GENERATING SYNTHETIC DATA FOR MECHANICAL PROPERTIES OF 3D PRINTED SPECIMENS USING GENERATIVE ADVERSARIAL NETWORKS

Alexandru Constantin STANCIU [1], Mihai BUTOLO[2], Nicolae GOGA[3], Florin BACIU[4], Miruna CIOLCA[5*], Anton HADAR[678]

This article explores the application of Generative Adversarial Networks (GANs) for creating synthetic data representing mechanical properties of 3D printed specimens. With the exponential growth of additive manufacturing in industrial production, understanding and optimizing process parameters that influence mechanical properties has become crucial. However, obtaining sufficient experimental data through physical testing is costly and time-consuming. This research demonstrates how conditional GANs can effectively generate synthetic data that maintains the statistical properties and relationships found in real experimental data. By expanding limited datasets, this approach enables better prediction models while reducing the need for extensive laboratory testing.

Keywords: Generative Adversarial Networks, synthetic data generation, additive manufacturing, mechanical properties, 3D printing, conditional GAN

1. Introduction

Additive Manufacturing (AM), commonly known as 3D printing, has experienced exponential growth in recent years across industrial production, educational applications, and hobby use. This technology offers the flexibility to produce parts and prototypes at reduced costs compared to traditional subtractive manufacturing methods. From aerospace and automotive components to

[1] PhD Student, Faculty of Industrial Engineering and Robotics, National University of Science and Technology POLITEHNICA Bucharest, Romania, e-mail: stancialexandru04@yahoo.com.
[2] PhD Student, Faculty of Industrial Engineering and Robotics, National University of Science and Technology POLITEHNICA Bucharest, Romania, e-mail: mihai.butolo@gmail.com.
[3] Professor, Faculty of Industrial Engineering in Foreign Languages, National University of Science and Technology POLITEHNICA Bucharest, Romania, e-mail: nicu.goga@upb.ro.
[4] Assoc. Professor, Department of Strength of Materials, Faculty of Industrial Engineering and Robotics, National University of Science and Technology POLITEHNICA Bucharest, Romania, e-mail: florin.baciu@upb.ro.
[5] Lecturer, Department of Strength of Materials, Faculty of Industrial Engineering and Robotics, National University of Science and Technology POLITEHNICA Bucharest, Romania, e-mail: miruna.ciolca@upb.ro. * Corresponding author.
[6] Professor, Department of Strength of Materials, Faculty of Industrial Engineering and Robotics, National University of Science and Technology POLITEHNICA Bucharest, Romania, e-mail: anton.hadar@upb.ro.
[7] Technical Science Academy of Romania, Dacia Blvd. 26, 030167 Bucharest, Romania.
[8] Academy of Romanian Scientists, Ilfov Street 3, 050045 Bucharest, Romania.

architectural designs [1] and complex medical models [2], 3D printing has transformed innovation and production processes.

The major advantages of this technology are substantial. Additive manufacturing enables the creation of complex geometric shapes that would be difficult or impossible to achieve through traditional methods. It significantly reduces product development time while offering the possibility of customizing each product without substantial additional costs [3]. Furthermore, it minimizes material waste since material is only added where necessary, in contrast to subtractive methods that remove material from a larger block.

Despite these advantages, a significant challenge remains in predicting and ensuring the mechanical properties of printed parts. Numerous factors influence these properties: material type, print settings (temperature, speed, infill density, path geometry), and post-processing conditions [4,5]. For polymeric materials, the internal structure and infill density significantly impact the elastic modulus, tensile strength, and yield strength. Industry standards require testing on standard specimens following ASTM or ISO guidelines [6]. However, obtaining sufficient experimental data to build robust prediction models is expensive and time-consuming. Each combination of material, printing parameters, and testing conditions requires producing multiple samples and conducting mechanical tests.

1.1. The Problem of Limited Experimental Data

In mechanical engineering and additive manufacturing, obtaining large volumes of experimental data presents significant challenges. Researchers face high costs associated with direct material expenses, equipment usage, and specialized personnel. The time-intensive nature of manufacturing specimens and conducting tests further compounds these difficulties. Moreover, inherent variations in manufacturing and testing processes necessitate multiple specimens to achieve statistically relevant results, adding another layer of complexity to data collection efforts. From a machine learning perspective, this limitation can be addressed through synthetic data generation techniques. A fundamental principle is that prediction model performance increases with more training data. Although obtaining large real datasets in engineering fields is often difficult due to laboratory procedure complexity and high equipment costs, synthetic data generation techniques can supplement the real sample.

1.2. Generative Adversarial Networks as a Solution

Generative Adversarial Networks (GANs) represent an innovative approach to synthetic data generation. By training two neural networks simultaneously – a Generator and a Discriminator – new data is created that closely resembles real data. The Generator learns to produce "fake" but realistic samples, while the Discriminator attempts to distinguish between real and fake data. The competition between these networks leads to continuous improvement in the quality of generated data. Compared to classical data augmentation methods, GANs excel at capturing complex relationships between variables. For 3D printed parts, this means the generated data respects both physical constraints and parameter correlations. While physical tests remain necessary, synthetic data can reduce the number of experiments, help explore more parameter combinations and improve prediction models.

