Apartment design

Arseny Ivanov
Anna Borisiuk
Aysel Mirzoeva
Nikita Boldyrev
Pham Quoc Viet

Problem

It is often difficult to come up with a design for an empty room

- We help visualize the interior
- Optimize furniture layouts and spatial arrangements to maximize the usability of the space.
- Reduce costs and time spent for design

Baseline model

Generates images using Stable Diffusion with ControNet

Key components

- Reading config.json to create ControlNetModel
- Initializing ‘StableDiffusionControlNetPipeline’ with pre-trained model
- Preprocessing input (resize, convert to tensor)
- Generating new image using pipeline with input image and text prompt
- Saving and displaying output image

```
def generate_design(pipeline, input_image, prompt):
    """
    Pipeline for image generation.
    """
    preprocess = transforms.Compose([
        transforms.Resize((512, 512)),
        transforms.ToTensor(),
    ])
    input_image_tensor = preprocess(input_image).unsq

    input_image_resized = transforms.Resize((512,
    input_image_resized.save("input_image.jpg")
    print("Output image saved")

    print(f"Input dtype: {input_image_tens                devi
    print(f"Pipeline UNet dtype: {pipeli                ype}")
    print(f"Pipeline device: {pipeline

    output = pipeline(prompt, image                ge_tensor)
    output_image = output.images[

    output_image.save("output                )
    print("Output image save                c_image.jpg")

    fig, axs = plt.subpl                igsize=(20,
    axs[0].imshow(input                ized)
    axs[0].set_title(                ge")
    axs[0].axis("of
    axs[1].imshow                ge)
    axs[1].set_t                ut image")
    axs[1].axi
    plt.show
```

Baseline results

Входное изображение

Выходное изображение

Final model

Produces good and valid results

Key components

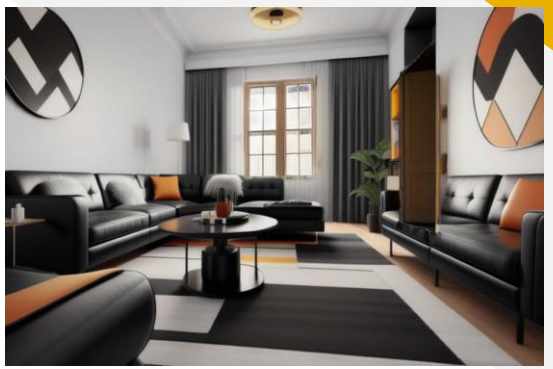
- Semantic segmentation
- Depth estimation
- Masking and preprocessing
- ControlNet Model
- Image synthesis
- Postprocessing

```
class ControlNetMultiPipeline:
    def __init__(self):
        os.environ['HF_HUB_OFFLINE'] = "True"

        depth_model = ControlNetModel.from_pretrained("models/controlnet_d
        segment_model = ControlNetModel.from_pretrained("models/own_contro

        self.pipeline = StableDiffusionControlNetInpaintPipeline.from_pret
            "models/Realistic_Vision_V5.1_noVAE",
            controlnet=[depth_model, segment_model],
            safety_checker=None,
            torch_dtype=torch.float16
        )
        self.pipeline.load_ip_adapter("models/models--h94--IP-
        self.pipeline.set_ip_adapter_scale(0.4)
        self.pipeline.scheduler = UniPCMultistepScheduler
        self.pipeline = self.pipeline.to(device)

        self.guide_pipeline = StableDiffusionXLPipe
            "models/models--segmind--SSD-1B", tor
        ).to(device)

        self.seed = 323 * 111
        self.negative_prompt = "window,
        self.excluded_items = ["window
        self.quality_suffix = "inter
        self.random_gen = torch.Ge

        self.seg_processor, s
        self.depth_extracto
        self.depth_model

    def generate_desi
```

