

Ziyang Zhang

Department of Mechanical and Aerospace
Engineering,
University of Central Florida,
Orlando, FL 32816
e-mail: ziyang.zhang@knights.ucf.edu

Laxmi Poudel

Department of Mechanical Engineering,
University of Arkansas,
Fayetteville, AR 72701
e-mails: lpoudel@email.uark.edu;
poudel@email.uark.edu

Zhenghui Sha

Department of Mechanical Engineering,
University of Arkansas,
Fayetteville, AR 72701
e-mail: zsha@uark.edu

Wenchao Zhou

Department of Mechanical Engineering,
University of Arkansas,
Fayetteville, AR 72701
e-mail: zhouw@uark.edu

Dazhong Wu¹

Department of Mechanical and Aerospace
Engineering,
University of Central Florida,
Orlando, FL 32816
e-mail: dazhong.wu@ucf.edu

Data-Driven Predictive Modeling of Tensile Behavior of Parts Fabricated by Cooperative 3D Printing

3D printing has been extensively used for rapid prototyping as well as low-volume production in aerospace, automotive, and medical industries. However, conventional manufacturing processes (i.e., injection molding and CNC machining) are more economical than 3D printing for high-volume mass production. In addition, current 3D printing techniques are not capable of fabricating large components due to the limited build size of commercially available 3D printers. To increase 3D printing throughput and build volume, a novel cooperative 3D printing technique has been recently introduced. Cooperative 3D printing is an additive manufacturing process where individual mobile 3D printers collaborate on printing a part simultaneously, thereby increasing printing speed and build volume. While cooperative 3D printing has the potential to fabricate larger components more efficiently, the mechanical properties of the components fabricated by cooperative 3D printing have not been systematically characterized. This paper aims to develop a data-driven predictive model that predicts the tensile strength of the components fabricated by cooperative 3D printing. Experimental results have shown that the predictive model is capable of predicting tensile strength as well as identifying the significant factors that affect the tensile strength. [DOI: 10.1115/1.4045290]

Keywords: cooperative 3D printing, tensile strength, predictive modeling, machine learning

1 Introduction

While 3D printing has many advantages over traditional manufacturing processes, two primary challenges remain: (1) limited build size and (2) low throughput [1,2]. For example, the build volume of the majority of desktop and industrial-grade fused deposition modeling (FDM) 3D printers is less than 150 mm × 150 mm × 150 mm and 800 mm × 800 mm × 800 mm, respectively [2]. While the build volume of industrial-grade 3D printers has increased over the past few years, large-scale and high-speed 3D printers are very expensive. For example, Ingersoll Machine Tools developed the largest 3D printer in the world in collaboration with the Oak Ridge National Laboratory. Ingersoll claimed that this machine is capable of printing a wide range of composite plastics. The maximum build volume of the machine is 6 m (L) × 2.4 m (W) × 1.8 m (H). This 3D printer fabricates parts at an order of magnitude larger than any other 3D printer does [2].

In addition to limited build size, the amount of build time it takes to fabricate a complex part can be a few hours or even several days. A 3D printing speed is dependent on the infill speed, outer wall speed, inner wall speed, and support infill speed. These speeds are also dependent on 3D printing process parameters such as printing temperature, filament thickness, layer thickness, build orientation, and the volume of support structures. The majority of desktop 3D printers on the market has a maximum 3D printing speed of 150 mm/s. The Delta WASP 20×40 Turbo 3D desktop 3D printer can achieve a maximum 3D printing speed of 500 mm/s. The HP Jet Fusion 4200 industrial-grade 3D printer can achieve a maximum 3D printing speed of 4500 cm³/h. The

big area 3D printer developed by the Oak Ridge National Laboratory can achieve a maximum feed rate of 80 pounds/h.

While several industrial-grade 3D printers with high speed and large build volume have been developed, few manufacturers have access to them because they are very expensive. To address this issue, a cooperative 3D printing process was recently developed [3]. Cooperative 3D printing is an additive manufacturing process where individual mobile 3D printers fabricate a portion of a part divided by a slicing algorithm. These mobile robots or extruders work cooperatively on printing the part, thereby increasing manufacturing throughput and build volume. While cooperative 3D printing has the potential to transform the 3D printing industry, the mechanical properties of the parts fabricated by this novel 3D printing process have not been examined. The main contributions of this research include:

- While finite element analysis (FEA) has been used to model the mechanical behavior of additively manufactured parts, it is difficult to model the effects of 3D printing process parameters on mechanical behavior using FEA. While cohesive elements in FEA can be used to model bonded interfaces, it is difficult to determine the boundary conditions at the bonded interface between the objects printed by multiple printheads in cooperative 3D printing. To address this issue, we introduced a data-driven predictive modeling approach that can predict the tensile strength of the specimens fabricated by cooperative 3D printing. We also compared the performance of the machine learning model with that of a linear regression model. The machine learning model outperforms the linear regression model significantly.
- To understand the process-property relationship in the components fabricated by cooperative 3D printing, we evaluated the effects of three 3D printing parameters on tensile strength with

¹Corresponding author.

Manuscript received July 18, 2019; final manuscript received October 18, 2019; published online October 25, 2019. Assoc. Editor: Mahesh Mani.

machine learning algorithms. Experimental results have shown that the effect of the angle of incline on tensile strength is significant. We compared the performance of the machine learning model with the analysis of variance (ANOVA), which is a traditional statistical method. The results generated by machine learning are consistent with those of ANOVA.

The remainder of this paper is organized as follows: Sec. 2 presents the related work on the modeling and simulation approaches to the prediction of the mechanical properties of specimens built by 3D printing processes. Section 3 presents an ensemble learning algorithm. Section 4 introduces the design of experiments and tensile tests. Section 5 presents the accuracy of the predictive model. Section 6 summarizes the conclusions.

2 Related Work

This section presents an overview of related work on the modeling and prediction of mechanical properties of additively manufactured parts as well as cooperative 3D printing.

2.1 Prediction of Mechanical Properties of Additively Manufactured Components. Melenka et al. [4] developed a volume averaging stiffness approach that can predict the tensile modulus of the Kevlar fiber-reinforced specimens. The MarkOne 3D printer was used to print the specimens with different volume of fibers. The predictive model was demonstrated to be effective especially when the fiber volume is high. Hayes et al. [5] developed a constitutive equation that can make predictions of the yield strength of the Ti-6Al-4V specimens fabricated by an electron beam 3D printing process. Experimental results have revealed that the predictive model can predict the yield strength with a prediction error of 5%. Tapia et al. [6] developed a predictive model based on Gaussian process to predict the porosity of metallic specimens fabricated by selective laser melting (SLM). Experimental results have indicated that the proposed method can predict the porosity of stainless steel specimens with a mean-squared error of 0.2593. Campoli et al. [7] developed a finite element (FE) model in order to predict the elastic constants including Young's modulus and Poisson's ratio of the specimens fabricated by SLM. Based on the experimental results, the FE model outperformed the analytical models of the mechanical properties. Mukherjee et al. [8] developed a heat transfer and fluid flow model that can predict the residual stresses and distortion in 3D printing. The predicted temperature and residual stress were compared with the experimental data. Experimental results have implied that the proposed model can achieve high prediction accuracy for the specimens made of Inconel 718 and Ti-6Al-4V.

