

SUPPLEMENTARY MATERIALS

1. USER STUDY DESIGN

We observe that *Stencil* demonstrates superior performance in generating novel layouts and accurately following complex prompts. To validate this claim, we conduct a user study comparing *Stencil* with the previous state-of-the-art, *DreamBooth*. Using the *DreamBench* dataset, we evaluate all live subjects across a set of specified prompts:

1. A [subject token] in a crowded cyberpunk city on a rainy day.
2. A [subject token] with hundreds of lanterns in the background.
3. A [subject token] jumping.
4. A [subject token] sleeping.

We include the full instruction set for our user study below. We conducted the study on 30 participants. The dataset consists of 9 subjects, each tested on 4 complex prompts. Each image is evaluated on subject consistency and text-to-image alignment.

Subject Consistency

1. Inspect the subject of the reference image. Select which of the images best reproduces the identity of the reference subject.
2. If you are unsure, or believe that the images equally resemble the reference subject, select 'Undecided'.

Text-to-Image Alignment

1. Select which of the images best follows the prompt [target prompt].
2. If you are unsure, or believe that the images equally follow the prompt, select 'Undecided'.

2. ADDITIONAL APPLICATIONS

In this section, we share various other applications of *Stencil* that were not covered in our main paper.

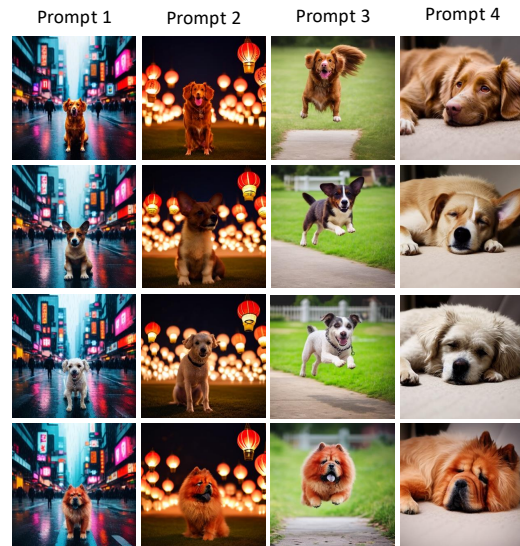


Fig. 1: Additional Qualitative Results. *Stencil* effectively adheres to complex prompts while maintaining layout consistency when using the same template image during context guidance

Age Progression/Regression (Fig. 2) Fine-tuning on images of the subject’s younger/older self allows us to generate age-regressed/age-progressed images of subject by reducing τ .

Subject Expression Manipulation (Fig. 3) *Stencil* can generate diverse unseen expressions of the subject, by following the prompt “A [emotion] [subject token]”.

Accessorization (Fig. 4) We can accessorize the subject by using the prompt “A [subject token] wearing [accessory].”

Perspective-conditioned Generation (Fig. 5) We can generate diverse images of the subject in different points-of-view, previously unseen in the reference images, using the prompt “A [subject token] seen from [angle].”

Pose Editing (Fig. 6) *Stencil* enables the generation of diverse poses by prompting “A [subject token] [pose].” We highlight that Dreambooth cannot achieve this to the same effect.

Style Transfer (Fig. 7) *Stencil* can also transfer the subject to various artistic styles by using the prompt “A [subject token] in [artistic style].” This approach preserves the subject’s core identity while adapting to the new style.

