SKIN DISEASE IMAGE CENERATION

INTRODUCTION

- [illegible]

- [illegible]

- [illegible]

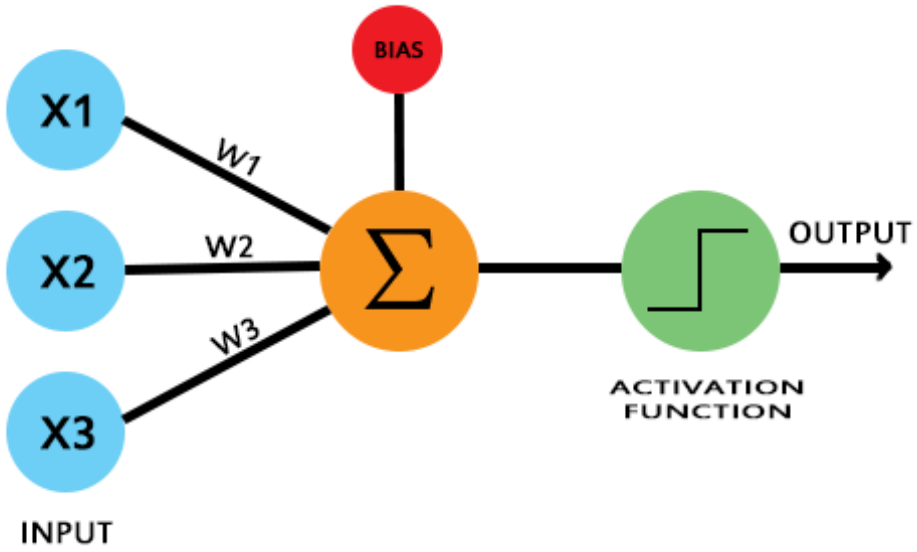
PURPOSE

- [illegible]
- [illegible]
- [illegible]
- [illegible]

HISTORY

- [illegible]

- [illegible]

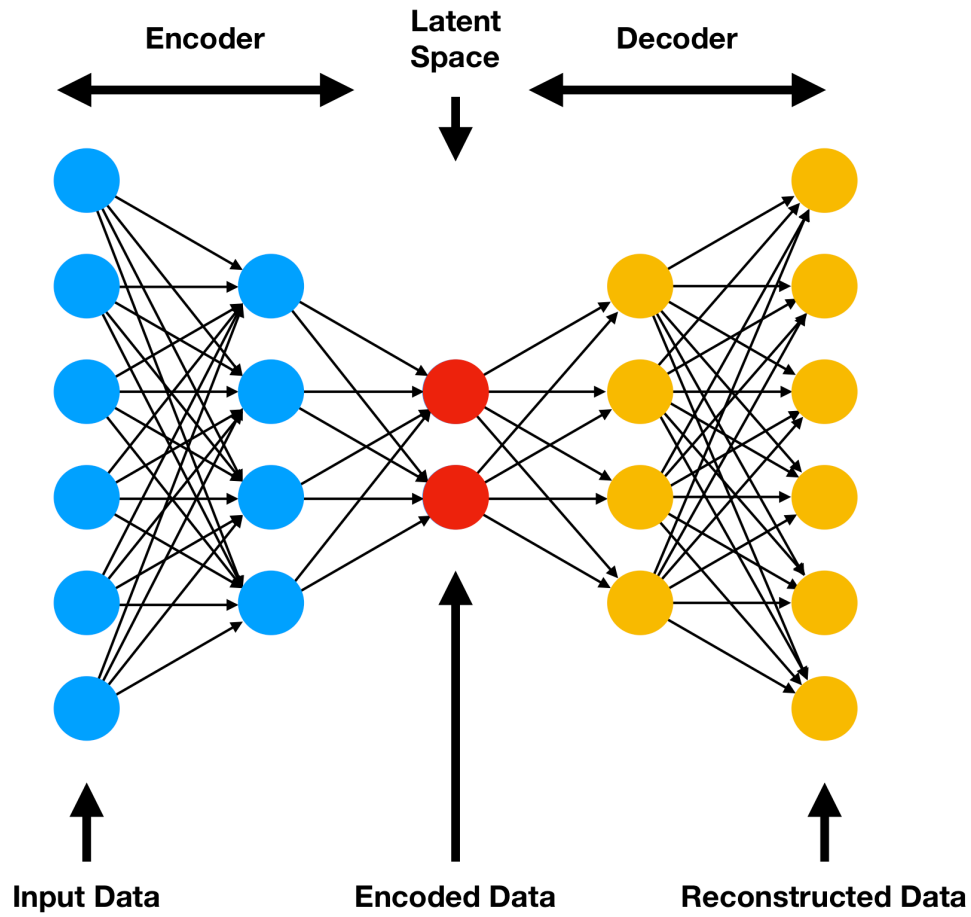


[illegible]

HISTORY (II)

- [illegible]

- [illegible]


Encoder
Latent Space
Decoder
Input Data
Encoded Data
Reconstructed Data

[illegible]

HISTORY (III)

- [illegible]

- [illegible]

DATASET

- [illegible]

- [illegible]

[illegible]

DATASET (II)

- [illegible]