2.2 Cooperative 3D Printing. McPherson and Zhou [3] developed a new chunk-based slicing algorithm and a cooperative printing platform in order to improve 3D printing throughput. The slicing algorithm split a printing task into small parallel subtasks such that multiple mobile robots can print a part cooperatively. A simulator was developed to validate the slicing algorithm. Poudel et al. [9] evaluated the tensile strength of dog-bone specimens fabricated by a cooperative 3D printing process. The effect of the slope angle, number of shells, and overlapping length between chunks on tensile strength was evaluated. A comparative study of the tensile strength of the specimens fabricated by the cooperative printing process and the standard FDM process was performed. Based on the experimental results, tensile strength decreases as the slope angle and overlapping length between chunks increase. Hunt et al. [10] developed a flying robot-assisted 3D printing prototype that is capable of depositing materials by a quadcopter with an on-board printing module. The flying robot-assisted 3D printing prototype was demonstrated by two printing scenarios.

In summary, while extensive work has been devoted to the modeling and prediction of the mechanical properties of a wide range of

materials fabricated by 3D printing processes, little research has been reported on the prediction of the tensile strength of specimens fabricated by cooperative 3D printing processes. To fill this research gap, this paper aims to develop a data-driven predictive modeling approach to estimate the tensile strength of polymers fabricated by the cooperative 3D printing process.

3 Ensemble Learning

Ensemble learning is a supervised machine learning algorithm that combines multiple learning algorithms (also known as base learners) in order to enhance prediction performance. The output of the ensemble learning algorithm has the following form:

$$\psi(\mathbf{W}) = \sum_{i=1}^K w_i \psi_i \quad (1)$$

where K represents the number of base learners and w_i is the weight of the i th base learner ψ_i . The optimal weight vector \mathbf{W}_0 is determined using non-negative least squares, which can be formulated as

$$\mathbf{W}_0 = \arg \min_{\omega} \left[\frac{1}{2} \|\mathbf{A}\omega - \mathbf{b}\| \right] \quad (2)$$

subject to $\omega \geq 0$

Here, \mathbf{A} represents the output of all base learners and \mathbf{b} represents the observed value. In this study, the non-negative least squares problem was solved using the Lawson–Hanson algorithm.

The selection of base learners is critical to the prediction performance. Equation (1) suggests that the output of ensemble learning has the form of a linear model. Each base learner can be considered as a single feature. In order to reduce the collinearity between different features, the selected base learners should be accurate and diverse. Selecting too many base learners in a model does not always improve prediction performance because it might cause the overfitting problem. In this study, the performance of multiple base learners, including gradient boosting, extreme gradient boosting, bootstrap aggregating, random forests, extremely randomized trees, kernel k-nearest neighbors, support vector regression, stepwise regression, multivariate adaptive regression splines, generalized linear models, Bayesian generalized linear models, and Lasso and ridge regression, was tested. Different combinations of these algorithms were evaluated in order to achieve the best performance. After comparing the performance of different combinations of base learners, three base learners, including Lasso, support vector regression (SVR), and extreme gradient boosting (XGBoost), were selected from three different machine learning categories. The final predictive model trained by the ensemble learning algorithm combines Lasso, SVR, and XGBoost to achieve better prediction accuracy. These base learners are briefly introduced in Secs. 3.1–3.3.

3.1 Lasso. Lasso was first introduced by Tibshirani [11]. As a shrinkage method, Lasso regularizes the coefficient for the least-squares estimates. For the data with the predictor variables $\mathbf{x}^i = (x_{i1}, \dots, x_{ip})$ and the response variables y_i , $i = 1, 2, \dots, N$, the objective function of Lasso can be written as

$$J = \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (3)$$

where β_0 and β_j represent the coefficients in the least squares estimates. λ is a tuning parameter. The first term of the objective function denotes the loss between the observed value and the fitted value. The second term of the objective function denotes a shrinkage penalty. The tuning parameter λ controls the shrinking effect. The advantage of Lasso over least squares is that the introduction of the penalty term can balance the trade-off between the variance and bias. In the hyperparameter tuning process, a total number of

100 different λ values were tested. The optimal value of λ was determined using 10-fold cross-validation.

3.2 Support Vector Regression. SVR is an instance-based machine learning algorithm for regression analysis [12]. SVR involves a margin function that separates hyperplanes. The region separated by hyperplanes is called the margin. The margin function can be written as

$$f(x, \omega) = \sum_{j=1}^m \omega_j g_j(x) + b \quad (4)$$

where $g_j(x)$ is the nonlinear transformation, ω_j is the weight for each transformation, and b is the intercept or bias term.

The margin is often referred to as the ε boundary. The ε -intensive loss function is calculated by the distance between these points and the predicted ε boundary. As a result, only response values that fall within the boundary can determine the fitted line while the outside data points have little impact on the result, which makes this algorithm robust and efficient. The ε -intensive loss function of SVR is described by

$$L_\varepsilon(y, f(x, \omega)) = \begin{cases} 0 & \text{if } |y - f(x, \omega)| \leq \varepsilon \\ |y - f(x, \omega)| - \varepsilon & \text{otherwise} \end{cases} \quad (5)$$

If a training dataset is not linearly separable, slack variables ξ_i and ξ_i^* ($i = 1, \dots, n$) are introduced into the loss function to form a soft margin. These two variables work as a tolerance term on the original ε boundary so that more data are included to decide the fitted line. The problem coming with such modification is a larger chance of overfitting, so the complexity of the model is further optimized by introducing the penalty terms. The norm value $\|\omega\|^2$ is for the weights regularization, and the box constraint C is introduced to penalize the slack variables. As a conclusion, the objective function of the SVR algorithm can be described as

$$J = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*)$$

Subject to

$$\begin{cases} y_i - f(x_i, \omega) \leq \varepsilon + \xi_i^* \\ f(x_i, \omega) - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n \end{cases} \quad (6)$$

To optimize the performance of SVR, the hyperparameters such as the boundary ε and the box constraint C in Eq. (6) need to be optimized. In this study, the values of the boundary ε ranged from 0.05 to 0.15 with a step of 0.01 and the values of the box constraint C ranged from 10 to 30 with a step of 2. The selection of hyperparameters and corresponding value ranges were discussed in detail in [13]. The grid search method was used to find the optimal values of the hyperparameters.

