Mahidol University
Wisdom of the Land

AUGMENTING MEDICAL IMAGE CLASSIFIERS WITH SYNTHETIC DATA FROM LATENT DIFFUSION MODELS

Msc. Pasit Supholkhan ID 6536115

07
AFFORDABLE
AND CLEAN ENERGY

Introduction

- The paper explores the use of **latent diffusion models** to **generate synthetic** images of **skin diseases** and demonstrates that **augmenting** model training **with** these **synthetic** images **improves performance** in **data-limited** settings.
- The **diversity** of disease appearance across different **skin tones** makes skin disease classification a **suitable** case study for **synthetic image** generation.

Methodology

Methodology

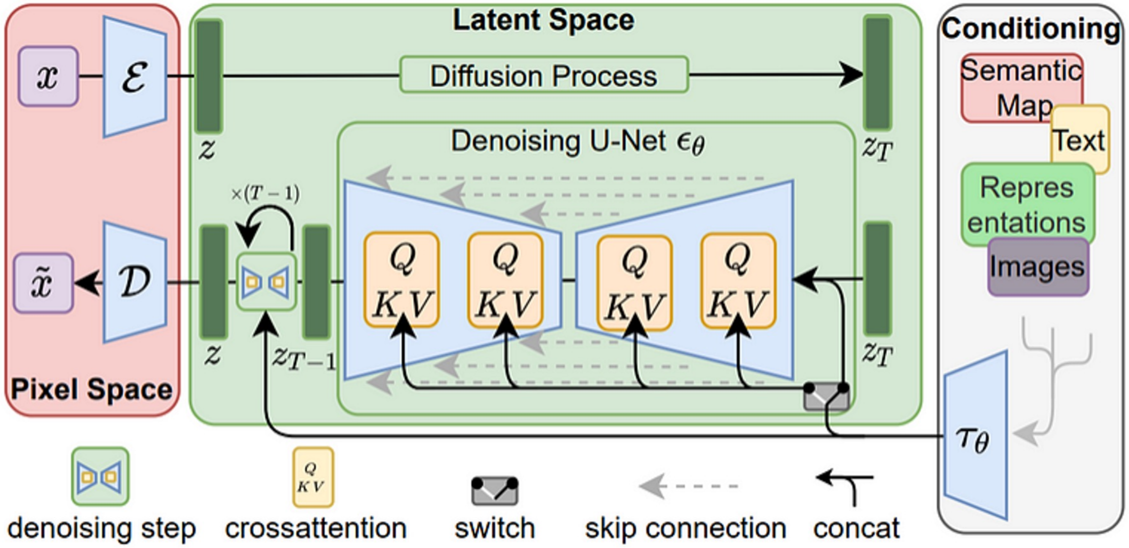
- The paper utilized **Stable Diffusion**, a large-scale latent diffusion model, for generating high-resolution synthetic images.
- **Three** different generation methods were employed: **text-to-image**, **inpainting**, and **inpainting followed by outpainting**. These methods involved conditioning the generative models on text prompts and masking/replacing specific regions of the images.
- Two variants of the generative model were considered: one using the pretrained diffusion model directly, and another fine-tuned on a small number of training images. The **fine-tuning** was performed using **DreamBooth** and the images used for fine-tuning were not included in the test sets.
- The experiments utilized the **Efficient-Net V2-M image classification** model **pretrained** on **ImageNet**. Training was performed for 30 epochs, and various image transforms were applied to both real and synthetic images.
- Synthetic data was **added only** to the **training** set and was produced independently of the images in the held-out test sets.

