Binarized Neural Network for Single Image Super Resolution

Jingwei Xin[1], Nannan Wang[2⋆], Xinrui Jiang[2], Jie Li[1], Heng Huang[3], and Xinbo Gao[1]

[1] State Key Laboratory of Integrated Services Networks, School of Electronic Engineering, Xidian University, Xi'an 710071, China
[2] State Key Laboratory of Integrated Services Networks, School of Telecommunications Engineering, Xidian University, Xi'an 710071, China
[3] School of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA
jwxintt@gmail.com, nnwang@.xidian.edu.cn
{leejie,xbgao}@mail.xidian.edu.cn
xrjiang@stu.xidian.edu.cn, heng.huang@pitt.edu

Abstract. Lighter model and faster inference are the focus of current single image super-resolution (SISR) research. However, existing methods are still hard to be applied in real-world applications due to the heavy computation requirement. Model quantization is an effective way to significantly reduce model size and computation time. In this work, we investigate the binary neural network-based SISR problem and propose a novel model binarization method. Specially, we design a bit-accumulation mechanism (BAM) to approximate the full-precision convolution with a value accumulation scheme, which can gradually refine the precision of quantization along the direction of model inference. In addition, we further construct an efficient model structure based on the BAM for lower computational complexity and parameters. Extensive experiments show the proposed model outperforms the state-of-the-art binarization methods by large margins on 4 benchmark datasets, specially by average more than 0.7 dB in terms of Peak Signal-to-Noise Ratio on Set5 dataset.

Keywords: Single Image Super-Resolution, Model Quantization, Binary Neural Network, Bit-Accumulation Mechanism.

1 Introduction

Single image super-resolution (SISR) aims to recover a high-resolution (HR) version from a low-resolution (LR) input image, which has been widely used in many fields, such as medical imaging [21], satellite imaging [25], security and surveillance [35] and so on. As a classical low-level problem, SISR is still an active yet challenging research topic in the field of computer vision due to its ill-poseness nature and high practical values.

⋆ Corresponding Author: Nannan Wang

Fig. 1: Results of binarized networks based on SRResNet.

Recently, convolutional neural network-based (CNN-based) super-resolution methods have been demonstrated state-of-the-art performance by learning a mapping from LR to HR image patches. Dong et al.[6] proposed the SRCNN model with only three convolution layers, which is the first deep learning method in super-resolution community. From then, researchers had carried out its study with different perspectives and obtained plentiful achievements, and various model structure and learning strategies are used in SR networks (*e.g.*, residual learning [10], recursive learning [11, 23, 24], skip connection [13, 27], channel attention [32] and so on). These CNN-based methods can often achieve satisfactory results, but the increasing model size and the computational complexity severely restrict their applications in the real world. Recently, there appear some lightweight approaches to reduce the computational complexity [1, 8]. However, it is still a huge burden for mobile devices with limited computing resources.

As a way to significantly reduce model size and computation time, binarized neural network (BNN) can replace the floating point operations with the bitcounting operations, and has shown excellent performance on many semantic-level tasks such as image classification and recognition. However, quantifying the weights and activations will lead to serious information loss in the process of network inference, which is unacceptable for super-resolution tasks because of its highly dependence on the accuracy of pixel values. Ma et al. [17] tried to apply the binarization to the residual blocks for image super resolution, and improve the performance by learning a gain term. However, this work only investigate the binary weights and full-precision activations model, the convolution calculation is not the bitcounting operation. Then the model's inference speed cannot be simplified enough.

In this paper, we introduce an efficient and accurate CNN-based SISR model by binarizing the weights and even the intermediate activations. Our binarization approach aims to find the best approximations of the convolutions using binary operations, and perform image super resolution task on the devices with limited computing resources. As shown in Fig.1, our method achieves better visual SR results compared with state-of-the-art methods, and could even be comparable to the performance of the full-precision convolutional model.

Overall, our contributions are mainly threefold: (1) To the best of our knowledge, this is the first work to introduce the binary neural network (both the weights and activations are binary values) to the field of image super resolution, in which the convolutions can be estimated by the bitcounting operations. As a result, our model could obtain about 58× speed up (the model size is also about 32× lighter) than that of an equivalent network with single-precision weight values. Its inference can be done very efficiently on CPU. (2) A bit accumulation mechanism is proposed to approximate the full-precision convolution with an iterative scheme of cascaded binary weights and activations. Whats more, it implements highly accurate one-bit estimation of filter and activation only relying on existing basic models and without introducing any other additional inference modules. (3) We construct an architecture of binary super-resolution network (BSRN) for highly accurate image SR problem. Experimental results show that the proposed BSRN can achieve the better SR performance with a lighter network structure and fewer operands.

2 Related Work

2.1 Single Image Super Resolution

Recently numerous deep learning based methods have been explored and shown dramatic improvements on the SISR tasks. Deep SISR network is first introduced by SRCNN [6] which is an end-to-end model with only three convolution layers. Considering the effectiveness of deep learning and the natural sparsity of images [30], Wang *et al.* [28] proposed a Sparse Coding Network (SCN) to make full use of the natural sparsity of images. Later on, Kim et al. [10] proposed a 20-layers network VDSR, which demonstrates significant improvement by increasing the network depth. After this, many others followed up with this strategy for network design. Tong et al. [27] adopted dense blocks to construct a 69-layers network SRDenseNet. Extended from it, Lim et al. [15] developed a more in-depth and broader residual network known as EDSR, which exhibits comparable performance for SR task. Zhang et al. [33] introduced a residual dense network RDN, which combines residuals and dense blocks to achieve higher image reconstruction performance with higher feature extraction capability. To overcome the gradient vanishing problem, residual channel attention network is adopted in RCAN [32], which proposes long and short skip connections in the residual structure to obtain deep residual network.

