

Date of current version October 13, 2020.

Digital Object Identifier None

Attribute-Aware Generative Design with Generative Adversarial Networks

CHENXI YUAN, MOHSEN MOGHADDAM

Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115, United States

Corresponding author: Mohsen Moghaddam (e-mail: mohsen@northeastern.edu).

ABSTRACT The designers' tendency to adhere to a specific mental set and heavy emotional investment in their initial ideas often limit their ability to innovate during the design ideation process. The shrinking time-to-market and the growing diversity of users' needs further exacerbate this gap. Recent advances in deep generative models have created new possibilities to overcome the cognitive obstacles of designers through automated generation or editing of design concepts. This paper explores the capabilities of generative adversarial networks (GAN) for automated, attribute-aware generative design of the visual attributes of a product. Specifically, a design attribute GAN (DA-GAN) model is developed for automated generation of fashion product images with the desired visual attributes. Experiments on a large fashion dataset signify the potentials of GAN for attribute-aware generative design, verify the ability of editing attributes with relatively higher accuracy and uncover several key challenges and research questions for future work.

INDEX TERMS Design automation; Generative modeling; Conceptual design and ideation

I. INTRODUCTION

The rapid development of artificial intelligence (AI) and automation technologies in recent years has created unprecedented, transformative capabilities for product design. Although human involvement is still an indispensable element of the creative design process, the shrinking product life cycles and the growing needs for massive design idea generation/exploration [1] and avoiding fixation on few ideas [2] inevitably demand augmented human performance through design automation. Technology-driven innovation using AI and machine learning has become an essential success factor for product design firms in the 21st century. In the fashion industry, for example, McKinsey & Company reports that over 140% of the global fashion industry profit is generated by the leading 20% of the fashion brands [3]. As a result, significant recent progress has been made in adopting AI and machine learning techniques for augmented and personalized design.

Deep generative models have been recently adopted for design automation with the goal of improving the performance of the design team through co-creation with AI [4]. In the fashion industry, deep generative design has recently received significant attention in view of the rapidly growing global need for mass-personalization and "fast-fashion". Recent applications of AI for design automation range from style matching [5]–[7] to trend forecasting [8], interactive

search [9], [10], style recommendation [11], [12], virtual try-on apps [13], and clothing type and style classification [14], [15]. The vision of AI and machine learning research in the fashion industry is to directly influence and enhance the purchasing behavior of customers, the garment design thinking and ideation processes, the user-centered design and mass-personalization knowledge, and the ability of the fashion industry to adapt their product development strategies accordingly [16].

This article investigates how generative adversarial networks (GAN) [17] can enable automated attribute-level editing of past successful designs to inform new product design and development processes. Attribute editing with GAN involves making translations/adjustments to images based on the target attributes to generate a new sample with desired attributes while preserving other details of the original image. Current GAN-based attribute editing research is predominantly centered on human face images [18]–[20]. The facial attribute editing task allows to edit a face image by manipulating one or multiple attributes of interest such as hair color, expression, mustache, and age [21], [22]. The attribute editing can play significant role in the new product ideation and design process. For fashion products, the analogous visual attributes of interest may include style type, sleeve length, color, and pattern, among others. The ability to manipulate the attributes of a prior design is particularly useful

in situations where customers are not satisfied with certain attributes or would like to explore various combinations of them [15].

Conditional GAN (cGAN) [23] is an extension of the original GAN formulation [17] which allows to generate samples conditioned on user-defined attributes that control the generative process. Among the variety of conditional GAN models proposed to date [24]–[26], attribute GAN (AttGAN) [27] has proven effective in generating realistic edited images with desired attributes on human face data. AttGAN can generate visually-appealing results with fine facial details in comparison with the state-of-the-art GAN models.

This article develops and tests a design-attribute GAN (DA-GAN) model to enable attribute-level editing of past, successful fashion product designs while preserving other visual aspects and attributes. The motivation behind this work is that although AttGAN has demonstrated great performance in facial attribute editing, there is no proof or indication that it can be directly applied for attribute-level editing of fashion data such as garment images with acceptable performance.

This article contributes to the current knowledge of generative design with GAN in two ways, as follows:

- 1) Preliminary experiments are conducted on a large fashion dataset consisting of 13,221 garment images along with 22 attribute values, which show that the great performance of AttGAN [27] on the human face editing task cannot be achieved on the fashion editing task. This finding indicates that AttGAN is not a “one-size-fits-all” attribute editing GAN model and is indeed sensitive to the task. This article hypothesizes the underlying reason to stem from the *relative size* of editing which, unlike human faces, corresponds to a large area of a garment image (e.g., entire sleeve or collar), and thus conducts an analytical assessment to investigate the poor performance of AttGAN on the fashion datasets under study.
- 2) A novel DA-GAN formulation is then proposed which is proven to address the identified limitations of AttGAN. The DA-GAN model is tested on the same fashion dataset, for editing the images with respect to five desired attributes including “vest”, “polo”, “hoodie”, “blouse” and “T-shirt” (e.g., selecting the attribute “vest” is desired to turn any type of shirt into a vest). Numerical experiments indicate significant improvement in successful editing of different attributes such as sleeve length, color, pattern and clothes type, while preserving the remainder of the original garment image.

The remainder of this article is organized as follows. Section II provides an overview of the related work, specifically attribute-aware GAN. Section III presents the DA-GAN methodology and propositions. Section IV presents the experimental results and analyses. Section V discusses the capabilities and limitations of GAN for attribute-aware generative design, and provides directions for future research.

II. RELATED WORK

This section provides a brief overview of GAN and its several extensions for attribute-aware image generation.

A. GENERATIVE ADVERSARIAL NETWORKS

Since its introduction in 2014, GAN [17] continues to attract growing interests in the deep learning community and has been applied to various domains such as computer vision [28]–[33], natural language processing [34], [35], time series synthesis [36], [37], and semantic segmentation [38], [39]. Specifically, GAN has shown significant recent success in the field of computer vision on a variety of tasks such as image generation [28], [29], image to image translation [30], [31], and image super-resolution [32], [33]. The standard GAN structure comprises two neural networks: a generator G and a discriminator D which are iteratively trained by competing against each other in a minimax game, where the generator attempts to produce realistic samples while the discriminator attempts to distinguish the fake samples from the real ones. The parameters of both networks are updated through backpropagation with the following learning objective:

$$\min_G \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}} [\log(1 - D(G(\mathbf{z})))] \quad (1)$$

where \mathbf{z} is a random or encoded vector, p_{data} is the empirical distribution of training images, and $p_{\mathbf{z}}$ is the prior distribution of \mathbf{z} (e.g., normal distribution).

In the standard GAN model, there is no control over the modes of the data being generated. In cGAN [23], however, the generative process is conditioned to generate images based on a user-defined vector of features. The generator learns to generate a fake sample with a specific condition or characteristics rather than a generic sample from unknown noise distribution. The learning objective of cGAN is as follow:

$$\min_G \max_D \mathbb{E}_{\mathbf{x}, \mathbf{y} \sim p_{\text{data}}} [\log D(\mathbf{x}, \mathbf{b})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}, \mathbf{b} \sim p_{\mathbf{b}}} [\log(1 - D(G(\mathbf{z}, \mathbf{b})), \mathbf{b})], \quad (2)$$

where \mathbf{b} is the extra information (e.g., class labels, attribute information) for a given real sample \mathbf{x} as input. cGAN allows to generate samples to be controlled by using the constrained variation \mathbf{b} .

