MICCAI2022 Singapore
25th International Conference on Medical Image Computing and Computer Assisted Intervention
September 18-22, 2022
Resorts World Convention Centre Singapore

Brain-Aware Replacements for Supervised Contrastive Learning in Detection of Alzheimer's Disease

Mehmet Saygin Seyfioglu, Zixuan Liu, Pranav Kamath, Sadjyot Gangolli, Sheng Wang, Thomas Grabowski, and Linda Shapiro

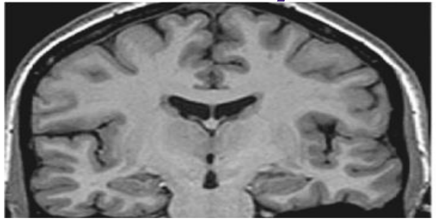
UW Medicine
UW SCHOOL OF MEDICINE

ELECTRICAL & COMPUTER ENGINEERING

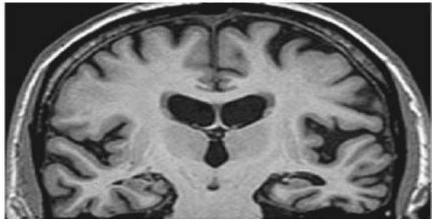
PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING

Motivation

Healthy

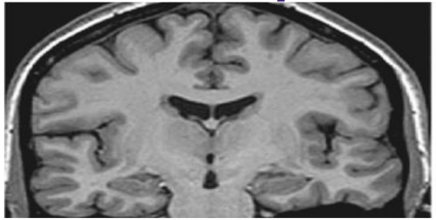


Alzheimer’s Disease (AD)

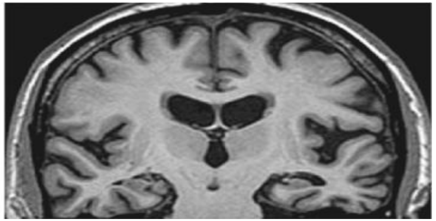


Motivation

Healthy



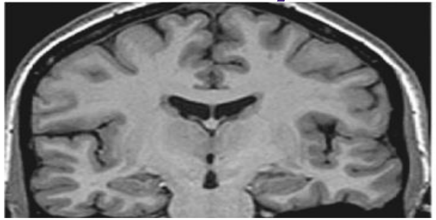
Alzheimer’s Disease (AD)



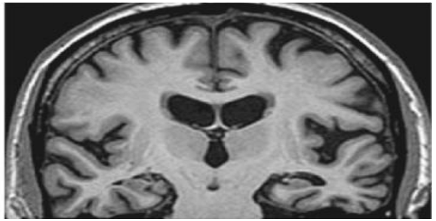
- Low sample support

Motivation

Healthy



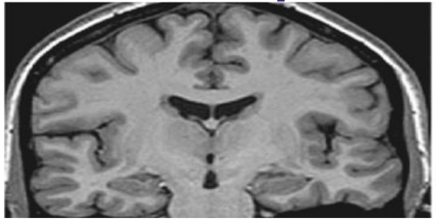
Alzheimer’s Disease (AD)

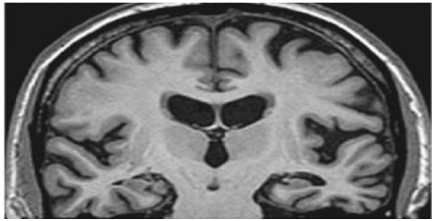


- Low sample support
- MRIs are high dimensional

Motivation

Healthy **Alzheimer's Disease (AD)**

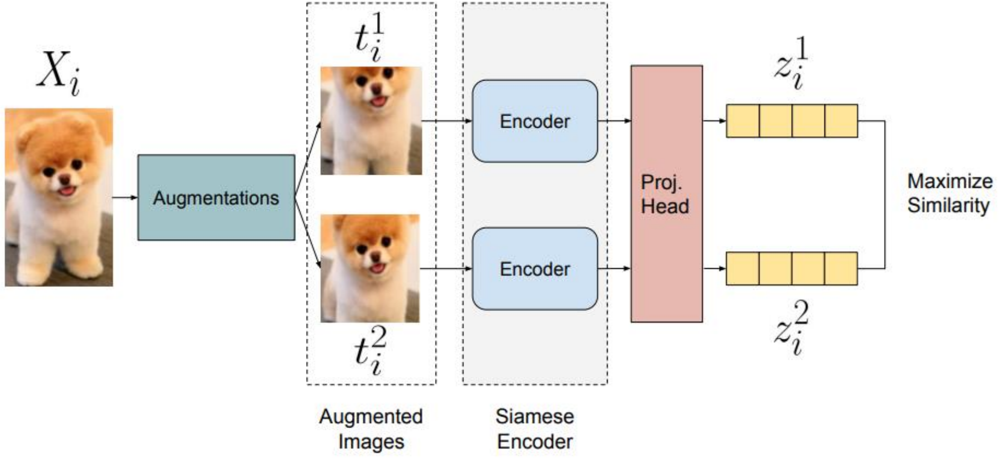


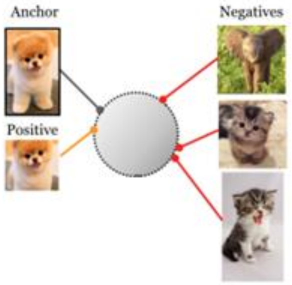


- Low sample support
- MRIs are high dimensional

Contrastive pre-training may help!

