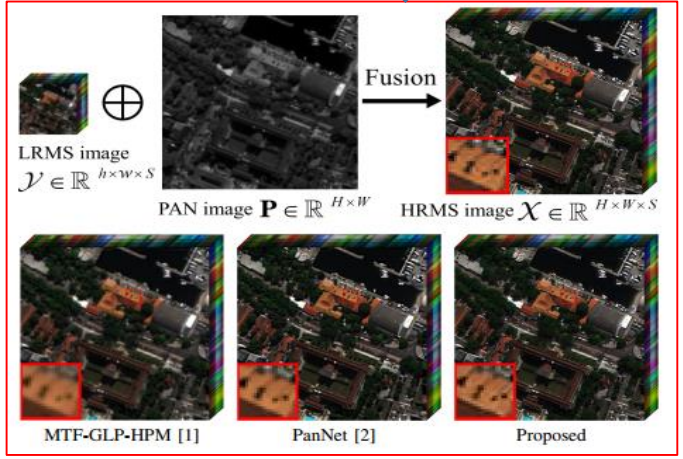
Model: (General form)

$$\min_{\mathbf{X}} f_{spec}(\mathbf{X},\mathbf{Y}) + \lambda f_{spat}(\mathbf{X},\mathbf{P}) + \alpha f_{PDI}(\mathbf{X},\mathbf{X}_{net}), \quad (2)$$

$$\min_{\mathbf{X}} \|\mathbf{XBS}-\mathbf{Y}\|_F^2 + \lambda \left\|\nabla\mathbf{X} - \nabla\widetilde{\mathbf{P}}\right\|_{2,1} + \alpha \|\mathbf{X}-\mathbf{X}_{net}\|_F^2$$

Algorithm: (ADMM)



Quantitative: (Simulated WV3, *Xnet* trained on WV3 and testd on WV3)

TABLE I
QUALITY METRICS FOR ALL THE COMPARED APPROACHES ON THE REDUCED RESOLUTION (SIMULATED) RIO AND TRIPOLI DATASETS, RESPECTIVELY (BOLD: BEST; UNDERLINE: SECOND BEST)

Method	*(a) Rio*					*(b) Tripoli*					Average
	ERGAS	*SAM*	*Q8*	*PSNR*	*SSIM*	*ERGAS*	*SAM*	*Q8*	*PSNR*	*SSIM*	time(s)
EXP	8.9481	6.3662	0.6775	29.3264	0.7872	4.8097	4.1088	0.8173	27.7309	0.7104	**0.03**
MTF-GLP-HPM	6.3748	5.7967	0.7996	31.0209	0.8867	2.9489	3.9725	0.9301	31.9819	0.8876	0.29
GLP-Reg-FS	6.6462	6.2046	0.7867	31.7901	0.8746	2.9339	3.8925	0.9312	31.9857	0.8900	0.38
CVPR19	7.0125	5.6664	0.7681	31.4024	0.8701	3.5465	3.8256	0.9017	30.4052	0.8515	20.26
DiCNN	3.8738	4.2262	0.8600	36.4007	0.9526	2.0734	3.1568	0.9649	34.8528	0.9390	0.22
PanNet	3.8588	4.4518	0.8643	36.3447	0.9529	2.1272	3.1887	0.9648	34.5557	0.9367	0.41
Proposed	**3.6266**	**4.0229**	**0.8718**	**36.7143**	**0.9556**	**1.9491**	**3.0014**	**0.9694**	**35.1881**	**0.9413**	4.22
Ideal value	**0**	**0**	**1**	**+∞**	**1**	**0**	**0**	**1**	**+∞**	**1**	-

WV3 Results: (Simulated, *Xnet* trained on WV3 and testd on WV3)

Fig. 4. The fusion results on the reduced resolution (simulated) Rio dataset (source: WorldView-3). Top row: the visual performance of the EXP, MTF-GLP-HPM, GLP-Reg-FS, CVPR19, DiCNN, PanNet, Proposed method, and the ground-truth (GT) image, respectively. Bottom row: the corresponding MAE maps using the GT image as reference. For a better visualization, we doubled the intensities of the MAE maps and added 0.5.

Data Generalization: (Simulated, *Xnet* trained on WV3 but tested on WV2)

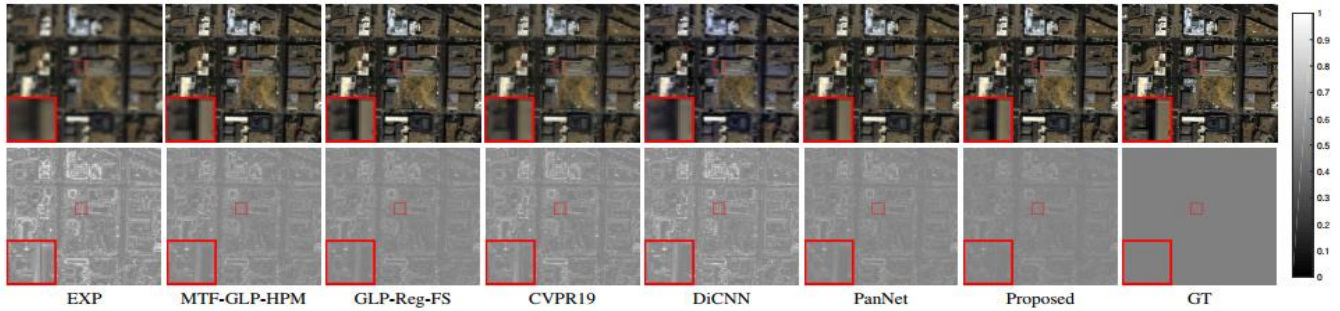


Fig. 5. The fusion results on the reduced resolution (simulated) WashingtonDC dataset (source: WorldView-2). Top row: the visual performance of the EXP, MTF-GLP-HPM, GLP-Reg-FS, CVPR19, DiCNN, PanNet, Proposed method, and the ground-truth (GT) image, respectively. Bottom row: the corresponding MAE maps using the GT image as reference. For a better visualization, we doubled the intensities of the MAE maps and added 0.5.

Data Generalization: (Simulated, *Xnet* trained on WV3 but tested on IKONOS + Pleiades)

Fig. 10. Top row: the fusion results on the reduced resolution (simulated) Toulouse datase (source: IKONOS). Bottom row: the fusion results on the reduced resolution (simulated) Pléiades2 dataset (source: Pléiades).

TABLE V
QUALITY METRICS ON THE REDUCED RESOLUTION (SIMULATED) TOULOUSE AND PLÉIADES2 DATASETS. (BOLD: BEST; UNDERLINE: SECOND BEST)

Dataset	PAN size	Method	*ERGAS*	*SAM*	*Q4*
Toulouse	512 × 512	PanNet	5.8152	5.5342	0.7180
		Proposed	**3.1948**	**3.5488**	**0.8943**
Pléiades2	256 × 256	PanNet	5.5887	5.9802	0.8417
		Proposed	**2.6649**	**4.0249**	**0.9522**
Ideal value			**0**	**0**	**1**