GPU TECHNOLOGY CONFERENCE

NVIDIA

Fashion Design with GANs: Disentangling Color, Texture, and Shape

Gökhan Yildirim
Calvin Seward
Urs Bergmann

11th of October 2018

Session: E8289

zalando research

Fashion E-Commerce @ zalando

~ 4.5 billion EUR
revenue 2017

> 200 million
visits per month

> 15,000
employees in Europe

> 70%
of visits via mobile devices

> 23 million
active customers

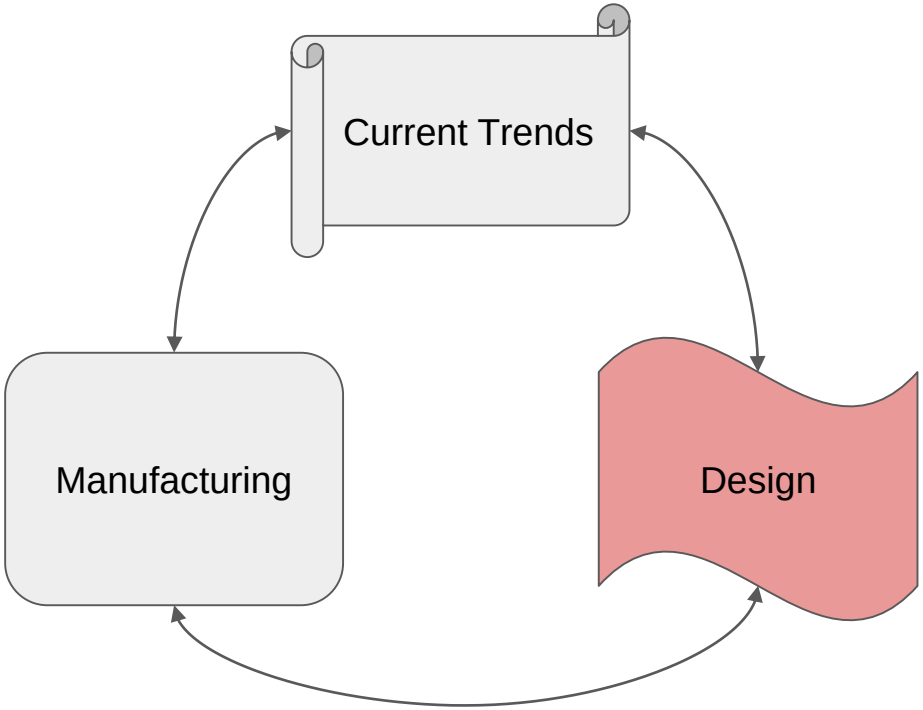
> 300,000
product choices

~ 2,000
brands

15
countries

zalando research

Fashion Design at Zalando



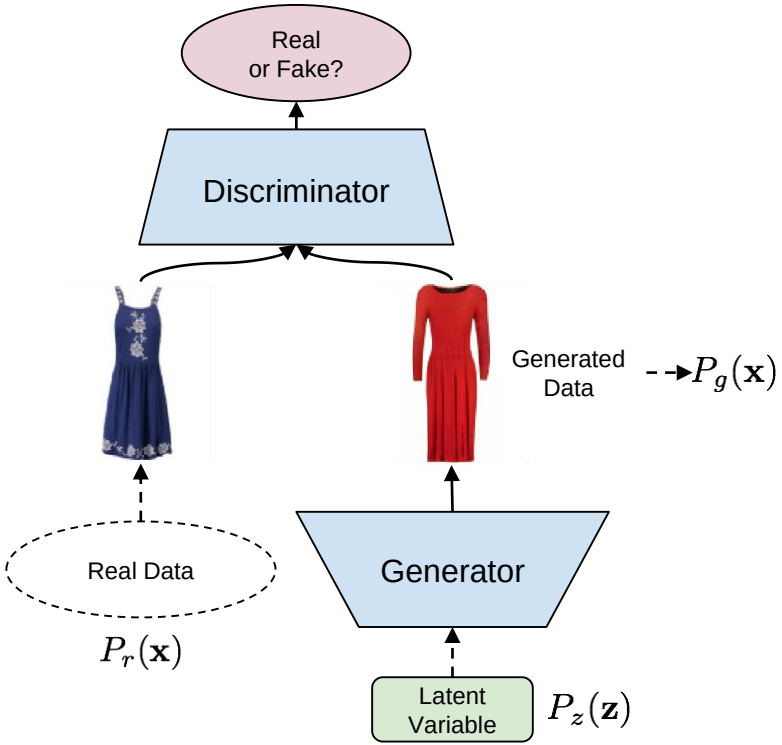
How can computers help speed-up the process?

