Key Engineering Materials
ISSN: 1662-9795, Vol. 1047, pp 11-20

Submitted: 2025-12-03
Revised: 2026-02-16
Accepted: 2026-02-25
Online: 2026-04-13

Machine Learning Topography Prediction and Optimization in Maskless Grayscale Laser Lithography

Adrian H. A. Lutey[1,a*], David Kuhness[2,b], Seyyedhossein Mckee[3,c], Marco Negozio[1,d], Markus Postl[2,e], Barbara Stadlober[2,f]

[1]Università degli studi di Parma, Dipartimento di Ingegneria dei Sistemi e delle Tecnologie Industriali, Parco Area delle Scienze 181/A, 43124 Parma, Italy

[2]Joanneum Research Forschungsgesellschaft, Franz-Pichler-Straße 30, A-8160 Weiz, Austria

[3]Università degli studi di Parma, Dipartimento di Ingegneria e Architettura, Parco Area delle Scienze 181/A, 43124 Parma, Italy

[a*]adrian.lutey@unipr.it, [b]David.Kuhness@joanneum.at, [c]seyyedhossein.mckee@unipr.it, [d]marco.negozio@unipr.it, [e]Markus.Postl@joanneum.at, [f]barbara.stadlober@joanneum.at

Keywords: maskless grayscale laser lithography, machine learning, artificial neural networks, process optimization.

Abstract. A machine learning (ML) framework was developed for the prediction of surface topography obtained with maskless grayscale laser lithography based on the spatial distribution of the applied laser energy dose, or virtual photomask. Artificial neural networks (ANNs) were employed, with the virtual photomask and its radial averages selected as input variables and the surface elevation selected as the output variable. Training of ANNs was carried out with data acquired from the production of models comprising a wide range of representative geometries. Hyperparameter optimization was performed by assessing the accuracy of trained ANNs, with the final configuration comprising a single hidden layer with 15 neurons and a Sigmoid activation function. The trained ANN was then employed within an iterative optimization algorithm to determine the best virtual photomask for the production of new objects by updating the virtual photomask based on the predicted error, thus automatically compensating for proximity effects and sharp dose transitions. The developed approach achieved a reduction in average build error from 2.8 µm to 1.3-1.5 µm compared to standard experimental approaches in a single build, improving not only accuracy but also greatly reducing time requirements for optimization of the process.

Introduction

Maskless grayscale laser lithography is a key manufacturing process for the production of microscale components and features such as micro-electromechanical systems (MEMS), micro-optical components and functional surfaces [1–3]. By locally modulating light intensity during exposure of a photoresist, continuous-relief structures can be produced with greater flexibility than traditional lithography, offering advantages in terms of versatility and potential applications [4,5]. Definition of the spatial distribution of the applied laser energy dose, or virtual photomask, is typically achieved by performing a series of single-line experiments, measuring the resulting elevation and defining an analytical contrast curve expressing the surface topography as a function of energy dose. This approach does not account for sharp dose transitions or proximity effects due to adjacent features, which can lead to significant deviations from the target geometry during the production of new objects [6]. Optimization of the virtual photomask therefore typically requires several iterations, requiring lengthy steps to be repeated due to difficulty in predicting outcomes within regions subject to sharp dose transitions or proximity effects [2].

Physics-driven models address these limitations by determining the equivalent absorbed laser energy dose and consequent changes in photoresist solubility, allowing fast and accurate correction of virtual photomasks for maskless grayscale laser lithography performed with established process setups and photoresists [7,8]. The applicability of such models, however, is limited by their complexity and the requirement for specialized software, as well as their sensitivity to calibration of

material properties and constants [7], particularly when changes to the process setup or photoresist are introduced. Machine learning (ML) offers a valid alternative to physics-driven modeling by automatically accounting for causal relationships between process parameters, material properties and process outcomes based on training data, allowing optimization of the virtual photomask required for a given object based on predicted outcomes [9,10]. ML models have seen application in traditional lithography for definition of optimized photomasks [11]. Artificial neural networks (ANNs) have been applied to maskless grayscale laser lithography for the prediction of optimal virtual photomasks based on U-Net convolutional neural networks [12]; however, this approach requires the design and training of complex deep neural networks while still not achieving accuracy up to the physical limits of the process.

Key issues surrounding the implementation of ML for topography prediction and optimization in maskless grayscale laser lithography include the development of standardized protocols for the acquisition of high-quality training datasets, selection of appropriate input variables and ML models, and development of optimization algorithms for the effective minimization of build error. The present work therefore addresses these issues by (i) employing tailored models for the acquisition of high-quality training and testing datasets, (ii) defining and evaluating appropriate input variables based on physical principles, (iii) implementing grid hyperparameter optimization of ANNs over appropriate ranges, (iv) training a final ANN representation of the process and (v) developing an iterative algorithm for virtual photomask optimization based on the trained ANN. The developed methodology is shown to be effective at optimizing virtual photomasks, reducing the average build error from 2.8 μm to 1.3-1.5 μm compared to standard experimental approaches in a single build.