DATA SOURCE

Data used in this paper

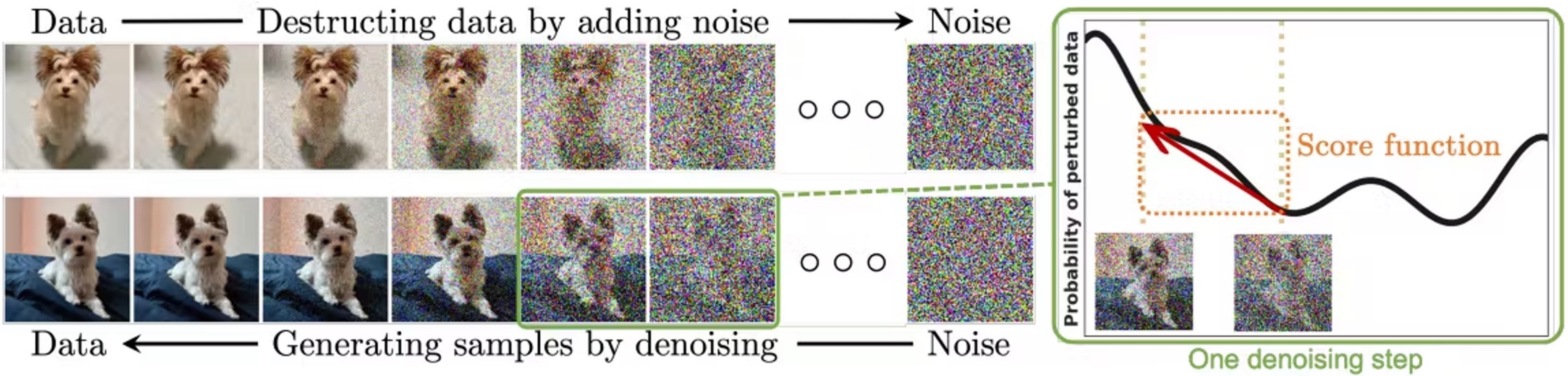
- **FiztPatrick 17k-9** datasets, **Sandford DDI**
- The paper generated and analyzed a new dataset of **458,920 synthetic** images using several generation strategies from latent diffusion models.
- The synthetic images were produced using **three** different **generation** methods: text-to-image, inpainting, and inpainting followed by outpainting.
- Image transforms, such as random **cropping**, horizontal **flips**, warping, and lighting adjustments, were applied to both real and synthetic images during the experiments.

