

INVERSE DESIGN WITH CONDITIONAL CASCADED DIFFUSION MODELS

Milad Habibi¹, Mark Fuge^{1,*}

¹Center for Risk and Reliability, Department of Mechanical Engineering, University of Maryland, College Park, MD

ABSTRACT

Adjoint-based design optimizations are usually computationally expensive and those costs scale with resolution. To address this, researchers have proposed machine learning approaches for inverse design that can predict higher-resolution solutions from lower cost/resolution ones. Due to the recent success of diffusion models over traditional generative models, we extend the use of diffusion models for multi-resolution tasks by proposing the conditional cascaded diffusion model (cCDM). Compared to GANs, cCDM is more stable to train, and each diffusion model within the cCDM can be trained independently, thus each model's parameters can be tuned separately to maximize the performance of the pipeline. Our study compares cCDM against a cGAN model with transfer learning.

Our results demonstrate that the cCDM excels in capturing finer details, preserving volume fraction constraints, and minimizing compliance errors in multi-resolution tasks when a sufficient amount of high-resolution training data (more than 102 designs) is available. Furthermore, we explore the impact of training data size on the performance of both models. While both models show decreased performance with reduced high-resolution training data, the cCDM loses its superiority to the cGAN model with transfer learning when training data is limited (less than 102), and we show the break-even point for this transition. Also, we highlight that while the diffusion model may achieve better pixel-wise performance in both low-resolution and high-resolution scenarios, this does not necessarily guarantee that the model produces optimal compliance error or constraint satisfaction.

1. INTRODUCTION

Optimizing designs, especially in the context of adjoint-based multi-physics shape and topology optimization (TO), can be a time-intensive process, often requiring costly iterations to achieve the desired designs [1]. For instance, one of the most commonly used methods to solve forward TO is Solid Isotropic

Material with Penalization (SIMP) method [2, 3]. This method is based on iterative optimization techniques over continuous, density-based methods to find the optimal material distribution. Despite the wide popularity of these methods, they suffer from two main pitfalls: (1) often requiring expensive simulations and (2) getting trapped in local optima due to their gradient-based solution approach. To overcome these challenges for TO problems, previous studies combine TO with inverse design methods driven by Machine learning. These methods complement adjoint methods by reducing the time and resources required to explore a broader range of design solutions compared to optimizing a single design with an iterative forward model. For instance, Habibi *et al.* [4] showed that ML-based inverse design can predict designs with 95% relative performance improvement for 2D heat conduction topology optimization problem compared to the SIMP method and more importantly it can accelerate the topology optimization even with some small amount of training data.

One of the main uses of ML-based design methods is addressing TO computational costs which scale with resolution. Several key questions arise in this context, including identifying techniques and models that can effectively reduce costs, determining the volume of data necessary for optimal model performance, and finding the right balance between gathering examples of low and high resolution and dimensions. A common technique to answer this challenge in the community is producing high-resolution images from low-resolution data which is also known as super-resolution [5]. Some researchers used Generative Adversarial Networks (GANs) trained on low-resolution TO datasets to perform upscaling to produce high-resolution images. For instance, Yu *et al.* implemented a two-stage refinement conditional GAN (cGAN) to predict optimal high-resolution structures (128x128), leveraging training on 64,000 low-resolution data samples (32x32) and corresponding high-resolution counterparts [6]. Additionally, other investigators have proposed leveraging transfer learning in conjunction with GANs or variational autoencoders (VAEs) for super-resolution. Behzadi *et al.*, for example, used transfer learning and conditional GANs to fine-tune a model on 11,000 fixed volume fraction low-resolution sam-

*Corresponding author: fuge@umd.edu

ples (80x40) and 1,500 high-resolution samples (up to 200x400), contributing valuable insights to the delicate balance between low and high-resolution examples in the transfer learning context [7].

In this study, we aim to extend Maz *et al.*'s comparison of the performance and training simplicity of condition diffusion models versus that of GANS in TO [8], but in the context of multi-resolution problems. Our primary objective is to quantify the extent to which diffusion-based generative models surpass their GAN-based counterparts in addressing low to high-resolution super-resolution challenges. Additionally, we explore the performance of each model under fixed-cost considerations, systematically examining their efficacy in handling both seen and unseen boundary conditions. Specifically, the contributions of this paper are as follows:

1. We describe a conditional cascaded diffusion model-based framework for solving low to high-resolution engineering design problems.
2. We demonstrate the extent to which cascaded diffusion-based generative models outperform GAN-based generative models on low to high-resolution inverse design problems.
3. For a given threshold of performance (e.g., constraint violation, compliance, etc.), we study the comparative sample efficiency of diffusion models versus GAN models for low to high-resolution problems, noting a critical high-resolution sample threshold below which the cGAN model outperforms the diffusion model, and vice versa.

2. RELATED WORKS

2.1 Inverse Design and Topology Optimization

Inverse design (ID) involves leveraging machine learning techniques to predict the optimal design, such as an airfoil mesh, based on specified input conditions, such as a desired Reynolds or Mach number [9]. ID studies focus on the mapping between the input conditions and optimal design which can be used in place of an existing optimizer or as an initial guess for further optimization (warmstart optimization) [9]. In the field of Mechanical engineering, ID has been explored in diverse applications, designing and characterizing materials [9–13], airfoil shape design [9, 14], and topology optimization [4, 15, 16].

Topology optimization (TO) finds the optimal distribution of the material in the design domain to optimize one or more objective functions while satisfying some constraints. Traditional TO methods rely heavily on Finite element analysis [1]. One of the most widely used formulations of TO problems are pseudo-density-based approaches such as the Solid Isotropic Material with Penalization method (SIMP) [2, 17]. The goal of structural TO is to identify the optimal material distribution within the design domain to minimize structural deformation, subject to a volume constraint and prescribed load condition [18].

The specific minimization problem is formulated as follows:

$$\begin{aligned}
 & \underset{x}{\text{minimize}} \quad c(x) = U^T K U \\
 & \text{subject to} \quad \frac{V(x)}{V_0} = f, \\
 & \quad K U = F, \\
 & \quad 0 \leq x_{\min} \leq 1
 \end{aligned} \tag{1}$$

where $c(x)$ represents structural compliance, U denotes global displacement, F is the force vector, and K is the global stiffness matrix. Additionally, $V(x)$ and V_0 represent material volume and design domain volume, respectively, with f representing the prescribed volume fraction. The vector x_{\min} is introduced to prevent the singularity condition by specifying minimum relative densities.

