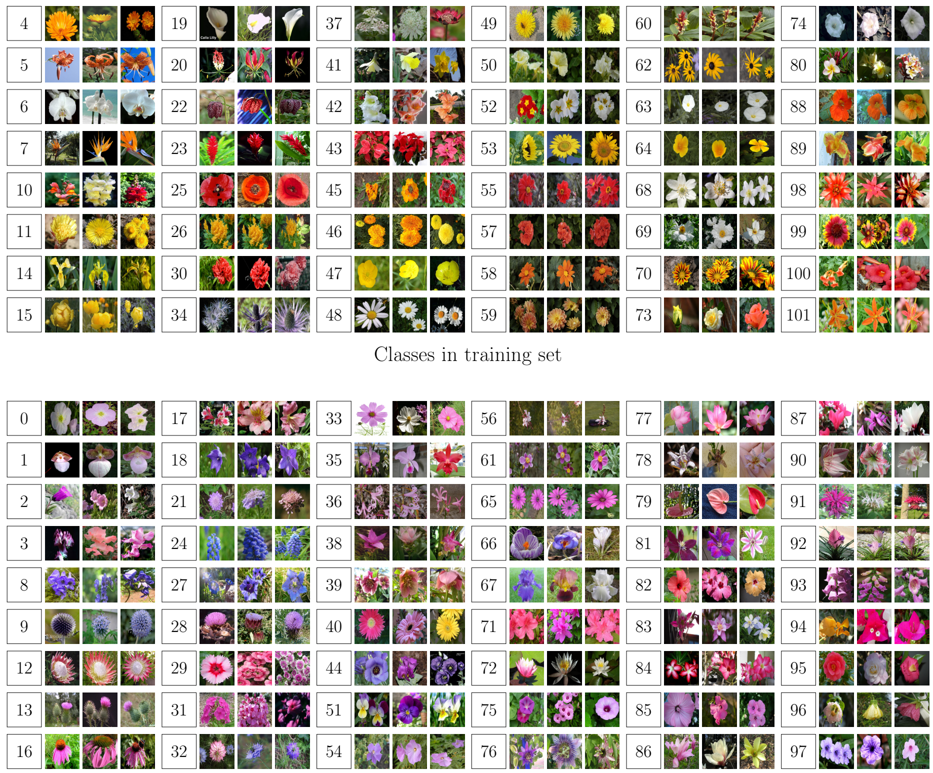
A Flower dataset split



Classes excluded from training set

Fig. 6: The class splits of the flower dataset, visualized with the class number and three examples for each class.

In Fig. 6, we illustrate the split of the flower dataset, featuring classes included in the training data at the top and manually excluded classes (only used during generation) at the bottom. For each class, three example images are provided to convey an impression of the classes. As detailed in Sec. 4.1, we manually excluded classes primarily showcasing the colors blue, purple, and pink, based on visual inspection of ten examples per class. This split aims to assess our model's ability to adapt to unseen colors. It is worth noting that some flower types may exhibit a variety of colors, as seen in class 40. Therefore, there is a possibility that some flowers with the colors blue, purple, and pink are included in the training set. However, such cases would be significantly underrepresented, allowing us to still evaluate the benefits of our method.

B Classifier-free guidance

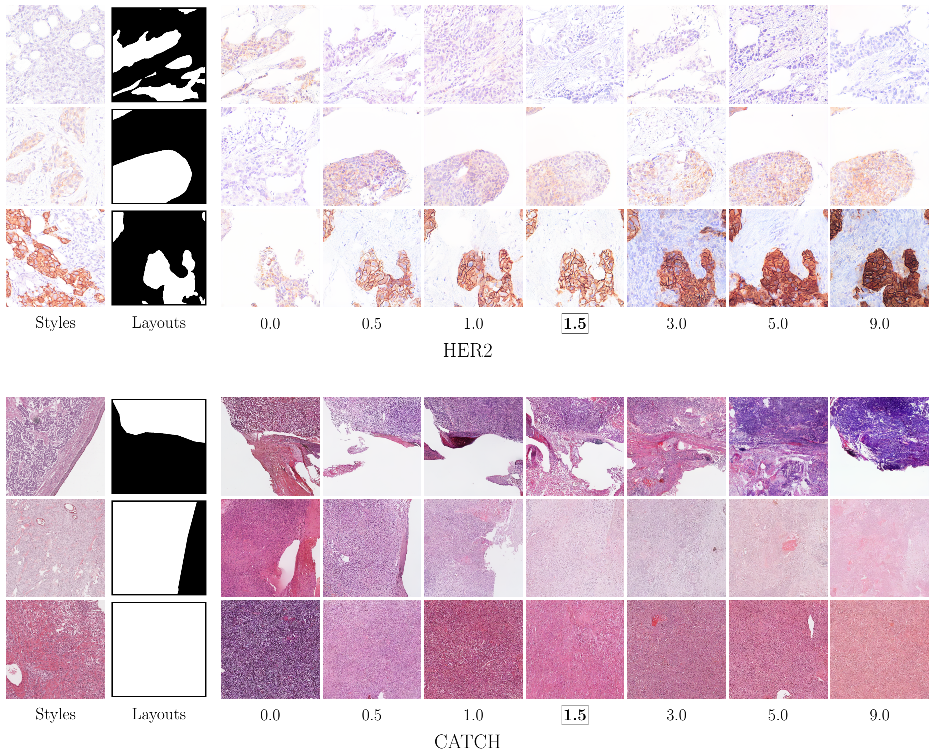


Fig. 7: Image generation results for the histopathological dataset with nearby style sampling and classifier-free guidance scales of 0.0, 0.5, 1.0, 1.5, 3.0, 5.0 and 9.0. For our work, we chose a classifier-free guidance scale of 1.5.