Latent Diffusion Model



Became known in 2021

Latent Diffusion Model

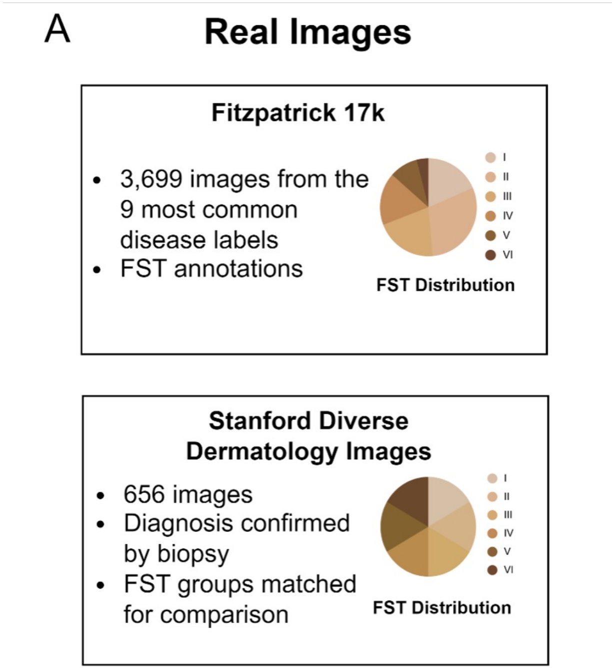


https://arxiv.org/abs/2209.00796

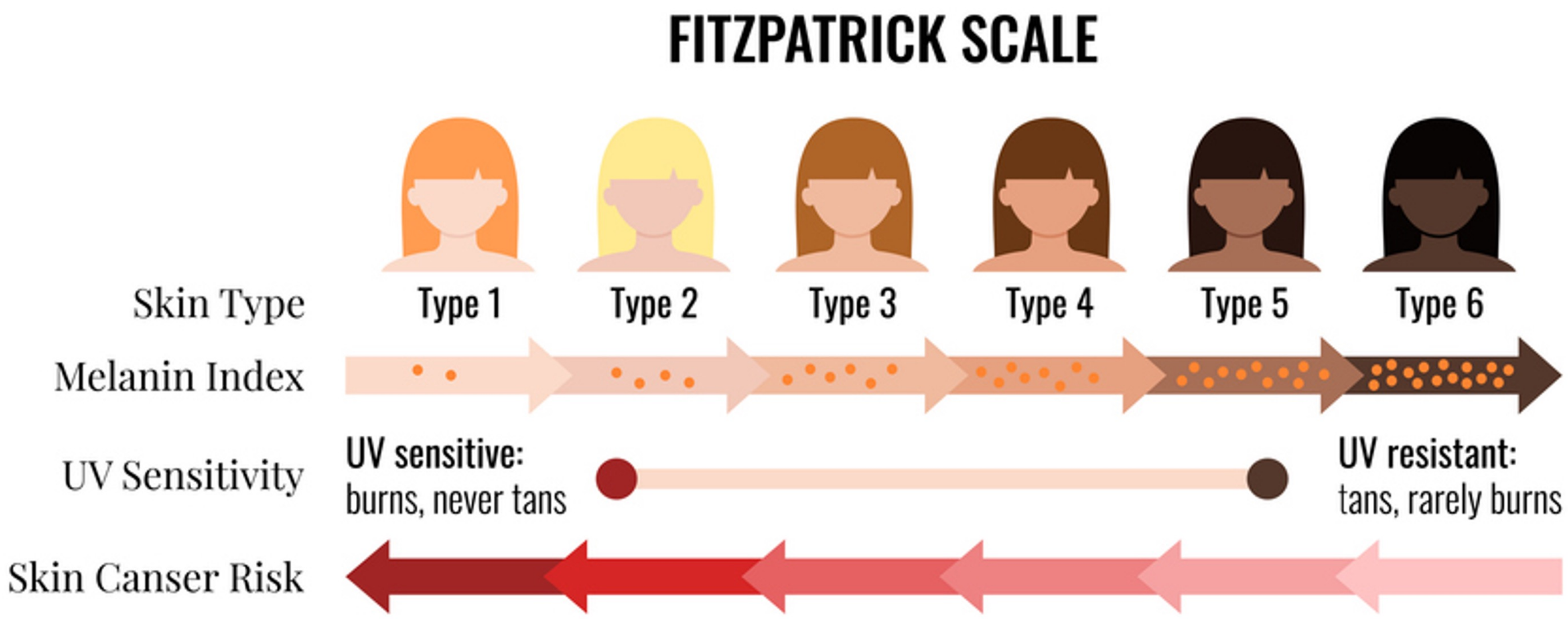
Datasets Real Images

9 most common skin conditions in Fitzpatrick 17k:

- psoriasis (624),
- squamous cell carcinoma (480),
- lichen planus (452),
- basal cell carcinoma (446),
- allergic contact dermatitis (398),
- lupus erythematosus (337),
- sarcoidosis (326),
- neutrophilic dermatoses (324), and photodermatoses (312),
- which we term the F17k-9 dataset

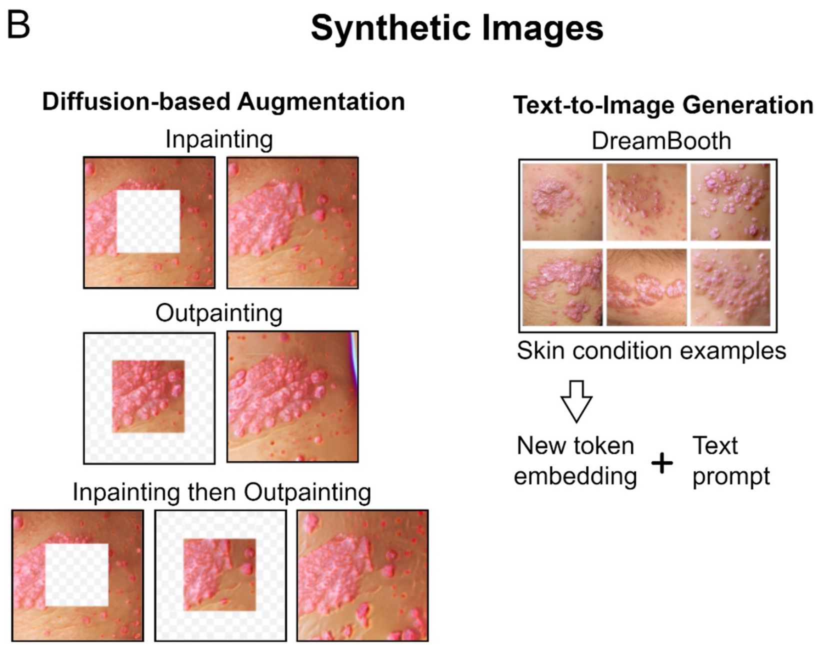


https://arxiv.org/abs/2308.12453


FITZPATRICK SCALE
Skin Type
Type 1
Type 2
Type 3
Type 4
Type 5
Type 6
Melanin Index
UV Sensitivity
UV sensitive:
burns, never tans
UV resistant:
tans, rarely burns
Skin Canser Risk

