25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

JIANGNAN UNIVERSITY

A Dual-branch Network for Infrared and Visible Image Fusion

Yu Fu， Xiao-Jun Wu
Jiangsu Provincial Engineering
Laboratory of Pattern Recognition and
Computational Intelligence,
Jiangnan University,
Wuxi, China, 214122
Email: yu_fu_stu@outlook.com

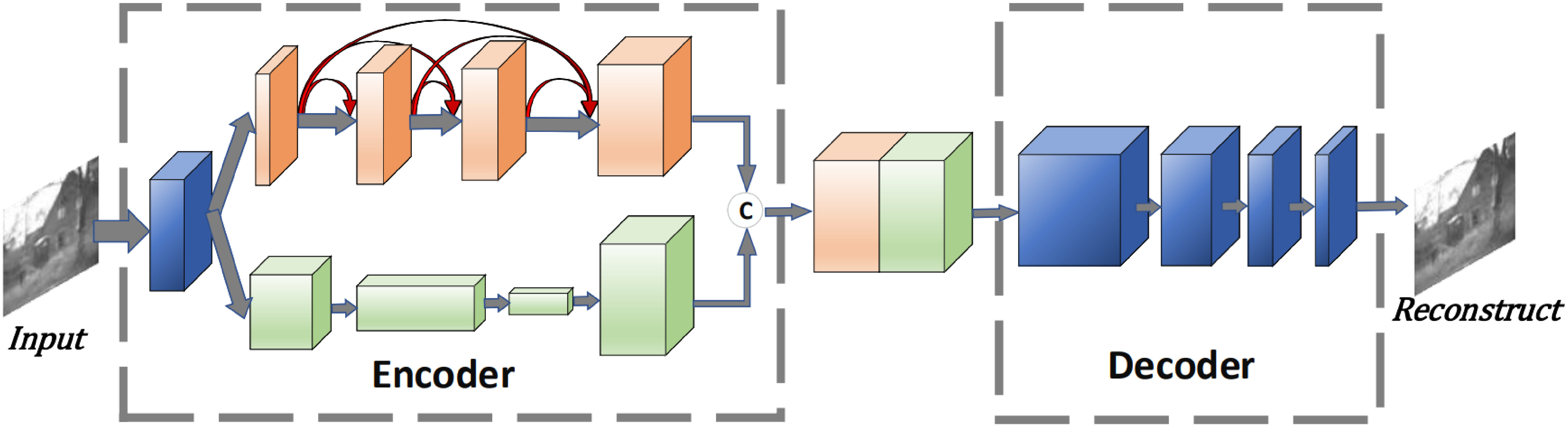
Task Definition

Infrared image

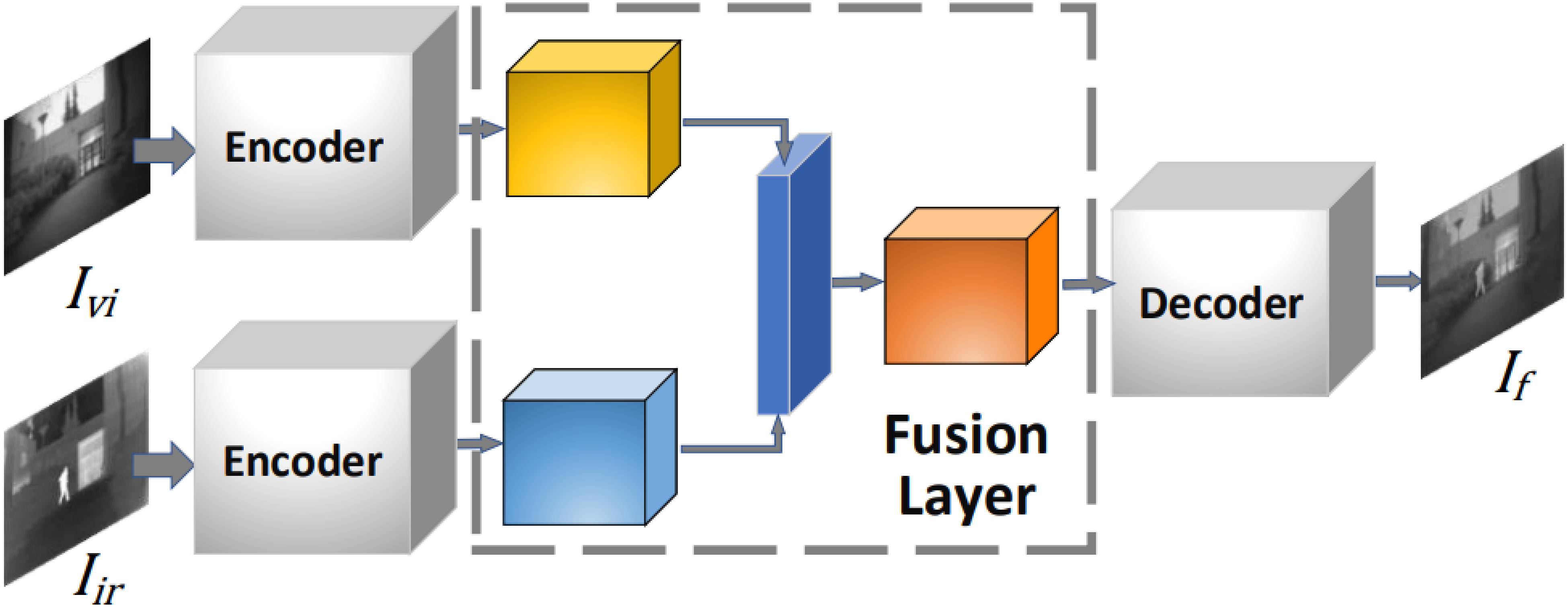
Visible image

Fused image

Framework



Test phase



Loss Function

$$L = L_{pixel} + \alpha L_{gradient} + \beta L_{color} + \gamma L_{perceptual}$$

Loss Function

L_{pixel} loss function can calculate the pixel error of the reconstructed image and the input image, the formula is given as follows:

$$L_{pixel} = MSE(I_{re}, I_{in})$$

$$MSE(X,Y) = \frac{1}{N}\sum_{n=1}^{N}(X_n - Y_n)^2$$

where I_{re} is the reconstructed image, I_{in} is the input image, and $MSE(x,y)$ is the mean square error function of x and y.

Loss Function

$L_{gradient}$ loss function can calculate the edge information loss of the reconstructed image and the input image, the formula is presented as follows:

$$L_{gradient} = MSE(Gradient(I_{re}), Gradient(I_{in}))$$

where, $Gradient(x)$ is the image sharpening using the Laplace operator to obtain the gradient map. The Laplace operator performs a mathematical convolution operation. Its definition and approximate discrete expression is given as follows:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}$$

$$\approx [f(x+1,y) + f(x-1,y) + f(x,y+1) f(x,y-1)] - 4f(x,y)$$

Loss Function

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

$$L_{color} = \frac{1}{255} \| Histogram(I_{re}) - Histogram(I_{in}) \|_2$$

where $Histogram(x)$ is the color histogram of x. We set the number of the histogram bins to 255, the maximum and minimum values of the histogram calculation are the maximum and minimum values between I_{re} and I_{in}. We calculate the norm of the difference between the two histograms.

