Predicting Manufactured Shapes of a Projection Micro-Stereolithography Process via Convolutional Encoder-Decoder Networks

Conference Paper · August 2018

DOI: 10.1115/DETC2018-85458

CITATIONS
4

READS
291

6 authors, including:

Yusen He
University of Iowa
25 PUBLICATIONS **690** CITATIONS
SEE PROFILE

Fan Fei
University of Iowa
11 PUBLICATIONS **31** CITATIONS
SEE PROFILE

Wenbo Wang
Arizona State University
3 PUBLICATIONS **28** CITATIONS
SEE PROFILE

Xuan Song
University of Iowa
42 PUBLICATIONS **1,338** CITATIONS
SEE PROFILE

Proceedings of the ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference
IDETC/CIE 2018
August 26-29, 2018, Quebec City, Quebec, Canada

DETC2018-85458

PREDICTING MANUFACTURED SHAPES OF A PROJECTION MICRO-STEREOLITHOGRAPHY PROCESS VIA CONVOLUTIONAL ENCODER-DECODER NETWORKS

Yusen He
Innovative Design and Art Laboratory
Center for Computer-Aided Design
University of Iowa
Iowa City, Iowa 52242
Email: yusen-he@uiowa.edu

Fan Fei
Wenbo Wang
Xuan Song
Additive Manufacturing-Integrated
Product Realization Laboratory
Center for Computer-Aided Design
University of Iowa
Iowa City, Iowa 52242
Email: fan-fei@uiowa.edu
Email: wenbo-wang@uiowa.edu
Email: xuan-song@uiowa.edu

Zhiyu Sun
Stephen Baek*
Innovative Design and Art Laboratory
Center for Computer-Aided Design
University of Iowa
Iowa City, Iowa 52242
Email: zhiyu-sun@uiowa.edu
Email: stephen-baek@uiowa.edu

ABSTRACT

Projection micro-stereolithography (P-μSLA) processes have been widely utilized in three-dimensional (3D) digital fabrication. However, various uncertainties of a photopolymerization process often deteriorates the geometric accuracy of fabrication results. A predictive model that maps input shapes to actual outcomes in real-time would be immensely beneficial for designers and process engineers, permitting rapid design exploration through inexpensive trials-and-errors, such that optimal design parameters as well as optimal shape modification plan could be identified with only minimal waste of time, material, and labor. However, no computational model has ever succeeded in predicting such geometric inaccuracies to a reasonable precision. In this regard, we propose a novel idea of predicting output shapes from input projection patterns of a P-μSLA process via deep neural networks. To this end, a convolutional encoder-decoder network is proposed in this paper. The network takes a projection image as the input and returns a predicted shape after fabrication as the output. Cross-validation analyses showed the root-mean-square-error (RMSE) of 10.72 μm in average, indicating noticeable performance of the proposed convolutional encoder-decoder network.

INTRODUCTION

Projection micro-stereolithography (P-μSLA), also namely digital light processing (DLP), is a process that fabricates 3D digital objects (e.g. CAD model) with desired characteristics by cross-linking the photopolymer resin and converting corresponding region of the photopolymer into solid parts [1–4]. When producing a CAD model, it contains dynamic resin changes during the fabrication process [5,6]. Hence, the prediction of the printed outcome matters to the architecture and the mechanical behavior of the fabricated object.

In practice, parts produced through P-μSLA processes suffer from a significant shape disparity between the ideal CAD model and the actual production result. Especially for micro-scale printing, the final output of a P-μSLA process usually differs from its original CAD model due to the non-uniform distribution of light intensity over the printing area, the dynamics of the resin changes, the projector resolution, light source, and

*Address all correspondence to this author.

the unpredictable photopolymerization process. For instance, inputting a star-shaped CAD model, the sharp angles of the input model are all becoming smoother in the projected outcome. Meanwhile, the area of the projected object often decreases compared with its original input CAD model. These phenomenons are often caused by the weaker light intensities on the edge of the inner image and short curing time. For manufacturing engineers, these shape deformations of micro-scale printing may easily surpass the engineering tolerance range and result disqualified production. Therefore, learning these shape disparities during the P-μSLA process is crucial to the accuracy and efficiency of the fabrication process.

Currently, there is no usable methods to predict the outcome shape of the P-μSLA process. The solid structure of the projected outcome is impacted by many manufacturing parameters and geometric characteristics of the original CAD model [7]. The laws of physics and mechanism of the interplay of these factors are always too complicated in practice for accurate simulation.