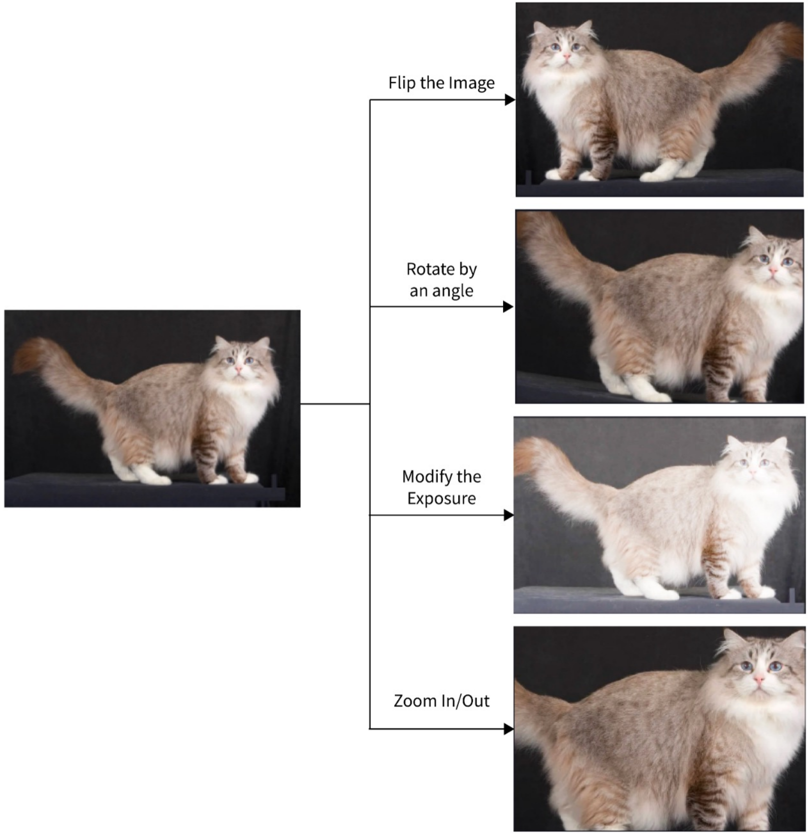
Synthetic Images

- text prompts with and without reference images, we generated synthetic images
  - inpainting
  - outpainting
  - in-then-outpainting
  - text-to-image