2. Related Work

Building upon the motivation for synthetic data described in section 1.1, researchers have developed increasingly sophisticated methods to address data limitations in engineering fields. Early approaches focused primarily on statistical methods such as bootstrapping, perturbation techniques, and simplified mathematical models to expand limited datasets. As computational capabilities advanced, physics-based simulation methods gained prominence, allowing researchers to model complex engineering phenomena based on first principles and generate synthetic data that complied with known physical laws.

The most recent paradigm shift has been toward machine learning-based generative models. Lin et al. highlighted the value of synthetic data generation for "sharing networked time series data" while addressing privacy considerations and practical constraints in data collection [7]. This transition marks an important evolution from rule-based generation to data-driven approaches that can capture complex, non-linear relationships in engineering data without requiring explicit formulation of all physical constraints.

2.1. GANs in Materials Science and Manufacturing

Since their introduction, GANs have been applied to numerous engineering domains beyond the general synthetic data generation approach described in section 1.2. In materials science, GANs have demonstrated success in microstructure modeling. Iyer et al. explored conditional GANs for capturing complex, multiphase microstructure images with applications to materials design [8]. Extending this concept to three dimensions, Hsu et al. implemented a GAN framework to learn and generate 3D microstructures of solid oxide fuel cell electrodes that were visually, statistically, and topologically realistic [9].

For manufacturing process-property relationships, which are central to our research, several notable applications have emerged. Xie et al. developed a "mechanistic data-driven framework integrating wavelet transforms and convolutional neural networks to predict location-dependent mechanical properties over fabricated parts based on process-induced temperature sequences" in metal additive manufacturing [10]. Their approach demonstrated how advanced machine learning techniques could create predictive models from sparse experimental data but focused primarily on metals rather than polymers and on prediction rather than synthetic data generation.

2.2. Specialized GAN Architectures for Structured Data

While standard GAN architecture works well for unstructured data like images, engineering applications often involve structured data with specific formats. For tabular data containing both categorical and continuous variables—like our dataset with categorical printing parameters and continuous mechanical properties—Xu et al. introduced CTGAN (Conditional Tabular GAN) [11]. Their architecture addresses the unique challenges of mixed-type data through mode-specific normalization for continuous features and conditional generation based on discrete variables.

For the time series data, which captures process dynamics in manufacturing, specialized GAN variants have been developed. TimeGAN by Yoon et al. [12] combines unsupervised adversarial training with supervised learning to better capture temporal dynamics. Zhang et al. [13] demonstrated another approach for time series synthesis in smart grid applications, using a modified GAN architecture to generate realistic load profiles while preserving temporal correlations. These specialized architectures demonstrate the versatility of the GAN framework and its adaptability to various data structures relevant to engineering applications.

2.3. Validation Approaches for Synthetic Engineering Data

A critical aspect of synthetic data generation – particularly for engineering applications where data will inform real-world designs – is rigorous validation. Beyond the statistical comparisons that will be presented in our results section, the literature describes several complementary approaches. Statistical validation methods typically compare distributions using metrics such as Kolmogorov-Smirnov tests, Jensen-Shannon divergence, and quantile-quantile plots to verify that synthetic data maintains the distributional characteristics of the original dataset.

Functional validation approaches go beyond statistical similarity to evaluate how synthetic data performs in downstream applications. This includes comparing the predictive accuracy of models trained on synthetic versus real data or assessing how design decisions made using synthetic data would compare to those made with real data. These validation approaches are essential for building confidence in synthetic data, especially in engineering contexts where physical safety and performance are at stake.

2.4. Research Gap and Contribution

Recent advancements have also explored the application of GANs for computationally efficient design in additive manufacturing. Hertlein et al. (2021) developed a conditional GAN framework that addresses the computational burden of repeated topology optimization during early-stage design for AM. Their approach learns latent similarities across different parameter combinations, including build orientations and loading conditions, enabling rapid generation of near-optimal designs that respect manufacturing constraints like overhangs. Their research demonstrated that the cGAN could not only predict viable designs efficiently but, interestingly, in approximately 9% of test cases, even generated structures with better compliance than traditional topology optimizations, sometimes improving performance by as much as 50%. This demonstrates that GANs not only serve as tools for synthetic data generation but can actively contribute to design innovation through their ability to explore the parameter space in ways that traditional methods might overlook [14]. The potential of such approaches for exploring design alternatives further underscores the importance of our research into synthetic data generation for improving mechanical property prediction in AM applications.

Despite the advances in GAN-based synthetic data generation, significant gaps remain in their application to additive manufacturing of polymers. Most existing research has focused on metal-based processes, with their distinct material behaviors and process-property relationships. The few studies addressing polymers have typically focused on single properties rather than capturing multiple mechanical responses simultaneously.