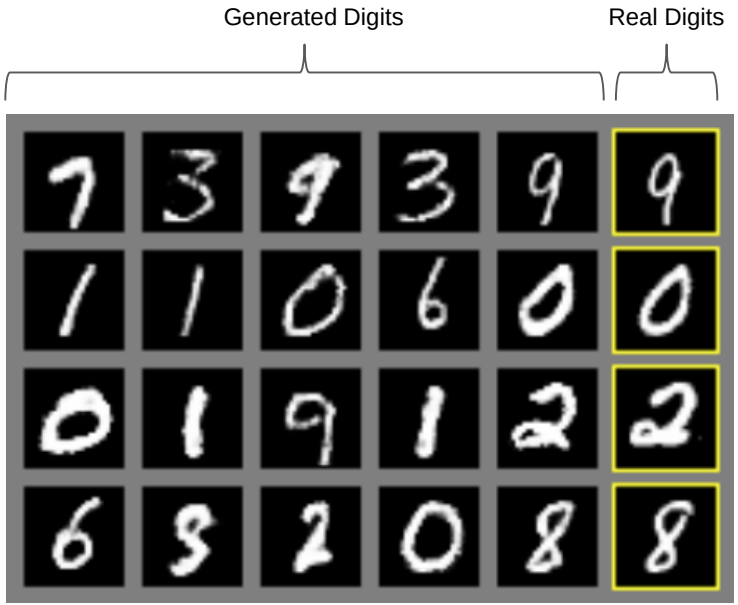
zalando research

Generative Adversarial Networks (GANs)

(Goodfellow et. al., 2014)

- Train two neural networks in an adversarial setting





(LeCun and Cortes, 1998)

zalando research

GANs and Fashion

Swapping Clothes in Photos

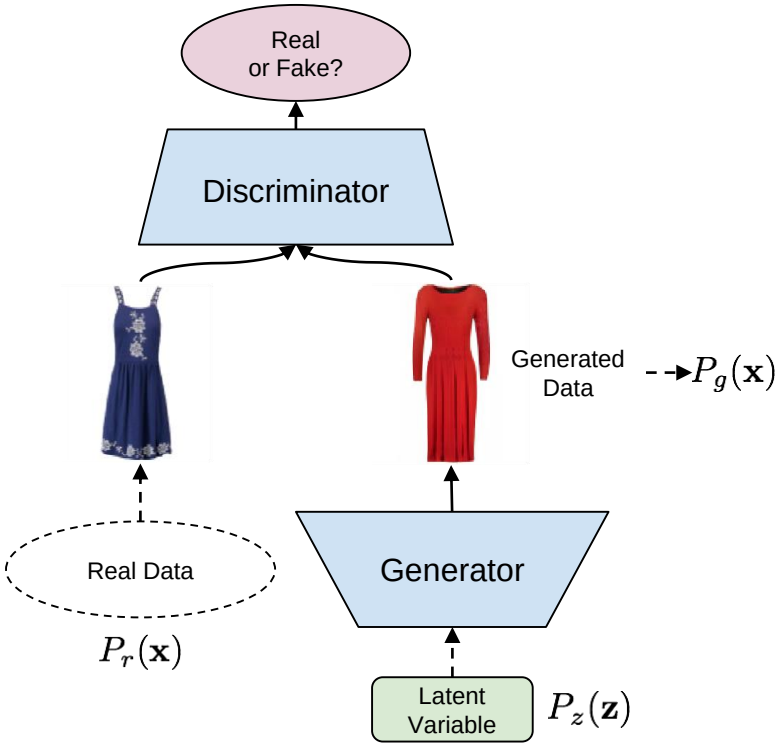
The Conditional Analogy GAN: Swapping Fashion Articles on People Images (Jetchev and Bergmann, 2017)