To overcome such challenges, we proposes a novel method utilizing modern deep learning algorithms to construct a map between the input CAD model as well as the manufacturing parameters and the output of P-μSLA process. The mapped relationship would enable us to accurately predict the shape of a manufactured part based on its input geometry and process parameters. In practice, the manufacturing engineers is capable of simulating the resultant shape at the design phase without the time-consuming and labor intensive trail-and-errors, and eventually come up with the optimal process parameters or design modification to achieve the most accurate production.

In this research, the P-μSLA process is hypothesized as a mathematical transformation from an input CAD geometry to a processed output. Such transformation process can be modeled using a convolutional encoder-decoder network. With any given CAD geometry and other manufacturing parameters, the fabricated shape can be predicted and illustrated in advance to the designers and manufacturing engineers. In addition, modification of the original design as well as process parameters can be achieved to reduce the manufacturing errors.

The hypothesis mentioned above has been tested in this paper and a remarkable correlation between the input shape and the output shape is achieved. Using binary images, the input images and actual output parts are segmented into two classes of pixels: object and background. With a pixel-wise classification layer, the convolutional encoder-decoder network is capable to study the shape disparities between the input and output objects. By validating the prediction results with the ground truth from fabricated output, the prediction results were highly accurate with on average 10.72 μm root-mean-square-error and 6.42 μm mean-absolute-error. Hence, it confirms the superior performance of the convolutional encoder-decoder network on fabricated results simulation and inverse process optimization.

BACKGROUND

In this research, the goal is to develop a deep learning framework to learn the sophisticated manufacturing process of projection micro-stereolithography (P-μSLA). Instead of using complicated formulas or equations to illustrate the process, a deep learning framework which performs as a black-box function with an input CAD model and a predicted outcome of the P-μSLA process enables us to achieve this task efficiently.

Manufacturing Prediction Method

Projection micro-stereolithography (P-μSLA) is a photopolymerization based additive manufacturing (AM) method that was first commercialized among all other AM processes. In P-μSLA process, a 2D image with micro-scale resolution is illuminated onto the photopolymer to create a solidified pattern for each layer of the 3D model. The major constraint of P-μSLA process is its shape deformation on the output pattern. Methods to predict and improve the results of P-μSLA process have been vigorously studied and can be classified into two categories: finite element analysis (FEA) approaches and image-based solidification models.

Applying FEA approaches, the deformation caused by the manufacturing uncertainties can be simulated for the P-μSLA process. Bugeda et al [8] used the FEA method to simulate the P-μSLA process that the volumetric shrinkage has been identified as the major reason for distortion. Chambers et al [9] considered the cured shrinkage as the main reason for the shape deformation of P-μSLA and formulated the shrinkage into FEA model to simulate such deformation. Hur and Youn [10] considered the thermal cooling effects in P-μSLA and used it to predict the deformation in fabricated parts. Tanaka et al [11] investigated the distortion in P-μSLA by incorporating polymerization shrinkage due to thermal effects in the FEA model. All methods above has achieved promising results in specific cases of P-μSLA fabrication. However, all these method are experiment-based procedures which is time consuming and labor expensive. Meanwhile, the FEA method is limited on the aspect of its robustness to solve the deformation problem for various fabrication tasks, and its high cost for computational analysis.

On the other hand, image-based projection method has been widely proposed to solve the shape deformation of P-μSLA process. To predict and improve the fabrication results of P-μSLA, Sun et al [7] introduced a solidification model to predict the fabrication results for line patterns. Jariwala et al [12] reported a model to predict solidified layer thickness based on experimental observation. Zhou et al [13] introduced a strategy to improve fabrication results by optimizing the gray scale of an illuminated 2D image. Overall, these methods have obtained higher prediction accuracy than experiment based approaches and enabled the designers to modify the input CAD model before the actual fabrication. However, their methods use case specific arbitrarily de-

signed 2D patterns in the fabrication process and the robustness of those approaches on massive 2D patterns is limited.

To mitigate such limitations, Kang et al [14] constructed a pixel based solidification model to increase the prediction accuracy of P-μSLA process. The intensity distribution of a single pixel in the illuminated image has been investigated to develop the solidification model. The intensity profile of an illuminated image was then expressed in terms of the mathematical model to estimate the solidified profile. The pixel-based solidification made a significant contribution to the image-based approaches but its limitations are related to its unsatisfactory accuracy around the corners and shape edges in the micro-scale fabrication cases.