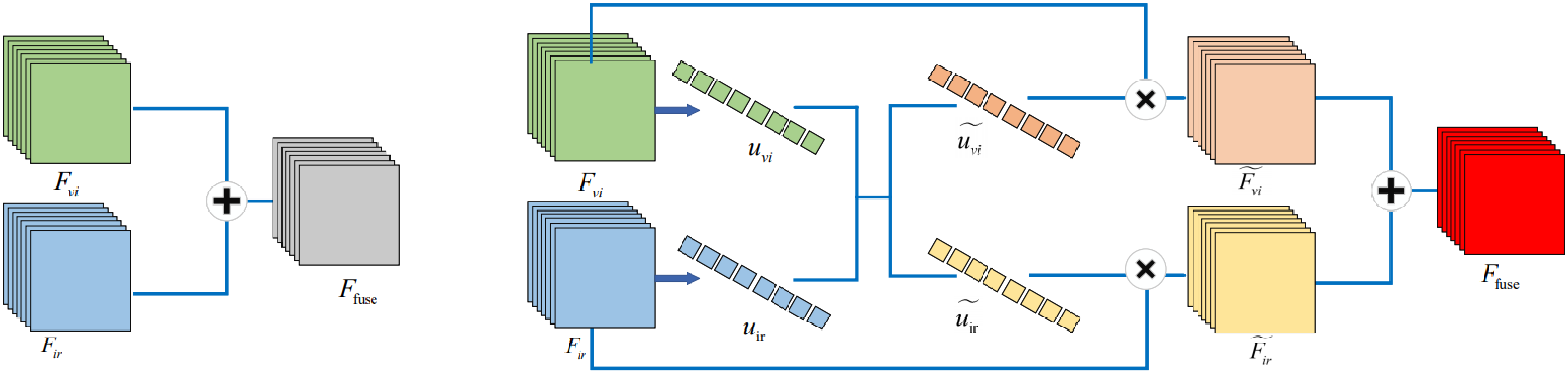
Loss Function

$L_{perceptual}$ loss function[17] can calculate the errors of features between the reconstructed image and the input image.

$$L_{perceptual} = \sum_{i=1}^{4} MSE(\phi_i(I_{re}), \phi_i(I_{in}))$$

where $\phi_i(x)$ represents the features of the ith layer obtained by inputting images I into a specific trained network. In this paper, we choose pre-trained vgg19 network as the feature extraction network.

Fusion Strategy



(a) addition strategy

(b) channel strategy

$$F_{fuse}(x,y) = F_{ir}(x,y) + F_{vi}(x,y)$$

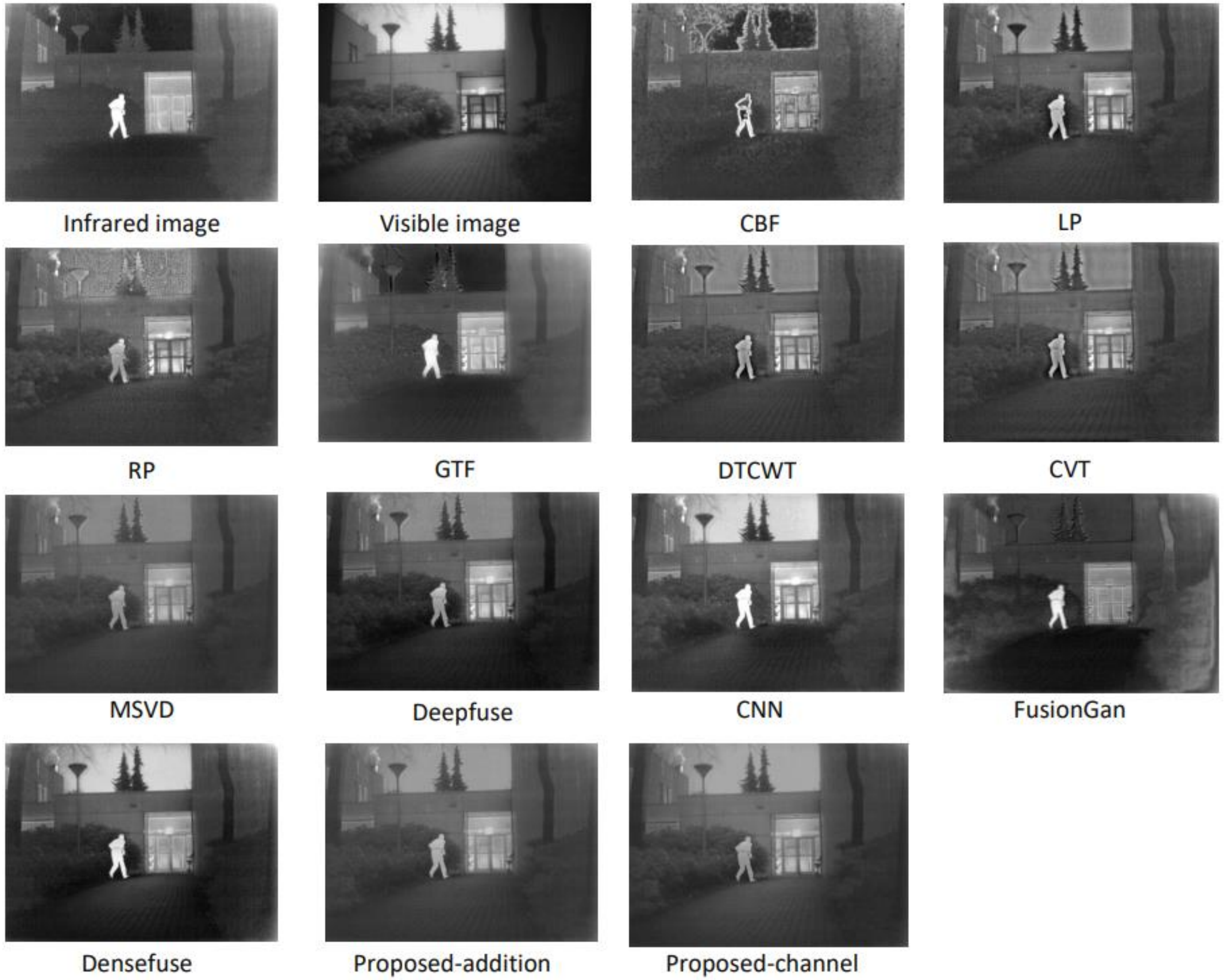
$$u_i = \frac{1}{H \times W} \sum_{m=1}^{H} \sum_{n=1}^{W} x_i(m,n)$$

