SAM3 Local Offline Segmentation Demo

A local offline image segmentation application powered by **SAM3 (Segment Anything Model 3)**, Meta's cutting-edge vision foundation model, with dual interaction modes.

Tech Stack

- **SAM3** - Meta's latest universal segmentation model with text-prompted zero-shot segmentation
- **PyTorch** - Deep learning inference framework with CUDA GPU acceleration
- **Gradio** - Rapid Web UI prototyping
- **PyQt5** - Cross-platform desktop GUI framework
- **Matplotlib** - Segmentation result visualization
- **PIL/Pillow + NumPy** - Image processing and numerical computation

Key Features

- 🚀 **Local LLM Deployment** - Fully offline, no cloud API required
- 🎯 **Open-Vocabulary Segmentation** - Segment any object using natural language descriptions
- ⚡ **Single Model Load** - Load once, infer efficiently throughout the session
- 🌐 **Dual Interface** - Web UI (Gradio) + Desktop GUI (PyQt5)

Two Demo Interfaces

1. Web Demo (Gradio)

```
pip install gradio matplotlib torch
python demo_gui.py
```

- Access via browser at `http://localhost:7860`
- Ideal for quick demos and remote access
- One-click example prompts

2. Desktop App (PyQt5)

```
pip install PyQt5 pillow numpy torch
python sam_qt_gui.py
```

- 📊 Real-time performance metrics (set_image / inference / visualization timing)
- 📂 Batch folder processing, A/D hotkeys to navigate images
- 🔍 Ctrl+scroll to zoom, drag to pan for detail inspection

Usage

1. Click "Load Model" to initialize SAM3
2. Upload or select an image
3. Enter a text prompt (e.g., "dog", "person in red shirt", "car")
4. Adjust confidence threshold
5. Run segmentation and view results

Requirements

- Python 3.8+
- CUDA GPU (recommended, CPU also supported)
- Model weights (~3GB)

> This project demonstrates a complete solution for deploying cutting-edge vision foundation models locally, with both Web and desktop interfaces.