3.3 Extreme Gradient Boosting. Extreme gradient boosting (XGBoost), introduced by Chen et al. [14], is an enhanced gradient tree boosting method. XGBoost generates a group of trees sequentially, and the new tree is decided by the previous trees. It has the power to ensemble the weak tree structures in order to produce a strong model. The prediction output built with a total of K trees can be described as

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in F \quad (7)$$

The set F represents all trees generated in the model and the function f is the weight score function of each tree within F . The

objective function of the XGBoost algorithm at the t -th iteration is

$$J^t = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + \varepsilon f_t(x_i)) + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (8)$$

where $\hat{y}_i^{(t-1)}$ represents the i -th prediction result in the $(t-1)$ -th iteration. ε denotes the shrinkage factor that controls overfitting. T denotes the total number of leaves in a decision tree, and ω is the weight value of each leaf. γ and λ are corresponding restraint parameters. Equation (8) is formed with two parts. The first part $\sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + \varepsilon f_t(x_i))$ is the loss function. It measures how close the predicted value is from the observed data. The term of $\hat{y}_i^{(t-1)}$ helps to create new trees based on the evaluation of the previous trees. The second part $\gamma T + (1/2)\lambda \sum_{j=1}^T \omega_j^2$ is the regularization term, which is used to penalize the complexity of the model.

In the XGBoost algorithm, the loss function in Eq. (8) is replaced by the second-order Taylor expansion, and it has the form of $\sum_{i=1}^n [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)]$ where g_i and h_i correspond to the first- and second-order gradient statistics of the loss function, respectively. To decide the optimal split at each node, a gain function is calculated with

$$\text{Gain} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad (9)$$

where G_L and H_L represent the sum of the first- and second-order gradient at the left branch after splitting. G_R and H_R are for the right branch after splitting. For each iteration, the splitting node is selected to make the gain function maximum.

To optimize the performance of XGBoost, four different hyperparameters were tuned in this study. For the number of trees, 500, 1000, and 2000 were used in the grid search method. The depth of a tree describes the length from the root to the top in a tree, and it was set as 4, 5, and 6 in the grid search. The stopping criterion, which represents the minimum number of observations in each node, was tested using 1, 2, and 3 in this study. The shrinkage factor ε was set as 0.1 and 0.01 in the grid search method. The selection of the hyperparameters and the corresponding values were determined based on the study in [15].

4 Experiment

4.1 Specimen Preparation. In this study, the cooperative 3D printing experiments were conducted on a MakerGear M3 printer with two independent dual printheads. Each specimen was first divided into two portions by an inclined plane, and then the toolpath was modified so that the dual printheads can fabricate the specimen chunk by chunk. SIMPLIFY3D, a slicing software package, was used for converting the digital model in STL format into the commands in G-code format that can be read by the 3D printer. Figure 1 shows the shape and dimension of the specimen according to ASTM D638 [16]. The specimen has a 30 deg dividing plane.

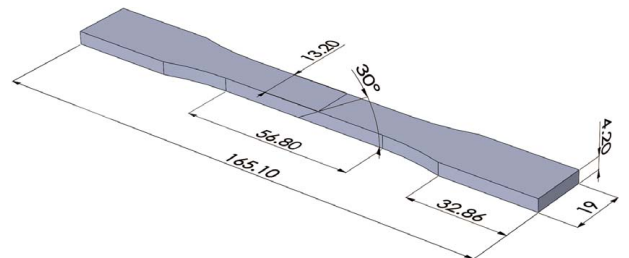


Fig. 1 Illustration of the specimen with a 30 deg dividing plane (unit in millimeters)

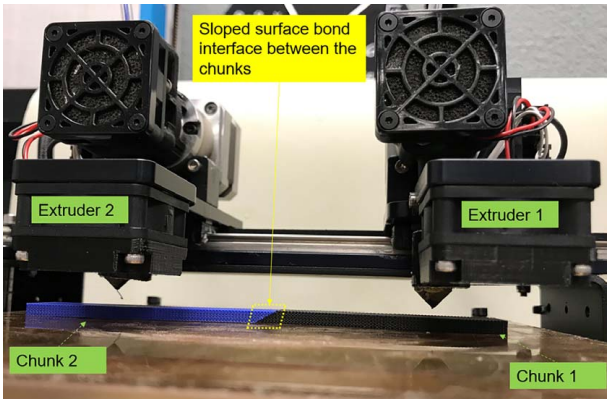


Fig. 2 Cooperative 3D printing process

Table 1 Parameter setting

Parameters	Values
Material	Poly lactide (PLA)
Nozzle temperature (°C)	195
Build plate temperature (°C)	60
First layer thickness (mm)	0.10
Subsequent layer thickness (mm)	0.20
Nozzle diameter (mm)	0.35
3D printing speed (mm/s)	58
Infill speed (mm/s)	46
Infill density	100%
Raster angle (deg)	45/−45
Air gap (mm)	0

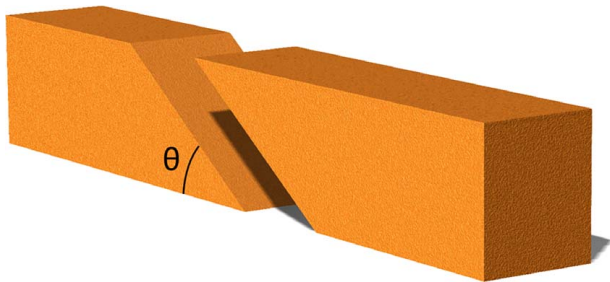


Fig. 3 Illustration of angle of incline

Figure 2 shows two chunks divided by an inclined plane. With a modified printing path, chunk-1 was first printed using extruder-1, and then chunk-2 was printed using extruder-2. In order to achieve cooperative printing using the conventional 3D printer, the toolpath for each chunk was generated individually using SIMPLIFY3D. The toolpath for each chunk was then combined together such that no collision between the printheads occurs. The G-code was modified in order to change the toolpath of the extruder-2 when it moved to the starting position in order to print the second chunk. Prior to uploading the G-code for printing, the modified toolpath was simulated in SIMPLIFY3D to ensure that there is no collision between the printed part and the extruder. Table 1 lists the 3D printing parameter setting for the experiments.