Convolutional Neural Networks

In recent years, artificial intelligence (AI) has made an unprecedented impact in the field of manufacturing. As one of the mostly used AI algorithm, the neural networks (NNs) which simulates the brain computation enables computers to recognize patterns by constructing a black-box function with an input layer, an output layer, and multiple hidden layers. In practical applications, the NNs try to find and generalize underlying rules and features of data, from observations on a training dataset, instead of using hard-coded or hand-crafted rules that are not robust at all.

Convolutional neural network (CNN) is an improved neural network with shift and translational invariance that is capable to recognize patterns directly from pixel images by incorporating both feature extraction and classification [15]. Proposed by Yan Lecun [16], a typical CNN involves four types of layers: convolutional, activation, pooling and fully connected (or dense) layers [17]. A convolutional layer is characterized by sparse local connectivity and weight sharing. The neurons in the convolutional layer only connects with a small local area in the input image which resembles the receptive field in the human visual system. Different neurons respond to different local areas of input image while the overlap of each other enables the network to obtain a better representation of the image. An activation layer always follows the convolutional layer to receive more complex properties of the input image. Pooling layers subsample the previous layer by aggregating small rectangular subsets of values and replace the input values with maximum or average value respectively for the purpose of sensitivity reduction from output to small input shifts. Last, the fully connected layers are put in place to produce the classification results [18]. The training procedure of CNNs is to minimize the loss function using stochastic gradient descent (SGD) which is similar to classical neural network (NN) [17]. The CNNs have been successfully deployed in many practical applications including image classification, image segmentation, object correspondence matching, and others [19–22].

Recently, fully convolutional network (FCN) [23] has attracted great attention in the field of conmputer vision for sematic segmentation. The FCN that takes input of arbitrary size and produces correspondingly-sized output with efficient inference and learning enables the network to classify the image pixels into different classes with sufficient accuracy. With any input image, the network is capable to segment the objects based on the initial pixel intensity values and their spatial location information.

NETWORK DESIGN

In this section, inspired by the success of SegNet for semantic segmentation [24], the architecture of the proposed deep convolutional encoder-decoder network that consists an encoder network and a decoder network for shape deformation prediction is presented. The encoder network is composed of convolutional layers, batch normalization layers, relu layers, and max-pooling layers [25]. The decoder network contains unpooling layers, transposed convolutional layers, batch normalization layers, relu layers. The end of the network includes a softmax layer and a pixel classification layer that classify the intensity values of each pixel of the predicted image into 0 and 1.

In this study, the structure of the network contains four convolutional layers, four batch normalization layers, four relu layers, and two max-pooling layers in the encoder network and four transposed convolutional layers, four batch normalization layers, four relu layers, and two unpooling layers in the decoder network. Also, a softmax layer and a pixel classification layer are stacked at the end of the decoder network. In addition, we added a skip connection between the first max-pooling layer and the last unpooling layer to reduce the initial information loss caused by multiple layers of convolution. The architecture of the proposed convolutional encoder-decoder network has been illustrated in Fig. 1. Overall, a well-trained convolution encoder-decoder network enables us to extract the features of shape deformations and predict the final printed outcome of the P-μSLA process.

DATASET

The P-μSLA system and its schematic diagram used in this paper are shown in Fig. 2 and Fig. 3. It contains two parts: the projection system and the stage system. The projection system includes a customized light engine with an LED light source of 405 nm wavelength and a digital micromirror device (DMD) with a resolution of 1280×800 pixels working as a dynamic mask to control the shape of light. It also has a CCD camera to observe the projected images for each layers. When all the pixels in the light engine are on, the projection system can provide UV light to polymerize the resin with an area of 2.00×1.25 mm. The stage system contains linear stages in three axes, with an absolute accuracy of 10 μm. In this article, only 2D patterns are used for manufacturing prediction. As a result, the movement