Additionally, while conditional GANs have shown promise for tabular data in general, their application to 3D printing parameter-property relationships has been limited. The specific challenge of generating synthetic mechanical property data that respects both the statistical distributions of individual properties and the complex relationships between printing parameters and multiple mechanical outcomes remain underexplored.

Our research addresses these gaps by implementing a conditional GAN specifically designed for polymer-based 3D printed components, capturing the complex relationships between printing parameters (infill pattern, density, and speed) and multiple mechanical properties (elastic modulus, tensile strength, and yield strength) simultaneously. By generating high-quality synthetic data that maintains both statistical and physical validity, this work provides a valuable methodology for expanding limited experimental datasets in additive manufacturing while preserving the essential characteristics needed for reliable prediction models.

3. Overview of Generative Adversarial Networks

Generative Adversarial Networks (GANs) represent one of the most significant innovations in artificial intelligence and machine learning in recent years. Introduced by Ian Goodfellow and collaborators in 2014 and described as a minimax two-player game [15], GANs revolutionized the generation of synthetic samples that mimic real data distributions.

3.1. Basic Principles: Generator vs. Discriminator

GANs consist of two competing neural networks: a Generator (G) that creates synthetic samples from random noise vectors, and a Discriminator (D) that classifies samples as real or synthetic. In conditional GANs, the Generator also incorporates categorical variables (e.g., infill density, printing speed) to produce targeted outputs like mechanical properties. The Discriminator evaluates both continuous and categorical inputs, outputting a probability indicating authenticity. During training, these networks engage in adversarial optimization - the Generator improves its synthetic samples based on the Discriminator's feedback, while the Discriminator continuously refines its classification criteria. This iterative process converges when the Generator produces high-fidelity samples indistinguishable from real data, effectively modeling the underlying data distribution.

3.2. Types of GANs Relevant to Engineering Applications

Several GAN variants have been developed since Goodfellow's original architecture, each offering specific advantages for different applications and data types. Below are presented some of the most relevant variants for engineering applications:

1) Conditional GAN (cGAN): incorporates conditional variables into both Generator and Discriminator inputs, enabling controlled generation of specific sample types. This approach excels with mixed tabular data and has been used to augment fatigue testing data for ferrous materials, improving ML model prediction accuracy (R^2) by 0.3-0.6 [16].

2) Wasserstein GAN (WGAN): employs Wasserstein distance to enhance training stability and prevent mode collapse. It has been successfully applied to model manufacturing form defects in cylindrical surfaces during turning processes [17].

3) Deep Convolutional GAN (DCGAN): leverages convolutional layers primarily for visual data generation and has achieved 98.43% accuracy in detecting water wall defects in thermal power plants using seam carving algorithms [18].

For our research, we selected a conditional GAN approach given the mixed nature of our data (categorical printing parameters and continuous mechanical properties) and the need to maintain relationships between these variables.

3.3. Advantages and Limitations in Engineering Applications

GANs offer key advantages in engineering applications such as supplementing limited experimental datasets [19,20], exploring untested parameter combinations, and reducing expensive physical testing costs. However, significant limitations exist. Training stability remains challenging, requiring careful hyperparameter tuning. We tested latent vector sizes (50-200), learning rates (10^{-4}-10^{-3}), Adam betas, and hidden layer configurations before achieving stable convergence with three hidden layers, batch normalization, dropout, gradient clipping, and learning rate schedulers. Generated data quality depends entirely on the real dataset's representativeness, propagating any existing flaws or biases. Moreover, synthetic data cannot replace experimental validation in safety-critical applications where physical testing remains essential.

4. Experimental Data and Methodology

4.1. Description of the Initial Dataset

The initial dataset consisted of 45 3D printed specimens using FDM technology with PLA material. The specimens were systematically designed to investigate the influence of process parameters on mechanical properties. Three main factors were varied: infill density (100%, 80%, and 60%), infill pattern (diagonal or hexagonal), and printing speed (40, 60, 80, and 100 mm/s). For each parameter combination, two specimens were printed according to a full factorial experimental plan. The resulting samples were tested under standard tensile loading using a universal testing machine (Instron 8872), in order to determine the elastic modulus, yield strength, and ultimate tensile strength.

Our base dataset (Table 1) contained several continuous properties that characterized the mechanical performance of each specimen. The longitudinal elastic modulus (measured in MPa) indicated material stiffness and resistance to deformation. Yield strength (MPa) represented the stress at which plastic deformation begins, while tensile strength (MPa) measured the maximum stress the material could withstand before failure.

Accompanying these mechanical properties, we took into consideration printing parameters that defined how each specimen was manufactured (infill density, infill pattern and printing speed).

Table 1

Sample of the initial dataset (5 specimens from the CSV file)

Code	Infill Density (%)	Infill Pattern	Printing Speed (mm/s)	Elastic Modulus (MPa)	Yield Strength (MPa)	Tensile Strength (MPa)
100G40_1	100	diagonal	40	2031.37	25.47	27.11
100G60_1	100	diagonal	60	2123.45	27.28	28.71
100G80_1	100	diagonal	80	2094.00	27.06	28.60
80G40_1	80	diagonal	40	1987.95	23.84	25.32
80G60_1	80	diagonal	60	2054.67	24.83	26.12

The dataset contained 45 specimens, each representing a unique combination of these parameters. This limited sample size presented challenges for developing robust prediction models, which typically require hundreds of examples for effective training.

