```
In [1]: import pandas as pd
        import os

        # Base directory
        base_dir = '/Users/Dell/Documents/TASK1/WikiArt'
        output_dir = os.path.join(base_dir, 'nga_portraits')
        os.makedirs(output_dir, exist_ok=True)

        # Load NGA metadata (for reference only, not for images)
        metadata = pd.read_csv('https://raw.githubusercontent.com/NationalGalleryOfArt/opendata/main/data/objects.csv', low_memory=False)
        portraits = metadata[(metadata['classification'] == 'Painting') &
                            (metadata['title'].str.lower().str.contains('portrait', na=False))].head(100)

        print("Sample portraits from metadata (for reference):")
        print(portraits[['title', 'attributioninverted']].head())

        # Use existing images in nga_portraits/ (no renaming needed)
        image_files = [f for f in os.listdir(output_dir) if f.endswith('.jpg')]
        if len(image_files) < 2:
            print("Insufficient images. Ensure 11+ portrait images are in WikiArt/nga_portraits/.")
            exit()

        image_paths = [os.path.join(output_dir, f) for f in image_files]
        # Assign simple labels (e.g., alternate between two artists for demo purposes)
        labels = [f"artist_{i % 2}" for i in range(len(image_files))]  # e.g., artist_0, artist_1, artist_0, ...

        print(f"Processed {len(image_paths)} images with filenames:")
        print(image_files[:5])  # Show first 5 filenames

        # Create pairs
        pairs = []
        pair_labels = []
        for i in range(len(image_paths)):
            for j in range(i + 1, len(image_paths)):
                pairs.append([image_paths[i], image_paths[j]])
                pair_labels.append(1 if labels[i] == labels[j] else 0)

        # Save to CSV
        if pairs:
            df = pd.DataFrame(pairs, columns=['image1', 'image2'])
            df['label'] = pair_labels
            df.to_csv(os.path.join(base_dir, 'nga_similarity_pairs.csv'), index=False)
            print(f"Created {len(df)} pairs in nga_similarity_pairs.csv")
            print(df.head())
        else:
            print("No pairs created. Check image count.")
```

```
Sample portraits from metadata (for reference):
                            title                attributioninverted
4              Portrait of a Man                   Andrea del Castagno
5       Profile Portrait of a Lady          Franco-Flemish 15th Century
8           Portrait of a Young Man  Antonello da Messina, Attributed to
10  Portrait of a Man with an Arrow                        Memling, Hans
11          Portrait of a Gentleman                          Mor, Antonis
Processed 10 images with filenames:
['the_apostle_paul_1942.9.59.jpg', 'family_portrait_1946.7.4.jpg', 'amsterdam_harbor_scene_2011.3.1.jpg', 'agrippina_and_germanicus_1963.8.1.jpg', 'the_alba_madonna_1937.1.24.jpg']
Created 45 pairs in nga_similarity_pairs.csv
                                               image1  \
0  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...
1  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...
2  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...
3  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...
4  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...

                                               image2  label
0  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...      0
1  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...      1
2  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...      0
```

```
3  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...        1
4  /Users/Dell/Documents/TASK1/WikiArt/nga_portra...        0
```

In [ ]: