```
import os
import torch
from diffusers import StableDiffusionPipeline

# --------------------------------
# Configuration
# --------------------------------
MODEL_ID = "runwayml/stable-diffusion-v1-5"
DEVICE = "cuda" if torch.cuda.is_available() else "cpu"
BASE_DIR = "chest_xray_dataset"
IMAGES_PER_CLASS = 10

classes = {
    "normal_lungs": "Chest X-ray of healthy lungs, normal anatomy",
    "pneumonia": "Chest X-ray showing bacterial pneumonia",
    "covid_opacities": "Chest X-ray with COVID-19 ground glass opacities",
    "lung_opacity": "Chest X-ray showing diffuse lung opacities",
    "pleural_effusion": "Chest X-ray showing pleural effusion",
    "pneumothorax": "Chest X-ray showing pneumothorax",
    "lung_nodules": "Chest X-ray showing lung nodules",
    "lung_fibrosis": "Chest X-ray showing pulmonary fibrosis",
    "cardiomegaly": "Chest X-ray showing enlarged heart",
    "medical_devices": "Chest X-ray with tubes and pacemaker",
    "imaging_artifacts": "Chest X-ray with motion blur and noise",
    "pa_view": "Chest X-ray PA view",
    "ap_view": "Chest X-ray AP view",
    "domain_shift": "Chest X-ray from different hospital scanner"
}

# --------------------------------
# Load Model
# --------------------------------
pipe = StableDiffusionPipeline.from_pretrained(
    MODEL_ID,
    torch_dtype=torch.float16 if DEVICE == "cuda" else torch.float32
)
pipe.to(DEVICE)

# --------------------------------
# Generate Dataset
# --------------------------------
os.makedirs(BASE_DIR, exist_ok=True)

for label, prompt in classes.items():
    class_dir = os.path.join(BASE_DIR, label)
    os.makedirs(class_dir, exist_ok=True)
```

```
        for i in range(IMAGES_PER_CLASS):
            print(f"Generating {label} image {i+1}")
            image = pipe(f"High resolution chest X-ray image. {prompt}").images[0]
            image.save(os.path.join(class_dir, f"{label}_{i+1}.png"))

    print("✅ Dataset generation completed")
```

```
Flax classes are deprecated and will be removed in Diffusers v1.0.0. We recommend migrating to PyTorch classes or pinning your version of Diffuser
Flax classes are deprecated and will be removed in Diffusers v1.0.0. We recommend migrating to PyTorch classes or pinning your version of Diffuser
/usr/local/lib/python3.12/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as secret in your
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
  warnings.warn(
model_index.json: 100%                                                  541/541 [00:00<00:00, 35.0kB/s]
Fetching 15 files: 100%                                                 15/15 [02:16<00:00, 20.12s/it]
scheduler_config.json: 100%                                             308/308 [00:00<00:00, 3.43kB/s]
config.json: 100%                                                       617/617 [00:00<00:00, 6.33kB/s]
preprocessor_config.json: 100%                                          342/342 [00:00<00:00, 4.02kB/s]
config.json:         4.72k/? [00:00<00:00, 89.9kB/s]
special_tokens_map.json: 100%                                           472/472 [00:00<00:00, 8.32kB/s]
merges.txt:         525k/? [00:00<00:00, 7.16MB/s]
safety_checker/model.safetensors: 100%                                  1.22G/1.22G [01:19<00:00, 7.92MB/s]
text_encoder/model.safetensors: 100%                                    492M/492M [02:16<00:00, 3.46MB/s]
tokenizer_config.json: 100%                                             806/806 [00:00<00:00, 27.4kB/s]
config.json: 100%                                                       743/743 [00:00<00:00, 16.9kB/s]
config.json: 100%                                                       547/547 [00:00<00:00, 13.2kB/s]
vocab.json:         1.06M/? [00:00<00:00, 12.3MB/s]
unet/diffusion_pytorch_model.safetensors: 100%                          3.44G/3.44G [02:15<00:00, 21.7MB/s]
vae/diffusion_pytorch_model.safetensors: 100%                           335M/335M [01:27<00:00, 2.23MB/s]
Loading pipeline components...: 100%                                    7/7 [00:21<00:00,  5.27s/it]
`torch_dtype` is deprecated! Use `dtype` instead!