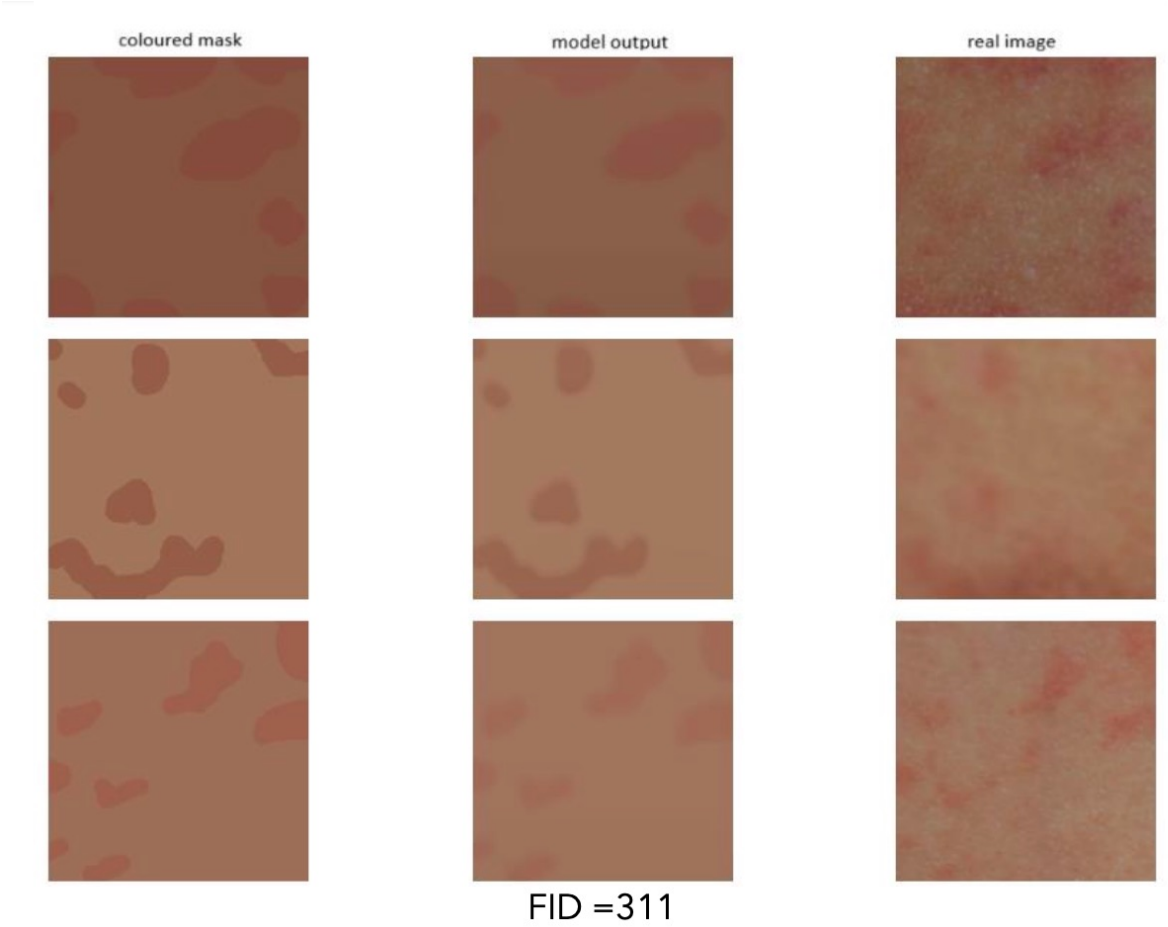
  - [illegible]

  - [illegible]

