LOCATEdit: Graph Laplacian Optimized Cross Attention for Localized Text-Guided Image Editing

Supplementary Material

1. Broader Impact

Our work advances the precision of text-guided image editing by ensuring that modifications are both spatially consistent and semantically faithful. This improvement has the potential to benefit a wide range of applications—from enhancing creative workflows in digital art and advertising to supporting critical tasks in medical imaging and scientific visualization—by reducing the need for extensive manual post-processing. At the same time, the increased reliability of automated editing tools underscores the importance of establishing robust ethical guidelines for their use, particularly in contexts where the authenticity of visual information is paramount. By delivering a method that better preserves the structural integrity of the source images, our approach paves the way for more trustworthy and accessible image editing solutions that can democratize creative technologies and support various high-stakes applications.

2. Proof of Lemma 1

Proof. To prove that $\mathbf{L}$ is PSD, we must show that for any $\mathbf{x} \in \mathbb{R}^n$, the quadratic form $\mathbf{x}^\top \mathbf{L}\mathbf{x}$ is nonnegative:

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} = \mathbf{x}^\top (\mathbf{D} - \mathbf{S}_{\text{sym}})\mathbf{x}.$$

Expanding this expression, we have:

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} = \mathbf{x}^\top \mathbf{D}\mathbf{x} - \mathbf{x}^\top \mathbf{S}_{\text{sym}}\mathbf{x}.$$

The degree matrix $\mathbf{D}$ is diagonal, with entries $\mathbf{D}(i,i) = \sum_{j=1}^{n} \mathbf{S}_{\text{sym}}(i,j)$. Therefore:

$$\mathbf{x}^\top \mathbf{D}\mathbf{x} = \sum_{i=1}^{n} \mathbf{D}(i,i)x_i^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \mathbf{S}_{\text{sym}}(i,j) \right) x_i^2.$$

The second term, $\mathbf{x}^\top \mathbf{S}_{\text{sym}}\mathbf{x}$, is given by:

$$\mathbf{x}^\top \mathbf{S}_{\text{sym}}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{S}_{\text{sym}}(i,j) x_i x_j.$$

Substituting these into the quadratic form, we get:

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \mathbf{S}_{\text{sym}}(i,j) x_i^2 \right) - \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{S}_{\text{sym}}(i,j) x_i x_j.$$

Reorganizing terms:

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{S}_{\text{sym}}(i,j) \left(x_i^2 + x_j^2 - 2x_i x_j \right).$$

This simplifies to:

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{S}_{\text{sym}}(i,j)(x_i - x_j)^2.$$

Since $\mathbf{S}_{\text{sym}}(i,j) \geq 0$ (by definition of the symmetrized self-attention matrix) and $(x_i - x_j)^2 \geq 0$, every term in the summation is nonnegative. Therefore:

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Thus, $\mathbf{L}$ is positive semidefinite. □

3. Proof of Theorem 1

3.1. Optimization Problem

We consider the following optimization problem:

$$\min_{x \in \mathbb{R}^{R^2}} J(x), \tag{1}$$

where the objective function is defined as

$$J(x) = (x - x^{(0)})^\top \mathbf{\Lambda}(x - x^{(0)}) + \lambda\, x^\top \mathbf{L}\, x.$$

Here, the <u>fidelity</u> term $(x - x^{(0)})^\top \mathbf{\Lambda}(x - x^{(0)})$ penalizes deviations from the initial mask $x^{(0)}$. $\mathbf{\Lambda}$ is a diagonal matrix of per-node weights, so it re-weights squared deviations from the initial saliency map $\mathbf{m}_0$, pulling each node back toward its original importance in proportion to how confident we are in that node's initial score. The second term is the classic graph-Laplacian regularizer, for every edge (i,j) that connects two spatially or semantically related nodes, it penalizes the squared difference $(\mathbf{m}_i - \mathbf{m}_j)^2$. By summing these penalties over all edges weighted by their connection strength $\mathbf{S}_{i,j}$, the regularizer forces neighboring regions to have similar saliency values, thereby enforcing <u>smoothness</u> across the graph.

3.2. Existence and Uniqueness of the Solution

To obtain the refined mask, we solve the minimization problem in Equation (1). The first term is strictly convex since $\mathbf{\Lambda}$ is positive definite, and the second term is convex because L is positive semidefinite. Thus, the overall objective $J(x)$ is strictly convex and has a unique minimizer.