$$\tilde{u}_{ir} = \frac{u_{ir}}{u_{ir} + u_{vi}} \ , \ \tilde{u}_{vi} = 1 - \tilde{u}_{ir}$$

$$\tilde{F}_{ir} = \tilde{u}_{ir} F_{ir} \ , \qquad \tilde{F}_{vi} = \tilde{u}_{vi} F_{vi}$$

$$F_{fuse} = \tilde{F}_{ir} + \tilde{F}_{vi}$$

Experiments



Experiments

	EI	SF	EN	SSIM	Nabf	MI
CBF	***52.3045***	***13.0921***	6.8275	0.6142	0.2669	13.6550
LP	44.7055	11.5391	6.6322	0.7037	0.1328	13.2644
RP	44.9054	12.7249	6.5397	0.6705	0.1915	13.0794
GTF	35.0073	9.5022	6.5781	0.6798	0.0710	13.1562
DTCWT	42.6159	11.1913	6.4799	0.6939	0.1428	12.9599
CVT	43.1558	11.2006	6.5005	0.6907	0.1644	13.0011
MSVD	27.9727	8.9758	6.2869	***0.7219***	***0.0378***	12.5738
Deepfuse	33.8768	8.3500	6.6102	0.7135	0.0610	13.2205
CNN	44.8334	11.6483	**7.0629**	0.6955	0.1286	**14.1259**
Fusion	30.6847	7.5492	6.5299	0.6211	0.1344	13.0597
Densefuse	36.4838	9.3238	6.8526	0.7108	0.0890	13.7053
ours-addition	25.3765	6.2758	6.2691	**0.7593**	**0.0010**	12.5382
ours-channel	**59.2286**	**21.9070**	***6.8793***	0.5990	0.1958	***13.7586***

Conclusion

- We propose a novel and effective deep learning network for infrared and visible light image fusion.
- We design a new fusion strategy to fuse two sets offeatures according to the importance of the channel.
- We designed a new loss function.
- Our proposed method has obtained the best value or the second best value on the six objective indicators.