Aiming to achieve better performance with less parameters, recursive learning have been employed in SISR. Kim *et al.* proposed a Deeply-Recursive Convolutional Network(DRCN) [11] for SISR task. Then, Tai *et al.* also proposed Deep Recursive Residual Network(DRRN) [23], which introduces a very deep model (52 layers) with residual learning and recursive module. The authors further proposed Memory Network (MemNet) [24], which could adaptively combine the multi-scale features by the memory blocks. Another research direction which is time-saving network designing. For instance, the deconvolution layer has been proposed in FSRCNN [7] and sub-pixel layer has been introduced in ESPCN

[20], which are the better up-sampling operator for accelerating super-resolution network. Ahn et.al. [1] introduced a cascading residual network CARN combining the efficient recursive scheme and multiple residual connections. In view of the above-mentioned methods which is heavily dependent on practical experience, He et.al [8] adopted an ordinary differential equation (ODE)-inspired design scheme to single image super resolution.

2.2 Quantitative Model

In the development of CNN, a great amount of efforts have been explored for model compression, which can speed-up the inference process of deep networks. Recently, published strategies for reducing precision (number of bits used for numerical representation) have achieved significant progress in computer vision tasks. Among the existing methods, Soudry et.al. [22] introduced a variational Bayesian approach to calculate the posterior distribution of the weights (the weights are constrained to +1 or -1). Courbariaux et.al. [4] proposed a BinaryConnect method, which binarizes network weights during the forward and updates the full-precision weights during the backward propagations. Extended from the BinaryConnect, Hubara et.al. [5] proposed a network named BinaryNet, weights and activations in BinaryNet are both binarized. Rastegari et.al. [19] proposed a similar model called XNOR-Net. XNOR-Net includes major steps from the original BNN, but adds a gain term to compensate for lost information during binarization. Zhou et.al. [34] tried to generalize quantization and take advantage of bitwise operations for fixed point data with widths of various sizes, and proposed the DoReFa-Net method. Lin et.al. [16] proposed ABC-Net to reconcile the accuracy gap between BNNs and full-precision networks. Little effort has been spent on model quantization for the image SR task. We design a binarization strategy for the SR task to make the large SR networks being configured on mobile device.

3 Proposed Approach

In this section, we present our proposed BNN-based SISR approach. After presenting the motivation of our approach in Sec. 3.1, we give details on our quantification of weights and activations in Sec. 3.2 and Sec. 3.3 respectively. The proposed binary SR framework is presented in Sec. 3.4.

3.1 Motivation

The process in existing deep learning-based SISR methods could be divided into three stages: feature extraction, nonlinear mapping and image reconstruction. Let x denote the LR input image and y as the final recovered HR image, the function of these models can be formulated as follows,

$$y = \mathcal{R}(\mathcal{M}(\mathcal{E}(x))), \tag{1}$$

where $\mathcal{E}, \mathcal{M}$ and $\mathcal{R}$ are the three aforementioned stages. Generally, $\mathcal{E}$ and $\mathcal{R}$ are composed of only one convolutional layer to realize the transformation from image to deep features and its inverse transformation. $\mathcal{M}$ realizes the mapping process from low-precision features to high-precision features through multiple cascading convolutional layers. The structure of $\mathcal{M}$ will directly determine the model's performance, parameters and computational complexity. Therefore, replacing the $\mathcal{M}$'s full-precision convolution with binary convolution can greatly reduce the model's consumption of computing and storage resources.

In BNN model, the full-precision convolution $W * A$ is estimated by the binary convolution $W^B \oplus A^B$. Where W^B and A^B are the binary weights and activations, $\oplus$ indicates the bitcounting operations. To find an optimal estimation, a straightforward way is to solve the following optimization problem:

$$\begin{aligned} J(W^B, A^B) &= \left\| W * A - W^B \oplus A^B \right\|, \\ W^{B^*}, A^{B^*} &= \underset{W^B, A^B}{arg \min} J(W^B, A^B), \end{aligned} \tag{2}$$

The key to solve this optimization problem lies in the generation of W^B and A^B. In the existing quantitative models, two gain terms α, β are usually introduced to compensate for the lost information during binarization ($W^B = \alpha |W|_{\text{Bin}}$, $A^B = \beta |A|_{\text{Bin}}$). Then the optimal solution (gain terms) can be obtained by calculating or learning from the weights and activations before binarization. However the improvement of convolution precision by gain term is limited, and the back propagation of network gradient is inefficient to update it. In this work, we propose a novel quantization method of filter weights and activations based on bit accumulation mechanism, which achieves better performance for the image SR task.