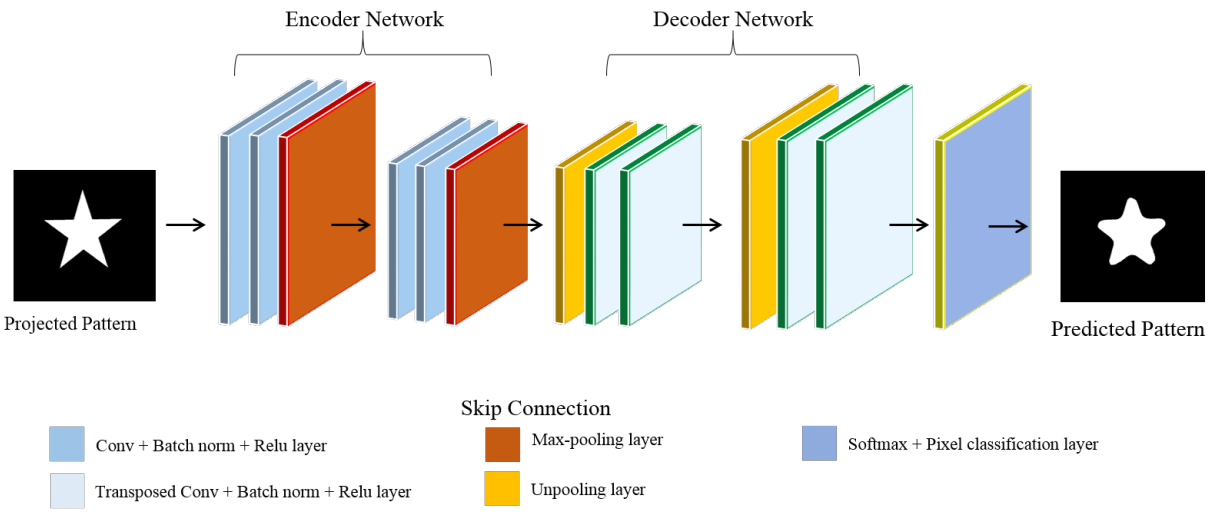

FIGURE 1. THE PROPOSED ARCHITECTURE OF CONVOLUTIONAL ENCODER DECODER NETWORK.

of the stage system is not required and the 2D patterns can be fabricated by projecting groups of required images one by one. Since the maximum area the UV light in this P-μSLA system is $2.00 \times 1.25\ mm^2$, 100 images with sizes less than that values are utilized to fabricate the 2D patterns. The material used to fabricate those 2D patterns is photosensitive resin (Formlabs, Inc.) with die (Sudan I, 0.5 wt.%). Curing time for each patterns is the same, which is 16 seconds. The projected patterns and their corresponding fabricated results are shown in Fig. 4.

Training and Validation

In this study, we use the 100 widely used patterns for P-μSLA process to train the convolutional encoder-decoder network. All patterns have been transformed into binary images in the size of 128×128 and only include the object and the background. The object in the image has pixel intensity value 1 and the background has the intensity value 0. Since the dataset is small, we applied image rotation and translation for both input and output images correspondingly. The rotation is conducted for every 5° until 180° and the translation for each rotation is five pixels per rotation. Hence, we obtained approximately 18500 images for both input and output images as the experimental dataset. The images generated from the first 70 patterns are used as the training and testing dataset and the images generated from the rest 30 patterns are used for prediction. During the training process, 10-fold cross validation is selected as the training strategy for the network in which of all training images are .

To assess the network performance, we apply two validation methods to validate the predicted output images with the actual printed outcome from P-μSLA process as the ground truth. Since the task is to measure the shape difference between the two images, the first validation method is to compute the Intersection over Union (IoU) [26] between the two images. The IoU which directly computes the percentages of pixels that are correctly predicted as object and background can represent the correctness of predictive shape deformation.

Another method to quantitatively measure the network performance is to compare the extracted boundaries from both the prediction and the groundtruth. In this study, the image erosion algorithm has been applied to extract the boundaries of the objects [27]. To compare the shape differences, we firstly applies k nearest neighbor (kNN) algorithm to construct a mapping for each pair of correspondence points in two images. The distance between two corresponding points are measured in μm as the standard length unit in micro-scale printing. The following criterias are utilized to measure the shape disparities based on the boundary points:

RMSE is the root-mean-square-error which indicate the square root of the average of the squared Euclidien distance between the points in two point sets. In comparison with MAE, it is more likely to be impacted by longer distances rather than shorter ones.

MAE is the mean-absolute-error which evaluates the arithmetic average of Euclidien distance between the points in two point sets.

$AE_{0.95}$ is the 95 percentile of absolute errors which indicates the ceiling of the majority Euclidien distances between two point sets.

