NANO · MICRO small
www.small-journal.com

A Data-Driven Approach to Complex Voxel Predictions in Grayscale Digital Light Processing Additive Manufacturing Using U-Nets and Generative Adversarial Networks

Jason P. Killgore, Thomas J. Kolibaba, Benjamin W. Caplins, Callie I. Higgins, and Jacob D. Rezac*

Data-driven U-net machine learning (ML) models, including the pix2pix conditional generative adversarial network (cGAN), are shown to predict 3D printed voxel geometry in digital light processing (DLP) additive manufacturing. A confocal microscopy-based workflow allows for the high-throughput acquisition of data on thousands of voxel interactions arising from randomly gray-scaled digital photomasks. Validation between prints and predictions shows accurate predictions with sub-pixel scale resolution. The trained cGAN performs virtual DLP experiments such as feature size-dependent cure depth, anti-aliasing, and sub-pixel geometry control. The pix2pix model is also applicable to larger masks than it is trained on. To this end, the model can qualitatively inform layer-scale and voxel-scale print failures in real 3D-printed parts. Overall, machine learning models and the data-driven methodology, exemplified by U-nets and cGANs, show considerable promise for predicting and correcting photomasks to achieve increased precision in DLP additive manufacturing.

1. Introduction

Additive manufacturing is leading a revolution in fabricating precision, architected materials with extensive customization. Variants of vat photopolymerization (VP) additive manufacturing, including digital light processing (DLP),[1] liquid crystal display VP,[2] and continuous liquid interface production [3] all leverage advanced digital masks to locally cure layers of liquid resin sequentially into a solid 3D part, an example of which is shown in **Figure 1**a. The availability of digital masks with 4K (e.g., 3840 × 2160 pixels) or 8K (e.g., 7680 × 4320 pixels) resolution now enables individuals to control millions of simultaneous, voxel-scale reactions in a single layer or billions to trillions of reactions in a single part. Despite the immense fabrication power of such light engines, numerous interactions between adjacent voxels can undermine the theoretical reaction control and the resultant part fidelity.[4–6]

Early commercial printers employed binary black (minimum light output) and white (maximum light output) digital masks.[2] Binary masks limit the ability of the printer to adjust exposure for complex voxel interactions. Grayscale masking utilizing the full bit depth of the light engine can extend the utility of vat photopolymerization.[7–9] Lower light intensity of the gray pixels compared to white pixels reduces the reaction rate, which correlates to voxel size, for the same exposure duration. The use of grayscale has thus become a powerful tuning control for optimal mask design. Grayscaling at the edge of local positive or negative features, referred to as pixel blending or antialiasing, can eliminate stairstep effects (i.e., pixelation artifacts) and enhance feature accuracy.[7,10,11] Further uses of grayscale seek to limit the degree of conversion of the photopolymerization reaction to introduce controlled mechanical heterogeneity in final parts[12,13] and can reduce stress concentration.[14] Such heterogeneity can be amplified by means of dual-cure resin systems,[15] wherein a secondary cure step can result in an elastic modulus range spanning orders of magnitude. Thus far, robust methods of optimizing mask design for geometric precision and mechanical control are still in their infancy.[16]

The prediction of voxel scale geometry given an arbitrary VP exposure mask is complicated by the numerous underlying physical-chemical phenomena that govern local photopolymerization at length scales equal to or less than the light engine pixel pitch.[4] The complexity of the reaction starts with the cumulative overlapping nature of light emission from adjacent mask pixels, meaning that the tails of light from one pixel can interact with the light from their neighbors.[6,7,17–20] Additionally, numerous species, including radicals, oxygen, oligomers, and initiators, can diffuse into or out of the illuminated pixels during exposure.[6] Diffusion length scales can approach or exceed the pixel pitch depending on resin and exposure characteristics.[21,22] In addition to mass transfer, the reactions are generally exothermic, and

J. P. Killgore, T. J. Kolibaba, B. W. Caplins, C. I. Higgins
Applied Chemicals and Materials Division
National Institue of Standards and Technology
Boulder, CO 80305, USA
E-mail: jason.killgore@nist.gov

J. D. Rezac
RF Technology Division
National Institue of Standards and Technology
Boulder, CO 80305, USA

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/smll.202301987

DOI: 10.1002/smll.202301987

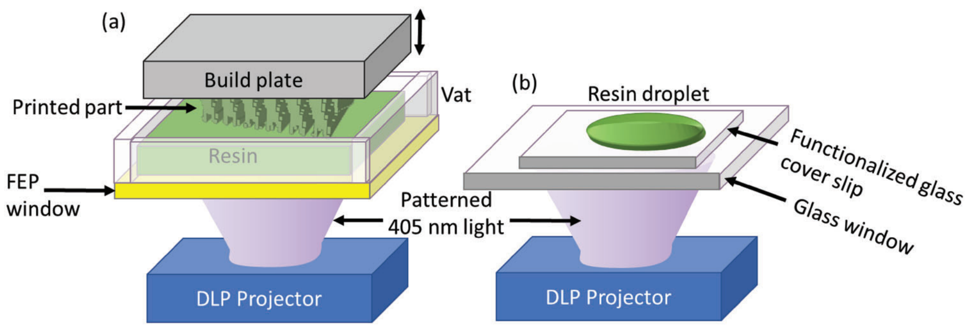

Figure 1. a) A typical bottom-up DLP vat-photopolymerization setup showing DLP projector, fluorinated ethylene propylene (FEP) window, vat, resin, part and build plate. For single-layer model training, the vat is replaced by a functionalized glass slide and resin droplet, as shown in (b).

thus the local generation of heat affects neighboring reaction rates.[4] To further complicate modeling, many of the underlying resin properties are inherently conversion dependent. For example, diffusivities can change dramatically between monomer, oligomer, just-gelled network, and a fully converted network, possibly above its glass transition temperature. Numerous recent efforts have pushed the ability to include the above phenomena to predict VP printing.[4,6,20–22] Because of the complexity of the reaction, often dozens of resin and light engine properties are needed to model a print, requiring significant experimental equipment and time. The multiphysics models provide immense insight into the cure reactions, but there is still a need for empirical models that provide accurate prediction without being subject to the dearth of material property information and the complexity of accurately accounting for all the underlying phenomena. Furthermore, as VP pushes hyperscale models with micron-scale feature sizes and meter-scale part sizes, the capacity of multiphysics models will be pushed to their limit.

Because VP can produce millions to trillions of reactions and interactions in a single part, the method is particularly well suited to big data, machine learning (ML), and artificial intelligence analysis approaches. Despite the intrinsic suitability of VP to ML modeling, only a few studies have applied ML tools to VP.[23–26] This adoption lags considerably behind applications of ML in metal AM.[23] You et al. applied machine learning to predict the lateral extent of polymerization of larger photopatterns[26]; however, to date, no VP ML studies have sought to quantitatively predict 3D voxel and sub-voxel scale geometry. The lack of development in this space is attributed to a lack of demonstrated VP characterization tools capable of generating the big, high-resolution data sets necessary for model training. A single characterization tool capable of measuring high-throughput, high-resolution voxel geometry could potentially yield data-driven models with predictive power at a fraction of the experimental burden of a multiphysics-based approach.

Characterizing voxel scale interactions requires microscopic tools to visualize and probe the part at sub-voxel scale (e.g., in this work, $\ll 80$ µm) resolution. Tools such as atomic force microscopy provide a high-resolution characterization of voxel morphology and properties,[5,21,22,27] but they sacrifice field of view and thus throughput. X-ray computed tomography provides robust 3D characterization of fully printed parts and has seen broad adoption in metal additive manufacturing.[28,29] However, the method has not yet been used to register sub-voxel scale resolution in VP parts. Recently, laser scanning confocal microscopy (LSCM) was used to characterize VP part surfaces in establishing relationships between light engine heterogeneity and printed voxel structure.[30] LSCM can image voxel geometry at a part surface with single-micron lateral resolution and a few-nanometer height resolution. LSCM also provides relatively high throughput, with commercial instruments capable of imaging a single field of view ($\approx$0.01 to 1 mm^2 depending on the microscope objective) in seconds, then automatically stitching together hundreds of fields of view.

With access to large, high-veracity datasets, several ML models are available to establish the mapping between input mask illumination and output voxel geometry. Given the representation of both input mask and output voxel geometry as 2D arrays or images, neural networks provide a promising framework to predict the mapping. Neural networks are a type of machine learning algorithm modeled after the structure and function of the human brain. Convolutional neural networks (ConvNets or CNNs) are a specialized type of neural network used for image and video recognition.[31] They are specifically designed to process data through multiple layers of neurons, using convolutional and pooling operations to identify patterns and features in images. The neurons in a layer are defined by weights and biases and are separated by non-linear activation functions that optionally allow information to pass between layers. Neural networks are trained with a backpropagation algorithm that seeks to minimize a loss function by adjusting weights and biases based on how much each node contributes to the error in the loss function (i.e., a measure of how well the predicted outputs match the ground truth targets). Training progresses for cycles through the entire dataset, referred to as epochs until a desired loss is achieved.

