

Generating High-Resolution Fashion Model Images Wearing Custom Outfits

ICCVW 2019

Gökhan Yildirim, Nikolay Jetchev, Roland Vollgraf, Urs Bergmann
Zalando Research

발표자: 최승환

DAVIAN Vision Paper Study

2020/01/28

Index

1. Introduction
 - Contributions
2. Background
 - StyleGAN
3. Dataset
4. Method
 - Unconditional Model
 - Conditional Model
5. Experiments
6. Appendix

Introduction

- Fashion e-commerce platforms simplify apparel shopping through search and personalization
- Visualizing an outfit on a human body can further enhance user experience
- Previous studies
 - Replacing a garment on an already existing image of a fashion model: **Virtual Try-ON**



- Generating low-resolution images from scratch by using pose and garment color

Introduction Contributions

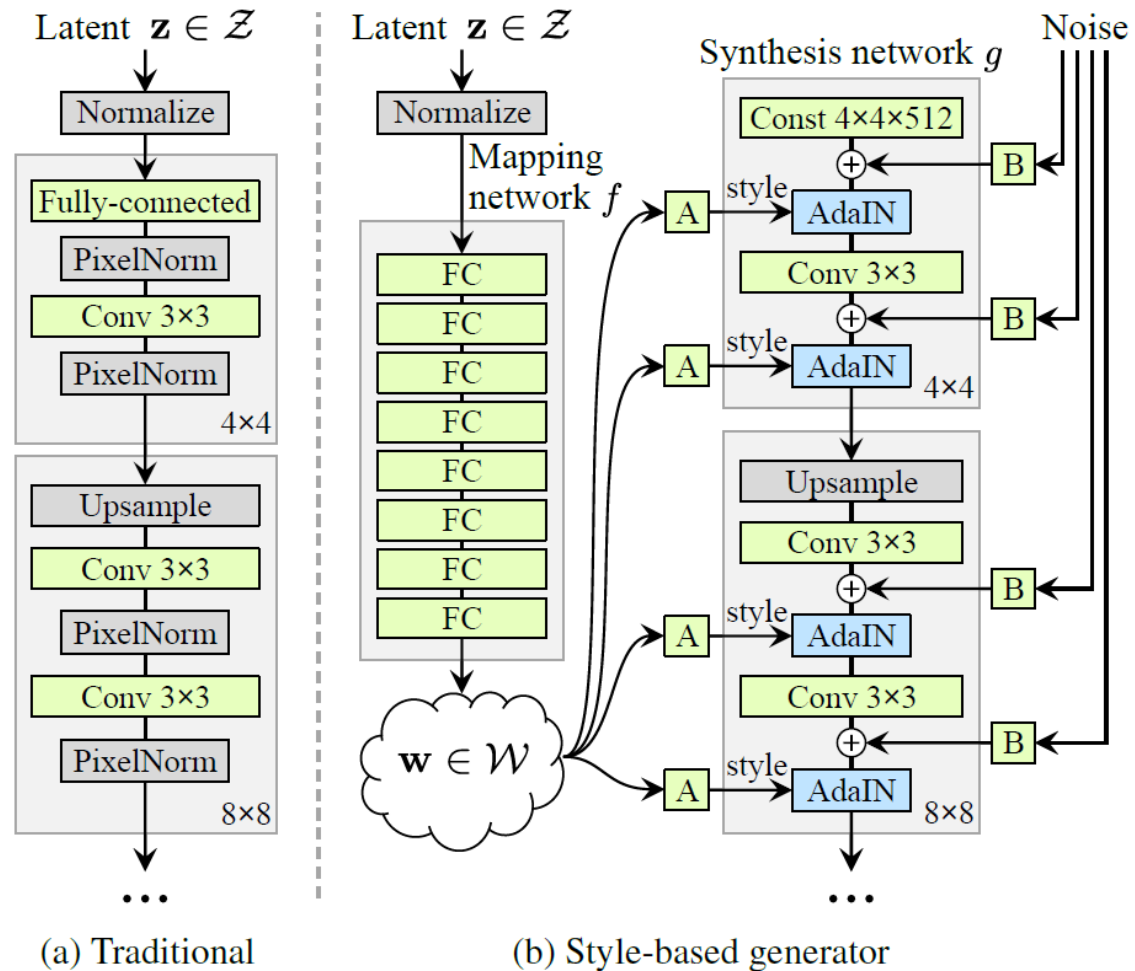
- The authors train the vanilla StyleGAN **on a set of fashion model images**
 - The network can transfer the outfit color and body pose of one model to another
- The authors **modify StyleGAN to condition the generation process** on an outfit and a human pose
 - The network can rapidly visualize custom outfits under different body poses and types

Background StyleGAN

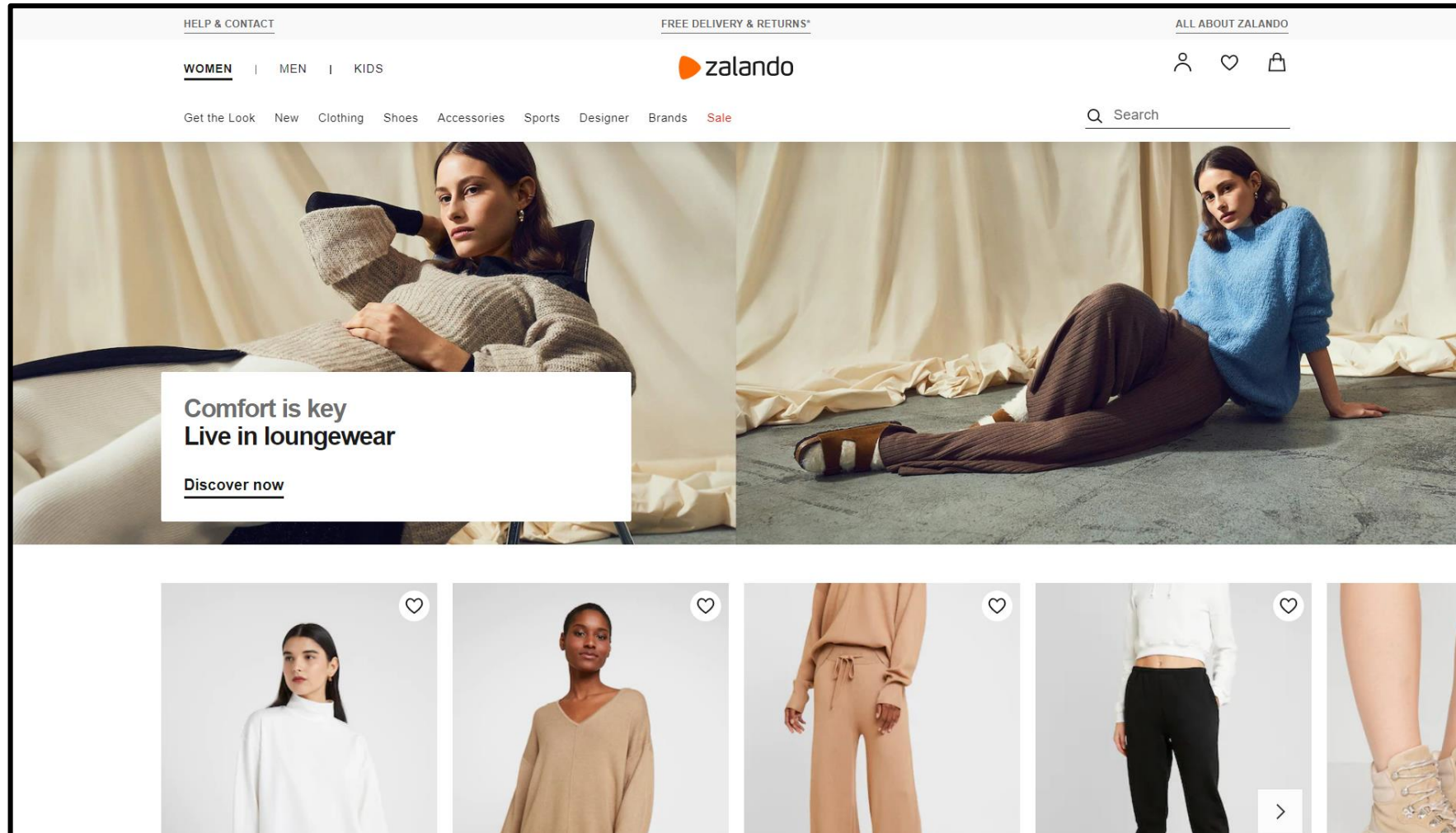
- A re-designed generator architecture, borrowing from style transfer literature
 - C. Adding AdaIN operations
 - F. Generating images using two random latent codes during training

Method	CelebA-HQ	FFHQ
A Baseline Progressive GAN [30]	7.79	8.04
B + Tuning (incl. bilinear up/down)	6.11	5.25
C + Add mapping and styles	5.34	4.85
D + Remove traditional input	5.07	4.88
E + Add noise inputs	5.06	4.42
F + Mixing regularization	5.17	4.40

Table 1. Fréchet inception distance (FID) for various generator designs (lower is better). In this paper we calculate the FIDs using 50,000 images drawn randomly from the training set, and report the lowest distance encountered over the course of training.