In cGAN, the generation of samples can be conditioned on class information [40], text description [24], [41], audio [42], [43], skeleton [44], [45], and attributes [46]. Using an encoder-decoder architecture, the conditions can be applied to conduct domain changes on images such as image editing [47], image segmentation [48] and image inpainting [49]. In the context of fashion design applications, researchers have applied GANs for a variety of applications such as: (1) textures filling [50] which allows users to try texture patches on a sketch to control the desired output texture, (2) texture transferring [51] where given a basic clothing image and a fashion style image, they generate a clothing image with the certain style in real time, (3) virtual try-on [15]

aimed at creating new clothing on a human body based on textual descriptions, (4) interactive image editing [52] where users can guide an agent to edit images via multi-turn via conversational language, (5) fashion recommendation [53] in which the model can be used for personalized design recommendation, and (6) clothes matching [54] where a multi-discriminator cGAN generates collocations of clothing pairs supervised by semantic attributes and implements clothing image translation between the specific domains based on an attribute-matching.

B. ATTRIBUTE-AWARE GAN

The introduction of GAN has created unprecedented capabilities for automated image generation and editing tasks. In the cGAN space, recent studies have focused on generating images from images [55], from text (e.g., captions) [24], [56], from long-paragraphs [57]), and from attributes [58]. Generating images from attributes, also known as attribute-aware image generation, is an important learning task that can automatically change various aspects of images with minimal human intervention. In this case, visual attribute vectors are regarded as the conditional information and embedded into both the generator and discriminator, encouraging synthesized images to be faithful to the visual attributes of the corresponding inputs [59], [60].

Among various attribute-aware image generation tasks, facial attribute editing has been widely studied due to the detailed description of human faces. IcGAN [25] introduces an encoder to cGANs forming an invertible conditional GANs (IcGAN) for facial attribute editing. IcGAN can modify real images of faces conditioned on arbitrary attributes by mapping a real image into a latent representation and an attribute vector. ResGAN [46] learns the corresponding residual image defined as the difference between images before and after the manipulation. The residual images are then added to the input images as the final outputs. In this way, the manipulation can be operated efficiently with modest pixel modification. SaGAN [61] introduces a spatial attention mechanism which ensures the manipulation of attributes only within attribute-specific regions while keeping the rest of irrelevant regions unchanged.

A major limitation of conventional cGAN is that the user-defined attributes/labels affect the editing of the entire image including the parts unrelated to the desired attributes/labels. Attribute GAN (AttGAN) [27] proposes an effective framework comprising an attribute classification loss, a reconstruction learning loss, and an adversarial learning loss, which is capable of editing specific facial attributes while preserving other “attribute-excluding” details of the original image. Eq. 3 presents the learning objective of the AttGAN generator and Eq. 4 presents the learning objective of the AttGAN discriminator and classifier:

$$\min_{G_{enc}, G_{dec}} \mathcal{L}_{enc, dec} = \lambda_1 \mathcal{L}_{rec} + \lambda_2 \mathcal{L}_{cls_g} + \mathcal{L}_{adv_g}, \quad (3)$$

$$\min_{D, C} \mathcal{L}_{dis, cls} = \lambda_3 \mathcal{L}_{cls_c} + \mathcal{L}_{adv_d}, \quad (4)$$

where \mathcal{L}_{rec} is the reconstruction loss for satisfactory preservation of attribute-excluding details, \mathcal{L}_{cls} is the classification constraint to guarantee the correct editing of the desired attributes, and \mathcal{L}_{adv} is the adversarial learning employed for visually realistic editing. λ_1, λ_2 and λ_3 are hyperparameters that control the importance of different terms and are tuned experimentally.

The majority of the attribute-aware image generation methods summarized above have been designed around facial attribute editing without any indication or proof of their applicability to other domains such as fashion product design. Fashion-AttGAN [62] introduces an attribute-aware fashion editing model based on AttGAN model. However, their attributes are limited to the color and sleeve length. To express the practicability and wide applicability of generative model on product design domain, the proposed DA-GAN (Design-Attribute GAN) builds upon AttGAN [27] with a new loss function formulation and by utilizing a different discriminator loss based on DCGAN [63]. The development of the DA-GAN is motivated by preliminary experiments that show AttGAN is not directly applicable for attribute-aware editing of garment images. The following section first conducts an analytical assessment of the AttGAN model to mathematically illustrate and prove the underlying reasons behind its poor performance on fashion datasets in the form of two propositions. A new formulation, DA-GAN, is then proposed to address those limitations accordingly.

III. METHODOLOGY

This section proposes a new model, DA-GAN, informed by an analytical assessment of the poor performance of AttGAN on fashion product images. AttGAN has shown great performance on facial image editing with binary attributes (e.g., {mustache, no-mustache}) and is used as our baseline model. A schematic of the proposed DA-GAN model is shown in Figure 1. Inspired by AttGAN’s success in human facial attribute editing, the authors first attempted to utilize the original AttGAN model for attribute-level editing of fashion product images. The preliminary observations showed that AttGAN model does not perform as expected on fashion data such as garment images. Specifically, the observation was that although AttGAN can reconstruct original fashion images, it is unable to generate new clear images with the desired attributes modified. The underlying reasons behind such poor performance on fashion data are elaborated and addressed next.

A. CLASSIFICATION-RECONSTRUCTION CONFLICT

To understand the reasons behind the poor performance of AttGAN of fashion data, an analytical assessment of the algorithmic aspects of AttGAN is conducted in this section. In Eq. (3), \mathcal{L}_{cls_g} is the attribute classification loss, employed to guide the generative process to learn and edit the desired attributes. The reconstruction loss \mathcal{L}_{rec} , on the other hand, is intended to enable the decoder to reconstruct the original input images so that the generated samples can preserve the

