Data augmentation using Stable Diffusion

Haider Ali

Stable Diffusion model

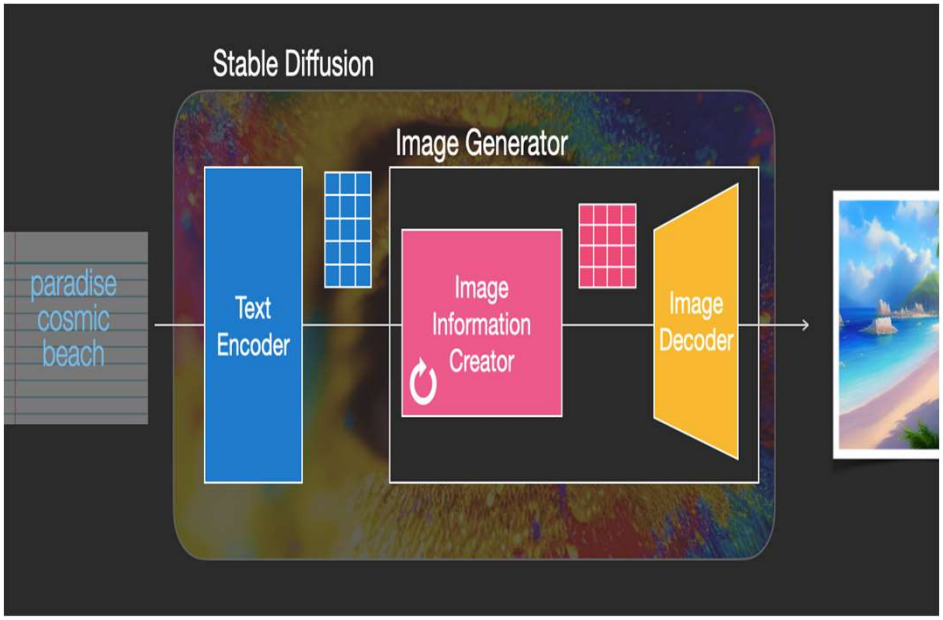


Image1: Components of a Stable Diffusion model

Credit: https://jalammar.github.io/illustrated-stable-diffusion/

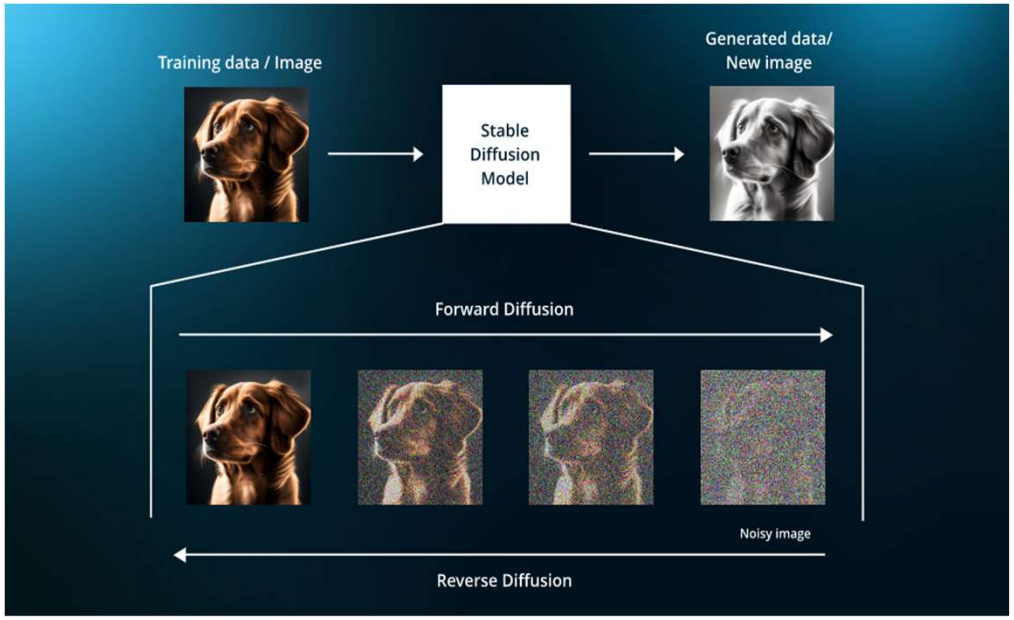


Image 2: Steps of the stable Diffusion model

Credit: https://www.leewayhertz.com/how-to-train-a-diffusion-model/

Dreembooth's Stable Diffusion

DreamBooth

Name/Path of the initial model. (Find model name here)