Self-supervised Contrastive Learning





$$L_{NCE} = -\sum_{i=1}^{n} log \frac{e^{\theta(t_1^i, t_2^i)}}{\frac{1}{b}\sum_{j=1}^{b} e^{\theta(t_1^i, t_2^j)}}$$

Issues with Self-supervised Contrastive Learning for AD

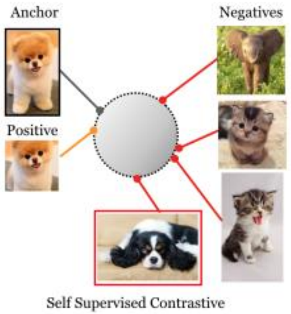
Negative sampling assumption is not suitable for AD detection

Issues with Self-supervised Contrastive Learning for AD

Negative sampling assumption is not suitable for AD detection

> Equidistant assumption risks false negatives

$$L_{NCE} = -\sum_{i=1}^{n} log \frac{e^{\theta(t_1^i, t_2^i)}}{\frac{1}{b}\sum_{j=1}^{b} e^{\theta(t_1^i, t_2^j)}}$$

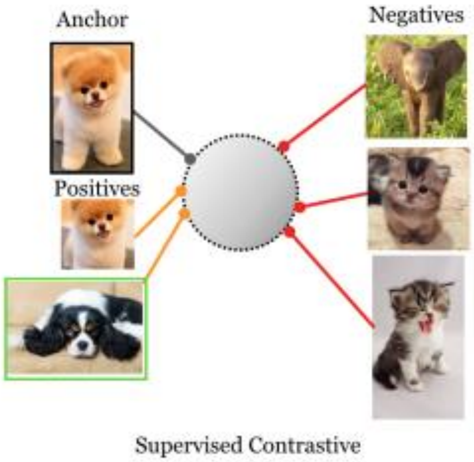


[1]

[1] Khosla, Prannay, et al. "Supervised contrastive learning." Advances in Neural Information Processing Systems 33 (2020): 18661-18673.

Supervised Contrastive Learning

> One way to fix the negative sampling issue is to use supervised-contrastive learning [1] during pre-training.

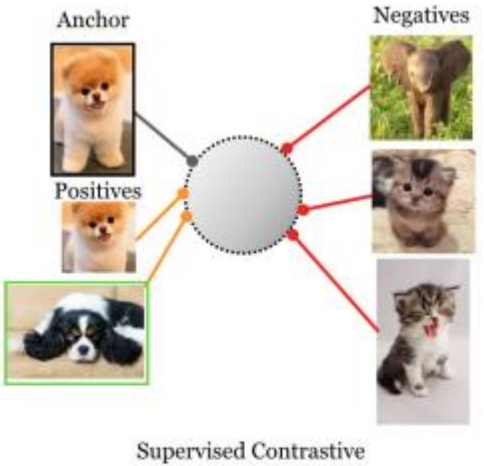


[1]

[1] Khosla, Prannay, et al. "Supervised contrastive learning." Advances in Neural Information Processing Systems 33 (2020): 18661-18673.

Supervised Contrastive Learning

> One way to fix the negative sampling issue is to use supervised-contrastive learning [1] during pre-training.



[1]

> However, this may exhaust the entropic capacity of the labels.

Summing Up

- Self-supervised Contrastive Learning's negative sampling assumption doesn't hold for AD data.
- Supervised Contrastive Learning exhausts all label information.

What to do?

Mixture Prediction with Synthetic Samples

- We can reformulate the contrastive objective as a mixture detection problem!
  - To that end, we need two main components:

Mixture Prediction with Synthetic Samples

- We can reformulate the contrastive objective as a mixture detection problem!
  - To that end, we need two main components:
    1. A way to generate mixtures (synthetic samples)

Mixture Prediction with Synthetic Samples

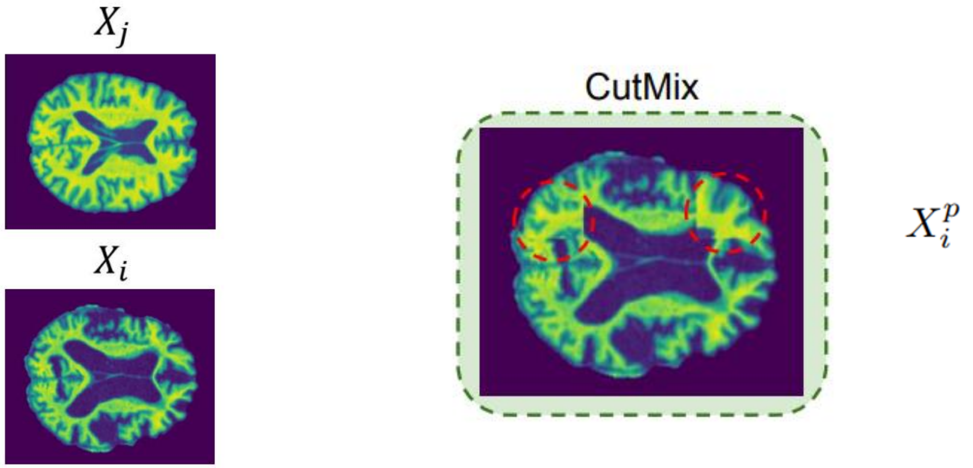
- We can reformulate the contrastive objective as a mixture detection problem!
  - To that end, we need two main components:
    1. A way to generate mixtures (synthetic samples)
    2. A soft-label capable contrastive loss

A Way to Generate Mixtures: CutMix

> CutMix [2] creates mixtures between the images and labels.

$$X_i^p = (1 - M) \odot X_i + M \odot X_j$$

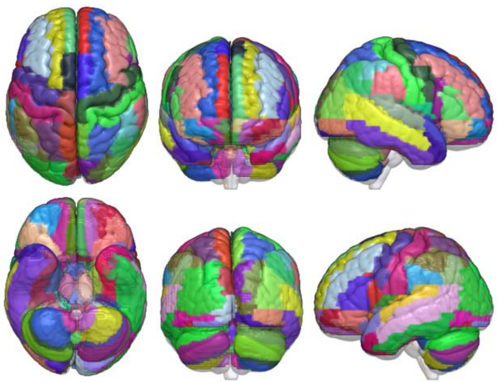
$$y_i^p = \lambda y_i + (1 - \lambda) * y_j$$



[2] Yun, Sangdoo, et al. "Cutmix: Regularization strategy to train strong classifiers with localizable features." *Proceedings of the IEEE/CVF international conference on computer vision*. 2019.

