(1) For the residual-aware models, they are using the concept of disentanglement in the image domain by simply defining the rainy image as the linear superposition of rain layer and background layer, and de-raining is achieved by formulating a network to learn the rain layer which is then used together with the rainy image to obtain the residual image. However, the simple linear superposition assumption is too coarse to describe the formation of images with diversified rainy components (i.e., the real-rainy images), thus leading the trained residual-aware models to be (i) easily overfitted to only some specific rain types covered by the training set and (ii) unable to perform well on real rainy images. In contrast, the factor disentanglement in DRLE-Net is achieved in the latent space with high semantics, and two networks are defined to learn the distribution of each factor in a non-linear way. Consequently, as demonstrated by the results in Sec 4.4 of the paper, our DRLE-Net is able to perform well on both synthetic and real images with diversified rainy conditions, and outperform the residual-aware models by a large margin;
(2) For raindrop segmentation aware models, they are not related to the disentanglement design in either image domain or semantic feature domain. Alternatively, they are formulated by using the guidance information (e.g., rainy components segmentation map in AGAN/ JORDER and density label in DID-MDN) to drive the filter in the de-raining network to be more activated on the rainy regions, such that all rainstreaks/raindrops can be attended for producing better de-raining results without artefacts of under-deraining. Good results of such models come from the usage of the guidance information to learn better semantic feature that will be used by the decoder to reconstruct the clean images. For a de-raining task, only the background related features are useful for reconstructing the rain-free images. However, without a factor separation process, the semantic features cover factors related to both rainy component (task-irrelevant factor) and background component (task-relevant factor). Consequently, the guidance information will take effect on both the task -relevant and -irrelevant factor learning procedure. To avoid this and find a better solution for only improving the learning of the task-relevant factor, we formulate a novel framework for joint factor disentanglement and enhancement. Results in Sec 4.4 of the paper also demonstrate the superiority of our framework over the segmentation aware models (JORDER/URML and AGAN).
(3) For the image-decomposition based models (DSC and GMM in Table3), the rainy image is separated into two subspaces represented by Gaussian Components or dictionary, then optimization methods (e.g., sparse coding or EM algorithm) are used to derive the representation of rainy and background components. This procedure can also be interpreted as disentanglement, and DRLE-Net is in spirit similar to such process. However, compared to the image-decomposition based models, DRLE-Net shows many advantages such as (a) easier to be optimized and much faster inference speed, (b) much better de-raining results, and (c) high potential of being used for solving other problems as analyzed in Sec 1.1 of the letter.
We will add the corresponding analysis to the “Introduction” of the paper in the camera-ready version.