Fusion from Decomposition: A Self-Supervised Decomposition Approach for Image Fusion

Pengwei Liang[1], Junjun Jiang[1]*, Xianming Liu[1], and Jiayi Ma[2]
Harbin Institute of Technology[1] Wuhan University[2]

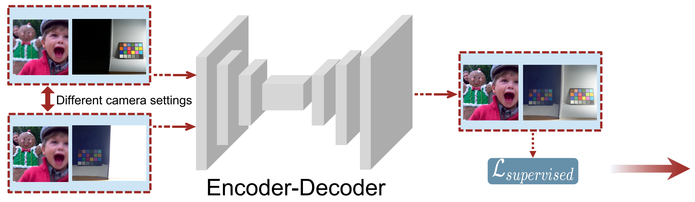
ECCV EUROPEAN CONFERENCE ON COMPUTER VISION TEL AVIV 2022 October 23-27, 2022

Introduction

Essence: The essence of image fusion is to integrate complementary information or best parts from source images.

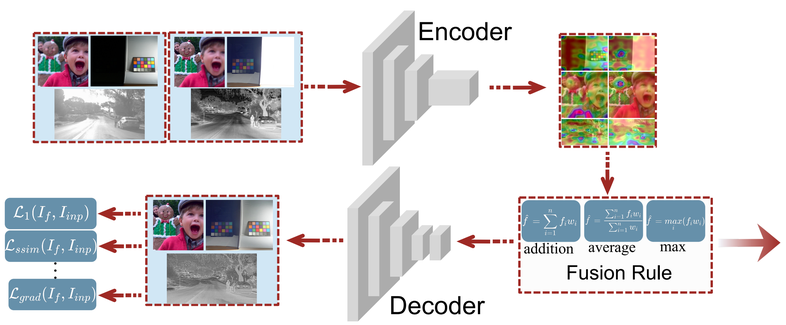
Under exposure image Over exposure image Fused image

Fusion evolving



(a) Direct Fusion via Supervised Learning

Disadvantage:
Obtaining the paired source images and ground-truth fused image would be difficult. It is impossible in some scenario such as **infrared-visible image fusion.**

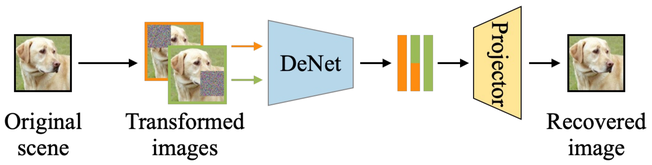


(b) General Fusion via Unsupervised Learning

Advantage:
Can be easily expanded for applicable scenarios.

Disadvantage:
The performance seriously depends on the human knowledge about the **auxiliary loss and fusion rule**.

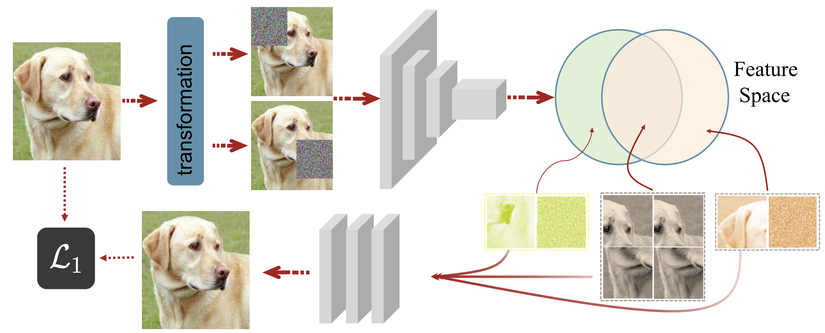
Contributions



We draw insight from the essence of image fusion, where the target fusion image can be generated from decomposed components of the source images

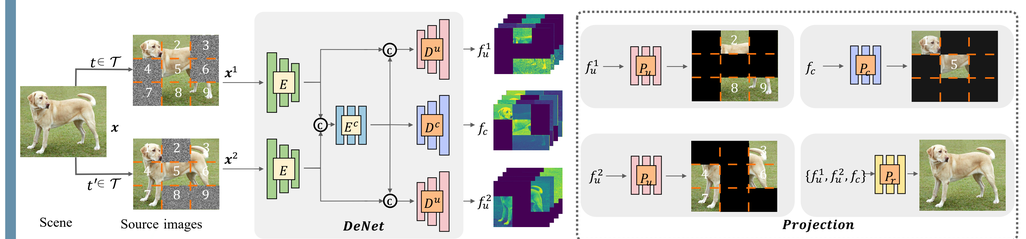
- ✓ We propose a novel image fusion method called DeFusion by **decomposing the source image based on a self-supervised learning framework**.
- ✓ We design **a pretext task**, called CUD, for image fusion, which does not rely on the existing supervised image fusion dataset, sophisticated loss functions and fusion rules.
- ✓ The proposed DeFusion is trained only with the COCO dataset and can be used as **a unified and versatile framework for various image fusion tasks** without any further fine-tuning

Overview



How image fusion is similar to self-supervised learning?

