https://doi.org/10.1038/s41378-023-00529-9

Microsystems & Nanoengineering
www.nature.com/micronano

ARTICLE | Open Access

Nondestructive monitoring of annealing and chemical–mechanical planarization behavior using ellipsometry and deep learning

Qimeng Sun[1], Dekun Yang[1], Tianjian Liu[2], Jianhong Liu[3], Shizhao Wang[4], Sizhou Hu[3], Sheng Liu[1,5]✉ and Yi Song[1,5]✉

Abstract
The Cu-filling process in through-silicon via (TSV-Cu) is a key technology for chip stacking and three-dimensional vertical packaging. During this process, defects resulting from chemical–mechanical planarization (CMP) and annealing severely affect the reliability of the chips. Traditional methods of defect characterization are destructive and cumbersome. In this study, a new defect inspection method was developed using Mueller matrix spectroscopic ellipsometry. TSV-Cu with a 3-µm-diameter and 8-µm-deep Cu filling showed three typical types of characteristics: overdishing (defect-OD), protrusion (defect-P), and defect-free. The process dimension for each defect was 13 nm. First, the three typical defects caused by CMP and annealing were investigated. With single-channel deep learning and a Mueller matrix element (MME), the TSV-Cu defect types could be distinguished with an accuracy rate of 99.94%. Next, seven effective MMEs were used as independent channels in the artificial neural network to quantify the height variation in the Cu filling in the z-direction. The accuracy rate was 98.92% after training, and the recognition accuracy reached 1 nm. The proposed approach rapidly and nondestructively evaluates the annealing bonding performance of CMP processes, which can improve the reliability of high-density integration.

Introduction

In recent years, three-dimensional (3D) integrated circuit (IC) technology with through-silicon via (TSV) has attracted significant attention because of its versatility, small size, and high performance. 3D IC is a technology that reduces the overall wire length and delay by vertically stacking multiple chips through high-density chip-to-chip interconnects[1–7]. TSV technology involves several processes, including etching holes in Si chips, depositing insulating/blocking/seeding layers, filling blind holes with Cu conductors, removing the backside Si and Cu overlay films via chemical–mechanical planarization (CMP) to expose Cu microcartridges, and ball bonding[8,9]. Several researchers have investigated the TSV manufacturing process. The microcharacterization and nanocharacterization of various processes have also received considerable attention in improving the reliability of 3D ICs.

Significant issues in TSV technology include residual stress, extrusion, cracking, delamination, and Cu leaks, leading to severe reliability problems, such as warping of chips or wafers, interface delamination, cracking, and weak internal wire contact. These problems mainly result from the significant difference in the coefficient of thermal expansion between the surrounding silicon substrate and the TSV filler metal[10–12]. Recently, studies on the Cu-filling process in TSV (TSV-Cu) protrusion have been conducted[13–15]. In these processes, detailed physical analyses are required because of the considerable impact of TSV-Cu protrusion on the reliability of the final 3D stacking. Many studies have been conducted to investigate

Correspondence: Sheng Liu (shengliu@whu.edu.cn) or
Yi Song (yi.song@whu.edu.cn)
[1]The Institute of Technological Sciences, Wuhan University, Wuhan, China
[2]School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China
Full list of author information is available at the end of the article

© The Author(s) 2023

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

annealing Cu protrusion and CMP dishing[16–23]. Cu protrusion and excessive dishing reduce the reliability of 3D stacking. In TSV manufacturing, it is crucial to control the CMP speed, annealing time, and temperature. Insufficient or excessive wafer polishing can lead to leaks and shorts, making the chips defective. Residual stress, interface delamination, and cracking occur when the annealing time or temperature is insufficient or excessive[6,24–26]. However, real-time nondestructive characterization methods for Cu protrusions and dishing remain limited.

Nondestructive characterization of TSV-Cu in a production line has not been extensively investigated. When many defects appear during the TSV process, the detection time obtained using conventional methods is unsatisfactory[5,9]. The period and critical size of the TSV-Cu structure prevail on the microscale for both vertical and horizontal dimensions. However, during CMP and annealing processes, the accuracy of the z-direction of Cu must be controlled at the nanoscale[5,15,24]. Cross-scale defect characterization in the z-direction of TSV-Cu structures is therefore a significant challenge.

In this study, an ellipsometry measurement method was developed for the characterization of TSV-Cu. The rigorous coupled-wave analysis (RCWA) algorithm was employed to calculate the reflection electric field of TSV-Cu with different dimensions of Cu filling in the z-direction at different wavelengths. The corresponding Mueller matrices were calculated by modulating the incident and reflective electric fields, as described in the "Methods" section. Deep learning with Mueller matrix datasets (see the "Methods" section for details) was used to distinguish the different TSV-Cu defects. The effects of the annealing temperature and time on the Cu grain size were experimentally studied, and the defect size of TSV-Cu was quantified using a multichannel deep-learning method. A single Mueller matrix element (MME) was replaced with multiple MMEs as a dataset to improve the stability and accuracy of the quantitative TSV-Cu size.

Results and discussion

Defect classification of TSV-Cu

Figure 1a shows an atomic force microscopy image of a 3-μm-diameter and 8-μm-deep TSV-Cu structure after annealing at 250 °C for 8.5 h. An optical model was established based on this structure, as shown in Fig. 1b. The model comprised 20 nm Ta as a blocking layer, 200 nm SiO_2 as an insulating layer, a silicon substrate, and a Cu-filling layer with a diameter of 3 μm and depth of 8 μm. The Cu-filling period was 6 μm, the incidence angle was fixed at 45°[27], and the wavelength range was