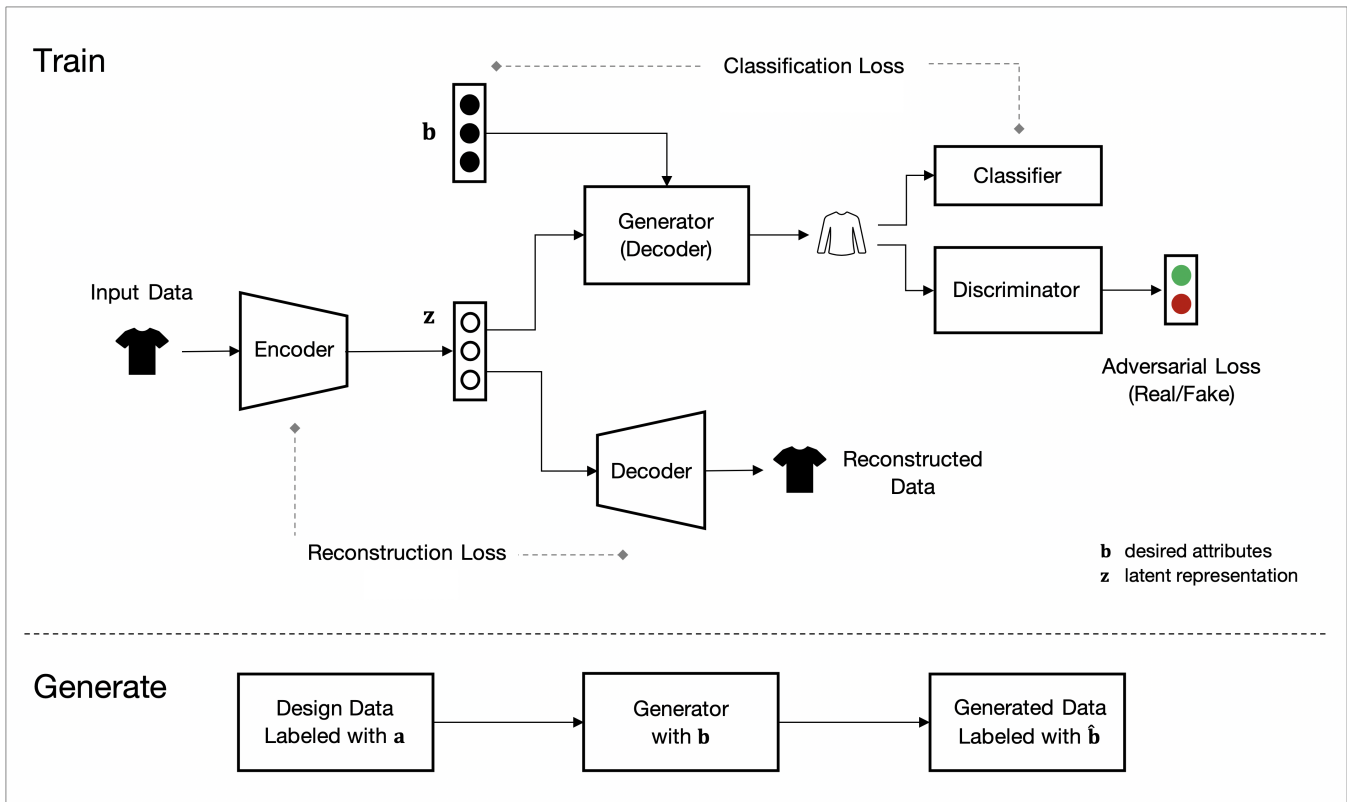


Figure 1: Overview of the proposed DA-GAN model for attribute-aware generative design (adapted from [27]).

attributes-excluding details. In the original AttGAN model, these two loss functions are both trained on the generator function. The problem associated with the poor performance of AttGAN on fashion data stems from an inherent conflict between these two loss functions.

The classification loss requires the generator to distinguish the desired attributes \mathbf{b} from the original images \mathbf{x}^a , by minimizing the summation of the binary cross entropy of the desired attributes and input images as follows:

$$\min_{G_{enc}, G_{dec}} \mathcal{L}_{cls_g} = \mathbb{E}_{\mathbf{x}^a \sim p_{data}, \mathbf{b} \sim p_{attr}} [\ell_g(\mathbf{x}^a, \mathbf{b})], \quad (5)$$

$$\ell_g(\mathbf{x}^a, \mathbf{b}) = \sum_{i=1}^n -b_i \log C_i(\mathbf{x}^b) - (1 - b_i) \log(1 - C_i(\mathbf{x}^b)), \quad (6)$$

where \mathbf{x}^b is the edited image expected to change the attributes of \mathbf{x}^a with respect to attributes \mathbf{b} . This is achieved by decoding latent representation \mathbf{z} conditioned on attributes \mathbf{b} : $\mathbf{x}^b = G_{dec}(\mathbf{z}, \mathbf{b})$, where \mathbf{z} is encoded from image \mathbf{x}^a with n binary attributes \mathbf{a} and is calculated as $\mathbf{z} = G_{enc}(\mathbf{x}^a)$. The generated image is thus formulated as $\mathbf{x}^b = G_{dec}(G_{enc}(\mathbf{x}^a), \mathbf{b})$.

The reconstruction loss, on the other hand, requires the generator to preserve original images as much as possible. This is accomplished by minimizing the Manhattan distance of the original attributes and the original images, as follows:

$$\min_{G_{enc}, G_{dec}} \mathcal{L}_{rec} = \mathbb{E}_{\mathbf{x}^a \sim p_{data}} [\|\mathbf{x}^a - \mathbf{x}^{\hat{a}}\|_1], \quad (7)$$

where $\mathbf{x}^{\hat{a}} = G_{dec}(G_{enc}(\mathbf{x}^a), \mathbf{a})$. The reconstruction loss enables the decoder to restore the original images conditioned on its own attributes \mathbf{a} from \mathbf{z} . Hence, the generator receives two tasks with exact opposite requirements during the training process, and the difficulty of achieving proper balance between these two conflicting tasks leads to poor performance in certain occasions, as elaborated in the following proposition.

Proposition 1. *The classification loss \mathcal{L}_{cls_g} and the reconstruction loss \mathcal{L}_{rec} are two conflicting objective functions assigned to the generator: $\mathcal{L}_{cls_g} = -\alpha \mathcal{L}_{rec}$, where α is a positive constant.*

Proof. The conflict stems from the assignment of both the classification loss and the reconstruction loss to the generator. The former requires the generator to generate new images with the maximum possible “distinction” from the original images while the latter requires the generator to generate new images with the maximum possible “similarity” to the original images. That is, for each input image \mathbf{x}^a , the reconstruction loss (the first term in Eq. 3) attempts to reconstruct the image by decoding $\mathbf{x}^{\hat{a}} = G_{dec}(G_{enc}(\mathbf{x}^a), \mathbf{a})$ from \mathbf{x}^a with respect to binary attributes \mathbf{a} . The classification loss (the second term in Eq. 3), on the other hand, attempts to decode a sample $\mathbf{x}^b = G_{dec}(G_{enc}(\mathbf{x}^a), \mathbf{b})$ from \mathbf{x}^a with respect to binary attributes \mathbf{b} . Yet, due to the binary nature of attribute vectors, certain elements of \mathbf{a} and \mathbf{b} are exactly the opposite

of each other. This, in turn, leads to a conflict between the the classification loss and the reconstruction loss in the form of $\mathcal{L}_{cls_g} = -\alpha\mathcal{L}_{rec}$. For example, \mathbf{a} may contain an attribute of “short sleeve” while \mathbf{b} may contain an attribute of “long sleeve” (e.g., $\mathbf{a} = [0, 1, \dots, 0, \dots]$, $\mathbf{b} = [0, 1, \dots, 1, \dots]$). In this case, the decoder (generator) faces a conflicting task of generating an image which simultaneously satisfies the requirements for both “long sleeve” and “short sleeve”. ■

B. ATTRIBUTE-SIZE TO IMAGE-SIZE RATIO

The classification loss (Eq. 5) requires the generator to guarantee the correct transformation of the desired attributes in the generated image by minimizing the cross entropy of the original image \mathbf{x}^a and the desired attributes \mathbf{b} . Our observation was that the relative sizes of attributes in \mathbf{b} to the image size vary significantly. For example, the sizes of facial attributes (e.g., {eyeglasses, no eyeglasses}, {mouth open, mouth closed}) are usually small, relative to the size of facial images, and the generated images with desired attributes have a high degree of similarity to the original ones. Smaller *attribute-size to image-size ratios* help the reconstruction loss achieve more desirable performance with less conflict with the classification loss. As the ratio of attribute size to image size increases (e.g., turn a “hoodie” into a “vest”), however, the *fake* image generated by the generator would require more significant distinction from the original image. In this case, the generator would need more flexibility to edit the image and the reconstruction loss would be more difficult to minimize for attribute-aware image generation tasks with relatively higher attribute-size to image-size ratios. This property is elaborated in the following proposition.

