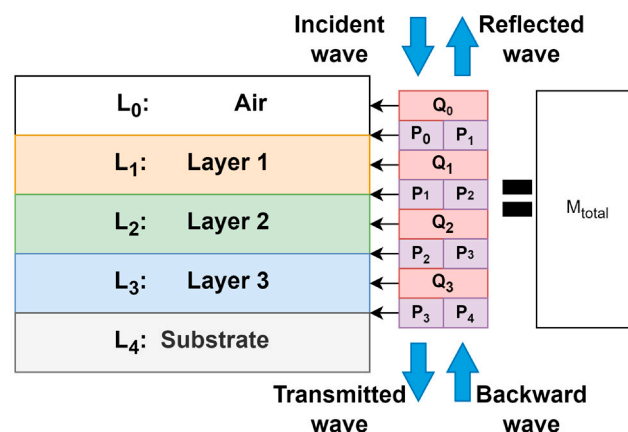


Figure 2. Schematic Diagram of the TMM for Calculating Multilayer Film Reflectivity.

The relationship between the fields in adjacent layers is given by:

$$P_n Q_n \begin{bmatrix} a_n \\ b_n \end{bmatrix} = P_{n+1} \begin{bmatrix} a_{n+1} \\ b_{n+1} \end{bmatrix} \quad (4)$$

For a system composed of N thin-film layers, the overall transfer matrix M can be obtained by multiplying the interface matrices and propagation matrices of each individual layer.

$$\begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = M \begin{bmatrix} a_{n+1} \\ b_{n+1} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} a_{n+1} \\ b_{n+1} \end{bmatrix} \quad (5)$$

$$M = Q_0^{-1} P_0^{-1} \left(\prod_{j=1}^N P_j Q_j^{-1} P_j^{-1} \right) P_{N+1} \quad (6)$$

From the elements of this overall transfer matrix M , Based on the boundary condition that there is no backward-propagating wave in the exit substrate $b_{n+1} = 0$, the field amplitudes at the incident interface are related to the transmitted amplitude $a_0 = M_{11} a_{n+1}$ and

$b_0 = M_{21}a_{n+1}$. Consequently, the complex reflection coefficient $r(k)$ of the entire multilayer film system can be precisely calculated.

$$r(k) = \frac{b_0}{a_0} = \frac{M_{21}}{M_{11}} \quad (7)$$

The thickness of each film layer is ultimately determined using the TMM. Figure 3 presents the simulated interference results for multilayer films with different thicknesses.

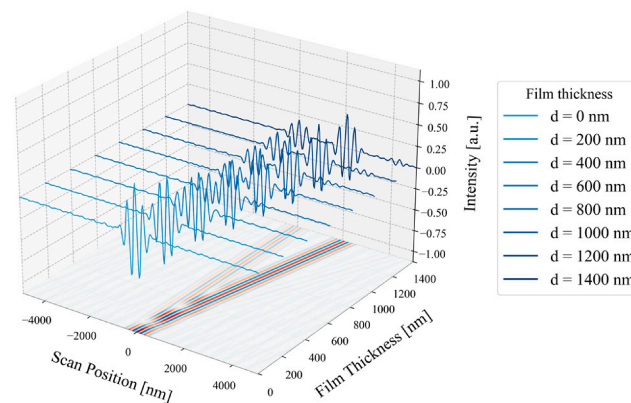


Figure 3. Interference Schematic of Multilayer Films with Different Thicknesses.

The accuracy and computational efficiency of TMM make it an ideal tool for generating high-quality simulation data. By performing large-scale sampling within a given parameter space and utilizing TMM, vast amounts of “parameter-spectrum” data pairs can be calculated. This forms an indispensable foundation for training and validating subsequent deep learning models. The accuracy of the forward model directly determines the performance upper limit of the inverse solving algorithm.

2.3. Deep Learning-Based Inverse Decoupling and Optimization Method

Inferring structural parameters from measurement signals is an inverse problem that poses significant challenges due to its ill-posed nature. Measurement noise and interferometric system aberrations further blur the distinctions between similar signals. Owing to the periodic nature of spectral responses, different combinations of thin-film structural parameters can produce spectra that are nearly indistinguishable within the measurement noise tolerance, leading to a “one-to-many” mapping ambiguity. To address this fundamental difficulty, we have designed a hybrid intelligent framework based on a “classify first, then specialize” strategy.

