Visible and infrared self-supervised fusion trained on a single example

Nati Ofir[1] **Jean-Christophe Nebel**[2]

[1]Weizmann Institute and The Open University of Israel [2]Kingston University London

We propose a self-supervised CNN for visible–NIR fusion, trained in seconds on a single pair using SSIM and edge-preservation losses with pseudo-labels. It preserves complementary details without large datasets, matches state-of-the-art, and enables efficient, robust fusion for dehazing and detection.

Image Fusion; Multispectral Imaging; Self-Supervised Learning

- Trains on a **single** RGB–NIR pair (no manual labels).
- Uses **SSIM** and **edge-preservation** losses as supervision.
- Preserves complementary structure and detail across modalities.
- Lightweight and fast: suitable for **on-demand** fusion.

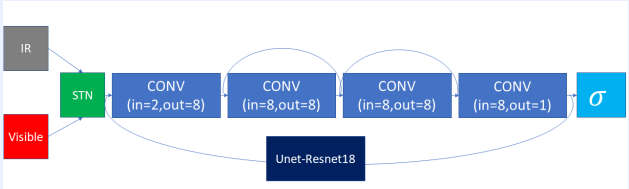
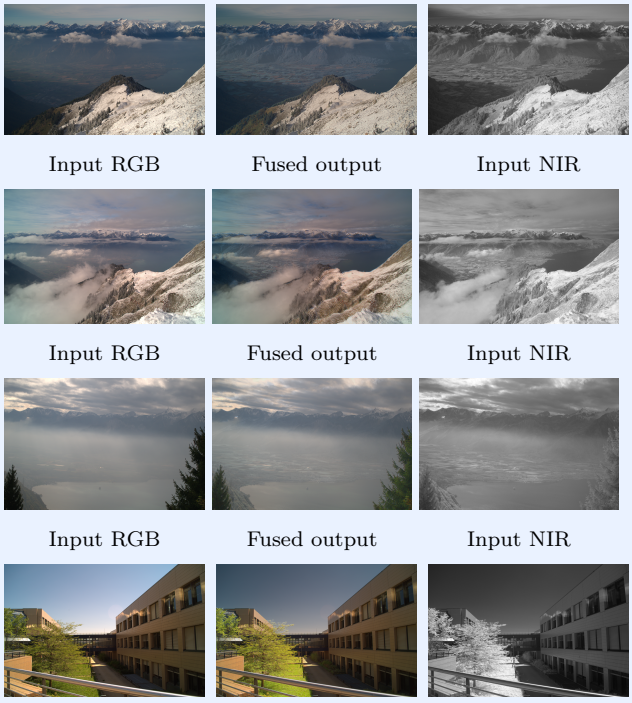


Figure: Proposed CNN architecture for visible–NIR fusion.



Examples: RGB–NIR fusion results from the VIS–NIR dataset.

Nati Ofir
nati.ofir@weizmann.ac.il / natiofir.github.io

Category	SuperPixel	PCA	Spectral	SSIM Loss	SSIM+EdgeLoss
Country	81.5	74.1	78.9	82.2	**82.3**
Mountain	89.9	89.6	88.1	90.1	**90.2**
Urban	93.8	**93.9**	92.7	93.4	93.5
Street	87.5	87.3	85.5	87.7	**87.8**

Structure-of-similarity (SSIM) scores of fused images for fast (non-heavy-training) methods.