Staging E-Commerce Products for Online Advertising using Retrieval Assisted Image Generation

Yueh-Ning Ku (Snap),*
Mikhail Kuznetsov (Amazon),*
Shaunak Mishra* (Amazon) – presenter,
Paloma de Juan (Yahoo Research)

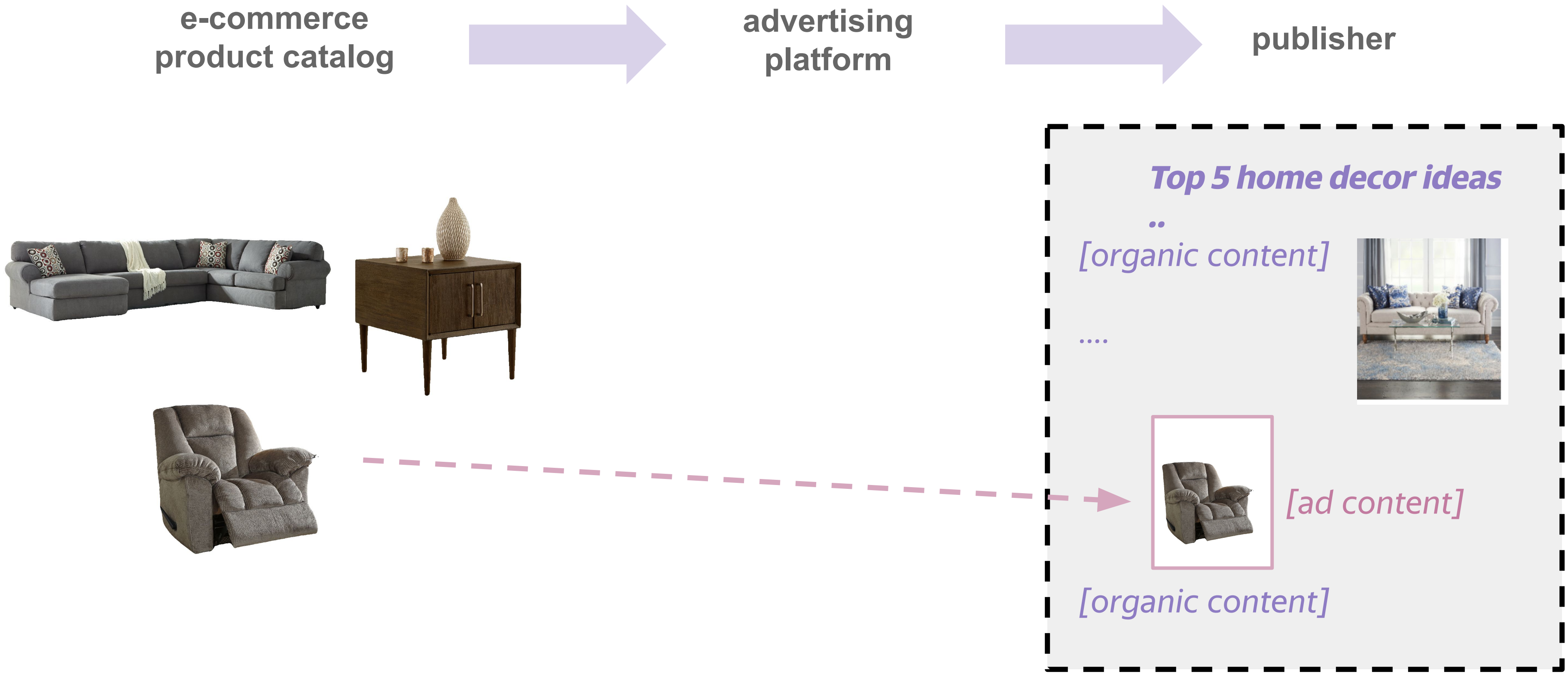
**work done at Yahoo Research*

AdKDD 2023

Advertising e-commerce products

e-commerce product catalog → advertising platform → publisher

Advertising e-commerce products



Advertising e-commerce products: staging

e-commerce
product catalog

advertising
platform

publisher

staged/lifestyle images blend well
with page content,
have higher engagement

Top 5 home decor ideas

..