2.3.1. Deep Learning-Based Intelligent Signal Classification

Intelligent signal classification forms the foundational step of this framework. Its primary objective is to automatically decompose complex white light interferometry signal sets, which contain multiple structural types, into several subsets with single-mode patterns. This provides efficient input for subsequent parameter inversion. This process is achieved through network training on simulated datasets, ensuring both accuracy and computational efficiency in classification.

In the spectral domain, the white light interference signals $F(k)$ can be expressed as:

$$F(k) = C(k)S(k)r(k)e^{-i2kh+\theta(k)} \quad (8)$$

The complex reflection coefficient of the sample $r(k)e^{i\theta(k)}$ carries the structural information, such as film thickness and refractive index encoded in the interference fringes.

The e^{-i2kh} links the topography height h of the measurement point to the spectral phase modulation. $S(k)$ represents the source spectrum, and $C(k)$ denotes the complex chromatic dispersion function determined by the hardware configuration, such as light source, objective lens and camera. These define the envelope and the phase gap of the signal.

In practical measurements, the system-specific terms $S(k)$ and $C(k)$ act as a transfer function that must be calibrated out to isolate the sample information. A standard sample (e.g., a silicon wafer with known properties) is used to calibrate the system function. The measured interference spectrum of the standard silicon wafer is given by

$$F_{Si}(k) = C(k)S(k)r_{Si}(k)e^{-i2kh_{Si}+\theta_{Si}(k)} \quad (9)$$

Consequently, the calibrated system function $H(k)$ is obtained as

$$H(k) = C(k)S(k) = \frac{F_{Si}(k)}{r_{Si}(k)e^{-i2kh_{Si}+\theta_{Si}(k)}} \quad (10)$$

This calibration process effectively converts raw optical intensity data into a standardized digital representation, removing hardware dependencies and allowing the neural network to focus purely on learning the structural features of the interference patterns. As illustrated in Figure 4, in the raw detected signal, dispersion imbalance between the reference and measurement arms induces a phase error, causing a misalignment between the interference phase and envelope signals (e.g., the Si substrate envelope peak no longer coincides with the phase peak in Figure 4b). After correction, both the envelope and phase peak profiles are restored, enabling subsequent matching with model-simulated signals.

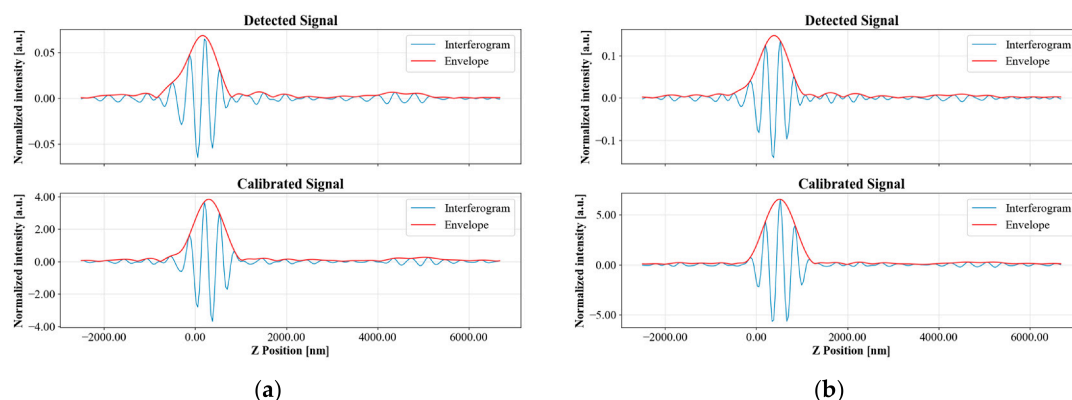


Figure 4. Comparison of detected and calibrated signals for different samples: (a) SiO₂ film; (b) Si substrate.