Methods and Materials

Models were employed to generate training and testing data for the developed ML approach. The model presented in Figure 1a was employed to generate training data with a variety of features in terms of height, aspect ratio, slope and curvature. The model was originally developed by Minetola et al. for LPBF [13], requiring scaling to 165 μm (l) × 165 μm (w) × 20 μm (h) for production via maskless grayscale laser lithography in the present study. High aspect ratio features exceeded the limits of the employed experimental setup, thus providing information regarding process limits within the training dataset. The model presented in Figure 1b was instead employed to generate an independent testing dataset with simple geometric forms of various heights.

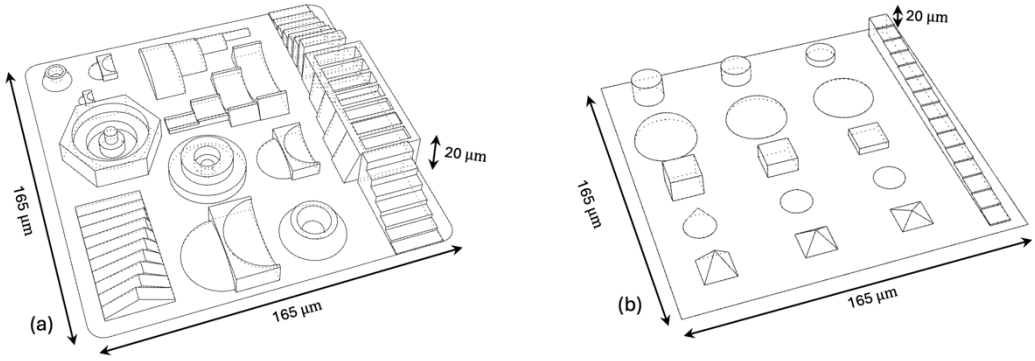

Fig. 1. (a) Training [13] and (b) testing models employed within the study.

Maskless grayscale laser lithography experiments were performed with a 405 nm continuous-wave diode laser, employing a series of filters to achieve a maximum laser power of 1 mW at the sample surface. Active control of laser power over the range 0-1 mW was achieved with an acousto-optic modulator (AOM) prior to focusing with a 10× objective to a spot size of approximately 2 μm. Samples were prepared by spin-coating 4" borosilicate glass wafers with mr-P 22G_XP positive photoresist (micro resist technology GmbH) for 20 s at 1000 RPM, prior to baking at 30°C for 30 min, 50°C for 15 min, 80°C for 60 min and 90°C for 360 min, and relaxation for 360 min. The final photoresist thickness was 50 μm. For each experiment, a virtual photomask was created by defining

the spatial distribution of the AOM voltage over the required build area. In order to determine the virtual photomask value at each point for a given target geometry, a series of single-line experiments was firstly performed to define an experimental contrast curve expressing the surface topography as a function of AOM voltage. Hatch spacing and scanning speed were held constant at 1 µm and 40 mm/s, respectively, for all experiments. After laser exposure, development of samples was performed for 30 min with TMAH-based ma-D 532/S (micro resist technology GmbH), followed by rinsing with deionized water and blow drying with nitrogen gas. Room temperature was 21 ± 1°C and humidity was 44 ± 3% at all times. Prior to measurement, all samples were sputter coated with Au-Pt to a thickness of 30 nm. Acquisition of the surface topography was performed with a Keyence laser scanning confocal microscope (LSCM) equipped with a 50× objective.

To obtain a training dataset, the training model (Figure 1a) was employed in upright and inverted positions, creating virtual photomasks based on the aforementioned experimental contrast curve, as well as a linear contrast curve assigning AOM voltages proportional to the model height. Photomasks created with both contrast curves were employed to ensure adequate representation of the physical process, including at low laser power below the interaction threshold and at high laser power up to system limits. Further to providing training data, experimental outcomes obtained with the experimental contrast curve were used as a baseline for assessing improvements in build accuracy with the developed ML approach. For acquisition of an independent testing dataset, the testing model (Figure 1b) was employed in both upright and inverted positions, creating a virtual photomask based on the experimental contrast curve. Further to providing testing data, experimental outcomes were also used as a baseline to assess improvements in build accuracy obtained with the developed ML approach.

ANN inputs were defined as the virtual photomask value (AOM voltage), together with a series of radial averages representing the radial distribution function of the virtual photomask at each point. The ANN output was the surface elevation acquired via LSCM. Individual radial averages of the virtual photomask were considered over the ranges 0 µm < r ≤ 2 µm, 2 µm < r ≤ 4 µm, 4 µm < r ≤ 8 µm, 8 µm < r ≤ 16 µm, 16 µm < r ≤ 32 µm and 32 µm < r ≤ 64 µm, where r is the radial distance from each point. Calculation of these input variables was achieved by performing two-dimensional convolutions of weighted Boolean masks and the virtual photomask. This approach allowed the absorbed energy dose within regions at and around each point to be accounted with a limited number of input variables, thus allowing simple ANNs to be employed. A flow diagram of the developed ML process is presented in Figure 2.

