CHAPTER 8

Generative Models as Fashion Designers

The great thing about fashion is that it always looks forward.

—Oscar de la Renta

Amazon made headlines in 2017 for a controversial idea that it introduced to the public consciousness. Amazon claimed the ability to train a **generative adversarial network** (**GAN**), a type of **generative model**, to design garments. For many fashion industry professionals, this announcement set off alarms. The threat that the role of the fashion designer would soon become obsolete hit close to home for everyone.

AI Fashion Designer

Fashion has always been a repetition of ideas, but what makes it new is the way you put it together.

—Carolina Herrera, fashion designer

The use of a generative model as a fashion designer refers to a process of taking in a dataset of images, and outputting images that are visually similar but generated by the model. The images that make up the input

L. Luce, *Artificial Intelligence for Fashion*,
https://doi.org/10.1007/978-1-4842-3931-5_8

dataset could be of garments trending on social media or other channels. The use of real-time data and generative models for design purposes could give companies like Amazon an advantage in understanding demand for garments before producing them. If you're interested in the **data-mining** aspect of this concept, read on to Chapter 9.

There are limitations to the current proposal for a computer-based fashion designer. Generative models can create images of garments, which is useful for providing an educated jumping-off point when it comes to trend research. However, any fashion designer reading this book knows that the reality of their job only begins with that image. An image of a garment is not a design, it is not a tech pack, and it is not a garment. These models do not have an inherent understanding of the real world; they are only able to identify patterns in data they are given.

GANs do, however, hold other potential for applications related to fashion. From graphics generation to automatic mapping of 2D images of garments onto images of people, the use cases are just beginning to be explored.

Artificial Creativity

Aside from Amazon's work in the area, other research has been exploring the possibility of generating fashion-based images as well. In their research paper, "DesIGN: Design Inspiration from Generative Networks," Othman Sbai et al. describe a method for generating garments by combining garment silhouettes as a mask and then transferring pattern and texture to that garment mask using a GAN. The goal of this research was not to propose the automation of fashion design, but to generate an inspirational machine assistant. Some of the most successful results from this research are shown in Figure 8-1.

Figure 8-1. *Images of garments generated by Sbai et al. from Facebook AI Research, Ecole des Ponts, and NYU*

The paper touches on some much larger philosophical topics in the field of AI as well. Creativity is a topic of controversy among those generating machine learning models that have creative motives. For many researchers, the idea that a machine can create original works of art and therefore embody a characteristic that is innately human is an ultimate achievement of artificial intelligence. For other researchers, it seems likely that creativity will never be achieved by a machine, but machines can provide excellent tools to the humans who are creating.

To date, in examples rendered, machines can only mimic creative works that they are fed. These include machine learning work that outputs images imitating great painters, the writing of pop songs, and doodling sketches. The reality of this work, as incredible as it is, is that it fails to meet our expectations when it comes labeling it as "art."

Mapping Garments onto Images of People

Think about all the different types of images that are created to sell a single garment style. Technical drawings, line plans, garment photos, garment-on-model photos, lifestyle photos, and blogger photos are all part of the strategy to get the idea of a garment into the hands of a consumer. Generative models provide one possible solution to reducing the cost of image generation in the fashion industry.

In "The Conditional Analogy GAN: Swapping Fashion Articles on People Images," Zalando researchers Nikolay Jetchev and Urs Bergmann propose a method for transforming 2D garment images onto images of people. An example of this work can be seen in Figure 8-2.

Figure 8-2. *A garment image (far right) being transformed onto a photograph of a person. The original source images are on the left.*

While the resulting images are still low fidelity and low resolution, this research shows promise for automating the process of photographing garment-on-model images. Photographing a season's garments can cost brands thousands of dollars per photoshoot. GANs have also been used for other types of photo editing, such as retouching.

Turning Sketches into Color Images

Another area of research using GANs is **image-to-image translations**. This includes converting a simple black-and-white sketch into a color image, a process usually described as edges to photo, as seen in Figure 8-3.