3.2 Quantization of Weights

The start point of our method is refining the precision of binary filters along the direction of model inference gradually. The processing of model inference is shown in Fig.2, we first accumulate the weight of the each past layer, then use the current weights to offset the accumulated weights in a positive or negative direction, lead to the accumulated weight can be quantified in a more accurate state (-1 or +1). Considering that different convolutional layers have different preferences for image features (color, edge, texture, etc.). We also set a group of combination coefficients α ($\alpha = [\alpha_1, \alpha_2, ..., \alpha_n]$) for the accumulation process.

$$W_n{}^B = Sign(BN(\alpha_1 W_1 + \alpha_2 W_2 + ... + \alpha_n W_n)) * E(|W_n|), \tag{3}$$

where W_n denotes the n_{th} full-precision weights, and $W_n{}^B$ denotes the actual convolution weights of the n_{th} binary filter after updating. $E(|W_n|)$ is the mean of absolute value of each output channel of weights. The utilization of $E(|W_n|)$ could increase the value range of weights and is beneficial to estimate the high-precision binary weights. $BN()$ is the batch normalization operation. $Sign()$ is the symbolic function which transform full-precision values to $+1$ or -1:

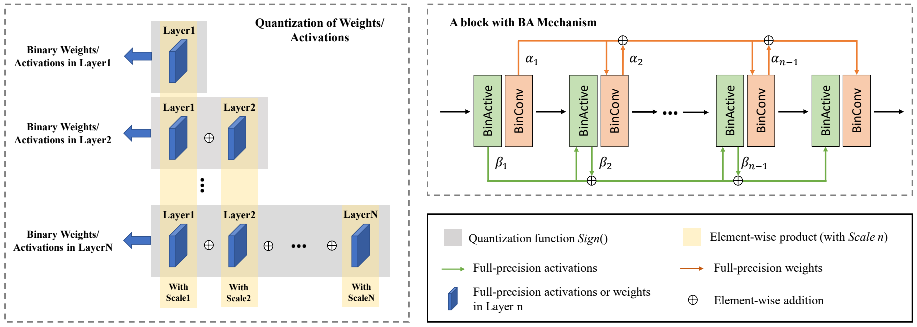

Fig. 2: Bit Accumulation Mechanism (BAM).

$$x^B = Sign(x) \begin{cases} +1, & \text{if} \quad x \geq 0, \\ -1, & \text{otherwise}, \end{cases} \tag{4}$$

An optimal estimation could be find by solving the following optimization problem:

$$\underset{W_n^B}{arg \min} J(W_n^B) = \left\| W_n - W_n^B \right\|^2 . \tag{5}$$

For W_n^B, full-precision weights W_n is given, and the gain term $E(|W_n|)$ could be determined from W_n^B. The optimization error directly depends on the choice of the combination coefficients α. Therefore, the equation (5) can be rewritten as the following optimization problem:

$$\alpha^* = \underset{W_n^B}{arg \min} J(W_n^B) \tag{6}$$

Straight-Through Estimator (STE) is defined for Bernoulli sampling with probability $p \in [0, 1]$, which could be thought of as an operator that has arbitrary forward and backward operations:

$$\begin{aligned} &Forward : q \sim Bernoulli(p). \\ &Backward : \frac{\partial c}{\partial p} = \frac{\partial c}{\partial q}. \end{aligned} \tag{7}$$

Here we adopt the STE method for back-propagate through W^B. Assume c as the cost function, A_o^B as the output tensor of a convolution respectively, the gradient at n_{th} filter in back propagation can be computed as follows:

$$\frac{\partial c}{\partial W_n} = \frac{\partial c}{\partial A_o} \frac{\partial A_o}{\partial {W_n}^B} \frac{\partial {W_n}^B}{\partial W_n} \overset{STE}{\approx} \frac{\partial c}{\partial A_o} \frac{\partial A_o}{\partial {W_n}^B} = \frac{\partial c}{\partial {W_n}^B}. \tag{8}$$

Bear in mind that during training, the full-precision weights are reserved and updated at every epoch, while in test-time only binary weights are used in convolution. The training details of our model are summarized in Algorithm 1.

Algorithm 1 Training a N-layers CNN-block with BAM:

Input: Full-precision activations A_1 and weights W from the N filters ($W = [W_1, W_2, ..., W_N]$), activation combination coefficients $\beta_1, \beta_2, ..., \beta_N$, cost function c and current learning rate η^t.
Output: updated weight W^{t+1}, updated activation combination coefficients $\beta_1, \beta_2, ..., \beta_N$, and updated learning rate η^{t+1}.
1: **for** $n = 1$ to N **do**
2: $A_n^B \leftarrow$ calculate the equation (9)
3: $\alpha \leftarrow$ solve the equation (6)
4: $W_n^B \leftarrow$ calculate the equation (3)
5: $A_{n+1} = BinaryConv(W_n^B, A_n^B)$
6: **end for**
7: $g^w \leftarrow$ **BinaryBackward**($\frac{\partial c}{\partial A_o}, W^B$), computed using the equation (8)
8: **for** $n = N$ to 1 **do**
9: $g_n^w \leftarrow$ **BackwardGradient**(g^w, W_n)
10: $\beta_n =$ **UpdateParameters**(${\beta_n}^{scale}, g_{\beta_n^{scale}}, \eta^t$)
11: $W_n^{t+1} =$ **UpdateParameters**(W_n^t, g_n^w, η^t)
12: **end for**
13: $\eta^{t+1} =$ **UpdateLearningrate**($\eta^t; t$)

3.3 Quantization of Activations

The operation of floating-point convolution could be implemented without multiplications when weights are binarized, but the computation is still much more than the bitcounting operation. Next we detail our approach to getting binary activations that are input to convolutions, which is of critical importance in replacing floating-point convolutions by bitcounting-based convolutions.