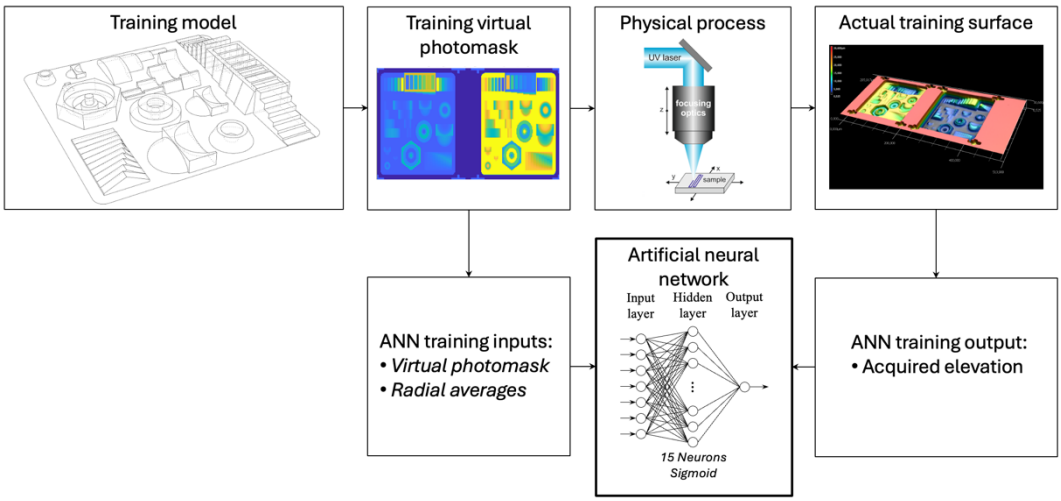

Fig. 2. Flow diagram of machine learning process for maskless grayscale laser lithography.

Training of ANNs was carried out with the *fitrnet* function in MATLAB, employing the Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm and a mean squared error (MSE) loss function. Hyperparameter optimization was performed by evaluating the MSE for prediction of outcomes in the independent testing dataset with 1-2 hidden layers, 1-25 neurons per layer and ReLU, sigmoid and tanh activation functions. The simplest configuration achieving consistently low MSE for the

independent testing dataset consisted of 1 hidden layer with 15 neurons and a sigmoid activation function. A schematic of this structure is shown in Figure 2. Final training was performed by partitioning the training dataset into 80/20% training/testing datasets to allow direct comparison of actual and predicted heights following a standard approach. The permutation importance of each input variable was also evaluated to determine the appropriateness of each for the prediction of surface topography. Once the final ANN had been trained, optimization of photomasks for both the training and testing models was performed by iteratively updating the virtual photomask based on the predicted topography obtained with the ANN. At each new iteration, the AOM voltage was corrected based on the predicted error at each point and surrounding regions, limiting corrected values to the usable voltage range of 0-5V. As a result, the optimization algorithm converged to an optimal virtual photomask achieving constant error corresponding to the physical limits of the process setup.

Results and Discussion

Three-dimensional surface profiles acquired after fabrication of the training and testing models in upright and inverted positions with the experimental contrast curve are shown in Figure 3. The majority of features are relatively well-defined; however, vertical walls and regions of high aspect ratio in both models (Figure 1) are sloped due to sharp dose transitions. Tightly-packed features such as rectangular columns and pockets are poorly defined in some regions due to severe proximity effects. As noted previously, features with high aspect ratio, exceeding process limits, were deliberately included to provide a diverse and complete training dataset, as well as to assess the performance of the optimization algorithm up to process limits, providing benchmark data for future studies.

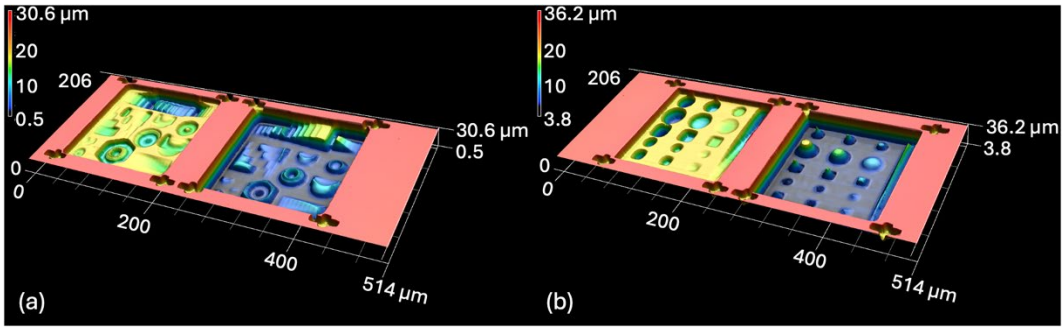

Fig. 3. Three-dimensional surface profiles acquired via LSCM for the (a) training and (b) testing models.