Proposition 2. *The reconstruction loss \mathcal{L}_{rec} is directly proportional to the attribute-size to image-size ratio.*

Proof. The proof is based on the idea of “image masking” from computer graphics. A masked image is simply an image where some of the pixel values are zeroed. The pixels with zero values are set to the background while the remaining, non-zero pixels are considered as the new actual image. Let \mathbf{s} be a binary image masking vector, where each element corresponds to one pixel in the image. Let also the non-zero values of \mathbf{s} represent the desired attribute to be edited in any arbitrary attribute-aware image generation task. An original image \mathbf{x}^a can be represented as:

$$\mathbf{x}^a = \mathbf{s} \circ \mathbf{x}^a + (\mathbf{1} - \mathbf{s}) \circ \mathbf{x}^a. \quad (8)$$

The reconstruction loss function (Eq. 7) can therefore be reformulated as:

$$\mathcal{L}_{rec} = \sum (\mathbf{s} \circ \|\mathbf{x}^a - \hat{\mathbf{x}}^a\|_1). \quad (9)$$

Assuming the dataset to be sufficiently large, $\|\mathbf{x}^a - \hat{\mathbf{x}}^a\|_1$ can be estimated by the mean distance between the pixel values of original images and new samples \mathbf{m} . Thus, Eq. 9 can be recast as follows:

$$\mathcal{L}_{rec} = \mathbf{1}^\top \mathbf{s} \circ \mathbf{m}. \quad (10)$$

Accordingly, it can be argued that $\mathcal{L}_{rec} \propto \mathbf{1}^\top \mathbf{s}$, where $\mathbf{1}^\top \mathbf{s}$ is an indicator of the attribute-size to image-size ratio. ■

Algorithm 1 DA-GAN.

```

1: Input: images  $\mathbf{X}$ , attributes  $\mathbf{A}$ , number of steps  $N$ 
2: for step  $\leftarrow 0$  to  $N$  do
3:   Sample batch  $\mathbf{x}^a \in \mathbf{X}$ ,  $\mathbf{a} \in \mathbf{A}$ ; randomly generate  $\mathbf{b}$ 
4:   for inner step  $\leftarrow 0$  to 5 do
5:     Minimize Eq. 4
6:   end for
7:   Minimize Eq. 11
8:   Minimize Eq. 12
9: end for
10: Output:  $G_{enc}$  and  $G_{dec}$ 

```

C. REFORMULATION OF LOSS FUNCTIONS

The aforementioned limitations of the original AttGAN model associated with the conflict between the classification and reconstruction loss functions (Section III-A) and sensitivity to the attribute-size to image-size ratio (Section III-B) are addressed by the proposed DA-GAN formulation. The proposed new formulation is primarily centered on the loss functions to enable more flexibility for the generator to generate images with larger attribute-size to image-size ratios while alleviating the conflict between the classification loss and the reconstruction loss. Accordingly, Eq. 3 is reformulated as follows:

$$\min_{G_{enc}, G_{dec}} \mathcal{L}_{enc, dec} = \lambda_1 \mathcal{L}_{rec} + \mathcal{L}_{adv_g}, \quad (11)$$

$$\min_{G_{enc}, G_{dec}} \mathcal{L}_{enc, dec} = \lambda_2 \mathcal{L}_{cls_g}. \quad (12)$$

The training procedure for the DA-GAN model is elaborated in Algorithm 1. It is worth noting that this new formulation is partially achieved through experimenting on the loss function formulations, where the reconstruction loss, the classification loss, and the adversarial loss were trained on the same function or independently. Our empirical analyses concluded that that breaking the classification loss out of the generator function leads to significant improvements in the quality of the generated images, notwithstanding the aforementioned challenges. Nevertheless, the DA-GAN model is still in at a preliminary stage and requires further development, as discussed next.

IV. EXPERIMENTS

This section presents the numerical experiments conducted on two large datasets to assess and validate the performance of DA-GAN against AttGAN as baseline.

A. DATASETS AND IMPLEMENTATION DETAILS

Two datasets that are rich in terms of both the number of images and the variety of attributes are used for conducting the experiments:

- 1) *Fashion dataset*. The AttGAN and DA-GAN model are tested on a fashion dataset, which contains 13221 images with the annotation of 22 binary attributes (i.e., with/without). Attributes with high frequency are chosen in all our experiments, including “vest”, “polo”, “stripe”, “short sleeve”, “long sleeve”, “red”, “yellow”, “blue”, “purple”, “black”, and “white”.
- 2) *Human face dataset*. CelebA [64] is a large dataset containing more than two hundred thousand images and 40 binary attributes for each image (i.e., with/without). Eight attributes are chosen for our experiments, including “bald”, “bangs”, “black hair”, “blond hair”, “brown hair”, “bushy eyebrows”, and “eyeglasses”.

The datasets are separated into training, test, and validation subsets. The experiments are conducted on TensorFlow, built upon an open-source code by [27]. The model is trained by Adam optimizer with the batch size of 32 and the learning rate of 0.0002.

B. RESULTS AND ANALYSES

In this section, AttGAN and DA-GAN are implemented on both datasets, and their performances are compared in terms of loss and quality of the generated images. Sensitivity analyses are also conducted on the parameters of both models.

a: AttGAN on Face and Fashion Images

The AttGAN model is implemented on the CelebA dataset and the fashion dataset for attribute-aware image generation with binary attributes, as shown in Figure 3. In spite of its great performance on all facial attribute editing tasks, our experiments show that AttGAN performs very poorly on garment images. The model does not preserve any garment patterns and is not even able to properly edit the images. The reason behind such poor performance by AttGAN on the fashion dataset is that the classification learning task is negatively influenced by the reconstruction learning task in the original AttGAN formulation. The encoder forces the reconstruction loss to be as small as possible, which in turn significantly limits the ability of the generator network to generate images with the desirable attributes edited. This is in part due to the inherent conflict between the classification loss and the reconstruction loss, and also because the attribute-size to image-size ratios are higher in the fashion images relative to the face images.

b: DA-GAN on Face and Fashion Images

To address the issues above, DA-GAN is implemented, where the classification loss is trained as an independent objective function to enhance the ability of the generator for attribute editing. As shown in Figure 2, DA-GAN significantly outperforms AttGAN on the fashion dataset in learning multiple attributes and changing the type of garment to “vest” or “polo”. However, it demonstrates worse performance than AttGAN on the face dataset—although attributes such as

Losses	CelebA Dataset		Fashion Dataset	
	AttGAN	DA-GAN	AttGAN	DA-GAN
Reconstruction loss	0.039	0.182	0.050	0.135
Classification loss	0.070	0.182	0.218	0.042

Table 1: The reconstruction and classification losses of DA-GAN and AttGAN.

“bold” or “eyeglasses” can be preserved by the DA-GAN model, its generated face images are too fake. In DA-GAN, the classification loss \mathcal{L}_{cls_g} is trained independently without any restrictions from the reconstruction loss and adversarial loss in the minimax game. This provides the model with more flexibility to generate good-quality, “fake” samples. This is necessary for the fashion attribute editing task because unlike the facial attributes, the attributes of garment products typically account for a relatively larger area of the image. This way, the generative model would have to generate more “wild” samples to incorporate the edited attributes in the original images. Facial attributes (e.g., “eyeglasses”, “bang”), on the other hand, have smaller attribute-size to image-size ratios, and thus the model not only needs to generate a new face with the desired attributes but also preserve more attribute-excluding details compared to garment images. This explains why AttGAN performs better than DA-GAN on the face dataset but worse on the fashion dataset.

c: Quantitative Analysis

From the visual analysis of , it is evident that part of the reason for the opposite performances of AttGAN and DA-GAN is that the attribute-size to image-size ratios are higher in the fashion images relative to the face images. This section conducts a quantitative analysis of the loss values to shed more light on this observation. The classification and reconstruction loss values are presented in Table 1. The losses are recorded after the convergence of the models. It is observed that AttGAN can generate clear face samples with the desired attributes successfully edited. The reconstruction loss of AttGAN is smaller than its classification loss. DA-GAN, on the other hand, show worse performance on the face dataset with relatively larger reconstruction loss.

