A Experimental Details

Here we present the datasets used during our experiments, as well as their settings that may differ from the offical settings.

- *Cityscape Denoising and Segmentation*, which exploits Cityscapes [8] dataset (2975 images for training and 500 images for testing) for denoising and segmentation. In regard to the semantic feature extraction, we use the penultimate layer of VGG19 [52] pre-trained on the ImageNet [9] dataset for object classification. For the high-level vision tasks, we use the HRNet48 [60] pre-trained on the Cityscapes dataset for semantic segmentation.
- *Face Super-resolution and Alignment*, which exploits CelebA [33] (168854 images for training and 1000 images for testing) and Helen [20] (2005 images for training and 50 images for testing) datasets for face super-resolution and face landmark detection (denoted as alignment accuracy). In regard to the semantic feature extraction, we use the penultimate layer of LightCNN-9 [63] pre-trained on the CelebA [33] dataset for face recognition. For the high-level vision tasks, we use hourglass [40] networks pre-trained on the CelebA dataset for face landmark detection.
- *Natural Image Restoration*, which exploits the real-world restoration datasets, *i.e.*, i-haze [3] (30 images for training and 5 images for testing) as well as o-haze [2] (40 images for training and 5 images for testing) for dehazing. The setting of semantic feature extraction is same as the first experiment.