Based on the structure of the sample under test, simulation datasets D1 and D2 are generated, corresponding to labels L1 and L2 respectively. Each sample in the datasets is constructed by discretizing the complex spectrum $F(k)$ into a two-channel input tensor of length, where the channels correspond to magnitude and phase part. A classification network, whose architecture is shown in Figure 5, is established. The network takes a multi-channel input tensor of size $2 \times L_{in}$. The network employs a multi-scale convolutional design: three parallel branches with kernel sizes of 3, 5, and 7 are used to capture high-frequency details, mid-frequency oscillations, and low-frequency trends in the spectral signal, respectively. This multi-scale structure enables the network to effectively recognize interference patterns across different spectral scales. Each branch outputs 8 feature channels, which are concatenated into a 24-channel feature map. These features are subsequently fused and refined through two 1D convolutional layers (reducing to 16 channels) with BatchNorm, ReLU activation, and a dropout rate of 0.3. Finally, the refined features are

aggregated via global pooling and passed through a classification head with a hidden layer, producing the final output logits. This multi-scale structure enables the network to effectively recognize interference patterns across different spectral scales.

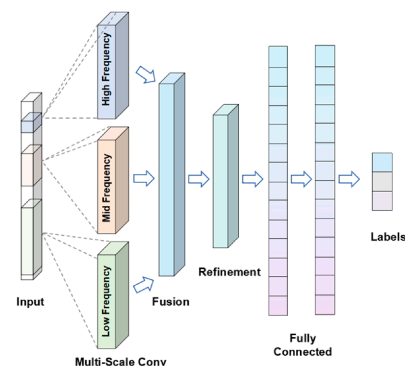


Figure 5. Classification Network Model.

By using a predefined dataset, the structural parameters within the independent dataset vary within a defined range. This approach avoids the issue of multiple solutions introduced by different thin-film microstructures.

2.3.2. Parameter Inversion via Jointly Trained Cascaded DNNs

Traditional parameter regression for inverse solving relies on unidirectional neural networks. Directly regressing structural parameters (such as layer thickness and refractive index) from interference spectra is a highly nonlinear and ill-posed inverse problem. Although traditional end-to-end regression models can fit training data, they often lack physical consistency, leading to poor generalization on out-of-distribution samples [10]. To incorporate physical priors and enhance model robustness, this paper proposes a bidirectional cascaded training framework, as illustrated in Figure 6. The architecture consists of two modules: the Inverse Network and the Forward Network.

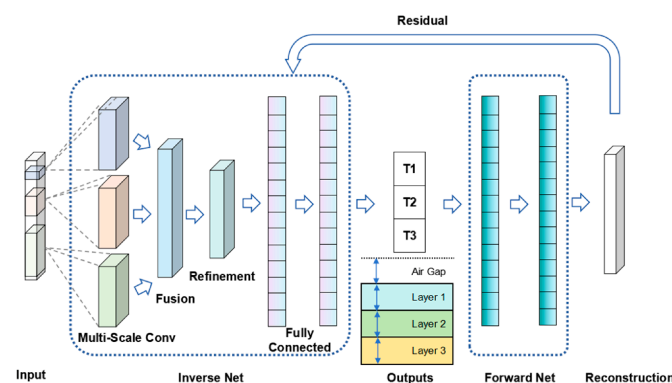


Figure 6. Schematic of Bidirectional Cascaded Network Architecture.

The inverse network takes the spectral input and adopts a multi-branch dilated convolutional architecture to capture spectral features at multiple scales. The network takes the spectral input $x = [A, \theta] \in \mathbb{R}^{2 \times L}$ (representing the magnitude and phase components of the complex spectrum). Three parallel branches with a fixed kernel size of 3 and dilation rates of 2, 4, 6 (yielding equivalent receptive fields of 5, 9 and 13) are utilized to extract high-frequency, mid-frequency, and low-frequency features, respectively. Each branch produces 16 channels, which are concatenated into a 48-channel feature map. This feature map is then refined by two 1D convolutional layers with BatchNorm, ReLU activation and

a dropout rate of 0.3. After flattening, a two-layer fully connected regressor generates the estimated structural parameter vector $\hat{p} = f(x) \in R^P$, thereby performing signal inversion.

The forward network takes the parameters $\hat{p}$ as input and outputs the corresponding complex amplitude components $[A, \theta] = g(p) \in R^{2 \times L}$. This network implicitly learns the mapping from the parameter space to the phase space, serving as an approximation of the true physical forward model based on the Transfer Matrix Method (TMM).