Clothing Texture Design

Fashion Style Generator (Jiang and Fu, 2017)

zalando research

Attribute Control and Disentangling



- Attribute control
  - Color
  - Texture
  - Shape
- Disentangle
  - Each attribute has a separate effect
  - Changing one should not change the other

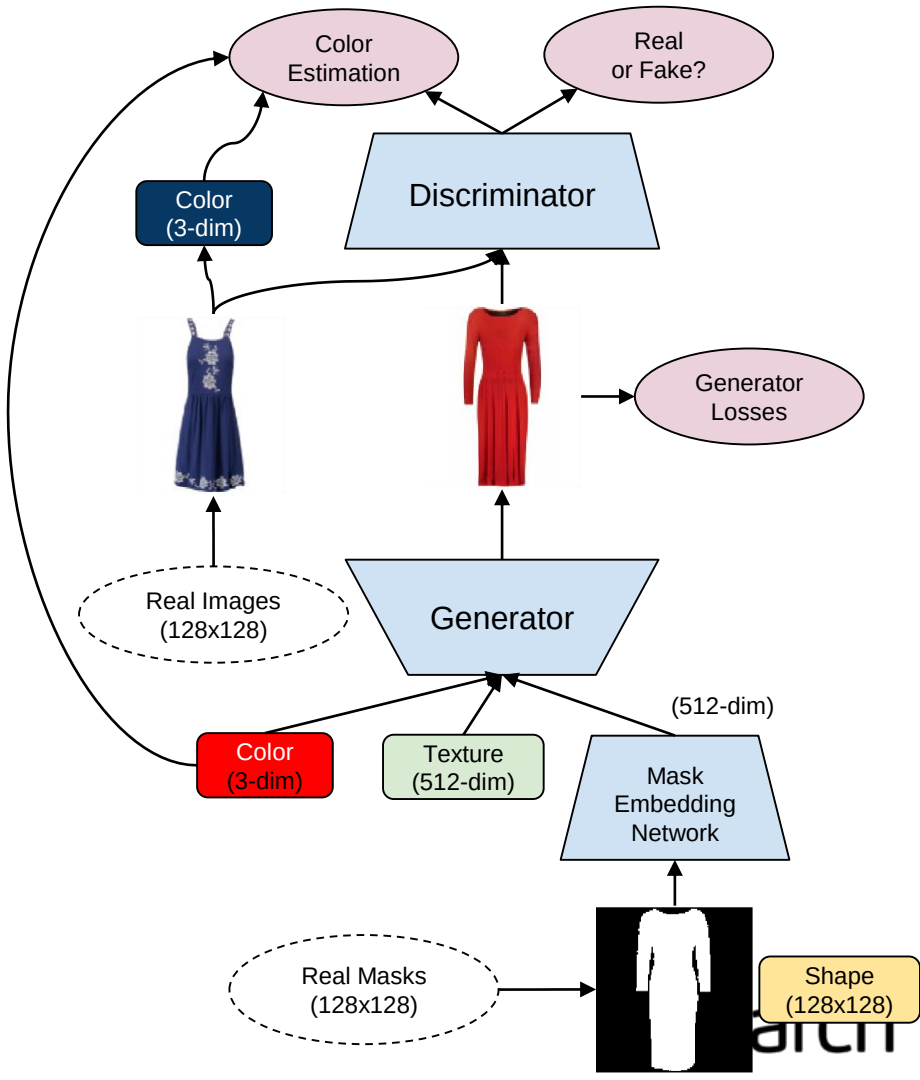
zalando research

Our Model

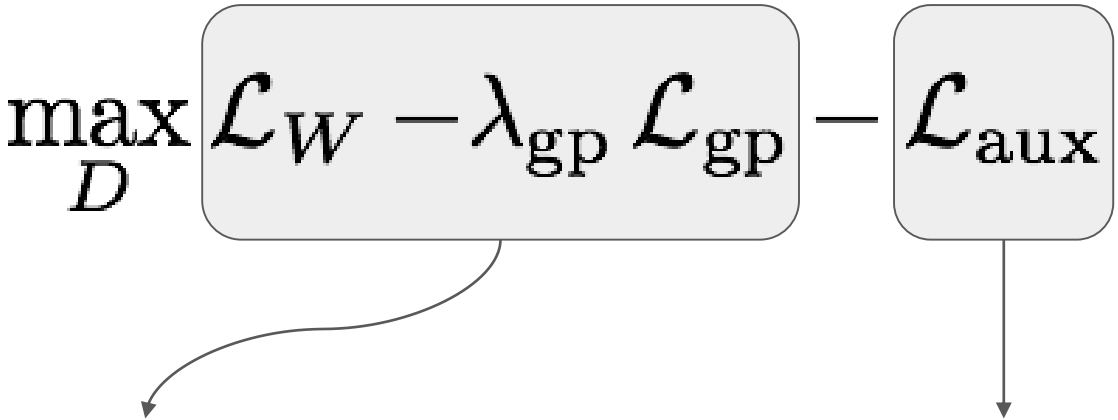
- Attribute Data
  - $\mathbf{c}$ Color
    - 3-dim
    - RGB
  - $\mathbf{t}$ Texture (local structure, pattern)
    - 512-dim
  - $\mathbf{s}$ Shape (segmentation mask)
    - Embedded first
    - 512-dim

$$\mathbf{x} = G(\mathbf{c}, \mathbf{t}, \mathbf{s})$$

Generated Image



Discriminator Losses



“Improved Training of Wasserstein GANs”
(Gulrajani et. al., 2017)

“Conditional Image Synthesis With Auxiliary Classifier GANs
(Odena et. al., 2017)

zalando research

Generator Losses

- Color Consistency
- Texture Consistency
- Shape Consistency
- Generator Color Check

zalando research

Color Consistency

$\mathbf{c}$ color

$\mathbf{t}$ texture

$\mathbf{s}$ shape

$$\mathbf{x}_1 = G(\mathbf{c}, \mathbf{t}_1, \mathbf{s}_1) \qquad \mathbf{x}_2 = G(\mathbf{c}, \mathbf{t}_2, \mathbf{s}_2)$$

Average Dress Color

$$A(\mathbf{x}, \mathbf{s}) = \frac{1}{|\mathbf{s}|} \sum_{i,j} \mathbf{s}(i,j) \cdot \mathbf{x}(i,j)$$

Color Consistency Loss $\longrightarrow \left|\left| A(\mathbf{x}_1, \mathbf{s}_1) - A(\mathbf{x}_2, \mathbf{s}_2) \right|\right|_2^2$

zalando research

Texture Consistency

- Laplacian matting matrix
  - (NxN image → N^2xN^2 matrix)
  - A Closed Form Solution to Natural Image Matting (Levin et. al., 2006)

Input Image (with markings)