MODEL_NAME: runwayml/stable-diffusion-v1-5

BRANCH: fp16

Enter instance prompt and class prompt.
Example 1: photo of zwx person, photo of a person
Example 2: photo of zwx toy, photo of a toy

instance_prompt: photo of cow teat

class_prompt: photo of a teat

training_steps: 800

learning_rate: 1e-6

Test image generation from model

prompt: random photo of cow teat

negative_prompt: Insert text here

num_samples: 4

guidance_scale: 8

num_inference_steps: 30

height: 512

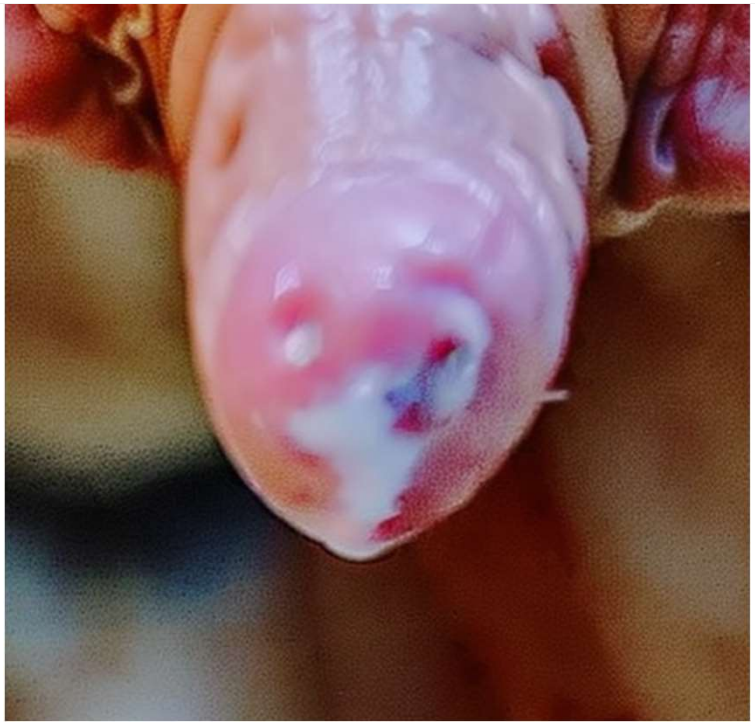
width: 512

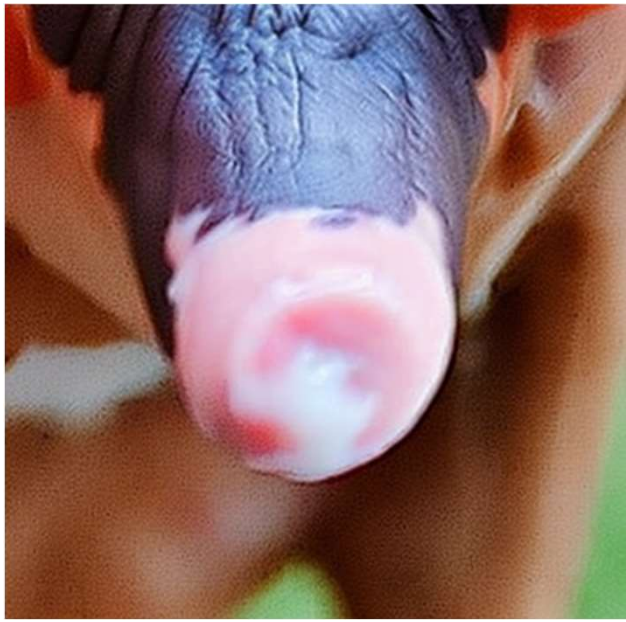
seed: 100

Show code

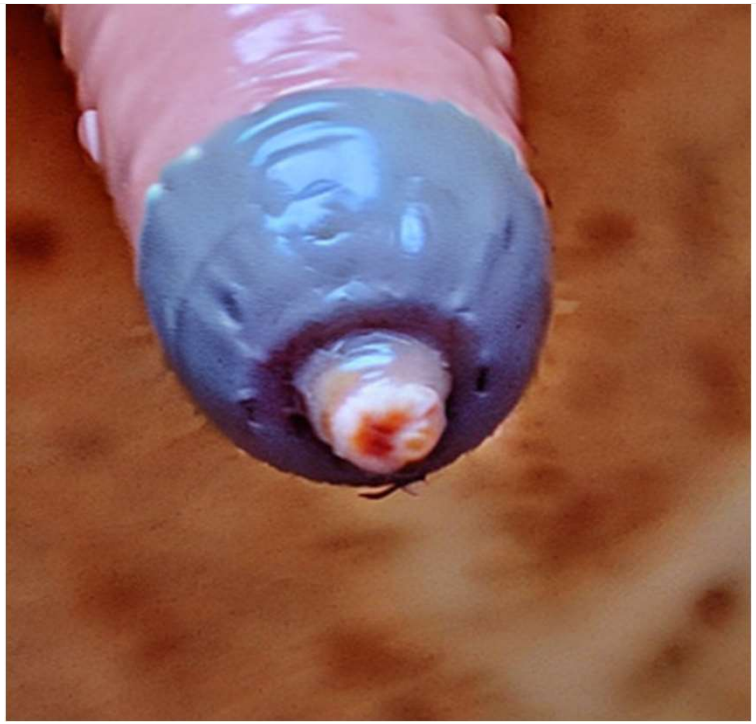
Dreembooth model: https://stable-diffusion-art.com/dreambooth/