As shown in **Figure 2**, by building the CNN with a contracting path that captures context and a symmetrical expanding path that enables precise localization, the CNN becomes a U-net.[32] Recently, U-nets have been further refined for image-to-image processing to translate one image (e.g., a line-drawing sketch of a clothing item) to a new image (e.g., a photo of the sketched clothing item).[33] Pix2pix is a type of generative model based on the concept of a conditional Generative Adversarial Network (cGAN).[33] The generator network in a cGAN is trained to produce outputs that are conditioned on a given input, whereas the discriminator network is trained to distinguish between the generated outputs and the ground-truth target outputs. During

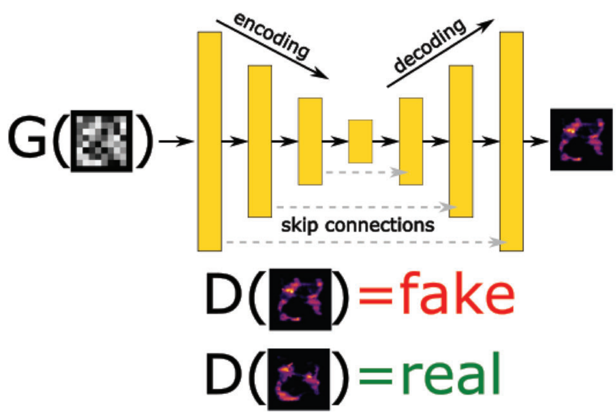

Figure 2. The conditional GAN architecture of pix2pix is based on a U-net with skip connections. The adversarial component arises from the incorporation of a discriminator *D* block in the training process. The generator *G* produces synthetic outputs that are compared to real outputs by the discriminator. As training progresses, it becomes more difficult for the discriminator to differentiate the real and generated (i.e., fake) prints. For application to DLP, the input images are grayscale photomasks and the output images are predicted height maps of a single layer of voxels.

training, the generator and discriminator networks play a two-player minimax game, with the generator attempting to produce outputs that are indistinguishable from the target outputs and the discriminator trying to correctly identify whether each sample is real or fake. In the case of pix2pix, the input to the generator is an image, and the output is a transformed version of the same image. The goal of the cGAN is to learn a mapping from input to output that is both realistic and conditioned on the input. The U-Net architecture used in pix2pix also includes skip connections, which allow information from earlier layers to bypass the series of down-sampling and up-sampling operations performed by the network. These connections help to preserve the spatial resolution and fine-grained details of the input image, allowing the network to generate high-quality outputs that closely match the input. Pix2pix has been used in a variety of increasingly challenging image-to-image translations, starting with tasks such as converting line drawings to photographs[33] and now pushing the envelope of medical image analysis.[34]

Here, we establish a high-throughput data workflow to characterize the geometry of 1500 unique single-layer voxel patterns produced by 100 000 randomly grayscaled pixel projections. The data are subsequently used to train and validate the ability of pix2pix cGAN and related non-adversarial U-nets to predict voxel geometry with micron-scale precision. We establish numerous metrics to validate the predictive capability of the models and the intrinsic variability in the prints. All U-net models are found to make compelling predictions of voxel geometry, with pix2pix cGAN showing the best performance at producing sharp features and capturing the spatial extent of adjacent voxel interactions. The models allow for a variety of virtual experiments on grayscale, single-layer printing to be carried out. Furthermore, the single-layer predictions can qualitatively inform intralayer and interlayer interactions and resulting geometry in real, multilayer 3D parts. Overall, the pix2pix model shows the promise of data-driven, ML approaches to predict and optimize VP additive manufacturing.

2. Results and Discussion

2.1. Data Curation and Workflow

A major hurdle in establishing voxel-scale ML for the prediction of VP has been the ability to register print and mask data with pixel or sub-pixel resolution. **Figure 3** shows an example of the workflow to register a series of 8 × 8 training masks to corresponding LSCM height maps of the training print. The choice to partition a larger print area into sub-prints was made based on the ability to provide a pixel-to-voxel registry in the post process. Gray level Y for individual pixels was randomly assigned 8-bit values between $Y = 0$ and 255. The exposure time was set to 5 s, which resulted in cure depth $C_d \approx 200$ µm in bulk (1 mm^2) working curve measurements at full intensity ($Y = 255$). The selection of the $n = 8$ pixels for $n \times n$ sub-mask dimensions was based on eight being much greater than the naïve voxel-to-voxel interaction distances determined by measuring cure depth as a function of n (i.e., center cure depth of the square patterns plateaued above $n = 3$). The choice of $n = 8$ has a few additional useful characteristics: it allows representation of moderately complex geometries for testing purposes, provides a large number of sub-masks in the LSCM's stitched field of view, and being base 2, provides scaling without interpolation up to the typical (e.g., 128 × 128 or 256 × 256 pixels) image sizes pix2pix was developed for. A total of 300 sub-masks, spaced n pixels apart, were fit onto an ≈25 × 50 mm pattern area. Interspersed evenly throughout the projection area were 14 full-intensity ($Y = 255$) fiducial sub-masks for image registration and alignment.

The masks were projected through a methacrylate-functionalized glass coverslip and into a ≈1 mm deep droplet of commercial photopolymer resin (as seen in Figure 1b) to produce the printed single-layer patterns shown in Figure 3b. The droplet thickness is expected to be sufficient to limit the availability of environmental oxygen at the resin-air interface to inhibit the reactions. The voxels adhere directly to the coverslip, and their geometry is not bounded by a build plate. This unbounded geometry may prove beneficial for predicting eventual multilayer 3D properties because it can better inform light penetration depth compared to a multilayer print with fixed layer thickness.[5] LSCM imaging of the voxels was performed on as-printed specimens with a 20× objective, resulting in sub-micron lateral and a few-nanometer height resolutions. To image the entire printed area required splitting the print into four regions with 75 sub-prints each. Imaging each of these four regions required ≈500 stitched fields of view, as shown in Figure 3e. The LSCM height map of the voxel patterns exhibits a significant range of variation because of the random gray sub-masks. Regions of no polymerization connect to voxels of varying heights. The fiducial prints are also clearly apparent from their larger heights and more uniform shape. A total of five coverslip specimens were printed (1500 sub-masks or ≈1×10^5 voxel interactions), with up to four specimens reserved for training and one test specimen reserved for independent verification.

Alignment between the mask and print was achieved by first scaling the mask to be similar in dimension to the LSCM height map of the print. Next, the centroids of the fiducial marks in the masks and prints were used to estimate a random sample

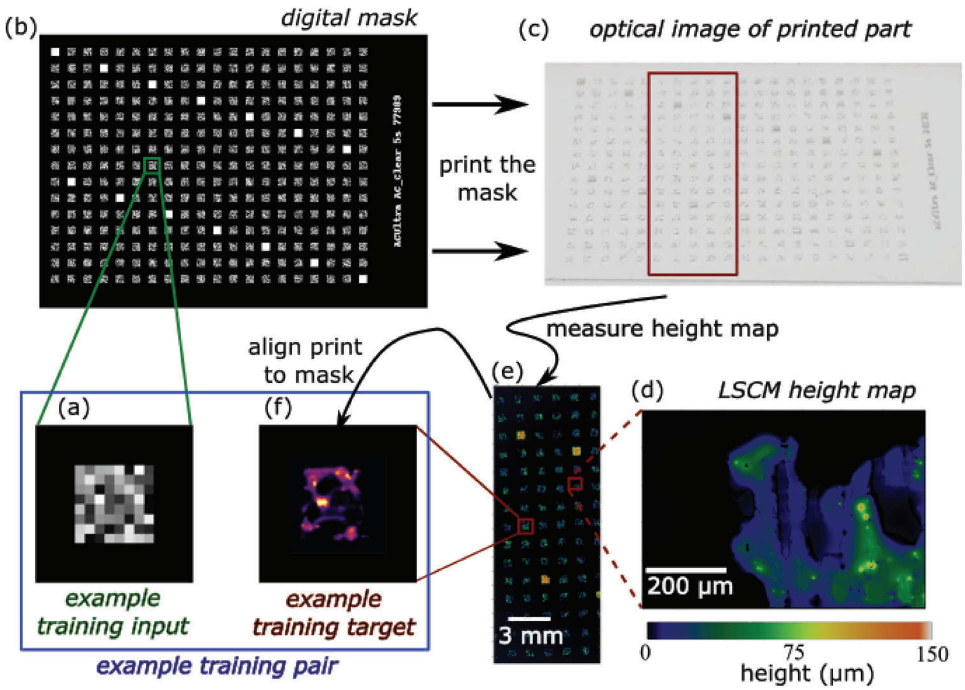

Figure 3. Workflow for measuring and processing gray-scale mask and corresponding print voxel geometry data. a) An example 8 × 8 pixels (8 × 8) mask. b) 300 8 × 8 random masks assembled in a grid, including 14 uniform white (8-bit gray level $Y = 255$) fiducial features. c) An optical micrograph of the print corresponding to the mask. d) A representative single field of view height map from the LSCM. e) The result of automatic stitching of 504 single fields of view to measure ¼ of the printed area from (b,c). Using reference features, the stitched LSCM map is aligned and registered to the input mask. The LSCM mask is then partitioned into a f) print that corresponds precisely with a) the associated pixels of the training mask.