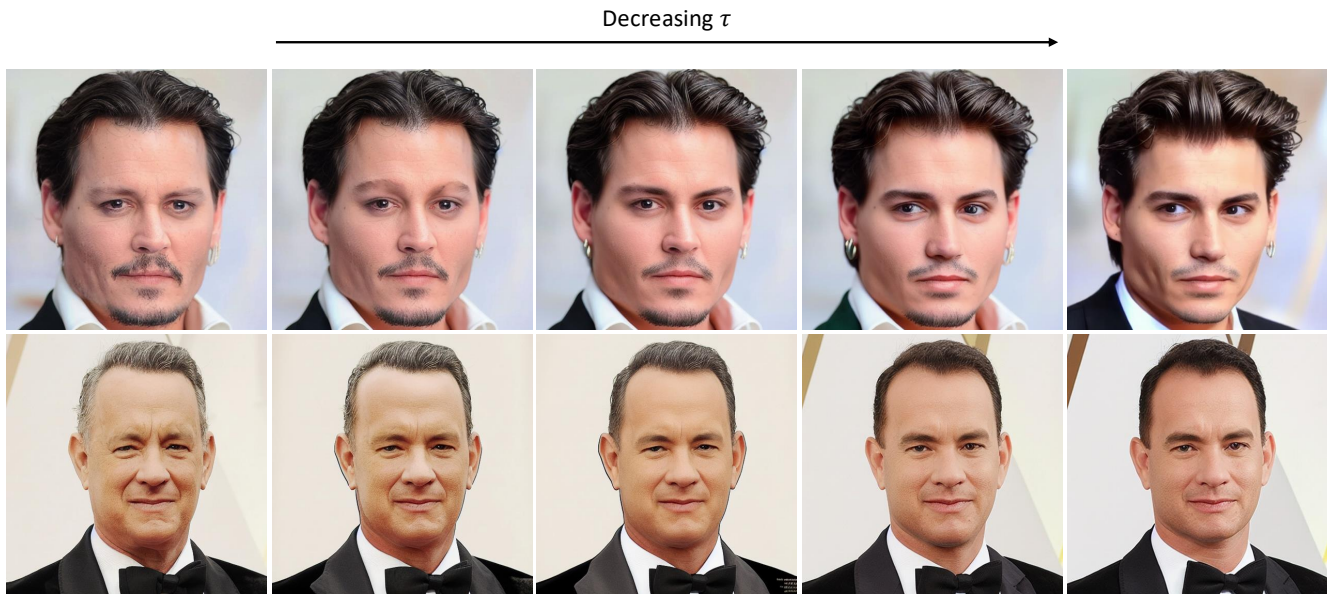


Fig. 2: Age Progression/Regression. We remark that manipulating τ does not distort the appearance of the background and foreground elements as much as it does to the face of the subject. We also note that the generated younger self shares a remarkable resemblance to what the subjects had looked like back then.

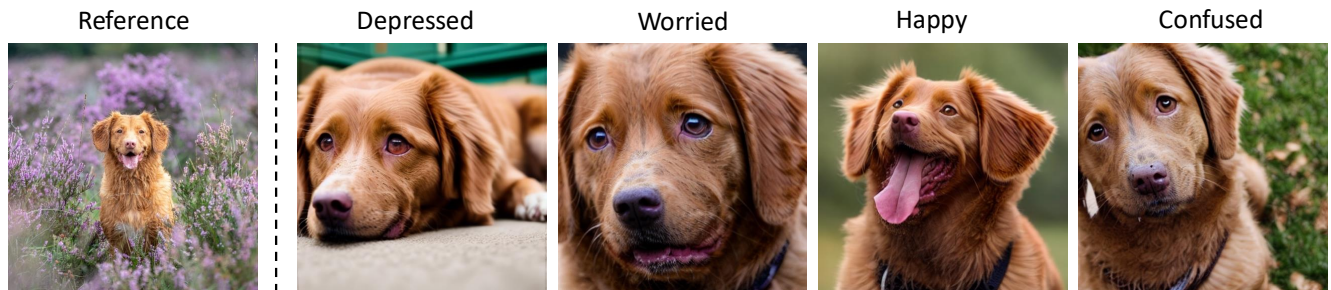


Fig. 3: Expression Manipulation. Stencil supports the generation of a diverse range of expressions of the subject while maintaining high subject fidelity.



Fig. 4: Accessorization. We remark that the generated layouts of the subject are unique from those of the reference image but the subject remains consistent with the reference.