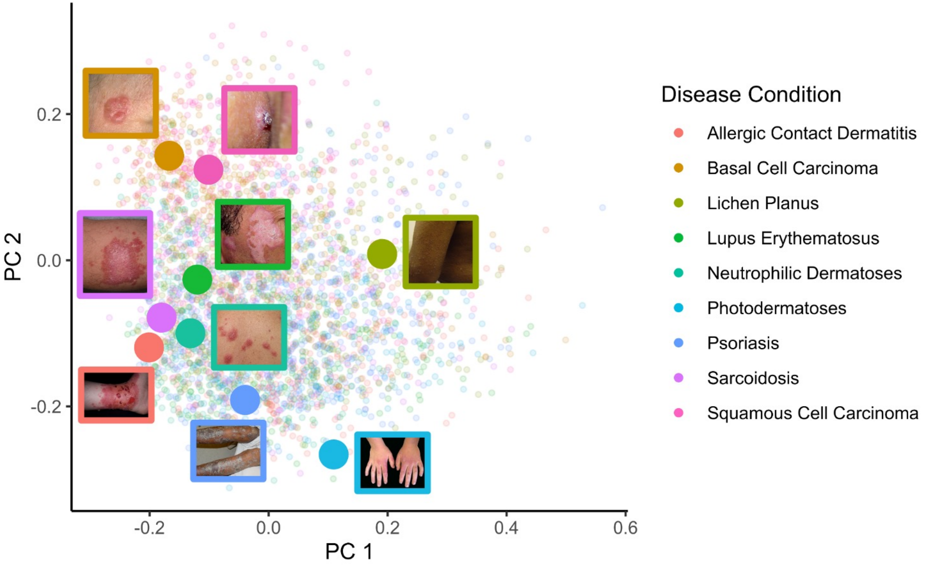


https://arxiv.org/abs/2308.12453

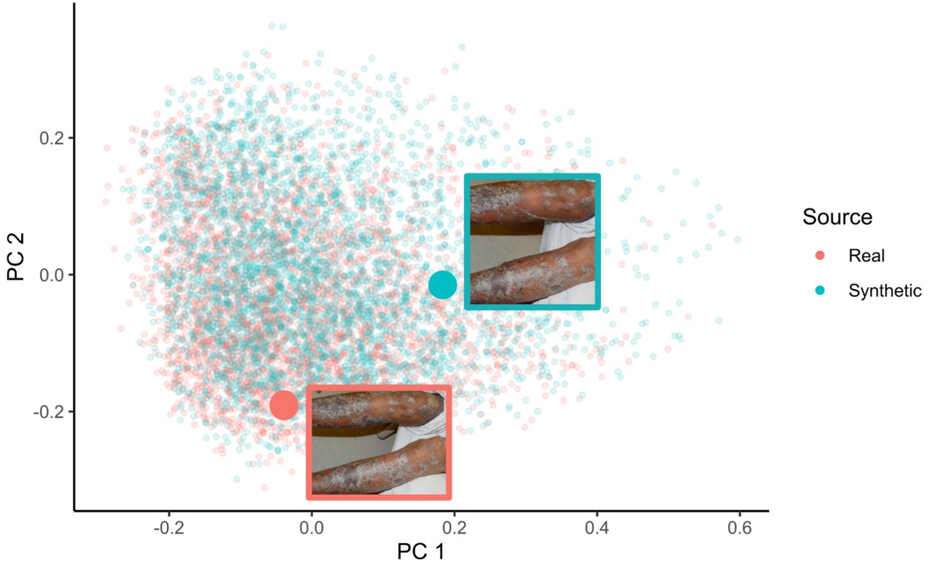
Data Augmentation



Generating synthetic images distribution



Real



Real vs Synthetic

https://arxiv.org/abs/2308.12453

Performance Evaluation

- Skin disease classifiers
- Synthetic augmentations,
    - comprising 10 synthetic images for each real image, or
- image transforms, comprising
    - random flipping, cropping, rotating, warping, and lighting changes

