Recommendations for Improving Model Performance

The objective of this project is to enhance the accuracy and usability of a computer vision pipeline for satellite image matching. While the current implementation uses a pre-trained LoFTR model, there are opportunities to improve performance by addressing its limitations and optimizing the workflow. Below are the key recommendations:

1. Fine-Tuning Pre-Trained Models

The current approach heavily relies on a pre-trained LoFTR model, which may not be fully optimized for Sentinel-2 satellite imagery. Fine-tuning the model on task-specific data can significantly enhance its accuracy and reliability.

Recommendations:

- Task-Specific Fine-Tuning: Train the LoFTR model on a dataset of Sentinel-2 image pairs to adapt its keypoint detection and matching capabilities to the domain of satellite imagery.

- Custom Loss Functions: Experiment with loss functions that emphasize domain-specific matching, such as weighted losses for high-confidence regions.

- Hyperparameter Tuning: Optimize learning rate, batch size, and other training parameters to improve convergence and accuracy.

2. Incorporating Multi-Modal Features

Satellite images often contain multi-spectral data beyond the visible spectrum. The current model processes RGB inputs, which might underutilize the full potential of Sentinel-2 imagery.

Recommendations:

- Multi-Spectral Input: Extend the input pipeline to include additional spectral bands (e.g., Near-Infrared and Shortwave Infrared) and modify the model to leverage this data.

- Feature Fusion: Employ techniques like early or late feature fusion to combine information from multiple bands effectively.

3. Enhanced Evaluation Metrics and Validation

The current workflow may not fully capture the model's performance in real-world scenarios. Expanding evaluation metrics and validation procedures can ensure robustness.

Recommendations:

- Geospatial Accuracy Metrics: Introduce evaluation metrics that quantify errors in terms of geospatial alignment, such as RMSE in pixel coordinates.
- Robust Validation Dataset: Curate a validation dataset that includes diverse geographical regions, seasons, and atmospheric conditions to ensure generalizability.
- Cross-Domain Testing: Validate the model's performance on satellite imagery from other sources (e.g., Landsat, PlanetScope) to assess adaptability.

References

- LoFTR: Detector-Free Local Feature Matching with Transformers
- Hugging Face Computer Vision Course - Feature Matching