- ✓ The source images are transformed from the same scene
- ✓ The key is to extract the useful embedding representation
- ✓ Predicting the fused image from corrupted images is a special MIM

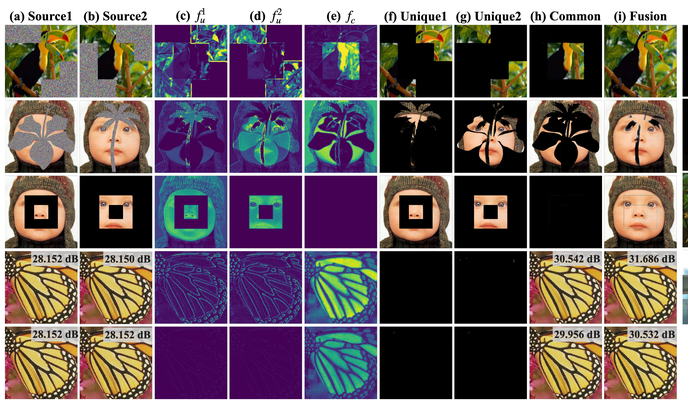


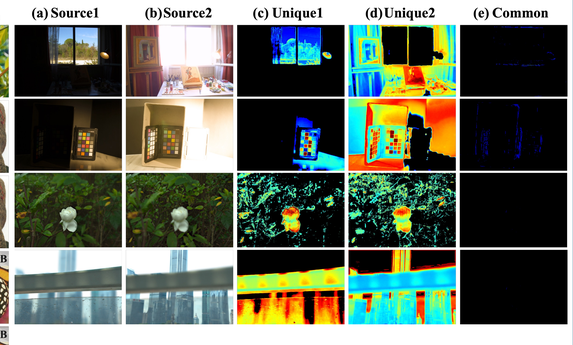
M_i denotes a random mask and n denotes Gaussian noise

An overview of our proposed self-supervised image decomposition and fusion method

Experiments

Visualizing Feature Embeddings





Quantitative comparisons

Method	MEFB [52] CE	QCV	SSIM	MEF-SSIM	IS	LPIPS	SICE [3] CE	QCV	SSIM	MEF-SSIM	IS	LPIPS
CU-Net	4.800	425.5	0.547	0.794	6.470	0.359	4.728	345.6	0.486	0.742	7.564	0.389
DeepFuse	4.993	363.0	0.544	0.796	6.346	0.380	5.262	189.3	0.523	0.797	**8.391**	0.322
IFCNN	4.943	247.7	0.573	0.818	6.776	0.335	4.551	290.2	0.492	0.697	8.453	0.372
MEFNet	**4.257**	593.4	**0.593**	0.796	6.432	0.321	5.102	505.7	**0.526**	0.711	8.068	0.358
PMGI	4.698	293.9	0.547	**0.822**	6.521	0.336	5.556	294.7	0.480	0.740	7.973	0.375
U2Fusion	4.526	**253.8**	0.526	0.815	**3.438**	**6.745**	**0.332**	209.5	0.488	**0.796**	8.314	**0.346**
DeFusion	2.881	262.3	0.608	0.827	6.587	**0.332**	2.830	**207.7**	0.571	0.788	7.869	0.353

	Method	Dataset[53] SSIM	Dataset[53] PSNR	Real-MFF noref [48] SSIM	Real-MFF noref [48] PSNR	Real-MFF [48] SSIM	Real-MFF [48] PSNR
Super-vised	IFCNN	**0.905**	**26.91**	**0.964**	**32.93**	0.983	36.92
Unsup-ervised	CU-Net	0.874	24.88	0.900	26.66	0.938	29.17
	MFFGAN	0.879	24.30	0.811	22.81	0.850	24.14
	PMGI	0.865	20.88	0.890	24.09	0.903	24.66
	U2Fusion	0.815	21.94	0.849	23.99	0.880	25.26
	DeFusion	0.928	28.13	0.969	33.61	**0.971**	**33.88**