Supporting Information

Feature-enhanced fiber bundles imaging based on light field acquisition

*Haogong Feng, Runze Zhu and Fei Xu**

S1. Network architectures

The network FBNet used in this work consists of two parts, the generator and the discriminator. u-FBNet and R-FBNet are the terms used when the generator selects U-Net and Resnet, respectively.

Two designs of generators in FBNet

Below, we provide two kinds of generators to extract image features to demonstrate that the light field information from the intra-core pattern contributes to the reconstruction of high-frequency features. Considering that there is a lot of low-level information shared between inputs and outputs, the first one (Fig. S1a) follows the general shape of a "U-Net"with skip connections between each layer i and layer $n-i$, where n is the total number of layers. The "U-Net" architecture consists of an encoder and a decoder. Each layer of the encoder is composed of a Convolution layer, a BatchNorm layer and a ReLU layer. The number of filters in each layer is 64, 128, 256, 512, 512, 512, 512, 512 in the order. Each layer of the decoder is composed of a Convolution layer, a BatchNorm layer, a Dropout layer and a ReLU layer with a dropout rate of 50%. The number of filters in each layer is 512, 512, 512, 512, 256, 128, 64 in the order. All convolutions are 4× 4 spatial filters applied with stride 2. Convolutions in the encoder downsample by a factor of 2, while in the decoder they upsample by a factor of 2. The final layer of the decoder uses a convolutional layer to map to the number of output channels, followed by a Tanh function. All ReLUs in the encoder are leaky, with slope 0.2, while ReLUs in the decoder are not leaky.

Another type of generator (Fig. S1b) is based on nine residual blocks, which has proven to be effective in extracting detailed features from images. Before entering the residual block, a convolutional layer with a convolutional kernel size of 7 and two downsampling layers are designed to preprocess the input image to a pixel size of 64*64, containing 256 channels. Each residual block contains, in order, a Convolutional layer, a BatchNorm layer, a ReLU layer, a Convolutional layer, a BatchNorm layer and a skip connections layer. The residual block does not change the pixel size of the feature image. This is followed by two upsampling layers (kernel_size=3, stride=2, padding=1, same as the downsampling layer). The final layer uses a convolutional layer to map to the number of output channels, followed by a Tanh function as shown in Fig. S1.