Accuracy gain from using synthetic data stack with image transforms

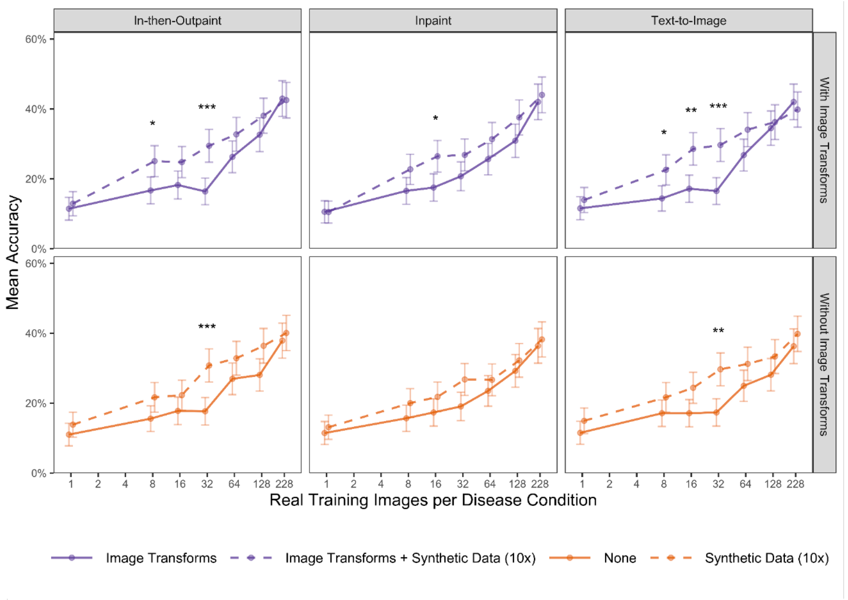
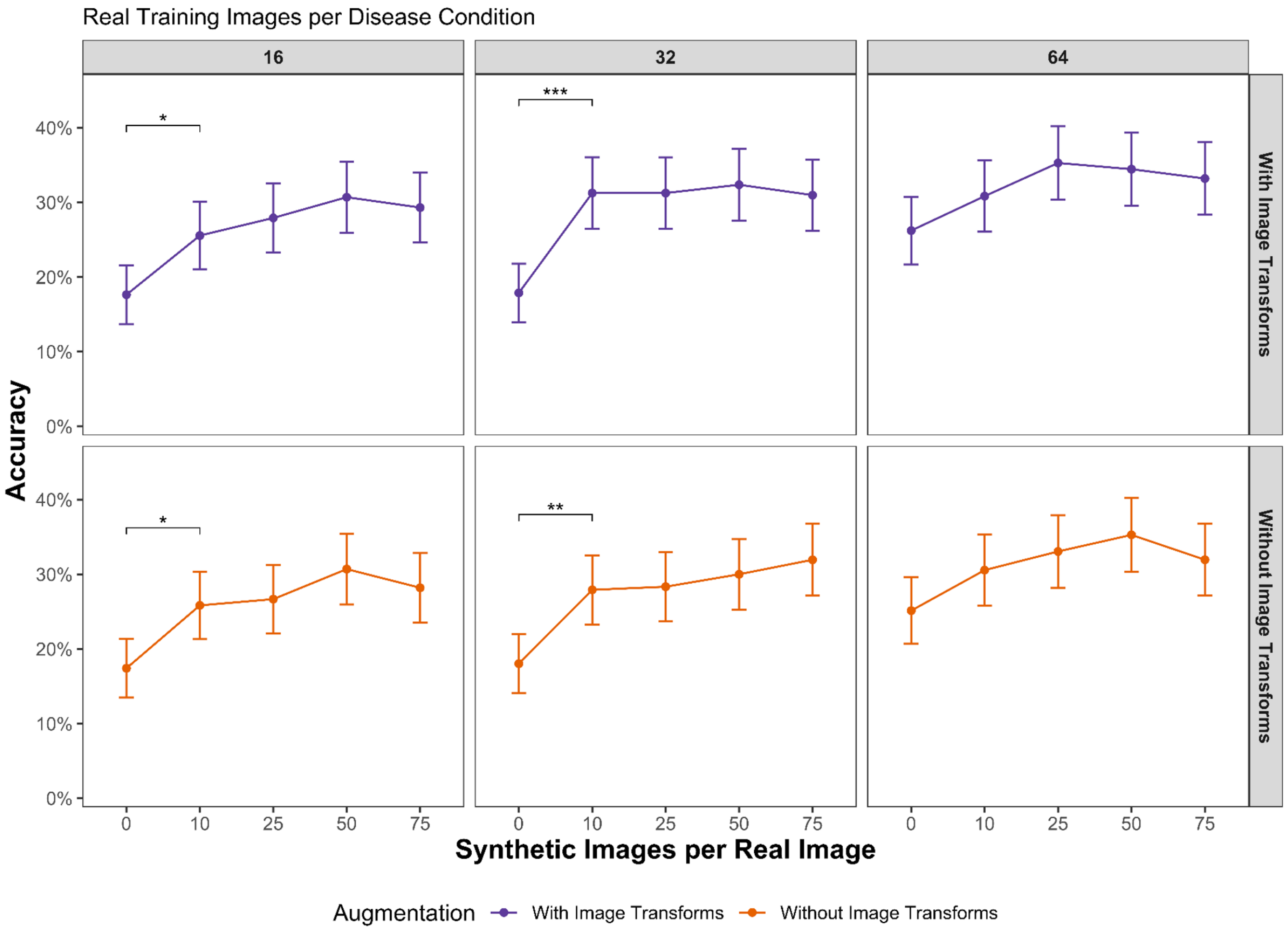
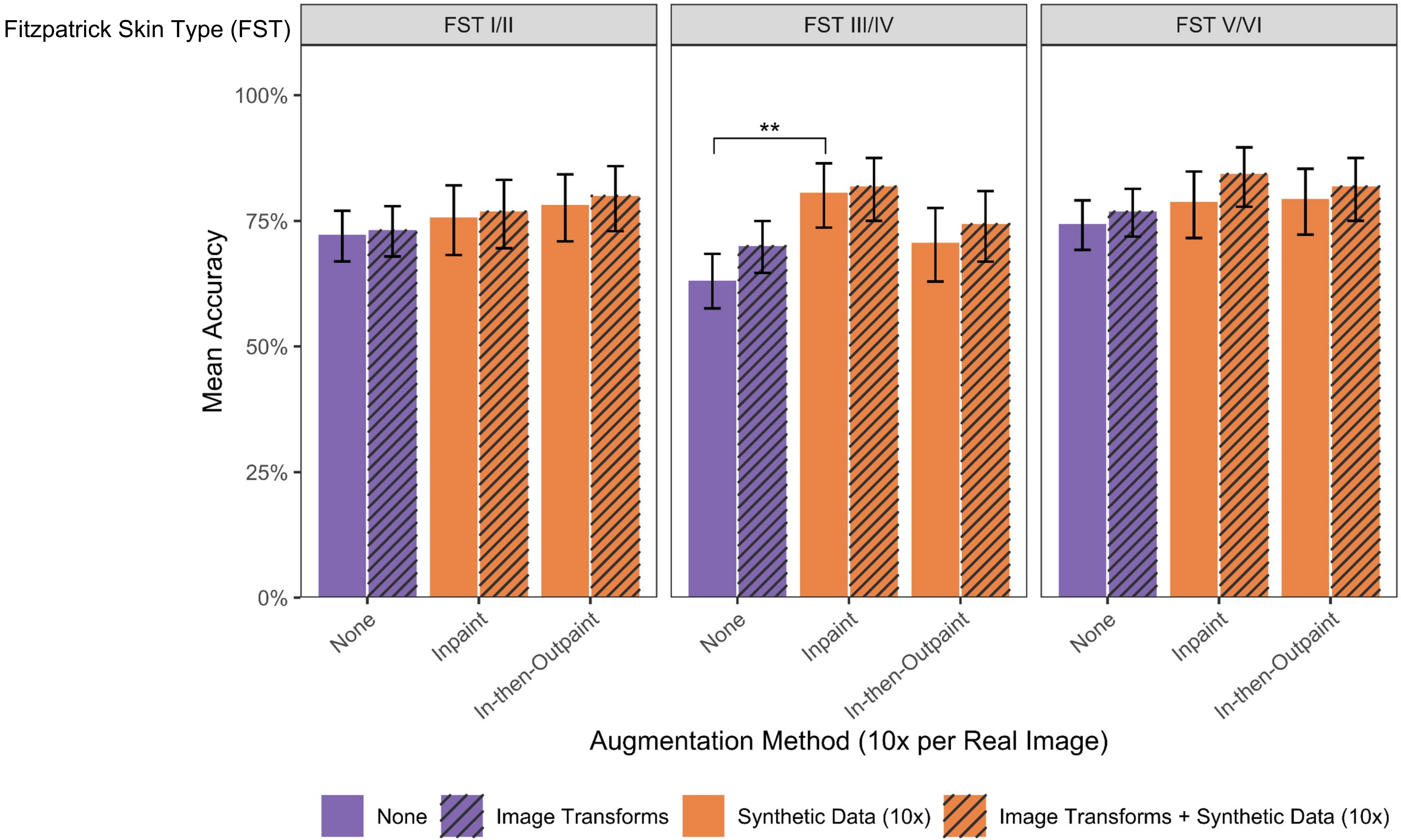