4.2. Exploratory Data Analysis

Before implementing the GAN, we conducted comprehensive exploratory data analysis to understand the characteristics of our dataset. Our statistical analysis revealed several important characteristics of the dataset. Descriptive statistics indicated that the elastic modulus averaged approximately 1661.67 MPa with a standard deviation of 534.84 MPa, while yield strength averaged 20.33 MPa (SD = 6.02) and tensile strength averaged 21.69 MPa (SD = 6.12). Distribution analysis through histograms showed slightly right-skewed distributions for the elastic modulus, and boxplots confirmed there were no significant outliers in the mechanical properties.

The correlation analysis exposed interesting relationships between mechanical properties. We observed strong positive correlations between yield strength and tensile strength ($r > 0.9$), suggesting that these properties increase proportionally. Moderate correlations also existed between elastic modulus and both strength measures, indicating partial interdependence between stiffness and strength characteristics.

When examining relationships with categorical variables, we found that higher infill densities consistently corresponded with increased mechanical properties across all measures. The influence of infill pattern varies depending on infill percentage, with more pronounced effects at lower densities. Printing speed demonstrated complex relationships with mechanical properties, sometimes showing inverse relationships at certain infill densities while having minimal impact on others. This analysis provided critical insights for GAN configuration and later helped validate the quality of synthetic data by ensuring these relationships were preserved.

4.3. GAN Architecture and Implementation

Our research employed conditional GAN architecture (Fig. 1) implemented in PyTorch, comprising several specialized components designed to handle both categorical and continuous variables in the context of additive manufacturing.

The implementation began with several preprocessing steps to prepare the data for the GAN. We applied one-hot encoding to transform categorical variables (infill pattern, density, and printing speed) into numerical format suitable for neural network processing. Continuous variables were normalized to mean 0 and standard deviation 1 to facilitate training convergence. These processed features were then combined into a single tensor dataset with appropriate batch sizing for efficient training.

For the Generator architecture, we designed a network that accepts a concatenated input of a random noise vector (dimension 10) and the one-hot encoded categorical variables. This combined input passes through three fully connected hidden layers with ReLU activations, with dimensions strategically decreasing through the network. The output layer produces three continuous variables corresponding to the normalized mechanical properties we aim to generate. The Discriminator architecture mirrors this approach but with different inputs. It accepts the concatenated continuous variables and one-hot encoded categorical variables, processing them through three fully connected hidden layers with ReLU activations. The final output is a single sigmoid-activated value representing the probability that the input data is real rather than generated.

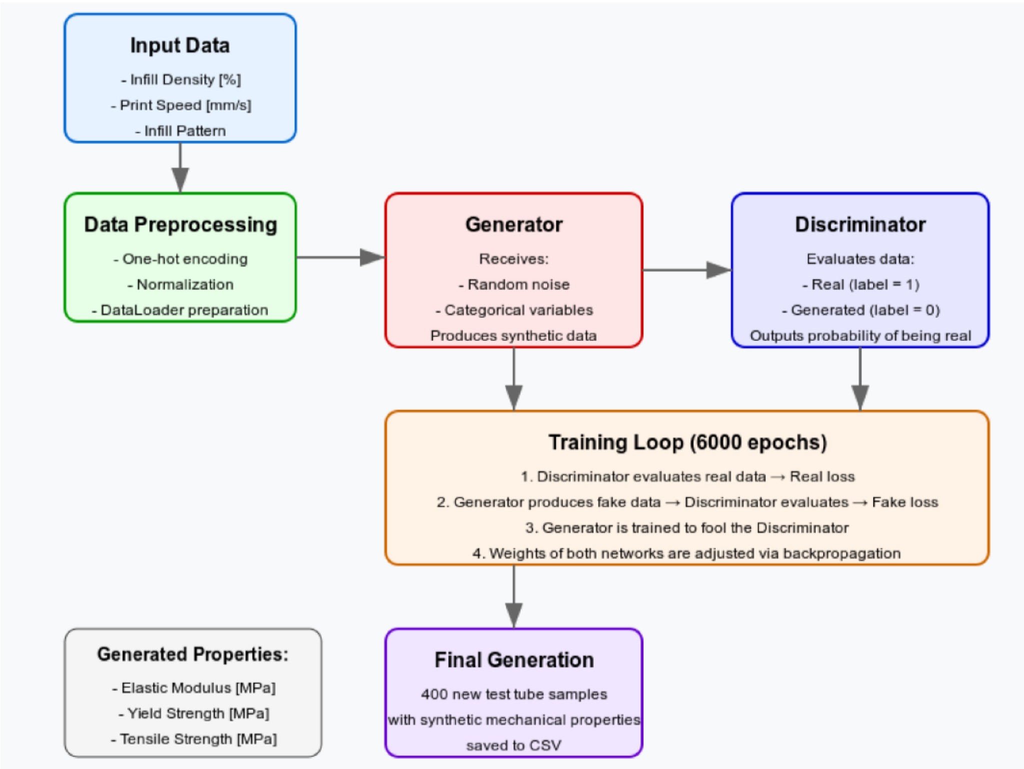

Fig. 1. Visualizing the GAN architecture for generating synthetic data for mechanical properties