The joint optimization objective for the cascaded network is defined as:

$$\mathcal{L}_{total} = \|\hat{p} - p\|^2 + \lambda \|\hat{\theta} - \theta\|^2 \quad (11)$$

Here, $\lambda > 0$ controls the strength of the physical constraint. This loss function updates both $f(x)$ and $g(p)$ simultaneously through backpropagation, ensuring that the inversion results not only fit the labels but also satisfy reconstructability—i.e., the predicted parameters must be able to generate phase signals consistent with the measurements via the forward model. Under this training mechanism, the inverse network is compelled to learn a mapping whose output parameters, when reconstructed through the physical model (forward network), must match the original spectrum. This eliminates ambiguous solutions during training and ensures the network converges to a physically consistent, unique solution. This architecture significantly enhances the stability and accuracy of the inverse solution and complements the aforementioned classification model, thereby strengthening the overall robustness of the framework.

3. Experimental Verification

To systematically validate the effectiveness of the proposed method and comprehensively evaluate its performance, we selected an etched silicon oxide on silicon wafer as the test sample. The measured area of this sample contains two material surfaces (SiO_2 and Si) along with transparent SiO_2 film layers ranging from 100 nm to 1000 nm in thickness. Based on this sample structure, we constructed forward physical models for interference reflection corresponding to the SiO_2 surface and the Si surface, using the Transfer Matrix Method and interferometric system simulation, respectively. By incorporating actual system parameters such as the light source spectrum and incidence angles, we generated over 200,000 sets of “structural parameter—reflection spectrum” data pairs through extensive random sampling. This dataset was used for subsequent model training. After splitting it into training and test sets in an 8:2 ratio, we trained the DNN described in Section 2.3, ultimately obtaining a deep learning model capable of material classification and film thickness inversion.

To improve the robustness of the model against measurements noise, gaussian white noise was added to the input interference spectra during training. The signal-to-noise ratio (SNR) is defined in Equation (12) as

$$SNR = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right) \quad (12)$$

where P_{signal} and P_{noise} denote the average power of the clean spectral signal and the added noise, respectively. The root-mean-square error (RMSE) between the predicted and ground-true values is used to quantitatively evaluate the inversion accuracy. In the noiseless case, the predicted residual is below 1 nm, as shown in Figure 7a. Under a test condition of $SNR = 30$ dB, the proposed method achieves a height or thickness regression RMSE of less than 2 nm, as shown in Figure 7b. Figure 7c shows the bias between predicted and true thickness values over a wide thickness range from 50 nm to 1300 nm. The errors remain

essentially zero throughout the entire range, confirming the high accuracy and negligible systematic bias of the proposed method.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{p}_i - p_i)^2} \quad (13)$$

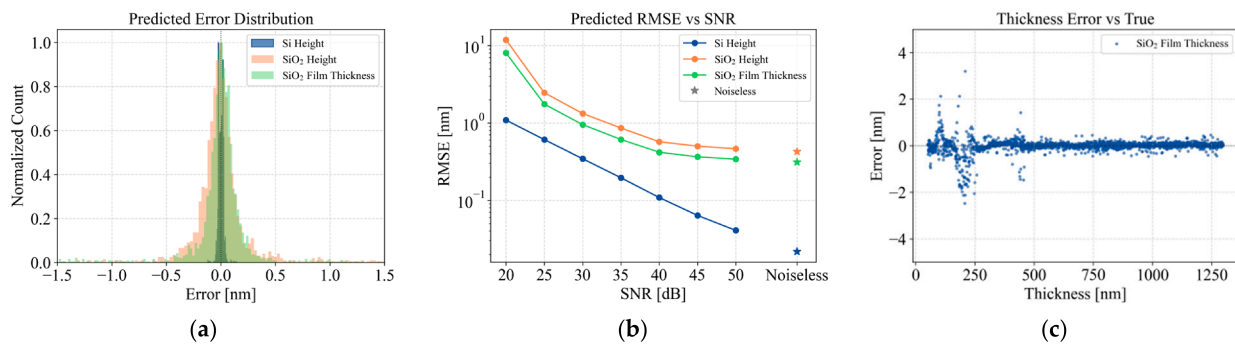


Figure 7. Analysis of the predicted error. (a) Distribution histograms of the predicted errors for height and thickness. (b) Predicted *RMSE* as a function of *SNR*. (c) Predicted error of the true film thickness.