Brain-Aware Replacements by Using a Brain Atlas

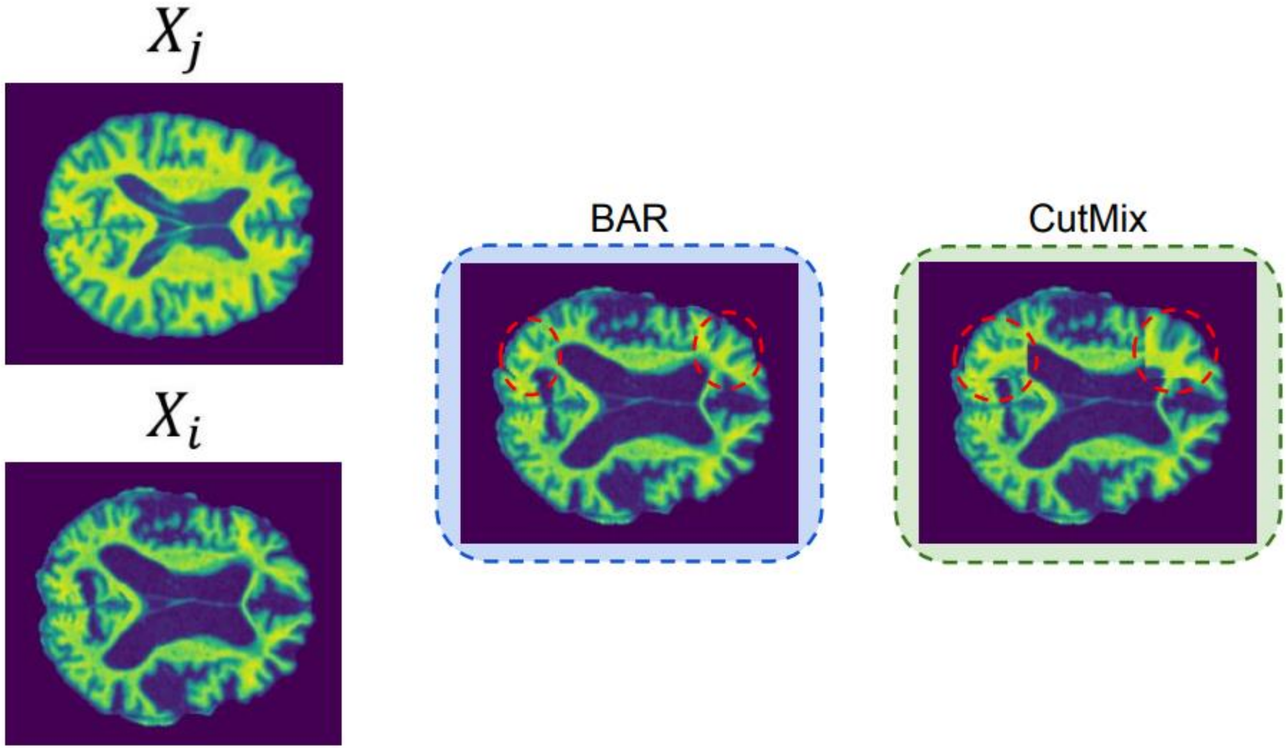
- Brain-Aware Replacements (BAR)
  - Utilizes anatomically relevant regions from the Automated Anatomical Labeling Atlas (AAL)



[3]

[3] Tzourio-Mazoyer, Nathalie, et al. "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain." Neuroimage 15.1 (2002): 273-289.

BAR vs CutMix



Mixture Prediction with Synthetic Samples

> To that end, we need two main components:

1. A way to generate mixtures (synthetic samples) ☑
2. A soft-label capable contrastive loss

2) Soft-Label Capable Contrastive Loss

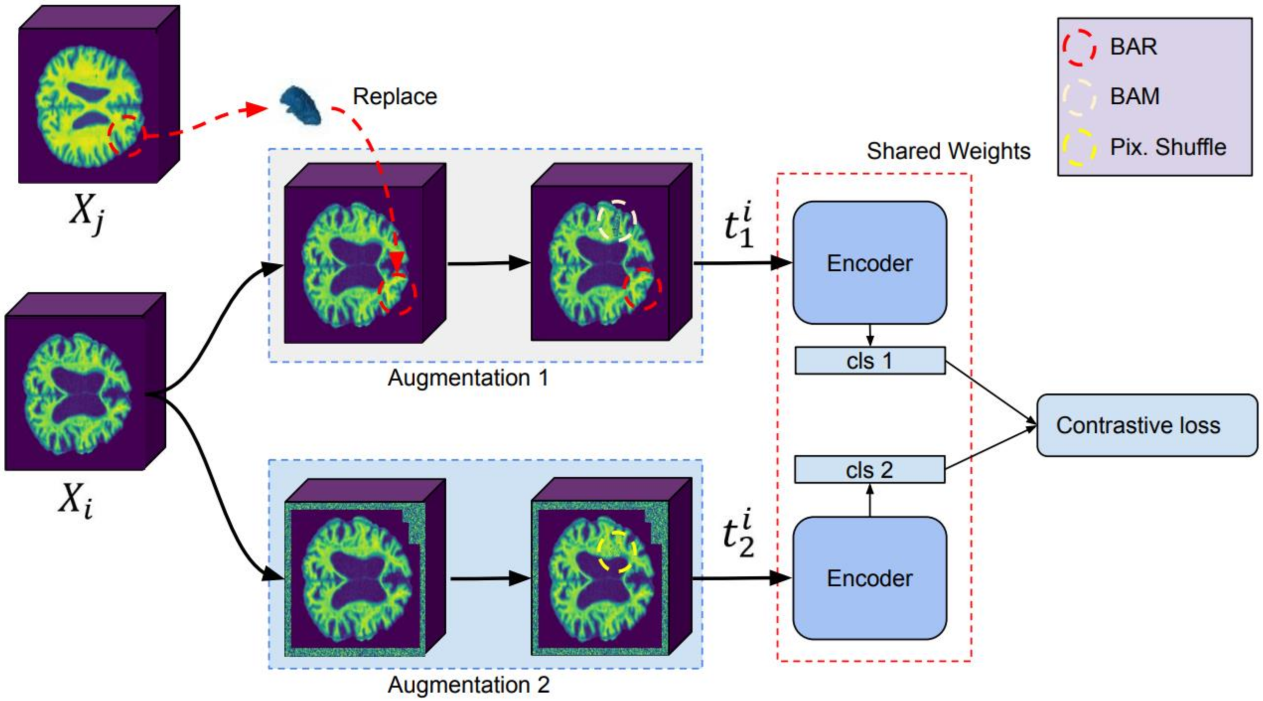
> Soft labels generated by BAR can be exploited with a slight modification on the supervised contrastive loss to learn the relative similarity between pairs.