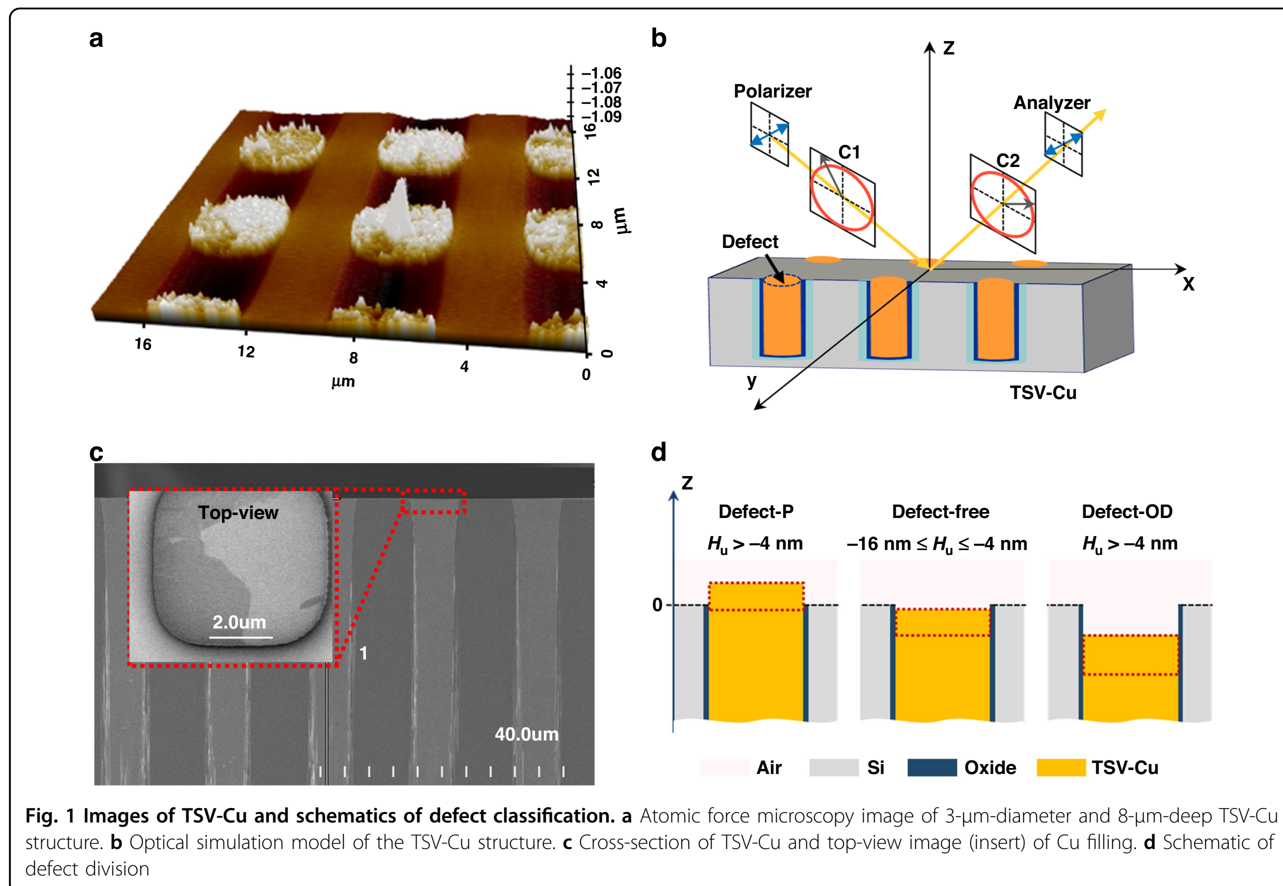

Fig. 1 Images of TSV-Cu and schematics of defect classification. a Atomic force microscopy image of 3-μm-diameter and 8-μm-deep TSV-Cu structure. **b** Optical simulation model of the TSV-Cu structure. **c** Cross-section of TSV-Cu and top-view image (insert) of Cu filling. **d** Schematic of defect division

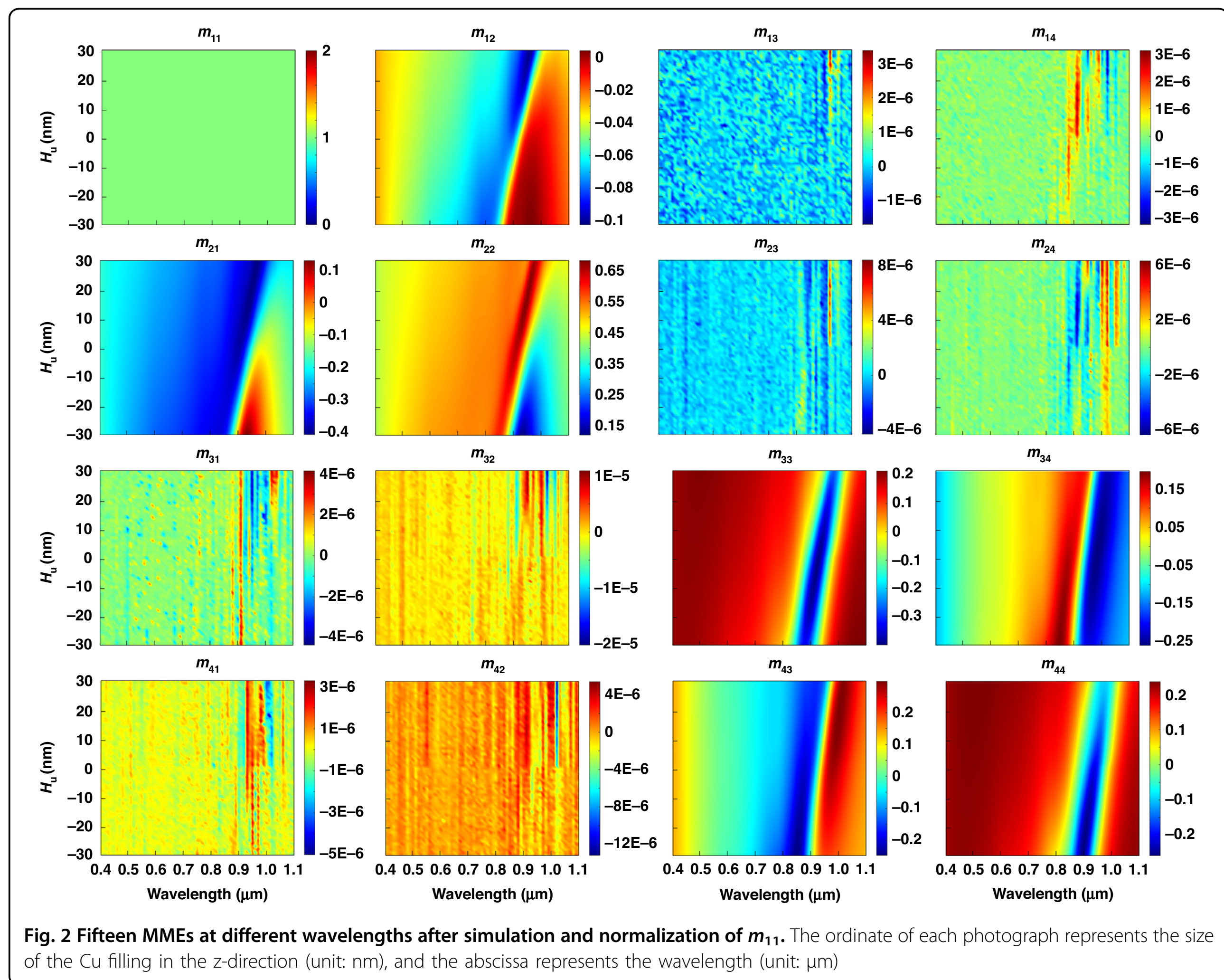

Fig. 2 Fifteen MMEs at different wavelengths after simulation and normalization of m_{11}. The ordinate of each photograph represents the size of the Cu filling in the z-direction (unit: nm), and the abscissa represents the wavelength (unit: μm)