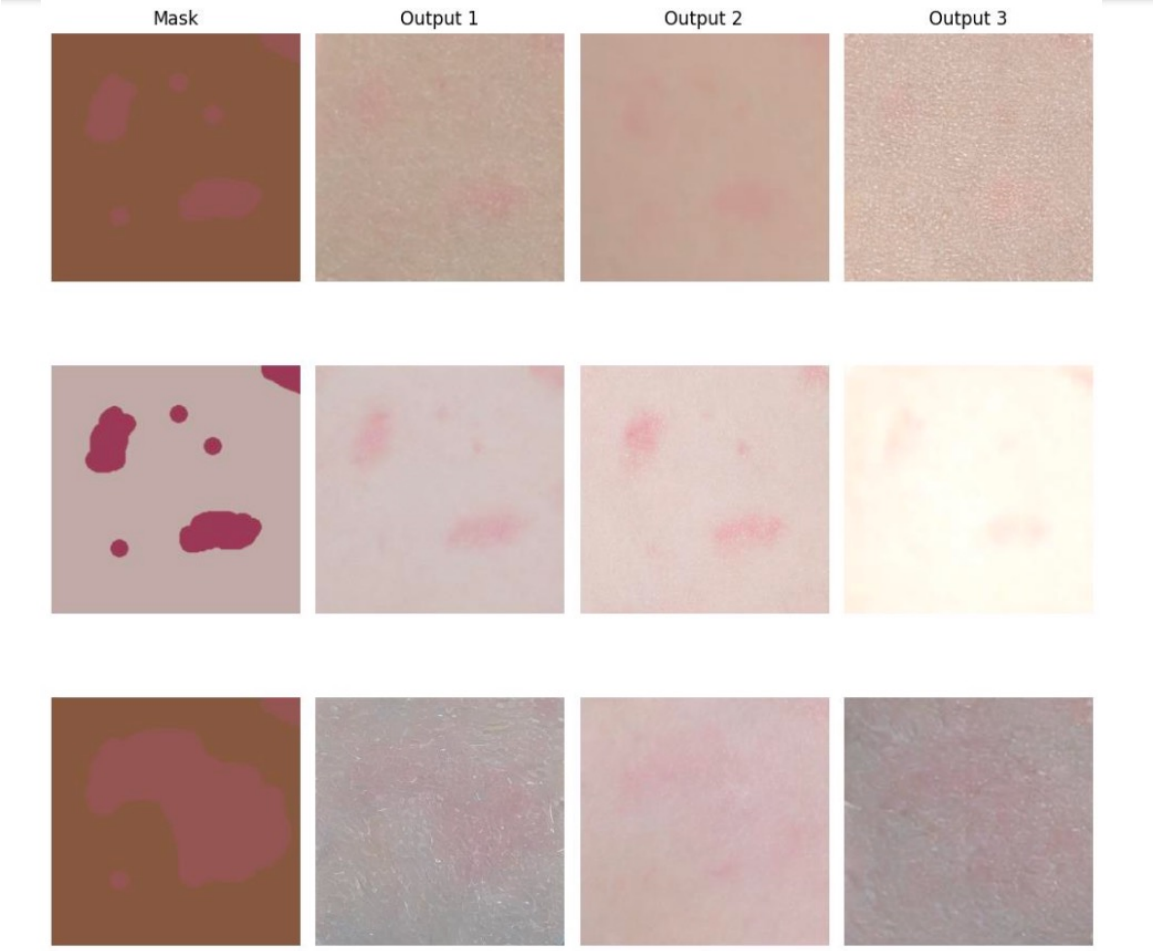
[illegible]

MODELS

DERMGAN



DERMDIFF



FID = 74

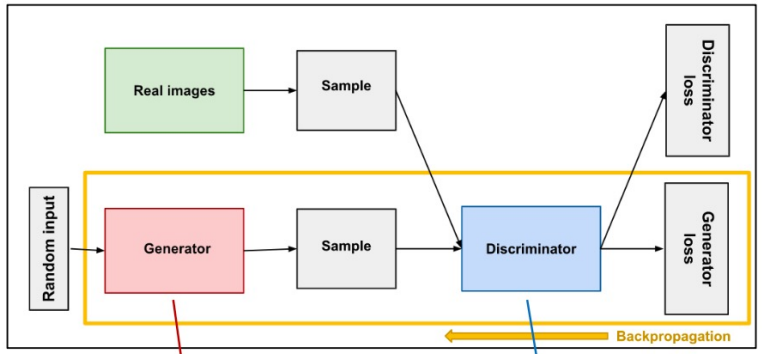
LIMITATIONS OF FIRST APPROACHES

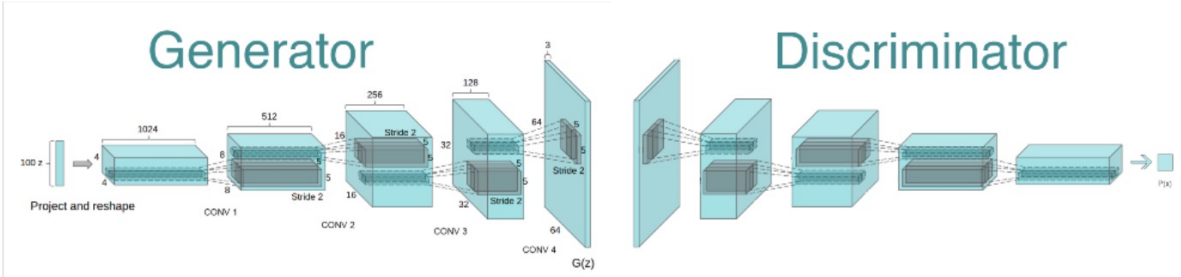
- **Using masks is not a flexible/generalizable approach**
- **color masks increases the preprocessing time.**
- **augmentation**

DCGAN

Generator [illegible]

Discriminator [illegible]