[3]

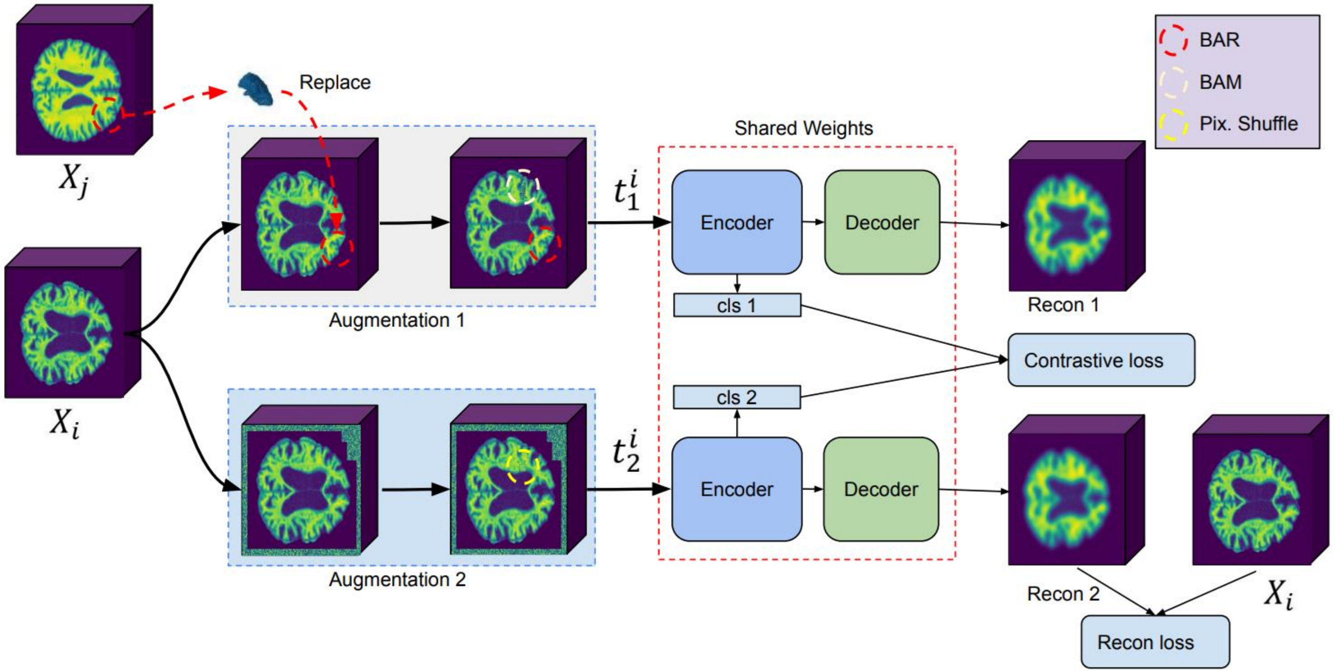
$$L^{c}_{NCE} = -\sum_{k=1}^{n} \frac{\varphi(y_k^p, y_i^p)}{\sum_{j=1}^{b} \varphi(y_j^p, y_i^p)} log \frac{e^{\theta(t_1^i, t_2^k)}}{\frac{1}{b}\sum_{j=1}^{b} e^{\theta(t_1^i, t_2^j)}}$$

[3] Dufumier, Benoit, et al. "Contrastive learning with continuous proxy meta-data for 3d mri classification." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021.