Estimated Matte

Texture Consistency

$\mathbf{c}$ color

$\mathbf{t}$ texture

$\mathbf{s}$ shape

$$N \times N \times 3 \longleftarrow \mathbf{x}_1 = G(\mathbf{c}_1, \mathbf{t}, \mathbf{s}_1) \qquad \mathbf{x}_2 = G(\mathbf{c}_2, \mathbf{t}, \mathbf{s}_2)$$

$$N^2 \times 3 \longleftarrow \mathbf{v}_1 \qquad \mathbf{v}_2$$

Texture Consistency Loss

$$\mathrm{Tr}(\mathbf{v}_1^T \mathbf{L}_2 \mathbf{v}_1 + \mathbf{v}_2^T \mathbf{L}_1 \mathbf{v}_2)$$

$$N^2 \times N^2 \longleftarrow \mathbf{L}_1 \longleftarrow \text{Laplacian Matrices} \longrightarrow \mathbf{L}_2$$

zalando research

Shape Consistency

$\mathbf{c}$ color

$\mathbf{t}$ texture

$\mathbf{s}$ shape

$$\mathbf{x} = G(\mathbf{c}, \mathbf{t}, \mathbf{s})$$

$\mathbf{s}$

Shape Consistency Loss $\longrightarrow B(\mathbf{x}, \mathbf{s}) = \frac{1}{|1 - \mathbf{s}|} \sum_{i,j} (1 - \mathbf{s}(i,j)) \cdot ||\mathbf{x}(i,j) - \mathbf{b}||_1$

Background Color $\longrightarrow \mathbf{b} = [1, 1, 1]$

Generator Color Check

$\mathbf{c}$ color

$\mathbf{t}$ texture

$\mathbf{s}$ shape

$$\mathbf{x}^{\mathbf{c}} = G(\mathbf{c}, \mathbf{t}, \mathbf{s})$$

Average Dress Color

$$A(\mathbf{x}^{\mathbf{c}}, \mathbf{s})$$

Generator Color Check → $$||\mathbf{c} - A(\mathbf{x}^{\mathbf{c}}, \mathbf{s})||_2^2$$

zalando research

Experiments & Results

- Modified the code from *
- Directly generate 128x128 pixel images
- NVIDIA P100 GPU → 1 week of training
- Dataset → 120,000 dresses from Zalando
- ADAM Optimizer (LR=0.001, B1=0, B2=0.99) (Kingma and Ba, 2014)

*Progressive Growing of GANs for Improved Quality, Stability, and Variation (Karras et. al., 2018)

Color Control

$$G(\mathbf{c}_i, \mathbf{t}, \mathbf{s})$$

- Changing color
- Texture and shape stay the same

zalando research

Texture Control

$$G(\mathbf{c}, \mathbf{t}_i, \mathbf{s})$$

- Changing texture
- Color and shape stay the same

zalando research

Shape Control

$$G(\mathbf{c}, \mathbf{t}, \mathbf{s}_i)$$

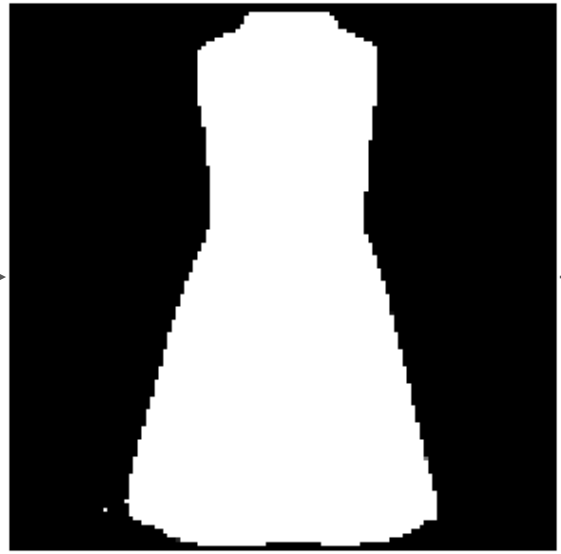
- Changing shape
- Color and texture stay the same