4.2 Design of Experiments. In the cooperative 3D printing process, each specimen was divided into different parts for multiple robots to print. Thus apart from the traditional 3D printing parameters, the factors introduced along with the cooperative printing process will also influence the tensile strength of the specimen. In this study, three parameters, the angle of incline, the overlapping length, and the number of shells, were studied for their impact on the tensile strength. As illustrated in Fig. 3, the angle of incline θ describes the slope angle of the plane that divides two parts. The design of the angle of incline will facilitate the deposition of material on the surface. Previous research [9] has shown that a smaller angle of incline is equivalent to a larger bonding surface, which makes the specimen stronger in tensile tests.

The overlapping length describes how much the layers from two parts overlap with each other at the bonding surface. Figure 4 shows both the top and lateral views of the bonding area with 0 and 0.3 mm overlapping length. As shown in Figs. 4(a) and 4(c), a zero overlapping length means that the layers from the left and right parts are fabricated exactly along the dividing plane. Figures 4(b) and 4(d) show a 0.3 mm overlapping length, meaning the layers are squeezed into the bonding surface to form a denser structure. According to the previous research [9], a more overlapped interface between two parts generally means more material is laid on the bonding area and thus a stronger bonding strength can be achieved.

In the traditional FDM printing process, adding shells to the specimens helps to increase the total strength. The existence of shells in the cooperative printing, however, shows a negative effect on the tensile strength according to the previous research [9]. Figure 5 shows the bonding surfaces of two specimens with zero and two shells.

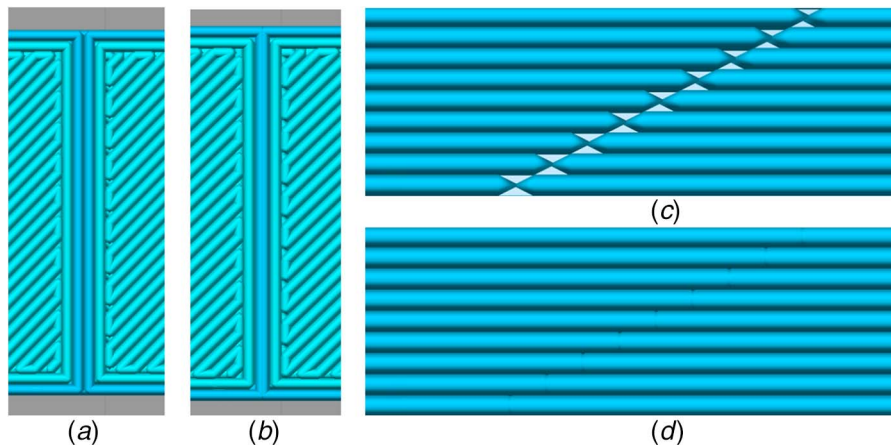


Fig. 4 Illustration of overlapping length: (a) top view of the bonding area with zero overlapping length, (b) top view of the bonding area with 0.3 mm overlapping length, (c) lateral view of the bonding area with zero overlapping length, and (d) lateral view of the bonding area with 0.3 mm overlapping length

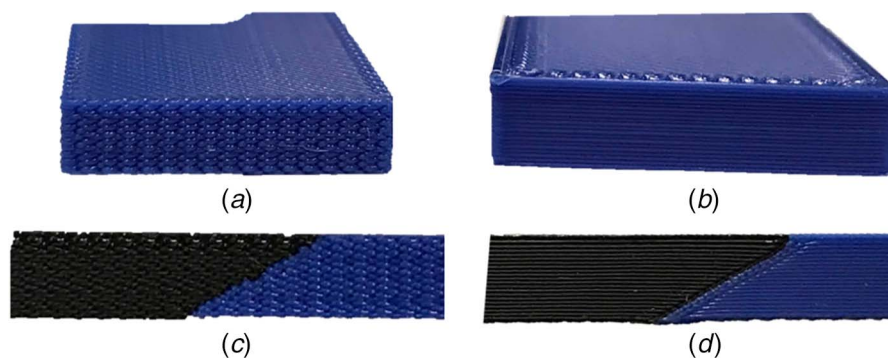


Fig. 5 Illustration of shells: (a) no shells, (b) two shells, (c) lateral view of bonding interface with no shells, and (d) lateral view of bonding interface with two shells

Some other factors that might affect the quality of 3D printing processes include raster orientation, infill density, air gap, and so on. The effects of these factors are not studied in this study. All of these parameters other than the three listed above were set as constant during the experiments.

To evaluate the effects of three selected 3D printing parameters on tensile strength, we designed an experiment as shown in Table 2. Each factor has four levels. The value of each level for each parameter was chosen based on the preliminary study in the previous work [9]. For each parameter combination in Table 2, three identical specimens were built to test the variations of the tensile strength. A total number of 192 data points were generated for model training and validation.

4.3 Tensile Testing. The tensile tests were conducted on an MTS[®] tensile testing machine with a 5KN load cell. A 2 mm/min loading rate was applied to the specimens. Figure 6 shows that the specimens broke at the bonding surface of two parts, which means that the tests result was indeed the strength at the bonding surface.

5 Results and Discussions

5.1 Tensile Strength. Figure 7 shows how the angle of incline affects the tensile strength when the overlapping length varies from 0 to 0.5 mm. Each tensile strength shown in Fig. 7 is the average value with a standard deviation of three repeated experiments. Figure 7 shows the tensile strength of the specimens with one and three shells. As indicated in Figs. 7(a)–7(d), the tensile strength tends to drop when the angle of incline increases. In Fig. 7(d),

when the overlapping length is 0.5 mm and the number of shells is one, specimens with a 30 deg angle of incline have an average tensile strength of 31.95 MPa, which is 106% higher than 15.51 MPa at the 45 deg angle of incline. Figure 7 also indicates that a smaller number of shells will make the tensile strength higher in most cases. As shown in Fig. 7(c), when the overlapping length is 0.3 mm and the angle of incline increases from 30 deg to 45 deg, the tensile strength for specimens with one shell is by average 7.8 MPa higher than those with three shells. According to the previous research [9], specimens with a smaller angle of incline have a larger contact area between two chunks, which makes the bonding surface stronger and thus a higher tensile strength is achieved.

Figure 8 shows the effect of the overlapping length on the tensile strength when the angle of incline varies from 30 deg to 45 deg. Each tensile strength shown in Fig. 8 is the average value with a standard deviation of three repeated experiments. Figure 8 shows the tensile strength of the specimens with one and three shells. The impact of the overlapping length on the tensile strength is not significant as shown in Fig. 8. In some cases, the specimens with larger overlapping lengths have larger tensile strength. As indicated in Fig. 8(c), when the angle of incline is 40 deg and the number of shells is one, the specimens with an overlapping length of 0.5 mm have an average tensile strength of 25.64 MPa, which is 65% higher than 15.57 MPa at the 0 overlapping length. However, for the specimens with 30 deg angle of incline and only one shell, the tensile strength decreases from 41.83 MPa to 31.95 MPa when the overlapping length increases from 0 to 0.5 mm. The reason why there is a weak correlation between the overlapping length and tensile strength is due to the uncertainty associated with the 3D printing process. Since each experiment in this work was repeated three times only, the high standard deviation as shown in Fig. 8 suggests that more repeated experiments should be performed in order to understand the impact of the overlapping length. The effect of the number of shells on the tensile strength can also be observed in Fig. 8. Almost all the specimens with only one shell have a larger tensile strength than those built with three shells.