During hyperparameter optimization, data obtained from production and analysis of the training model were employed to train ANNs with 1-2 hidden layers, 1-25 neurons per layer and ReLU, sigmoid and tanh activation functions. The performance of each ANN was evaluated in terms of the MSE achieved for prediction of outcomes in the independent testing dataset obtained from production and analysis of the testing model. The results of hyperparameter optimization are presented in Figure 4, where values of MSE generally decreased with increasing number of neurons before reaching constant values of approximately 1-1.5 μm^2 with both one and two hidden layers. All activation functions performed similarly; however, sigmoid provided faster and more stable convergence to low values of MSE with both one and two hidden layers. In view of adopting the simplest configuration achieving acceptable results, the final ANN was selected as having 1 hidden layer with 15 neurons and a sigmoid activation function, yielding an MSE of 1.1 μm^2 for the independent testing dataset.

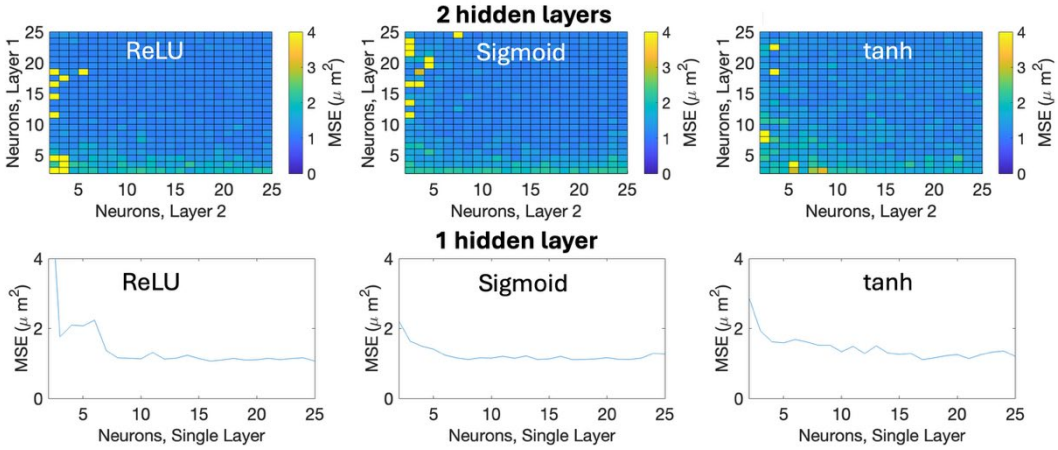

Fig. 4. ANN hyperparameter optimization: Mean squared error with independent testing dataset with 1-2 hidden layers, 1-25 neurons per hidden layer and ReLU, Sigmoid and tanh activation functions.

A comparison of the actual and predicted height in the testing dataset obtained by partitioning training data into 80/20% training/testing datasets is presented in Figure 5, together with the permutation importance of each input variable. It should be noted that negative height values were obtained due to the use of a positive photoresist. Good alignment between predicted and actual height values can be observed over the entire range, with only a limited number of points exhibiting large deviations. Values of mean permutation importance provided insight into the relative importance of each input variable, with greatest values observed for radial averages over the ranges 0 μm < r ≤ 2 μm, 2 μm < r ≤ 4 μm and 4 μm < r ≤ 8, followed by the AOM voltage at each point (r = 0 μm). Radial averages at distances greater than 8 μm were instead found to be of limited influence on the resulting surface topography. These outcomes nonetheless implied that the surface topography was greatly influenced by the incident laser energy dose within regions well beyond the focused laser spot of diameter of approximately 2 μm. The model therefore implicitly accounted for variations in converted initiator and photoresist solubility due to the absorbed energy dose [14,15]. This result confirms the importance of correctly accounting for sharp dose transitions and proximity effects during optimization of virtual photomasks for maskless grayscale laser lithography.

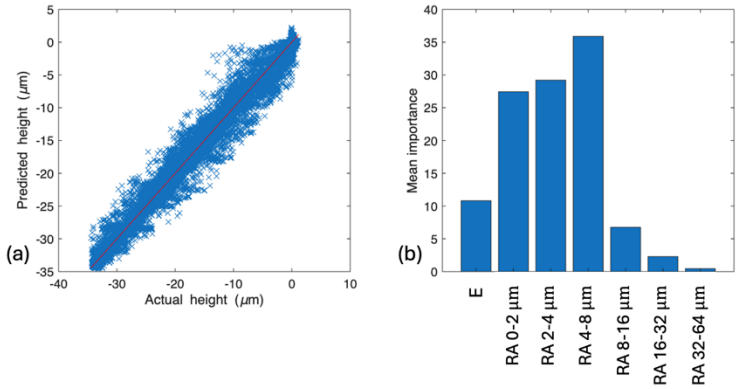

Fig. 5. (a) Comparison of actual height and predicted height with final trained ANN using testing dataset obtained by partitioning training data into 80/20% training/testing datasets. (b) Permutation importance of ANN input parameters with the same dataset. RA: Radial average range.