In the experimental measurement, we used a Skyverse CYPRESSIFM-U960 3D surface profiling metrology equipment (Skyverse Technology Co., Ltd., Shenzhen, China) to perform vertical scanning on the sample to detect its interference signals, which were then converted into spectral signals. To validate the metrological performance, a NIST-traceable film standard featuring nominal with SiO₂ thicknesses of 100 nm and 1000 nm was utilized. Figure 8 shows the white light interferometry image of the test sample and its corresponding scanned interference signal. Figure 8a displays the surface topography image of the sample, clearly marking the silicon substrate region (Si) and the silicon oxide film regions. Figure 8b presents the vertical scanning interference signal acquired along the red dashed line position indicated above. Its horizontal axis strictly corresponds to the scan position in Figure 8a, while the vertical axis represents the scanning height. The colored fringes in the figure reflect the periodic variation in light intensity with the optical path difference at each pixel point. Different material and film thickness regions exhibit distinctly characteristic interference fringe patterns in the interferogram.

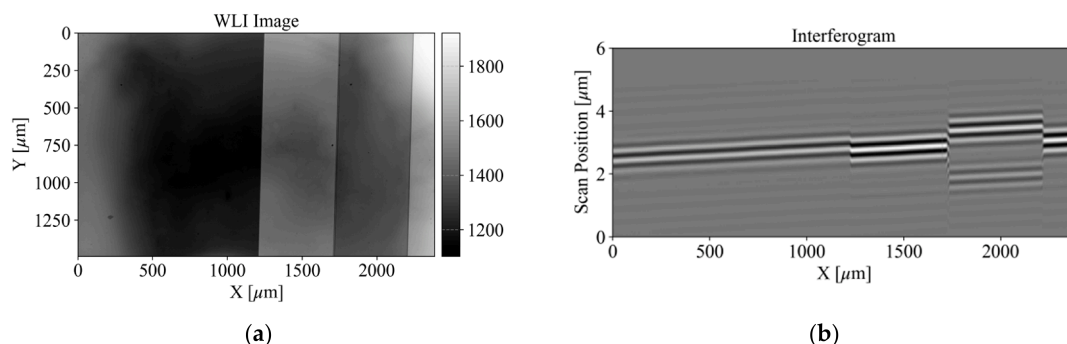


Figure 8. Test Sample. (a) Test Sample Image; (b) Scanned Interference Signal of the Test Sample.

The spectral signals are input into the classification model to predict the class label for each pixel within the region. The corresponding optical models are then applied to different film layer categories. Figure 9 displays the classification label map predicted by the trained DNN classification model. Figure 9a presents the output of the trained DNN model classifying the spectral signal at each pixel in the measured area. Two distinct regions are clearly delineated: gray corresponds to the “SiO₂/Si” multilayer structure area, and

black corresponds to the exposed “Si” substrate area. The spatial boundary between the two regions is sharp and aligns with the designed sample structure, effectively validating the model’s capability for pixel-level automatic classification of heterogeneous thin-film materials. This provides the mask basis for subsequent parameter inversion by category. Based on the classification result from Figure 7a, Figure 9b applies the respective optical models to the interference signals from different category regions for thickness inversion, thereby reconstructing the three-dimensional topography of the sample surface. A higher SiO₂ step platform and a lower Si substrate platform are visible, separated by a distinct height discontinuity boundary. This result quantitatively presents the sample’s topographic features in three-dimensional space, fully demonstrating the proposed “classify first, then invert” method’s capability for automatic identification and nanoscale 3D measurement of complex heterogeneous structures.

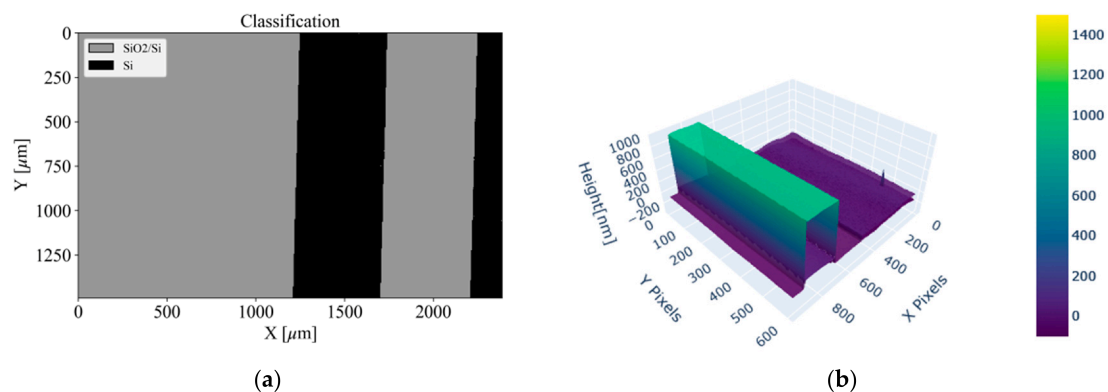


Figure 9. Material Classification and 3D Topography Reconstruction Results of the Sample. (a) Classification Label Map; (b) 3D Topography Map.