400–1000 nm. The scanning electron microscopy images of the TSV-Cu cross-section are shown in Fig. 1c. The red dotted line represents the areas where different types of defects appeared in this study. The top view of the Cu-filled TSV is shown by the red dotted line (Fig. 1c). During the annealing process, when the temperature is lower than 200 °C, the height of the upper surface of the Cu-filling ($\boldsymbol{H_u}$) changes due to the significant mismatch of the coefficient of thermal expansion (CTE) between the Cu-filling and silicon substrate. Due to the higher CTE of Cu, the expansion of Cu is greater than that of silicon. At this temperature, there is almost no change in the microstructure of Cu-filling. The recrystallization of Cu-filling occurs at approximately 200–250 °C, at which point the Cu grain is refined. Once the annealing temperature exceeds 250 °C, the grain grows significantly as driven by the total free energy[3]. $\boldsymbol{H_u}$ is positively correlated with the grain size[14]. Therefore, $\boldsymbol{H_u}$ increases subsequentially. During the CMP process, the removal of the barrier layer and insulating layer is slower than that of Cu filling. Through this process, $\boldsymbol{H_u}$ decreases subsequentially. When $\boldsymbol{H_u}$ is excessively large or small, there is a risk of cracks and voids, respectively. Under the effect of poor bonding quality, the electric and thermal resistance of TSV-Cu increase, and the surface bonding strength decreases. Therefore, a reasonable $\boldsymbol{H_u}$ process window (−16 to −4 nm) is proposed in this work to achieve void-free and reliable interconnection[7,13]. A schematic of the defect types is shown in Fig. 1d. The upper surface of Si was regarded as the reference plane of 0 nm. When $\boldsymbol{H_u}$ exceeded 0 nm or was lower than 4 nm in the z-direction, it was considered a protrusion defect (defect-P). $\boldsymbol{H_u}$ ranging between 4 and 16 nm was considered free from defects (defect-free). $\boldsymbol{H_u}$ exceeding 16 nm indicated an overdishing defect (defect-OD). Defect-P and defect-OD caused extrusion cracking and incomplete contact during annealing bonding, respectively. The actual process window (4–16 nm) was expected to have a higher upper limit because no increase in the dielectric bond strength was assumed.

Figure 2 shows the 4 × 4 MMEs at different wavelengths after ellipsometry and normalization of m_{11}. The ordinate

of each photograph represents $\boldsymbol{H}_{\mathbf{u}}$, and the abscissa represents the wavelength. The off-diagonal MMEs were nearly zero because the structure of TSV-Cu was isotropic[28]. The sensitivity of the MME, m_{12}, with a change in the structure, increased at 0.7–1.0 μm wavelengths. The sensitivities of m_{21} and m_{22} with structural change increased at 0.85–1.05 μm wavelengths. The sensitivities of m_{33} and m_{44} with structural change increased at wavelengths of 0.80–1.05 μm. In addition, the sensitivities of m_{34} and m_{43} with structural change increased at 0.80–0.95 μm wavelengths. We selected seven effective MMEs (m_{12}, m_{21}, m_{22}, m_{33}, m_{34}, m_{43}, and m_{44}) as the sample set for defect classification, ignoring the off-diagonal MMEs. The effective MMEs corresponding to 13 structures with $\boldsymbol{H}_{\mathbf{u}}$ ranging from −4 to −16 nm (intervals of 1 nm), −17 to −29 nm (intervals of 1 nm), and −3 to 9 nm (intervals of 1 nm) and the 300 noise dataset of each structure were adopted as the artificial neural network (ANN) datasets representing the defect-free, defect-OD, and defect-P topologies, respectively.

Figures 3a and 3b show the cross-entropy loss and verification accuracy of different MMEs in the defect classification trained network, respectively. The training convergence speed of m_{12} was faster than those of the other elements, with an increase in training time. The Mueller matrices can be decomposed using the Lu–Shipman polar decomposition method[29]:

$$D = \frac{\sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2}}{m_{11}} \tag{1}$$

where D is the dichroic scalar of the Mueller matrices after polar decomposition, and $D \approx m_{12}$. Therefore, the diattenuation of TSV-Cu with different $\boldsymbol{H}_{\mathbf{u}}$ values showed obvious changes. In defect classification training using m_{12}, the verification accuracy at the seventh epoch was 99.80%, and the corresponding cross-entropy loss was 0.00043. Hence, only a single MME was required as a sample set to complete the defect classification using deep learning. t-distributed stochastic neighbor embedding (t-SNE) was employed to analyze the defect types, as shown in Fig. 3c[30]. Different scatter colors represent different defect types, and each point represents the simulation data with 10% noise randomly generated by m_{12} (Fig. 2). The separation of the three clusters was more evident (Fig. 3c), indicating that the defect types could be distinguished precisely. m_{12}, m_{21}, m_{22}, m_{33}, m_{34}, m_{43}, and m_{44} achieved defect classification accuracy values of 99.53%, 99.75%, 98.69%, 99.80%, 99.94%, 99.77%, and 99.66%, respectively.