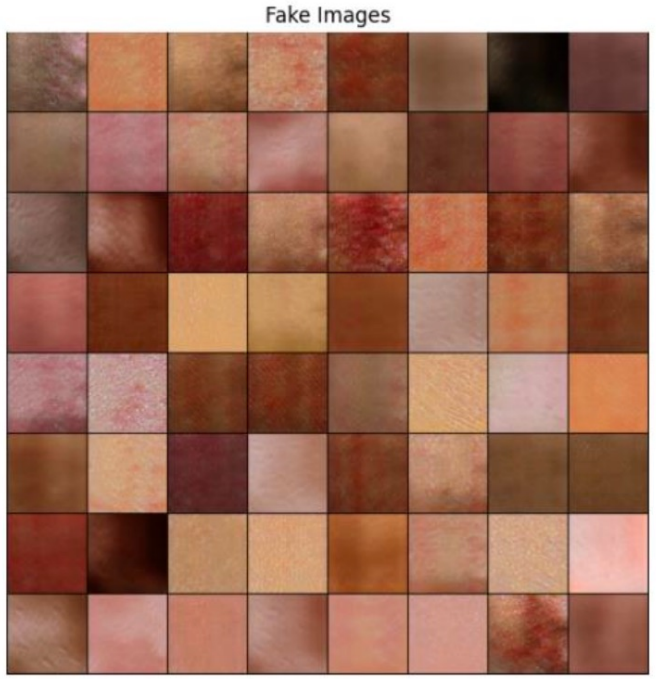


- Architecture is fully convolutional.
- Resize-convolutions used instead of Transpose Convolutions (to avoid checkerboard patterns).
- One sided label smoothing for the descriminator (1 →0.9), to avoid descriminator overconfidence.

```
maximize log(D(G(z)))
```

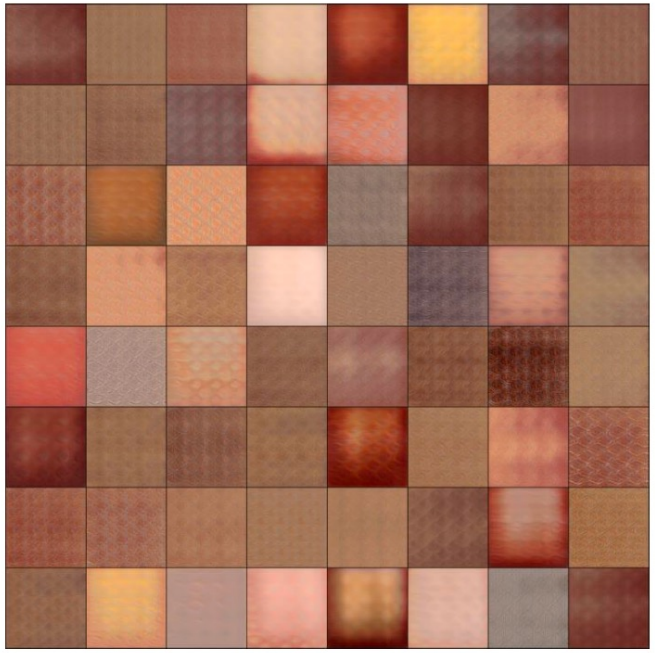
```
maximize log(D(x)) + log(1 - D(G(z)))
```

PROBLEMS OF DCGAN



128x128 patches

0.08 sec/image (Nvidia T4)
FID = 98



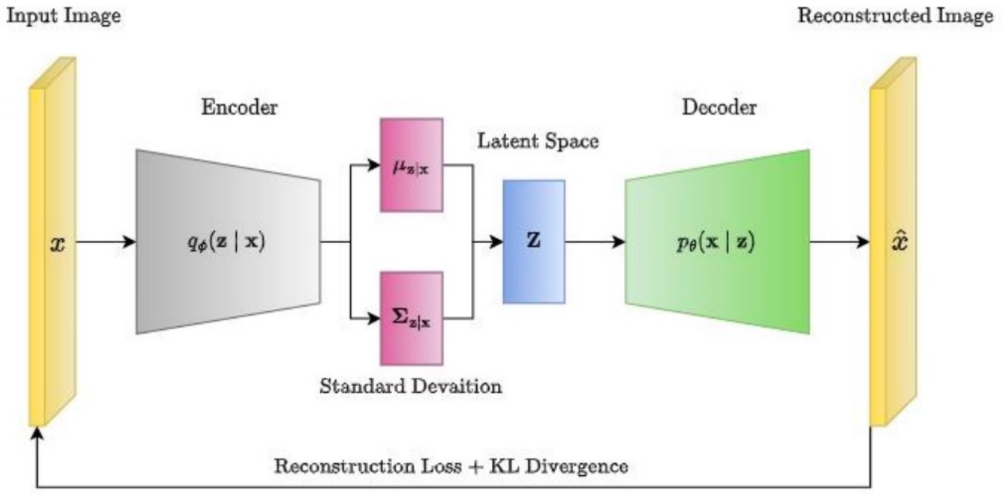
256x256 patches

0.10 sec/image (Nvidia T4)
FID = 358

CVAE

- Class of neural networks used to learn probabilistic representations of complex data in a continuous latent space.
- Composed by two main parts and **Decoder**.
- Encoder composed of Convolutional Blocks and Dense Layers.
- Decoder composed of Dense Layers and Resize-Convolutional Blocks.
- Two additional Dense Layers (estimate mean and std)

Encoder



$$L(\theta, \phi, x) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] + \beta \cdot D_{\mathrm{KL}}(q_\phi(z|x)||p(z))$$