Figure 10 presents the inversion verification results for three typical structures: a 100 nm SiO₂ layer, a 1000 nm SiO₂ layer, and a bare Si surface. In each subplot, the red dashed line represents the theoretical interference signal (target curve) calculated from the forward physical model, while the blue solid line denotes the curve fitted from the measured signal using the inversion method proposed in this study. The horizontal axis indicates the scan position, and the vertical axis shows the normalized intensity.

The comparison reveals that, under varying film thicknesses (100 nm and 1000 nm SiO₂) and different materials (SiO₂ and Si), the fitted curves (blue lines) exhibit high agreement with the target curves (red lines) in terms of the period, phase, and amplitude of the interference oscillations. This result quantitatively validates that the described inversion model possesses accurate signal reconstruction capability and good generalizability across different materials and film thicknesses. It provides reliable methodological support for the precise extraction of nanoscale thin-film parameters based on white-light interference spectra.

The reference site of this standard was certified at 95.2 ± 1.4 nm (expanded uncertainty, 95% confidence level), the mean thickness obtained from measurements was 96.3 nm, which falls within the certified uncertainty interval. The step height of this 100 nm and 1000 nm region was independently measured using a commercial stylus profiler with an uncertainty specification of $0.5 \text{ nm} + 0.4\%H$, yielding a value of 964.8 nm. In comparison, the height determined by our proposed interferometric method was 967.7 nm, which falls well within the expanded uncertainty of stylus profiler. Moreover, the repeatability, expressed as the standard deviation ($\sigma = 0.7$ nm) from repeated dynamic scanning measurements.

This chapter systematically validates the effectiveness of the “classify first, then invert” method by generating a large-scale “structure-spectrum” training dataset through a

transfer matrix forward model and utilizing experimental data from scanning white light interferometry. The process achieves high-precision classification and three-dimensional topography reconstruction of heterogeneous samples containing SiO₂ layers of varying thicknesses on Si substrates. The core inversion algorithm demonstrates excellent accuracy in signal comparisons of typical structures, providing a reliable technical pathway for the quantitative measurement of transparent thin-film structures.

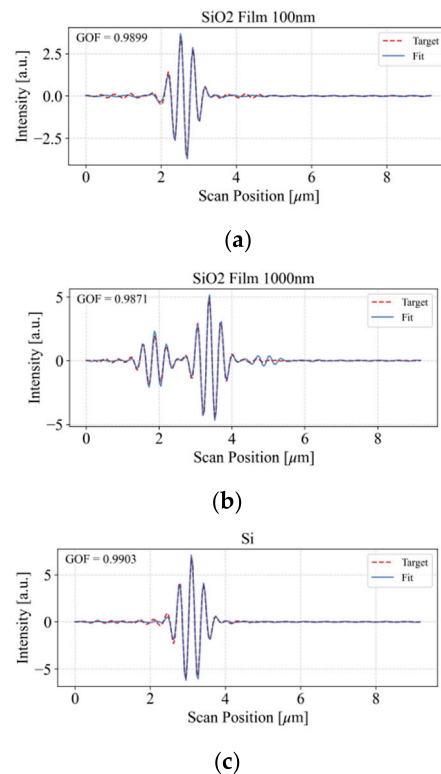


Figure 10. Signal Inversion Verification. Schemes follow another format. (a) Typical Structure—100 nm SiO₂; (b) Typical Structure—1000 nm SiO₂; (c) Typical Structure—Pure Si Surface.