Mixture Learning Framework with BAR



Proposed Pre-Training Framework

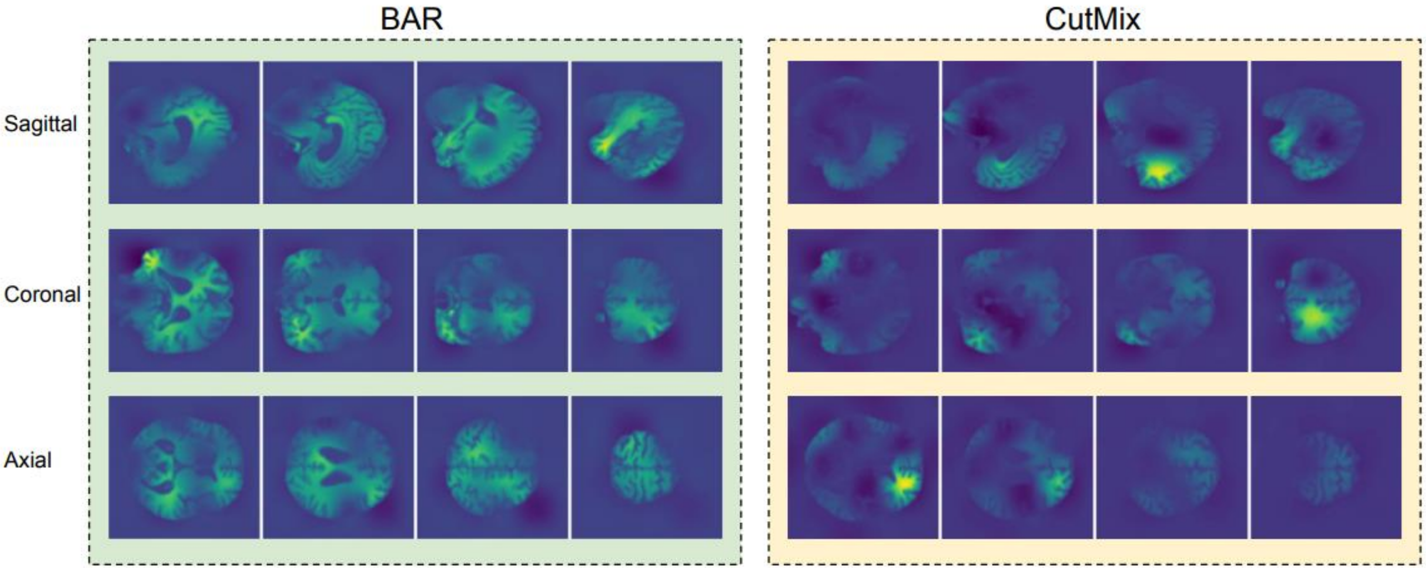


Results

> We used ADNI on all our experiments

Framework	Method	Precision	Recall	Accuracy
No Pre Training	ViT from scratch	74.38 ± 7	85.6 ± 3.1	80.83 ± 3
Self Supervised Pre-Training + Fine Tuning	Contrastive	78.42 ± 4.5	81.18 ± 1.6	80.1 ± 1.9
	Recon	78.6 ± 5	85.57 ± 1.1	82.69 ± 2.5
	Contrastive + Recon	80.2 ± 4.1	85.77 ± 2	83.4 ± 1.7
Supervised Pre-Training +Fine Tuning	CutMIX	83.06 ± 4.8	87.08 ± 3.5	85.29 ± 2.8
	CutMIX + Recon	84.6 ± 3.8	87.9 ± 2.2	86.4 ± 1
	BAR	84.7 ± 3.3	87.6 ± 2.1	86.3 ± 1.1
	BAR + Recon	$\mathbf{86.24 \pm 3}$	$\mathbf{88.08 \pm 2.3}$	$\mathbf{87.22 \pm 0.8}$

Attention Visualization for the AD case

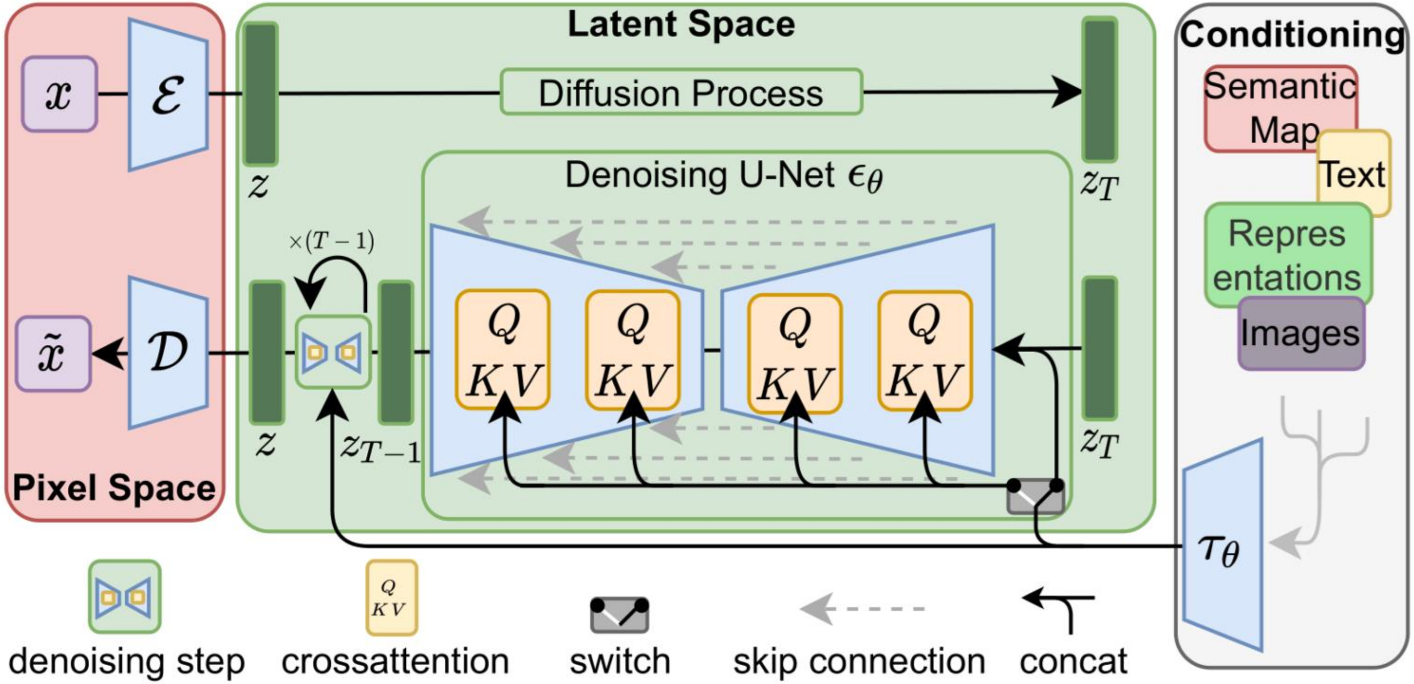


A Bonus Talk on Diffusion Models

- Text-to-image diffusion models have shown unprecedented success in recent research!
- Stable Diffusion
- Dall-E
- PARTI

"a photograph of an astronaut riding a horse"