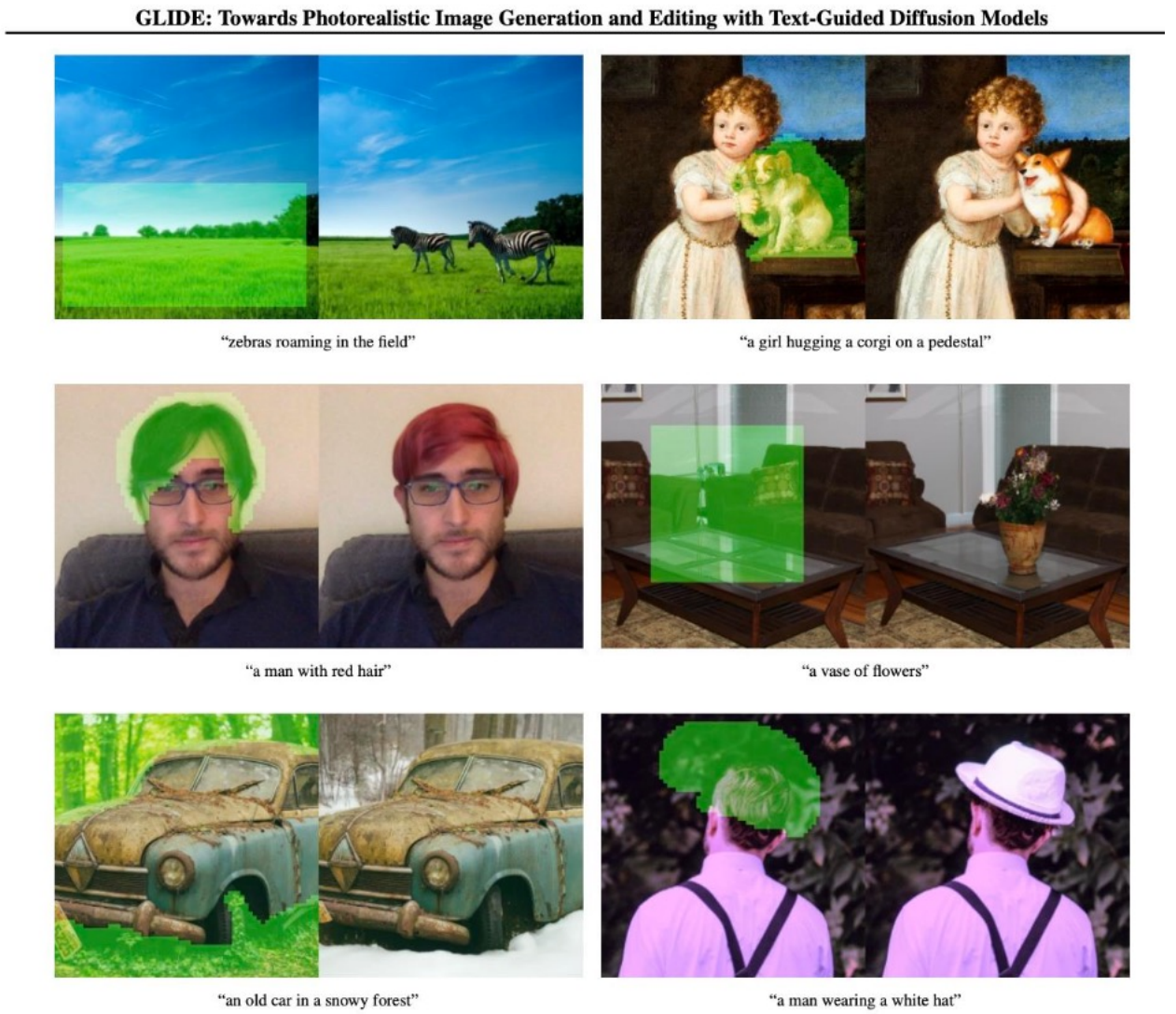
CVAE (II)

LIMITATIONS OF VARIATIONAL AUTOENCODERS

Hyperparameter Sensitivity:

Blurry Outputs:

GLIDE


GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models
"zebras roaming in the field"
"a girl hugging a corgi on a pedestal"
"a man with red hair"
"a vase of flowers"
"an old car in a snowy forest"
"a man wearing a white hat"

GLIDE (II)

- [illegible]