consensus (RANSAC) 2D affine transformation. The transform was then applied to the print height map to register with the mask. The print and mask were partitioned into sub-prints and sub-masks, maintaining four border pixels around the sub-masks for defined boundary conditions, with a proportional border on the sub-prints. From the representative sub-masks and sub-prints in Figure 3, the lateral extent of the printed area is smaller than the corresponding mask due to under-polymerization on the pattern edges and cure shrinkage. It is also clear that all features printed from the random masks are lower height than the fiducial prints even though pixels approaching $Y = 255$ are present in the random masks, thus confirming that pixel interactions are contributing to local voxel cure depth. The input photomasks, raw LSCM maps, and processed training pairs are freely available at https://doi.org/10.18434/mds2-2950.

To assess the uniformity of the data sets across a given slide and between slides, statistical comparisons were applied to cropped versions of all 70 of the fiducial prints from the five slides. The fiducial marks exhibit a mean height (taken as the 95th percentile height of the voxel height map) of 215.3 µm with a standard deviation of 12.0 µm, lower bound of 154.2 µm and upper bound of 228.0 µm. In order to provide a comprehensive comparison of the print reproducibility, the 70 fiducial prints were assembled into 2415 unique pairings and the paired images were compared. The four statistical metrics used in this study to compare height maps of fiducial prints are the Pearson correlation coefficient (CC), 95th percentile height similarity (h95), structural similarity index metric (SSIM), and root mean square error (RMSE). **Figure 4** shows violin plots of the distributions of

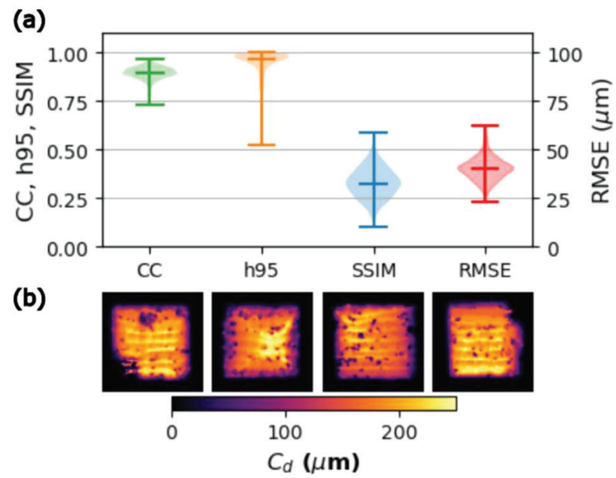

Figure 4. a) A statistical comparison of the 70 fiducial prints generated in this study. Fiducial prints are compared pairwise (2415 unique pairs) based on four metrics, Pearson correlation coefficient (CC), 95th percentile height similarity (h95), structural similarity index metric (SSIM), and root mean square error (RMSE). b) Four randomly selected fiducial prints indicative of the variability between the prints.

the various metrics across all the fiducial pairings. The CC values are calculated on flattened versions of the height maps and indicate strongly correlated images, with a median value of 0.895 and a tight distribution around that value. The h95 metric is designed to assess the similarity of extreme height features, which are particularly important in establishing interlayer adhesion in DLP printing. The value of h95 is given by 1 minus the absolute value of the difference between the 95th percentile heights of both height maps in a pair, normalized by the 95th percentile height of the first map in the pair. The fiducial marks exhibit 97% median h95 similarity between pairs. Notably, extreme outliers do exist, and h95 similarity as low as ≈50% was observed. The SSIM has been widely used for image-to-image translation tasks because it not only measures pixel-level similarity but also takes into account the structural and perceptual similarities between the images.[35,36] The median SSIM value of 0.33 represents the structural variations between fiducials, such as random voids, edge/boundary variations, and total cure depth variations, which can be seen in Figure 4b. We use SSIM here on the fiducials as a benchmark for later predictions of model performance. Finally, RMSE is a useful metric because it predicts the height error expected for a given pixel. The median RMSE value of 40 µm initially appears large given the strong CC and h95 reproducibility of the prints, however, these fiducial prints are typically > 200 µm high, and they do exhibit some defects such that when discrepancies do occur, they are often the full height between substrate and top surface. Overall, the high degree of reproducibility indicated by CC and h95 values between fiducial prints suggests that spatial variations within a slide (e.g., from vignetting of the light source) and between slides (e.g., due to sample preparation) were not unacceptably large, and it is reasonable to use data from a series of slides to predict print performance at an arbitrary location on a different slide.

2.2. Training of the Machine Learning Models

A publicly available implementation of pix2pix[37] for Pytorch (V1.13.1) was adapted for machine learning of the mapping between grayscale masks and voxel prints. The model was modified to optionally disable the discriminator block and implement the traditional scalar loss functions of mean squared error (denoted below by MSE) and mean absolute error (denoted below by L1) for comparison with pix2pix cGAN training. To train the models, up to 1144 sub-masks and sub-prints were loaded with 90% of data used for training and 10% used for validation (1030 training pairs and 114 validation pairs). All the fiducial prints were removed from the training and validation data to avoid biasing those predictions. Numerous model and algorithm hyperparameters were explored to optimize the print prediction quantitatively and qualitatively. While a globally optimized model was outside the scope of this proof-of-concept, key performance parameters were still identified. Of the conditions explored, the best compromise of prediction quality, computational burden, and model complexity was achieved with 1) a relatively small batch size of 8 to avoid overtraining, 2) a larger, single-layer discriminator patch to "see" a larger spatial extent of pixel interactions, and 3) 128 × 128 pixel input and output image size to balance print fidelity and overall image size. Given the physical expectation of rotational invariance of the prints and to improve the veracity of the model predictions, random vertical and horizontal mirror augmentations were applied identically to sub-masks and sub-prints.

In order to understand model training progression, the four metrics discussed in the fiducial print comparison (i.e., CC, h95, SSIM, and RMSE) were evaluated between 5 and 200 training epochs for 300 test prints and corresponding predictions. The masks and test prints were from a fully independent set of mask-print pairs that were not included in the training or validation. The results of the analysis are shown in **Figure 5**. As expected, increased training generally improves predictive performance, but optima for the different statistical quantities can occur at different epochs. Overall, the results indicate accurate predictions, but multiple metrics must be balanced when assessing overall model performance. The statistical analyses can be complemented with selected user inspections of representative prints and predictions, as shown in **Figure 6**.

The MSE and L1 loss models are trained with their corresponding loss functions enabled and the discriminator block disabled. These models are computationally less expensive than the cGAN, but they still produce compelling predictions of voxel geometry. From Figure 5, the rate of learning and SSIM performance is slightly better for the L1 loss compared to MSE loss functions. In contrast, MSE provides slightly improved CC and h95 performance. For the exemplary print in Figure 6, both models capture the key features such as the diagonal channel, the small moderate-height island in the upper left corner, and the two-height-step nature of the larger lower-right feature. The L1 model identifies these features sooner during training and appears more accurate after the full 200 epochs.

By introducing the discriminator block in the pix2pix model, predictions improve further. Particularly, the h95 similarity is higher for pix2pix compared to MSE and L1. From Figure 6, patchiness attributed to the discriminator is apparent at lower epochs, but this quickly transitions into a sharper-edged prediction. One concern regarding the use of pix2pix for DLP prediction was the model's ability to translate lateral shrinkage from the mask to the print. The presence of skip connections in the pix2pix model generally preserves the spatial input information (e.g., preserving the silhouette of a segmented building or sketch).[33] Nonetheless, the pix2pix model predicts similar lateral shrinkage as the real test print. The consistent ability of pix2pix to predict print geometry warrants further investigation into more complex input masks, such as finer pixel pitch, and more process parameters like exposure time.