Taking the gradient with respect to x yields:

$$\nabla J(x) = 2\,\mathbf{\Lambda}(x - x^{(0)}) + 2\lambda\,\mathbf{L}\,x. \tag{2}$$

Setting $\nabla J(x) = 0$ gives:

$$\mathbf{\Lambda}(x - x^{(0)}) + \lambda \, \mathbf{L} \, x = 0. \quad (3)$$

Rearranging, we obtain:

$$(\mathbf{\Lambda} + \lambda \, L) \, x = \mathbf{\Lambda} \, x^{(0)}. \quad (4)$$

Since $\mathbf{\Lambda} + \lambda \, \mathbf{L}$ is positive definite, it is invertible, and the unique solution is

$$x^{*} = (\mathbf{\Lambda} + \lambda \, \mathbf{L})^{-1} \, \mathbf{\Lambda} \, x^{(0)}.$$

The positive semidefiniteness of L ensures the convexity of the regularization term, thereby guaranteeing the existence and uniqueness of the solution.

3.3. Additional Qualitative Results

This section presents qualitative results for refined masks achieved through graph Laplacian regularization and compares the editing outcomes with existing image editing methods.

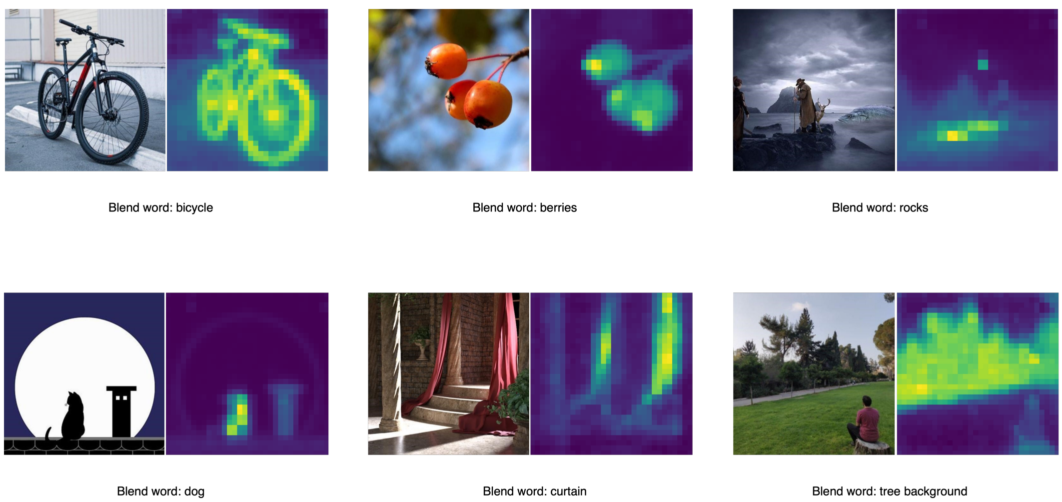


Figure 1. Refined masks after Graph Laplacian Regularization