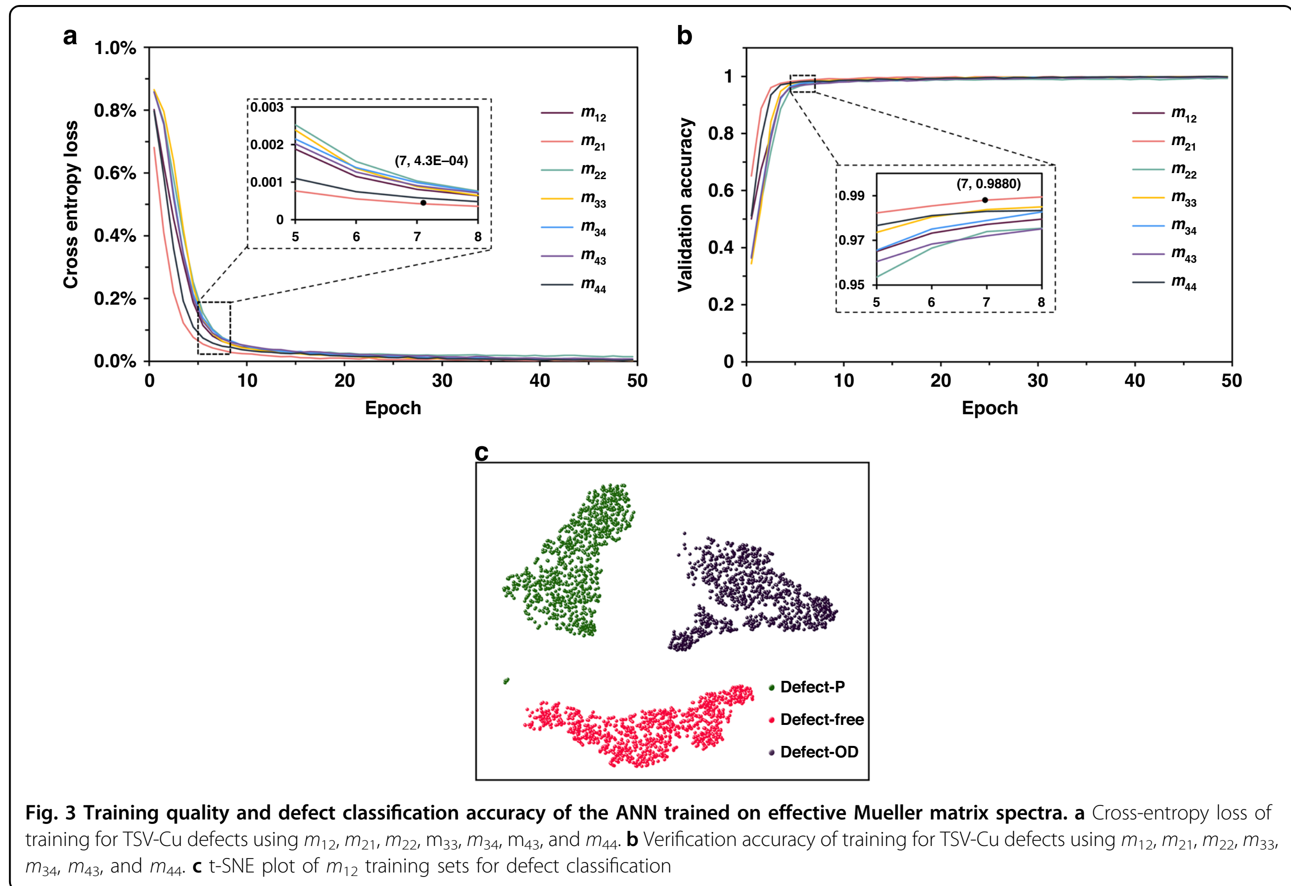

Fig. 3 Training quality and defect classification accuracy of the ANN trained on effective Mueller matrix spectra. a Cross-entropy loss of training for TSV-Cu defects using m_{12}, m_{21}, m_{22}, m_{33}, m_{34}, m_{43}, and m_{44}. **b** Verification accuracy of training for TSV-Cu defects using m_{12}, m_{21}, m_{22}, m_{33}, m_{34}, m_{43}, and m_{44}. **c** t-SNE plot of m_{12} training sets for defect classification

Quantification of TSV-Cu defect size using multichannel deep learning

During the TSV process, annealing was applied to make the Cu grains more homogeneous and to release the residual stress. The Cu grain sizes were measured using electron backscatter diffraction at different annealing temperatures and times to show the need for defect quantification and required quantization accuracy (Fig. 4). Higher temperatures and longer times facilitate grain growth, produce Cu protrusions, and release residual stress. Minute grains appeared when the temperature and time were 250 °C and 40 h, respectively. Two or more nuclei were formed during the extended annealing times. Therefore, the fusing and growth times of the grains are insufficient when the annealing time is extremely short. When the annealing time was very long, small grains were formed owing to repeated nuclei precipitation, leading to inadequate annealing. When annealed at 250 °C, the average Cu grain size increased by 60 nm from 9 to 40 h. Therefore, it is necessary to adjust the annealing time and temperature to control $\boldsymbol{H}_{\mathbf{u}}$ during annealing bonding. The proposed metrology method should be able to distinguish 10% of the process window to effectively monitor the TSV-Cu process. In this case, the recognition accuracy in the z-direction should exceed 1.2 nm to more accurately characterize $\boldsymbol{H}_{\mathbf{u}}$ in TSV-Cu and be defect-free within the 12 nm process window. In this section, we attempted to increase the recognition accuracy to 1 nm.

The training and testing accuracy and stability of deep learning must be improved to achieve a distinguishing accuracy level of 1 nm. In this study, a multichannel deep learning method was applied, in which seven effective MMEs were utilized. Because the actual characterization system causes random errors, random noise was added to the MMEs of each structure. Figure 5a shows the cross-entropy loss of a single channel (one MME) and multiple channels (seven MMEs) under 10% random noise, and Fig. 5b depicts the validation accuracy. The training object consisted of 61 labels, including $\boldsymbol{H}_{\mathbf{u}}$ ranging from −30 to 30 (in intervals of 1 nm). The convergence rate of the multichannel training cross-entropy loss and validation accuracy was faster than that of single-channel training. Figure 5c depicts the test accuracy of the single-channel and multichannel methods with different random noises. The test accuracy of the single channel was significantly influenced by the increase in noise, whereas its influence on the multichannel was less. When the random noise increased from 10% to 30%, the test accuracy of the multichannel changed by only 1.01%. When the random noise increased to 30%, the multichannel deep-learning test accuracy reached 98.92%. t-SNE was employed to quantify the size to intuitively evaluate the difference in the TSV-Cu size (Fig. 5d). In the t-SNE scatter plot, clusters with good separation indicate that the TSV-Cu size can be distinguished clearly, whereas the adjacent and overlapping clusters indicate very similar datasets. The effects of different morphologies and different measurement conditions on the measurement sensitivity were discussed using seven MMEs (Fig. 6). One such effect is the aspect ratio of TSV-Cu. The diameter of the Cu-filling is 3 μm, and the aspect ratios are 0.5, 1, 2, 4, 6, 8, and 10 (Fig. 6a). Another such effect is the critical dimension difference between the top and bottom ($\Delta_{\text{Top-Bottom}}$) of the Cu-filling. The diameter of the Cu-filling is 3 μm, the aspect ratio is 1, and the $\Delta_{\text{Top-Bottom}}$ values are 2 μm, 1.6 μm, 1.2 μm, 0.8 μm, 0.4 μm, 0.2 μm, and 0 μm (Fig. 6b). The third is that the azimuth angles of incident light include 0°, 30°, 60°, 90°, 120°, 150°, and 180° (Fig. 6c). As seen from Table 1, with the change in the TSV morphologies and measurement conditions, the test accuracies fluctuate by no more than 1%. In addition, all the cross-entropy losses and validation accuracies converge to 0 and 1, respectively, as shown in Fig. 6. Hence, multichannel deep learning using seven MMEs can quantify the $\boldsymbol{H}_{\mathbf{u}}$ of TSV-Cu with a quantization accuracy of 1 nm and is not affected by random errors, aspect ratios, $\Delta_{\text{Top-Bottom}}$ or azimuths.

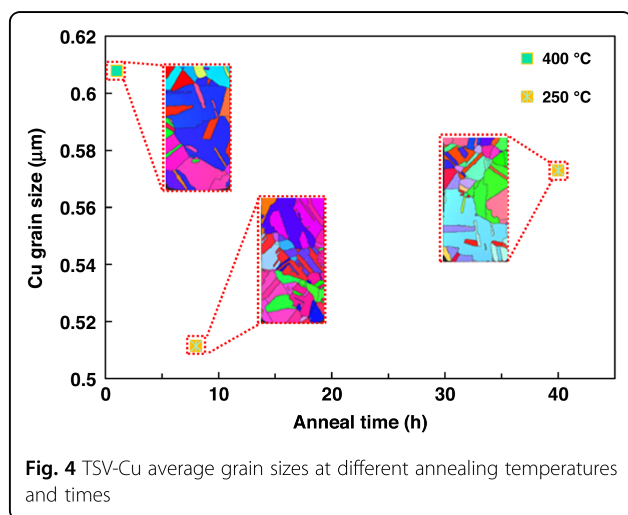

Fig. 4 TSV-Cu average grain sizes at different annealing temperatures and times

Conclusion

The polarization of light is highly sensitive to TSV-Cu particles of different sizes. In this study, three defect process windows of TSV were selected, and the Mueller matrix of TSV-Cu was calculated using Fourier series and ellipsometry. The effective MME and single-channel deep learning distinguished the defect types of TSV-Cu with an accuracy of more than 99%. Multichannel deep learning based on seven MMEs quantified the $\boldsymbol{H}_{\mathbf{u}}$ of TSV-Cu and achieved a recognition accuracy level of 1 nm. Random noise significantly impacts quantization accuracy, reaching 98.92% when random noise increases to 30%.