Generating normal_lungs image 1
100%                                    50/50 [00:08<00:00,  7.41it/s]
Generating normal_lungs image 2
100%                                    50/50 [00:07<00:00,  7.06it/s]
Generating normal_lungs image 3
100%                                    50/50 [00:06<00:00,  7.25it/s]
Generating normal_lungs image 4
100%                                    50/50 [00:07<00:00,  7.22it/s]
Generating normal_lungs image 5
```

```
100%                                      50/50 [00:07<00:00,  7.16it/s]
Generating normal_lungs image 6
100%                                      50/50 [00:07<00:00,  7.06it/s]
Generating normal_lungs image 7
100%                                      50/50 [00:07<00:00,  7.05it/s]
Generating normal_lungs image 8
100%                                      50/50 [00:07<00:00,  6.98it/s]
Generating normal_lungs image 9
100%                                      50/50 [00:07<00:00,  6.90it/s]
Generating normal_lungs image 10
100%                                      50/50 [00:07<00:00,  6.80it/s]
Generating pneumonia image 1
100%                                      50/50 [00:07<00:00,  6.76it/s]
Generating pneumonia image 2
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating pneumonia image 3
100%                                      50/50 [00:07<00:00,  6.52it/s]
Generating pneumonia image 4
100%                                      50/50 [00:07<00:00,  6.37it/s]
Generating pneumonia image 5
100%                                      50/50 [00:08<00:00,  6.37it/s]
Generating pneumonia image 6
100%                                      50/50 [00:07<00:00,  6.51it/s]
Generating pneumonia image 7
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating pneumonia image 8
100%                                      50/50 [00:07<00:00,  6.66it/s]
Generating pneumonia image 9
100%                                      50/50 [00:07<00:00,  6.69it/s]
Generating pneumonia image 10
100%                                      50/50 [00:07<00:00,  6.69it/s]
Generating covid_opacities image 1
100%                                      50/50 [00:07<00:00,  6.01it/s]
Generating covid_opacities image 2
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating covid_opacities image 3
100%                                      50/50 [00:07<00:00,  6.57it/s]
Generating covid_opacities image 4
```

```
100%                                                50/50 [00:07<00:00,  6.52it/s]
Generating covid_opacities image 5
100%                                                50/50 [00:07<00:00,  6.55it/s]
Generating covid_opacities image 6
100%                                                50/50 [00:07<00:00,  6.57it/s]
Generating covid_opacities image 7
100%                                                50/50 [00:07<00:00,  6.54it/s]
Generating covid_opacities image 8
100%                                                50/50 [00:07<00:00,  6.56it/s]
Generating covid_opacities image 9
100%                                                50/50 [00:07<00:00,  6.63it/s]
Generating covid_opacities image 10
100%                                                50/50 [00:07<00:00,  6.62it/s]
Generating lung_opacity image 1
100%                                                50/50 [00:07<00:00,  6.60it/s]
Generating lung_opacity image 2
100%                                                50/50 [00:07<00:00,  6.64it/s]
Generating lung_opacity image 3
100%                                                50/50 [00:07<00:00,  6.62it/s]
Generating lung_opacity image 4
100%                                                50/50 [00:07<00:00,  6.63it/s]
Generating lung_opacity image 5
100%                                                50/50 [00:07<00:00,  6.60it/s]
Generating lung_opacity image 6
100%                                                50/50 [00:07<00:00,  6.57it/s]
Generating lung_opacity image 7
100%                                                50/50 [00:07<00:00,  6.58it/s]
Generating lung_opacity image 8
100%                                                50/50 [00:07<00:00,  6.57it/s]
Generating lung_opacity image 9
100%                                                50/50 [00:07<00:00,  6.57it/s]
Generating lung_opacity image 10
100%                                                50/50 [00:07<00:00,  6.60it/s]
Generating pleural_effusion image 1
100%                                                50/50 [00:07<00:00,  6.60it/s]
Generating pleural_effusion image 2
100%                                                50/50 [00:07<00:00,  6.58it/s]
```

```
Generating pleural_effusion image 3
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating pleural_effusion image 4
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating pleural_effusion image 5
100%                                      50/50 [00:07<00:00,  6.56it/s]
Generating pleural_effusion image 6
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating pleural_effusion image 7
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating pleural_effusion image 8
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating pleural_effusion image 9
100%                                      50/50 [00:07<00:00,  6.57it/s]
Generating pleural_effusion image 10
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating pneumothorax image 1
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating pneumothorax image 2
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating pneumothorax image 3
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating pneumothorax image 4
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating pneumothorax image 5
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating pneumothorax image 6
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating pneumothorax image 7
100%                                      50/50 [00:07<00:00,  6.54it/s]
Generating pneumothorax image 8
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating pneumothorax image 9