Traditional TO methods usually suffer from two limitations: (1) often requiring costly iterations and (2) getting trapped in local optima due to their gradient-based solution approach. To address these limitations of traditional TO many researchers developed various deep-learning methods [19–21]. For instance, Sharpe and Seepersad introduce conditional generative adversarial networks (cGANs) to generate a compact latent representation of structures resulting from traditional TO methods [22]. Furthermore, Guo *et al.* used variational autoencoder and style transfer to enhance topology optimization capabilities [23]. Additionally, Nie *et al.* expand the generalizability of using the condition generative model by introducing TopologyGAN for more variant physical and boundary conditions [24]. Moreover, some researchers incorporate physics and generative models to solve the TO problems [13, 25, 26]. Additionally, reinforcement learning-based generative design frameworks have been explored by some researchers to improve the diversity of topology-optimized designs [27, 28].

The aforementioned works showcase diverse approaches to incorporating deep learning into design optimization. All are either supervised or unsupervised algorithms that rely on the use of enough amount of training data to predict acceptable results. But how much of this expensive high-resolution data is enough? There are a few papers that consider the limitation of high-resolution training data amounts [7, 29], typically leveraging transfer learning. In this paper, we aim to quantify and understand how training size affects ML-based inverse design model performance, but beyond the GAN-based models studied previously.

2.2 Generative models for Inverse Design using Super-Resolution

One of the main challenges in solving inverse design problems, especially TO problems is the computational cost which scales with the resolution. To tackle this issue, researchers often refer to super-resolution techniques from computer vision [19, 30]. Yu *et al.* [6] used a two-stage refinement, employing a conditional generative adversarial network (GAN) for upscaling and predicting near-optimal structures, while Li *et al.* employed a Super-Resolution GAN (SRGAN) for refining high-resolution structures [31].

While various GANs and Variational Autoencoders (VAEs) have been applied in these methods, an alternative approach in-

volves transfer learning, a technique transferring knowledge from one domain to another effectively [32, 33]. Behzadi and Ilies introduced a transfer learning method based on a convolutional neural network (CNN) to the topology optimization community [29]. Their approach involved training a feed-forward CNN on low-resolution data, transferring knowledge by locking model weights, and adding layers for upscaling output resolution. Subsequent work by Behzadi and Ilies proposed a design exploration framework leveraging transfer learning and conditional GANs, surpassing previous generalization capabilities [7].

Dhariwal and Nichol demonstrated that diffusion models excel in image generation, outperforming GANs, and offer easier training and adaptability to various tasks [34]. Building on this, Mazé and Ahmed showcased that a topology-based diffusion model significantly outperforms state-of-the-art conditional GANs, reducing average errors in physical performance by a factor of eight and producing fewer infeasible samples [8]. Inspired by these findings, our study addresses the different problem of super-resolution in diffusion models, and answers the question “to what extent do diffusion-based generative models surpass GAN-based counterparts in low to high-resolution problems, and is the adoption of diffusion models appropriate for super-resolution tasks?”

2.3 Diffusion Models

Diffusion models, as introduced by Sohl-Dickstein *et al.* [35], represent a class of deep generative models. These models include two main components: the forward and backward diffusion processes. The forward diffusion process entails the addition of noise to sample vectors over T steps to a sample from data distribution $x_0 \sim q(x_0)$, generating a sequence of noisy samples x_1, x_2, \dots, x_T , following the Markov chain:

$$q(x_t | x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I) \quad (2)$$

Where $q(x_t | x_{t-1})$ the forward process posteriors, β_t is variance schedule at time t . As the forward diffusion progresses, the initial data sample x_0 gradually loses its features, a phenomenon controlled by the variance schedule denoted as $(\beta_t)_{t=1}^T \in (0, 1)$.

The reverse process seeks to approximate the true posterior using a parametric Gaussian process. This involves recreating the true sample from a Gaussian noise input, expressed as

$$p_\theta(x_{t-1} | x_t) = \mathcal{N}(\mu_\theta(x_t), \Sigma_\theta(x_t)) \quad (3)$$

In the reverse process the new image is generated by sampling a data from pure noise and removing noise gradually using reverse process equation. The reverse process involves training the neural network model to predict the mean and variance of the denoising process. In our model we assumed variance of the denoising process Σ_θ is fixed and constant and we learn the mean variable μ_θ [36]. In summary, the whole intuition behind the diffusion model is training models to reverse a noise process and mapping a Gaussian noise distribution to the data distribution. Readers interested in finding out more mathematical details about diffusion models are directed to [37].

Diffusion models, as demonstrated by Ho *et al.* [36], show superior image quality and greater stability compared to GANs.

While Denosing Diffusion Models (DDMs) are relatively novel in the context of engineering design, there have been notable studies exploring their application in creating two-dimensional structures. One noteworthy example is the work of Mazé and Ahmad [8], who introduced a condition diffusion model-based architecture, TopoDiff. This innovative approach was employed for performance and manufacturability-aware topology optimization, outperforming state-of-the-art conditional GANs by reducing the average physical performance by a factor of eight. Building on this advancement, Giannone *et al.* [25] proposed a learning diffusion model framework that eliminates the need for expensive finite element preprocessing, resulting in improved design performance. These studies highlight diverse methodologies for integrating diffusion models into design optimization, yet a crucial aspect remains unexplored. None of these works have delved into the potential of super-resolution within diffusion models as compared to GANs for engineering designs which we aim to explore in this study.

3. METHODOLOGY

To address the contributions outlined in the introduction we divide the methodology into the following sections: (1) introducing the cascaded pipeline, (2) what are the diffusion model parameters and its architecture, (3) what cGAN model we benchmarked against our proposed diffusion model, and (4) how we measure and evaluate our models.

3.1 Conditioned Cascading Pipelines

Inspired by previous studies [38, 39], we introduce a framework for addressing multi-resolution engineering challenges, illustrated in Fig 1. Our core approach lies in conditional cascaded diffusion (cCDM), which involves a sequence of generative diffusion models. At the lower resolution, we use a generative diffusion model conditioned on specified boundary conditions. Following the methodology of prior works [8, 24], we perform channel-wise concatenation with channels representing the volume fraction, strain energy density, von Mises stress, and applied loads on the domain’s boundary. These conditionally augmented features are then integrated with the low-resolution image. Moving to higher resolutions, our framework incorporates a super-resolution diffusion model conditioned by channel-wise concatenation. This concatenation involves the low-resolution image predicted by the preceding diffusion model, which is further processed through bilinear upsampling to achieve the desired resolution. Also, the conditioning at the high resolution includes the given boundary conditions. One important advantage of employing these cascading models is their independence during training, allowing for separate tuning of hyperparameters at each stage. This flexibility improves the overall performance of the pipeline, as parameters can be optimized individually for each resolution level. We describe the specific network architectures and diffusion parameters in later sections.

3.2 U-Net Architecture and Diffusion Parameters

We used a UNET architecture as the basis for our diffusion models implementation, obtained from the following publicly available repository [40]. We made specific modifications based

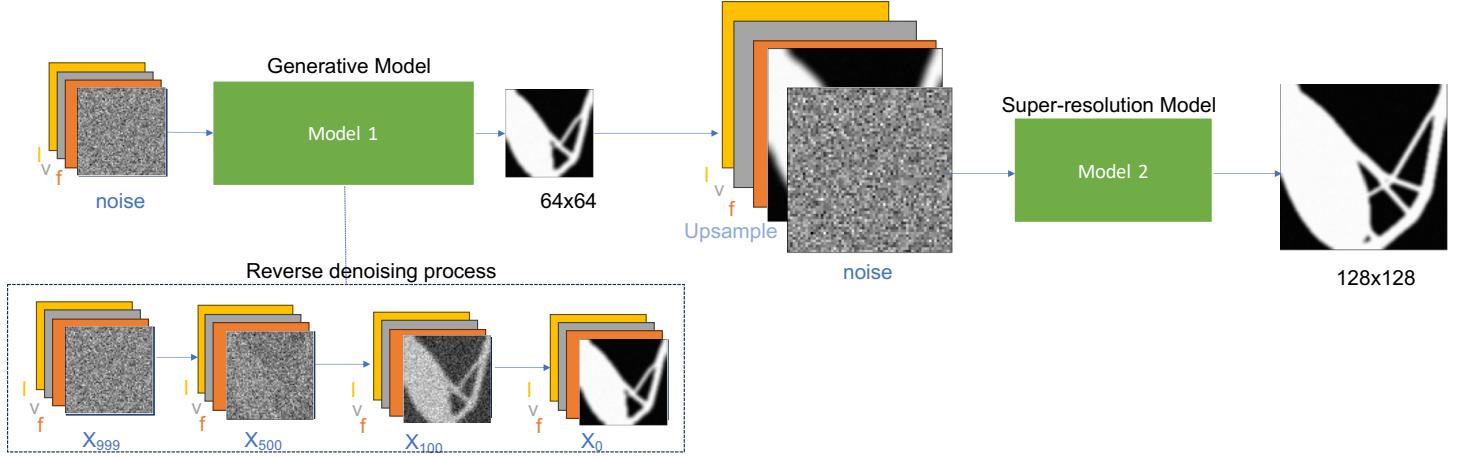


FIGURE 1: PROPOSED CCDM PIPELINE FOR TO SUPER-RESOLUTION. MODEL 1 IS A STANDARD CONDITIONAL DIFFUSION MODEL FOR TO. L IS THE LOAD APPLIED, REPRESENTED WITH AN ARROW ON THE TOPOLOGY; V IS THE VOLUME FRACTION; F REPRESENTS THE PHYSICAL FIELD INCLUDING THE STRESS AND STRAIN FIELD. MODEL 2 IS A SUPER-RESOLUTION MODEL CONDITIONED ON BOUNDARY CONDITIONS AND UPSAMPLED PREDICTION OF MODEL 1.

on our needs, notably, we removed the attention block within the UNET architecture to reduce the computational complexity and because including the attention block worsened performance in our initial tests. For the baseline model, we used a linear variance β schedule linearly increasing from 10^{-4} to 0.02 as it is suggested in the previous study [36] with a total of 1000 steps where the 1st step represents the lowest amount of noise, and 1000 is the highest.

3.3 Conditional Generative Adversarial Networks and Transfer Learning

We adapted the Conditional Generative Adversarial Network (cGAN) architecture from Behzadi *et al.* [7], sourced from their public GitHub repository, to benchmark against our CCDM. Our modifications prioritized improving stability and performance over our dataset. To address challenges encountered during model training we used a Wasserstein loss, which reduced mode collapse issues we found during testing relative to the original implementation in [7].

3.4 Dataset

To test our models and be comparable to the existing state-of-the-art we created a dataset via a classical structure compliance topology optimization problem based on the SIMP approach using the code provided by Andreassen *et al.* which we detail below [41]. This problem aims to minimize the structural compliance of a beam (represented by a square domain) while satisfying the constraints on material volume and boundary conditions, including force magnitude, location, and direction.

Consistent with past work, we selected two sets of boundary conditions, as illustrated in Fig. 2(a), using the first set for both training and testing, while the second set served as unseen boundary conditions exclusively for testing purposes. The force magnitude was set to 5000N, and the design domains included volume fractions (v) ranging from 0.30 to 0.60, force locations (h) ranging from 0 to 1, and force directions (α) on the right side of the design domain.

For the seen boundary conditions, we divided the volume fraction range into 20 segments and the force location and direction ranges into 10 segments, resulting in 2000 designs for each boundary condition. For the unseen boundary conditions, we used 10 segmentations for volume fraction and force location and direction, resulting in 1000 designs for each boundary condition. Within the seen boundary conditions category, we employed a random selection process to allocate 200 from each seen boundary condition designs for testing purposes, while the 1500 designs were dedicated to training the model which resulted in 9000 training data and 1200 testing seen boundary conditions. Similarly, for the unseen boundary conditions category, we randomly selected 1000 designs for testing, with an equal distribution of designs across different unseen boundary conditions. These unseen boundary conditions served as a critical benchmark for evaluating the model's ability to test the generalizability of the models. Also, we generated the dataset at two distinct resolutions: a low resolution of 64x64 and a high resolution of 128x128.

Inspired by previous studies [8, 24] highlighting the improved prediction accuracy achieved by incorporating physical fields such as von Mises stress and strain energy density instead of explicitly representing boundary conditions, we introduced these fields as additional input channels in the generative models. Additionally, force and volume fraction channels were included to enable the network to identify relationships between physical boundary conditions and fields, focusing on structural patterns that serve as effective load paths.

The physical representation dataset comprises five channels:

- The first channel is uniform and contains the volume fraction constraint;
- The second channel includes the von Mises stress values of the design domain under the given constraints;
- The third channel is the strain energy density of the design considering both load constraints and boundary conditions. Strain energy is defined as $W = \frac{\sigma_x \epsilon_x + \sigma_y \epsilon_y + 2\sigma_{xy} \epsilon_{xy}}{2}$;

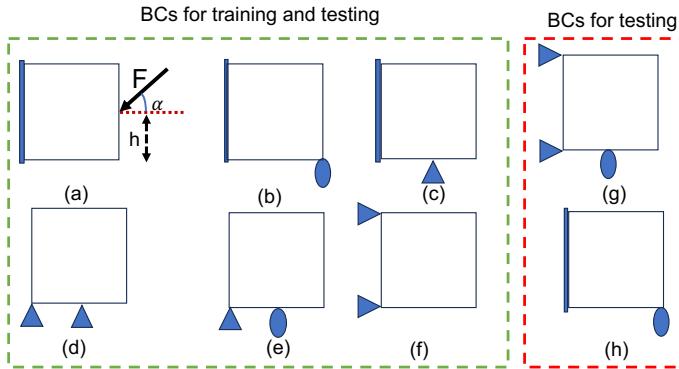


FIGURE 2: THE PHYSICAL LAYOUT OF TOPOLOGY OPTIMIZATION PROBLEM OF BEAMS AND CORRESPONDING TRAINING AND TESTING BOUNDARY CONDITION. DESIGN DOMAIN OF A CANTILEVER BEAM WITH DESIGN PARAMETERS OF FORCE LOCATION (H) AND DIRECTION (α) ON THE RIGHT SIDE OF THE BEAM. BOUNDARY CONDITIONS: (A)–(F) ARE USED FOR TRAINING AND TESTING AND (G)–(H) ARE UNSEEN BCS ONLY USED FOR TESTING.

- The fourth channel represents the load constraints in the x-direction. Each node indicates the force applied in the x-direction direction (or 0 if there's no force applied);
- The fifth channel represents the load constraints in the y-direction. Each node indicates the force applied in the y-direction direction (or 0 if there's no force applied);

Since our study involved the comparison of different models, variations existed in the input data for each model. Specifically, the GAN model was trained on the physical representation dataset, while the diffusion model on low resolution incorporated an extra channel in the black-and-white image, representing the optimal topology. Furthermore, the diffusion model on higher resolution included two extra channels, incorporating topology optimization on higher resolutions and upsampled topology predictions from the low-resolution diffusion model.

3.5 Evaluation Metrics

In this study, we use pixel-wise error as a final evaluation metrics as well as physical objectives since in designing generative models minimizing objective values is more important than classical computer vision metrics. As result, we define and use the following three metrics for our case studies evaluations:

- Mean Squared Error (MSE), which is defined as:

$$MSE = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m (y_{ij}^{pred} - y_{ij}^{true})^2, \quad (4)$$

where n and m are the number of rows and columns, y_{ij}^{pred} and y_{ij}^{true} are the predicted and reference values of the element located in the i th row and j th column, respectively.

- Volume fraction error (VFE) relative to the input volume fraction is defined as:

$$VFE = \frac{|VF(y^{pred}) - VF(y^{true})|}{VF(y^{true})} \quad (5)$$

where $VF(y^{true})$ and $VF(y^{pred})$ represent, respectively, the prescribed volume fraction and the volume fraction of the topology generated by our diffusion model.

- Compliance error (CE) in relation to the ground truth is defined as:

$$CE = \frac{C(y_f^{pred}) - C(y_f^{true})}{C(y_f^{true})} \quad (6)$$

where $C(y_f^{true})$ and $C(y_f^{pred})$ represent the compliance of the feasible topology generated by the SIMP method and the feasible topology generated by the generative models which do not violate the volume fraction constraints, respectively. It is important to note that a negative compliance error indicates that generative models produce a topology with lower compliance than the ground truth. Furthermore, we considered topologies that satisfy volume fraction constraints and can be considered feasible designs.

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 How well do different models perform in low-resolution?

To assess the performance of the diffusion model and the cGAN model for low-resolution topology optimization, we used the two test sets described in Section 3.4. Our evaluation focuses on comparing the models' performance using pixel-wise error metrics as well as physical values.

Figure 3 shows examples of a few structures obtained with (1) the conditional GAN, (2) the conditional diffusion model, and (3) the SIMP method, which we label as the “Ground Truth” (GT), for randomly selected constraints from seen and unseen boundary conditions. As expected, the diffusion model exhibits superior performance qualitatively in terms of pixel-wise metrics in comparison to the cGAN model, aligning with previous research that showed diffusion models can outperform GANs on image synthesis [34]. However, our previous study has shown that optimizing solely for MSE values does not necessarily guarantee optimal design outcomes [15].

To quantitatively validate this observation, we summarize the performance of the generated structures across seen and unseen boundary conditions in Table 1. Our analysis of the MSE indicates that the diffusion model outperforms the cGAN model in capturing finer details and minimizing pixel-wise errors across both seen and unseen boundary conditions. Interestingly, both models exhibit comparable performance in maintaining volume fraction constraints. In addition, the diffusion model outperforms the cGAN model in minimizing compliance for more difficult test cases (unseen boundary condition test cases) by producing smaller median compliance errors. It should be noted we used the median for CE values since the median is less sensitive to outliers. This indicates that the topologies generated by the diffusion model are producing less structural compliance, highlighting its efficacy in producing feasible optimal designs.

Model	Boundary Condition	Average MSE ($\times 10^{-4}$)	Average VFE ($\times 10^{-4}$)	Median CE ($\times 10^{-2}$)
Diffusion Model	Seen	6.67 ± 2.00	81.67 ± 4.02	0.34 ± 0.04
	Unseen	324.94 ± 19.12	448.83 ± 30.60	4.18 ± 8.54
cGAN Model	Seen	13.58 ± 2.28	183.31 ± 11.13	0.09 ± 0.15
	Unseen	327.15 ± 17.24	421.37 ± 24.11	5.86 ± 425.85

TABLE 1: PERFORMANCE METRICS COMPARISON IN LOW RESOLUTION (64X64) ON SEEN AND UNSEEN BOUNDARY CONDITIONS. VALUES AFTER \pm INDICATE THE 95 % CONFIDENCE INTERVAL AROUND AVERAGES/MEDIANS. THE VALUES IN BOLD FEATURE ARE THE BEST ONES FOR EACH BOUNDARY CONDITION.

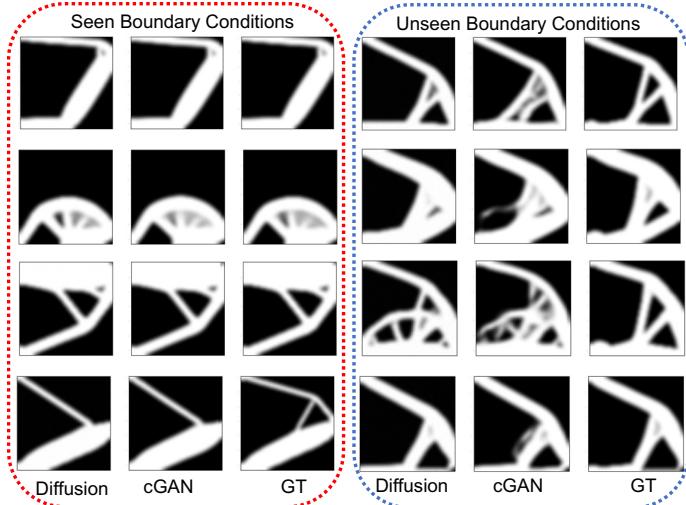


FIGURE 3: COMPARISON OF GENERATED STRUCTURES ON RANDOMLY SELECTED SAMPLES FROM SEENS AND UNSEEN BOUNDARY CONDITIONS IN THE LOW-RESOLUTION DATA. THE GT COLUMN REPRESENTS THE GROUND TRUTH GENERATED BY THE SIMP METHOD.

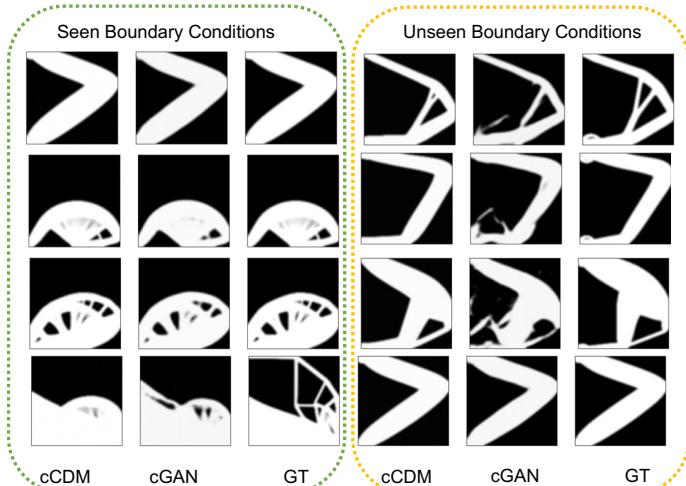


FIGURE 4: COMPARISON OF GENERATED STRUCTURES ON RANDOMLY SELECTED SAMPLES FROM SEENS AND UNSEEN BOUNDARY CONDITIONS IN THE HIGH-RESOLUTION DATA. THE GT COLUMN REPRESENT THE GROUND TRUTH GENERATED BY SIMP METHOD.

4.2 How well do different models perform in high-resolution with limited training data?

We evaluated the performance of the cCDM and the transfer of knowledge from low-resolution to high-resolution in the cGAN model. For this evaluation, we trained the cCDM using a limited number of training data (350 samples) from each seen boundary condition, resulting in a total of 2100 training samples for high-resolution. To provide a benchmark, we fine-tuned the cGAN model with transfer learning using the same training dataset.

Figure 4 showcases examples of structures generated by the diffusion model and the cGAN model, along with SIMP method for randomly selected constraints from both seen and unseen boundary conditions. Consistent with expectations, qualitative analysis reveals the superior performance of the cCDM in terms of pixel-wise metrics compared to the cGAN model with transfer learning across both seen and unseen boundary condition test cases.

To quantify these comparisons, Table 2 provides a comprehensive overview of our evaluation metrics. The results clearly demonstrate the diffusion model's superiority over the cGAN model in all evaluated metrics for high-resolution topology optimization on both seen and unseen boundary conditions. Specifically, the diffusion model achieves lower MSE, VFE, and CE values, indicating its superior ability to capture finer details, preserve volume fraction constraints, and minimize compliance errors in super-resolution task.

4.3 How well do models perform in high-resolution as a function of training data size?

One of the motivations to address this question stems from the high cost associated with generating high-resolution training data using the SIMP method. In response, we investigated the performance of the cGAN with transfer learning and cascaded diffusion models for high-resolution structures as we decreased the amount of high-resolution training data. The experiment aimed to assess how the performance of both models is impacted by the reduction in training data and to compare their adaptability to limited training resources, which can be the dominant cost.

The number of training data varied from 24 to 2100. We excluded seen test boundary conditions and gradually reduced the size of the high-resolution training data. Each subsequent training set was obtained by randomly removing data from the next largest set. For example, the high-resolution dataset initially included 2100 designs, with 350 designs from each seen boundary condition. Subsequently, the dataset containing 1050 designs

Model	Boundary Condition	Average MSE ($\times 10^{-4}$)	Average VFE ($\times 10^{-4}$)	Median CE ($\times 10^{-2}$)
cCDM	Seen	44.96 \pm 6.36	99.11 \pm 5.32	0.39 \pm 0.20
	Unseen	465.14 \pm 23.57	464.65 \pm 29.27	10.44 \pm 5.40
cGAN Model	Seen	98.09 \pm 8.52	210.54 \pm 11.30	3.20 \pm 60533.52
	Unseen	579.00 \pm 21.78	583.75 \pm 29.70 $\times 10^{-3}$	17.29 \pm 20.30

TABLE 2: PERFORMANCE METRICS COMPARISON IN HIGH RESOLUTION (128X128) ON SEEN AND UNSEEN BOUNDARY CONDITIONS. VALUES AFTER \pm INDICATE THE 95 % CONFIDENCE INTERVAL AROUND AVERAGES/MEDIAN

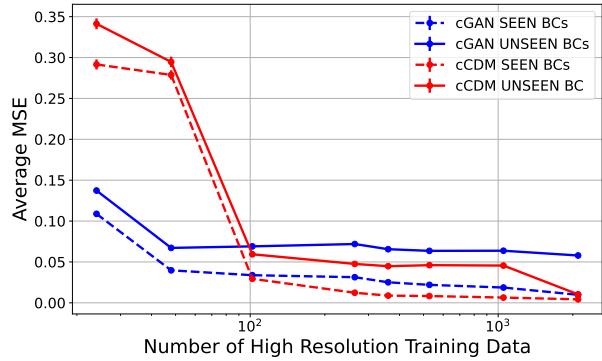


FIGURE 5: MEAN MSE OF THE CGAN MODEL WITH TRANSFER LEARNING VERSUS THE CASCADED DIFFUSION MODEL ACROSS TEST DATA WITH SEEN AND UNSEEN BOUNDARY CONDITIONS AS THE NUMBER OF HIGH-RESOLUTION TRAINING DATA CHANGES. ERROR BARS REPRESENT THE 95% CONFIDENCE INTERVALS AROUND THE AVERAGE MSE VALUES, PROVIDING INSIGHTS INTO THE UNCERTAINTY OF THE MEASUREMENTS.

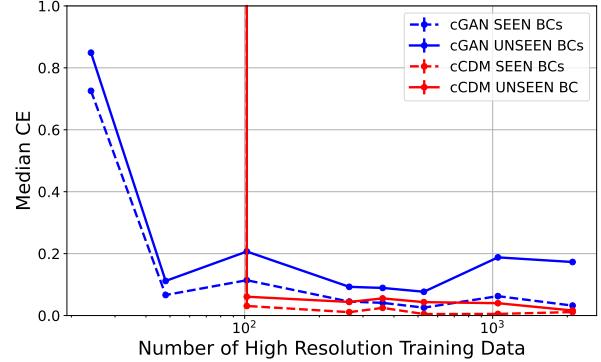


FIGURE 7: MEDIAN OF CE OF THE CGAN MODEL WITH TRANSFER LEARNING AND CASCADED DIFFUSION MODELS ACROSS TEST DATA WITH SEEN AND UNSEEN BOUNDARY CONDITIONS FOR FEASIBLE DESIGNS AS THE NUMBER OF HIGH-RESOLUTION TRAINING DATA DECREASES ARE SHOWED.