The proposed bit accumulation mechanism is not only applicable to the quantification of weights, but also applicable to the quantization of activations ass shown in Fig.2. Here we also make linear combination of the past multi-layer's activations to approximate the current layer's activations.

$$ {A_n}^B \approx Sign(BN(\beta_1 A_1 + \beta_2 A_2 + ... + \beta_n A_n)), \tag{9} $$

Considering that the process of image super resolution depends heavily on the details of the activation, the rough quantification ($Sign()$) of the activation can lead to a sharp decline in model performance. Instead of setting the combination weights of the activations to a single scale, we use a two-dimensional array to represent these weights.

$$ \beta_n \leftarrow Update({\beta_n}^{scale} E(A_n), g_{\beta_n^{scale}}, \eta) \tag{10} $$

where ${\beta_n}^{scale}$ is a scale factor, and only ${\beta_n}^{scale}$ will be updated during the training process. $g_{\beta_n^{scale}}$ and η are the gradient and learning rate. $E(A_n)$ is the activation statistics along the channel dimension, which could reflect the spatial information of the input image.

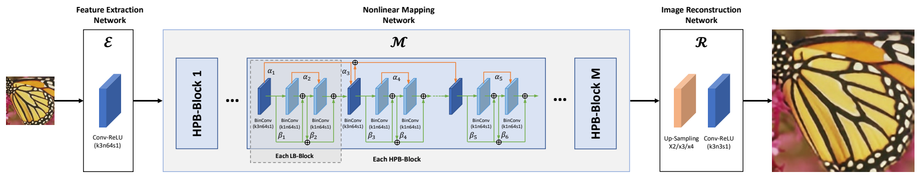

Fig. 3: Network architecture of our proposed BSRN.

3.4 Binary Super Resolution Network

As described in section 3.1, our binary super resolution network (BSRN) also consists of three subnetworks: a feature extraction network $\mathcal{E}$, a nonlinear mapping network $\mathcal{M}$ and an image reconstruction network $\mathcal{R}$. $\mathcal{E}$ contains only one convolutional layer and $\mathcal{R}$ contain only one convolutional layer, while $\mathcal{R}$ contains an additional upsampling layer. Consistent with existing binarization methods, all convolution operations in $\mathcal{E}$ and $\mathcal{R}$ are full-precision.

Particularly, we construct a high-precision binarization (HPB) block structure to form our nonlinear mapping network $\mathcal{M}$, which could achieve the approximation of high precision convolution by accumulation the multiple binary convolution. The details are shown in Fig.3. HPB block is also composed of several cascaded local binarization (LB) block, and each LB-block consists of one 3×3 convolution layer and two 1×1 convolution layers. The novelty of HPB block lies in two ranges of bit accumulation. Firstly, the short-range accumulation exists in each LB-block. The filters and activations of intermediary layers are accumulated into the higher layers, and finally converge on the last 1×1 convolution layer. Then, the long-range accumulation occurs in the first convolutional layer of each LB-block, and only filters are accumulated in the same way as the short-range accumulation.

4 Experiments

4.1 Datasets

DIV2K is a high quality image dataset which consists of 800 training images, 100 validation images and 100 testing images and it is widely used for super-resolution in recent years. Following[14, 1, 3, 8], we use 800 training images from DIV2K dataset[26] as training set. To illustrate the performance of our proposed methods, we conduct test experiments on four standard datasets, Set5[2], Set14[31] , BSD100[18] and Urban100[9] which includes 5, 14, 100 and 100 images respectively.

Three upscaling factors are evaluated, including $\times 2$, $\times 3$ and $\times 4$. The input LR image is generated by bicubic down sampling the HR image with $\times 2$, $\times 3$ and $\times 4$ scale respectively. The size of input patches are 48×48, and the output patch size is 96×96, 144×144 and 192×192 for $\times 2$, $\times 3$ and $\times 4$ upscaling factor respectively.

4.2 Implementations

Data augmentation is performed on aforementioned 800 training images, which are randomly rotated by 90°, 180°, 270° and flipped horizontally. Batch size is set to 16. Adam is utilized to optimize the network. The momentum parameter is set to 0.5, weight decay is set to 2×10^{-4}, and the initial learning rate is set to 1×10^{-4} and will be divided a half every 200 epochs. All of our models are implemented under the PyTorch environment with Python 3.6 on Ubuntu 16.04 system with a 12G NVIDIA Titan Xp GPU.

For assessing the quality of SR results, we employ two objective image quality assessment metrics: Peak Signal to Noise Ratio (PSNR) and structural similarity (SSIM)[29]. All metrics are performed on the Y-channel (YCbCr color space) of center-cropped, removaling of a s-pixel wide strip from each $\times s$ upscaling image border.

4.3 Evaluation

We choose two simple and practical super-resolution networks to evaluate model quantization methods: VDSR[10] and SRResNet[13]. VDSR and SRResNet can be regarded as the most typical methods of convolution-cascade model and block-cascade model respectively. Experimental evaluation on these two networks can more intuitively reflect the performance of model quantization. The performance evaluation of the proposed bit accumulation mechanism (BAM) is carry out on these two SR network, the compared methods including BNN[5] DoReFa-Net[34] and ABC-Net[16].