Diffusion Models



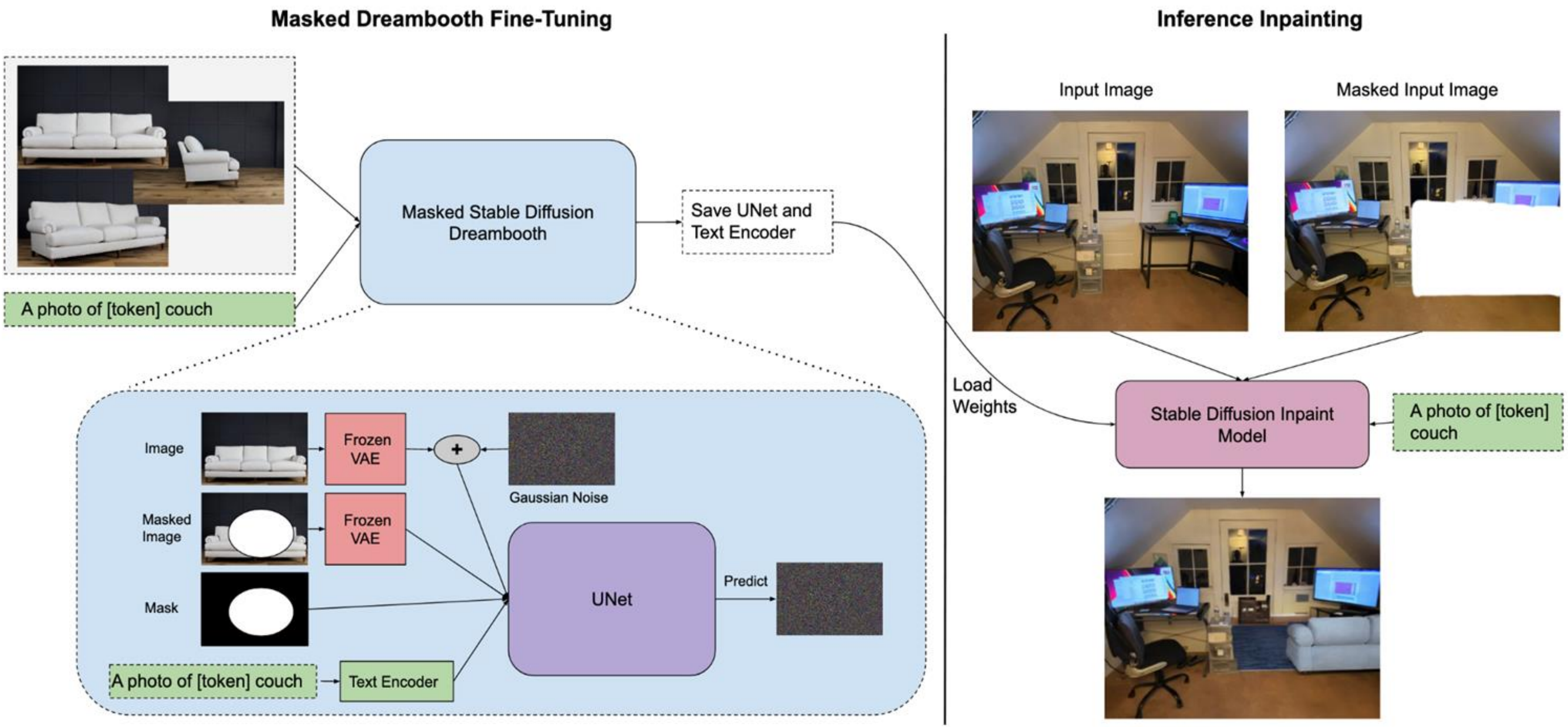
Diffusion Model Applications for E-Commerce

- Augmented reality allows Virtual Try-on but it's very expensive as it requires some manual labor.
- Can we do Virtual-Try On without 3D modeling directly on 2d images using the diffusion models?

DREAMbooth-inPAINT(DreamPaint)

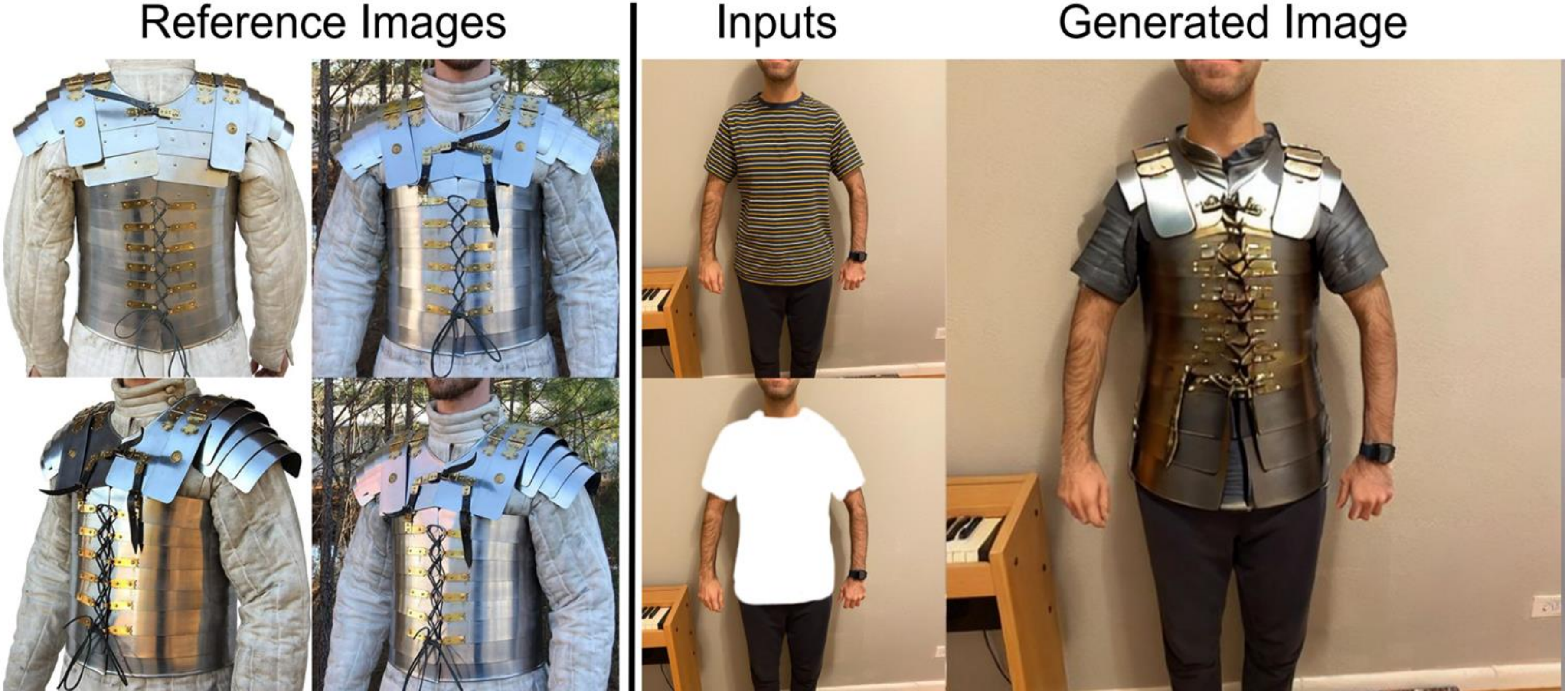
- We propose DreamPaint, a framework to intelligently inpaint any e-commerce product on any user-provided context image!