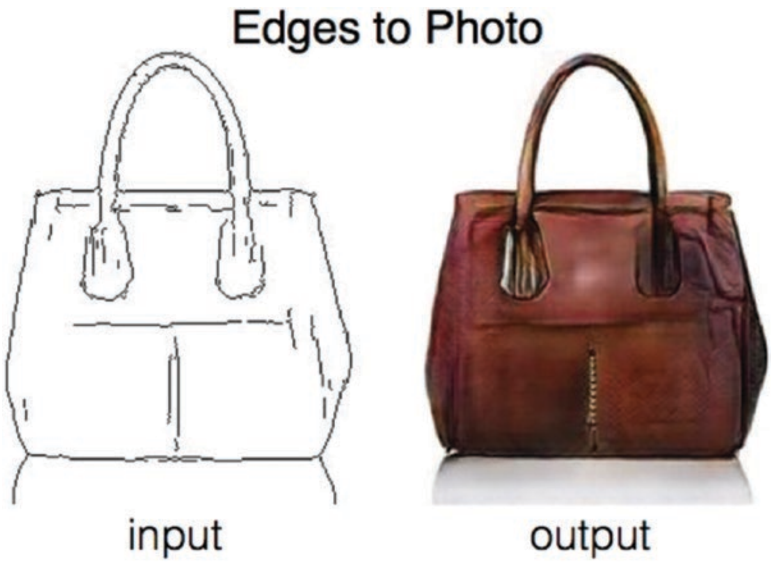

Figure 8-3. *An image of a black-and-white sketch on the left and an output photo-like image on the right generated by* ***pix2pix***

This kind of image-to-image translation is done using a conditional GAN. Less of the setup needs to be hand-engineered than was previously, making the model more accessible to use by a wider audience.

How Generative Models Work

On the most basic level, a generative model refers to a method of computer-driven generation of images, video, and music, for example. This is different from other types of machine learning because the output is a sort of recombinant variation of the training data.

The high-level concept is shown in Figure 8-4. A generative model takes in a series of input images, and outputs images that are similar but completely machine generated.

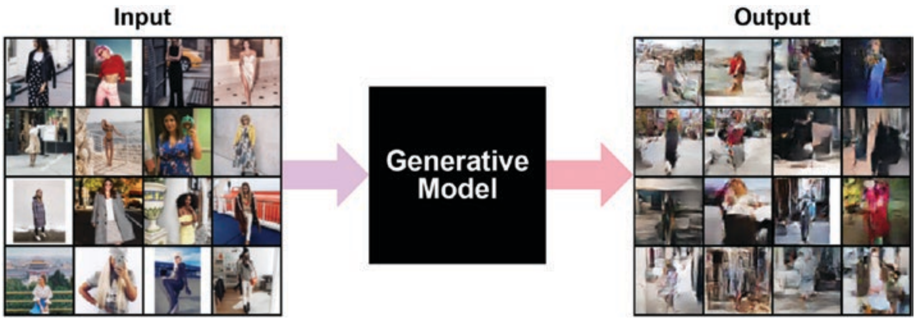

Figure 8-4. *A high-level view of what a generative model does. It takes in an input dataset and generates images that are "fake" but look similar to the original data.*

Generative adversarial networks (GANs), the most popular generative model and the method used by Amazon, are a subset of generative models. Put simply, a GAN is a series of two unsupervised dueling neural networks. In a GAN, one of the neural networks generates images based on patterns it recognizes in the input dataset, and the other neural network classifies those images as real or fake. The information is passed back through the generative network in order for the model to improve.

There are other types of generative models, including **variational autoencoders** (**VAEs**) and **autoregressive models**. A VAE relies on probabilistic modeling to generate outputs, but often the outputs of a VAE tend to be replications of the original dataset rather than creating something unique. Autoregressive models have remained probably the lesser explored of these generative models but are starting to become more popular.

Limitations

> *In a machine age, dressmaking is one of the last refuges of the human, the personal, the inimitable.*
>
> —Christian Dior, fashion designer

There is little reason to believe that a generative model will be able to write a tech pack anytime in the near future, for instance. It's also important to keep in mind that this is currently a topic being explored in the research community, and though the field is moving quickly, generative models are not quite ready for commercialization.

The AI fashion designer does touch upon some larger issues that many industries are calling into question. What is the role of human labor in this new era, as machines are taking over so many of the tasks that we are currently responsible for? What are the jobs we will hold as this becomes more true? Will the fashion designer be a profession of the past? What about other jobs in the fashion industry? This subject will be explored and speculated on in greater detail in Chapter 12.

Why GANs?

GANs show promise to AI researchers in part because of their ability to accomplish complex tasks using unsupervised techniques. GANs are also compelling because they provide potential solutions for complex automation problems in areas such as these:

- Photography
  - *Image in-painting and retouching*: Completing images with missing patches or even potentially retouching
  - *Increasing image resolution*: Converting images from low resolution to high resolution by filling in information

- Design
    - *Aggregating trends*: Generated by a target audience into a visual summary
    - *Style transfer*: Applying the style of a particular aesthetic quality
    - *Medium translation*: Taking images from a sketch-based medium to a photograph, and vice versa
- Storyboarding
    - *Text-to-image generation*: Eventually, trained GANs could be good enough to deliver images from a text-based input.

Admittedly, some of these applications are speculative, and this list is not comprehensive. The capabilities of technology like GANs and all of AI start and stop at what humans invest in building. Progress will require a motivated individual or team to pursue and accomplish many of these tasks. Without those people, it could remain dormant.

Implementation Example: AI Fashion Blogger

GANs are, at least currently, best suited for use cases that require images as the output. In the case of fashion design, this is challenging, because a garment is the final output rather than an image.

There are other avenues to apply GANs to the fashion industry; for example, the fashion blogger. The images in Figure 8-5 are output from a GAN I trained on fashion blogger data, images I scraped from blogger Instagram accounts.

Figure 8-5. *Examples of 64px images output from DCGAN trained on fashion blogger data*

Every image in the fashion industry has more to it than meets the eye. Fashion bloggers make up an industry of their own. Fashion blogger images could not exist without a complex network created by the people who made them.

On the flip side, fashion blogging wasn't an industry before the proliferation of two technologies: the smart mobile device and social media. It will be interesting to see the resulting outcomes from the introduction of AI to this ecosystem as a creator and how it will reshape the economy behind blogging.

How It Works

The easiest way to understand how a GAN works is through a description of a simple example. Figure 8-6 shows a dataset of 200 images. These images might be images from fashion bloggers' Instagram accounts. That data is used to train the first neural network, a generative neural network, which creates similar-looking images from scratch. The output dataset is a collection of 200 images that were created by the generative neural network (G). These are not real images of bloggers, but were created by the neural network to look like them. Once that dataset has been created, the second neural network takes the data as an input and returns the probability that the image is from the initial dataset. In other words, the second model classifies the image as real or fake. The goal of the GAN is to generate images that are convincing enough that the second neural network believes the generated images are examples from the real world rather than images that were created by the first neural network.

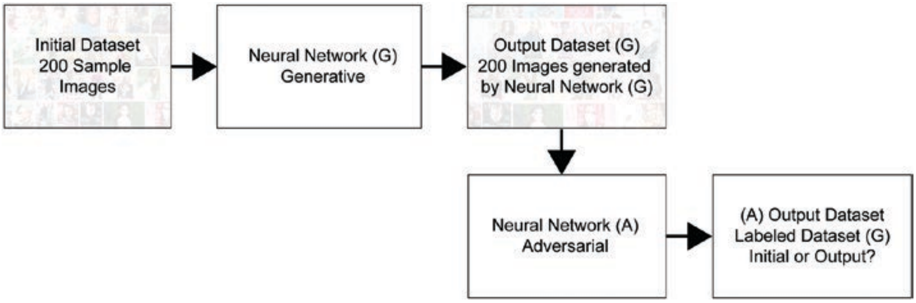

Figure 8-6. *A diagram of the example GAN*

Training GANs

Because GANs are made up of two neural networks, training a GAN is similar to training any other neural network. Refer to Chapter 4 for more information about how neural networks are trained.