Figure 9 shows the effect of the number of shells on the tensile strength when the overlapping length varies from 0 to 0.5 mm. Each tensile strength shown in Fig. 9 is the average value with a standard deviation of three repeated experiments. Figure 9 shows the tensile strength of the specimens with 30 deg and 40 deg angle of incline. Generally, when the number of shells increases, the tensile strength drops correspondingly. As shown in Fig. 9(b), the tensile strength for the specimens with 30 deg angle of incline decreases from 36.37 MPa to 24.83 MPa when the number of shells increases from 0 to 3. The curves shown in Fig. 9 also indicates that the impact of number of shells on the tensile strength is not as significant as the factor angle of incline, but more significant than the overlapping length. Based on the results shown in Figs. 7–9, the specimens fabricated with a smaller angle of incline, and no shells tend to have a larger

Table 2 Levels of three different parameters

Parameters	Level 1	Level 2	Level 3	Level 4
Angle of incline (deg)	30	35	40	45
Overlapping length (mm)	0	0.1	0.3	0.5
Number of shells	0	1	2	3



Fig. 6 Specimens after tensile tests

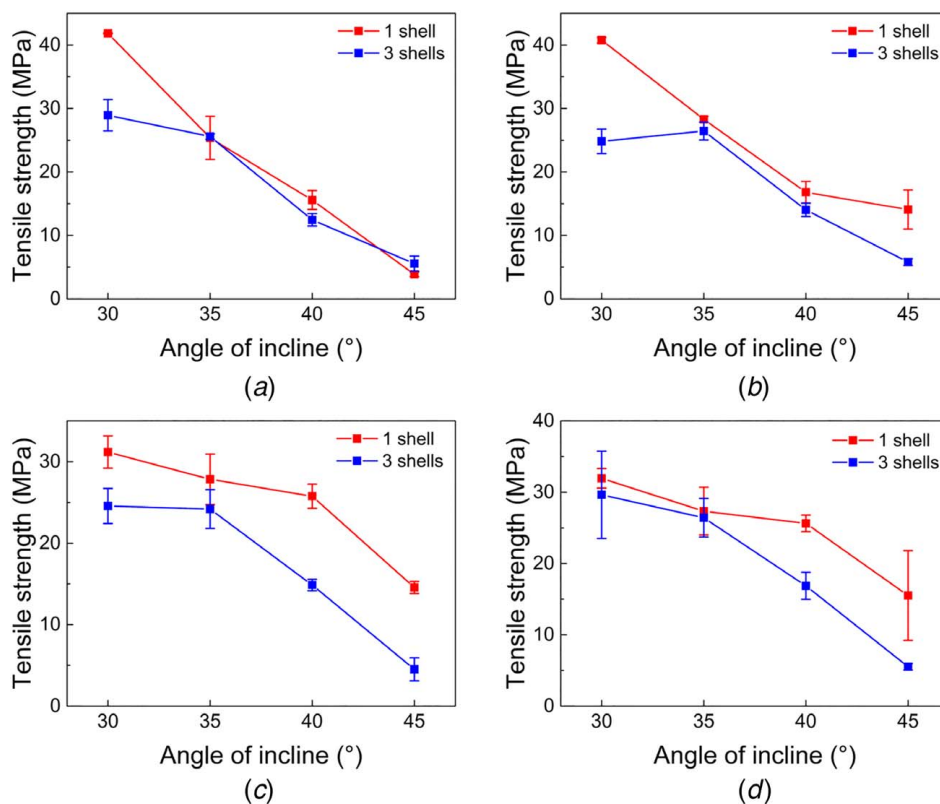


Fig. 7 Effect of angle of incline on tensile strength for specimens with different overlapping lengths: (a) 0 mm, (b) 0.1 mm, (c) 0.3 mm, and (d) 0.5 mm

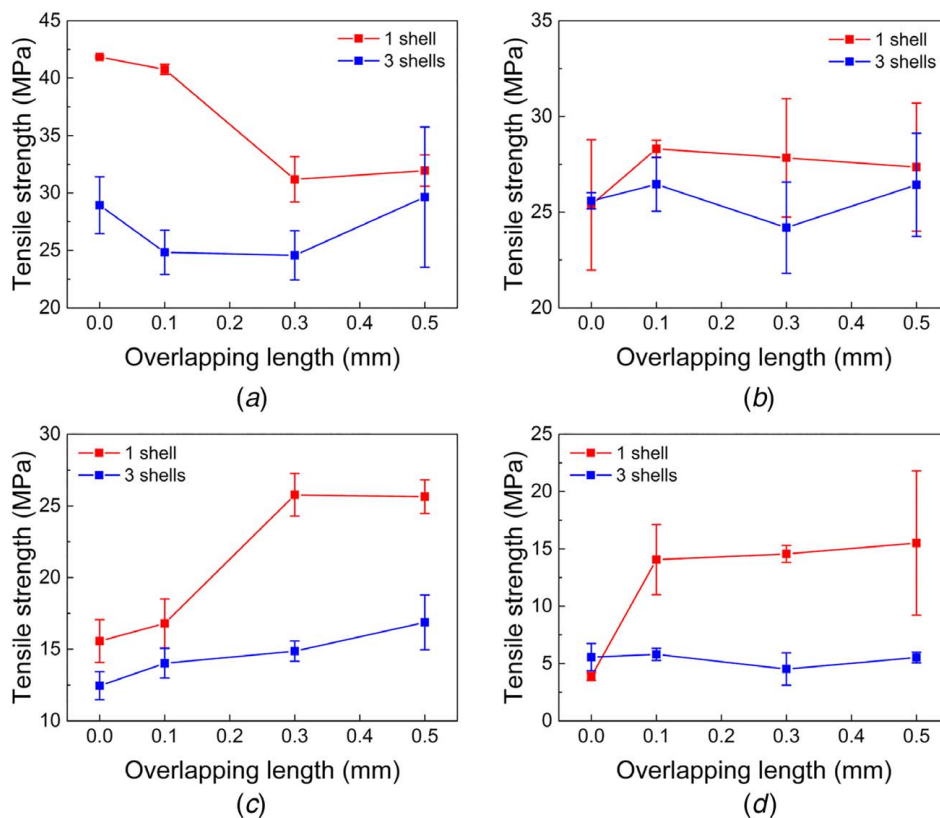


Fig. 8 Effect of overlapping length on tensile strength for specimens with different angle of incline: (a) 30 deg, (b) 35 deg, (c) 40 deg, and (d) 45 deg