Our training configuration was carefully designed to address the challenges of a small dataset. We selected Binary Cross Entropy (BCELoss) as the loss function due to its effectiveness for binary classification problems. We employed Adam optimizers with a learning rate of 0.001 for both networks. The batch size was set to 5, constrained by our limited dataset size, and we ran the training for 6000 epochs to ensure proper convergence despite the small amount of training data.

The training process followed the standard adversarial approach with alternating updates to each network. In each iteration, we first trained the Discriminator on a batch of real data with the target of labeling these samples as real. We then generated fake data using the current state of the Generator and trained the Discriminator to label these samples as fake. Following this, we trained the Generator with the objective of producing data that the Discriminator would label as real. This adversarial process continued for the full 6000 epoch, with the two networks essentially competing against each other and improving throughout the training period.

After completing the training process, we proceeded to generate 400 synthetic specimens using our trained model. This generation phase involved randomly selecting categorical parameter combinations that represented various printing configurations, then generating appropriate noise vectors which were passed through the trained Generator network. The continuous output produced by the Generator was subsequently denormalized to restore them to their original scale and meaningful physical units. Finally, we assembled complete synthetic specimens by combining these denormalized mechanical properties with their corresponding categorical parameters, ensuring the resulting dataset matched the format and structure of the original experimental data for seamless integration and analysis.

5. Results and Evaluation

We conducted a thorough statistical evaluation (Table 2) comparing the real dataset (45 specimens) with the synthetic dataset (400 specimens):

Table 2

Descriptive statistics comparison

Property	Real Mean	Synthetic Mean	Real SD	Synthetic SD
Elastic Modulus (MPa)	1661.67	1554.51	534.84	394.70
Yield Strength (MPa)	20.33	19.00	6.02	5.02
Tensile Strength (MPa)	21.69	20.49	6.12	5.14

Range comparison: Elastic Modulus: Real [1166-2984 MPa], Synthetic [912-2911 MPa]; Yield Strength: Real [14.64-33.12 MPa], Synthetic [14.28-33.31 MPa]; Tensile Strength: Real [15.58-34.92 MPa], Synthetic [15.45-34.91 MPa].

These results indicate that the synthetic data closely matched the central tendencies of the real data, with slightly reduced standard deviations suggesting more conservative variation in the generated samples.

5.1. Distribution Analysis

The Kolmogorov-Smirnov (KS) test, measuring maximum distance between empirical cumulative distribution functions, formally evaluated distribution similarity. Results showed Elastic Modulus (KS-statistic = 0.0986, p-value = 0.7874), Yield Strength (KS-statistic = 0.1694, p-value = 0.1734), and Tensile Strength (KS-statistic = 0.1225, p-value = 0.5355). All p-values exceeded 0.05, confirming no statistically significant differences between real and synthetic distributions and providing quantitative evidence of successful GAN modeling.

5.2. Visual Assessment

Visual comparison through histograms and boxplots confirmed the statistical findings. Histogram comparisons (Fig. 2) revealed strong similarities between the real and synthetic distributions.