Dataset



Zalando Website

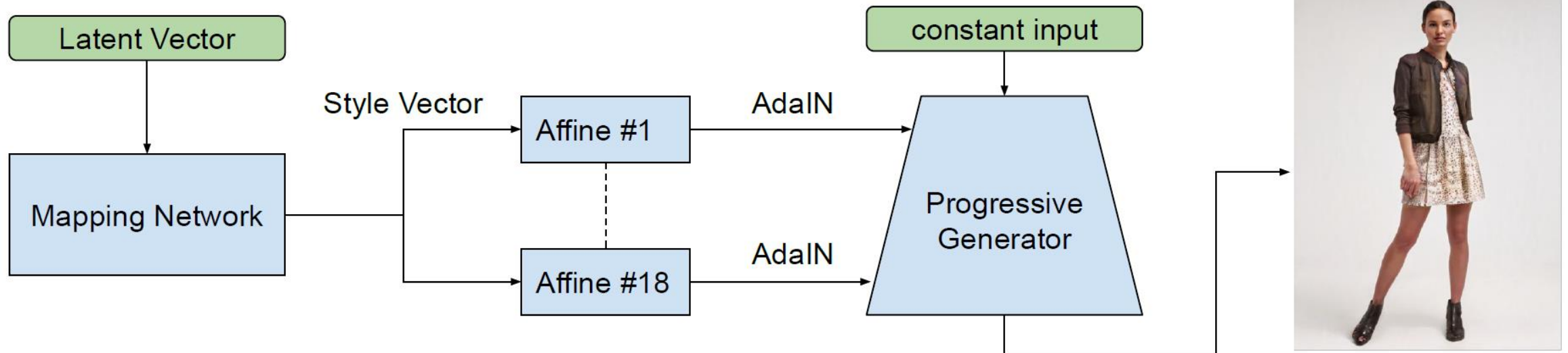
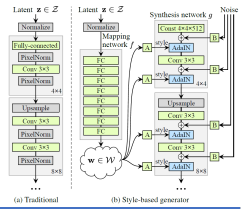
Dataset



- A proprietary dataset of model-outfit-pose images
 - Outfit: a set of maximum 6 articles
 - Pose: 16 extracted keypoints using a deep pose estimator
 - A resolution of 1024 x 768 pixels