Evaluation on VDSR: Considering that the operation of our bit accumulation method needs to work in multiple convolutional layers, we divide the middle 18 convolutional layers of VDSR into 6 blocks on average. Each block contains three convolutional layers. The weight and activation accumulation process in each convolution block is consistent with sections 3.2 and 3.3. Other methods quantify each convolutional layer at the middle of VDSR separately. One should be noted is that ABC-Net simulates a full-precision convolution through multiple one-bit convolutions. Here we set the number of one-bit convolution to 3, that is, the model size of VDSR_ABC is nearly three times larger than other methods.

Table.1 shows the quantitative comparisons of the performances over the benchmark datasets. For binarization of weight and activation, there is no compensation process for quantization error in VDSR_BNN and VDSR_DoReFa methods. The information carried in the activations is limited, and the image's high frequency details are difficult to predict. These two methods achieve inferior performance than VDSR_ABC and our proposed VDSR_BAM. VDSR_ABC approximates the full-precision convolution by linear combination of multiple binary convolutions. Its performance is significantly higher than VDSR_DoReFa. However, the single convolution process in VDSR_ABC is not significantly improved compared with that of VDSR_DoReFa. Besides, the model parameters and computational operands of VDSR_ABC are 3 times (in this work) higher

Methods	Scale	Set5		Set14		B100		Urban100	
		PSRN	SSIM	PSRN	SSIM	PSRN	SSIM	PSRN	SSIM
VDSR	×2	37.53	0.959	33.05	0.913	31.90	0.896	30.77	0.914
Bicubic	×2	33.66	0.930	30.24	0.869	29.56	0.843	26.88	0.840
VDSR_BNN	×2	34.43	0.936	30.94	0.882	30.05	0.856	27.54	0.860
VDSR_DoReFa	×2	34.70	0.933	31.22	0.876	30.25	0.849	28.25	0.865
VDSR_ABC	×2	35.35	0.939	31.71	0.886	30.68	0.861	28.77	0.878
VDSR_BAM	×2	**36.60**	**0.953**	**32.41**	**0.905**	**31.32**	**0.886**	**29.43**	**0.895**
VDSR	×3	33.66	0.921	29.77	0.831	28.82	0.798	27.14	0.828
Bicubic	×3	30.39	0.868	27.55	0.774	27.21	0.739	24.46	0.735
VDSR_BNN	×3	31.01	0.874	28.15	0.791	27.57	0.755	25.01	0.758
VDSR_DoReFa	×3	31.79	0.895	28.68	0.806	27.98	0.766	25.53	0.782
VDSR_ABC	×3	32.01	0.898	28.86	0.808	28.08	0.770	25.80	0.787
VDSR_BAM	×3	**32.52**	**0.907**	**29.17**	**0.819**	**28.29**	**0.782**	**26.07**	**0.799**
VDSR	×4	31.35	0.884	28.01	0.767	27.29	0.725	25.18	0.752
Bicubic	×4	28.42	0.810	26.00	0.703	25.96	0.668	23.14	0.658
VDSR_BNN	×4	29.02	0.827	26.55	0.724	26.29	0.685	23.55	0.685
VDSR_DoReFa	×4	29.39	0.837	26.79	0.728	26.45	0.689	23.81	0.696
VDSR_ABC	×4	29.59	0.841	29.63	0.730	26.51	0.687	23.96	0.699
VDSR_BAM	×4	**30.31**	**0.860**	**27.46**	**0.749**	**26.83**	**0.706**	**24.38**	**0.720**

Table 1: Quantitative evaluation of VDSR-based state-of-the-art model quantization methods.

than other methods. Benefit from BAM's ability to retain and compensate for lost information during the quantization process, our VDSR_BAM exceeds all previous methods on four benchmark datasets.

Evaluation on SRResNet: SRResNet is a modular (residual-block) network structure. Then our BAM can be directly applied to each module. The setup of other model quantization methods is consistent with evaluation on VDSR. Each residual block in SRResNet_ABC contains six one-bit convolutions. The comparison results are shown in Table.2.

Compared with VDSR-based models, the performance of model quantization methods with SRResNet has been significantly improved. This is not only attributed to the stronger learning ability of nonlinear mapping of the residual block, but also to its highly gain brought by the increase of model parameters. Especially, the performance of SRResNet_DoReFa improved significantly. The main reason is that the updating and generating modes of weights and activations in VDSR is monotonous, especially under the interference of binaryzation function $Sign()$. The difficulty of weights updating in gradient back propagation is greatly increased. While the gain of the residuals to the gradient back propagation is beyond doubt, and its skip connection also can effectively enrich the