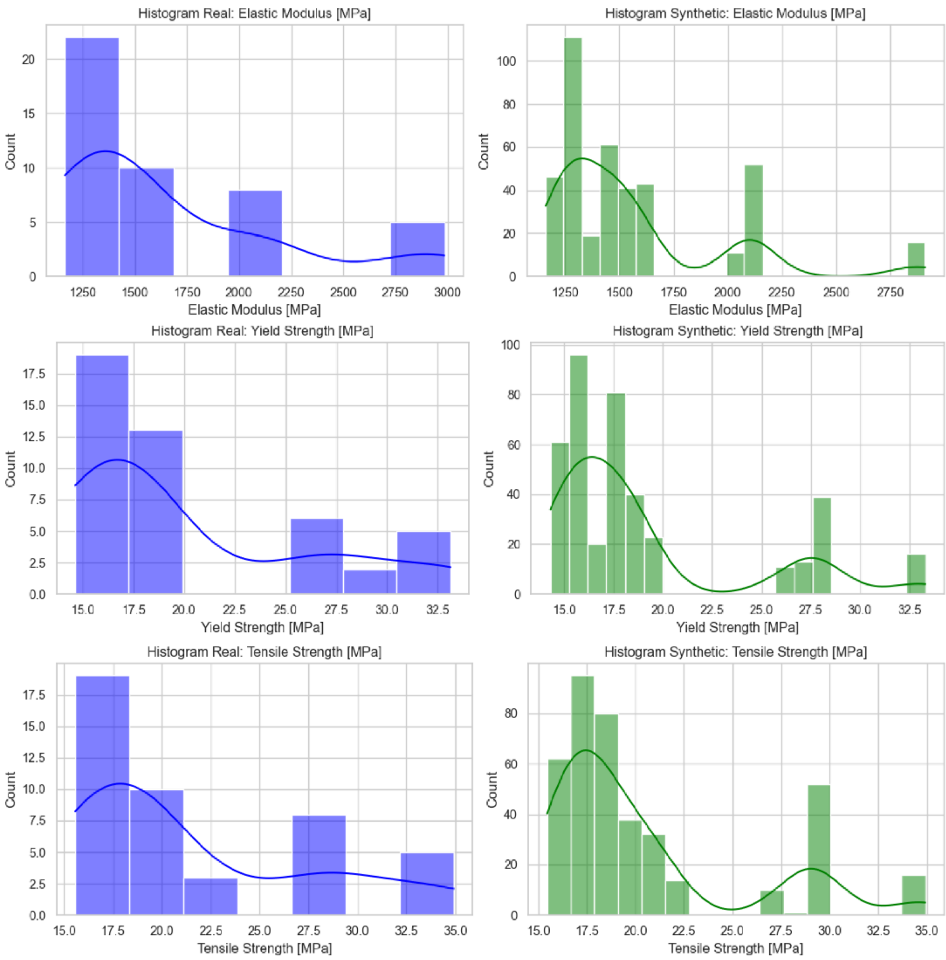

Fig. 2. Histogram comparisons between real and synthetic data

Both datasets exhibited comparable distribution shapes across all three mechanical properties, with the synthetic data successfully maintaining the slight right skew observed in the elastic modulus of the real specimens. The peak distributions aligned particularly well for the strength properties, demonstrating the GAN's ability to capture the central tendencies of the data.

When examining boxplot comparisons, we found that median values were closely aligned between real and synthetic data, indicating successful modeling of the central tendency. The interquartile ranges, representing the middle 50% of data, showed similar spreads in both datasets, though the synthetic data exhibited slightly tighter distributions. An interesting observation was that the synthetic data contained fewer extreme values or outliers, suggesting the GAN tended toward more conservative estimates within the established patterns. Despite this conservative tendency, the whisker ranges representing non-outlier extremes remained comparable between the two datasets. These visual confirmations support the conclusion that the synthetic data preserved the essential distributional characteristics of the original dataset.

5.3. Relationship Preservation

Beyond univariate distributions, the synthetic data preserved inter-variable relationships. The strong correlation between yield and tensile strength ($r > 0.9$) and moderate correlations between elastic modulus and strength properties remained consistent across both datasets. Categorical relationships were also maintained: higher mechanical properties with increased infill density, differential effects of infill patterns across density levels, and printing speed influences were all accurately reproduced.

This multi-dimensional validation is particularly important for engineering applications, as it ensures the synthetic data captures not just individual property distributions but also the complex interactions between printing parameters and mechanical outcomes.

6. Discussion

The GAN-generated synthetic data demonstrated high fidelity to the original experimental data across multiple evaluation metrics. The close alignment of means, standard deviations, and distributions indicates that the Generator successfully learned the underlying patterns in the real data.

The slightly narrower distributions observed in the synthetic data (smaller standard deviations) suggest a conservative bias in the Generator. This is a common phenomenon in GANs, particularly with small training datasets, and represents a trade-off between generating diverse samples and maintaining physical plausibility. In engineering applications, this conservative tendency may be beneficial, as it reduces the risk of generating physically impossible or unsafe values.

The preservation of relationships between variables is particularly noteworthy, as it indicates the GAN captured not just surface-level statistics but deeper structural patterns in the data. This is crucial for engineering applications where the interactions between parameters determine real-world performance.

6.1. Practical Applications

The validated synthetic dataset enables several practical applications in the field of additive manufacturing. The expansion from 45 to 400 specimens enabled the training and validation of predictive machine learning models. In particular, the augmented dataset was successfully used to develop (i) a regression model for predicting the elastic modulus, which achieved an accuracy of ~99%, and (ii) a classification model for printing speed, which reached ~98% accuracy. These results demonstrate that the generated data can effectively support ML training and yield high-performing models. Nevertheless, further studies with experimental data are needed to confirm the robustness and generalizability of these findings.

In terms of design space exploration, engineers can leverage synthetic data to investigate parameter combinations that were not tested experimentally. This capability allows optimization algorithms to identify promising parameter regions for targeted experimental validation, reducing the need for exhaustive physical testing. Engineers can more comprehensively assess the impact of parameter variations on mechanical outcomes, enabling more informed design decisions without the costs associated with extensive physical testing. Nevertheless, it should be noted that unseen parameter combinations may also lead to unforeseen consequences in certain domains and applications, as further discussed in the limitations section below. Perhaps most significantly, this approach offers substantial cost and time savings in the development process. By reducing the number of required physical tests, organizations can realize significant resource savings in materials, equipment time, and specialized labor. Preliminary designs can be rapidly evaluated using synthetic data before committing to physical prototyping, accelerating the iteration process. Additionally, testing resources can be strategically focused on validating only the most critical conditions identified through synthetic data analysis, optimizing the use of laboratory resources.