In the training phase and during the image generation process, we employed classifier-free guidance for the style query images. This approach proves beneficial as the model learns the style-unconditional distribution of the training data for the omitted style query images, and it also extracts valuable style information when style query images are provided.

For image generation where the requested style lies outside the training style distribution, we utilized classifier-free guidance to compel the model to produce samples beyond the training style distribution. The classifier-free guidance scale determines how far we push the reconstructed image away from the learned style distribution.

Examples of generated images for the histopathological datasets under different classifier-free guidance scales are presented in Fig. 7. Lower classifier-free guidance scales yield less style-accurate generated images, while higher scales sometimes result in oversaturated images. Based on visual assessment of the generated images, we selected a classifier-free guidance scale of 1.5 for our experiments, demonstrating accurate styles without oversaturation.

C Generations with missing style information

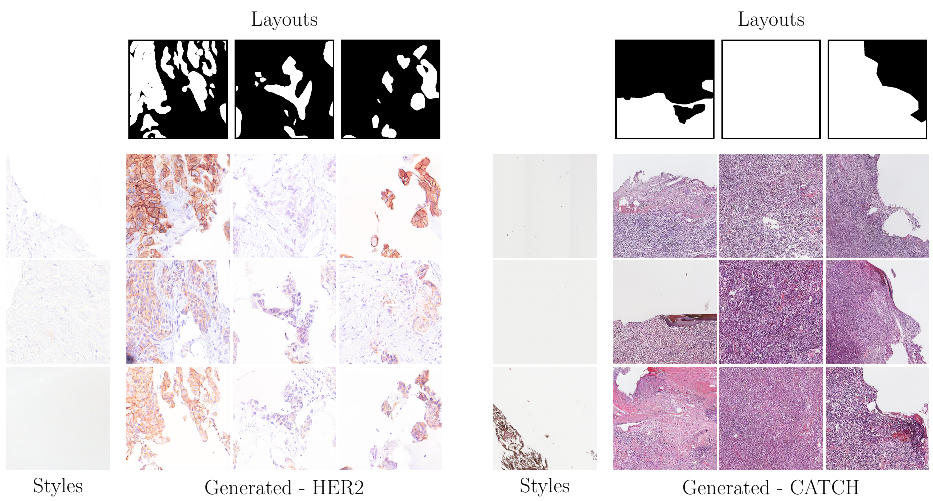


Fig. 8: Image generation results for the histopathological dataset with nearby style sampling. Shown are cases where tumor is present in the layout (white: tumor, black: non-tumor tissue and slide background), but no tumor tissue is present in the style images.

In Fig. 8, we provide examples of image generation for the histopathological datasets in cases where style information is missing in the style query images. The generated images exhibit realistic tissues, but the style is determined by the model and reflects styles from the training style distribution.

This scenario primarily arises in nearby style sampling, where only a single style query image is used. We see these cases as noncritical as they do not result in invalid images, and they highlight that the model falls back to plausible styles if it cannot extract style information from the style query images.

To ensure that the model respects the style information of valid style query images and does not recreate known styles, we incorporate classifier-free guidance, as discussed in Appendix B.

D Flower segmentation results

The segmentation results for the flower dataset are presented in Tab. 3. Across all setups, the optimal results were attained when training without synthetic

	Synthetic data	Mean IoU	IoU Variance	Mean IoU	IoU Variance	Mean IoU	IoU Variance
		960 Images		480 Images		144 Images	
Flowers	None	**87.80 (0.06)**	4.20 (0.12)	**87.79 (0.26)**	**4.14 (0.06)**	**87.05 (0.18)**	**4.41 (0.17)**
	Style Transfer	86.67 (0.30)	4.29 (0.17)				
	Semantic DM	87.26 (0.17)	**4.18 (0.07)**	87.41 (0.09)	4.21 (0.15)	85.91 (0.17)	5.23 (0.25)
	Augmented (ours)	86.49 (0.40)	4.48 (0.29)	86.26 (0.17)	4.32 (0.08)	85.83 (0.21)	4.86 (0.33)

Table 3: Segmentation results for the flower dataset, with different amounts of training data and synthetic images.

data, with mean IoU scores consistently exceeding 87. The introduction of synthetic images into the training data did not yield improvements in mean IoU scores, although all reported scores remained at high levels, with none dropping below 85. No clear trend in IoU variance between images was evident across the experiments.

We argue that the lack of benefit from synthetic data in the flower dataset is attributable to the task's simplicity, as evidenced by the high IoU scores even at lower/lowest data settings. The ImageNet-pretrained encoder of our segmentation UNet appears capable of adapting to the segmentation task without necessitating the additional information provided by synthetic images. Additionally, we argue that the diffusion models could overfit to the layouts, due to the limited number of training examples and the distinct shapes of some flower types, leading to less diverse generated images. For the histopathological datasets, this problem does not exist, since even for low amounts of data, no connection between images and layouts exists.