To gain physical insight into the effective performance of the trained ANN, Figure 6 presents virtual photomasks obtained for the training and testing models using the experimental contrast curve, together with the experimental and predicted surfaces in both cases. As noted previously, experimental surfaces exhibited clear deviations in topography from their respective models (Figure 1) due to sharp dose transitions at the edges of vertical elements and proximity effects taking place within closely-packed features such as rectangular columns and pockets in the upper region of the

training model. The surface topography predicted by the trained ANN generally reproduced the same effects accurately, with only minor differences between the experimental and predicted topography for both models at the center of complex features such as hexagonal prisms and truncated cones. Outcomes were nonetheless in line with physics-driven models [16] and considered acceptable at effectively accounting for the overall energy dose within all regions, including those subject to sharp dose transitions and proximity effects.

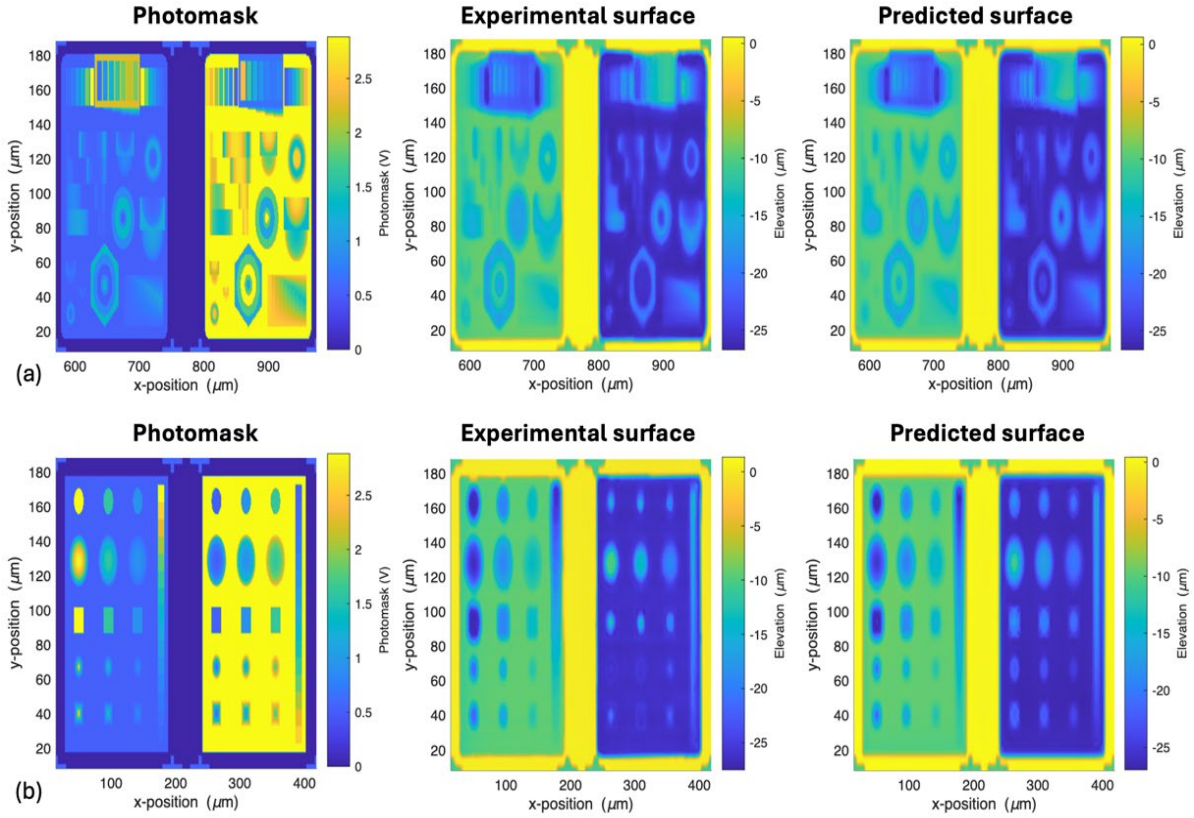

Fig. 6. Photomask and comparison of experimental and predicted surfaces with final trained ANN for virtual photomasks employed to produce (a) training model and (b) testing model.

Figure 7 provides a comparison of virtual photomasks obtained with the experimental contrast curve and the ML optimization algorithm for the training and testing models. Noting the differences in color scale, iterative correction of the virtual photomasks based on the predicted error led to profound changes in the spatial distribution of the AOM voltage for production of both objects. Large differences can be observed within base regions between features, particularly for upright structures on the righthand side where the AOM voltage increased from 2.7 V to approximately 4 V for optimized virtual photomasks, indicating macroscopic correction of the incident laser power. Furthermore, aggressive correction of sharp dose transitions and proximity effects can be observed at the edges of vertical elements and within closely-packed features. Sunken features, requiring higher laser power than adjacent base regions (due to the use of a positive photoresist), exhibited higher values around edges to compensate for the lower energy dose within adjacent base regions. The adjacent base regions, in turn, exhibit lower values to compensate for the higher energy dose at the edges of the features. For the same reason, raised features, requiring lower power than adjacent base regions, exhibit lower values around edges, while adjacent base regions exhibit higher values. Corrections were implemented up to process limits, exploiting the entire operating range of the AOM (0-5 V), suggesting that virtual photomask correction was implemented up the physical limits of the experimental setup. Conceptually, these outcomes are in line with proximity effect correction currently employed within physics-driven models [7].

