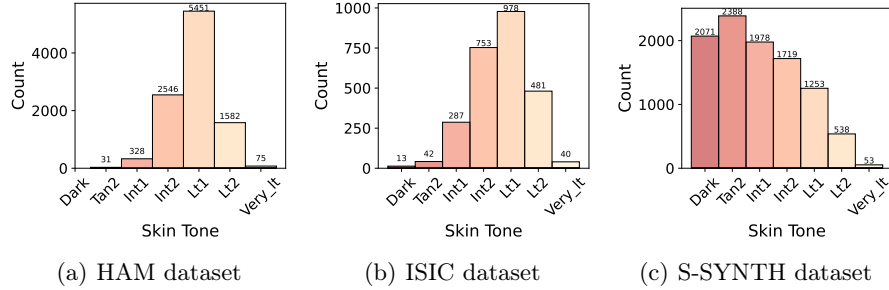# 1 Supplementary Materials

(a) HAM dataset (b) ISIC dataset (c) S-SYNTH dataset

Fig. 1. Distribution of skin tone category based on the mean ITA score of the non-diseased region for (a) HAM dataset with 10015 images, (b) ISIC dataset with 2594 images, and (c) S-SYNTH dataset with 10000 images. The categories are estimated based on [1], however, the examples categorized as "dark" and "tan1" are combined due to the limited number of examples in each group for the real-patient datasets.

Growing step	$S \sim U(1,2)$
Probabilities update	$G(\mu = 0, \sigma = 0.5)$
Cancer probability	$C_p = 0.0001$
Cancer iterations	$C_i = 10$
Maximum cancer recursion	$C_r = 2$

Table 1. Assigned values and probabilities to the skin growing lesion parameters.
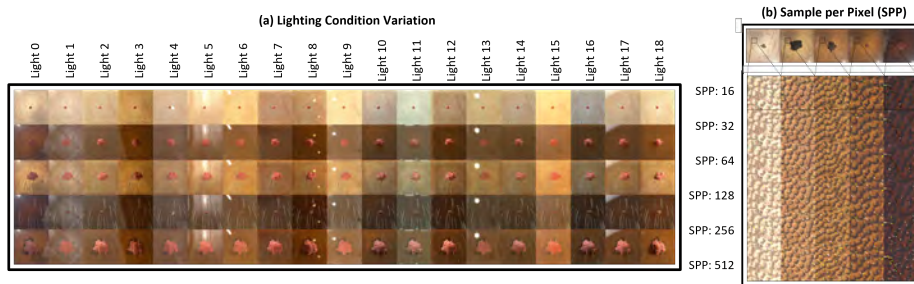

Fig. 2. Examples of S-SYNTH images generated with variations of (a) lighting conditions, (b) samples per pixel (SPP).
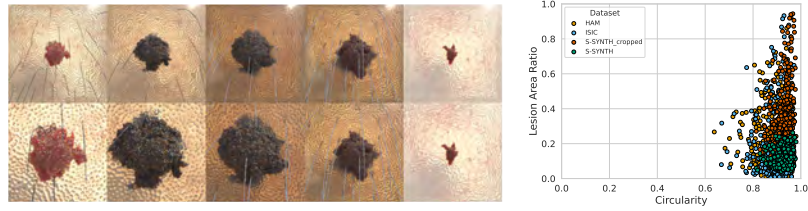
(a) Examples of uncropped (top row) and cropped S-SYNTH images (bottom row)

(b) Distribution of lesion relative area and lesion circularity for the real and synthetic images

Fig. 3. (a) Visualization of S-SYNTH images before and after random crop of 0-60% of the original image size around the center (b) Distribution of lesion relative area and lesion circularity for a subset of HAM (500 images), ISIC (259 images), S-SYNTH before cropping (500 images), and S-SYNTH after cropping (500 images).
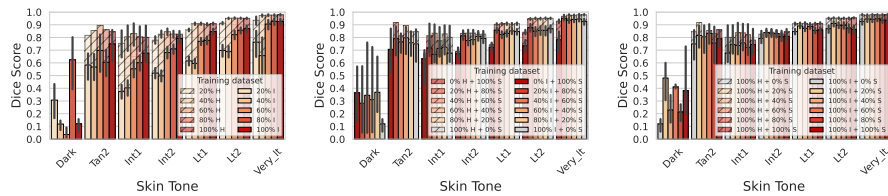
(a) Different proportions of real data

(b) Real data replaced with synthetic data

(c) Synthetic data added to real data

Fig. 4. Changes of model performance when the training data is composed of (a) different subsets of the real images, (b) different subsets of synthetic images substituted for real images (c) different subsets of synthetic images added to real images. The dice scores are stratified based on the skin tone of the test dataset. H: HAM, I: ISIC, S: S-SYNTH.

References

1. Newton M Kinyanjui, Timothy Odonga, Celia Cintas, Noel CF Codella, Rameswar Panda, Prasanna Sattigeri, and Kush R Varshney. Estimating skin tone and effects on classification performance in dermatology datasets. *arXiv:1910.13268*, 2019.