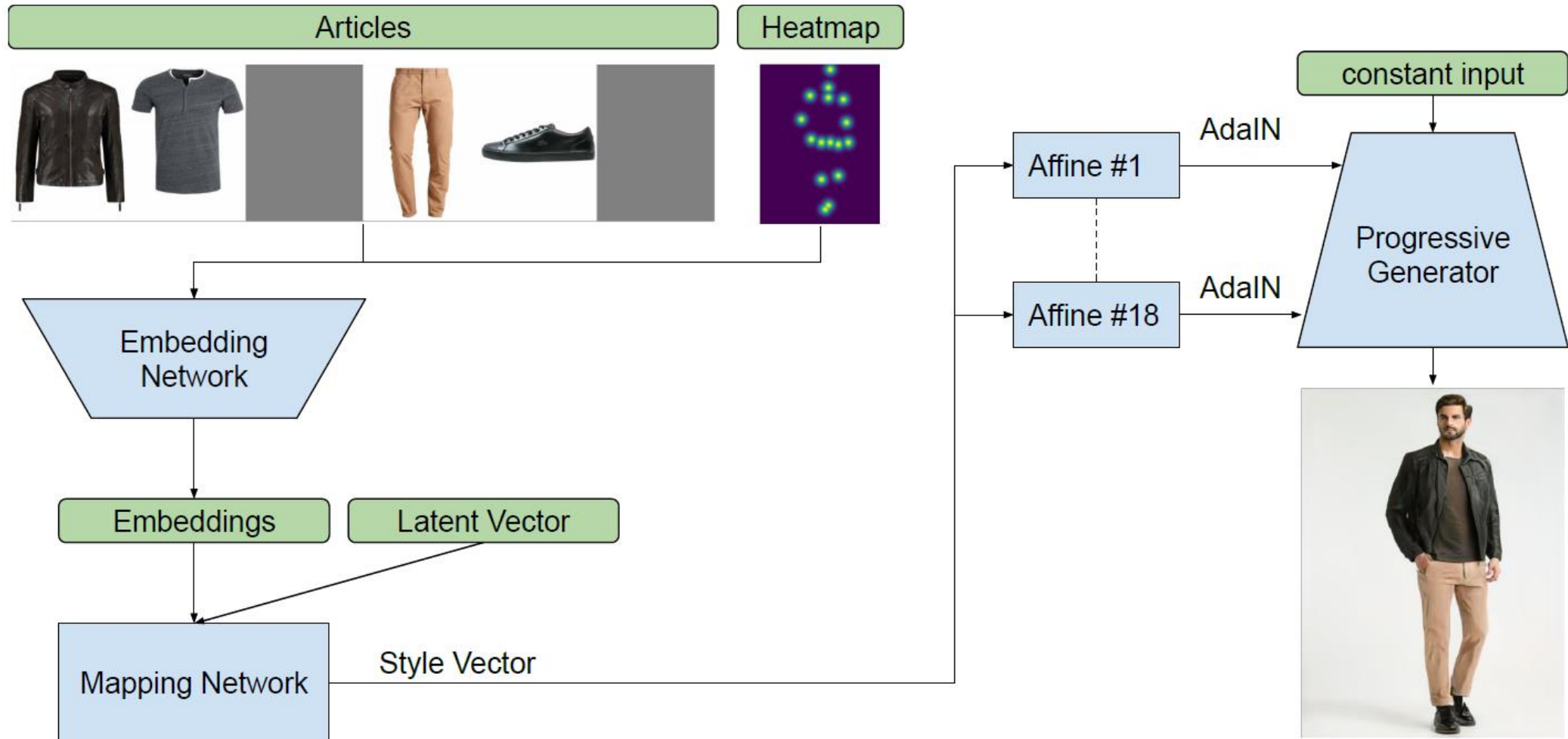
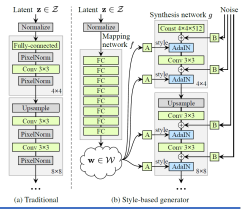
Method

Unconditional Model



Method

Conditional Model



Experiments

- Unconditional model
 - Not only articles, but also human body parts are realistically generated



Model images that are generated by the unconditional StyleGAN

Experiments

	Color Transfer	Pose Transfer
Source	13-18	1-3
Target	1-12	4-18

Table 1: Layers to broadcast the style vector.

- Unconditional model
 - Not only articles, but also human body parts are realistically generated



Transferring the colors of an outfit or a body pose to a different generated model

Experiments

outfit #1



- Conditional model
 - Articles are correctly rendered on the generated bodies



Generated model images with outfit #1

Experiments

outfit #2



- Conditional model
 - Articles are correctly rendered on the generated bodies



Generated model images with outfit #2

Experiments



- Conditional model
 - Articles are correctly rendered on the generated bodies



Generated model images with outfit #2 and the jacket from outfit #1


Experiments


	FID Score	Training Epochs
Unconditional	5.15	115
Conditional	9.63	115

Table 2: FID Score for the models.

Appendix

- <https://research.zalando.com/welcome/mission/research-projects/>


HOME RESEARCH ▾ TEAM JOBS CONTACT



Generative Fashion Design

Modern generative machine learning models allow the sampling of data from a very high-dimensional distribution. A prime example domain is images. Can these methods be used to assist fashion designers and accelerate fashion design? In this project we're investigating answers to this question.


[View More](#)



Swapping Of Fashion On People Images With Generative Modelling

Our work focuses on using state-of-the-art-image modelling techniques in order to flexibly modify the appearance of human fashion images. Our fashion applications of generative modeling require the generation of human body poses conditioned on specific inputs -- e.g. the swapping of clothes that people are wearing, or changing the body pose of a human image while keeping a plausible appearance.

[View More](#)



Feidegger Dataset

The FEIDEgger (fashion images and descriptions in German) dataset is a new multi-modal corpus that focuses specifically on the domain of fashion items and their visual descriptions in German. The dataset was created as part of ongoing research at Zalando into text-image multi-modality in the area of fashion.

[View More](#)