Training GANs is still not a widely understood process. This GAN fashion blogger example was trained using a **PyTorch** implementation of a GAN called **deep convolution generative adversarial networks** (**DCGANs**). The first thing to do to train a GAN is to find or collect a large, high-quality dataset. Figure 8-7 shows the learning process using this method.

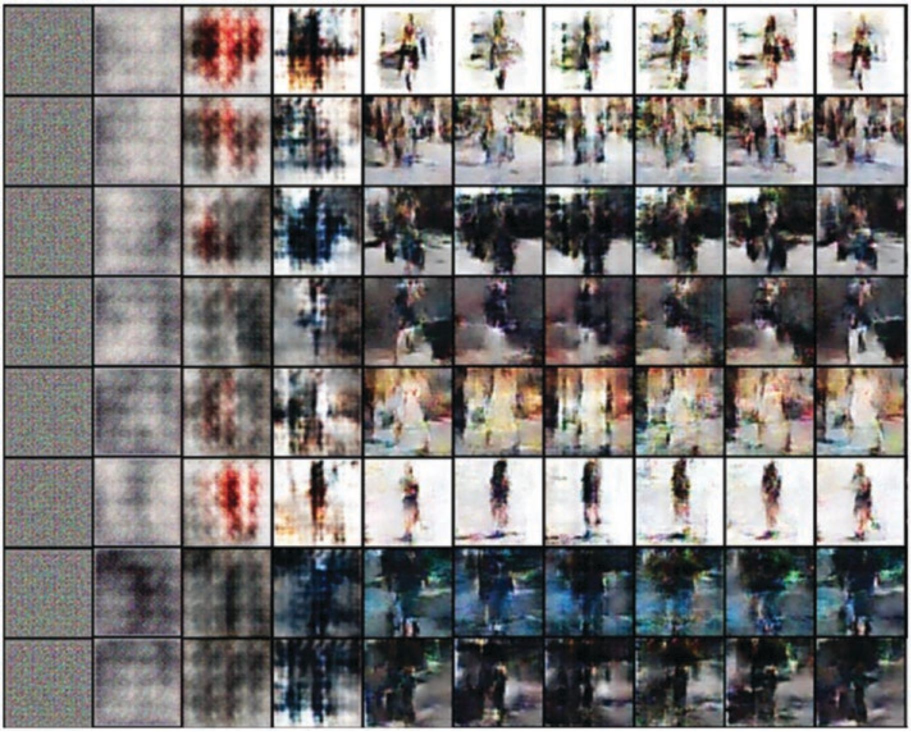

Figure 8-7. *A series of images that illustrate the learning process, from left to right. As the GAN trains, in every epoch it becomes better and better at generating realistic fashion blogger images.*

Each time the GAN iterates through an epoch, it learns how to improve the images, leading the discriminator network to believe that the images are real. An **epoch** is an iteration through the entire dataset. This terminology may be seen as less necessary or even potentially unnecessary in examples with large datasets, but in the case of this DCGAN, crucial to getting quality results.

Improving Results

The most straightforward variable to control while training these networks is the size and quality of the dataset. Expanding the dataset will improve the quality of images output by the network for a variety of reasons.

The more images there are, the less likely you are to run into issues of **overfitting**. Overfitting, in this case, occurs a when the data outputs too closely fit the inputs; sometimes this is caused by having too few images in the dataset. Hence, the network reproduces the same imagery in all of its output images. Figure 8-8 shows an example of overfitting; the same face appears over and over again in the generated images.