4. Discussion and Outlook

The deep learning-based framework for inverse decoupling and optimization of white light interferometry signals, proposed in this study, has yielded promising results in simulations, offering a powerful tool to address long-standing challenges in semiconductor thin-film measurement. By combining unsupervised sparse dictionary classification with supervised cascaded network solving, this method not only demonstrates significant potential to surpass traditional algorithms in measurement accuracy and speed, but more importantly, provides an effective paradigm for tackling the pervasive and challenging non-unique inverse problem in optical metrology. Its core advantage lies in its “data-driven” nature, which enables automatic learning and decoupling of highly nonlinear, interdependent physical parameters from massive datasets. This eliminates the need for complex and time-consuming manual feature engineering and model design, greatly enhancing the algorithm’s robustness and automation level. For industrial applications requiring high-throughput, high-precision online inspection, such as semiconductor manufacturing, this technology signifies a substantial advancement. It compresses the time for full-field, array-based measurement tasks from several minutes or longer to the order of seconds. This makes real-time process control for each full wafer possible, holding immeasurable value for improving chip yield and shortening development cycles.

However, we must also clearly recognize the challenges and limitations for the successful application of this data-driven method. First, model performance is highly depen-

dent on the quality and comprehensiveness of the training dataset. The upper limit of model accuracy is constrained by the precision of the forward physical model (e.g., TMM). Any discrepancies between the physical model and the real world—such as inaccuracies in material dispersion models, or unmodeled factors like interfacial transition layers or roughness—may be “learned” by the neural network, leading to systematic measurement errors. Second, the model’s generalization capability is another critical challenge. A model trained on a specific material system (e.g., SiO₂ on Si) within a certain parameter range may exhibit significantly degraded performance when applied directly to entirely new materials (e.g., Si₃N₄) or structures outside the training range (e.g., thicker or thinner films). Enhancing model generalization, enabling it to “infer broadly from one instance,” is a necessary hurdle to overcome for moving from the laboratory to industrial application. Finally, the “black-box” nature of deep learning models results in a certain lack of physical interpretability. While they can provide accurate predictions, they struggle to intuitively offer confidence intervals for solutions, reveal intrinsic physical relationships between parameters, or assess measurement uncertainty like traditional fitting methods can. This may become a drawback in scenarios requiring high-reliability verification.

Looking forward, to build more robust and intelligent optical measurement systems, this research can be further deepened and expanded in the following directions. First, the application of Physics-Informed Neural Networks (PINNs) can be explored. Embedding the Maxwell equations governing light propagation in multilayer films, or their simplified forms (e.g., the TMM formulation), as regularization terms into the loss function would introduce physical constraints during training. This approach not only holds promise for reducing dependency on large-scale labeled data but also fundamentally ensures the physical self-consistency of the network’s output, significantly enhancing the model’s generalization capability and credibility. Second, hybrid modeling strategies combining deep learning with probabilistic models (e.g., Bayesian Neural Networks or Gaussian Process Regression) can be investigated. This would allow the model to provide confidence intervals or probability distributions for predictions alongside high-precision point estimates, enabling quantitative assessment of measurement uncertainty and providing critical support for industrial decision-making. Furthermore, efforts should be directed towards promoting the application of transfer learning and domain adaptation methods. By combining massive simulation data with limited amounts of high-quality experimental data, models can effectively transfer general knowledge acquired in simulation environments to real-world scenarios, rapidly adapting to various non-ideal factors in actual measurements (such as systematic errors and environmental noise drift). This would ultimately lead to the realization of intelligent measurement systems that can be robustly deployed on industrial production lines.

5. Conclusions

This study systematically proposes and investigates a novel deep learning-based inverse decoupling and optimization method to address the core challenges of low computational efficiency, severe parameter coupling, and solution non-uniqueness faced by white light interferometry when measuring sub-micron multilayer transparent films. The innovation lies in constructing a “classify first, then specialize” hybrid intelligent framework. This framework first employs a DNN to automatically classify complex interference signals, thereby decomposing the challenging multimodal inverse problem into a series of simpler, single-mode sub-problems. Subsequently, for each specific signal category, a cascaded neural network architecture is utilized for parameter inversion. This architecture effectively overcomes the inherent “one-to-many” mapping difficulty in optical inverse problems by cascading an inverse prediction network with a pre-trained and fixed forward physical

model network, and by using spectral reconstruction error as part of the loss function. This fundamentally ensures the convergence of the solving process and the uniqueness and physical consistency of the final solution.