Methods	Scale	Set5		Set14		B100		Urban100	
		PSRN	SSIM	PSRN	SSIM	PSRN	SSIM	PSRN	SSIM
SRResNet	×2	37.76	0.958	33.27	0.914	31.95	0.895	31.28	0.919
Bicubic	×2	33.66	0.930	30.24	0.869	29.56	0.843	26.88	0.840
SRResNet_BNN	×2	35.21	0.942	31.55	0.896	30.64	0.876	28.01	0.869
SRResNet_DoReFa	×2	36.09	0.950	32.09	0.902	31.02	0.882	28.87	0.880
SRResNet_ABC	×2	36.34	0.952	32.28	0.903	31.16	0.884	29.29	0.891
SRResNet_BAM	×2	**37.21**	**0.956**	**32.74**	**0.910**	**31.60**	**0.891**	**30.20**	**0.906**
SRResNet	×3	34.07	0.922	30.04	0.835	28.91	0.798	27.50	0.837
Bicubic	×3	30.39	0.868	27.55	0.774	27.21	0.739	24.46	0.735
SRResNet_BNN	×3	31.18	0.877	28.29	0.799	27.73	0.765	25.03	0.758
SRResNet_DoReFa	×3	32.44	0.903	28.99	0.811	28.21	0.778	25.84	0.783
SRResNet_ABC	×3	32.69	0.908	29.24	0.820	28.35	0.782	26.12	0.797
SRResNet_BAM	×3	**33.33**	**0.915**	**29.63**	**0.827**	**28.61**	**0.790**	**26.69**	**0.816**
SRResNet	×4	31.76	0.888	28.25	0.773	27.38	0.727	25.54	0.767
Bicubic	×4	28.42	0.810	26.00	0.703	25.96	0.668	23.14	0.658
SRResNet_BNN	×4	29.33	0.826	26.72	0.728	26.45	0.692	23.68	0.683
SRResNet_DoReFa	×4	30.38	0.862	27.48	0.754	26.87	0.708	24.45	0.720
SRResNet_ABC	×4	30.78	0.868	27.71	0.756	27.00	0.713	24.54	0.729
SRResNet_BAM	×4	**31.24**	**0.878**	**27.97**	**0.765**	**27.15**	**0.719**	**24.95**	**0.745**

Table 2: Quantitative evaluation of SRResNet-based state-of-the-art model quantization methods.

information of the activations in forward inference processing. In some cases, *e.g.* x3 on B100, x4 on Set14 and x4 on B100, the gap between our SRResNet_BAM and the full-precision SRResNet is shrunk to no more than 0.3dB.

Qualitative Evaluation: As demonstrated in Fig.4, we present a subjective comparison with state-of-the-art model quantization methods based on VDSR. We enlarge the texture on the general real-world images to compare the subjective visual effects of different SR methods. It is obvious that the compared SR methods fails to extract realistic details from LR inputs and they are prone to produce a blurry texture. Our model could reveal the most accurate and realistic details and generates the correct direction of texture.

Fig.5 presents the results of different quantization SR methods based on SRResNet. Parts including lines or holes in the buildings are magnified for more obvious comparison. It is observed that after 4 minification and magnification by bicubic, the direction of the line and the outline of the hole is hard to be distinguished. Most methods can effectively recover the lines in the regions close to the shooting point. However, most of them can do nothing for the regions far from the shooting point. Benefit from the retention of information in the BAM

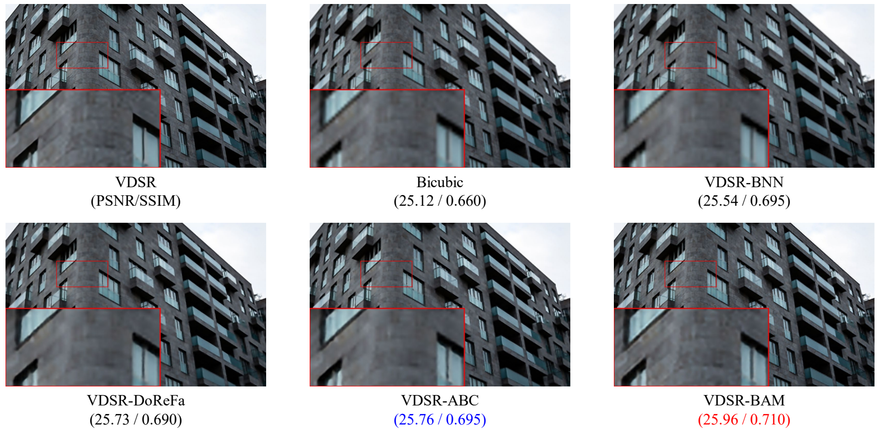

Fig. 4: VDSR-based subjective quality assessment for $\times 4$ upscaling on the structured image: $img001$ from Urban100.

quantization process, our proposed SRResNet_BAM method could recover more accurate details for these regions.

4.4 Model Analysis

In this section, we first evaluated the performance of our proposed BSRN model. Then we investigate the effects of different quantization approaches. Finally, we compare our method with existing quantized convolution-based super resolution method [17].

Evaluation on Proposed BSRN Framework: The purpose of our work is to reduce model parameters and improve the reasoning speed so that it can be applied to mobile devices such as mobile phones. We restrict our study to the binary networks with a low number of parameters and do not further investigate potentially beneficial extensions such as width and depth of network [32] or different loss functions [12, 13].

In our experiment, we set the number of HPB-Block (see Fig.3) as 20. Each HPB-Block contains 2 LB-Blocks. The number of feature channels is set to 48. Then, based on our proposed BAM quantization method, BSRN is compared with the other two super-resolution networks (VDSR and SRResNet). The results are shown in Table.3. It is obvious that our model can achieve superior performance with lower model parameters and calculation operands.