zalando research

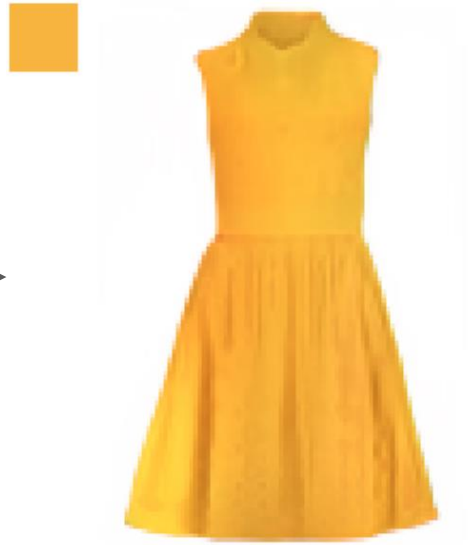
DEMO

Fashion Design

Real Article



Estimated Shape Mask



Reconstructed Article

zalando research

Fashion Design

Failure Cases

Real Article

Estimated Shape Mask

Reconstructed Article

What’s Next?

- Improve upon the design model
  - Allow for multiple colors/color histograms
  - More attributes
  - Direct texture input

GPU TECHNOLOGY CONFERENCE

NVIDIA

Thank You!

zalando research

Preserving Local Structure

Deep Photo Style Transfer
(Luan et. al., 2017)

(a) Input image

(b) Neural Style

(c) CNNMRF

(d) Our result

(e) Reference style image

zalando research

Laplacian Matrix

$$\mathbf{L}(i,j) = \sum_{k|(i,j)\in w_k} \left(\delta_{ij} - \frac{1}{|w_k|} \left(1 + (I_i - \mu_k)(\Sigma_k + \frac{\epsilon}{|w_k|})^{-1}(I_j - \mu_k) \right) \right)$$

$|w_k|$ Size of the neighborhood

μ_k Mean color vector (within neighborhood)

Σ_k Color covariance matrix (within neighborhood)

Controlling the Color

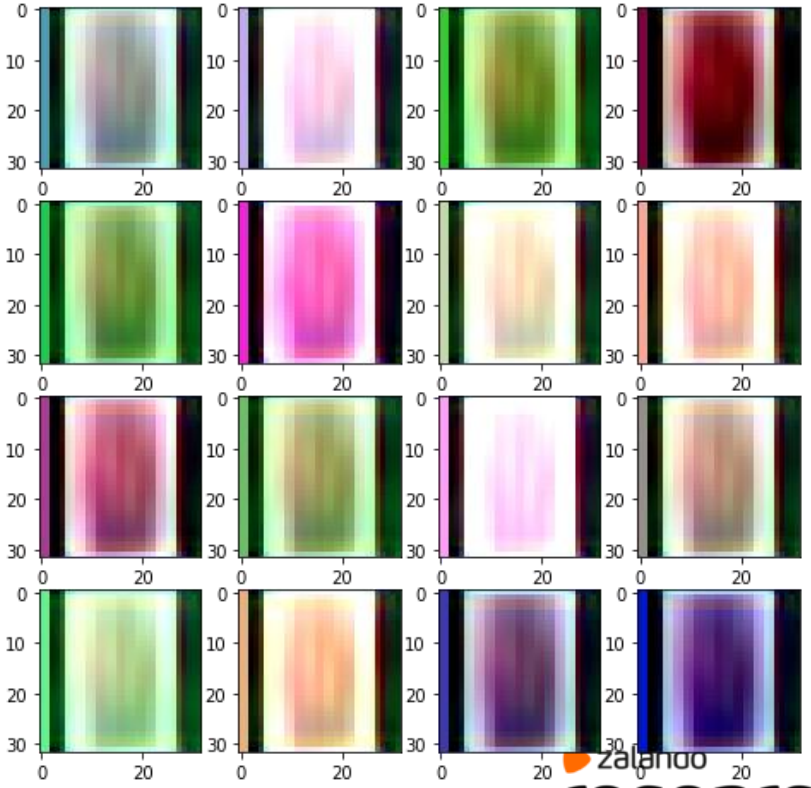
- Start small (32x32)
- Latent Vector + Color input (RGB)
- Pro:
  - We get the desired color
- Con:
  - Color input changes the shape

Same latent vector - Different colors

Adding Shape Consistency

- Shape consistency with Laplacian Matting matrices

- Weighting between color and shape control



zalando research