In order to explore the potential of reducing the experimental burden of microscopic training-print characterization, the pix2pix models were also trained on significantly smaller training data sets of only 10 or 100 training pairs. The LSCM imaging of the samples consumes a considerable fraction of the total expense of this data-driven methodology. Furthermore, LSCM may not be appropriate for all voxel characterization in vat photopolymerization. For example, lower throughput techniques such as atomic force microscopy can provide even higher resolution imaging of voxel patterns in situ to the print environment and simultaneously provide other important information, such as mechanical property characterization.[5,21,27] In addition to lessening microscopy time, reduced data set size significantly reduces training time. With 100 training prints, the model still identifies the lateral pattern within 5 epochs, although the height range of the print is compressed compared to the real print. By ≈100 epochs, the 100-print model is comparable to the result from the full data set at 10 epochs and overall in good agreement on the major features of the ground truth. Further reduction of the data

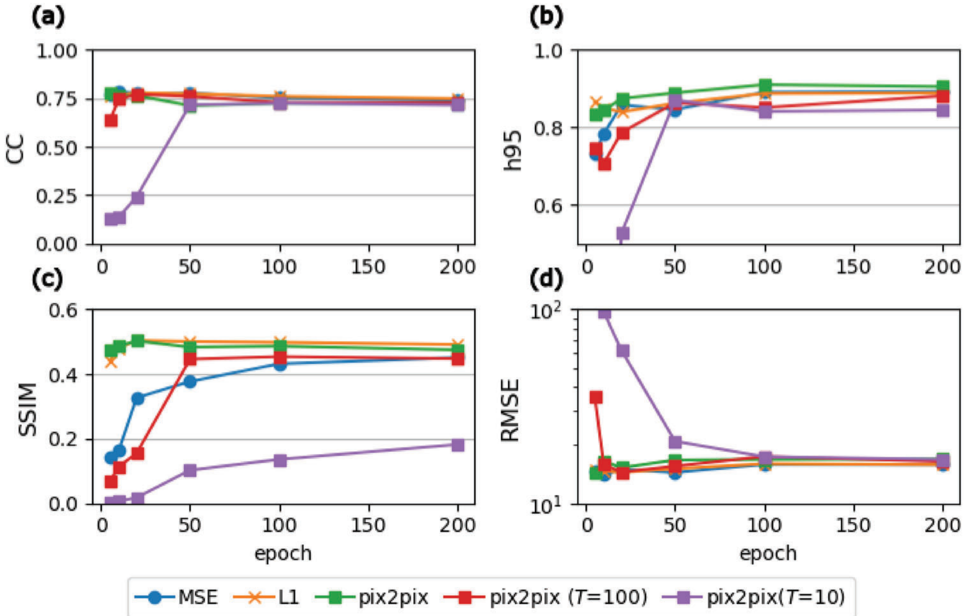

Figure 5. Statistical summary of model training to predict 300 independent test prints as a function of training epoch. Results from non-adversarial U-nets with scalar MSE and L1 loss functions and cGAN pix2pix are shown. For pix2pix, additional training with a reduced number of training pairs (T) is presented. Plots show median values for 4 statistical quantities: a) CC, b) h95, c) SSIM and d) RMSE.

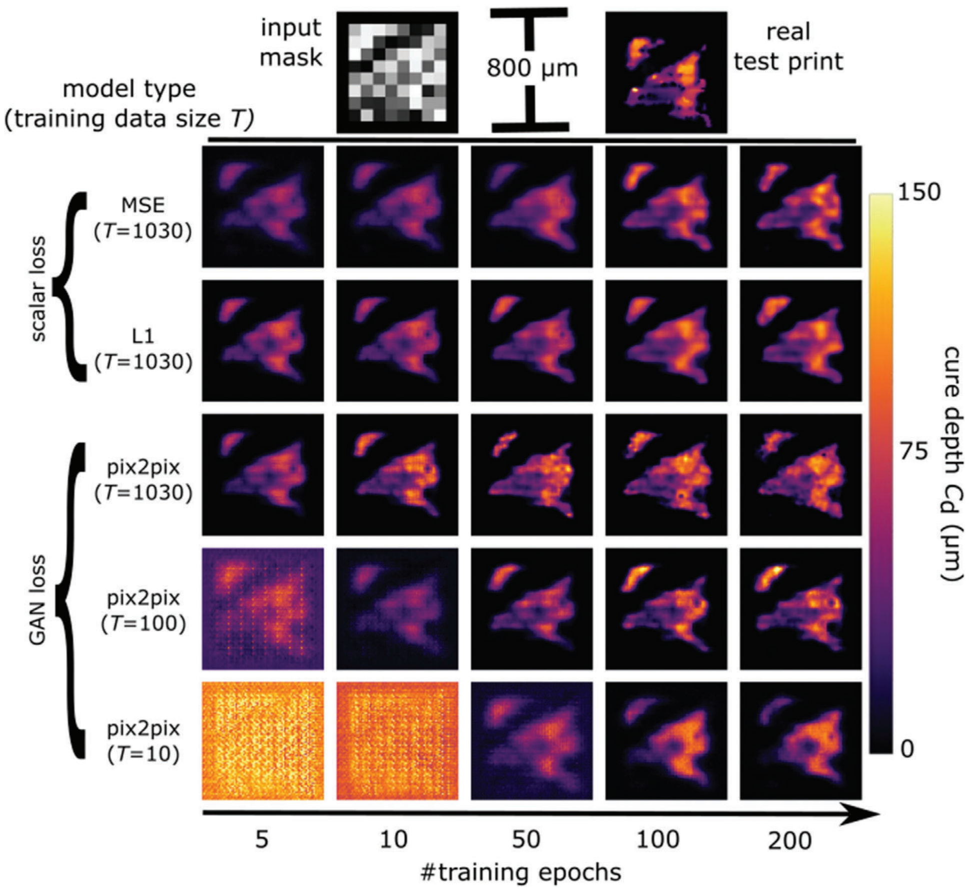

Figure 6. Progression of training for a representative grayscale mask and print. Results from non-adversarial U-nets with scalar MSE and L1 loss functions and cGAN pix2pix are shown. The pix2pix model performs well even when the training data size T is reduced to 10 or 100 training masks and prints. The structural similarity index metric SSIM is also shown between each prediction and the test print.

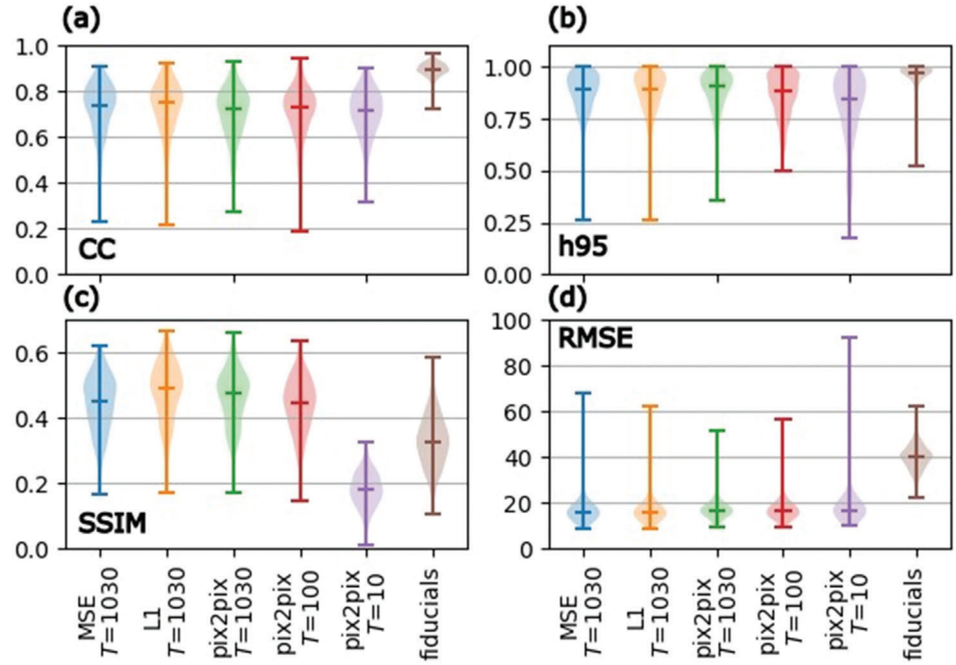

Figure 7. Statistical comparisons of model predictive performance after 200 training epochs. Violin plots show median, distribution and extrema for 4 statistical quantities: a) CC, b) h95, c) SSIM and d) RMSE for the 5 trained models (200 epochs). Distributions from the fiducial prints are also shown for reference.

set to only 10 prints adversely affects the print prediction over the observed 200 epochs, but additional refinement with more epochs may be possible. However, even with the limited data set, accurate predictions of the lateral extent of the cure are possible, and some physically consistent variations in cure depth are observed. Deliberate engineering of the input masks with more sophistication than the present random grayscale may provide a more efficient future pathway to accurate model training with a smaller number of training pairs.

