

NTIRE 2019 Image Dehazing Challenge

Factsheet: FastNet

–Feature Forwarding for Efficient Image Dehazing–

March 25, 2019

1 Team details

- Challenge name:
NTIRE 2019 Image Dehazing Challenge
- Team name:
FastNet (test results submitted as team tk)
- Team leader name:
Peter Morales
- Team leader address, phone number, and email:
244 Wood St, Lexington, MA 02420; Office S1-313
781 981 8738
peter.morales@ll.mit.edu
- Rest of the team members:
Tzofi Klinghoffer
- Team website URL:
<https://pmm09c.github.io/ntire-dehazing/>
- Affiliation:
MIT Lincoln Laboratory

- Affiliation of the team and/or team members with NTIRE2019 sponsors:
None
- User names and entries on the NTIRE2019 Codalab competitions:
pmm09c, tzofi
- Best scoring entries of the team during development/validation phase:
NA for current method.
- Link to the codes/executables of the solution(s): https://github.com/pmm09c/ntire-dehazing/releases/tag/ntire_submission

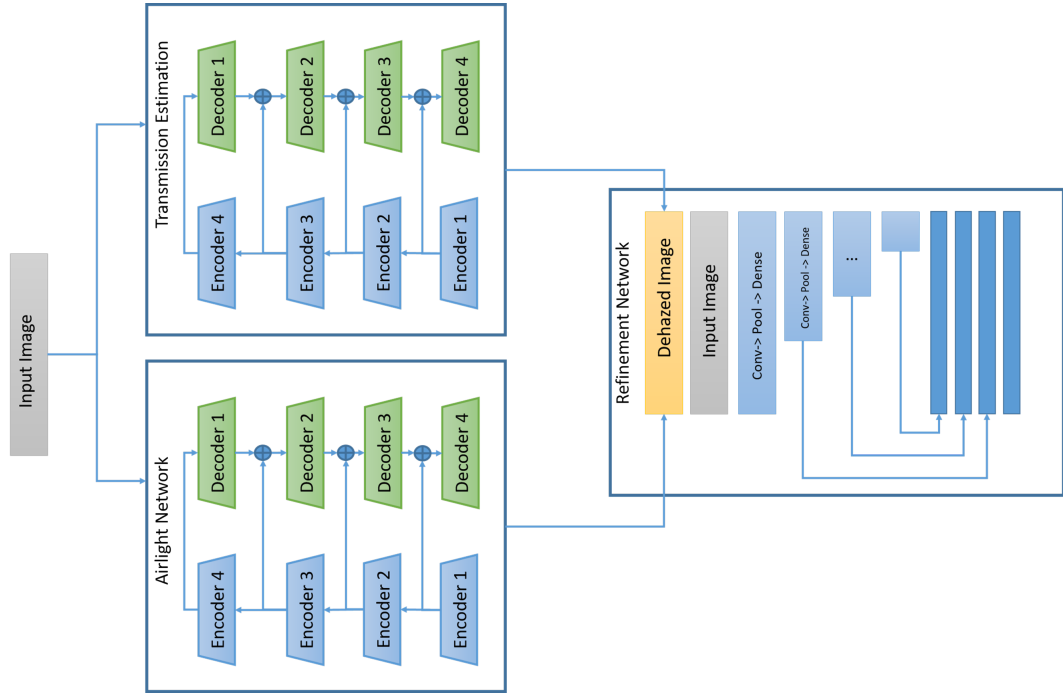
2 Contribution details

- Title of the contribution :
Feature Forwarding for Efficient Image Dehazing
- General method description:
This approach utilizes a feed-forward convolutional neural network trained only on the competition’s provided dataset. The network uses a pre-trained feature encoder-to-decoder convolutional architecture, feeding to dense convolutional refinement layers. Two approaches were attempted. One approach estimates airlight and transmission maps with two separate encoder-to-decoder networks that feed into the refinement layers. This model is named DualFastNet in our repository. The second approach uses a single encoder-to-decoder network that feeds into the refinement layers. This model is named FastNet50 in our repository. Our competition test results were generated from the DualFastNet architecture.
- Description of the particularities of the solutions deployed for each of the challenge competitions or tracks :
Our top performing model is the DualFastNet architecture. In this model, the encoder-to-decoder architectures used for estimating airlight and transmission maps are adapted LinkNet models [1] that use a pre-trained ResNet18 model in the encoding layers. The original LinkNet

architecture was developed for image segmentation and the output layers were modified to output a feature map rather than a single channel segmentation map. The output of these layers are used to compute the dehazed image using the following atmospheric scattering model: $I(z) = J(z) + A(z)(1 - t(z))$, where $I(z)$ is the hazy image, $J(z)$ is the dehazed image, $A(z)$ is the airlight, and $t(z)$ is the transmission map. The dehazed image estimated from this model is concatenated with the input, hazy image and fed into the refinement layers. The refinement layers were adapted from Zhang et al.’s Densely Connected Pyramid Dehazing Network [2].

Unlike Zhang’s architecture, our network does not utilize U-Net, a discriminator, or need to be trained separately on airlight and transmission map data for convergence. In fact, our model can be easily be trained from scratch as a result of the airlight and transmission map encoder-decoders not needing to be trained separately. Loss is enforced on the output image only, rather than on the output image, transmission map, and airlight estimation. Additionally, utilizing LinkNet encoder-decoders for our feature extraction allows us to achieve speed-efficient performance, processing a single image within 0.025 seconds on a Titan RTX GPU. More training data would improve PSNR, SSIM, and qualitative performance even further.