Models

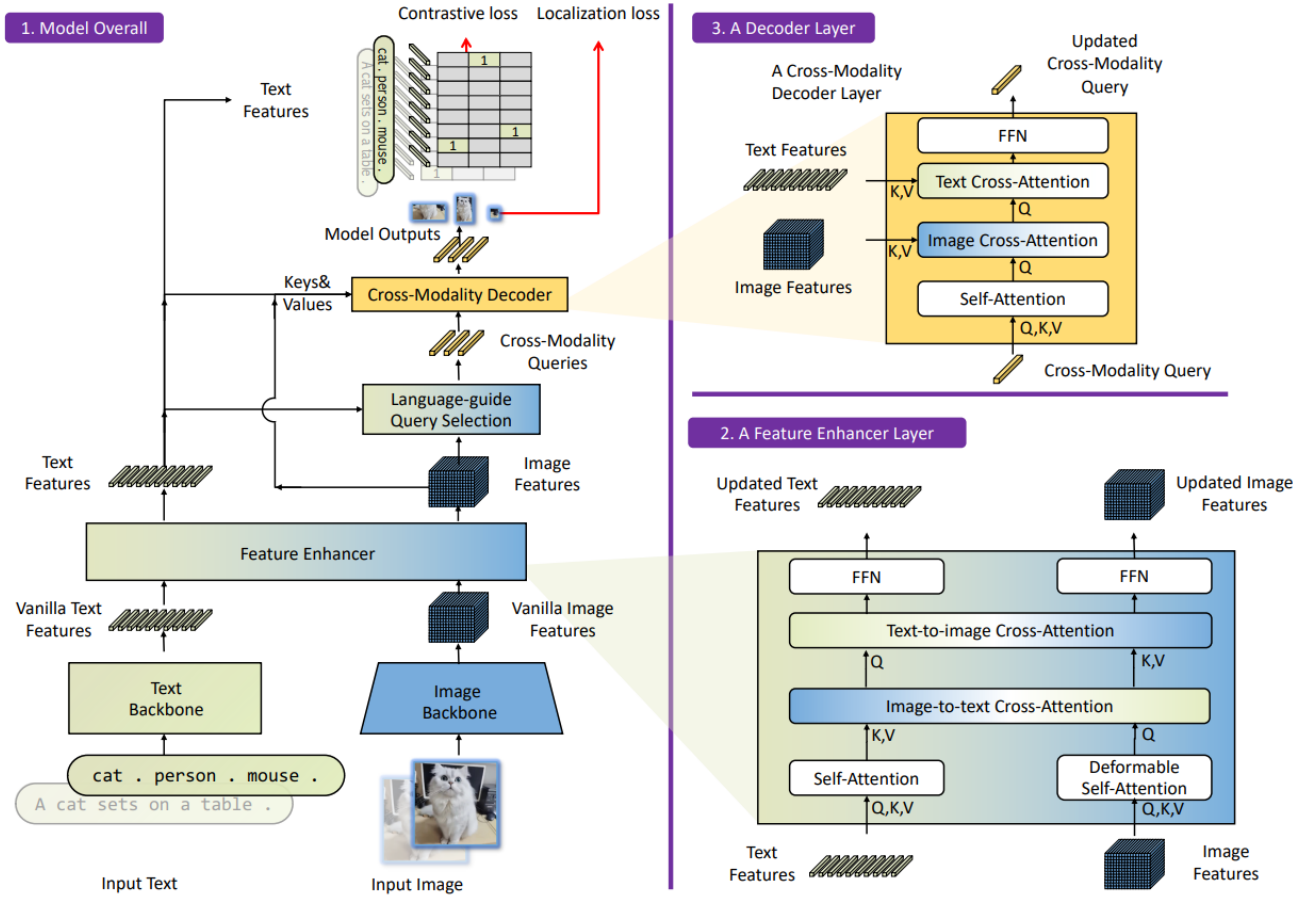
Model name	Role	Output
Realistic_Vision_V5.1_noVAE	Primary generator	Synthesized image based on the provided inputs
Own_controlnet	Segmentation guidance	Guides the generator to align with labeled semantic regions
Openmmlab—upernet-convnext-small	Semantic segmentation	Segmentation map of input image for pixel-level guidance
Models—segmind—SSD-11B	Quality enhancer	High-quality intermediate image with enhanced detail
Models—LiheYoung—depth—anything-large-hf	Depth estimation	Depth map to define the spatial structure of the scene
Models—h94—IP-Adapter	Image processing adapter	Refined and aligned intermediate features
controlnet_depth	Depth-based conditional guidance	Conditions the generator on depth for 3D realism in synthesized output