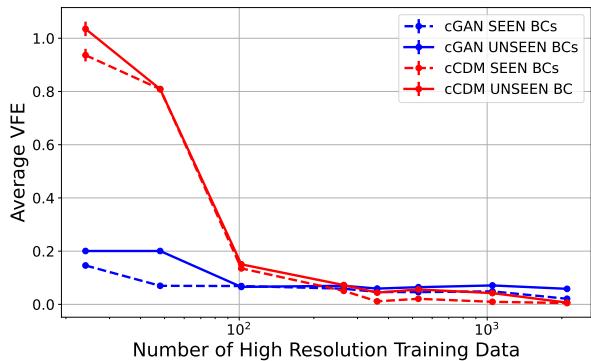


FIGURE 6: MEAN OF VFE OF THE CGAN MODEL WITH TRANSFER LEARNING AND CASCADED DIFFUSION MODELS ACROSS TEST DATA WITH SEEN AND UNSEEN BOUNDARY CONDITIONS AS THE NUMBER OF HIGH-RESOLUTION TRAINING DATA DECREASES ARE SHOWED. ERROR BARS REPRESENT THE 95% CONFIDENCE INTERVALS AROUND THE AVERAGE VFE VALUES, PROVIDING INSIGHTS INTO THE UNCERTAINTY OF THE MEASUREMENTS.

was created by removing 175 designs from each previously seen boundary conditions.

Figure 5 illustrates the MSE performance of the cGAN with transfer learning and cCDM across seen and unseen boundary conditions as the number of high-resolution training data decreases. The cCDM outperformed the cGAN model, maintaining its superiority even as the amount of training data decreased for both sets of test data where the training dataset contains more than 102 samples. When the number of high-resolution data is limited (less than 102), however, we observe a reversal in performance, with the cGAN model with transfer learning outperforming the cascaded diffusion model.

Similarly, Figure 6 shows how volume fraction error changes across both models as the number of training data decreases. Remarkably, the performance remains relatively stable up to 360 training data. However, beyond this threshold, a notable decrease in performance is observed, showcasing the sensitivity of both models to a low amount of training data. Notably, the cCDM consistently demonstrates superiority over the cGAN model, when the training dataset contains more than 102 samples. However, we observe a reversal in performance when the amount of high-resolution data is limited (less than 102), with the cGAN model with transfer learning outperforming the cascaded diffusion model under these conditions.

Figure 7 illustrates the median values of CE since median value is less sensitive to outliers. We observed that the median value of the cCDM demonstrates a smoother trend, with a gradual

increase as the number of training data decreases. It is worth noting that we only considered feasible designs that do not violate volume fraction constraints for compliance error calculation in Fig 7. For completeness, we also compare the median CE across all designs (both including versus excluding infeasible designs) and provide this in the supplement material. Figure 7 highlights the consistent superiority of the diffusion model over the cGAN model when the training dataset contains more than 102 samples. However, similar to other metrics the reversal in performance is observed when the number of high-resolution data is limited (less than 102), with the cGAN model with transfer learning outperforming the cCDM under these conditions. Another important observation is that with only 48 designs, the median value of CE for the cascaded diffusion model is eight orders of magnitude higher than that of the cGAN model. Moreover, with 24 training data, the cascaded diffusion model failed to predict any feasible designs. However, as demonstrated in the supplementary material, considering all predicted designs (ignoring constraint violations) can lead to misleading conclusions, such as the best performance with a training size of 24.

4.4 What is the effect of just the high-resolution model if we use a low-resolution oracle prediction?

While the above results show how both generative models perform (Low resolution plus subsequent High Resolution), it is possible that deficiencies in the low-resolution model could be impacting the overall results. In this section, we try to isolate the effect of just the super-resolution in cCDM by providing ground truth low-resolution samples generated by the adjoint solver (SIMP method) in place of the low-resolution design produced by the first stage of the conditional generative model—we refer to this setup as the “Super Resolution (SR) Model using SIMP,” in contrast to our joint cCDM model. The evaluation focuses on assessing whether the low-resolution generative model’s accuracy influences the performance of the multi-resolution model. The motivation for this experiment is both scientific and practical: we wish to understand the isolated causal effect of the Super Resolution element of the pipeline, and in practice, it might be cheap to generate good SIMP TO solutions such that the ML methods may only be needed for higher resolution predictions.

Table 3 compares the error metrics under both scenarios and shows that when using the upsampled low-resolution data, both the pixel-wise error and physical metrics improve, except for the compliance error in seen boundary conditions. Remarkably, for compliance error in seen boundary conditions, the use of upsampled low-resolution data results in predicting almost 200 more feasible designs. These observations suggest that improving the performance of the low-resolution generative model could be a good start for improving the cascading diffusion pipeline.

4.5 Limitations and Future Work

While our study demonstrates the effectiveness and generalizability of the cCDM compared to the cGAN model with transfer learning, several challenges remain. One important pitfall of the diffusion model, in contrast to the cGAN model, is its computational inference time. For instance, our low-resolution diffusion model on NVIDIA GeForce RTX 3080 Ti required 4.81 seconds,

whereas the high-resolution model took 10.93 seconds to generate a single topology. In comparison, the cGAN model only took 2.09 seconds to produce similar results on the same hardware. Reducing the computational time of diffusion models is actively under investigation [42] which could help for effectiveness of cCDM.

Another significant limitation of our work is its focus on 2D structural problems. Future research directions could include extending our model to address three-dimensional or more complex problems, thereby broadening its applicability and impact. Additionally, a key challenge in using generative models for topology optimization lies in incorporating objective values and ensuring feasibility without imposing additional computational costs. While some researchers have made progress in this area [7, 8, 25], further investigations are needed to streamline this process and minimize associated costs.

Furthermore, future work could explore quantifying the tradeoff between the cost of low-resolution and high-resolution samples, particularly regarding their impact on generative model performance. Understanding this tradeoff could provide information regarding the optimal ratio of sample types, thereby maximizing the effectiveness of generative models for multiresolution tasks.

5. CONCLUSIONS

We proposed the cCDM to investigate the performance of the diffusion model in the context of the multi-resolution for engineering design problem. We benchmarked our model against cGAN model with transfer learning. Our results demonstrate that the cCDM outperforms the cGAN model in terms of capturing finer details, preserving volume fraction constraints, and minimizing compliance errors in multi-resolution task when there are enough (more than 102 designs) high-resolution training data. Furthermore, we studied the impact of training data size on both models’ performance. While both models showed a decrease in performance with reduced training data, the cCDM lost its superiority to cGAN model with transfer learning in limited training data and we found the breakeven point for this change. While we showed that the diffusion model might produce the best pixel-wise performance in low-resolution and high-resolution it does not guarantee the best performance in design metrics. Also, our experiment isolating the effect of the low-resolution generative model on the downstream super resolution tasks showed that meaningful improvements in the low-resolution model can significantly improve the joint accuracy of our cascaded pipeline. In conclusion, our study underscores the potential of cCDM as effective frameworks for multi-resolution inverse design problem tasks offering superior performance and generalizability compared to cGAN model with transfer learning when there exists enough high-resolution training data.