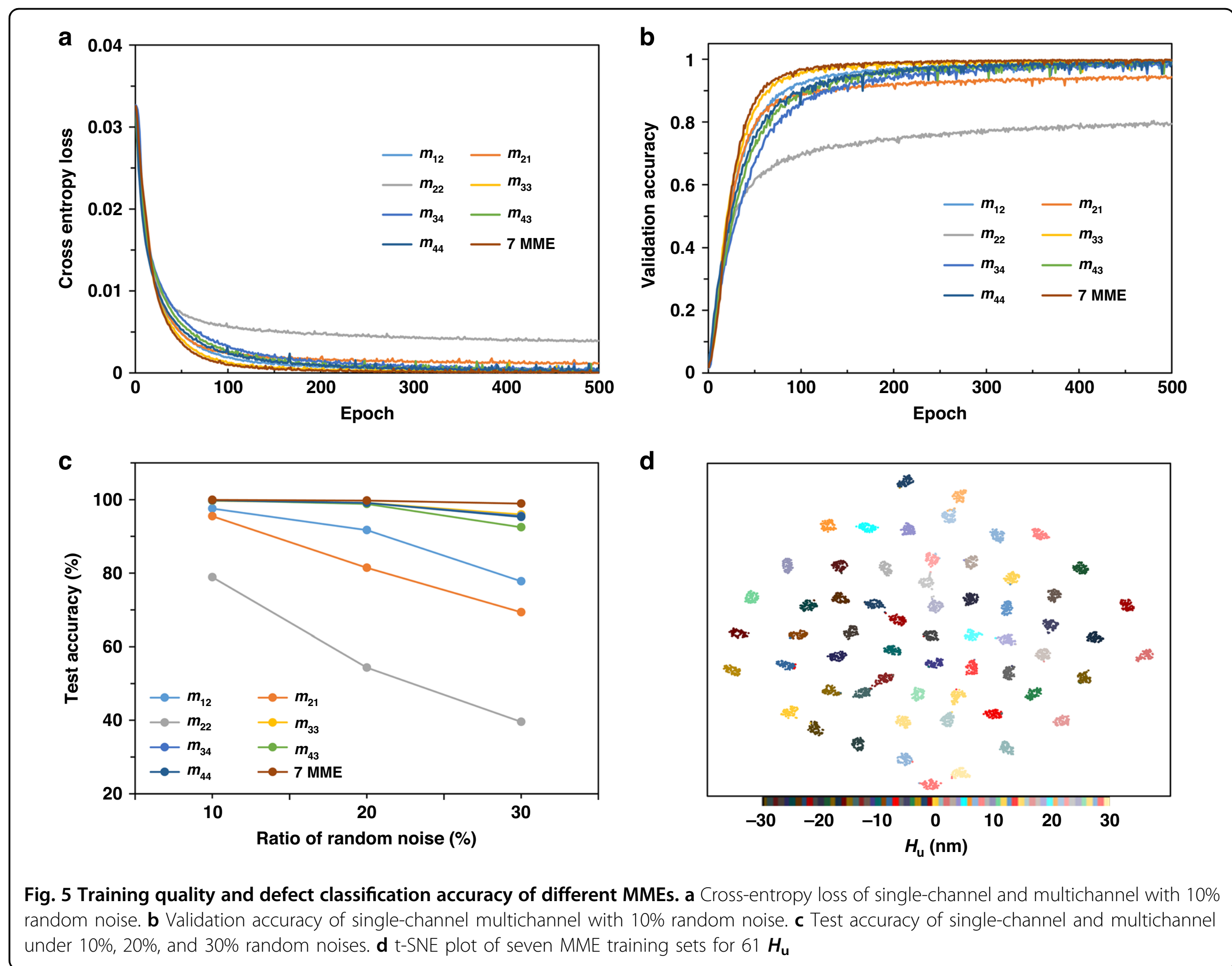

Fig. 5 Training quality and defect classification accuracy of different MMEs. a Cross-entropy loss of single-channel and multichannel with 10% random noise. **b** Validation accuracy of single-channel multichannel with 10% random noise. **c** Test accuracy of single-channel and multichannel under 10%, 20%, and 30% random noises. **d** t-SNE plot of seven MME training sets for 61 $\boldsymbol{H}_{\mathbf{u}}$

Multichannel deep learning is more accurate and stable and less affected by random noise, aspect ratios, $\Delta_{\text{Top-Bottom}}$, and azimuths, which can improve the reliability of high-density integration in micro- and nanofabrication technologies.

Methods

Optical simulation and calculation of the Mueller matrix

The modulation method of the incident optical wave can be considered natural light transmitting a stationary linear polarizer and a quarter-wave plate in turn, and the fast-axis rotation angle is θ. The Stokes vector of natural light is $S_0=\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$. The Mueller matrix of the linear polarizer is $M_P=\frac{1}{2}\begin{bmatrix}1&1&0&0\\1&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$ and that of the quarter-wave plate is $M_{C1}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&-1&0\end{bmatrix}$. The rotation matrix is $R(\theta)=\begin{bmatrix}1&0&0&0\\0&\cos 2\theta&\sin 2\theta&0\\0&-\sin 2\theta&\cos 2\theta&0\\0&0&0&1\end{bmatrix}$, where θ is the fast-axis rotation angle of the first quarter-wave plate with values of 0°, 5°, 10°, 15°,..., 170°, 175°, and 180°. Therefore, the Stokes vector of incident light on the sample, S_{in}, can be defined as follows.

$$S_{\text{in}} = R(5\theta)M_{C1}R(-5\theta)M_P S_0 \tag{2}$$

In addition, the polarization angle and phase difference of the incident optical wave can be expressed as follows:

$$\Psi = -\text{Sign}\{\sin 2\theta\}\ \arcsin\sqrt{\frac{1-\cos^2 2\theta}{2}} \tag{3}$$