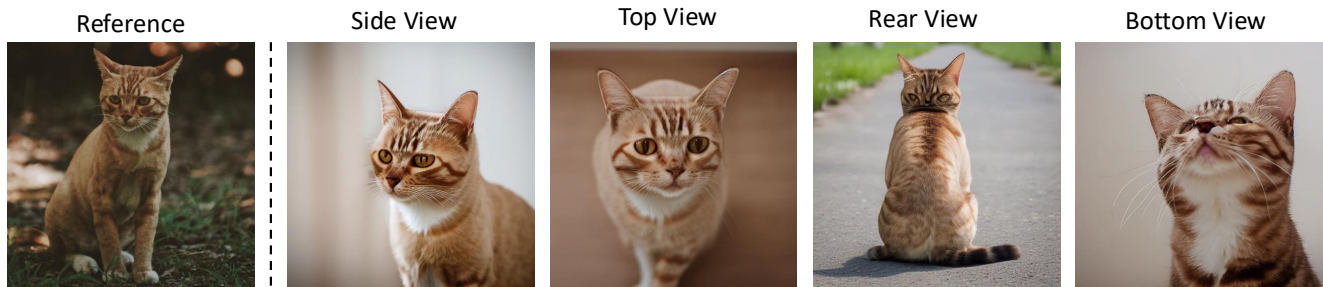


Fig. 5: Perspective-conditioned Generation. We remark that no images from the reference set was taken from the rear or bottom view. This highlights Stencil’s incredible generalization capabilities.

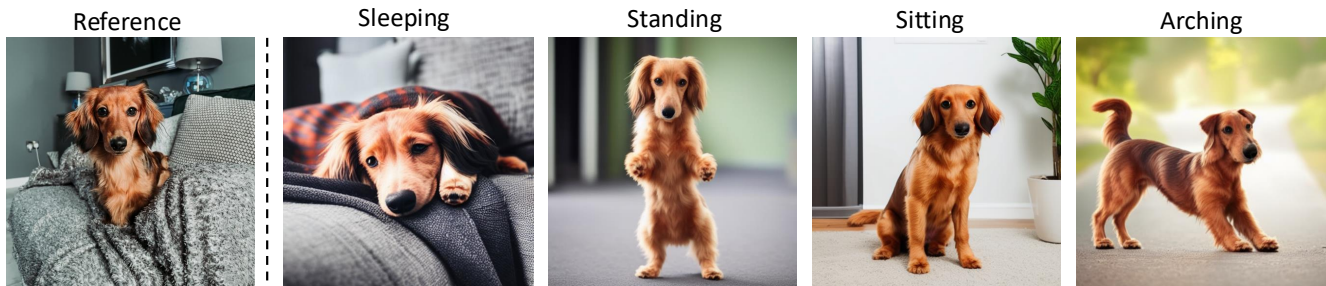


Fig. 6: Pose Editing. Stencil can generate diverse unseen poses of the subject that is beyond the generation capabilities of the fine-tuned model. This highlights the ability of the stencil generator to introduce new priors to the fine-tuned model.

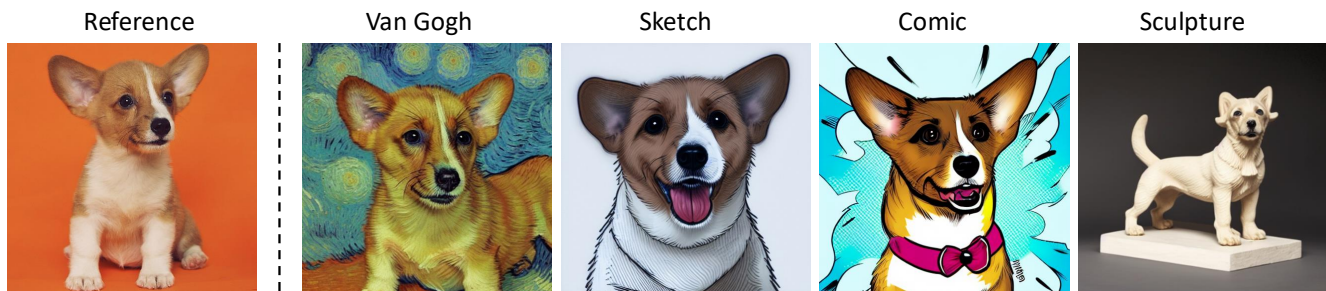


Fig. 7: Style Transfer. Stencil enables the seamless transfer of the subject to various artistic mediums, such as paintings and sculptures while maintaining key visual characteristics.

“...in the jungle”



“...in the snow”



“...on the beach”



“...on a cobblesstone street”



“... on top of green grass with sunflowers around it”



Fig. 8: DreamBench Qualitative Results Part 1.

“... on a dirt road”

“...on top of the sidewalk in a crowded street”

“...with a wheat field in the background”

“...with a mountain in the background”

“...on top of a purple rug in the forest”



Fig. 9: DreamBench Qualitative Results Part 2.