Figure 3: Performance of models trained for nine-way skin classification while varying the number of real images (1 to 228), use of synthetic images (0 or 10 synthetic images per real image), and use of an image transform augmentation utility (yes or no). Model accuracy is averaged across five runs on a balanced, held-out test set of 360 images. Synthetic augmentation was performed using one of three methods: inpaint only, in-then-outpaint, and text-to-image. Error bars

https://arxiv.org/abs/2308.12453



Saturation point


Fitzpatrick Skin Type (FST)
FST I/II
FST III/IV
FST V/VI
100%
75%
50%
25%
0%
Mean Accuracy
**
None
Inpaint
In-then-Outpaint
Augmentation Method (10x per Real Image)
None
Image Transforms
Synthetic Data (10x)
Image Transforms + Synthetic Data (10x)

Limitations of this paper

- The study focused specifically on **skin disease classifiers** and may not be directly applicable to other medical image classifiers or AI algorithms in different domains.
- The **performance gains** achieved by augmenting model training with **synthetic** data were **smaller** compared to **adding real** images.
- The analysis was limited to a **specific set of generative methods** and skin conditions, and the results may not generalize to other generative models or disease categories.
- The paper acknowledged that synthetic data could contain **hallucinations**, highlighting the need for careful evaluation and validation of the generated

Mahidol University
Wisdom of the Land

Thank You

07
AFFORDABLE
AND CLEAN ENERGY