The authors propose that the generator of AttGAN merely yields great performance in terms of reconstruction, and has limited flexibility to generate good-quality fake images with significant distinction from the original images. Experiments on the fashion dataset show that the DA-GAN model with better classification loss can generate good quality samples that are considerably different from the original images. On the fashion dataset, the reconstruction loss of AttGAN becomes too large because of the relatively higher attribute-size to image-size ratios, and thus it is unable to successfully carry out the attribute-aware image generation task on the garment images. The results tell us that GAN models are not necessarily directly transferable across different domains, and a state-of-the-art model such as AttGAN may perform very poorly on a similar task from a different domain.

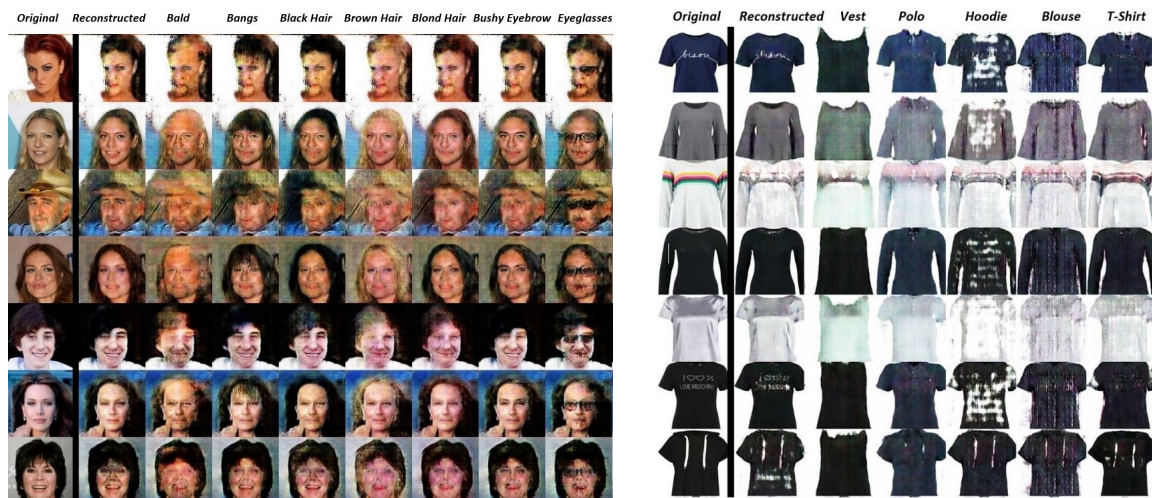


Figure 2: DA-GAN for attribute-aware image generation: face (left) vs. fashion (right).



Figure 3: AttGAN for attribute-aware image generation: face (left) vs. fashion (right).

d: Sensitivity Analysis

Informed by the previous AttGAN work [27], λ_2 is set to 10 in the experiments of implementing AttGAN on the face dataset. The goal of our sensitivity analysis is to observe the model performance in terms of reconstruction loss and classification loss with different value of λ_1 , because these two terms show high fluctuations during the training process and significantly impact the model performance. λ_2 is set to 3 in the experiments of implementing DA-GAN on the fashion dataset. The best empirical value of λ_1 is 100, as shown in Table 2. It is observed that the reconstruction loss decreases as λ_1 grows, and that the classification loss increases slightly. However, the classification loss has tremendous increase when the weight equal to 500. In order to keep the balance between the two loss terms, the best empirical value for both λ_2 and λ_3 is 100.

Models	λ_1	Reconstruction Loss	Classification Loss
AttGAN $\lambda_2 = 10$	500	0.027	0.275
	100	0.039	0.071
	50	0.116	0.032
	10	0.237	0.017
	5	0.506	0.011
DA-GAN $\lambda_2 = 3$	500	0.129	0.186
	100	0.135	0.042
	50	0.174	0.038
	10	0.309	0.025
	5	0.421	0.018

Table 2: The reconstruction loss and the classification loss with different λ_1 values.

e: Extended Attributes & Effects of Dataset Imbalance

Figure 4 shows the performance of the DA-GAN model on the attribute-editing task with eleven distinct attributes, including garment type (“vest”, “polo”), garment pattern (“slim horizontal stripes”), length of sleeves (“short sleeves”, “long sleeves”), and multiple colors (“red”, “yellow”, “blue”,

“purple”, “black”, “white”). As shown, the model can successfully generate new images with the desired colors and patterns, and change the length of sleeves. However, it is not able to learn the latent pattern associated with the attribute “polo”.

To solve this problem, another experiment is conducted on a narrowed dataset. The original dataset has over 12,000 images; however, unlike sleeve length and color that are the indispensable attributes of any garment type, attributes such as “polo” are relatively rare and thus cause the data to be imbalanced. Hence, 5,782 images with attributes of cloth type category (e.g., vest, polo, blouse, t-shirt, hoodie, one-piece dress) are selected from the original dataset. The DA-GAN model is then retrained on this narrowed dataset to generate samples with the desired attributes “vest” and “polo”. Results show that the DA-GAN model yields better performance on the narrowed dataset (Figure 5, right) than it does on the original dataset (Figure 5, left). With the narrowed dataset where each image is guaranteed to contain clothes type attributes, the model is more likely to capture these attributes and edit accordingly. This is further proof to the fact that training a cGAN model is highly sensitive to the balance of different attributes in the dataset. The imbalanced distribution of attributes hinders the ability of the model to learn the attributes with low frequency in the dataset.

C. PERFORMANCE ASSESSMENT

Attribute-aware generative models must have ability to edit the desired attributes with high accuracy and generate clear images with attribute-excluding features preserved. This section presents the results of a quantitative performance assessment of DA-GAN compared to AttGAN and Fashion-AttGAN as baseline models. An attribute classifier is trained on fashion dataset for attributes classification with cross-entropy loss. All three models are tested using this classifier, which examines whether the model has accurately edited the desired attributes and preserved the rest of them. The accuracy of each model is calculated as the ratio of all successful cases to all test samples.

The test dataset used for assessment contains 782 images. Eleven attributes are selected to be tested, including “vest”, “polo”, “stripe”, “short sleeve”, “long Sleeve”, “red”, “yellow”, “blue”, “purple”, “black”, and “white”. The classification accuracy results with respect to different attribute types are illustrated in Figure 6. Results show that DA-GAN achieves an accuracy of 63% on “vest”, 53% on “polo”, and 69% on “Stripe”, which are almost 2-5 times higher than the accuracy of Fashion-AttGAN and 6 times higher than the accuracy of AttGAN. Further, Fashion-AttGAN shows slightly higher accuracy than DA-GAN on sleeve length editing; yet, both significantly outperform AttGAN on this attribute type with around 70% accuracy compared to AttGAN’s accuracy of 22%-30%. In terms of color editing, all three models are able to successfully perform various colors editing with relatively high accuracy, ranging from 80% to 95%.

To sum up, DA-GAN is proven to have the ability to edit

all attributes with relatively higher accuracy on garment data. Although DA-GAN and Fashion-AttGAN perform comparably on short/long sleeve editing, DA-GAN yields higher accuracy on attributes “vest”, “polo” and “slim horizontal stripes”. AttGAN [27] is proven a successful model for editing facial images with respect to attributes such as “bald”, “eyeglass”, and “mustache”. Those attributes constitutes a small area of the image and AttGAN has capability to generate new edited images while preserving the attribute-excluding parts of the image. Since the attributes of garment images typically constitute a large area of the image, the original AttGAN model loses the balance between generating new attributes and preserving other parts of the image, leading to poor performance on garment dataset. The performance assessment results verify the ability of DA-GAN in editing attributes of images with high attribute-size to image-size ratios.