[organic content]

....

[ad content]

[organic content]

Staging e-commerce products

stage physically?

- ❖ $$$
- ❖ can’t scale

can generative AI help?

Staging products using image generation

staged background generation

task 1

vanilla staging
(unstaged → staged
via background gen.)

Staging products using image generation

staged background generation

task 1

vanilla staging
(unstaged → staged
via background gen.)

task 2

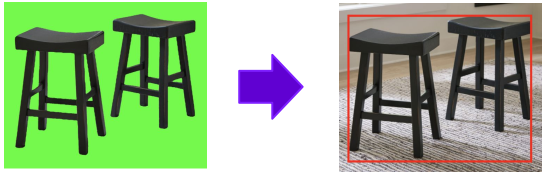
copy-paste staging
(copy staging from other
product images + in-paint)

Staging products using image generation

staged background generation

task 1

vanilla staging
(unstaged → staged
via background gen.)



task 2

copy-paste staging
(copy staging from other
product images + in-paint)

task 3 (bonus 😊)

image → parallax animation

Task 1: vanilla staging (using pix2pix)

Pix2pix [1] is a conditional generative adversarial network (GAN)

[1] Isola et. al, “Image-to-Image Translation with Conditional Adversarial Networks”, CVPR 2017

Task 1: vanilla staging (using pix2pix)

Pix2pix [1] is a conditional generative adversarial network (GAN)

hallucination!

(larger gen. area → higher risk of hallucinations)

[1] Isola et. al, “Image-to-Image Translation with Conditional Adversarial Networks”, CVPR 2017

Staging products using image generation: task 2

staged background generation

task 1

vanilla staging
(unstaged → staged
via background gen.)

task 2

copy-paste staging
(copy staging from other
product images + in-paint)

task 3 (bonus 😊)

**image → parallax
animation**

Copy-paste staging: core idea

product image with staging

using same staging for an unstaged product (red sofa)

in-painting focus on smaller regions easier

Copy-paste staging workflow (retrieval assisted gen.)

pool of staged product images

unstaged product image (input)

retrieve similar product images with staging (inception-V3)

remove products from similar images with staging

in-paint the empty mask in background after product removal

copy-paste input product onto in-painted similar product background

sofa 1

sofa 2

sofa 1

Copy-paste staging workflow (retrieval assisted gen.)

pool of staged product images

inspired from retrieval augmented generation (RAG) in text generation

retrieve similar product in... with staging (inception-V3)