Ablation Study on Quantization of Weights and Activations: Table.4 presents the ablation study on the effect of our BAM quantization method. In this table, BSRN-W1 is BSRN without linear combination coefficients in the process of weight quantization, that is, $\alpha_1 = \alpha_2 = ... = \alpha_n = 1$. BSRN-W2

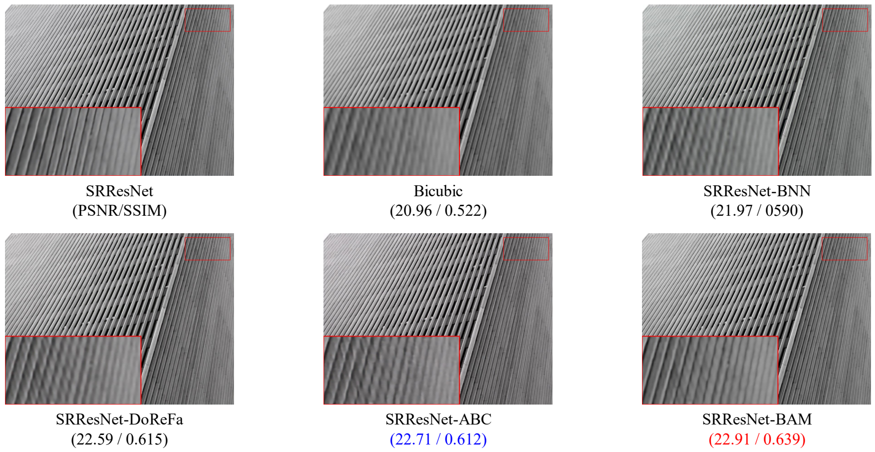

Fig. 5: SRResNet-based subjective quality assessment for $\times 4$ upscaling on the structured image: $img045$ from Urban100.

Methods	Paras	MAC	Set5		Set14		B100		Urban100	
			PSRN	SSIM	PSRN	SSIM	PSRN	SSIM	PSRN	SSIM
VDSR_BAM	$668K$	$616.9G$	30.31	0.860	27.46	0.749	26.83	0.706	24.38	0.720
SRResNet_BAM	$1547K$	$127.9G$	31.24	0.878	27.97	0.765	27.15	0.719	24.95	0.745
BSRN	$1216K$	$85G$	**31.35**	**0.880**	**28.04**	**0.768**	**27.18**	**0.720**	**25.11**	**0.749**

Table 3: Comparison between VDSR_BAM and SRResNet_BAM and BSRN on $\times 4$. MAC is the number of multiply-accumulate operations. We assume that the generated SR image is 720P (1280×720).

refers to the process of weight quantization of BSRN without bit accumulation operation, *i.e.* ${W_n}^B = Sign(BN(\alpha W_n)) * E(|W_n|)$. Corresponding to BSRN-W1 and BSRN-W2, BSRN-A1 and BSRN-A2 quantize the activations according to the same way.

From the results, we can see that the processing mode of activation has a great impact on the performance of the model, which also indicates that for the image super-resolution task, the ability to activate information can directly determine the performance of the model. In addition, the bit accumulation operation has a greater gain on model performance than the combination coefficients. This allows us to achieve a satisfactory performance with less parameters and without combination coefficients.

Comparison with Ma et.al [17]: Recently, Ma et al. [17] proposed an image super-resolution work based on model quantization. It is not a complete binary neural network, but a network model with a binary filters and full-precision activations. In this section, we evaluate the performance of the our model under the binary filters and with full-precision activations. The results are shown in Table.5. The superiority of our performance can be clearly seen from the results.

Models	BSRN-W1	BSRN-W2	BSRN-A1	BSRN-A2	BSRN
PSNR	31.25	31.14	31.22	30.92	31.35

Table 4: Effects of the quantization of weights and activations measured on the Set5 $\times 4$ dataset.

Methods	Scale	Set5		Set14		Urban100	
		PSRN	SSIM	PSRN	SSIM	PSRN	SSIM
SRResNet_Ma et.al.	$\times 2$	35.66	0.946	31.56	0.897	28.76	0.882
SRResNet_BAM	$\times 2$	**37.51**	**0.956**	**33.03**	**0.912**	**30.79**	**0.915**
SRResNet_Ma et.al.	$\times 4$	30.34	0.864	27.16	0.756	24.48	0.728
SRResNet_BAM	$\times 4$	**31.57**	**0.883**	**28.16**	**0.769**	**24.30**	**0.755**

Table 5: Quantitative evaluation with the work of Ma et.al on SRResNet.

In general, the experiments not only illustrate the effectiveness of design binarization method but suggest the reasonability of bit-accumulation perspective.

5 Conclusions

In this paper, we proposed a BNN-based model for SISR, in which a novel binarization method named bit-accumulation mechanism and a lightweight network structure are designed to approximate the full-precision CNN. The evaluation and analysis in this paper indicates that the presented method can gradually refine the precision of quantization along the direction of model inference and significantly improve the model performance. Extensive experiments compared with the state-of-the-art methods demonstrated the superiority of the our proposed binarization method. We believe that this bit-accumulation mechanism could be more widely applicable in practice. Moreover, it is readily to be used in other machine vision problems, especially image reconstruction related tasks, such as image/video deblurring, compression artifact removal, image/video restoration and so on.

Acknowledgement

This work was supported in part by the National Key Research and Development Program of China under Grant 2018AAA0103202, in part by the National Natural Science Foundation of China under Grant 61922066, Grant 61876142, Grant 61671339, Grant 61772402, Grant U1605252, and Grant 61976166, in part by the National High-Level Talents Special Support Program of China under Grant CS31117200001, in part by the Fundamental Research Funds for the Central Universities under Grant JB190117, in part by the Xidian University Intellifusion Joint Innovation Laboratory of Artificial In- telligence, and in part by the Innovation Fund of Xidian University.