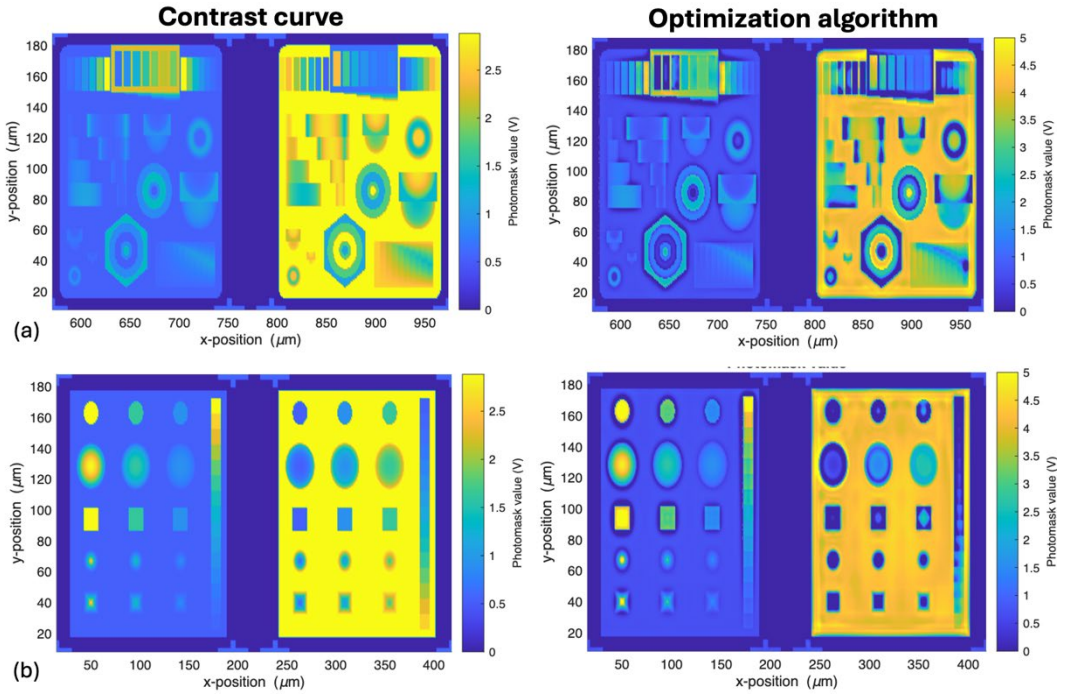

Fig. 7. Comparison of virtual photomasks obtained with experimental contrast curve and machine learning optimization algorithm for (a) training model and (b) testing model.

The surface topography of both training and testing models obtained with the experimental and optimized virtual photomasks are presented in Figure 8, together with the build error expressed in terms of the minimum Euclidean distance between the acquired and model surfaces. Results achieved with the experimental contrast curve exhibited evidence of proximity effects and sharp dose transitions due to non-uniform laser exposure and absorbed energy [6]. Clear differences in build accuracy can be observed between surfaces obtained with the experimental contrast curve and the optimization algorithm. In the former case, base regions between features exhibited constant error of up to 6-8 μm, while in the latter case, error was <2 μm within the same regions. This outcome suggests that the optimization algorithm successfully implemented macroscopic corrections within flat regions, automatically increasing the AOM voltage from 2.7 V to approximately 4 V within the same regions (Figure 7). While the same outcome could also be effectively achieved experimentally by correcting the contrast curve, this approach would require several trials, with additional time and resource consumption. Correction for sharp dose transitions and proximity effects was also observed at the edges of vertical elements and within rectangular columns and pockets on surfaces obtained with the optimization algorithm. Such a correction could not be achieved experimentally by correcting the contrast curve alone but would require adjustment of the virtual photomask within specific regions based on observed outcomes, with considerable time and resource consumption for each new build. Noting that some features exceeded process limits, residual error was nonetheless observed for surfaces produced with the optimization algorithm due to the physical limitations of the experimental setup, as well as LSCM measurement errors resulting from the higher aspect ratio obtained with the optimization algorithm. Such errors were most evident along the perimeter of upright models in each case. Use of the optimization algorithm led to a reduction in the mean Euclidean distance between the acquired and model surfaces from 2.8 μm to 1.3 μm for the training model and from 2.8 μm to 1.5 μm for the testing model compared to outcomes obtained with the experimental contrast curve, with maximum residuals generally <1 μm within central flat regions.

State-of-the-art physics-driven models and proximity effect correction techniques have been shown to reduce the root mean square (RMS) error from 0.629 μm to 0.189 μm for staircase structures [7], achieving maximum residuals of <0.5 μm within central flat regions [7]. While quantitative comparison with previous works is not possible due to differences in lithography setup, photoresist and target geometry, it should be noted that the step height of staircase structures in [7] was 2 μm and

the maximum overall height 10 µm, whereas the maximum step and overall heights in the present study were both 20 µm, leading to differences in terms of the physical limitations of the process itself. Few studies have proposed ML models for optimization of laser lithography, with quantitative comparison again not possible due to fundamental differences in experimental setup and measurement metrics. Nonetheless, state-of-the-art models and correction techniques have been shown to exhibit up to 10% deviation in surface elevation for relatively smooth structures without vertical walls [12]. Excluding such features, outcomes obtained in the present study generally provide similar or greater accuracy with a more straight-forward approach in terms of implementation.