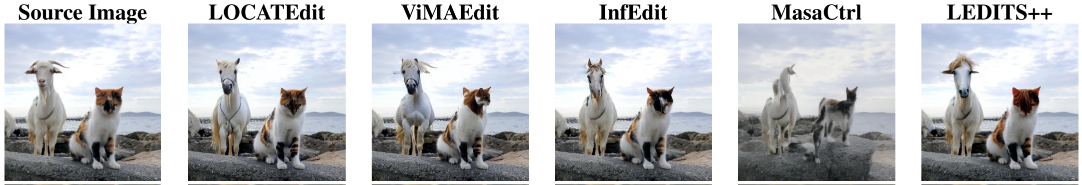
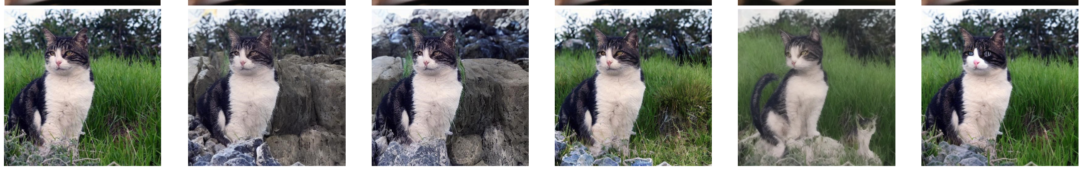
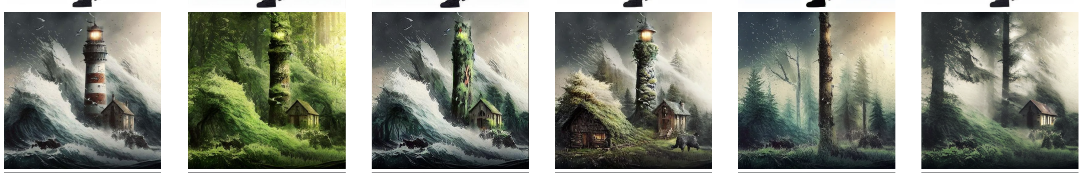
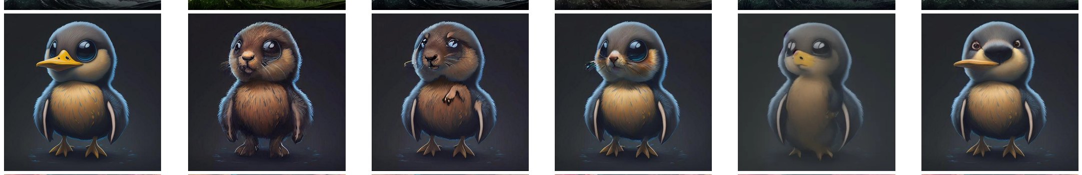
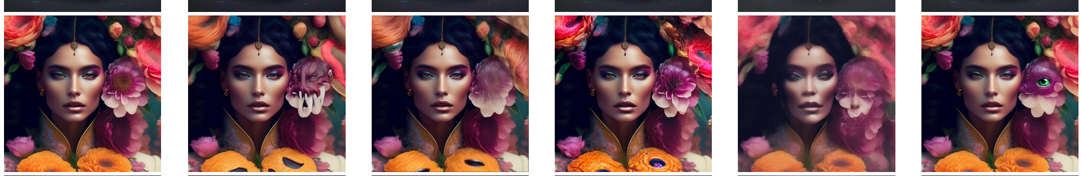
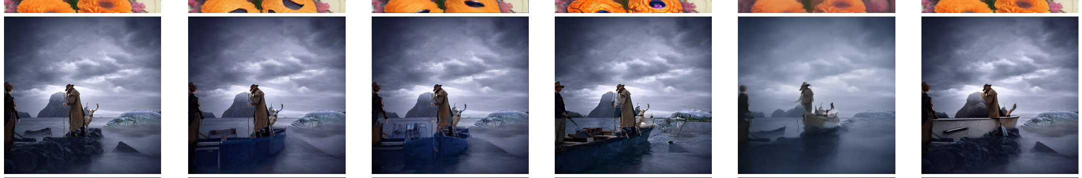
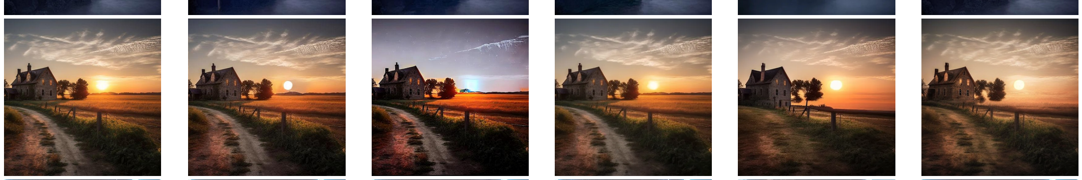
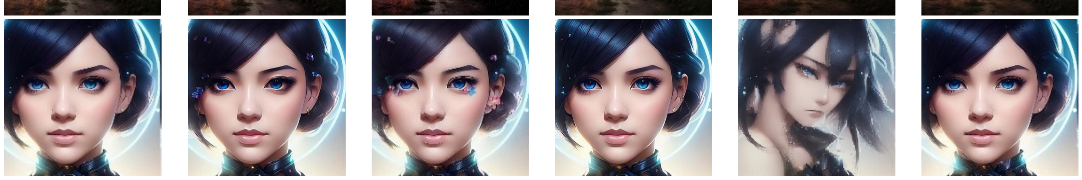
	Source Image	LOCATEdit	ViMAEdit	InfEdit	MasaCtrl	LEDITS++
a photo of ~~goat~~ **horse** and a cat standing on rocks near the ocean						
a ~~brown~~ **white** tea cup and a book						
a cat sitting in the ~~grass~~ **rocks**						
a woman with black hair and a white shirt is holding a ~~phone~~ **coffee**						
~~sea~~ **forest** and house						
a cute little ~~duck~~ **marmot** with big eyes						
a woman with ~~flowers~~ **monster** around her face						
the two people are standing on ~~rocks~~ **boat** with a fish						
the ~~sun~~ **moon** over an old farmhouse						
An asian woman with blue thick-lashed eyes and **flowers on her** black hair						

Table 1. Additional qualitative results on PIE-Bench