ACKNOWLEDGEMENTS

We acknowledge the support from the National Science Foundation through award #1943699. We also acknowledge several helpful conversations with Mohammad Mahdi Behzadi that allowed us to accurately replicate the results from their original

Model	Boundary Condition	Average MSE ($\times 10^{-4}$)	Average VFE ($\times 10^{-4}$)	Median CE ($\times 10^{-2}$)
SR Model using SIMP	Seen	42.69 ± 6.51	49.16 ± 2.87	1.13 ± 0.12
	Unseen	103.38 ± 9.65	66.9 ± 4.85	1.62 ± 0.26
cCDM	Seen	44.96 ± 6.36	99.11 ± 5.32	0.39 ± 0.20
	Unseen	465.14 ± 23.57	464.65 ± 29.27	10.44 ± 5.40

TABLE 3: PERFORMANCE METRICS COMPARISON IN HIGH RESOLUTION (128X128) ON SEEN AND UNSEEN BOUNDARY CONDITIONS FOR MODEL ISOLATED GENERATIVE MODEL AND CASCDED MODEL. VALUES AFTER \pm INDICATE THE 95 % CONFIDENCE INTERVAL AROUND AVERAGES/MEDIANS

paper. We also acknowledge Cashen Diniz for helpful discussions regarding the implementation of diffusion models.

REFERENCES

[1] Sigmund, Ole and Maute, Kurt. “Topology optimization approaches: A comparative review.” *Structural and multidisciplinary optimization* Vol. 48 No. 6 (2013): pp. 1031–1055.

[2] Bendsøe, Martin P. “Optimal shape design as a material distribution problem.” *Structural optimization* Vol. 1 (1989): pp. 193–202.

[3] Tcherniak, Dmitri. “Topology optimization of resonating structures using SIMP method.” *International Journal for Numerical Methods in Engineering* Vol. 54 No. 11 (2002): pp. 1605–1622.

[4] Habibi, Milad, Wang, Jun and Fuge, Mark. “When Is it Actually Worth Learning Inverse Design?” *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 87301: p. V03AT03A025. 2023. American Society of Mechanical Engineers.

[5] Park, Sung Cheol, Park, Min Kyu and Kang, Moon Gi. “Super-resolution image reconstruction: a technical overview.” *IEEE signal processing magazine* Vol. 20 No. 3 (2003): pp. 21–36.

[6] Yu, Yonggyun, Hur, Taeil, Jung, Jaeho and Jang, In Gwon. “Deep learning for determining a near-optimal topological design without any iteration.” *Structural and Multidisciplinary Optimization* Vol. 59 No. 3 (2019): pp. 787–799.

[7] Behzadi, Mohammad Mahdi and Ilieş, Horea T. “Gantl: Toward practical and real-time topology optimization with conditional generative adversarial networks and transfer learning.” *Journal of Mechanical Design* Vol. 144 No. 2 (2022): p. 021711.

[8] Mazé, François and Ahmed, Faez. “Diffusion models beat gans on topology optimization.” *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), Washington, DC*. 2023.

[9] Chen, Qiuyi, Wang, Jun, Pope, Phillip, Chen, Wei Wayne and Fuge, Mark. “Inverse Design of 2D Airfoils using Conditional Generative Models and Surrogate Log-Likelihoods.” *Journal of Mechanical Design* (2021): pp. 1–22.

[10] Lee, Xian Yeow, Balu, Aditya, Stoecklein, Daniel, Ganapathysubramanian, Baskar and Sarkar, Soumik. “A case study of deep reinforcement learning for engineering design: Application to microfluidic devices for flow sculpting.” *Journal of Mechanical Design* Vol. 141 No. 11 (2019).

[11] Kim, Baekjun, Lee, Sangwon and Kim, Jihan. “Inverse design of porous materials using artificial neural networks.” *Science advances* Vol. 6 No. 1 (2020): p. eaax9324.

[12] Kim, Baekjun, Lee, Sangwon and Kim, Jihan. “Inverse design of porous materials using artificial neural networks.” *Science advances* Vol. 6 No. 1 (2020): p. eaax9324.

[13] Challapalli, Adithya, Patel, Dhrumil and Li, Gouqiang. “Inverse machine learning framework for optimizing lightweight metamaterials.” *Materials & Design* Vol. 208 (2021): p. 109937.

[14] Sekar, Vinothkumar, Zhang, Mengqi, Shu, Chang and Khoo, Boo Cheong. “Inverse design of airfoil using a deep convolutional neural network.” *Aiaa Journal* Vol. 57 No. 3 (2019): pp. 993–1003.

[15] Bernard, Shai, Wang, Jun and Fuge, Mark. “Mean Squared Error May Lead You Astray When Optimizing Your Inverse Design Methods.” *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 86229: p. V03AT03A004. 2022. American Society of Mechanical Engineers.

[16] Giannone, Giorgio and Ahmed, Faez. “Diffusing the optimal topology: A generative optimization approach.” *arXiv preprint arXiv:2303.09760* (2023).

[17] Bendsøe, Martin Philip and Kikuchi, Noboru. “Generating optimal topologies in structural design using a homogenization method.” *Computer methods in applied mechanics and engineering* Vol. 71 No. 2 (1988): pp. 197–224.

[18] Sigmund, Ole. “A 99 line topology optimization code written in Matlab.” *Structural and multidisciplinary optimization* Vol. 21 (2001): pp. 120–127.

[19] Regenwetter, Lyle, Nobari, Amin Heyrani and Ahmed, Faez. “Deep generative models in engineering design: A review.” *Journal of Mechanical Design* Vol. 144 No. 7 (2022): p. 071704.

[20] Sosnovik, Ivan and Oseledets, Ivan. “Neural networks for topology optimization.” *Russian Journal of Numerical Analysis and Mathematical Modelling* Vol. 34 No. 4 (2019): pp. 215–223.

[21] Rawat, Sharad and Shen, MH Herman. “Application of adversarial networks for 3D structural topology optimization.” Technical report no. SAE Technical Paper. 2019.

[22] Sharpe, Conner and Seepersad, Carolyn Conner. “Topology design with conditional generative adversarial networks.”

International design engineering technical conferences and computers and information in engineering conference, Vol. 59186: p. V02AT03A062. 2019. American Society of Mechanical Engineers.

[23] Guo, Tinghao, Lohan, Danny J, Cang, Ruijin, Ren, Max Yi and Allison, James T. "An indirect design representation for topology optimization using variational autoencoder and style transfer." *2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*: p. 0804. 2018.