6.2. Limitations and Considerations

While the synthetic data demonstrated high quality, several limitations should be acknowledged. Regarding generalization boundaries, it's important to recognize that the synthetic data can only represent patterns present in the original experimental data. Extrapolation beyond the bounds of the training data should be approached with caution, as novel material behaviors or extreme conditions may not be accurately captured by the model.

Validation requirements remain an important consideration when working with synthetic data. These generated datasets should not completely replace experimental testing, especially for safety-critical applications where physical performance must be rigorously verified. A validation strategy using targeted physical tests remains necessary to confirm that real-world performance aligns with predictions. Synthetic data should be viewed as a valuable supplement too, rather than a replacement for real-world testing.

GAN-specific limitations also warrant consideration. Training stability continues to be challenging with small datasets like the one used in this study, requiring careful monitoring and parameter tuning. More advanced GAN architectures might yield further improvements in data quality and diversity, suggesting potential for future refinement. Domain expertise remains essential throughout the process, both for properly configuring the GAN and for validating the physical plausibility of the results generated. This underscores that these methods augment rather than replace engineering judgment.

7. Conclusions

This research demonstrated that conditional GANs can successfully generate synthetic mechanical property data for 3D printed specimens that are statistically indistinguishable from experimental data. The comprehensive analysis of both real and generated datasets revealed several significant findings:

1. The cGAN effectively integrated categorical printing parameters (infill pattern, layer height, print orientation) with continuous mechanical properties (elastic modulus, tensile strength, elongation at break), enabling controlled output generation based on manufacturing conditions.
2. The synthetic dataset (400 virtual specimens) preserved the real dataset's (45 physical specimens) statistical characteristics—means, standard deviations, correlations, and higher-order moments—achieving 9-fold data amplification without additional material consumption or testing.
3. Complex non-linear relationships and multi-parameter interactions affecting mechanical performance were accurately captured, surpassing traditional statistical modeling capabilities.
4. Kolmogorov-Smirnov tests confirmed no significant distributional differences between real and synthetic data ($p > 0.05$), with Q-Q plots and distribution overlays providing additional validation of sample fidelity.

These results validate using GANs to augment limited experimental datasets in additive manufacturing and other engineering domains where data collection is resource-intensive, time-consuming, and materially expensive.

The methodology demonstrated offers several practical benefits for researchers, engineers, and manufacturers working in additive manufacturing:

- **Resource efficiency**: Reduces physical testing requirements, conserving materials, energy, and equipment usage which is particularly valuable for expensive materials or limited testing facilities.
- **Accelerated development**: Enables faster design iteration through virtual testing, potentially reducing development cycles from months to weeks.
- **Enhanced modeling**: Provides sufficient data for robust machine learning models that predict properties with higher accuracy and better generalization to novel conditions.
- **Parameter exploration**: Facilitates investigation of untested or experimentally challenging parameter combinations, identifying optimal processing conditions and configurations.

- **Uncertainty quantification**: Enables comprehensive assessment of property variability, supporting design decisions that account for statistical variations rather than mean values alone.

These benefits directly address the persistent challenges faced in additive manufacturing, where material and process optimization remain critical for expanding industrial applications and meeting stringent performance requirements across aerospace, automotive, medical, and consumer product sectors.

Building on this work, several promising research directions emerge. Advanced GAN architectures like Wasserstein GAN with gradient penalty could enhance training stability with limited data sets, while variants optimized for tabular data or ensemble approaches might generate more diverse synthetic data. Parameter space expansion presents another opportunity, incorporating additional printing variables (nozzle temperature, layer height, extrusion width) and material composition factors for wider applicability. Validation methodologies need refinement through systematic approaches comparing synthetic data to physical tests and establishing optimal synthetic-to-real data ratios for industrial implementation. Finally, this approach could extend to multi-property optimization, developing Pareto frontiers to visualize trade-offs between competing mechanical properties and integrating with process simulation for comprehensive digital twins. These advancements would address broader challenges in engineering domains where experimental data is limited or costly.

R E F E R E N C E S

[1] *Abdulhameed, Osama, et al.*, Additive manufacturing: Challenges, trends, and applications, Advances in Mechanical Engineering, vol. **11**, no. 2, 2019, pp. 1–27, https://doi.org/10.1177/1687814018822880.

[2] *Giannopoulos, Andreas A., et al.*, Applications of 3D printing in cardiovascular diseases, Nature Reviews Cardiology, vol. **13**, no. 12, 2016, pp. 701–718, https://doi.org/10.1038/nrcardio.2016.170.

[3] *Fayazfar, Haniyeh, et al.*, A critical review of powder-based additive manufacturing of ferrous alloys: Process parameters, microstructure and mechanical properties, Materials & Design, vol. **144**, 2018, pp. 98-128, https://doi.org/10.1016/j.matdes.2018.02.018.