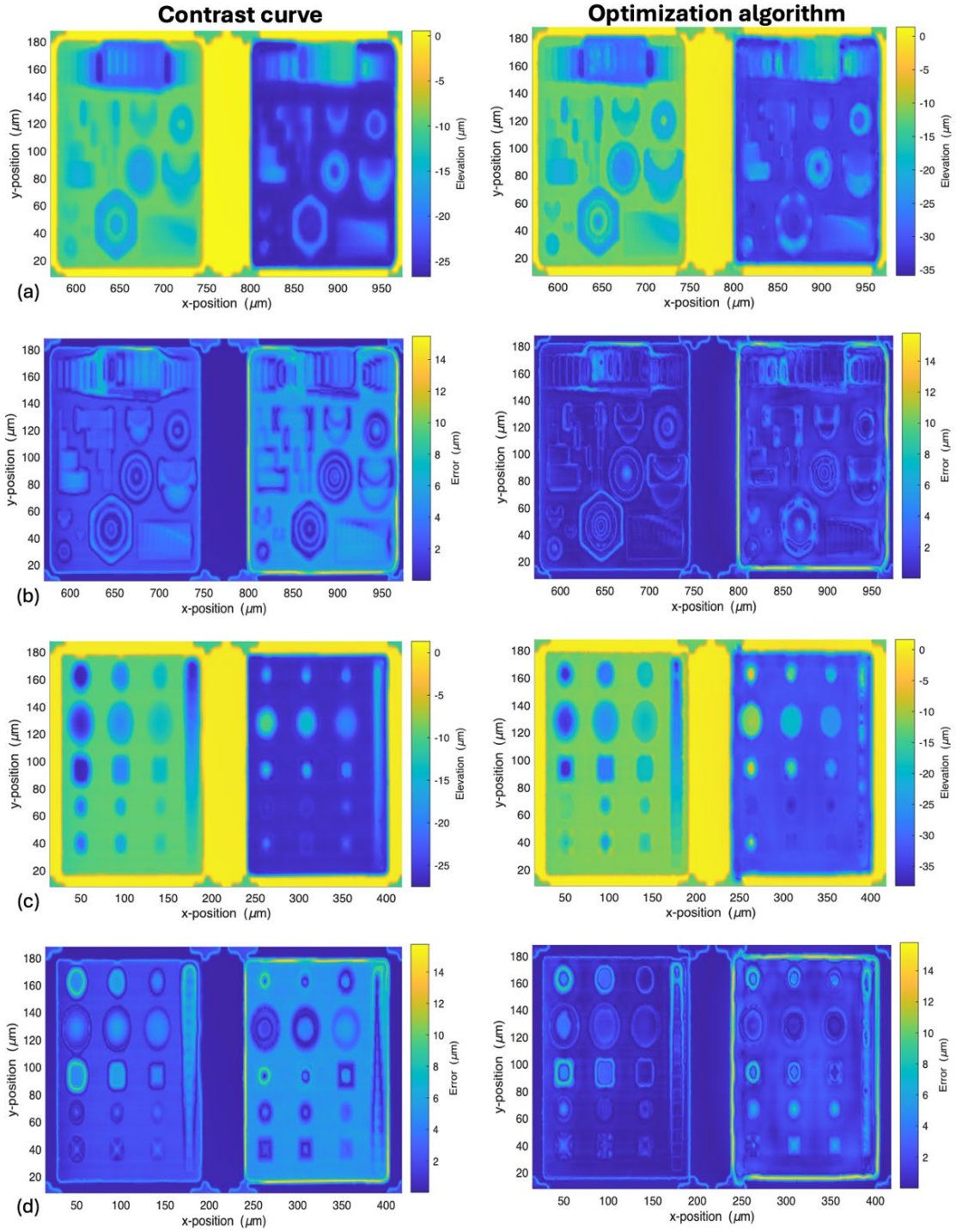

Fig. 8. Comparison of surface topography and error (minimum Euclidean distance) obtained with experimental contrast curve and ML optimization algorithm: (a) Training model, topography, (b) training model, error, (c) testing model, topography and (d) testing model, error.

Conclusion

A machine learning framework for predicting and optimizing surface topography in maskless grayscale laser lithography was developed, achieving reductions in build error from 2.8 μm to 1.3-1.5 μm compared to standard experimental approaches for defining virtual photomasks based on single-line experiments and an experimental contrast curve. Key aspects of the developed approach were the use of training and testing benchmarks for the generation of diverse and complete datasets for training ANNs, as well as grid hyperparameter optimization over an appropriate range. Once training was complete, a simple iterative optimization algorithm was implemented to correct for sharp dose transitions and proximity effects along edges and within closely-packed features. The main limitations of the developed approach include difficulties in obtaining high-quality training data within critical regions due to surface acquisition artefacts within steep/vertical regions, as well as the ability to deal with changes in photoresist, process parameters and experimental setup. However, these are essentially the limitations as for physics-driven models, while machine learning has the potential to reduce time and cost requirements associated with training and calibrating models for new configurations. While the effectiveness of the approach was demonstrated for a single maskless grayscale laser lithography setup and photoresist with constant hatch spacing and scanning speed, it could potentially be applied to a greater variety of cases by enlarging the training dataset and introducing more input variables to account for additional process parameters. Further investigation is therefore required into expanding and improving machine-learning models and optimization algorithms for maskless grayscale laser lithography.