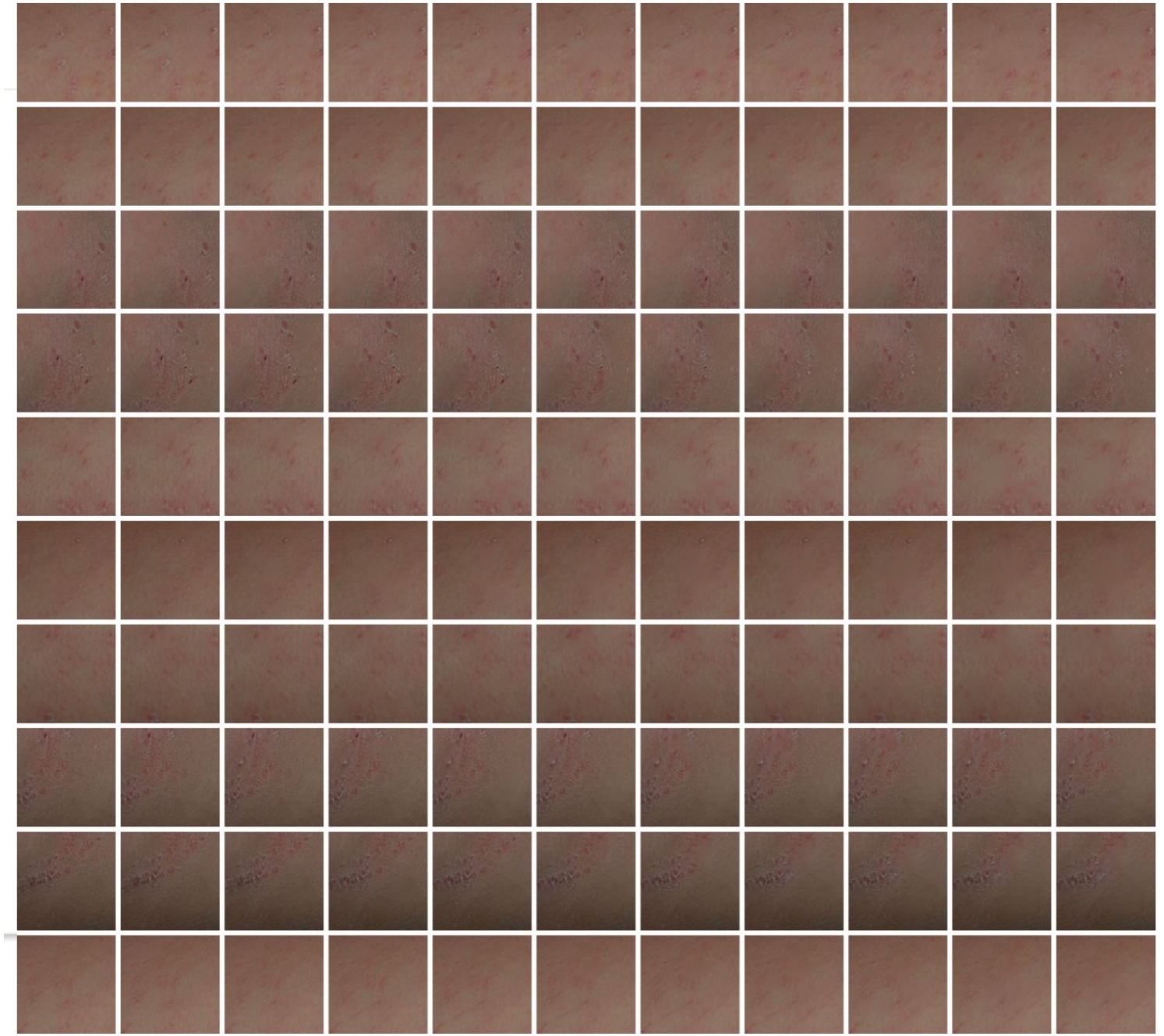
[illegible]

- [illegible]

- [illegible]

- [illegible]

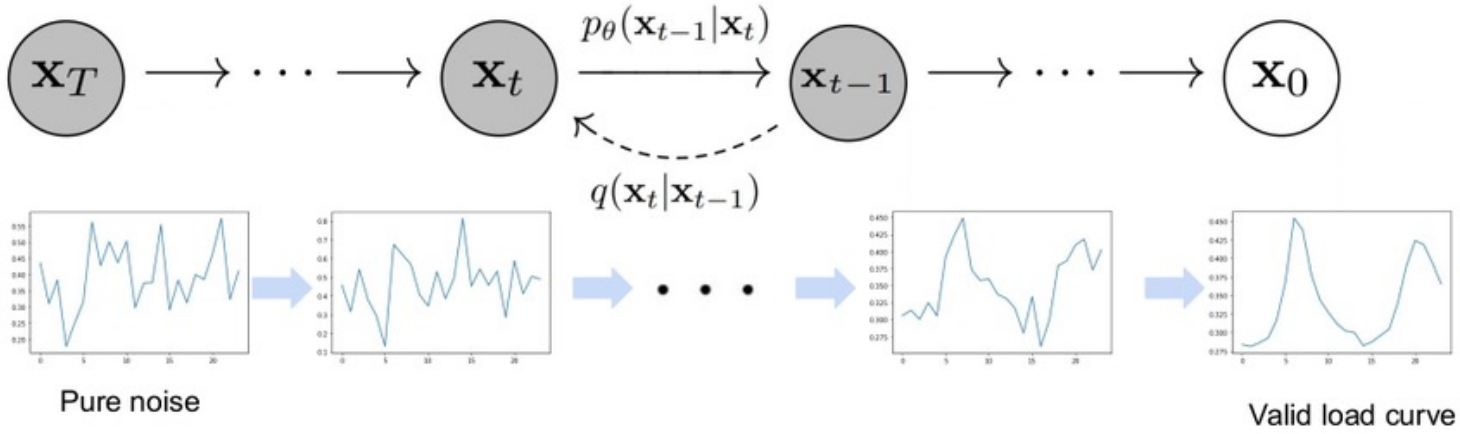
THEIR RESULTS



DENOISING DIFFUSION PROBABILISTIC MODELS

- [illegible]