DreamPaint



Masked Dreambooth Fine-Tuning
Inference Inpainting
A photo of [token] couch
Masked Stable Diffusion Dreambooth
Save UNet and Text Encoder
Image
Masked Image
Mask
Frozen VAE
Frozen VAE
+
Gaussian Noise
UNet
Predict
A photo of [token] couch
Text Encoder
Load Weights
Input Image
Masked Input Image
Stable Diffusion Inpaint Model
A photo of [token] couch

Example 1



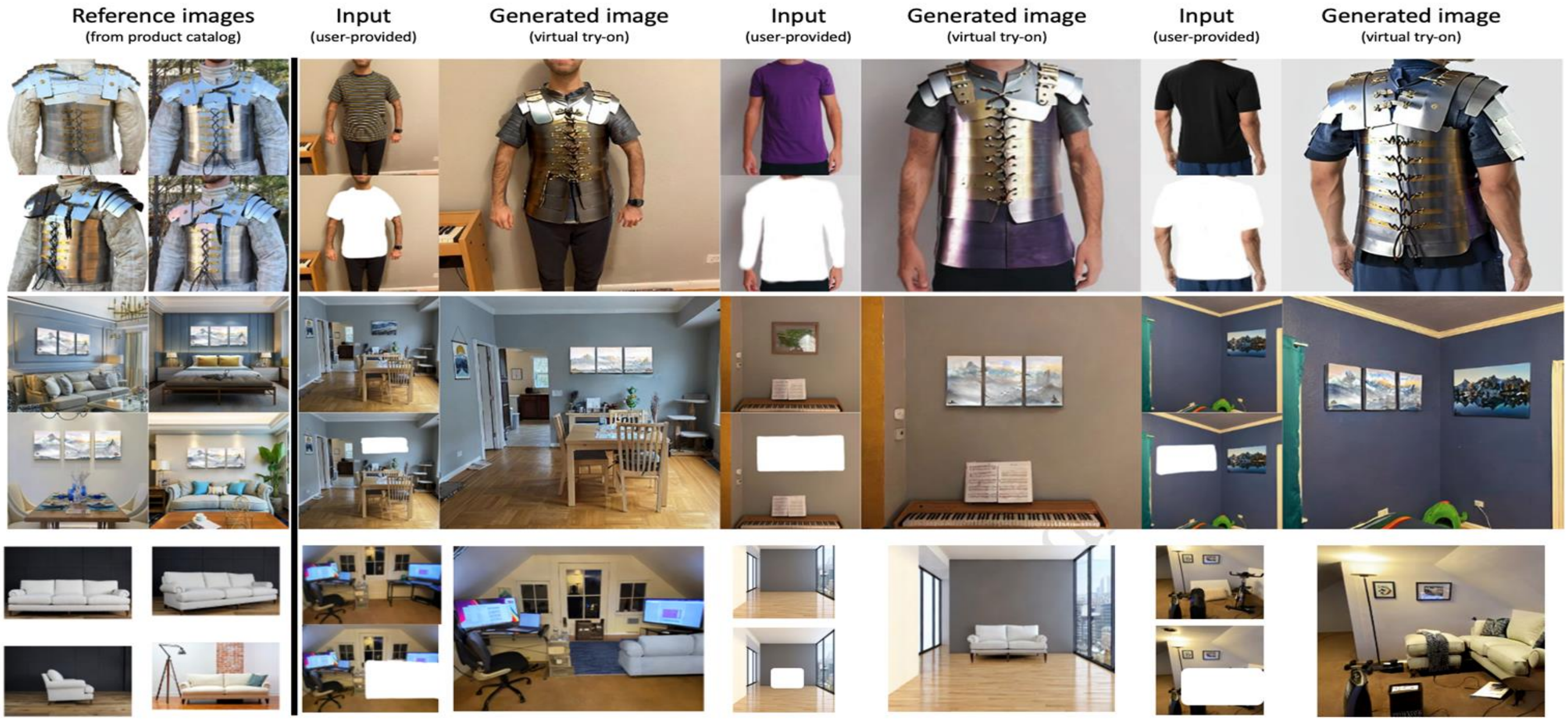
Example 2

Reference Images

Inputs

Generated Image

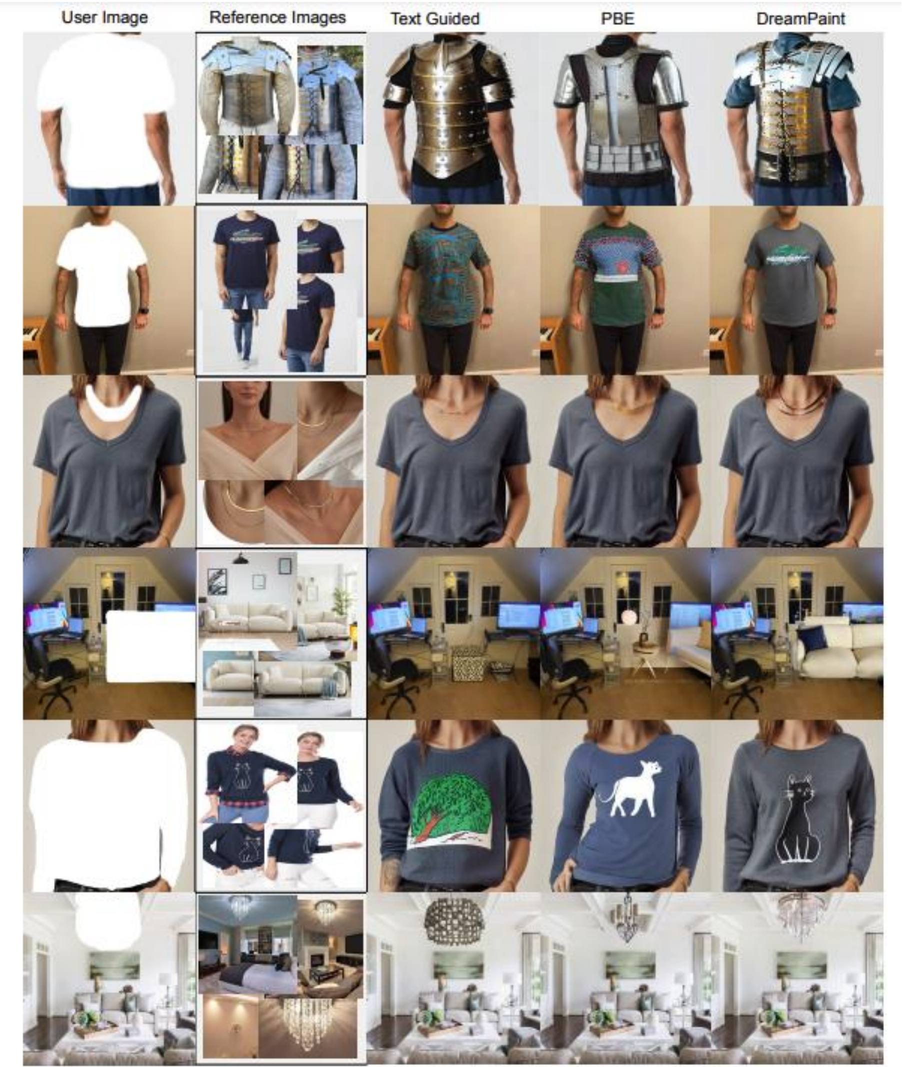
More Examples



Compared to SOTA

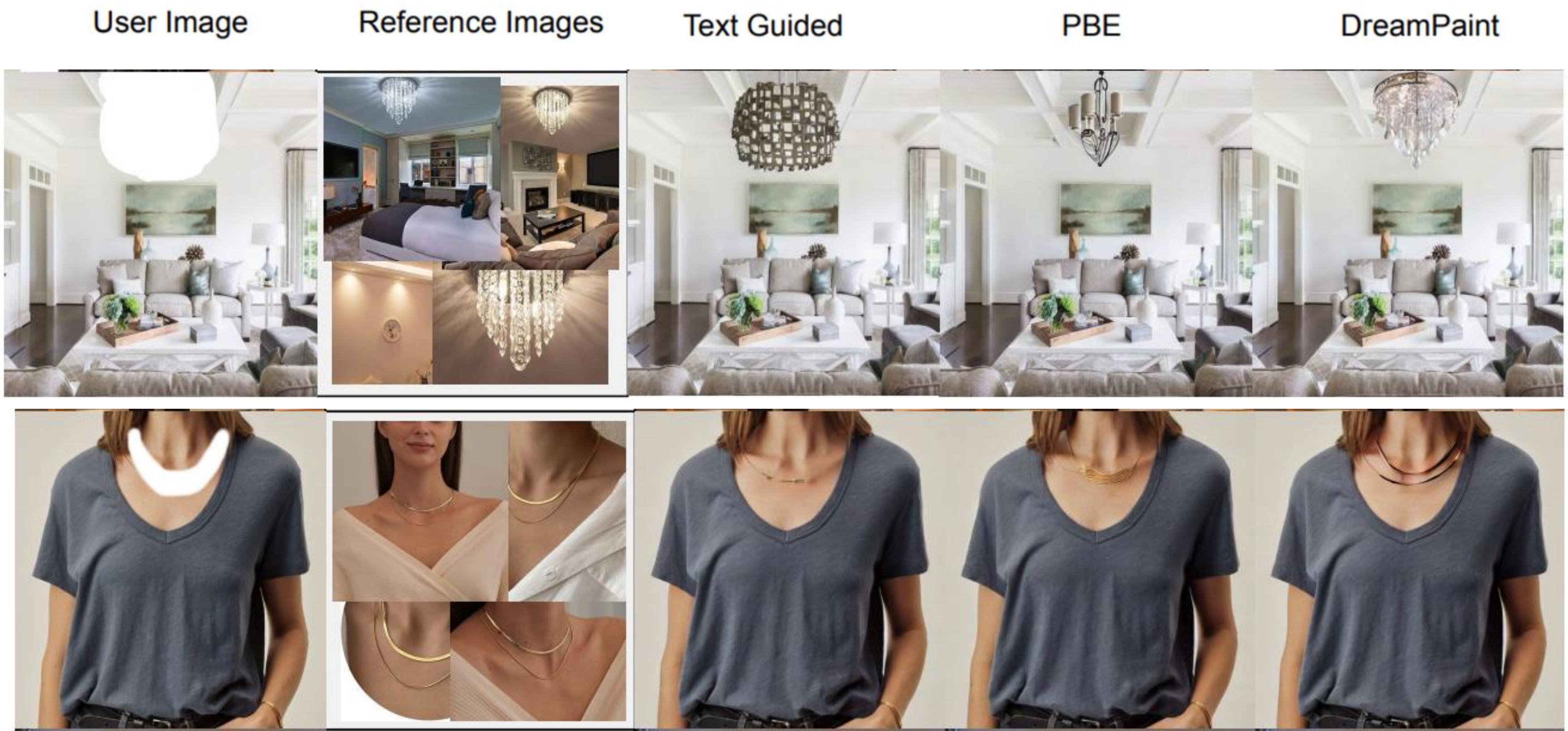
Inpainting is either done by text guided models, or image guided models.

Text Guided column shows results generated by the asin title (which does not have the capacity to fully describe the item) or a SOTA model, Paint By Example (PBE), which uses a single exemplar image, but mostly omits the fidelity that is required in the e-commerce setting.



Compared to SOTA

Compared to SOTA



Quantitative and Human Survey Results

Table 1: Quantitative comparison of CLIP score between our approach and the baselines.

Method	CLIP Score(↑)
SD Inpaint with Text Guidance	0.62
Paint by Example	0.68
Ours	0.70

Table 2: Average user ratings for each method measuring similarity to the reference image and how harmoniously the generated image.

Method	Similarity(↓)	Harmony(↓)
SD Inpaint with Text Guidance	4.41	2.75
Paint by Example	3.82	2.57
Ours	2.68	2.33