V. CONCLUDING REMARKS AND FUTURE WORK

This paper proposes the DA-GAN model for attribute-aware generative design of fashion products. The DA-GAN model has the ability to automatically edit garment images conditioned on certain user-defined, visual attributes. The performance of the generative model is experimented and tested on two large datasets, against the AttGAN model as a baseline. Two propositions are presented to assess and alleviate the limitations of the original AttGAN model associated with the conflict between classification learning and reconstruction learning, and the sensitivity of the model to the attribute-size to image-size ratios. A new formulation is then proposed, which is proven to be capable of addressing the aforementioned limitations of AttGAN. The performance assessment results verify the ability of the proposed DA-GAN in accomplishing attribute-aware image editing tasks with high accuracy, especially in tasks with high attribute-size to image-size ratios.

Despite the high attributes editing accuracy, the DA-GAN model yields a relatively high average preservation error of other attributes; a key issue which must be addressed in future research. Further development of the proposed DA-GAN model is indeed needed to generate images with higher resolution, improve the stability of the DA-GAN model, and broaden the scope of the proposed methodology beyond visual/image-based generative design. Evaluation of GAN’s performance and human involvement in machine-based design procedure will also be important areas to explore. In view of these research opportunities and the pressing need for enabling efficient and scalable solutions for attribute-aware generative design, the authors recommend the following critical research thrusts for future work.

a: Domain-Transferability of DA-GAN

An important observation through our experiments was that GAN models are generally sensitive to the problem domains and therefore need careful revision and hand-engineering of the algorithms based on the dataset and the target task.



Figure 4: DA-GAN on an extended set of fashion attributes.

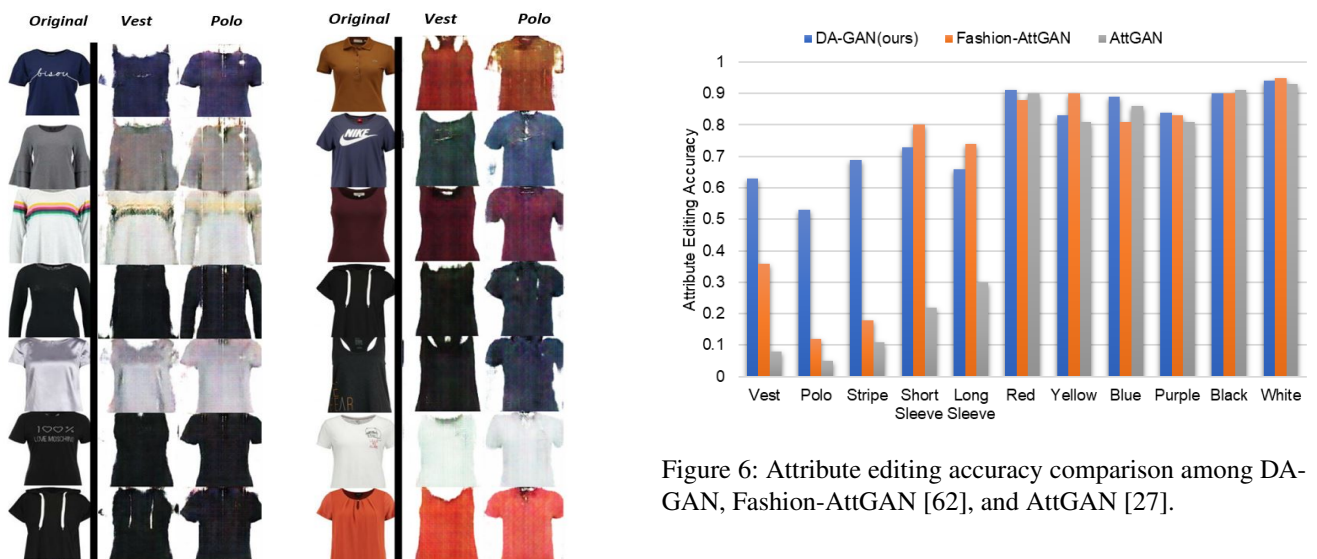


Figure 5: DA-GAN trained on the entire dataset (left) and narrowed dataset (right).

GAN is not a generic, domain-agnostic deep learning technique to be directly utilized in industrial applications such as generative design. As illustrated in the experiments, although AttGAN performs exceptionally well on the facial attribute editing task, it fails to carry out the same task on

a different dataset from the fashion domain. On the other hand, the proposed DA-GAN model is capable of generating good-quality fashion images with edited attributes, but acts poorly on the face dataset. Further, the model parameters of GAN need to be set up and tuned empirically, based on the outcomes of the training process. Tuning the parameters manually is a tedious way to improve model performance and effectiveness, and increases the complexity and cost of deployment and implementation in practical applications

such as generative design. The “black box” nature of GAN makes the process of discovering overfitting and other kinds of network architecting and training failures even more cumbersome. Future research must address these challenges to enable GAN-based generative design methods and tools that are scalable and transferable across different domains with minimal modeling and implementation efforts.

b: Sample Efficiency of DA-GAN

CelebA [64] is a large-scale face attributes dataset with more than 200K celebrity images, each with 40 attribute annotations. The images in this dataset cover large pose variations and background clutters. Unlike the face dataset that contains large diversities, large quantities, and rich annotations, it is hard to collect a similar large dataset with high-resolution and annotations in the fashion industry domain. In this work, only 13k images were used to train the GAN models. It is widely known in the field of deep learning [65] that small training data is likely to cause overfitting. Further, finding clean data with large quantities and high quality is another challenge in different generative design domains. Data imbalance is yet another problem. In experiments of DA-GAN, images with labels of garment types such as “polo” and “hoodie” are remarkably less frequent than labels like “color” and “sleeve length”. Thus, the model is likely to have limited ability to capture type related attributes because the data is insufficient and imbalanced. Those attributes would not be identified, generalization would exacerbate, and the model would keep learning the frequent, popular attributes and ignore the less-frequent ones. Future research must address these limitations by building novel, pretrained GANs that minimize the dependency on “big data” and the need for training from scratch.

c: Generative Design of Form and Function

Since the inception of generative models, from variational autoencoders and Boltzmann machines to GAN, there has been an exponential growth in generative modeling research and innovation. The majority of recent GAN literature, however, is concerned with theoretical developments applied on “toy problems” such as editing human faces, birds, and cats with limited practical applications. Part of the reason is that such datasets are clean, well-organized, freely-available, and enormous. Nevertheless, Deep generative models have been recently adopted for design automation [66]–[68] with the goal of improving designers’ performance through co-creation with AI. Specifically, GAN has shown tremendous success in a variety of generative design tasks, from topology optimization [69] to material design [70] and shape parametrization [68]. In line with Osborn’s rules for brainstorming [71], these generative models have proven effective in increasing the quantity of ideas at the designer’s disposal to inspire her exploration and avoid investing too heavily in few ideas. Despite significant recent progress, two major knowledge gaps limit the ability of state-of-the-art generative design models to effectively assist designers in early-

stage product development processes. First, current literature merely focuses on the generative design of “form”, disregarding other non-visual aspects associated with its “function” (e.g., architecture, materials, performance). Second, there is a lack of a standardized method of assessing the performance of the generated design concepts [69]. Few recent studies propose assessment mechanisms based on form-function relationships [68] (e.g., physics-based simulators); however, those mechanisms are domain-specific and applicable to a limited set of functional attributes (e.g., aerodynamic performance). Future research must build novel, verifiable GAN-based generative design techniques capable of conditioning the design concepts on both visual and functional attributes.

ACKNOWLEDGMENT

This work was partially supported by the FY20 TIER 1—Seed Grant/Proof of Concept Program sponsored by Northeastern University, Boston.