MaxAE is the maximum-absolute-error which computes

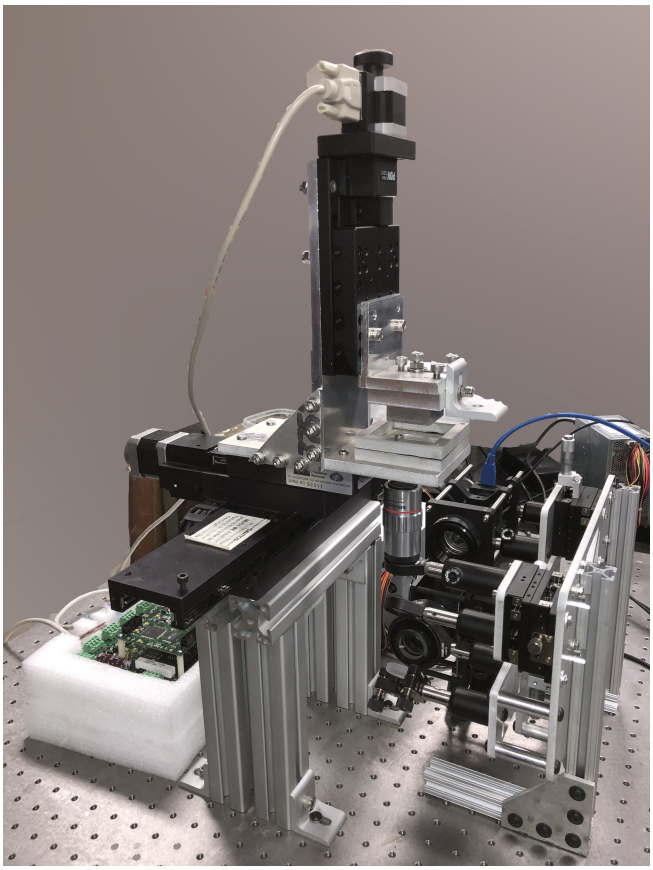

FIGURE 2. PICTURE OF THE P-μSLA SYSTEM

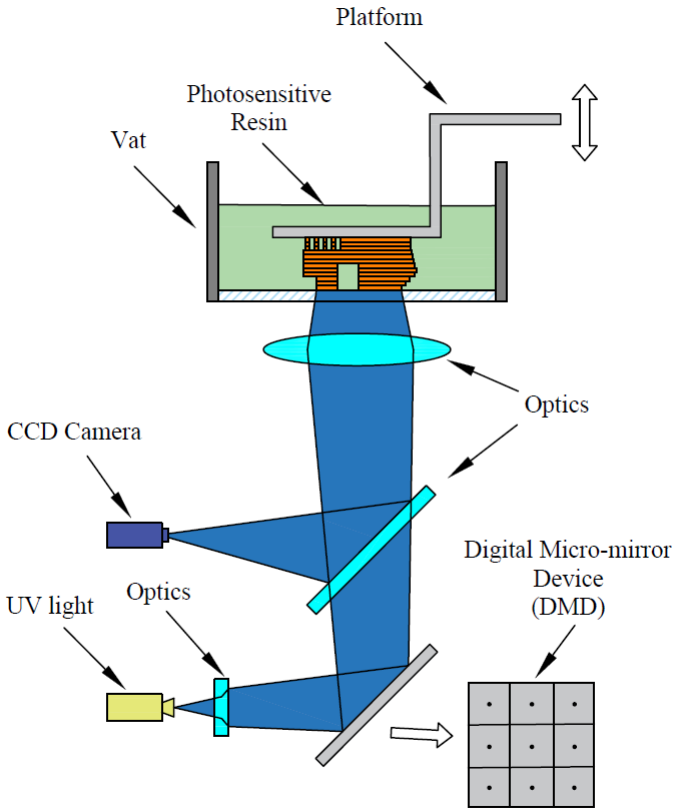

FIGURE 3. SCHEMATIC DIAGRAM OF THE P-μSLA SYSTEM

the maximum Euclidien distances between points. It measures the maximum distance between the corresponding points in two point sets.

BHD is the bidirectional 95-percentile Hausdorff distance which measures distance between two point sets. It identifies the point that is the 95th percentile farthest from any point of another point set. It measures the distance from this point to its nearest neighbor in another point set. The BDH is very sensitive to outliers points that result long distances.

