Contributions

1. A lateral inhibition network (LIN) is proposed for single MR image SR. With a small amount of parameters and computational overhead, LIN can achieve accurate and fast SR reconstruction.

2. We conduct a inhibition unit (LIU) to impose inhibitory regulation on features explicitly, which is motivated by the lateral inhibition mechanism.

3. We propose to fuse the shallow features with different receptive field sizes, which increases the diversity of the extracted features and provide more effective evidence for nonlinear inference and image reconstruction.

4. We experimentally verify that combining the lateral inhibition mechanism with the proposed shallow feature extraction strategy contributes to improving the performance of deep models.

Explicit Visual Inhibition

A famous computing model for simulating visual inhibition is the Hartline-Ratliff Equation [1].

\[ \phi_i = \phi(\ell_i - \sum_{j \neq i} w_{ij} \cdot \max(0, \ell_i - \ell_j)) \]  

(1)

where \( \ell_i \) denotes the \( i-th \) element of the reversed feature map \( \ell \), and \( \ell_i \) is the \( i-th \) element of input feature \( \ell \). \( w_{ij} \) indicates the inhibition coefficient of the \( j-th \) neuron on the \( i-th \) neuron, while \( t_j \) denotes the threshold that the \( j-th \) neuron must reach to inhibit the \( i-th \) neuron. \( \phi(\cdot) \) is a linear correction function, which we add to the original Hartline-Ratliff Equation. The symbols, i.e., $\mathbb{Q}$ and $\mathbb{Q}_i$, in the following figure are element-wise operations.

Inhibition tail (IT) of a lateral inhibition unit (LIU): The upper branch imitates the weight tensor \( w \) in the Hartline-Ratliff Equation, while the lower branch simulates the threshold tensor \( t \).

Network Architecture

The network architecture is shown in the following figure. Feature extraction is composed of a set of parallel dilated 3×3 conv layers followed by a 1×1 conv and a 3×3 conv layer. The outputs of these dilations convolutions are concatenated together along the channel direction. The extracted shallow feature is fed into the nonlinear mapping part of the network, which consists of a series of cascaded lateral inhibition blocks (LIBs). Here, residual learning is adopted to stabilize model training. Image reconstruction includes a upsacle module, which is usually followed by a 3×3 conv layer.

The overall network structure: The overall architecture of our LIN network. The dilated 3×3 convolutions for feature extraction have different dilation rates to collect features in the receptive fields with different sizes.

Ablation Investigation and Model Analysis

We keep the backbone of the network unchanged and adjust FE and IT accordingly. For FE, the comparative case is a single 3×3 conv layer that is denoted as "F", and a group of dilated 3×3 conv layers is denoted as "IT". As for FE, "IT" stands for removing the IT from a LIU, the opposite is denoted as "IT".

The valid curves of these different configurations are shown in the left figure. We also analyze the impact of the number of LIB and LIU on model performance. We fix \( n \) at 4 and set \( m = 2, 3, 4 \), and fix \( m \) at 4 and set \( n = 1, 2, 3, 4 \). The right figure shows the testing results of these configurations.

Ablation study and model analysis: The valid curves of different configurations of the network for studying the feature extraction and visual inhibition mechanism (left). The testing performance comparison between the models with different numbers of LIB and LIU (right).

Quantitative and Qualitative Evaluation

Quantitative comparison between typical SR models. The maximum values in each comparative cell are marked in blue (PSNR (dB) / SSIM).

As can be seen from both quantitative and visual comparisons above, the proposed LIN model can achieve better SR performance with fewer model parameters and computational overhead.