References

- [1] Luis A Vasconcelos, Carlos C Cardoso, Maria Sääksjärvi, Chih-Chun Chen, and Nathan Crilly. Inspiration and fixation: the influences of example designs and system properties in idea generation. *Journal of Mechanical Design*, 139(3), 2017.
- [2] Steven P Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L Schwartz, and Scott R Klemmer. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 17(4):1–24, 2010.
- [3] Imran Amed, Johanna Andersson, Achim Berg, Martine Drageset, Saskia Hedrich, and Sara Kappelmark. *The State of Fashion: Renewed optimism for the fashion industry*, 2018.
- [4] Christian E. Lopez, Scarlett R. Miller, and Conrad S. Tucker. Exploring biases between human and machine generated designs. *Journal of Mechanical Design, Transactions of the ASME*, 141(2), 2019.
- [5] Lukas Bossard, Matthias Dantone, Christian Leistner, Christian Wengert, Till Quack, and Luc Van Gool. Apparel classification with style. In *Asian conference on computer vision*, pages 321–335. Springer, 2012.
- [6] Yannis Kalantidis, Lyndon Kennedy, and Li-Jia Li. Getting the look: clothing recognition and segmentation for automatic product suggestions in everyday photos. In *Proceedings of the 3rd ACM conference on International conference on multimedia retrieval*, pages 105–112, 2013.
- [7] Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, and Xiaoou Tang. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1096–1104, 2016.
- [8] Ziad Al-Halah, Rainer Stiefelhofen, and Kristen Grauman. Fashion forward: Forecasting visual style in fashion. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 388–397, 2017.
- [9] Bo Zhao, Jiashi Feng, Xiao Wu, and Shuicheng Yan. Memory-augmented attribute manipulation networks for interactive fashion search. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1520–1528, 2017.
- [10] Adriana Kovashka, Devi Parikh, and Kristen Grauman. Whittlesearch: Image search with relative attribute feedback. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2973–2980. IEEE, 2012.
- [11] Yujie Lin, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Jun Ma, and Maarten de Rijke. Explainable fashion recommendation with joint outfit matching and comment generation. *arXiv preprint arXiv:1806.08977*, 2, 2018.
- [12] Edgar Simo-Serra, Sanja Fidler, Francesc Moreno-Noguer, and Raquel Urtasun. Neuroaesthetics in fashion: Modeling the perception of fashionability. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 869–877, 2015.
- [13] Xintong Han, Zuxuan Wu, Zhe Wu, Ruiqi Yu, and Larry S Davis. Viton: An image-based virtual try-on network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7543–7552, 2018.

- [14] Xiaodan Liang, Liang Lin, Wei Yang, Ping Luo, Junshi Huang, and Shuicheng Yan. Clothes co-parsing via joint image segmentation and labeling with application to clothing retrieval. *IEEE Transactions on Multimedia*, 18(6):1175–1186, 2016.
- [15] Shizhan Zhu, Raquel Urtasun, Sanja Fidler, Dahua Lin, and Chen Change Loy. Be your own prada: Fashion synthesis with structural coherence. In *Proceedings of the IEEE international conference on computer vision*, pages 1680–1688, 2017.
- [16] C. Giri, S. Jain, X. Zeng, and P. Brumiaux. A detailed review of artificial intelligence applied in the fashion and apparel industry. *IEEE Access*, 7:95376–95396, 2019.
- [17] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.
- [18] S. Liu, D. Li, T. Cao, Y. Sun, Y. Hu, and J. Ji. Gan-based face attribute editing. *IEEE Access*, 8:34854–34867, 2020.
- [19] Z. Tang, J. Yang, Z. Pei, X. Song, and B. Ge. Multi-process training gan for identity-preserving face synthesis. *IEEE Access*, 7:97641–97652, 2019.
- [20] D. Huang, X. Tao, J. Lu, and M. N. Do. Geometry-aware gan for face attribute transfer. *IEEE Access*, 7:145953–145969, 2019.
- [21] J. Zhao, X. Xie, L. Wang, M. Cao, and M. Zhang. Generating photographic faces from the sketch guided by attribute using gan. *IEEE Access*, 7:23844–23851, 2019.
- [22] Jiayuan Zhang, Ao Li, Yu Liu, and Minghui Wang. Adversarially regularized u-net-based gans for facial attribute modification and generation. *IEEE Access*, 7:86453–86462, 2019.
- [23] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.
- [24] Scott Reed, Zeynep Akata, Xinchun Yan, Lajanugen Logeswaran, Bernt Schiele, and Honglak Lee. Generative adversarial text to image synthesis. *arXiv preprint arXiv:1605.05396*, 2016.
- [25] Guim Perarnau, Joost Van De Weijer, Bogdan Raducanu, and Jose M Álvarez. Invertible conditional gans for image editing. *arXiv preprint arXiv:1611.06355*, 2016.
- [26] Takuhiro Kaneko, Kaoru Hiramatsu, and Kunio Kashino. Generative attribute controller with conditional filtered generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6089–6098, 2017.
- [27] Zhenliang He, Wangmeng Zuo, Meina Kan, Shiguang Shan, and Xilin Chen. AttnGAN: Facial attribute editing by only changing what you want. *IEEE Transactions on Image Processing*, 28(11):5464–5478, 2019.
- [28] J. Wang, P. Li, J. Deng, Y. Du, J. Zhuang, P. Liang, and P. Liu. Ca-gan: Class-condition attention gan for underwater image enhancement. *IEEE Access*, pages 1–1, 2020.
- [29] K. Xu, J. Cao, K. Xia, H. Yang, J. Zhu, C. Wu, Y. Jiang, and P. Qian. Multichannel residual conditional gan-leveraged abdominal pseudo-ct generation via dixon mr images. *IEEE Access*, 7:163823–163830, 2019.
- [30] W. Zhang. Generating adversarial examples in one shot with image-to-image translation gan. *IEEE Access*, 7:151103–151119, 2019.
- [31] F. Xiong, Q. Wang, and Q. Gao. Consistent embedded gan for image-to-image translation. *IEEE Access*, 7:126651–126661, 2019.
- [32] Y. Xi, J. Zheng, W. Jia, X. He, H. Li, Z. Ren, and K. Lam. See clearly in the distance: Representation learning gan for low resolution object recognition. *IEEE Access*, 8:53203–53214, 2020.
- [33] S. Huang, H. Zhou, Y. Liu, and R. Chen. High-resolution crowd density maps generation with multi-scale fusion conditional gan. *IEEE Access*, 8:108072–108087, 2020.
- [34] L. Shao, C. Liang, K. Wang, W. Cao, W. Zhang, G. Gui, and H. Sari. Attention gan-based method for designing intelligent making system. *IEEE Access*, 7:163097–163104, 2019.
- [35] Tong Che, Yanran Li, Ruixiang Zhang, R Devon Hjelm, Wenjie Li, Yangqiu Song, and Yoshua Bengio. Maximum-likelihood augmented discrete generative adversarial networks. *arXiv preprint arXiv:1702.07983*, 2017.
- [36] Cristóbal Esteban, Stephanie L Hyland, and Gunnar Rätsch. Real-valued (medical) time series generation with recurrent conditional gans. *arXiv preprint arXiv:1706.02633*, 2017.
- [37] Yonghong Luo, Xiangrui Cai, Ying Zhang, Jun Xu, et al. Multivariate time series imputation with generative adversarial networks. In *Advances in Neural Information Processing Systems*, pages 1596–1607, 2018.
- [38] Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. Semantic segmentation using adversarial networks. *arXiv preprint arXiv:1611.08408*, 2016.
- [39] X. Chen, L. Xu, H. Wei, Z. Shang, T. Zhang, and L. Zhang. Emotion interaction recognition based on deep adversarial network in interactive design for intelligent robot. *IEEE Access*, 7:166860–166868, 2019.
- [40] Augustus Odena, Christopher Olah, and Jonathon Shlens. Conditional image synthesis with auxiliary classifier gans. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2642–2651. *JMLR.org*, 2017.
- [41] Y. Cai, X. Wang, Z. Yu, F. Li, P. Xu, Y. Li, and L. Li. Dualattn-gan: Text to image synthesis with dual attentional generative adversarial network. *IEEE Access*, 7:183706–183716, 2019.
- [42] Y. Zhao, S. Takaki, H. Luong, J. Yamagishi, D. Saito, and N. Minematsu. Wasserstein gan and waveform loss-based acoustic model training for multi-speaker text-to-speech synthesis systems using a wavernet vocoder. *IEEE Access*, 6:60478–60488, 2018.
- [43] Lele Chen, Sudhanshu Srivastava, Zhiyao Duan, and Chenliang Xu. Deep cross-modal audio-visual generation. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, pages 349–357, 2017.
- [44] Liqian Ma, Xu Jia, Qianru Sun, Bernt Schiele, Tinne Tuytelaars, and Luc Van Gool. Pose guided person image generation. In *Advances in Neural Information Processing Systems*, pages 406–416, 2017.
- [45] Amit Raj, Patsorn Sangkloy, Huiwen Chang, Jingwan Lu, Duygu Ceylan, and James Hays. Swapnet: Garment transfer in single view images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 666–682, 2018.
- [46] Wei Shen and Rujie Liu. Learning residual images for face attribute manipulation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4030–4038, 2017.
- [47] L. Teng, Z. Fu, and Y. Yao. Interactive translation in echocardiography training system with enhanced cycle-gan. *IEEE Access*, 8:106147–106156, 2020.
- [48] Y. Li and L. Shen. cc-gan: A robust transfer-learning framework for hep-2 specimen image segmentation. *IEEE Access*, 6:14048–14058, 2018.
- [49] W. Cai and Z. Wei. Piigan: Generative adversarial networks for pluralistic image inpainting. *IEEE Access*, 8:48451–48463, 2020.
- [50] Wenqi Xian, Patsorn Sangkloy, Varun Agrawal, Amit Raj, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. Texturegan: Controlling deep image synthesis with texture patches. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8456–8465, 2018.
- [51] Shuhui Jiang and Yun Fu. Fashion style generator. In *IJCAI*, pages 3721–3727, 2017.
- [52] Yu Cheng, Zhe Gan, Yitong Li, Jingjing Liu, and Jianfeng Gao. Sequential attention gan for interactive image editing via dialogue. *arXiv preprint arXiv:1812.08352*, 2018.
- [53] Wang-Cheng Kang, Chen Fang, Zhaowen Wang, and Julian McAuley. Visually-aware fashion recommendation and design with generative image models. In *2017 IEEE International Conference on Data Mining (ICDM)*, pages 207–216. *IEEE*, 2017.
- [54] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *July 2019*.
- [55] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.
- [56] R. Yanagi, R. Togo, T. Ogawa, and M. Haseyama. Query is gan: Scene retrieval with attentional text-to-image generative adversarial network. *IEEE Access*, 7:153183–153193, 2019.
- [57] Yitong Li, Zhe Gan, Yelong Shen, Jingjing Liu, Yu Cheng, Yuexin Wu, Lawrence Carin, David Carlson, and Jianfeng Gao. Storygan: A sequential conditional gan for story visualization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6329–6338, 2019.
- [58] Mandar Dixit, Roland Kwitt, Marc Niethammer, and Nuno Vasconcelos. Aga: Attribute-guided augmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7455–7463, 2017.
- [59] Yunfan Liu, Qi Li, and Zhenan Sun. Attribute-aware face aging with wavelet-based generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11877–11886, 2019.
- [60] Yunfan Liu, Qi Li, Zhenan Sun, and Tieniu Tan. A3gan: An attribute-aware attentive generative adversarial network for face aging. *arXiv preprint arXiv:1911.06531*, 2019.