RESULT AND DISCUSSION

From the 30 patterns we used to predict the shape deformations through P-μSLA process, pairs of images including the projected patterns and the groundtruths which are the actual printed patterns are randomly selected to validate the performance of the proposed approach. In this research, due to the limit of space, we randomly select 8 patterns from the reserved 30 to validate the effectiveness of the convolutional encoder decoder network.

Fig. 4 illustrats the result of such analysis in detail. The actual printed outcomes (groundtruth) and the predicted outcomes (prediction) are all compared correspondingly. By comparing the predicted shapes with the actual printed shapes, it is obvious that the convolutional encoder decoder network is capable to study and simulate the shape deformation caused by the uncertainties within the P-μSLA process. Meanwhile, the randomly sampled points from the boundaries of the two images are overlapped and illustrated in Fig. 4. It is significant that the majority of the boundary points from the predicted patterns overlaid on the points sampled from the boundaries of groudtruth image which indicates the accurate shape prediction from the network.

To validate the network performance quantitatively, the IoU which measures the percentage of correctly classified pixels which represent the ”object” are computed and illustrated in Table 1. Among them, the majority of the IoU values are greater than 0.90 which indicate that over 91% of the pixels are accurately classified as ”object” through the trained network.

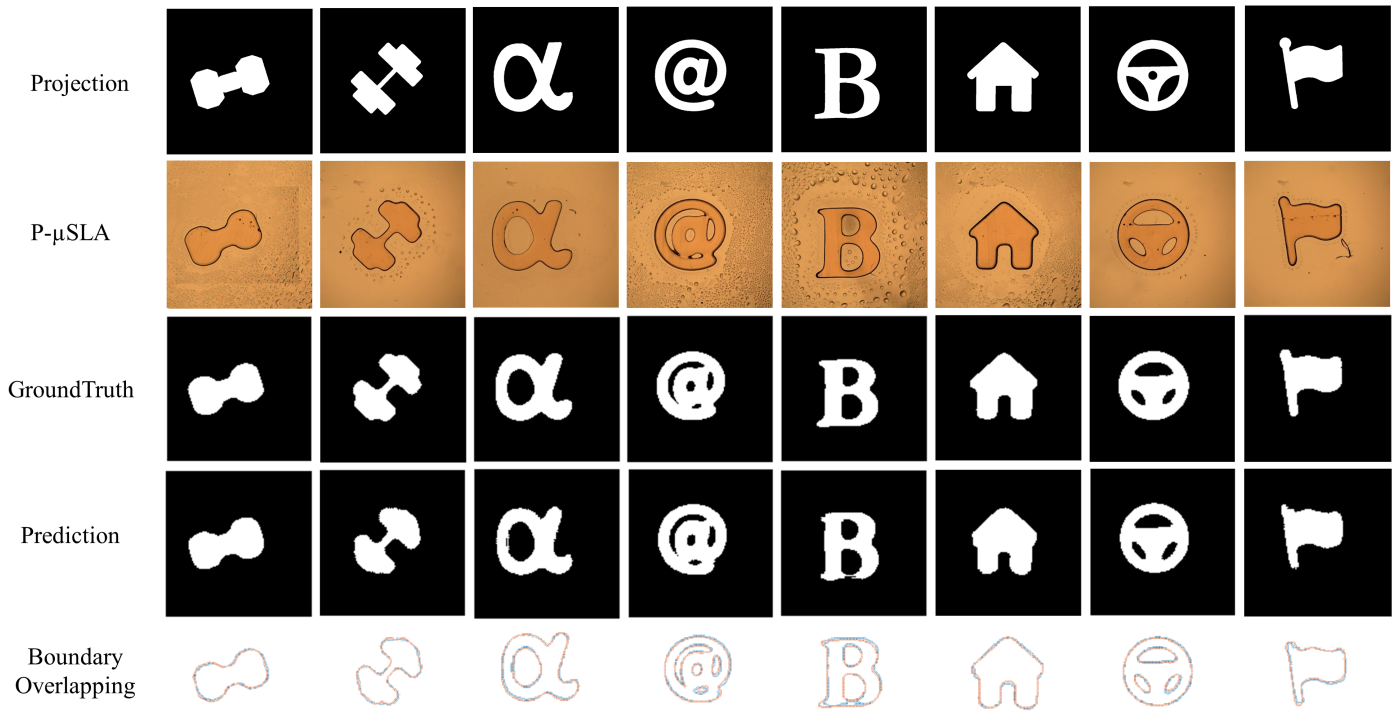

FIGURE 4. COMPARISON WITH PREDICTION AND GROUNDTRUTH OF VALIDATION PATTERNS.