- References:
A. Chaurasia, et al. LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation. <https://arxiv.org/abs/1707.03718>.
H. Zhang, et al. Densely Connected Pyramid Dehazing Network. <https://arxiv.org/abs/1803.08396>.
- Representative image / diagram of the method(s)



3 Global Method Description

[* Indicates method used in competition test results.]

- Total method complexity:

Model Parameters

- **FastNet** 11,554,167 parameters
- **FastNet50** 28,782,647 parameters
- ***DualFastNet** 23,072,725 parameters

Run Time

- **FastNet** 0.01 seconds
- **FastNet50** 0.016 seconds
- ***DualFastNet** 0.025 seconds

Run time can be greatly increased by batching images on input and / or optimizing using NVIDIA TensorRT from native PyTorch.

- Which pre-trained or external methods / models have been used :
 - **FastNet** ResNet18 on pre-trained ImageNet
 - **FastNet50** ResNet50 on pre-trained ImageNet
 - ***DualFastNet** ResNet18 on pre-trained ImageNet
- Which additional data has been used in addition to the provided NTIRE training and validation data (at any stage, if any) :
None
- Training description :
We utilize the Adam Optimizer with an initial learning rate of 0.001. Many loss functions were compared, but test results are from a Dual-FastNet model trained using only MSE loss on the hazy input image and dehazed output image. Each sample during training could be augmented at random with image rotations and / or image crops.
- Testing description:
The NTIRE validation images were used for quantitative evaluation, and NTIRE test images were used for qualitative evaluation. The final submission was trained on provided training data as well as 2 images from the provided validation set.
- Quantitative and qualitative advantages of the proposed solution :
The technique proposed yields PSNR scores that are competitive with other results on the Validation Leaderboard when run on 3 of the validation images deemed to be "more difficult" (46, 48, and 50). PSNR scores are greater than 16. Run time is also very efficient, running at about 0.025 seconds per test image.

Qualitatively, our method performs well and is able to create detail in areas of the imagery with high concentrations of haze.
- Results of the comparison to other approaches (if any) :
No comparison has yet been done with other neural network approaches. Models are currently training on the He, Zhang dataset

(which draws from the NUY Depth dataset) for fair comparison and comparisons will be included in any submitted paper.

When comparing with non-neural network approaches, we found our approach qualitatively outperformed the current methods, such as: Berman, Dana et. al's "Non-local Image Dehazing" method, He, Kaiming et. al's "Single Image Haze Removal Using Dark Channel Prior" method, and Galdran, Adrian et. al's "On the Duality Between Retinex and Image Dehazing" method. These methods did not seem to tackle areas of the scene with high concentration of haze well.

- Results on other standard benchmarks such as I-HAZE, O-HAZE (if any) :

In progress: we are training on the He, Zhang dataset (which draws on the NYU-Depth dataset) and withholding validation images, as well as I-HAZE and O-HAZE, for PSNR/SSIM metric evaluation.

- Novelty degree of the solution and if it has been previously published: To the best of our knowledge, this is a novel approach with no previous publications that use our technique. The most similar approaches are based on either style transfer architectures like Conditional Generative Adversarial Networks, U-Net, or He Zhang's DCPDN. Our method achieves better run time efficiency than any other neural network approach for image dehazing that we have seen.

4 Ensembles and fusion strategies

- Describe in detail the use of ensembles and/or fusion strategies (if any). :
None
- What was the benefit over the single method? :
NA
- What were the baseline and the fused methods? :
NA

5 Technical details

- Language and implementation details (including platform, memory, parallelization requirements) :
PyTorch, Ubuntu 18.04, 2x NVIDIA Titan RTX GPUs
- Human effort required for implementation, training and validation?:
Implemented neural network models, training environment, and test environment, as well as performed multiple experiments comparing loss functions, hyper-parameters, and model.
- Training/testing time? Runtime at test per image :
0.025 seconds per image (faster if run in batches or with NVIDIA TensorRT)
- Comment the robustness and generality of the proposed solution(s)? Is it easy to deploy it for other sets of downscaling operators? :
Our relatively small model is easy to deploy and can be greatly sped up with TensorRT. The model is also generalizable and performs well on images vastly different than those in the training set, per initial results.
- Comment the efficiency of the proposed solution(s)? :
Compared with most neural network solutions, our approach is extremely efficient, but SSIM may suffer as a tradeoff.

6 Other details

- Planned submission of a solution(s) description paper at NTIRE2019 workshop. :
Our submission will outline and describe in detail our methods; comparison with other methods; Run Time, SSIM, PSNR, and qualitative based results on validation data imagery; comparison of loss functions used and results from each; and discussion of interesting findings and abnormalities in the data and models associated with training neural networks (for example, large batch size proved to be paramount in our experiments).
- General comments and impressions of the NTIRE2019 challenge. :
It was a pleasure to participate in and very well run.

- What do you expect from a new challenge in image restoration and enhancement? :
We find it interesting when there is so much haze that the problem goes beyond physics based approaches and requires in-painting. We'd hope to see a similar contest next year with even more training data.
- Other comments: The leaderboards didn't always work.