Model results

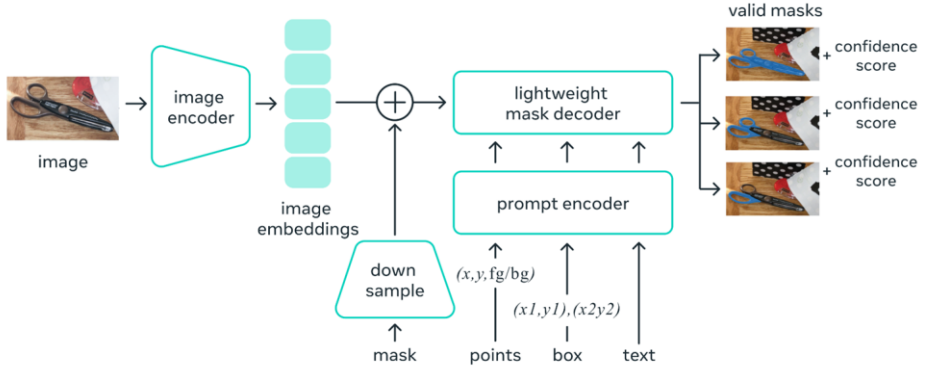


Grounding DINO + SAM + SD-1.5

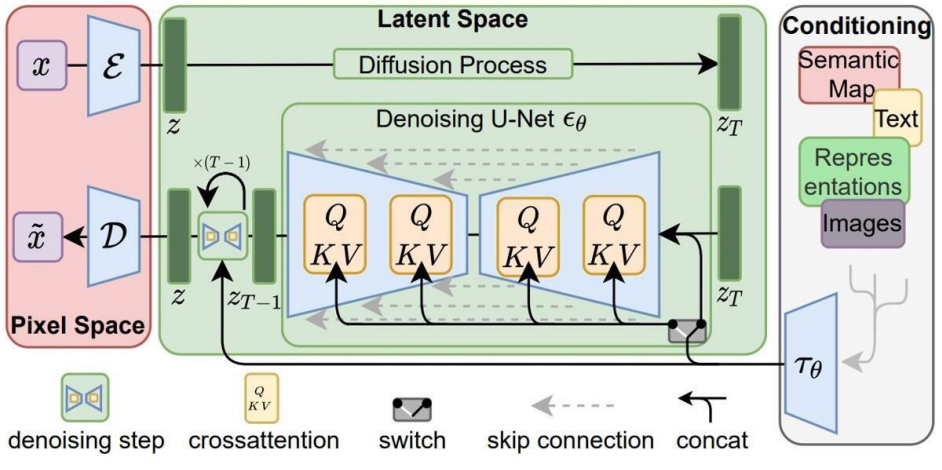
Grounding DINO for Open Vocabulary Object Detection



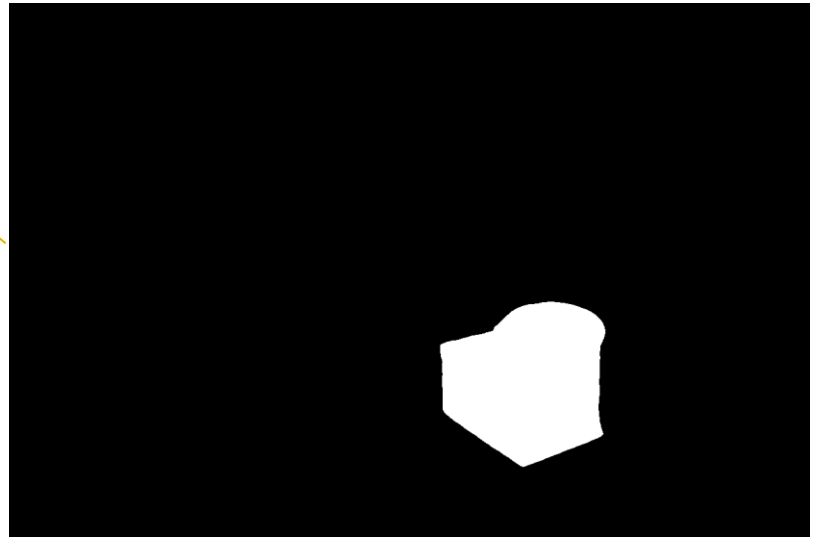
Universal segmentation model



Stable Diffusion for Image Editing



Experiments and Results



Thank you for your attention!