Project Report: Interactive User-Guided Image Colorization Application

Overview

This project presents an interactive image colorization system that allows users to guide the colorization of grayscale images by selecting specific pixel locations and assigning desired colors. The implementation features a complex deep neural network with a U-Net style architecture enhanced by residual blocks and skip connections to produce high-quality colorizations influenced by user-provided color hints.

Objectives

- Develop an interactive user interface for uploading grayscale images.
- Enable users to specify multiple pixel locations and assign individual colors as hints.
- Design a deep learning model that dynamically incorporates user hints to guide colorization.
- Display output images with vibrant colorization reflecting user input.
- Show relevant metrics such as image resolution, unique grayscale levels, and hint point counts.

Model Architecture

- **Input:** The model takes grayscale luminance (L channel) along with two-color hint channels (ab channels) and a hint mask, stacked to form a 4-channel input tensor.
- **Encoder:** Four convolutional layers progressively downsample the input, extracting hierarchical features. Each layer includes batch normalization, ReLU activations, and residual blocks to improve learning capability.
- **Bottleneck:** A deeper convolutional block captures complex feature representations.
- **Decoder:** Transposed convolution layers perform upsampling, concatenated with encoder skip connections to preserve spatial details, followed by residual blocks for refinement.

- **Output:** The model outputs two ab color channels in LAB color space, normalized within [-1, 1], which are resized to the original image size and combined with input luminance for final colorization.

User Hints and Interaction

- Users input pixel coordinates (x, y) and select a color hex code for each hint point in an interactive data table.
- The system converts each RGB hint color to LAB ab channels normalized appropriately.
- The hint points and corresponding colors are integrated into mask and ab hint channels as input to the model.
- The UI provides buttons to dynamically add or reset hint points.

Technical Stack

- **Deep Learning Framework:** PyTorch, used for model definition, inference, and tensor manipulations.
- **Image Processing:** OpenCV and PIL handle image conversions and processing, including color space transformations.
- **User Interface:** Gradio provides the interactive web interface for image upload, hint input, and result display.
- **Data Management:** Pandas is used to handle the table of hint points, enabling smooth UI-data integration.

Features and Functionality

- Upload grayscale images that are resized to 256x256 for uniform model input.
- Input multiple user-defined hue hints with independent colors.
- Interactive buttons to add new hint rows or reset all hints.
- Real-time model inference on user input to produce colorized images.
- Metrics displayed include:
    - Image resolution.
    - Unique grayscale levels in the input image.

- Count of user hint points.

Limitations and Considerations

- The model is provided without pretrained weights; effective colorization quality depends on training on extensive datasets.
- Input image and hint coordinates must correspond correctly to the standardized resolution.
- Realistic color propagation depends on prompt, distributed, and accurate hint placement.
- Computational performance requires appropriate hardware for interactive speed.

Potential Improvements

- Integrate pretrained weights from state-of-the-art colorization models to improve visual quality.
- Implement training procedures with color datasets to fine-tune this model architecture.
- Enhance UI with freehand drawing of color hints or brush tools.
- Provide multi-scale processing for high-resolution images.
- Include feedback and undo/redo options for hint management.

Conclusion

This project successfully builds a complex interactive framework for user-guided image colorization, combining advanced neural network architecture with an intuitive user interface. It allows creative and precise control over colorization while displaying informative metrics about the processed images. With further training and enhancements, it has strong potential for practical applications in photo restoration, artistic rendering, and interactive graphics.