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating pneumothorax image 10
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating lung_nodules image 1
100%                                      50/50 [00:07<00:00,  6.57it/s]
```

```
Generating lung_nodules image 2
100%                                      50/50 [00:07<00:00,  6.64it/s]
Generating lung_nodules image 3
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating lung_nodules image 4
100%                                      50/50 [00:07<00:00,  6.57it/s]
Generating lung_nodules image 5
100%                                      50/50 [00:07<00:00,  6.64it/s]
Generating lung_nodules image 6
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating lung_nodules image 7
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating lung_nodules image 8
100%                                      50/50 [00:07<00:00,  6.56it/s]
Generating lung_nodules image 9
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating lung_nodules image 10
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating lung_fibrosis image 1
100%                                      50/50 [00:07<00:00,  6.56it/s]
Generating lung_fibrosis image 2
100%                                      50/50 [00:07<00:00,  6.57it/s]
Generating lung_fibrosis image 3
100%                                      50/50 [00:07<00:00,  6.57it/s]
Generating lung_fibrosis image 4
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating lung_fibrosis image 5
100%                                      50/50 [00:07<00:00,  6.54it/s]
Generating lung_fibrosis image 6
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating lung_fibrosis image 7
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating lung_fibrosis image 8
100%                                      50/50 [00:07<00:00,  6.54it/s]
Generating lung_fibrosis image 9
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating lung_fibrosis image 10
100%                                      50/50 [00:07<00:00,  6.56it/s]
```

```
Generating cardiomegaly image 1
100%                                      50/50 [00:07<00:00,  6.56it/s]
Generating cardiomegaly image 2
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating cardiomegaly image 3
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating cardiomegaly image 4
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating cardiomegaly image 5
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating cardiomegaly image 6
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating cardiomegaly image 7
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating cardiomegaly image 8
100%                                      50/50 [00:07<00:00,  6.64it/s]
Generating cardiomegaly image 9
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating cardiomegaly image 10
100%                                      50/50 [00:07<00:00,  6.59it/s]
Potential NSFW content was detected in one or more images. A black image will be returned instead. Try again with a different prompt and/or seed.
Generating medical_devices image 1
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating medical_devices image 2
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating medical_devices image 3
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating medical_devices image 4
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating medical_devices image 5
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating medical_devices image 6
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating medical_devices image 7
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating medical_devices image 8
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating medical_devices image 9
```

```
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating medical_devices image 10
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating imaging_artifacts image 1
100%                                      50/50 [00:07<00:00,  6.64it/s]
Generating imaging_artifacts image 2
100%                                      50/50 [00:07<00:00,  6.56it/s]
Generating imaging_artifacts image 3
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating imaging_artifacts image 4
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating imaging_artifacts image 5
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating imaging_artifacts image 6
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating imaging_artifacts image 7
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating imaging_artifacts image 8
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating imaging_artifacts image 9
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating imaging_artifacts image 10
100%                                      50/50 [00:07<00:00,  6.62it/s]
Generating pa_view image 1
100%                                      50/50 [00:07<00:00,  6.58it/s]
Generating pa_view image 2
100%                                      50/50 [00:07<00:00,  6.60it/s]
Generating pa_view image 3
100%                                      50/50 [00:07<00:00,  6.61it/s]
Generating pa_view image 4
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating pa_view image 5
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating pa_view image 6
100%                                      50/50 [00:07<00:00,  6.59it/s]
Generating pa_view image 7
100%                                      50/50 [00:07<00:00,  6.59it/s]
```

Generating pa view image 8

```
import shutil

dataset_folder = "chest_xray_dataset"
zip_name = "chest_xray_dataset.zip"

shutil.make_archive(
    base_name="chest_xray_dataset",
    format="zip",
    root_dir=dataset_folder
)

print("✅ Dataset zipped as chest_xray_dataset.zip")
```

```
✅ Dataset zipped as chest_xray_dataset.zip
```

Generating ap view image 4

```
from google.colab import files
files.download("chest_xray_dataset.zip")
```

Generating ap view image 6

100% 50/50 [00:07<00:00, 6.63it/s]

```
from google.colab import drive
```