2.3. Predictions of Single-Layer Printing

To compare the predictive power of the different models, **Figure 7** shows violin plots of the CC, h95, SSIM, and RMSE values between the 300 prints and predictions from the independent test slide. For benchmarking, the results of the fiducial print intercomparison are also shown. For CC, h95, and SSIM, we seek values of 1 with a tight distribution. However, the intrinsic variability of the prints sets an upper bound for predictive performance, so at best, the prediction is a denoised version of the real print. Cross-correlation values range from 0.72 to 0.75 for the different models, which is lower than the CC value of 0.89 on the fiducial prints. Despite similar median values between models, the minimum values are improved with the more sophisticated $T = 1030$ and $T = 10$ pix2pix models. The median h95 values between print and prediction are also similar between models, with the $T = 10$ pix2pix having a value of 0.85 and the other models between 0.88 and 0.91. These were again lower than the experimental fiducial h95 value of 0.97. The highest minimum value was achieved for the $T = 100$ pix2pix, but the overall distribution was tighter for the $T = 1030$ pix2pix model. SSIM values were generally well below 1.0, but surprisingly, the model predictions generally outperformed the inherent SSIM variability in the fiducial prints. Only the $T = 10$ pix2pix underperformed the fiducials. We attribute the ability of the models to overperform the fiducials to the increased amount of deliberate structure present in the random masks compared to the uniform fiducial masks. Even with similar types of defects, the inherent structure remains apparent in model predictions. Finally, median RMSE values ranged from 16 to 17 µm, which was considerably lower than the 40 µm RMSE value of the experimental fiducials, although the mean heights are also much lower for the random prints. Overall, the statistical analysis indicates that the models had a strong predictive ability, with uncertainties that compared favorably to the repeated fiducial prints. The analysis also indicates the importance of considering numerous metrics when assessing overall model performance. If only considering CC, the $T = 10$ pix2pix appears to be one of the best models, with a very small microscopy cost. However, when considering the other metrics, that model overwhelmingly performs the worst. Furthermore, we must consider the distribution of prediction metrics across numerous test patterns rather than just considering a mean or median value.

To complement the statistical analysis of print prediction, **Figure 8** shows a comparison of six selected test masks (five random, plus a fiducial), real test prints, and pix2pix predicted model prints obtained after 200 training epochs with $T = 1030$ training pairs. All prints are from the test slide that was not included in the training. Overall, the model produces faithful predictions of diverse, complex geometries, picking up subtle features such as narrow, sub-pixel width ribs, wider channels, and taller features correlated with clusters of bright mask pixels. In

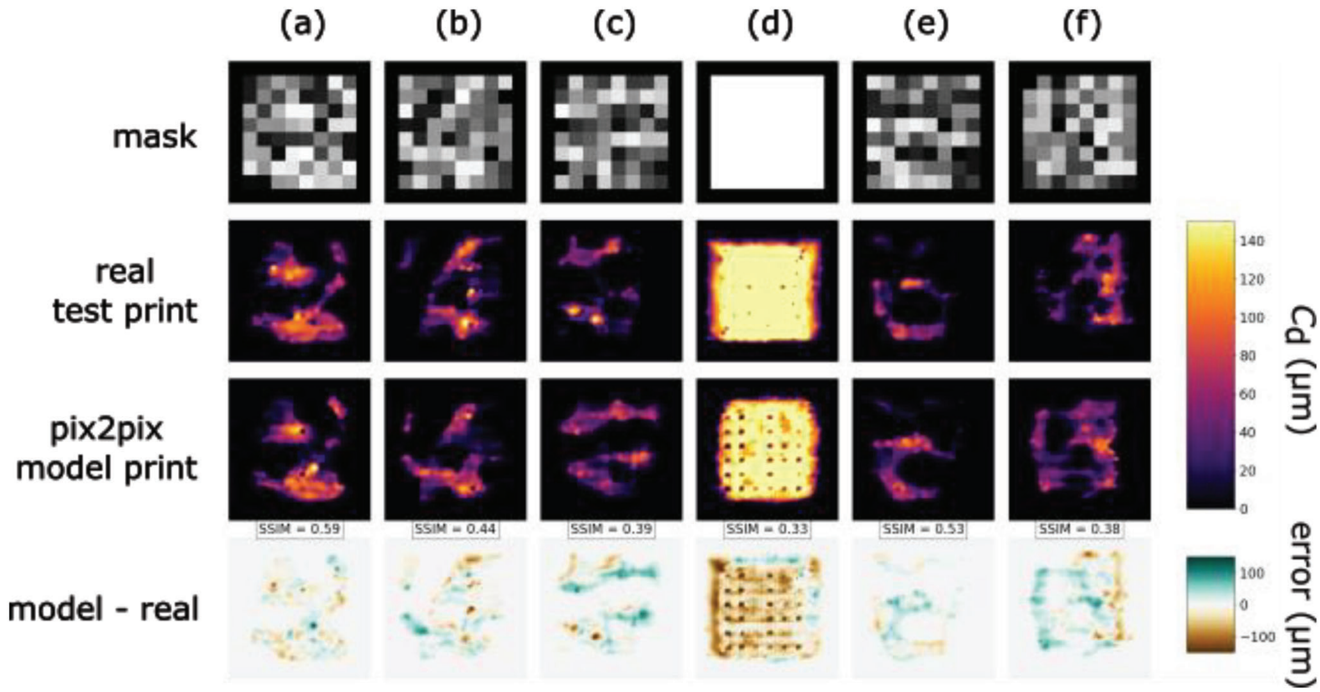

Figure 8. Comparison of real test print and pix2pix predicted print cure depths C_d after 200 epochs of training with $T = 1030$ training image pairs. Test pairs (a-f) were chosen randomly from the 300 prints on the fully independent test slide. Bottom row shows the cure depth error between the model and the test print. Error is defined as $error = C_d(\text{model}) - C_d(\text{real})$.

comparing model predictions to real prints, some discrepancies may be attributed to the denoising of the printing process by the model, whereas other discrepancies may be the model's inability to fit certain features. The bottom row indicates the cure depth error between model prediction and real print. Some errors may arise from slight misalignment, and however generally, the errors are distributed between over and under-estimation. The largest errors are seen at the edges of tall features, where a slight underestimation of lateral size can result in a very large error. Large errors are also observed on some isolated sharp features where the model might predict their location shifted by a pixel. The SSIM values range from 0.33 to 0.59, indicating good agreement between the model and predictions and representative of the broader statistical distribution in Figure 7c. The prediction of maximum voxel height bears particular importance to layer-by-layer printing. If voxel height is less than layer thickness, there is no mechanism to adhere the present layer to the previous layer, leading to geometric defects that could further propagate into later layers. If voxel height is much greater than layer thickness, the overlapping light exposures can result in unintended mechanical property gradients or the loss of negative features in a printed part.

Figure 8d shows a prediction for a uniform $Y = 255$ 8×8 fiducial mask. While the uniform print is less complex than most of the modeled pixel/voxel interactions in the study, it has high importance because real prints often have large regions of contiguous illuminated pixels. Notably, as shown in Figure S1, Supporting Information, the pix2pix model with $T = 1030$ training pairs significantly outperformed all other examined models in predicting the deeper cure depth of the Y = 255 square. Nonetheless, the predicted heights are still lower than the experimentally measured fiducial prints. The ability to predict isolated single-pixel cure depth, and larger cure depths arising from pixel-family interactions is attributed to the patch size of the discriminator. In traditional CNNs and U-nets, individual pixels are considered independent of adjacent pixels, and the result is unstructured loss. By comparison, the discriminator in pix2pix considers the interactions between adjacent pixels, resulting in a structured loss, which allows the model to predict larger length-scale interactions. The fiducial prints have double the average intensity of a random print, requiring significant extrapolation for prediction. Biasing some of the training data towards higher intensities may provide a more accurate prediction of extreme features in the future.