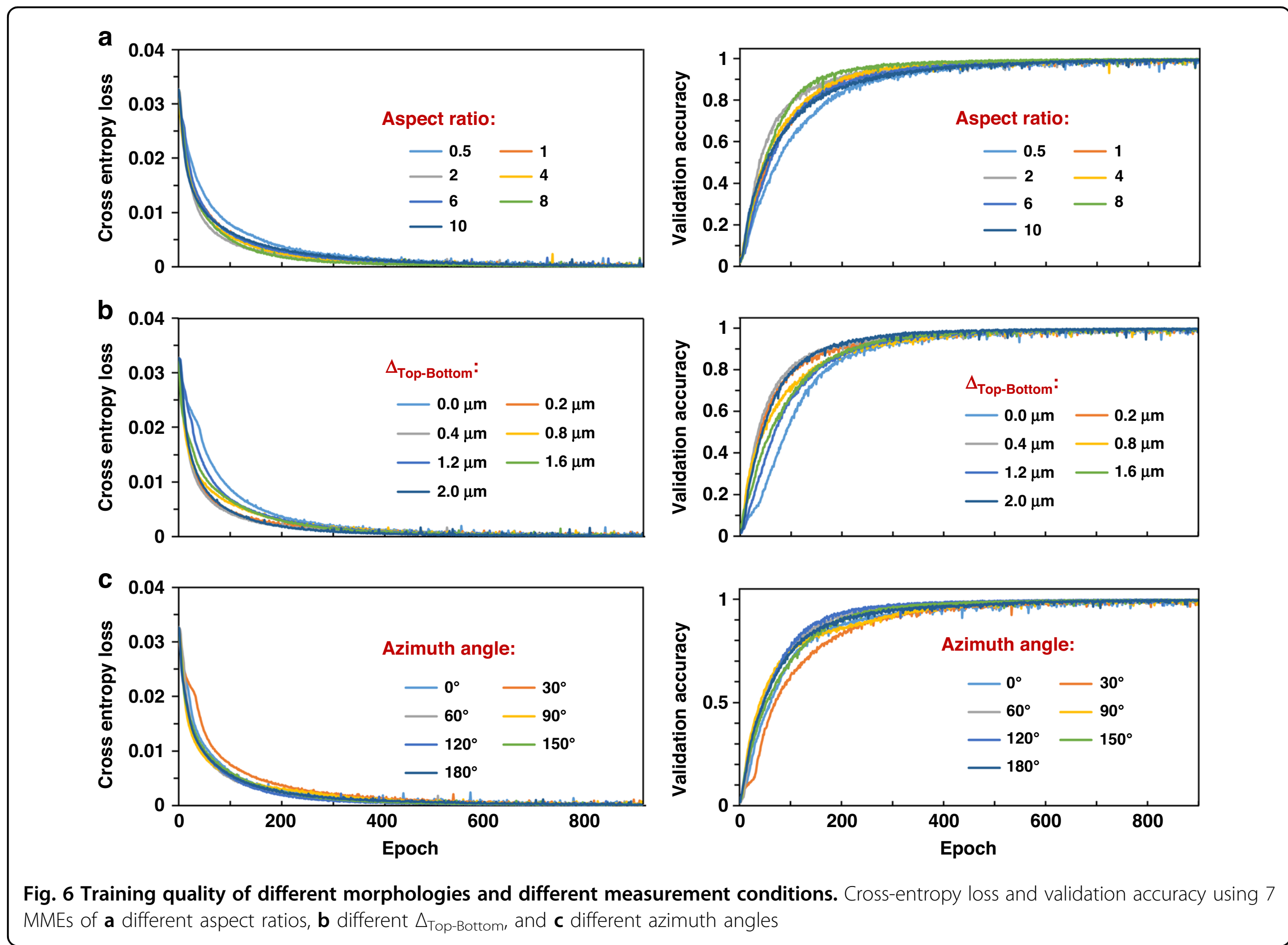

Fig. 6 Training quality of different morphologies and different measurement conditions. Cross-entropy loss and validation accuracy using 7 MMEs of **a** different aspect ratios, **b** different $\Delta_{Top\text{-}Bottom}$, and **c** different azimuth angles

Table 1 The test accuracies with different morphologies and different measurement conditions

Aspect ratio	0.5	1	2	4	6	8	10
Test accuracy (%)	99.40	99.96	99.62	99.89	99.69	99.32	99.98
$\Delta_{Top\text{-}Bottom}$ (μm)	0	0.2	0.4	0.8	1.2	1.6	2
Test accuracy (%)	99.78	99.82	98.87	99.65	99.98	99.94	99.96
Azimuth angle (°)	0	30	60	90	120	150	180
Test accuracy (%)	99.75	99.58	99.98	99.96	99.96	99.93	99.54

MPL max-pooling layer

$$\Phi = \frac{\arccos[\cos 2(\theta + 90^\circ)]}{\sqrt{1+\cos^2 2(\theta+90^\circ)}} \quad (4)$$

where Sign$\{x\}$ = 1 if $x > 1$, Sign$\{x\}$ = 0 if x = 0, and Sign$\{x\} < 0$ if $x < 0$. Ψ is the polarization angle of the incident optical wave, where $\Psi = 0°$ is p-polarized and $\Psi = 90°$ is s-polarized. Φ is the phase difference between the incident optical waves, and θ is the fast-axis rotation angle of the first quarter-wave plate.

The reflected electric fields, E_x and E_y, and the corresponding phases of incident light after passing through TSV-Cu were calculated using the RCWA algorithm[31]. The RCWA algorithm was integrated into RSoft Diffractmod (RSoft Design Group, Inc., USA). The relationship between the reflected field and wavelength of the structural, optical model representing the periodic grating was calculated. $M_{TSV-Cu} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix}$ is the Mueller matrix of TSV-Cu. The Stokes vector reflected by the sample, S_s, can be expressed as follows:

$$S_S = \begin{bmatrix} S_{s1} \\ S_{s2} \\ S_{s3} \\ S_{s4} \end{bmatrix} = \begin{bmatrix} E_x^2 + E_y^2 \\ E_x^2 - E_y^2 \\ 2E_xE_y\cos\delta \\ 2E_xE_y\sin\delta \end{bmatrix} = M_{TSV-Cu}R(5\theta)M_{C1}R(-5\theta)M_P S_0 \quad (5)$$

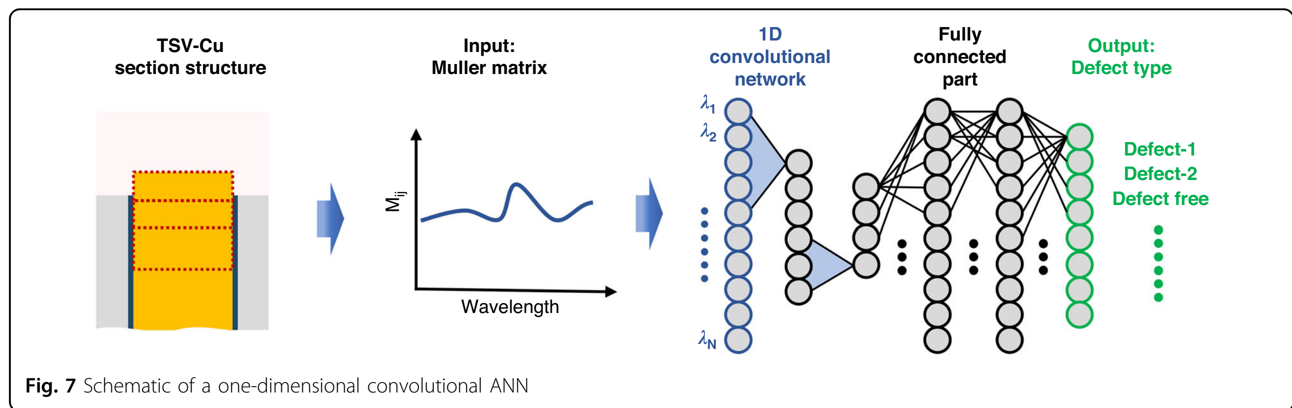

Fig. 7 Schematic of a one-dimensional convolutional ANN

Table 2 Number of filters and kernel size of each channel in each CNN layer

Layer	Number of filters in each channel	Kernel size	Kernel size of MPL
First	32	7	2
Second	32	5	2
Third	1	3	2

where δ is the phase difference between E_x and E_y. The light reflected by the sample is remodulated as S_{out}:

$$S_{out} = \begin{bmatrix} S_{out1} \\ S_{out2} \\ S_{out3} \\ S_{out4} \end{bmatrix} = M_A R(3\theta) M_{C2} R(-3\theta) S_S \quad (6)$$

where M_A is the Mueller matrix of the analyzer, $M_A = M_P$. M_{C2} is the Mueller matrix of the second quarter-wave plate, $M_{C2} = M_{C1} \cdot S_{out1}$. is the total light intensity detected by the detector. In this study, S_{out1} was expanded using the Fourier series. Because the rotation angle ratio of the incident and outgoing quarter-wave plates was 5:3, the highest order of the Fourier series was 32. Sixteen MMEs of TSV-Cu were calculated, and all elements were normalized to m_{11}[32].