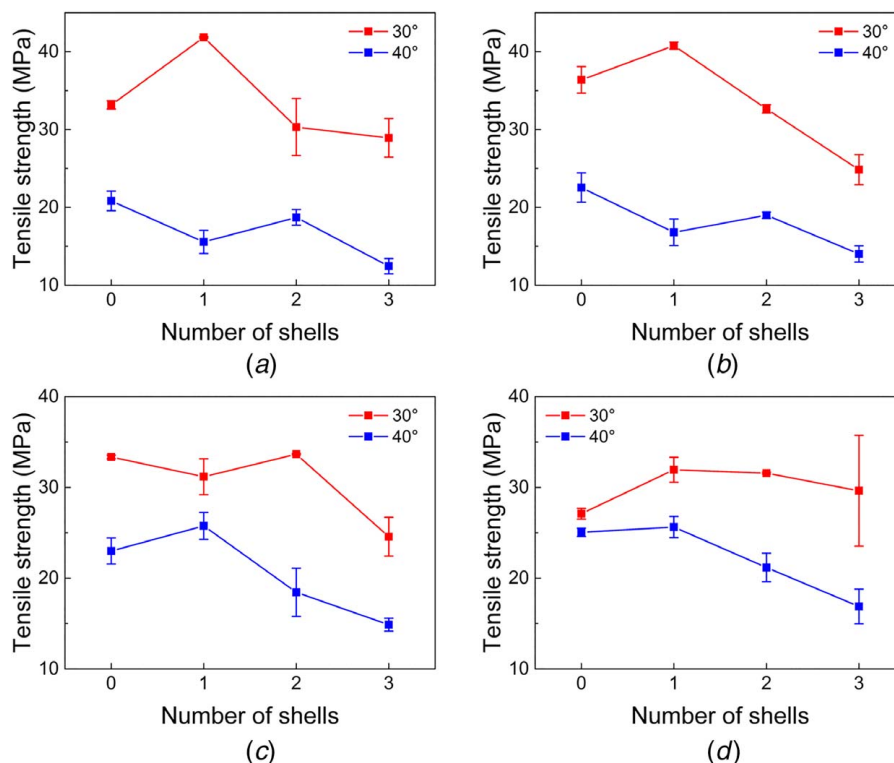


Fig. 9 Effect of number of shells on tensile strength for specimens with different overlapping length: (a) 0 mm, (b) 0.1 mm, (c) 0.3 mm, and (d) 0.5 mm

tensile strength. In addition, Figs. 7–9 show that the angle of incline is the most important factor to influence the tensile strength while the overlapping length is the least important parameter.

5.2 Performance Evaluation. Three different kinds of accuracy measurement were used to measure the discrepancy between the observed value and the model prediction. Root-mean-squared error (RMSE) is defined as $RMSE = \sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n}$ where y_i represents the observed value and \hat{y}_i represents the predicted value. The RMSE will be small if the predicted value is close to the observed value and will be large if the model fails to make a good prediction. Relative error (RE) is defined as $RE = (1/n) \sum_{i=1}^n |\hat{y}_i - y_i| / y_i$. RE describes the error relative to the size of the measurement. It falls between 0 and 1 considering how inaccurate the model is. Coefficient of determination (R^2) is defined as $R^2 = 1 - SS_{res} / SS_{tot}$. Here SS_{res} is the residual sum of squares, which has the form of $\sum_{i=1}^n (\hat{y}_i - y_i)^2$. SS_{tot} is the total sum of squares, and it can be expressed with $\sum_{i=1}^n (\bar{y}_i - y_i)^2$ where \bar{y}_i denotes the mean of the observed values. R^2 describes the percentage of the response variation that is explained by the model. A close to 100% R^2 means that the model can explain most of the variability.

The observed dataset contains 192 data points. For each data point, there is one output of tensile strength and three input features, which are the angle of incline, overlapping length, and the number of shells. In order to evaluate the model performance, the dataset is split into two parts, one for training and validation and the other for testing. A 10-fold cross-validation method is applied to the training set for the hyperparameter tuning and the base learner coefficients optimization. In this research, three different sizes of training set were selected to fit the model in order to evaluate the performance under different circumstances. As mentioned in Sec. 4.2, for each parameter combination in Table 2, three identical specimens were built to test the variations of the tensile strength. Thus, the 192 data points can be separated into 64 groups. Within each group,

there are three data points that have the same input features. 50%, 70%, and 90% of the 64 groups were randomly selected as the training set. The remaining 50%, 30%, and 10% of the experimental data were used as the test set.

Table 3 shows the RMSE, RE, and R^2 of the prediction model with different training data size. For the situation of 50% training data size, there are only 96 data points for training and a small RE of 15.246% is achieved with the prediction model. An 84.064% R^2 also indicates that most of the variability can be explained by the model. When the data size increases to 90%, the RE decreases to 12.974% while the R^2 increases to 95.945%. This result shows that the predictive model is more accurate as the training data size increases.

Table 4 lists the weights of each base learner with different training data size. The 10-fold cross-validation was used to determine the weights. When the training data size changes, the

Table 3 Prediction accuracy with different training data

Training data size	RMSE	RE (%)	R^2 (%)
50%	3.627	15.246	84.064
70%	4.339	15.558	82.684
90%	2.648	12.974	95.945

Table 4 Weights of the base learners for different training data

Training data size	Weights		
	XGBoost	Lasso	SVR
50%	0.383	0.106	0.511
70%	0.370	0.058	0.572
90%	0.238	0.095	0.667

hyperparameters tuned for each base learner will be different, and the weights for each base learner change correspondingly. The weights for Lasso are all below 0.11 in three different training data size. They are far less than the weights of the other two base learners, which indicate that Lasso is not as contributive as XGBoost and SVR in the ensemble learning model. Figures 10–12 show the relationship between the observed value and the predicted value in three different training data sizes. The dotted lines in Figs. 10–12 are 45 deg lines where the observed values match the predicted values. In all three cases, all the dots fall closely

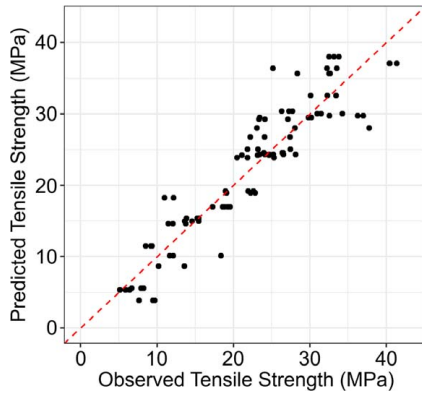


Fig. 10 Predicted versus observed tensile strength with 50% training data. The predictive model is built by ensemble learning combining Lasso, SVR, and XGBoost as base learners.