TABLE 1. INTERSECTION OVER UNION OF THE RESTULS PRESENTED IN FIGURE 5.

Patterns	Dumbbell1	Dumbbell2	Alpha	At	B	House	Icon	Flag
IoU	0.931	0.916	0.906	0.915	0.851	0.943	0.911	0.914

Meanwhile, the distance between the boundary points of predicted outcome and groundtruth are computed and presented in Table 2. The average RMSE is 10.72 μm and the average MAE is 6.42 μm for all patterns respectively. In addition, the large shape disparities measured by $AE_{0.95}$, MaxAE, and BHD all reflect that only small portion of outliers exist and more than 95% of the distances between the prediction and groundtruth boundaries are within 25 μm. Hence, considering the scale of P-μSLA printing area which is $2.00 \times 1.25\ mm^2$, the prediction error is absolutely within the engineering tolerance which validated the accuracy of the predicted outcome.

CONCLUSION

This paper proposed a novel method using the convolutional encoder-decoder network to simulate the shape deformation of the P-μSLA process. With the shape disparities caused by process uncertainties, the shape deformation has been successfully predicted by using the projected patterns as the input and the printed patterns as the output. Experimental results indicated that the network has successfully studied the shape deformation of the P-μSLA process. Hence, the proposed method can be beneficial for manufacturing engineers and designers to further optimize the P-μSLA process.

While only a few applications of the method is illustrated with respect to the P-μSLA process, its simplicity and effectiveness will lead to its broad applications in multiple manufacturing

TABLE 2. BOUNDARY DISTANCE BETWEEN GROUNDTRUTH AND PREDICTED OUTCOME COMPUTED VIA RMSE, MAE, $AE_{0.95}$, MAXAE, AND BIDIRECTIONAL 95% HAUSDORFF DISTANCE IN MICROMETERS.

Patterns	RMSE	MAE	$AE_{0.95}$	MaxAE	BHD
Dumbbell1	7.883	4.208	16.563	23.423	16.563
Dumbbell2	10.351	6.039	23.389	25.300	23.423
Alpha	13.256	8.805	25.300	30.239	25.300
At	13.175	8.698	23.423	27.047	25.300
B	11.893	7.394	23.419	38.250	23.423
House	8.063	4.285	13.523	31.715	13.827
Icon	11.434	6.855	23.423	28.688	23.423
Flag	9.711	5.107	23.611	31.715	23.611

fields. For example, the power of deep neural networks would enable engineers to understand the intrinsic of many complex manufacturing process. These board applications would reveal exciting potential technical revolutions in the near future.

REFERENCES

[1] Song, X., Chen, Y., Lee, T. W., Wu, S., and Cheng, L., 2015. "Ceramic fabrication using mask-image-projection-based stereolithography integrated with tape-casting". *Journal of Manufacturing Processes,* ***20***, pp. 456–464.

[2] He, L., and Song, X., 2018. "Supportability of a high-yield-stress slurry in a new stereolithography-based ceramic fabrication process". *JOM,* ***70***(3), pp. 407–412.

[3] Song, X., Chen, Z., Lei, L., Shung, K., Zhou, Q., and Chen, Y., 2017. "Piezoelectric component fabrication using projection-based stereolithography of barium titanate ceramic suspensions". *Rapid Prototyping Journal,* ***23***(1), pp. 44–53.

[4] Chen, Z., Song, X., Lei, L., Chen, X., Fei, C., Chiu, C. T., Qian, X., Ma, T., Yang, Y., Shung, K., et al., 2016. "3d printing of piezoelectric element for energy focusing and ultrasonic sensing". *Nano Energy,* ***27***, pp. 78–86.

[5] Pan, Y., Zhao, X., Zhou, C., and Chen, Y., 2012. "Smooth surface fabrication in mask projection based stereolithography". *Journal of Manufacturing Processes,* ***14***(4), pp. 460–470.

[6] Song, X., and Chen, Y., 2012. "Joint design for 3-d printing non-assembly mechanisms". In ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 619–631.