[24] Nie, Zhenguo, Lin, Tong, Jiang, Haoliang and Kara, Levant Burak. "Topologygan: Topology optimization using generative adversarial networks based on physical fields over the initial domain." *Journal of Mechanical Design* Vol. 143 No. 3 (2021): p. 031715.

[25] Giannone, Giorgio, Srivastava, Akash, Winther, Ole and Ahmed, Faez. "Aligning Optimization Trajectories with Diffusion Models for Constrained Design Generation." *arXiv preprint arXiv:2305.18470* (2023).

[26] Cang, Ruijin, Yao, Hope and Ren, Yi. "One-shot generation of near-optimal topology through theory-driven machine learning." *Computer-Aided Design* Vol. 109 (2019): pp. 12–21.

[27] Jang, Seowoo, Yoo, Soyoung and Kang, Namwoo. "Generative design by reinforcement learning: enhancing the diversity of topology optimization designs." *Computer-Aided Design* Vol. 146 (2022): p. 103225.

[28] Khan, Arbaaz, Midha, Chetan and Lowther, David. "Reinforcement learning for topology optimization of a synchronous reluctance motor." *IEEE Transactions on Magnetics* Vol. 58 No. 9 (2022): pp. 1–4.

[29] Behzadi, Mohammad Mahdi and Ilieş, Horea T. "Real-time topology optimization in 3d via deep transfer learning." *Computer-Aided Design* Vol. 135 (2021): p. 103014.

[30] Farsiu, Sina, Robinson, Dirk, Elad, Michael and Milanfar, Peyman. "Advances and challenges in super-resolution." *International Journal of Imaging Systems and Technology* Vol. 14 No. 2 (2004): pp. 47–57.

[31] Li, Baotong, Huang, Congjia, Li, Xin, Zheng, Shuai and Hong, Jun. "Non-iterative structural topology optimization using deep learning." *Computer-Aided Design* Vol. 115 (2019): pp. 172–180.

[32] Zhuang, Fuzhen, Qi, Zhiyuan, Duan, Keyu, Xi, Dongbo, Zhu, Yongchun, Zhu, Hengshu, Xiong, Hui and He, Qing. "A comprehensive survey on transfer learning." *Proceedings of the IEEE* Vol. 109 No. 1 (2020): pp. 43–76.

[33] Weiss, Karl, Khoshgoftaar, Taghi M and Wang, DingDing. "A survey of transfer learning." *Journal of Big data* Vol. 3 No. 1 (2016): pp. 1–40.

[34] Dhariwal, Prafulla and Nichol, Alexander. "Diffusion models beat gans on image synthesis." *Advances in neural information processing systems* Vol. 34 (2021): pp. 8780–8794.

[35] Sohl-Dickstein, Jascha, Weiss, Eric, Maheswaranathan, Niru and Ganguli, Surya. "Deep unsupervised learning using nonequilibrium thermodynamics." *International conference on machine learning*: pp. 2256–2265. 2015. PMLR.

[36] Ho, Jonathan, Jain, Ajay and Abbeel, Pieter. "Denoising diffusion probabilistic models." *Advances in neural information processing systems* Vol. 33 (2020): pp. 6840–6851.

[37] Luo, Calvin. "Understanding diffusion models: A unified perspective." *arXiv preprint arXiv:2208.11970* (2022).

[38] Ho, Jonathan, Saharia, Chitwan, Chan, William, Fleet, David J, Norouzi, Mohammad and Salimans, Tim. "Cascaded diffusion models for high fidelity image generation." *The Journal of Machine Learning Research* Vol. 23 No. 1 (2022): pp. 2249–2281.

[39] Saharia, Chitwan, Ho, Jonathan, Chan, William, Salimans, Tim, Fleet, David J and Norouzi, Mohammad. "Image super-resolution via iterative refinement." *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 45 No. 4 (2022): pp. 4713–4726.

[40] Rampas, Dominic. "Diffusion Models pytorch." <https://github.com/dome272/Diffusion-Models-pytorch> (2022). Accessed: 2022.

[41] Andreassen, Erik, Clausen, Anders, Schevenels, Mattias, Lazarov, Boyan S and Sigmund, Ole. "Efficient topology optimization in MATLAB using 88 lines of code." *Structural and Multidisciplinary Optimization* Vol. 43 (2011): pp. 1–16.

[42] Ma, Hengyuan, Zhang, Li, Zhu, Xiatian and Feng, Jianfeng. "Accelerating score-based generative models with preconditioned diffusion sampling." *European Conference on Computer Vision*: pp. 1–16. 2022. Springer.

SUPPLEMENTAL MATERIAL

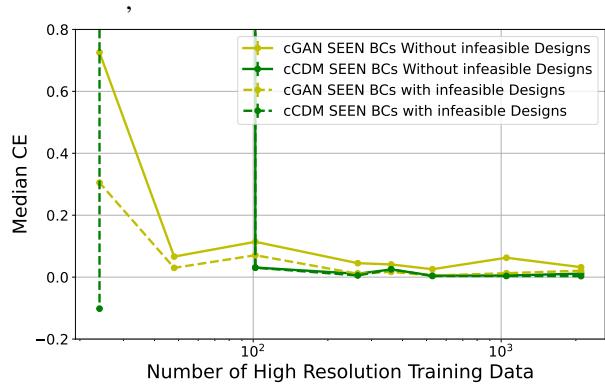


FIGURE 8: COMPARISON THE MEDIAN OF CE BETWEEN THE CGAN MODEL WITH TRANSFER LEARNING AND THE CASCADED DIFFUSION MODELS ACROSS SEEN TEST DATASETS WHEN THE MODELS' PREDICTION INCLUDES/EXCLUDES INFEASIBLE DESIGNS.

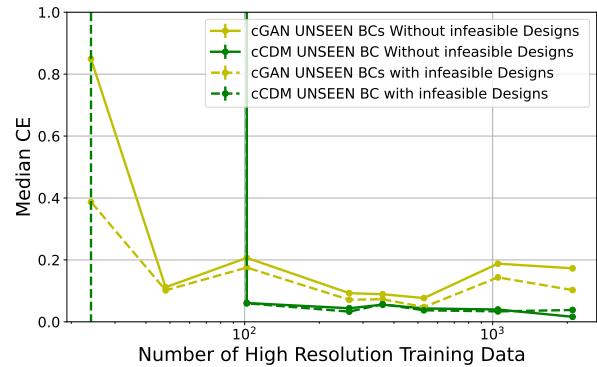


FIGURE 9: COMPARISON OF THE MEDIAN OF CE BETWEEN THE CGAN MODEL WITH TRANSFER LEARNING AND THE CASCADED DIFFUSION MODELS ACROSS UNSEEN TEST DATASETS WHEN THE MODELS' PREDICTION INCLUDES/EXCLUDES INFEASIBLE DESIGNS.