Score 1 (hypokeratosis) generated images



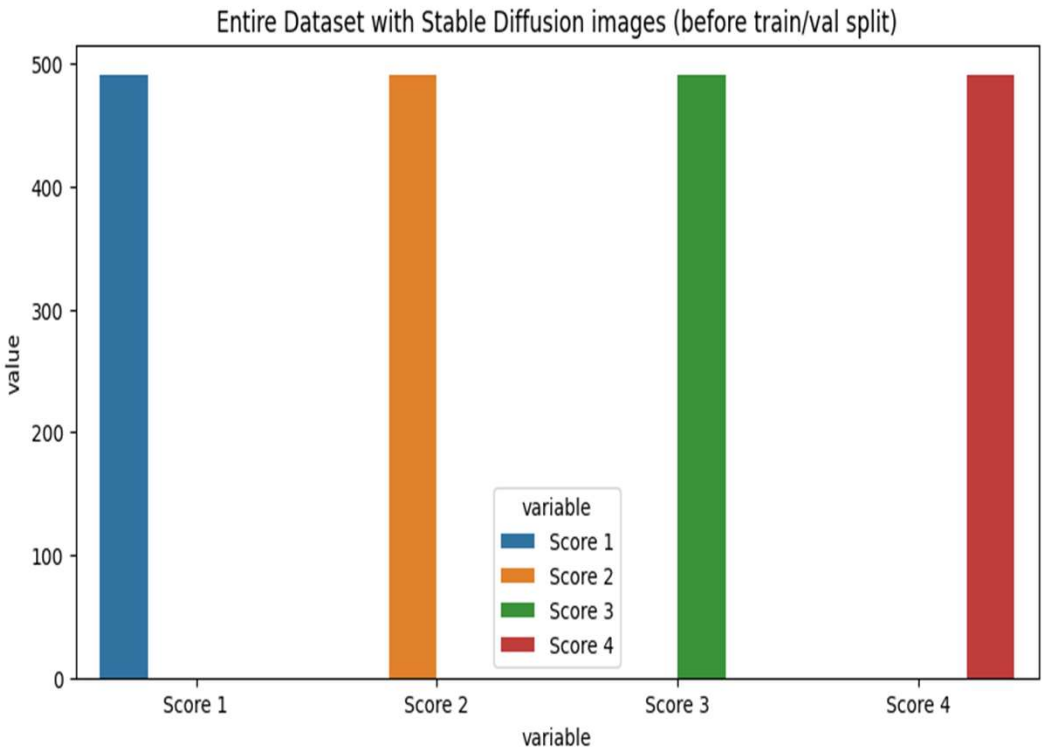


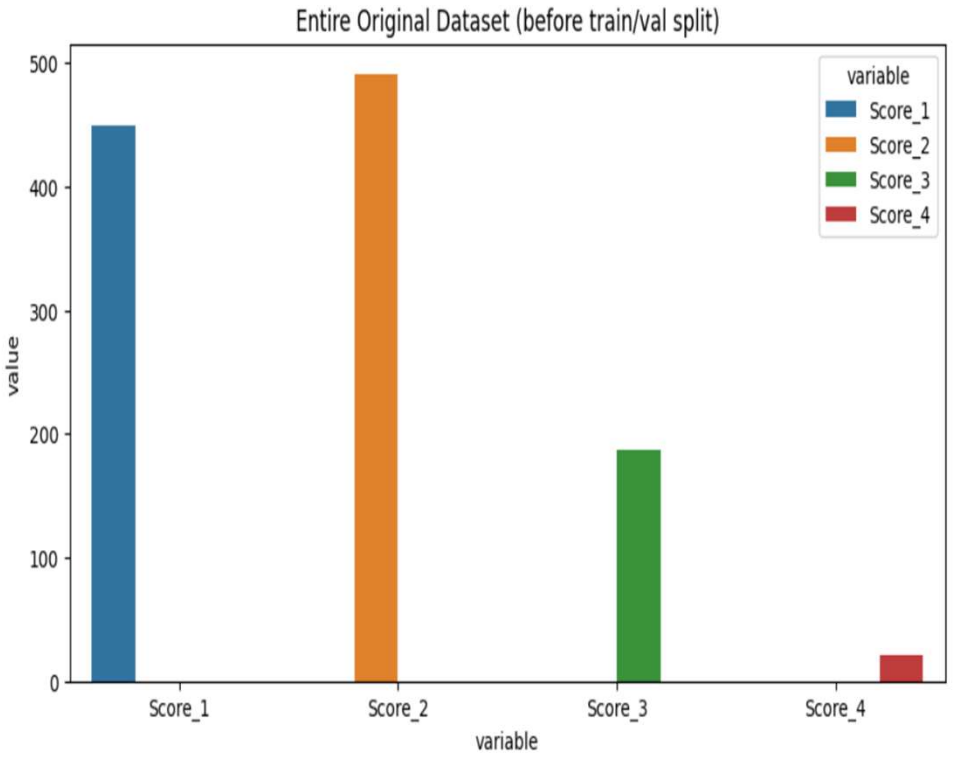
Score 3 generated images



Score 4 (Hyperkeratosis) generated images

Comparison of images according to classes





Dataset and Hyper-parameters

- Loss functions: Used CrossEntropy loss for augmented dataset and Focal loss for original dataset
- Learning rate: 0.001, 0.007, 0.005, 0.01.
- Used both Adam and SGD.

Augmented dataset split:

Training images: 1571

Validation images: 393

Original dataset split:

Training images: 919

Validation images: 393

Model comparison

Model	Valid Accuracy	Test Accuracy
ResNet 50	65.74%	63.42%
ResNet 50 SD	81.77%	63.42%
Custom model	66.66%	61.05%
Custom model SD	83.59%	61.05%
ResNet 18	64.35%	61.05%
ResNet18 SD	82.03%	60.78%
VIT SD	71.35%	60.52%
VGG 19	49.07%	59.47%
Inception	64.35%	59.47%
VIT	64.35%	59.21%
Inception SD	75.26%	48.94%
VGG 19 SD	70.83%	48.42%

Conclusion and future work

- There is not much a difference between the model's accuracies of original and generated datasets.
- It can be concluded that adding extra images or balancing the dataset won't improve the model's accuracy.
- Advanced techniques can be used such as Separable Confident Transductive Learning to increase achieve better accuracy.

Thanks

Accuracy is not solely dependent on the amount of data, but on the quality, relevance, and understanding of the data at hand.

Questions?