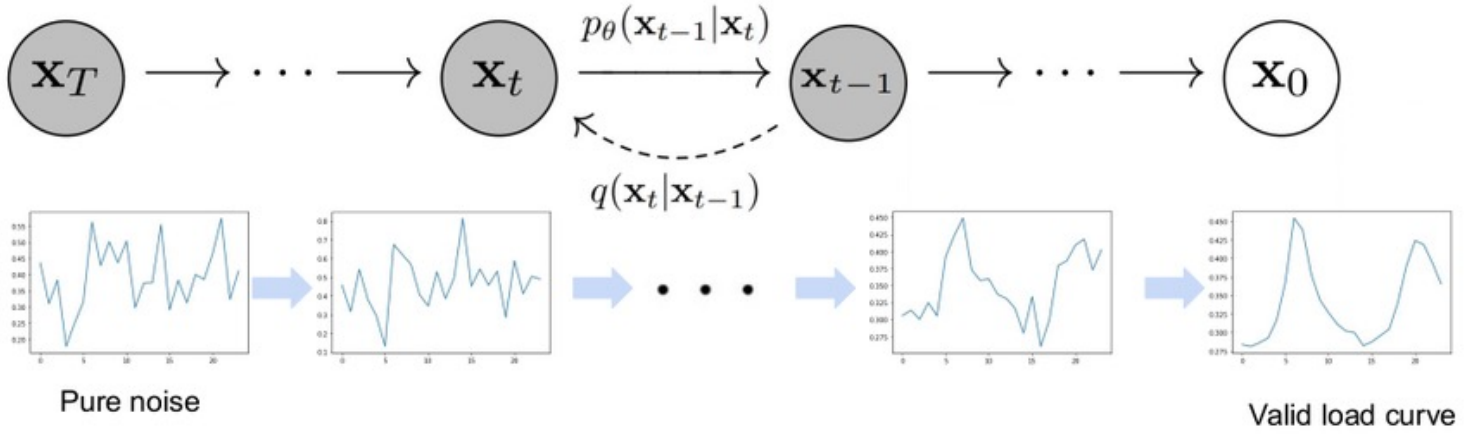
- [illegible]



DENOISING DIFFUSION PROBABILISTIC MODELS (II)

- [illegible]

- [illegible]



STABLE DIFFUSION XL

- [illegible]

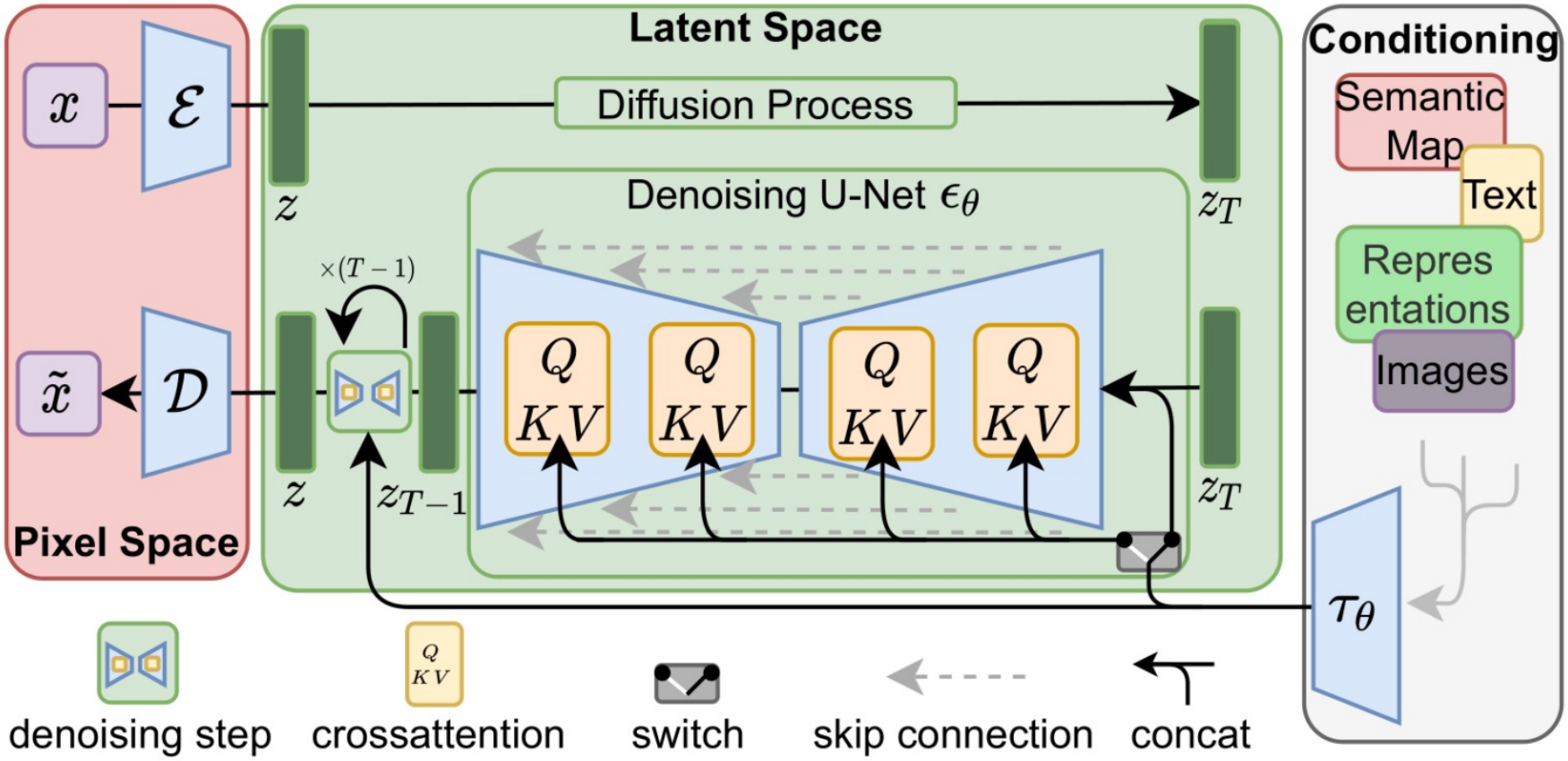
- **Key Features:** [illegible]
  - [illegible]