Through systematic experimental measurement and analysis, the effectiveness of this method has been comprehensively validated across multiple dimensions. In classification, the model accurately distinguishes different material surfaces and film thickness regions, providing reliable classification masks for subsequent 3D topography reconstruction. Regarding parameter inversion accuracy, for various interference signals, the spectral curves reconstructed by this method show high agreement with theoretical signals, significantly improving analytical reliability and efficiency. This confirms the powerful capability of the proposed hybrid framework in decoupling complex parameters.

In summary, the deep learning method proposed in this study successfully provides a viable path to break through the technical bottlenecks of traditional white light interferometry signal inversion. By deeply integrating the powerful data-driven capability of deep learning with the prior knowledge of optical physical models, it achieves high-precision, highly robust inverse solving of structural parameters and 3D topography reconstruction for sub-micron multilayer transparent films. This research holds significant theoretical value and offers an intelligent solution with broad application prospects for online process control and intelligent quality inspection in high-precision fields such as semiconductor manufacturing.

Author Contributions: Conceptualization, Y.M. and C.C.; methodology, L.C.; software, C.C.; formal analysis, Y.M.; investigation, X.T.; resources, L.C.; data curation, C.C.; writing—original draft preparation, Y.M. and C.C.; writing—review and editing, J.Z., G.W. and Z.L.; visualization, Y.M.; supervision, J.Z., G.W. and Z.L.; project administration, X.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The datasets presented in this article are not readily available due to commercial confidentiality requirements. Requests to access the datasets should be directed to the corresponding author and will be subject to approval by the collaborating company.

Conflicts of Interest: Authors Yanzhong Ma, Chi Chen, Lu Chen, Ji Zhang, and Xiaojun Tian were employed by the company Skyverse Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

WLI	White Light Interferometry
DNN	Deep Neural Network
TMM	Transfer Matrix Method

References

1. Chen, H.; Zhang, L.; Sun, W.; Chen, B.; Cao, Z.; Wu, Q. White light interferometry micro measurement algorithm based on principal component analysis. *Chin. Opt.* **2023**, *16*, 637–644. [[CrossRef](#)]
2. Chu, H.; Kim, D.; Park, J.; Kang, S.; Son, J.; Lee, H.; Yoo, H.; Kim, S.; Kim, S.; Kim, Y.-J. Nanometer-resolution white-light scanning interferometry for surface-profiling of hybrid bonding samples for advanced semiconductor packaging. *Appl. Surf. Sci.* **2025**, *689*, 162466. [[CrossRef](#)]
3. Guo, T.; Hu, C.; Chen, J.; Fu, X.; Hu, X. Vertical Scanning White-Light Interferometry for Dimensional Characterization of Microelectromechanical System Devices. *Acta Opt. Sin.* **2007**, *27*, 665–670.

4. Scholz, G.; Yang, L.; Schake, M.; Fortmeier, I. Concept for improving the form measurement results of aspheres and freeform surfaces in a tilted-wave interferometer. *J. Sens. Sens. Syst.* **2024**, *13*, 89–97. [[CrossRef](#)]
5. Huang, Y.; Gao, J.; Zhang, L.; Deng, H.; Chen, X. Fast template matching method in white-light scanning interferometry for 3D micro-profile measurement. *Appl. Opt.* **2020**, *59*, 1082–1091. [[CrossRef](#)] [[PubMed](#)]
6. Yuan, L.; Guo, T.; Tang, D.; Liu, H.; Guo, X. Simultaneous film thickness and refractive index measurement using a constrained fitting method in a white light spectral interferometer. *Opt. Express* **2021**, *30*, 349–363. [[CrossRef](#)] [[PubMed](#)]
7. Duan, M.; Li, J.; Bi, S.; Zong, Y.; Yu, C.; Guo, R. Dynamic white-light interferometry via multiwavelength tilt iteration and sliding local least squares. *Opt. Lett.* **2021**, *46*, 5810–5813. [[CrossRef](#)] [[PubMed](#)]
8. You, H.; Du, L.; Xu, X.; Zhao, J. Deep learning methods for solving non-uniqueness of inverse design in photonics. *Opt. Commun.* **2024**, *554*, 130122. [[CrossRef](#)]
9. de Groot, P.; de Lega, X.C. Signal modeling for low-coherence height-scanning interference microscopy. *Appl. Opt.* **2004**, *43*, 4821–4830. [[CrossRef](#)] [[PubMed](#)]
10. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.