Funding Declaration

The authors would like to kindly acknowledge financial support received from the OPTIMAL project (Horizon Europe, grant agreement No. 101057029).

References

[1] A. Grushina, Direct-write grayscale lithography, Advanced Optical Technologies 8 (2019) 163–169. https://doi.org/10.1515/aot-2019-0024.

[2] J. Dunkel, F. Wippermann, A. Brückner, A. Bräuer, A. Tünnermann, Laser lithographic approach to micro-optical freeform elements with extremely large sag heights, Opt. Express 20 (2012) 4763. https://doi.org/10.1364/OE.20.004763.

[3] R. He, S. Wang, G. Andrews, W. Shi, Y. Liu, Generation of Customizable Micro-wavy Pattern through Grayscale Direct Image Lithography, Sci Rep 6 (2016) 21621. https://doi.org/10.1038/srep21621.

[4] J. Loomis, D. Ratnayake, C. McKenna, K.M. Walsh, Grayscale lithography—automated mask generation for complex three-dimensional topography, J. Micro/Nanolith. MEMS MOEMS 15 (2016) 013511. https://doi.org/10.1117/1.JMM.15.1.013511.

[5] Z. Zhu, S. Wei, R. Liu, Z. Hong, Z. Zheng, Z. Fan, D. Ma, Freeform surface design for high-efficient LED low-beam headlamp lens, Optics Communications 477 (2020) 126269. https://doi.org/10.1016/j.optcom.2020.126269.

[6] J. Du, F. Gao, Y. Zhang, Y. Guo, C. Du, Z. Cui, Modified proximity function for OPC in laser direct writing, in: C.J. Progler (Ed.), Santa Clara, CA, 2000: p. 1047. https://doi.org/10.1117/12.388940.

[7] J. Erjawetz, D. Collé, G. Ekindorf, P. Heyl, D. Ritter, A. Reddy, H. Schift, Bend the curve – Shape optimization in laser grayscale direct write lithography using a single figure of merit, Micro and Nano Engineering 15 (2022) 100137. https://doi.org/10.1016/j.mne.2022.100137.

[8] T. Onanuga, M. Rumler, A. Erdmann, Simulation flow and model verification for laser direct-write lithography, J. Micro/Nanolith. MEMS MOEMS 16 (2017) 1. https://doi.org/10.1117/1.JMM.16.3.033511.

[9] D. Weichert, P. Link, A. Stoll, S. Rüping, S. Ihlenfeldt, S. Wrobel, A review of machine learning for the optimization of production processes, Int J Adv Manuf Technol 104 (2019) 1889–1902. https://doi.org/10.1007/s00170-019-03988-5.

[10] W. "Grace" Guo, Q. Tian, S. Guo, Y. Guo, A physics-driven deep learning model for process-porosity causal relationship and porosity prediction with interpretability in laser metal deposition, CIRP Annals 69 (2020) 205–208. https://doi.org/10.1016/j.cirp.2020.04.049.

[11] P. Kareem, Y. Shin, Synthesis of Lithography Test Patterns Using Machine Learning Model, IEEE Trans. Semicond. Manufact. 34 (2021) 49–57. https://doi.org/10.1109/TSM.2021.3052302.

[12] X. Sun, S. Yin, H. Jiang, W. Zhang, M. Gao, J. Du, C. Du, U-Net convolutional neural network-based modification method for precise fabrication of three-dimensional microstructures using laser direct writing lithography, Opt. Express 29 (2021) 6236. https://doi.org/10.1364/OE.416871.

[13] P. Minetola, F. Calignano, M. Galati, Comparing geometric tolerance capabilities of additive manufacturing systems for polymers, Additive Manufacturing 32 (2020) 101103. https://doi.org/10.1016/j.addma.2020.101103.

[14] F.H. Dill, W.P. Hornberger, P.S. Hauge, J.M. Shaw, Characterization of positive photoresist, IEEE Trans. Electron Devices 22 (1975) 445–452. https://doi.org/10.1109/T-ED.1975.18159.

[15] M.D. Smith, J.D. Byers, C.A. Mack, The lithographic impact of resist model parameters, in: J.L. Sturtevant (Ed.), Santa Clara, CA, 2004: p. 322. https://doi.org/10.1117/12.537581.

[16] N. Xie, D. Jones, G. Lopez, Comprehensive modeling of the lithographic errors in laser direct write, Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena 37 (2019) 061603. https://doi.org/10.1116/1.5122660.