[7] Sun, C., Fang, N., Wu, D., and Zhang, X., 2005. "Projection micro-stereolithography using digital micro-mirror dynamic mask". *Sensors and Actuators A: Physical,* ***121***(1), pp. 113–120.

[8] Bugeda, G., Cervera, M., Lombera, G., and Onate, E., 1995. "Numerical analysis of stereolithography processes using the finite element method". *Rapid Prototyping Journal,* ***1***(2), pp. 13–23.

[9] Chambers, R., Guess, T., and Hinnerichs, T., 1995. A phenomenological finite element model of part building in the stereolithography process. Tech. rep., Sandia National Labs., Albuquerque, NM (United States).

[10] Hur, S., and Youn, J., 1998. "Prediction of the deformation in stereolithography products based on elastic thermal shrinkage". *Polym.-Plast. Technol. Eng.,* ***37***(4), pp. 539–563.

[11] Tanaka, F., 2000. "Thermal distortion analysis during photopolymerization in stereolithography". In Proc. of the Eighth International Conference on Rapid Prototyping Tokyo, 2000, pp. 87–92.

[12] Jariwala, A. S., Ding, F., Zhao, X., and Rosen, D. W., 2009. "A process planning method for thin film mask projection micro-stereolithography". In ASME 2009 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 685–694.

[13] Zhou, C., and Chen, Y., 2009. "Calibrating large-area mask projection stereolithography for its accuracy and resolution improvements". In Proceedings of Solid Freeform Fabrication Symposium, Austin, Texas.

[14] Kang, H.-W., Park, J. H., and Cho, D.-W., 2012. "A pixel based solidification model for projection based stereolithography technology". *Sensors and Actuators A: Physical,* ***178***, pp. 223–229.

[15] Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2012. "Imagenet classification with deep convolutional neural networks". In Advances in neural information processing systems, pp. 1097–1105.

[16] LeCun, Y., Bengio, Y., et al., 1995. "Convolutional networks for images, speech, and time series". *The handbook of brain theory and neural networks,* ***3361***(10), p. 1995.

[17] Du, B., Xiong, W., Wu, J., Zhang, L., Zhang, L., and Tao, D., 2017. "Stacked convolutional denoising auto-encoders for feature representation". *IEEE transactions on cybernetics,* ***47***(4), pp. 1017–1027.

[18] Masci, J., Meier, U., Cireşan, D., and Schmidhuber, J., 2011. "Stacked convolutional auto-encoders for hierarchical feature extraction". In International Conference on Artificial Neural Networks, Springer, pp. 52–59.

[19] Bruna, J., and Mallat, S., 2013. "Invariant scattering con-

volution networks". *IEEE transactions on pattern analysis and machine intelligence,* ***35***(8), pp. 1872–1886.

[20] Ciregan, D., Meier, U., and Schmidhuber, J., 2012. "Multi-column deep neural networks for image classification". In Computer vision and pattern recognition (CVPR), 2012 IEEE conference on, IEEE, pp. 3642–3649.

[21] Sun, Z., He, Y., Gritsenko, A., Lendasse, A., and Baek, S., 2017. "Deep spectral descriptors: Learning the point-wise correspondence metric via siamese deep neural networks". *arXiv preprint arXiv:1710.06368*.

[22] Ouyang, T., He, Y., Li, H., Sun, Z., and Baek, S., 2017. "A deep learning framework for short-term power load forecasting". *arXiv preprint arXiv:1711.11519*.

[23] Long, J., Shelhamer, E., and Darrell, T., 2015. "Fully convolutional networks for semantic segmentation". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431–3440.

[24] Ronneberger, O., Fischer, P., and Brox, T., 2015. "U-net: Convolutional networks for biomedical image segmentation". In International Conference on Medical image computing and computer-assisted intervention, Springer, pp. 234–241.

[25] Badrinarayanan, V., Kendall, A., and Cipolla, R., 2017. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation". *IEEE transactions on pattern analysis and machine intelligence,* ***39***(12), pp. 2481–2495.

[26] Ahmed, F., Tarlow, D., and Batra, D., 2015. "Optimizing expected intersection-over-union with candidate-constrained crfs". In Proceedings of the IEEE International Conference on Computer Vision, pp. 1850–1858.

[27] Haralick, R. M., Sternberg, S. R., and Zhuang, X., 1987. "Image analysis using mathematical morphology". *IEEE transactions on pattern analysis and machine intelligence*(4), pp. 532–550.