Deep learning

In this study, we developed a machine learning approach to classify defects. The convolutional neural network (CNN) was employed, matching a fully connected network (FCN) (Fig. 7). The "leaky ReLU" activation function, a preferred method for solving pattern recognition problems, was used, and it consisted of three layers[33,34]. Each CNN layer was connected to a max-pooling layer with batch normalization. The number of filters and kernel size of each channel in each CNN layer are listed in Table 2. The FCN consisted of two layers with 64 and 32 neurons and utilized "tanh" activations. Mueller matrices of different $\boldsymbol{H}_{\mathbf{u}}$ at different wavelengths were introduced into the ANN input layer as parallel channels, and the kernel size of each channel was seven. Each defect type was assigned to a neuron in the ANN output layer. During the training of the ANN, a Mueller matrix was introduced into the ANN input layer, and different defect types were introduced into the output layer. Hundreds or tens of sample sets for each group of structures were not produced because of the cost and time in the actual TSV-Cu production process. During the process of sample measurement, there are two random errors, including the sample manufacturing error and measurement error[35,36]. The manufacturing error of surface displacement fluctuation of 0.1 nm is inevitable in the TSV-Cu process[23]. It is difficult to obtain specific measurement error because the measurement error depends on the instrument and environment. In practice, the spectral change caused by measurement error is usually smaller than that caused by manufacturing error. The spectrum of TSV-Cu surface fluctuation ($\boldsymbol{H}_{\mathbf{u}} \pm 0.1$ nm) was analyzed, and all effective spectral change rates were less than 10% (Fig. 8). It can be seen from Fig. 8 that only a small number of spectral change rates are close to 10%. Therefore, 60 structures were simulated, each with a Cu pillar size interval of 1 nm in the z-direction, and 300 sets of 10% noise from the Mueller matrix of each structure at different wavelengths were generated as the training dataset. Seventy percent of the Mueller matrices were used for training. The remaining 30% were used for testing. In the readout scheme, the Mueller matrix was fed into the ANN after training and propagated forward through the network to retrieve the defect types distinguished by the z-direction structure. Finally, the cross-entropy loss and validation accuracy were determined to verify the training quality of the ANN.

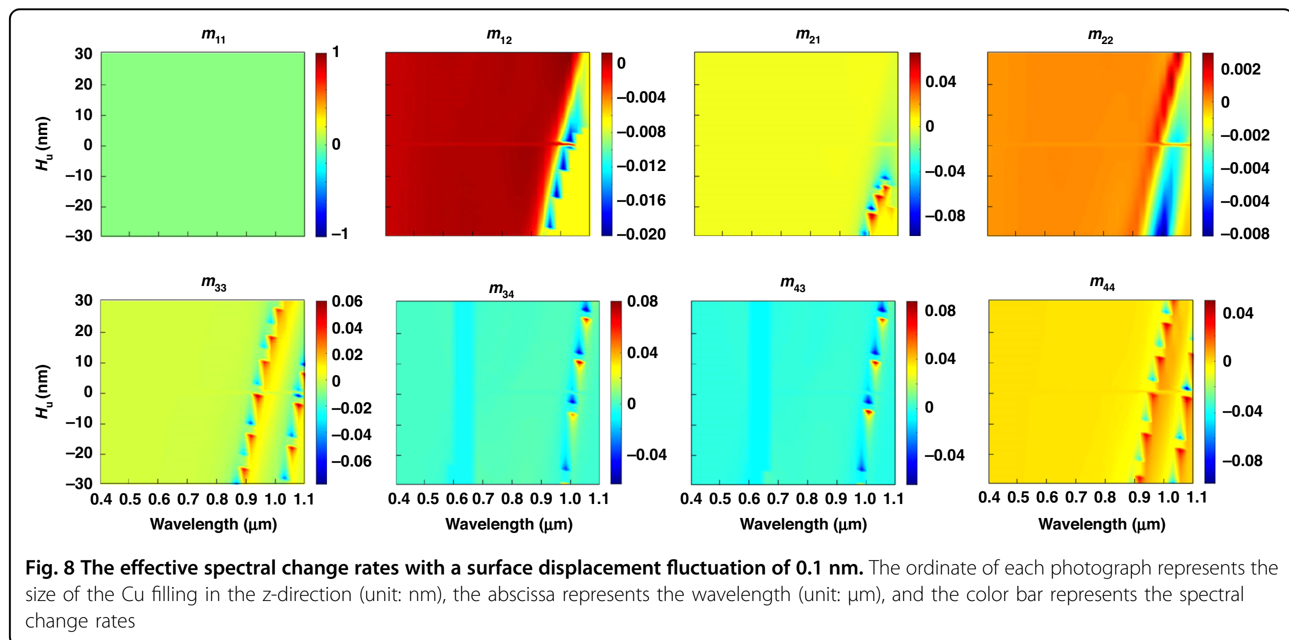

Fig. 8 The effective spectral change rates with a surface displacement fluctuation of 0.1 nm. The ordinate of each photograph represents the size of the Cu filling in the z-direction (unit: nm), the abscissa represents the wavelength (unit: µm), and the color bar represents the spectral change rates

Acknowledgements
This work is supported by the National Key Research and Development Program of China (2022YFB3206000) and the Key Research and Development Program of Hubei (2021BAA173).

Author details
[1]The Institute of Technological Sciences, Wuhan University, Wuhan, China. [2]School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China. [3]Hongyi Honor College of Wuhan University, Wuhan, China. [4]School of Power and Mechanical Engineering, Wuhan University, Wuhan, China. [5]The School of Microelectronics, Wuhan University, Wuhan, China

Author contributions
Q.S. and D.Y. conceptualized the work and provided the deep learning method. J.L., S.W., and S.H. designed the detailed experiments and characterizations with support from T.L. Q.S. prepared the manuscript with support from all co-authors. Y.S. and S.L. discussed the results and commented on the manuscript. Y.S. supervised the research.

Conflict of interest
The authors declare no competing interests.