- [61] Gang Zhang, Meina Kan, Shiguang Shan, and Xilin Chen. Generative adversarial network with spatial attention for face attribute editing. In Proceedings of the European Conference on Computer Vision (ECCV), pages 417–432, 2018.
- [62] Qing Ping, Bing Wu, Wanying Ding, and Jiangbo Yuan. Fashion-attgan: Attribute-aware fashion editing with multi-objective gan. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages 0–0, 2019.
- [63] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.
- [64] Ziwei Liu, Ping Luo, Xiaoang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In Proceedings of the IEEE international conference on computer vision, pages 3730–3738, 2015.
- [65] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning. MIT Press, 1 edition, 2016.
- [66] D. Ha and D. Eck. A Neural Representation of Sketch Drawings, 2018.
- [67] Alexander Burnap, Ye Liu, Yanxin Pan, Honglak Lee, Richard Gonzalez, and Panos Y. Papalambros. Estimating and exploring the product form design space using deep generative models. In Proceedings of the ASME Design Engineering Technical Conference, volume 2A-2016. American Society of Mechanical Engineers (ASME), dec 2016.
- [68] Dule Shu, James Cunningham, Gary Stump, Simon W. Miller, Michael A. Yukish, Timothy W. Simpson, and Conrad S. Tucker. 3D Design Using Generative Adversarial Networks and Physics-based Validation. *Journal of Mechanical Design*, 142(July):1–51, 2019.
- [69] Sangeun Oh, Yongsu Jung, Seongsin Kim, Ikjin Lee, and Namwoo Kang. Deep generative design: Integration of topology optimization and generative models. *Journal of Mechanical Design, Transactions of the ASME*, 141(11), 2019.
- [70] Min Yang, Qiang Qu, Xiaojun Chen, Chaoxue Guo, Ying Shen, and Kai Lei. Feature-enhanced attention network for target-dependent sentiment classification. *Neurocomputing*, 307:91–97, 2018.
- [71] A.F. Osborn. *Applied Imagination*. Scribner's, New York, 1953.



MOHSEN MOGHADDAM, Ph.D., is an Assistant Professor of Mechanical and Industrial Engineering at Northeastern University. Prior to joining Northeastern, he was with the GE-Purdue Partnership in Research and Innovation in Advanced Manufacturing as a Postdoctoral Associate. He received his PhD from Purdue University in 2016. His areas of research interest include cyber-physical manufacturing, human-technology collaboration, user-centered design, artificial intelligence, and machine learning. He is coauthor of over 40 refereed journal articles and two books including *Revolutionizing Collaboration Through e-Work, e-Business, and e-Service* (Springer, 2015), and *Best Matching Theory & Applications* (Springer, 2017). He also served as a reviewer for several international journals such as *Int. J. Production Economics*, *Int. J. Production Research*, *J. Intelligent Manufacturing*, *Computers & Industrial Engineering*, *Decision Support Systems*, *Computers in Industry*, *IEEE Trans. on Industrial Informatics*, *IEEE Trans. on Systems, Man, & Cybernetics*, and *IEEE Trans. on Automation Science and Engineering*. His scholarly research is supported by the U.S. National Science Foundation (NSF) and industry.

...



CHENXI YUAN received the B.E. degree in Mechanical Engineering from Northwest University, China, in 2015 and the M.S. degree in Industrial and System Engineering from University of Florida, USA, in 2017. She is currently pursuing the Ph.D. degree in Industrial Engineering with Northeastern University. Her research interests include machine learning, artificial intelligence, and data science.