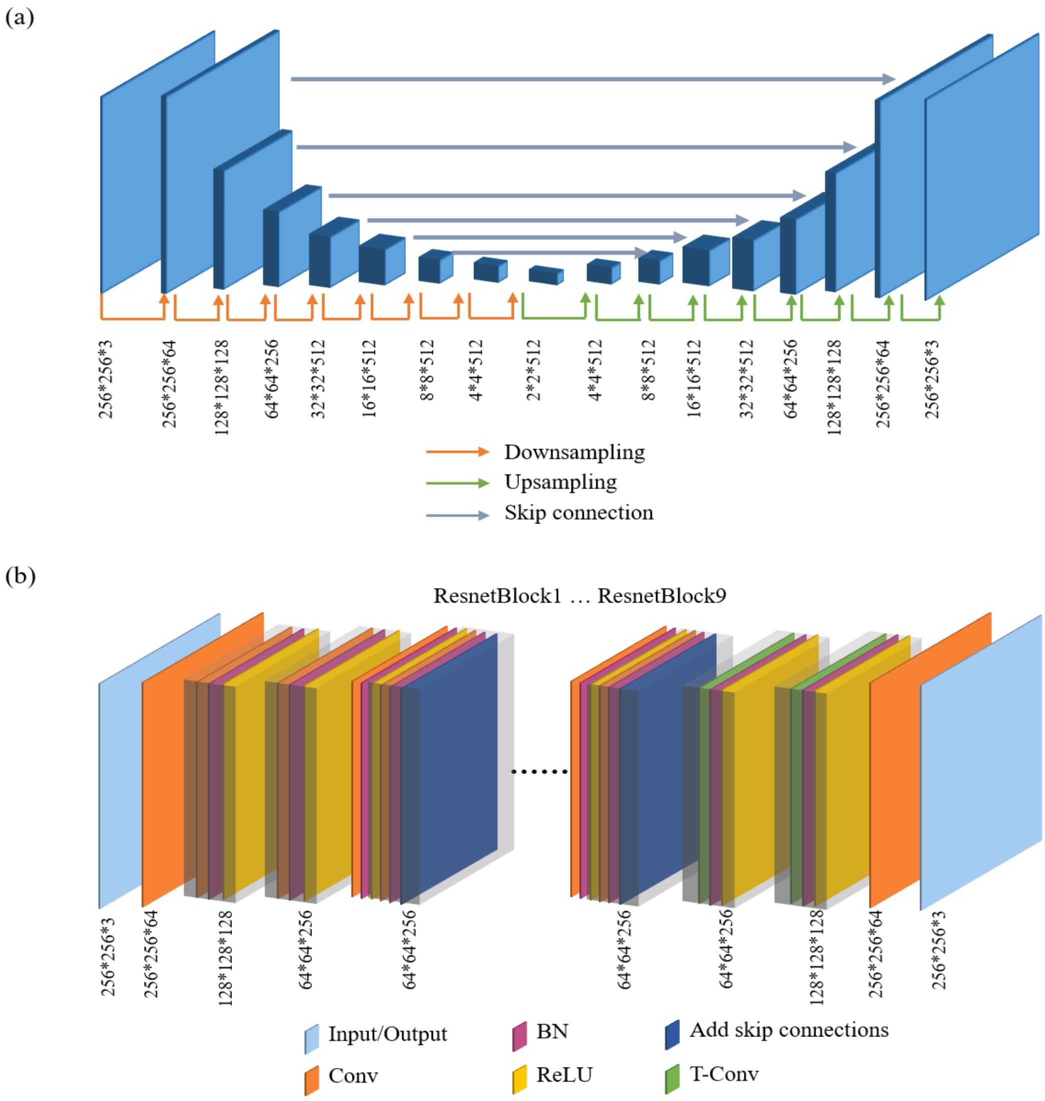

Fig. S1. Schematic diagram of the structure of generators in U-FBNet and R-FBNet. The generator in U-FBNet. (b) The generator in R-FBNet.

Discriminator architectures

Discriminator use modules of the form convolution-BatchNorm-ReLU, too. The number of filters in each layer is 64, 128, 256, 512 in the order. A convolution is applied to map to a 1-dimensional output, followed by a Sigmoid function, after the last layer.

S2. Test results for edge improvement using U-FBNet

Below, we show the results of the reconstruction using U-FBNet. The reconstructed maps based on the dataset of in-core pattern (FM-U) have spatial frequency information that are closer to the ground truths.

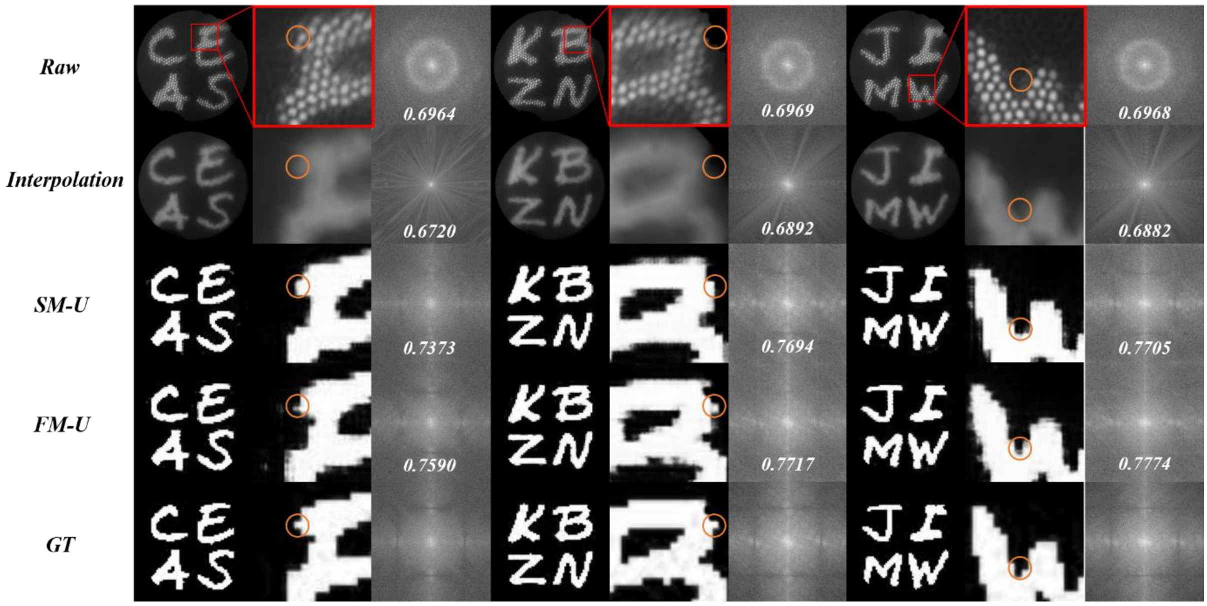

Fig. S2. Part of the test results for edge improvement using U-FBNet. The first row displays the original images recorded by the image sensor containing the mode pattern, their regions of interest and the spatial frequency spectrums. The results of reconstruction by traditional interpolation method are shown in the second row. The third and fourth rows present the reconstruction results by U-FBNet for different datasets. Some boundary features are marked with orange circles. The ground truths are listed in the last row as a comparison. The similarity coefficients of their spectrograms with the corresponding GTs are recorded in white.

S3. Test results for textural features enhancement using R-FBNet

We show the results of the reconstruction using R-FBNet. Similar to the results shown in Fig. 5, artifacts appear for the reconstruction of fine features when the pattern within the core is not considered.