[4] *Ćwikła, G., et al.,* The influence of printing parameters on selected mechanical properties of FDM/FFF 3D-printed parts, IOP Conference Series: Materials Science and Engineering, vol. **227**, no. 1, https://dx.doi.org/10.1088/1757-899X/227/1/012033.

[5] *Galeta, Tomislav, et al.*, Influence of Structure on Mechanical Properties of 3D Printed Objects, Procedia Engineering, vol. **149**, 2016, pp. 100-104, https://doi.org/10.1016/j.proeng.2016.06.644.

[6] ISO/ASTM 52900:2021(en) Additive manufacturing — General principles — Fundamentals and vocabulary. https://www.iso.org/obp/ui/en/#iso:std:iso-astm:52900:ed-2:v1:en.

[7]. *Lin, Z., Jain, A., Wang, C., Fanti, G., Sekar, V.*, Using GANs for Sharing Networked Time Series Data: Challenges, Initial Promise, and Open Questions, Proceedings of the ACM Internet Measurement Conference, 2020, pp. 464-483. https://dl.acm.org/doi/10.1145/3419394.3423643

[8] *Iyer, A., Dey, B., Dasgupta, A., Chen, W., Chakraborty, A*., A Conditional Generative Model for Predicting Material Microstructures from Processing Methods, arXiv preprint arXiv:1910.02133, 2019, 10.48550/arXiv.1910.02133
[9] *Hsu, T., Epting, W.K., Kim, H., Abernathy, H.W., Hackett, G.A., Rollett, A.D., Salvador, P.A., & Holm, E.A.*, Microstructure Generation via Generative Adversarial Network for Heterogeneous, Topologically Complex 3D Materials, JOM, vol. **73**, 2021, pp. 90-102. https://link.springer.com/article/10.1007/s11837-020-04484-y
[10]. *Xie, X., Bennett, J., Saha, S., Lu, Y., Cao, J., Liu, W.K., & Gan, Z,* Mechanistic data-driven prediction of as-built mechanical properties in metal additive manufacturing, Computational Materials, vol. **7**, 2021, 86. https://www.nature.com/articles/s41524-021-00555-z
[11] *Xu, L., Skoularidou, M., Cuesta-Infante, A., Veeramachaneni, K.*, Modeling Tabular data using Conditional GAN, Advances in Neural Information Processing Systems, vol. **32**, 2019. https://papers.nips.cc/paper/8953-modeling-tabular-data-using-conditional-gan
[12] *Yoon, J., Jarrett, D., & van der Schaar, M.* Time-series Generative Adversarial Networks, Advances in Neural Information Processing Systems, vol. **32**, 2019. https://proceedings.neurips.cc/paper/2019/file/c9efe5f26cd17ba6216bbe2a7d26d490-Paper.pdf
[13] *Zhang, C., Kuppannagari, S.R., Kannan, R., Prasanna, V.K.*, Generative Adversarial Network for Synthetic Time Series Data Generation in Smart Grids, IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, 2018, pp. 1-6. https://ieeexplore.ieee.org/document/8587464
[14] *Hertlein, N., Buskohl, P.R., Gillman, A., Vemaganti, K., Anand, S.*, Generative adversarial network for early-stage design flexibility in topology optimization for additive manufacturing, Journal of Manufacturing Systems, **59**, 2021, 675-685. https://doi.org/10.1016/j.jmsy.2021.04.007
[15] *Goodfellow, Ian J., et al*., Generative Adversarial Nets, Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014), vol. **2**, 2014, pp. 2672-2680, https://doi.org/10.48550/arXiv.1406.2661.
[16] *Li, Si-Geng, et al*., Data-driven approach to predict the fatigue properties of ferrous metal materials using the cGAN and machine-learning algorithms, Advances in Manufacturing, vol. **12**, no. 3, 2024, pp. 447-464, https://doi.org/10.1007/s40436-024-00491-3.
[17] *Qie, Yifan, et al*. Wasserstein generative adversarial networks for form defects modeling, Procedia CIRP, vol. **114**, 2022, pp. 7-12, https://doi.org/10.1016/j.procir.2022.10.002.
[18] *Geng, Zhiqiang, et al*., Intelligent Small Sample Defect Detection of Water Walls in Power Plants Using Novel Deep Learning Integrating Deep Convolutional GAN, IEEE Transactions on Industrial Informatics, vol. **19**, no. 6, 2022, pp. 7489-7497, https://doi.org/10.1109/TII.2022.3159817.
[19] *Gao, Yun, et al*., FEM Simulation-Based Generative Adversarial Networks to Detect Bearing Faults, IEEE Transactions on Industrial Informatics, vol. **16**, no. 7, 2020, pp. 4961-4971, https://doi.org/10.1109/TII.2020.2968370.
[20] *Pan, Tongyang, et al*, Generative adversarial network in mechanical fault diagnosis under small sample: A systematic review on applications and future perspectives, ISA Transactions, vol. 128, 2022, pp. 1-10, https://doi.org/10.1016/j.isatra.2021.11.040.