An advantage of the well-trained pix2pix model is the ability to perform virtual print experiments with very little effort compared to repeated printing and characterization. A wide range of virtual experiments are possible with the model, examples of which are shown in **Figure 9**. Figure 9a shows the prediction of cure depth for $n \times n$ patterns with uniform gray illumination given by Y. The curve is analogous to the working curve (WC) measurements prevalent in VP research and industry,[38] except that the gray level is used instead of the exposure dose. A notable feature of the curve is the inflection point from zero cure depth to non-zero cure depth. In WC measurements, the zero cure depth transition point is known as the critical exposure dose E_c. Here, a critical gray intensity Y_c is considered. Early assumptions in VP asserted that exposure energy was independent of feature size.[39] Here, the model predicts that Y_c is strongly size dependent, with no cure (i.e., $Y_c > 255$) for a small (i.e., $n = 1$ or 2 pixels) pattern compared to $Y_c = 50$ for an 8×8 uniform exposure. Figure 9b expands the control of cure depth by applying a gray padding border with

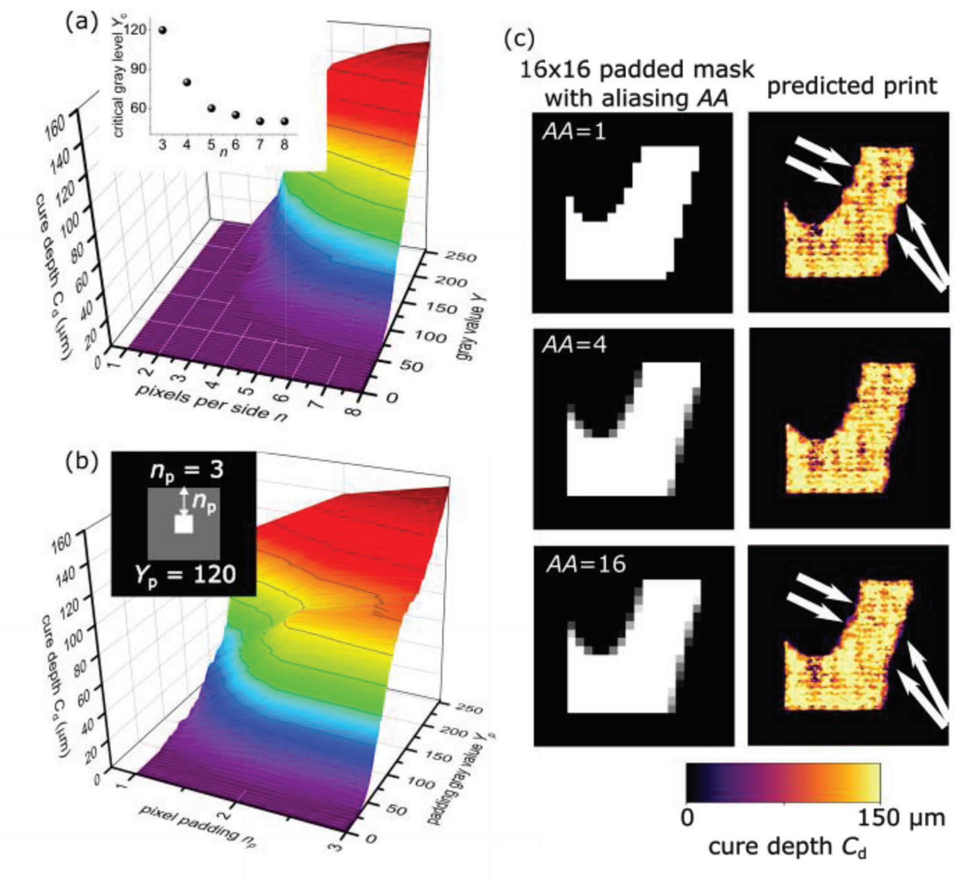

Figure 9. Virtual experiments performed with the pix2pix cGAN model. a) The feature size dependent cure depth versus square $n \times n$ pattern size and grayscale. b) The effects of gray padded border illumination of width n_p and intensity Y_p around a center $n = 2$ and Y = 255 feature. The border provides a means of providing finer pattern control. In (a,b), cure depth is determined from the 99th percentile of the height map prediction. c) cGAN predictions of antialiasing efficacy for 3 different antialiasing strengths (AA = 1, AA = 4or AA = 16).

intensity Y_p and width n_p to the perimeter of a $Y = 255$, $n = 2$ pixel center feature. The presence of a gray border, even at intensities below the minimum Y_c value in Figure 9a, can be used to tune the cure profile. The maximal values in Figure 9a,b are experimentally verified by the fiducial prints summarized in Figure 3. Both Figures 9a,b indicate that the length scale of pixel interaction is increased at lower light intensities. This is attributed to the increased time to gelation and hence increased time for species diffusion to occur. This observation has consequences regarding the choice of n for the training masks and padding. To further refine low-intensity interactions, training on even larger n values may be necessary.

Figure 9c applies the pix2pix model to the most common present use of grayscale DLP, antialiasing.[8] An n = 16 mask is populated with a curved feature, resulting in pixelation steps to represent the curvature. Without antialiasing (i.e., AA = 1), the stair step profile is clearly evident in the predicted print. In contrast, by increasing the intensity of the antialiasing setting in the slicing software (AA = 4 or 16), the stairstepping is largely eliminated. AA values of 4 or 16 result in slightly different lateral extent of cure between one another, thus, the model can optimize antialiasing to minimize stairstepping while maintaining dimensional control. While the model could help choose suitable AA values for the print from a discrete set of choices, the model is also invertible, which could allow for target geometries to be directly input.[33] Notably, the antialiasing study is performed on an n = 16 pixel pattern rather than the n = 8 size of the training data set. A known benefit of the pix2pix model is its applicability to larger images (given base 2-pixel dimensions) than it was trained on.[33] Validation of large mask performance is shown in Figure S6, Supporting Information. The larger masks were validated by partitioning the test mask into 32 by 32 pixel regions (256 × 256 after scale-up for mask and print), then performing the same statistical analyses as above. Validation values of 0.77, 0.90, 0.73, and 11 µm were obtained for CC, h95, SSIM, and RMSE, respectively. The model performs very well on the larger masks, outperforming even the smaller photomasks due to the grid-like structure of the larger partition. Figure S6b, Supporting Information shows an example large mask print and prediction, which visually confirms the accurate model performance. This applicability to larger masks has benefits when predicting real print outputs, given that real printers typically have pixel dimensions of hundreds or thousands in a given direction.

NANO · MICRO small
www.small-journal.com

A demonstration of print prediction with pix2pix cGAN in a real 3D-printed DLP part is shown in **Figure 10**. From Figure 1, a key difference between multilayer 3D prints and single-layer training prints is the type of window material. For multilayer printing, the window must exhibit very low adhesion to the layers, which is achieved here using fluorinated ethylene propylene (FEP) film. FEP relies on low surface energy to prevent adhesion, compared to other methods like CLIP that complement low surface energy with high oxygen permeability to inhibit the reaction.[3] Because the functionalized glass used in single-layer training and the FEP used in multilayer printing have low oxygen permeability, similar reaction kinetics in the adjacent resin are expected. Multilayer parts also provide an opportunity for interactions between layers, such as monotonic heat build-up from the exothermic reactions,[40,41] swelling and shrinkage-induced deformation,[42] oligomer trapping, and arbitrary z-boundary conditions depending on previous layer print geometry. Thus, the single-layer prediction can only provide qualitative insight into the final multilayer part geometry. Nonetheless, the single-layer predictions can inform over- and under-polymerization within and between layers, which could be predictors of multilayer failures.

A pixelated gyroid lattice geometry (shown in Figure 10a) was chosen for demonstration because of the relevance of lattice structures in AM and the broad feature size variation present in the part (from single pixel-wide features to large pixel clusters). To assess both layer-scale and pixel-scale print failures, the gyroid part was replicated across the build plate while varying the gray level Y and the end layer of the part z_{last}. Variation of z_{last} allows for the final layer of the part to represent what would otherwise be an internal layer in the part. The masks associated with the last six layers ($z = 31$ to 36) are shown in Figure 10b, noting that only $Y = 255$ is shown, but other Y levels from 60 to 250 were also printed. Parts were printed with four burn-in layers with 30 s exposure per layer, followed by the same 5 s exposure used in model training for layers $z = 5$ through $z = z_{\text{last}}$. The layer thickness was fixed at 50 µm throughout the print. Between each layer, an atypically slow retraction speed for the build plate was used to give the system ≈120 s to dissipate any heat and local oligomers generated in the reaction before exposing the next layer. Transient phenomena at the scale of the full part are not trained in the model and thus were sought to be minimized. In principle, layer-layer interactions could be trained from multilayer prints with 3D CNNs. Conversely, some effects could be encoded into a second layer, such as controlling and measuring vat temperature or deliberately creating pre-gelation oligomers. As shown in Figure S8, Supporting Information, printing with a more typical 5 s between layers resulted in considerably more overall polymerization of features and was not well predicted by the pix2pix model, which was trained on single, isolated layers.