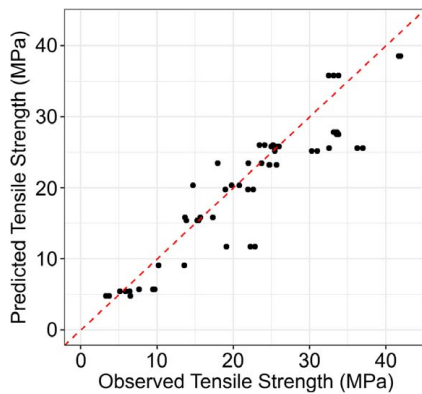


Fig. 11 Predicted versus observed tensile strength with 70% training data. The predictive model is built by ensemble learning combining Lasso, SVR, and XGBoost as base learners.

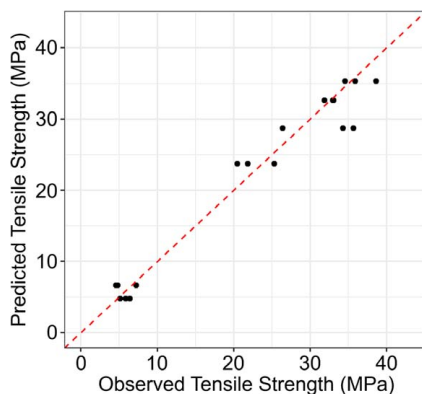


Fig. 12 Predicted versus observed tensile strength with 90% training data. The predictive model is built by ensemble learning combining Lasso, SVR, and XGBoost as base learners.

Table 5 Variable importance score

Features	Importance score by XGBoost	Importance score by SVR
Angle of incline	0.851	0.740
Number of shells	0.083	0.166
Overlapping length	0.066	0.094

toward these lines, which indicates that the predictions agree with the observed values. All the results shown in Table 3 and Figs. 10–12 demonstrate that the ensemble learning method is effective in making predictions accurate.

Table 5 shows the variable importance scores of the three factors determined by XGBoost and SVR, respectively. For XGBoost, the variable importance score is calculated based on how often the variable is selected as a splitting feature in the tree structure. For SVR, the variable importance score is determined based on how much the output varies as the input varies [17]. As shown in Table 5, the rankings of the variable importance scores calculated by XGBoost and SVR are consistent. The angle of incline and the overlapping length are the most and least important factors that affect the tensile strength of the specimens, respectively. The numerical result has good agreement with the experiment result shown in Figs. 7–9. Therefore, if the parts fabricated by the cooperative 3D printing process have a smaller angle of incline, then the tensile strength of the parts is greater. In addition, the parts with fewer shells have greater tensile strength. The effect of the overlapping length on tensile strength is not significant.

5.3 Analysis of Variance. To compare the ensemble learning method with traditional statistical methods, the variable importance and prediction accuracy were also evaluated using the ANOVA and regression analysis. The ANOVA statistics for the regression model are listed in Table 6. The degree of freedom (DF) represents the amount of information used for analysis. The adjusted sum of squares (Adj SS) for a variable measures the variation of the response data that can be explained by this variable. The Adj SS for error measures the variation of the response data that cannot be explained by all variables. The adjusted mean squares (Adj MS) takes into account the degrees of freedom when measuring the variations. The F -value is the ratio between the Adj MS for a variable and the Adj MS for error. The F -value measures the statistical significance of a variable. Based on the F -values shown in Table 6, the angle of incline is the most significant variable that affects the tensile strength while the overlapping length is the least significant variable.

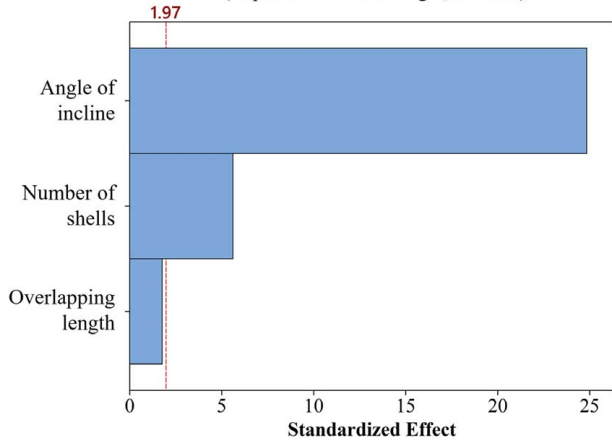
The simple linear regression analysis is performed to understand the relationship between the variables (i.e., angle of incline, overlapping length, and number of shells) and the response (i.e., tensile strength). The parameters of the regression analysis are listed in Table 7. The coefficient indicates how the variable affects the response. The T -value is the ratio of the coefficient and the corresponding standard error. The P -value measures the probability against the null hypothesis. Both T and P values describe the significance of the variable. The variable significance based on the T and P values shown in Table 7 is in good agreement

Table 6 Analysis of variance

Source	DF	Adj SS	Adj MS	F -value
Angle of incline	1	12,649.8	12,649.8	616.84
Overlapping length	1	64.0	64.0	3.12
Number of shells	1	646.0	646.0	31.50
Error	188	3855.4	20.5	N/A

Table 7 Regression analysis parameters

Term	Coefficient	T-value	P-value	VIF
Constant	78.32	34.17	<0.001	N/A
Angle of incline	-1.4520	-24.84	<0.001	1.00
Overlapping length	3.01	1.77	0.079	1.00
Number of shells	-1.641	-5.61	<0.001	1.00

Pareto Chart of the Standardized Effects(response is tensile strength, $\alpha = 0.05$)**Fig. 13 Pareto chart of the standardized effects on the tensile strength**

with the ANOVA results in Table 6. Specifically, the angle of incline and the overlapping length are the most and least significant variables, respectively. The variance inflation factor (VIF) quantifies the multicollinearity within a model. A VIF larger than five suggests the existence of multicollinearity in the analysis. Although a high level of multicollinearity might not weaken the prediction performance in most cases, it will interfere in determining the precise effect of each variable. The coefficient estimates with a high VIF also become unstable as the result of increasing variance. In this study, including the quadratic and interaction terms in the model makes most of VIFs larger than 100, suggesting a very high level of multicollinearity. Therefore, only simple linear regression was selected in the analysis.