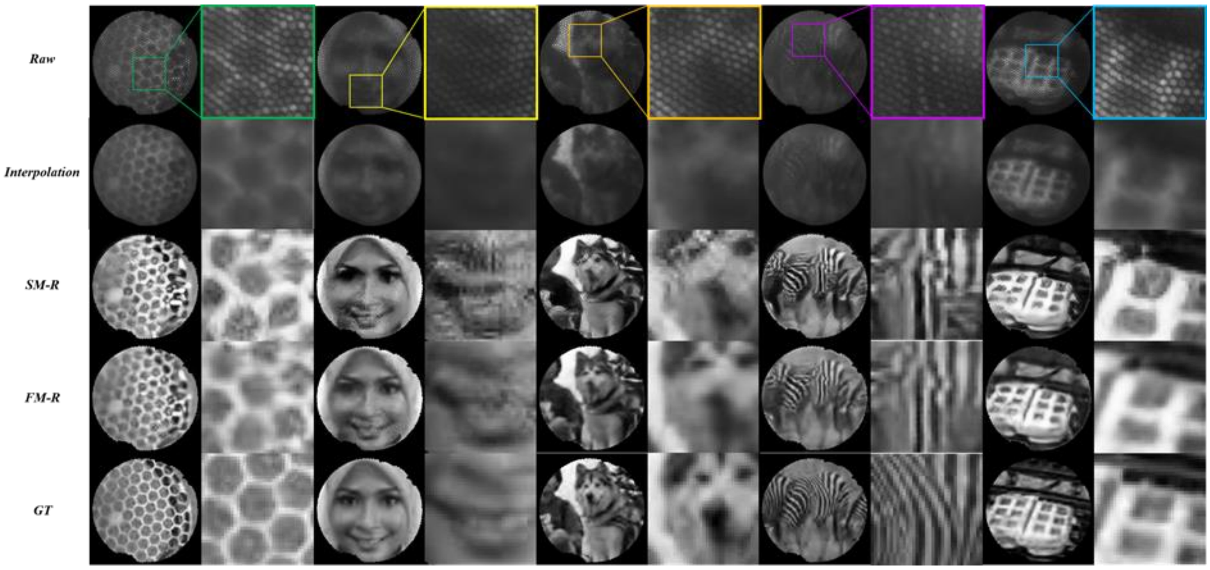

Fig. S3. Results of training and testing of SM and FM for textural features through R-FBNet. The first, second and last rows illustrate the raw acquisition maps, the reconstruction maps by traditional methods and the real images. The third and fourth rows show the results of the reconstructions using SM and FM as the dataset, respectively. The regions of interest are enlarged by the color window, corresponding to real images size of about 40 μm at the distal.

S4. Test results for image colorization using U-FBNet

We show the results of the reconstruction using U-FBNet. The reconstructed saturation maps of the red and blue channels are closer to that of the GTs, and the green channel is less well restored. The reasons for this are analyzed in Section 3.3 in the original manuscript.

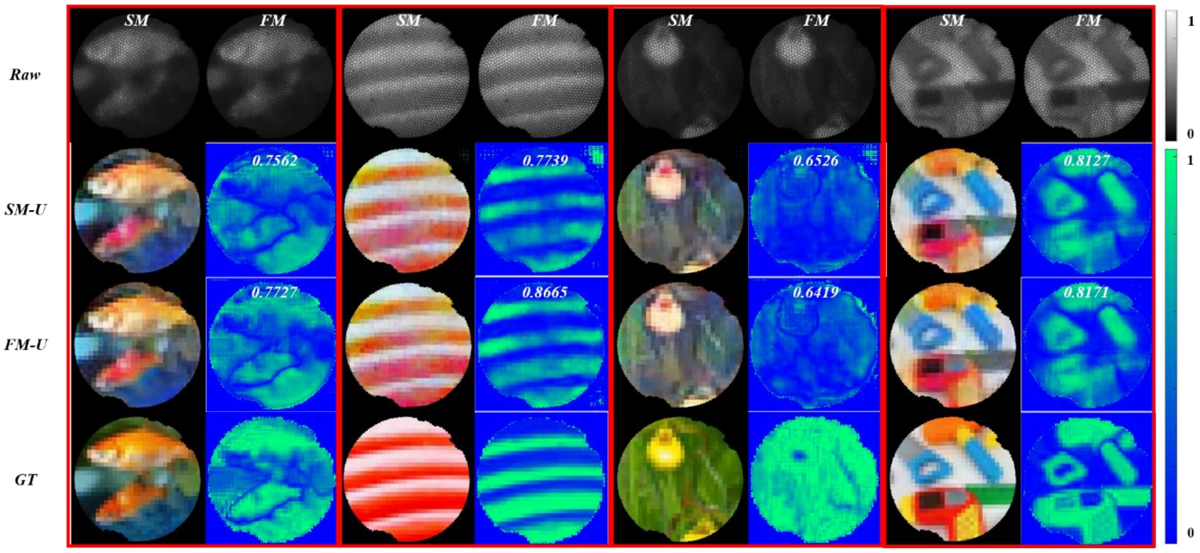

Fig. S4. Results of training and testing of SM and FM for image colorization based on U-FBNet. The first row illustrates the raw SM and FM datasets used for image coloring. The colored maps and their saturation maps are shown in the second and third rows. The last row lists the GTs and their saturation graphs as a comparison. The white numbers indicate the similarity between the saturation of the reconstructed maps and that of GTs.