

Image Similarity Finder

A tool with both GUI and command-line interfaces that finds visually similar images across directories, regardless of size, format, or minor modifications.

Features

- Find images similar to a reference image across multiple directories
- Works with different image sizes and aspect ratios
- Supports various image formats (JPG, PNG, BMP, TIFF, WebP, GIF)
- Adjustable similarity threshold for fine-tuning results
- Configurable number of results to display
- User-friendly graphical interface with image preview
- Command-line interface for automation and scripting
- Type-safe implementation with Pydantic models
- Robust error handling and validation
- Easy installation and uninstallation

Installation

1. Download the installation package
2. Make the installer executable:

```
chmod +x install.sh
```

3. Run the installer:

```
./install.sh
```

The installer will: - Install required Python packages (numpy, pillow, opencv-python, scikit-learn) - Set up the tool in `~/.image-similarity-finder/` - Create a command-line executable at `~/.local/bin/imagesim`

Usage

Graphical User Interface

Launch the GUI with:

```
imagesim --gui
```

or simply:

```
imagesim -g
```

The GUI provides: - Visual image selection - Directory browsing - Adjustable threshold with slider - Results with similarity scores - Image preview

Command-line Interface

Basic usage

```
imagesim path/to/reference_image.jpg path/to/search/directory
```

Search multiple directories

```
imagesim reference_image.jpg dir1 dir2 dir3
```

Adjust similarity threshold (0-1, where 1 is identical)

```
imagesim reference_image.jpg directory --threshold 0.6
```

Limit number of results

```
imagesim reference_image.jpg directory --max-results 5
```

How It Works

The tool uses computer vision techniques to find similar images:

1. **Feature Extraction:** Each image is converted into a feature vector using Histogram of Oriented Gradients (HOG)
2. **Normalization:** Feature vectors are normalized to ensure consistent comparison
3. **Similarity Calculation:** Cosine similarity measures how similar the vectors are
4. **Result Ranking:** Images are ranked by similarity score and returned in descending order

Examples

Find similar landscape photos:

```
imagesim vacation/sunset.jpg ~/Pictures --threshold 0.75
```

Find all variations of a logo across multiple folders:

```
imagesim assets/logo.png ~/Documents ~/Downloads ~/Desktop --threshold 0.8
```

Uninstallation

To remove the tool completely:

```
~/image-similarity-finder/uninstall.sh
```

Or use the separate uninstaller:

```
./uninstall.sh
```

Requirements

- Python 3
- pip (Python package manager)
- Required Python packages (automatically installed):
 - numpy: For numerical operations
 - pillow: For image processing
 - opencv-python: For computer vision algorithms
 - scikit-learn: For similarity calculations
 - tkinter: For the graphical user interface
 - pydantic: For data validation and modeling

Troubleshooting

Command not found

If you see “command not found” when trying to run `imagesim`, your `~/local/bin` directory might not be in your PATH. Add it by running:

```
echo 'export PATH="$PATH:$HOME/.local/bin"' >> ~/.bashrc
source ~/.bashrc
```

Permission errors

If you encounter permission errors during installation, try:

```
pip install --user numpy pillow opencv-python scikit-learn
```

Performance considerations

- Processing large images or searching through many directories may take time
- For faster results with large datasets, consider using a lower similarity threshold