remove products from similar images with staging

in-paint the empty mask in background after product removal

copy-paste input product onto in-painted similar product background

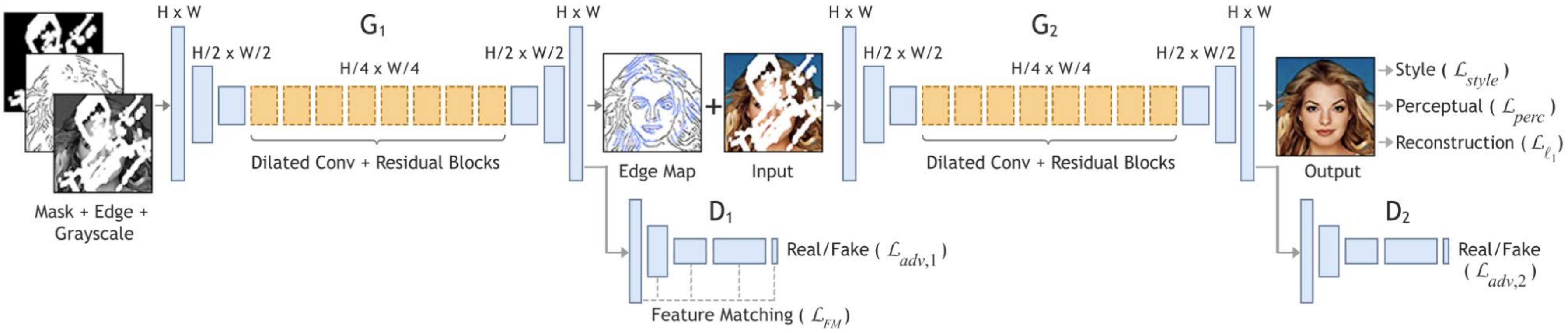
sofa 1

sofa 2

sofa 1

Image inpainting

- ❖ Image inpainting is the task of reconstructing missing regions in an image, e.g. object removal, image restoration, manipulation.
- ❖ We propose to use an adapted EdgeConnect [2] model to fill the gap between the empty mask (from the staged product) and the target (unstaged) product.
  - ➢ EdgeConnect: generated edges and then generates color and texture



[2] Nazeri et. al, "EdgeConnect: Structure Guided Image Inpainting using Edge Prediction", ICCV 2019.

Image inpainting

- Image inpainting is the task of reconstructing missing regions in an image, e.g. object removal, image restoration, manipulation.
- We propose to use an adapted EdgeConnect [2] model to fill the gap between the empty mask (from the staged product) and the target (unstaged) product.
  - EdgeConnect: generated edges and then generates color and texture
  - **Our adaptation: weighted boundary loss to focus on boundaries**

ground truth

conventional free-form mask

higher penalty for boundaries

[2] Nazeri et. al, “EdgeConnect: Structure Guided Image Inpainting using Edge Prediction”, ICCV 2019.

Copy-paste staging demo

Copy-paste staging demo

Results: human evaluation

Human auditors were given three tasks (100 samples per task)

audit task	score
vanilla staging (pix2pix) better than ground truth	0%
copy-paste staging (our approach) better than ground truth	3%
copy-paste staging better than vanilla staging (pix2pix)	76%

Results: FID

Experiments on data from Yahoo (sample of ~ 2000 furniture product images).

Frechet Inception Distance (FID) [3] calculates the (feature distribution) distance between target domain and generated domain; **the smaller the better.**

	baseline FID (EdgeConnect)	our approach FID (EdgeConnect + weighted boundary loss)
copy-paste staging	38.44	**37.44**

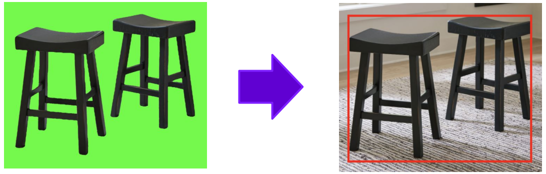
[3] Heusel et. al, “GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium” , NeurIPS 2017.

What about the bonus?

staged background generation

task 1

vanilla staging
(unstaged → staged
via background gen.)



task 2

copy-paste staging
(copy staging from other
product images + in-paint)

task 3 (bonus 😊)

image → parallax animation

Image to parallax animation

Link to video:
https://www.dropbox.com/s/9at5gz24ukhf2gi/product_staging_image_to_parallax_demo.mp4?dl=0

Image to parallax animation

Link to video:
https://www.dropbox.com/s/9at5gz24ukhf2gi/product_staging_image_to_parallax_demo.mp4?dl=0

Conclusion

staged background generation

task 1

vanilla staging
(unstaged → staged
via background gen.)

task 2

copy-paste staging
(copy staging from other
product images + in-paint)

task 3 (bonus 😊)

image → parallax animation

- copy-paste better than vanilla (FID, human eval.); need online test for further validation
- room for improvement in terms of shadows/lighting, hallucinations
- retrieval based ideas can be extended to recent stable diffusion based models