The Pareto chart as shown in Fig. 13 illustrates the standardized effects of three variables on the tensile strength. The standardized effects are based on the T -values shown in Table 7. The reference line marked with 1.97 suggests whether the effect is statistically significant or not. In comparison with the machine learning method, both ANOVA and regression analysis make the same conclusion about the variable significance. The linear regression model based on the regression analysis can be expressed as

$$\begin{aligned} \text{Tensile strength} = & 78.32 - 1.4520 \times \text{Angle of incline} \\ & + 3.01 \times \text{Overlapping length} - 1.641 \\ & \times \text{Number of shells} \end{aligned} \quad (10)$$

The ability to provide interpretable results with an explicit solution is the advantage of traditional statistical methods over machine learning.

The prediction accuracy using the traditional linear regression method was also evaluated. As mentioned in Sec. 5.2, the 192 data points were separated into 64 groups. Within each group, there are three data points that have the same input features. 50%, 70% and 90% of the 64 groups were selected as the training set to fit the linear regression model. The remaining 50%, 30%, and

Table 8 Prediction accuracy for different training data size using linear regression method

Training data size	RMSE	RE (%)	R^2 (%)
50%	5.28	26.64	74.67
70%	4.19	24.92	81.14
90%	4.76	26.47	86.90

10% of the experimental data were used as the test set to evaluate the model performance. The selected data for each condition was kept the same with those used in machine learning method so that the performance comparison is valid. The prediction accuracy of the linear regression method is shown in Table 8. When 90% data were used for training dataset, the RE increases about 104.02% and the R^2 decreases about 9.43% comparing with the machine learning method. Similarly, the RE increases and the R^2 decreases for both 70% and 50% training data size conditions. Based on the prediction accuracy shown in Tables 3 and 8, the ensemble learning method outperformed the linear regression model. For example, the RMSEs of the ensemble learning model are 3.627, 4.339, and 2.648, while the RMSEs of the linear regression model are 5.28, 4.19, and 4.76 when 50%, 70%, and 90% of the total data are used as training data.

6 Conclusions

In this paper, we developed a data-driven predictive modeling approach to the prediction of the tensile strength of components fabricated by the cooperative 3D printing process. The data-driven predictive modeling approach combines three base learners, including Lasso, SVR, and XGBoost, in order to improve the accuracy of the predictive model of the tensile strength. A linear regression model was also built to conduct a comparative study. The experimental results have shown that the ensemble learning model achieved much higher prediction accuracy than the linear regression model. In addition, the effects of three 3D printing process parameters (i.e., angle of incline, overlapping length, and number of shells) on tensile strength were evaluated using XGBoost and SVR, as well as ANOVA. The results have shown that the effect of the angle of incline on tensile strength is significant. We also found that the effects of three 3D printing process parameters on tensile strength revealed by the machine learning algorithms and ANOVA are consistent.

References

- [1] Li, L., Tirado, A., Nlebedim, I., Rios, O., Post, B., Kunc, V., Lowden, R., Lara-Curzio, E., Fredette, R., and Ormerod, J., 2016, "Big Area Additive Manufacturing of High Performance Bonded NdFeB Magnets," *Sci. Rep.*, **6**(1), p. 36212.
- [2] Duty, C. E., Kunc, V., Compton, B., Post, B., Erdman, D., Smith, R., Lind, R., Lloyd, P., and Love, L., 2017, "Structure and Mechanical Behavior of Big Area Additive Manufacturing (BAAM) Materials," *Rapid Prototyping J.*, **23**(1), pp. 181–189.
- [3] McPherson, J., and Zhou, W., 2018, "A Chunk-Based Slicer for Cooperative 3D Printing," *Rapid Prototyping J.*, **24**(9), pp. 1436–1446.
- [4] Melenka, G. W., Cheung, B. K., Schofield, J. S., Dawson, M. R., and Carey, J. P., 2016, "Evaluation and Prediction of the Tensile Properties of Continuous Fiber-Reinforced 3D Printed Structures," *Compos. Struct.*, **153**, pp. 866–875.
- [5] Hayes, B. J., Martin, B. W., Welk, B., Kuhr, S. J., Ales, T. K., Brice, D. A., Ghamarian, I., Baker, A. H., Haden, C. V., and Harlow, D. G., 2017, "Predicting Tensile Properties of Ti-6Al-4V Produced Via Directed Energy Deposition," *Acta Mater.*, **133**, pp. 120–133.
- [6] Tapia, G., Elwany, A., and Sang, H., 2016, "Prediction of Porosity in Metal-Based Additive Manufacturing Using Spatial Gaussian Process Models," *Addit. Manuf.*, **12**, pp. 282–290.
- [7] Campoli, G., Borleffs, M., Yavari, S. A., Wauthle, R., Weinans, H., and Zadpoor, A. A., 2013, "Mechanical Properties of Open-Cell Metallic Biomaterials Manufactured Using Additive Manufacturing," *Mater. Des.*, **49**, pp. 957–965.

- [8] Mukherjee, T., Zhang, W., and DebRoy, T., 2017, "An Improved Prediction of Residual Stresses and Distortion in Additive Manufacturing," *Comput. Mater. Sci.*, **126**, pp. 360–372.
- [9] Poudel, L., Sha, Z., and Zhou, W., 2018, "Mechanical Strength of Chunk-Based Printed Parts for Cooperative 3D Printing," *Proc. Manuf.*, **26**, pp. 962–972.
- [10] Hunt, G., Mitzalis, F., Alhinai, T., Hooper, P. A., and Kovac, M., 2014, "3D Printing with Flying Robots," Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, May 31–June 7, IEEE, New York, pp. 4493–4499.
- [11] Tibshirani, R., 1996, "Regression Shrinkage and Selection via the Lasso," *J. Royal Stat. Soc.: Series B (Methodological)*, **58**(1), pp. 267–288.
- [12] Vapnik, V., 2013, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, New York.
- [13] Duan, K., Keerthi, S. S., and Poo, A. N., 2003, "Evaluation of Simple Performance Measures for Tuning SVM Hyperparameters," *Neurocomputing*, **51**, pp. 41–59.
- [14] Chen, T., and Guestrin, C., 2016, "Xgboost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, Aug. 13–17, ACM, pp. 785–794.
- [15] Probst, P., Bischl, B., and Boulesteix, A.-L., 2019, "Tunability: Importance of Hyperparameters of Machine Learning Algorithms," *J. Mach. Learn. Res.*, **20**(53), pp. 1–32.
- [16] International, A., 2015, *ASTM D638-14, Standard Test Method for Tensile Properties of Plastics*, ASTM International.
- [17] Cortez, P., and Embrechts, M. J., 2013, "Using Sensitivity Analysis and Visualization Techniques to Open Black box Data Mining Models," *Inf. Sci.*, **225**, pp. 1–17.