- **Advantages:** [illegible]
  - [illegible]

STABLE DIFFUSION XL (II)

- **Image Quality:** [illegible]
  - [illegible]
- **Performance:** [illegible]
  - [illegible]
- **Speed and Efficiency** [illegible]

ARCHITECTURE



APPLICATIONS OF SDXL

- **Image Generation:** [illegible]
- **Custom Fine-Tuning:** [illegible]
- **Inpainting and Edits:** [illegible]

[illegible]

DREAMBOOTH

- [illegible]

- [illegible]

- **Key Features:**
  - [illegible]
  - [illegible]

ADVANTAGES OF DREAMBOOTH

- **High Fidelity:** Maintains key visual features of the subject.
- **Flexibility:** Generates images of the subject in various contexts.
- **Applications:** Suitable for subject recontextualization, artistic modifications, and more.

LORA

- **LoRA** [illegible]

- **Key Features:**
  - [illegible]
  - [illegible]

[illegible]

BENEFITS OF LORA

- **Computational Efficiency:** [illegible]
- **Memory Savings:** [illegible]
- **Improved Generalization:** [illegible]
- **Adaptability:** [illegible]

THE MODELS

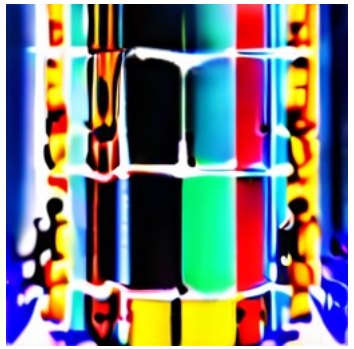
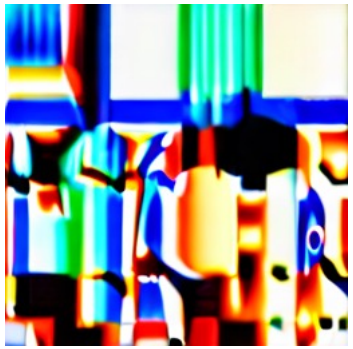
- [illegible]
- [illegible]
- [illegible]
- [illegible]

THE ENVIRONMENT

- [illegible]
- [illegible]
- [illegible]

THE IMAGES

- [illegible]

- [illegible]

THE IMAGES

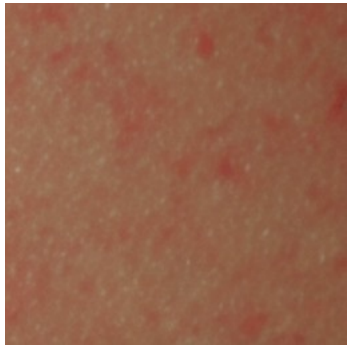
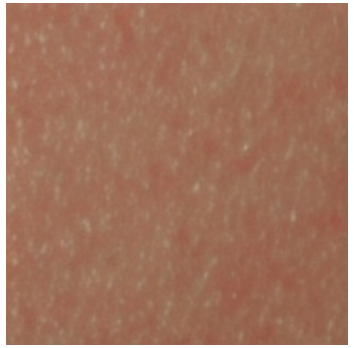
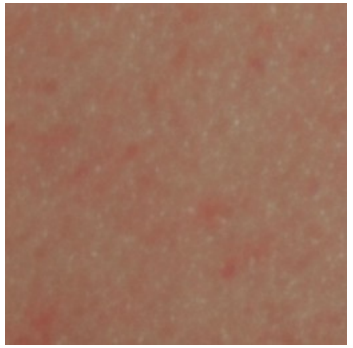
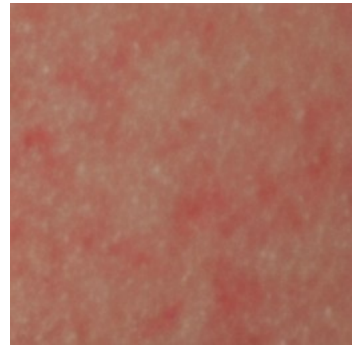
- [illegible]




- [illegible]

THE IMAGES

- [illegible]

- [illegible]

RESULTS

- [illegible]
- [illegible]

CONCLUSIONS

- [illegible]
- [illegible]
- [illegible]

FUTURE STEPS

- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]

THANK YOU FOR YOUR ATTENTION