Figure 8-8. *An example of overfitting, expressed through the repetition of the same face over and over again*

For context, a researcher would consider CIFAR-10, a dataset of images of 80 million images, to be a large dataset. In the case of the fashion blogger, I was using a dataset of around 3,000 images to start.

The Future of GANs

This chapter just skims the surface of what can be done using generative models. While many of the image examples appear fairly low-resolution, some networks already can output images at higher resolutions and with more lifelike results. A couple of examples of networks capable of this are **StackGANs** and **progressive growing GANs** (**PGGANs**); sample PGGAN outputs are shown in Figure 8-9.

Figure 8-9. *Output images generated by PGGAN, released by Nvidia in late 2017. The model was trained on celebrity headshots.*

Today, these images are expensive to produce in terms of compute power and researcher hours, but generative models have quickly become a topic of interest in the machine learning community. Their use is growing and expanding to solve problems in new industries.

Summary

While today the promise of an AI fashion designer might be more hype than production ready, generative models have potential for a wide variety of applications in the fashion industry. From graphics generation and tools for creativity by Sbai et al. to automatic mapping of 2D images of garments onto images of people by Jetchev and Bergmann, the use cases in fashion are just beginning to be explored.

Generative models are still a topic of research but have gained traction quickly. For a deeper dive on generative models, I recommend the following resources. You can find even more in the annotated bibliography at the end of the book.

- "Generative Models" by Andre J. Karpathy et al., (OpenAI Blog, Nov. 28, 2017): `blog.openai.com/generative-models/`
- *Deep Learning* by Ian Goodfellow et al. (MIT Press, 2016): `www.deeplearningbook.org`
- "Photo Editing with Generative Adversarial Networks (Part 1)" by Greg Heinrich (Nvidia Developer Blog, April 20, 2017): `https://devblogs.nvidia.com/photo-editing-generative-adversarial-networks-1/`

Terminology from This Chapter

Autoregressive models—Train a neural network based on the conditional distribution of each pixel, given the previous pixel. This method is similar to methods used in NLP, but iterate over pixels instead of characters.

Data mining—A process of analyzing a large dataset in order to uncover patterns that might be prevalent. More on this topic in Chapter 9.

Deep convolution generative adversarial networks (**DCGANs**)—A type of GAN first described by Alec Radford et al. in 2016. Since then, many implementations of this neural network system have surfaced from the machine learning community.

Epoch—An iteration through a dataset. Iterating through a dataset more times, especially for small datasets, can help increase the accuracy of the outputs in GANs.

Generative adversarial network (**GAN**)—A type of generative model in which two neural networks are dueling, resulting in a higher-accuracy output in an unsupervised learning process.

Generative model—Often refers to a model that generates images, though generative models can also be used to create other types of data.

Image-to-image translation—Similar to a language translation in that it refers to a process of taking one style of images and converting it to another style.

Overfitting—Occurs when the data outputs too closely fit the inputs. In the abstract, it means that the underlaying algorithm is tailored to fit a specific set of characteristics and has less-broad application.

Pix2pix—Shorthand for a network architecture release by Isola et al. that allows users to do general image-to-image translations using a conditional GAN architecture.

Progressive growing GANs (**PGGANs**)—A GAN architecture concept that came out of Nvidia in late 2017. The outputs are 1024 × 1024px images that are indistinguishable from real images; their example was trained on celebrity headshots.

PyTorch—A Python (general-purpose programming language) framework for machine learning. In general, frameworks help software engineers to move faster in developing and refining machine learning models.

StackGAN—A method of using a system of multiple stacked GANs in order to generate higher-quality outputs, like higher-resolution images.

Variational autoencoders (**VAEs**)—A type of generative model made up of an encoder, a decoder, and a loss function. They work similarly to GANs in that they are also made up of a set of neural networks, measuring loss over the network generation and improving the model continually.