Received: 22 October 2022 Revised: 4 March 2023 Accepted: 27 March 2023
Published online: 28 April 2023

References

1. Pahwa, R. S. et al. Automated void detection in TSVs from 2D X-ray scans using supervised learning with 3D X-ray scans. in *2021 IEEE 71st Electron. Compon. Technol. Conf. (ECTC)*, 842–849 (2021).
2. Smallwood, D. C., McCloskey, P., O'Mathuna, C., Casey, D. P. & Rohan, J. F. Methods for latent image simulations in photolithography with a polychromatic light attenuation equation for fabricating VIAs in 2.5D and 3D advanced packaging architectures. *Microsyst. Nanoeng.* **7**, 39 (2021).
3. Li, Y. D. et al. Constitutive modeling of annealing behavior in through-silicon vias-copper. *Mater. Charact.* **179**, 111359 (2021).
4. Khorramdel, B. et al. Inkjet printing technology for increasing the I/O density of 3D TSV interposers. *Microsyst. Nanoeng.* **3**, 17002 (2017).
5. Sonawane, D. & Kumar, P. Role of grain boundary sliding in structural integrity of Cu-filled through Si by during isothermal annealing. *J. Electron. Mater.* **50**, 767–778 (2021).
6. Zhang, M. et al. Protrusion of through-silicon-via (TSV) copper with double annealing processes. *J. Electron. Mater.* **51**, 2433–2449 (2022).
7. Ren, H. X., Yang, Y. T., Ouyang, G. Q. & Iyer, S. S. Mechanism and process window study for die-to-wafer (D2 W) hybrid bonding. *ECS J. Solid State Sci. Technol.* **10**, 064008 (2021).
8. Lau, J. H. Overview and outlook of through-silicon via (TSV) and 3D integrations. *Microelectron. Int.* **28**, 8–22 (2011).
9. Zhang, Y. Z., Ding, G. F., Wang, H. & Cheng, P. Microstructure of electrodeposited Cu microcylinders in high-aspect-ratio blind holes and crystallographic texture of the Cu overburden film. *J. Mater. Sci. Technol.* **32**, 67–73 (2016).
10. Che, F. X., Li, H. Y. Y., Zhang, X. W., Gao, S. & Teo, K. H. Development of wafer level warpage and stress modeling methodology and its application in process optimization for TSV wafers. *IEEE Trans. Comp. Packag. Manuf. Technol.* **2**, 944–955 (2012).
11. Tu, K. N. Reliability challenges in 3D IC packaging technology. *Microelectron. Reliab.* **51**, 517–523 (2011).
12. Ryu, S. K. et al. Impact of near-surface thermal stresses on interfacial reliability of through silicon vias for 3D interconnects. *IEEE Trans. Device Mater. Reliab.* **11**, 35–43 (2011).
13. Wolf, I. D. et al. Cu pumping in TSVs: Effect of pre-CMP thermal budget. *Microelectron. Reliab.* **51**, 1856–1859 (2011).
14. Si, C. et al. Protrusion of electroplated copper filled in through silicon vias during annealing process. *Microelectron Reliab.* **63**, 183–193 (2016).
15. Bourzgui, S. et al. in *Reflection, Scattering, and Diffraction from Surfaces VI* (2018).
16. Yadong, L. et al. Constitutive modeling of annealing behavior in through silicon vias-copper. *Mater. Charact.* **179**, 111359 (2021).
17. Che, F. X. et al. Study on Cu protrusion of through-silicon via. *IEEE Trans. Compon. Packaging Manuf. Technol.* **35**, 732–739 (2013).
18. Ming, S. et al. Study on copper protrusion of through-silicon via in a 3-D integrated circuit. *Mater. Sci. Eng. A* **755**, 66–74 (2019).
19. Huang, Y. J., Pan, C. L., Lin, S. C. & Guo, M. H. Machine-learning approach in detection and classification for defects in TSV-based 3-D IC. *IEEE Trans. Compon. Pac. Manuf. Technol.* **8**, 699–706 (2018).
20. Lekha, P. et al. Preparation of spherical ceria coated silica nanoparticle abrasives for CMP application. *Appl. Surf. Sci.* **357**, 1306–1312 (2015).
21. Ke, L. et al. Comprehensive characterization of TSV etching performance with phase-contrast X-ray microtomography. *J. Synchrotron Radiat.* **27**, 1023–1032 (2020).

22. Lin, T. C. et al. Inhibiting the detrimental Cu protrusion in Cu through-silicon-via by highly (111)-oriented nanotwinned Cu. *Scr. Mater*. **197**, 113782 (2021).
23. Shizhao, W. et al. Surface action mechanism and design considerations for the mechanical integrity of Cu/low K BEOL interconnect during chemical mechanical polishing process. *Microelectron. Reliab*. **134**, 0026–2714 (2022).
24. An, Z. L. et al. Mechanically strengthened graphene-Cu composite with reduced thermal expansion toward interconnect applications. *Microsyst. Nanoeng*. **5**, 20 (2019).
25. Van Olmen, J. et al. 3D stacked IC demonstrator using hybrid collective die-to-wafer bonding with Cu through silicon vias (TSV). in *Proc. IEEE Int. Conf. 3D Syst. Integr*. 1–5 (2009).
26. Wang, J. J., Ma, L. M., Feng, J. Y., Wang, Y. S. & Guo, F. Study on thermal shock and annealing behavior of $Sn_3Ag_{0.5}Cu$-TSV prepared by modified molten metal infiltration method. *J. Electron. Mater.* **51**, 4054–4062 (2022).
27. Chen, X. G., Jiang, H., Zhang, C. W. & Liu, S. Y. Toward understanding the detection of profile asymmetry from Mueller matrix differential decomposition. *J. Appl. Phys.* **118**, 225308 (2015).
28. Dixit, D. et al. Nonconventional applications of Mueller matrix-based scatterometry for advanced technology nodes. *J. Micro-Nanolithogr. MEMS MOEMS* **17**, 034001 (2018).
29. Lu, S. & Chipman, R. A. Interpretation of Mueller matrices based on polar decomposition. *J. Opt. Soc. Am. A* **13**, 5 (1996).
30. Laurens, V. D. M. & Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
31. Moharam, M. G., Grann, E. B., Pommet, D. A. & Gaylord, T. K. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.* **12**, 1068–1076 (1995).
32. Collins, R. W. & Koh, J. Dual rotating-compensator multichannel ellipsometer: instrument design for real-time Mueller matrix spectroscopy of surfaces and films. *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.* **16**, 1997–2006 (1999).
33. Albawi, S., Mohammed, T. A. & Al-Zawi, S. Understanding of a convolutional neural network. in *2017 Int. Conf. Eng. Technol., IEEE*, 1–6 (2017).
34. He, K. M., Zhang, X. Y., Ren S. Q. & Sun J. Deep residual learning for image recognition. in *2016 IEEE Conf. on Comp. Vis. and Pat. Recog. (CVPR)*, 770–778 (2016).
35. Peiting, Z., Fei, P., Dekun, Y., Zhidan, L. & Yi, S. A Laplace sensitivity operator enhances the calculation efficiency of OCD metrology. *Opt. Express* **31**, 2147–2160 (2023).
36. Dekun, Y. et al. High optical storage density using three-dimensional hybrid nanostructures based on machine learning. *Opt. Lasers Eng.* **161**, 0143–8166 (2023).