Figure 10c shows the resultant build plate from the printed array varying Y and z_{last}. Notably, for $Y < 140$, only the burn layers print. At $Y = 140$, extremely fragile structures were printed, but they failed partially or completely during rinsing in isopropanol. For $Y > 140$, parts generally printed to completion, however, defects in lattice geometry were apparent. Figure 10d shows the pix2pix prediction of cure depth for $z = 36$ (z = 31 to 35 are in Supporting Information). Because printed parts are assembled on a build plate opposite the window, the model predictions must be considered based on how the current layer with predicted cured depth would adhere to a preceding layer whose thickness was fixed by the layer height. If C_d is less than layer thickness for a contiguous feature, we expect that feature to fail to adhere to the previous layer. If C_d is greater than layer thickness, we expect adhesion to be possible, however, we risk over-polymerizing deliberate void space in the previous layer. If C_d varies between less than and greater than layer thickness, intermittent adhesion of that feature may occur. Within layers 31 through 36, layer $z = 36$ exhibits the smallest average feature size and is predicted to have the shallowest cure depth. The cure is negligible up to $Y = 100$. At $Y = 140$, we predict that only intermittent pixels of the largest features will reach a cure depth close to the 50 µm layer thickness. The partial printing of the full part at $Y = 140$ is attributed to the ability of larger features to cure deeper and provide intermittent adhesion, as well as possible swelling of the part in the monomer, which might slightly reduce the programmed layer height. As Y is increased to 160 and above, increasing numbers of features can reach 50 µm cure depth, however, the single pixel-wide features do not reliably reach 50 µm until $Y > 220$.

While the model proves useful in making qualitative, layer scale predictions of print failure, the utility of the subvoxel scale geometry training with pix2pix is the ability to make voxel scale predictions for whatever arbitrary layers compose the part. Figure 10e shows LSCM height maps of the last printed layer when z_{last} is varied from 31 to 36 and Y is either 160 or 255. The accompanying pix2pix predictions indicate where the cure depth is expected to achieve the 50 µm layer thickness versus where it will come up short. Considerable variation in cure depth is observed between $z = z_{\text{last}} = 31$ and $z = z_{\text{last}} = 36$. The smaller average feature widths at higher z result in locally shallower cure depths, with many features failing to reach 50 µm cure depth. These predictions are consistent with the LSCM height maps, which show feature replication at $z_{\text{last}} = 31$, albeit with a warpage of the final morphology. For $z_{\text{last}} = 32$, the presence of narrower features in the mask results in regions of cure depth ≪ layer thickness. In the height maps, these regions show up as defects, where larger features abruptly terminate when mask dimensions reduce to a single pixel. For $z_{\text{last}} > 32$, the prevalence of failed features increases, and the parts exhibited considerable deformation, which we attribute to increased stress from lower average conversion in the part. In contrast to the $Y = 160$ result, at $Y = 255$, nearly all illuminated features are predicted to cure deeper than the layer thickness. Narrow features only just reach 50 µm cure depth, but wide features exceed 100 µm cure depth. In the LSCM height maps, the $Y = 255$ part surfaces show the complete nature of the patterns, with much wider features than corresponding features at $Y = 160$. Overall, the example shows the pitfall of printing full parts with uniform irradiance. To print fine features, large features will be over-polymerized, but for printing accurate large features, the fine features will fail to print at all. A key use of grayscale voxel models going forward will be to design optimal masks for parts with varying length scale features.

3. Conclusion

A high throughput method of characterizing voxel scale geometry of test patterns printed with grayscale digital light processing was developed. The method uses laser scanning confocal

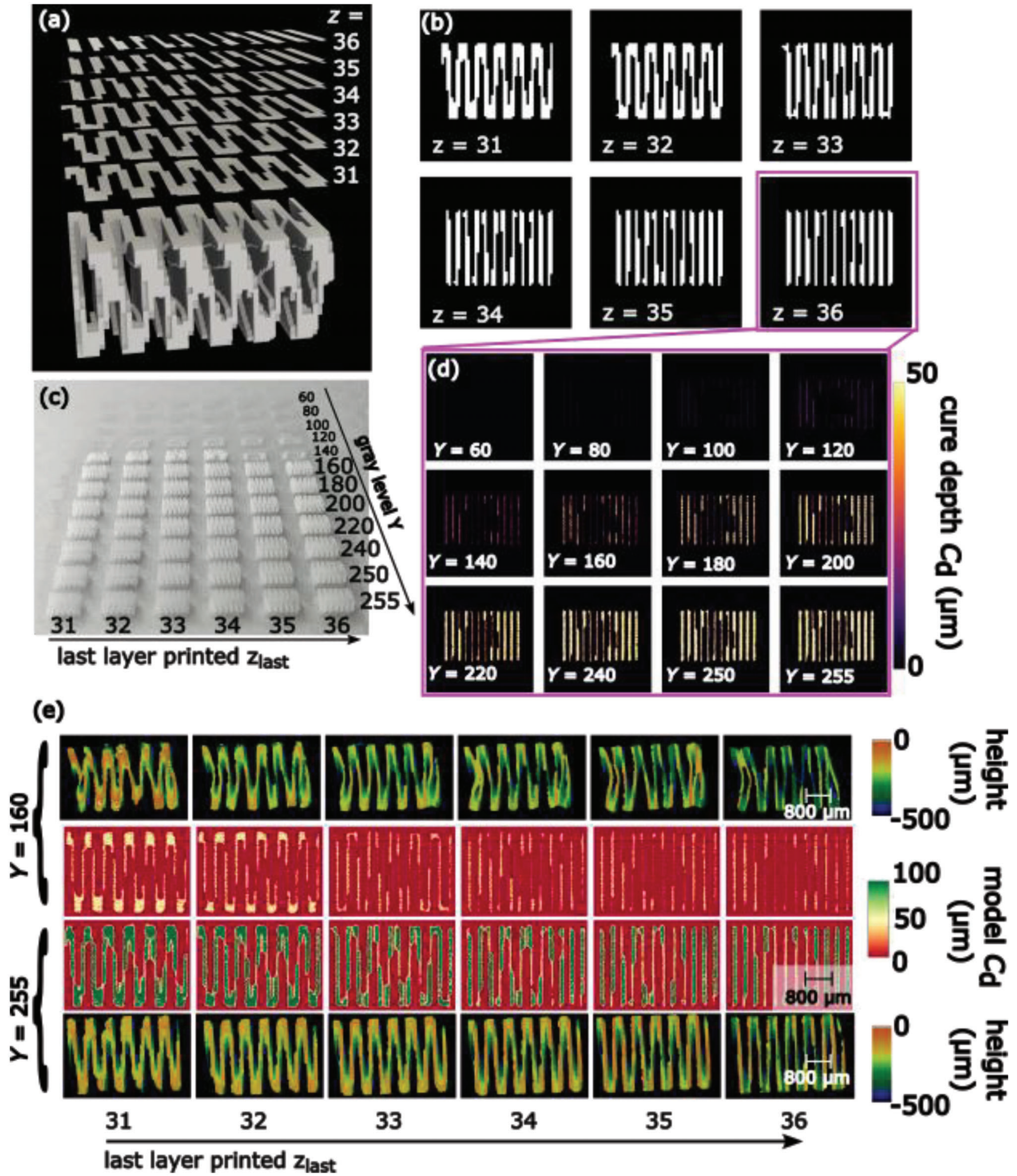

Figure 10. Application of pix2pic cGAN to layer prediction in real 3D printed lattices. a) Pixelated geometry of the gyroid lattice model, with the last six layers exploded for clarity. b) Detailed photomasks of the last six layers ($z = 31$ through $z = 36$), which exhibit progressively smaller average feature sizes. c) Photograph of build plate from the DLP printer where the gyroid part was printed at varying gray scale Y and sequentially increasing last layer z_{last}. A wait time of 120 s between layers was used to mitigate heat buildup and diffusion effects. d) The pix2pix prediction of layer $z = 36$ for increasing Y from 60 to 255. The transition from $C_d \gtrapprox 50$ μm to $C_d \lessapprox 50$ μm coincides with the transition from failed print to generally successful print in (c). e) The experimental height map (top and bottom rows) and predicted cure depth (middle two rows) variation in voxel scale cure on layers $z = 31$ through $z = 36$ at $Y = 160$ or $Y = 255$. Experimental variation is assessed based on the LSCM height maps taken when $z = z_{last}$. Height scale for LSCM maps is set to the last 500 μm (≈10 layers) of print.

microscopy to measure the geometry of tens of thousands of interacting voxels. The printed voxels are aligned to their corresponding grayscale masks and then the mask-print pairs are used as training data for machine learning models. Neural network models based on U-net architecture are shown to provide accurate predictions of the complex patterns that arise from neighboring voxels interacting with one another. The pix2pix cGAN is used as an exemplary adversarial network to capture the sharp features and spatial interactions that arise in DLP printing. In pix2pix, the discriminator patch can consider numerous interacting pixels and voxels during training, making it well suited to capture the effects of light non-uniformity, mass transport, heat transfer, and reaction kinetics that dictate final part geometry. The trained model successfully predicted numerous pixel patterns that were not included in the training data. The model was also used to perform virtual experiments and to predict layer scale and voxel scale failures in real 3D printed parts. Overall, the use of data-driven methods in DLP predictions can improve mask design, promote new process discovery, and ultimately enable the production of higher-performance parts for critical applications.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This research was performed while Thomas J. Kolibaba held a National Research Council Research Associateship Award at the National Institute of Standards and Technology. The contribution of NIST, an agency of the US government; not subject to copyright. The authors acknowledge the MIMIC facility at CU Boulder (RRID:SCR_019307) for the use of the laser scanning confocal microscope.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available in “A Data-Driven Approach to Complex Voxel Predictions in Grayscale Digital Light Processing Additive Manufacturing Using U-nets and Generative Adversarial Networks” at https://doi.org/10.18434/mds2-2950, reference number 2950.