References

1. Ahn, N., Kang, B., Sohn, K.A.: Fast, accurate, and lightweight super-resolution with cascading residual network. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 252–268 (2018)
2. Bevilacqua, M., Roumy, A., Guillemot, C., Alberi-Morel, M.L.: Low-complexity single-image super-resolution based on nonnegative neighbor embedding. BMVA press (2012)
3. Choi, J.S., Kim, M.: A deep convolutional neural network with selection units for super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 154–160 (2017)
4. Courbariaux, M., Bengio, Y., David, J.P.: Binaryconnect: Training deep neural networks with binary weights during propagations. In: Advances in neural information processing systems. pp. 3123–3131 (2015)
5. Courbariaux, M., Hubara, I., Soudry, D., El-Yaniv, R., Bengio, Y.: Binarized neural networks: Training deep neural networks with weights and activations constrained to+ 1 or-1. arXiv preprint arXiv:1602.02830 (2016)
6. Dong, C., Loy, C.C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence **38**(2), 295–307 (2015)
7. Dong, C., Loy, C.C., Tang, X.: Accelerating the super-resolution convolutional neural network. In: European conference on computer vision. pp. 391–407. Springer (2016)
8. He, X., Mo, Z., Wang, P., Liu, Y., Yang, M., Cheng, J.: Ode-inspired network design for single image super-resolution. In: 2019 IEEE Conference on Computer Vision and Pattern Recognition (July 2019)
9. Huang, J.B., Singh, A., Ahuja, N.: Single image super-resolution from transformed self-exemplars. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5197–5206 (2015)
10. Jiwon, K., Jung Kwon, L., Kyoung Mu, L.: Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1646–1654 (2016)
11. Kim, J., Kwon Lee, J., Mu Lee, K.: Deeply-recursive convolutional network for image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1637–1645 (2016)
12. Lai, W.S., Huang, J.B., Ahuja, N., Yang, M.H.: Deep laplacian pyramid networks for fast and accurate super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 624–632 (2017)
13. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4681–4690 (2017)
14. Li, J., Fang, F., Mei, K., Zhang, G.: Multi-scale residual network for image super-resolution. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 517–532 (2018)
15. Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K.: Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp. 136–144 (2017)
16. Lin, X., Zhao, C., Pan, W.: Towards accurate binary convolutional neural network. In: Advances in Neural Information Processing Systems. pp. 345–353 (2017)

17. Ma, Y., Xiong, H., Hu, Z., Ma, L.: Efficient super resolution using binarized neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 0–0 (2019)
18. Martin, D., Fowlkes, C., Tal, D., Malik, J., et al.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Iccv Vancouver: (2001)
19. Rastegari, M., Ordonez, V., Redmon, J., Farhadi, A.: Xnor-net: Imagenet classification using binary convolutional neural networks. In: European conference on computer vision. pp. 525–542. Springer (2016)
20. Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z.: Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1874–1883 (2016)
21. Shi, W., Caballero, J., Ledig, C., Zhuang, X., Bai, W., Bhatia, K., de Marvao, A.M.S.M., Dawes, T., O Regan, D., Rueckert, D.: Cardiac image super-resolution with global correspondence using multi-atlas patchmatch. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 9–16. Springer (2013)
22. Soudry, D., Hubara, I., Meir, R.: Expectation backpropagation: Parameter-free training of multilayer neural networks with continuous or discrete weights. In: Advances in Neural Information Processing Systems. pp. 963–971 (2014)
23. Tai, Y., Yang, J., Liu, X.: Image super-resolution via deep recursive residual network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3147–3155 (2017)
24. Tai, Y., Yang, J., Liu, X., Xu, C.: Memnet: A persistent memory network for image restoration. In: Proceedings of the IEEE international conference on computer vision. pp. 4539–4547 (2017)
25. Thornton, M.W., Atkinson, P.M., Holland, D.: Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping. International Journal of Remote Sensing **27**(3), 473–491 (2006)
26. Timofte, R., Agustsson, E., Van Gool, L., Yang, M.H., Zhang, L.: Ntire 2017 challenge on single image super-resolution: Methods and results. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 114–125 (2017)
27. Tong, T., Li, G., Liu, X., Gao, Q.: Image super-resolution using dense skip connections. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 4799–4807 (2017)
28. Wang, Z., Liu, D., Yang, J., Han, W., Huang, T.: Deep networks for image super-resolution with sparse prior. In: Proceedings of the IEEE international conference on computer vision. pp. 370–378 (2015)
29. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., et al.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4), 600–612 (2004)
30. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. IEEE transactions on image processing **19**(11), 2861–2873 (2010)
31. Zeyde, R., Elad, M., Protter, M.: On single image scale-up using sparse-representations. In: International conference on curves and surfaces. pp. 711–730. Springer (2010)

32. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 286–301 (2018)
33. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2472–2481 (2018)
34. Zhou, S., Wu, Y., Ni, Z., Zhou, X., Wen, H., Zou, Y.: Dorefa-net: Training low bitwidth convolutional neural networks with low bitwidth gradients. arXiv preprint arXiv:1606.06160 (2016)
35. Zou, W.W., Yuen, P.C.: Very low resolution face recognition problem. IEEE Transactions on image processing **21**(1), 327–340 (2011)