Keywords

3D printing, digital light processing, machine learning, neural networks, pix2pix, stereolithography, vat photopolymerization

Received: March 8, 2023
Revised: May 19, 2023
Published online:

[1] T. Akahane, K. Takahashi, *IEEJ Trans. Sens. Micromach.* **2001**, *121*, 275.
[2] A. Bertsch, J. Y. Jézéquel, J. C. André, *J. Photochem. Photobiol., A* **1997**, *107*, 275.
[3] J. R. Tumbleston, D. Shirvanyants, N. Ermoshkin, R. Janusziewicz, A. R. Johnson, D. Kelly, K. Chen, R. Pinschmidt, J. P. Rolland, A. Ermoshkin, E. T. Samulski, J. M. DeSimone, *Science* **2015**, *347*, 1349.
[4] M. M. Emami, M. Jamshidian, D. W. Rosen, *J. Manuf. Sci. Eng.* **2021**, *143*, 091005.
[5] A. C. Uzcategui, C. I. Higgins, J. E. Hergert, A. E. Tomaschke, V. Crespo-Cuevas, V. L. Ferguson, S. J. Bryant, R. R. McLeod, J. P. Killgore, *Small Sci.* **2021**, *1*, 2000017.
[6] S. M. Montgomery, C. M. Hamel, J. Skovran, H. J. Qi, *Extreme Mech. Lett.* **2022**, *53*, 101714.
[7] C. Zhou, Y. Chen, *J. Manuf. Processes* **2012**, *14*, 107.
[8] K. Mostafa, A. J. Qureshi, C. Montemagno, *Proceedings of the ASME 2017 International Mechanical Engineering Congress and Exposition. Volume 2: Advanced Manufacturing*, Tampa, Florida, USA **2017**, V002T02A035.
[9] E. Guven, Y. Karpat, M. Cakmakci, *Addit. Manuf.* **2022**, *57*, 102954.
[10] R. M. Greene, R. M. Schmidt, Sub-pixel Grayscale Three-dimensional Printing, USA **2016**, 10354445B2.
[11] L. Yue, S. Macrae Montgomery, X. Sun, L. Yu, Y. Song, T. Nomura, M. Tanaka, H. Jerry Qi, *Nat. Commun.* **2023**, *14*, 1251.
[12] G. I. Peterson, J. J. Schwartz, D. Zhang, B. M. Weiss, M. A. Ganter, D. W. Storti, A. J. Boydston, *ACS Appl. Mater. Interfaces* **2016**, *8*, 29037.
[13] J. E. Hergert, A. C. Uzcategui, A. Muralidharan, V. Crespo-Cuevas, V. L. Ferguson, R. R. McLeod, *Adv. Eng. Mater.* **2022**, *24*, 2101543.
[14] C. T. Forte, S. M. Montgomery, L. Yue, C. M. Hamel, H. J. Qi, *J. Appl. Mech.* **2023**, *90*, 071003.
[15] X. Kuang, J. Wu, K. Chen, Z. Zhao, Z. Ding, F. Hu, D. Fang, H. J. Qi, *Sci. Adv.* **2019**, *5*, aav5790.
[16] A. S. Limaye, D. W. Rosen, *Rapid Prototyping J.* **2007**, *13*, 76.
[17] H.-W. Kang, J. H. Park, D.-W. Cho, *Sens. Actuators, A* **2012**, *178*, 223.
[18] M. M. Emami, D. W. Rosen, *Addit. Manuf.* **2020**, *36*, 101595.
[19] T. Deng, W. Liu, W. Xie, J. Huang, A. Tang, *The Int. J. Adv. Manuf. Technol.* **2021**, *113*, 3011.
[20] S. M. Montgomery, F. Demoly, K. Zhou, H. J. Qi, *Adv. Funct. Mater.* **2023**, *33*, 2213252.
[21] C. I. Higgins, J. P. Killgore, F. W. DelRio, S. J. Bryant, R. R. McLeod, *Soft Matter* **2020**, *16*, 4131.
[22] T. E. Brown, V. Malavé, C. I. Higgins, A. P. Kotula, B. W. Caplins, E. J. Garboczi, J. P. Killgore, *ACS Appl. Polym. Mater.* **2021**, *3*, 290.
[23] M. D. Xames, F. K. Torsha, F. Sarwar, *J. Intell. Manuf.* **2022**, *34*, 2529.
[24] Y. Yang, M. He, L. Li, *J. Cleaner Prod.* **2020**, *251*, 119710.
[25] Q. Huang, Y. Wang, M. Lyu, W. Lin, *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 1486.
[26] S. You, J. Guan, J. Alido, H. H. Hwang, R. Yu, L. Kwe, H. Su, S. Chen, *J. Manuf. Sci. Eng.* **2020**, *142*, 081002.
[27] H. Gojzewski, Z. Guo, W. Grzelachowska, M. G. Ridwan, M. A. Hempenius, D. W. Grijpma, G. J. Vancso, *ACS Appl. Mater. Interfaces* **2020**, *12*, 8908.
[28] B. Bircher, S. Wyss, D. Gage, A. Küng, C. Körner, F. Meli, *Proceedings of euspen SIG Advancing Precision in Additive Manufacturing*, St. Gallen, **2021**, p. 4.
[29] E. Cakmak, P. Bingham, R. W. Cunningham, A. D. Rollett, X. Xiao, R. R. Dehoff, *Mater. Charact.* **2021**, *173*, 110894.
[30] B. W. Caplins, C. I. Higgins, T. J. Kolibaba, U. Arp, C. C. Miller, D. L. Poster, C. J. Zarobila, Y. Zong, J. P. Killgore, *Addit. Manuf.* **2023**, *62*, 103381.
[31] S. Albawi, T. A. Mohammed, S. Al-Zawi, presented at *2017 Int. Conf. on Engineering and Technology (ICET)*, Antalya, Turkey, August **2017**, pp. 1–6.
[32] O. Ronneberger, P. Fischer, T. Brox, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th Int. Conf.*

(Eds: N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi), Springer, Munich, Germany **2015**, p. 234.
[33] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, presented at *2017 IEEE Conf. on Computer Vision and Pattern Recognition*, Honolulu, HI, July **2017**, p. 1125.
[34] M. Mehmood, N. Alshammari, S. A. Alanazi, A. Basharat, F. Ahmad, M. Sajjad, K. Junaid, *J. King Saud Univ., Comput. Inf. Sci.* **2022**, *34*, 4358.
[35] C. Wang, C. Xu, C. Wang, D. Tao, *IEEE Trans. Image Process.* **2018**, *27*, 4066.
[36] W. Zhou, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, *IEEE Trans. Image Process.* **2004**, *13*, 600.
[37] pytorch-CycleGAN-and-pix2pix, https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix (accessed: January 2023).
[38] J. Bennett, *Addit. Manuf.* **2017**, *18*, 203.
[39] P. F. Jacobs, in *Proc. of Solid Free Form Fabrication Symp.* (Eds: H. L. Marcus, J. J. Beeman, J. W. Barlow, D. L. Bourell, R. H. Crawford), Texas Scholarworks, Austin, TX **1992**, pp. 192–211.
[40] J. Bachmann, E. Gleis, S. Schmölzer, G. Fruhmann, O. Hinrichsen, *Anal. Chim. Acta* **2021**, *1153*, 338268.
[41] D. A. Walker, J. L. Hedrick, C. A. Mirkin, *Science* **2019**, *366*, 360.
[42] K. Sekmen, T. Rehbein, M. Johlitz, A. Lion, A. Constantinescu, *Mech. Mater.* **2023**, *179*, 104566.