Zac Anderson
https://github.com/zacanderson/
2021 April 8

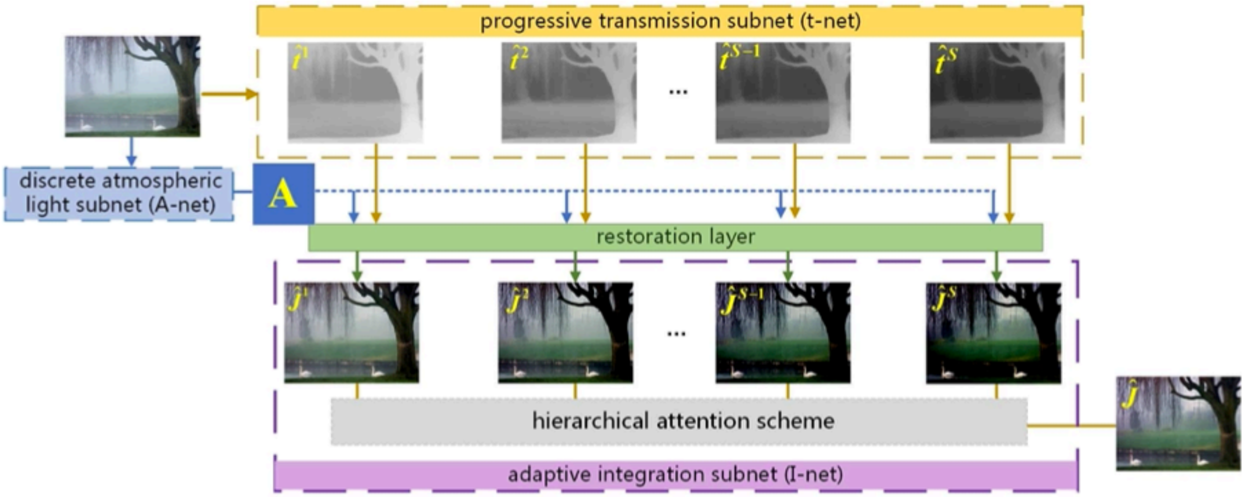
LAP-Net : Synopsis

Sourced from ICCV research paper by Yunan Li, et. al

This paper delves into the fascinating topic of image dehazing: the process in which fog and environmental conditions are reduced through a trained convolutional neural network, improving image clarity and providing access to details not previously accessible. Dehazing has been attempted before this paper with varying levels of success - previous attempts used single-layer CNNs, which can work very well situationally, but tend to see efficacy trail off in real-world situations. LAP-Net aims to ameliorate many dehazing concepts by utilizing a hierarchical integration scheme, where supervision of the image is the responsibility of numerous training layers. These results are pooled into the scheme, and subsequently combined into a single image, thus completing the role of LAP-Net.

As mentioned previously, real-world situations are difficult for single-layer CNNs due to the immense amount of situational data recorded. The training set needs specific instructions on which features to look for in order to reduce the image's noise, leading to improved image clarity; however, training and weighting all of these variables in a single neural network often leads to muddled results. The hierarchical integration scheme is the main focus of this paper, due to its ability to run multiple CNNs, each aiming to improve a certain field of the image. By having multiple CNNs with specific tasks, coupled with the ability of the HIS to combine their results, the final result of LAP-Net is able to see superior results with both datasets and real-world examples.

LAP-Net is made up of several intermediary steps and networks that gather this specific data to be translated by the HIS. The next section of this report will review those methods and their anticipated outputs in reference to the overall performance of LAP-Net.

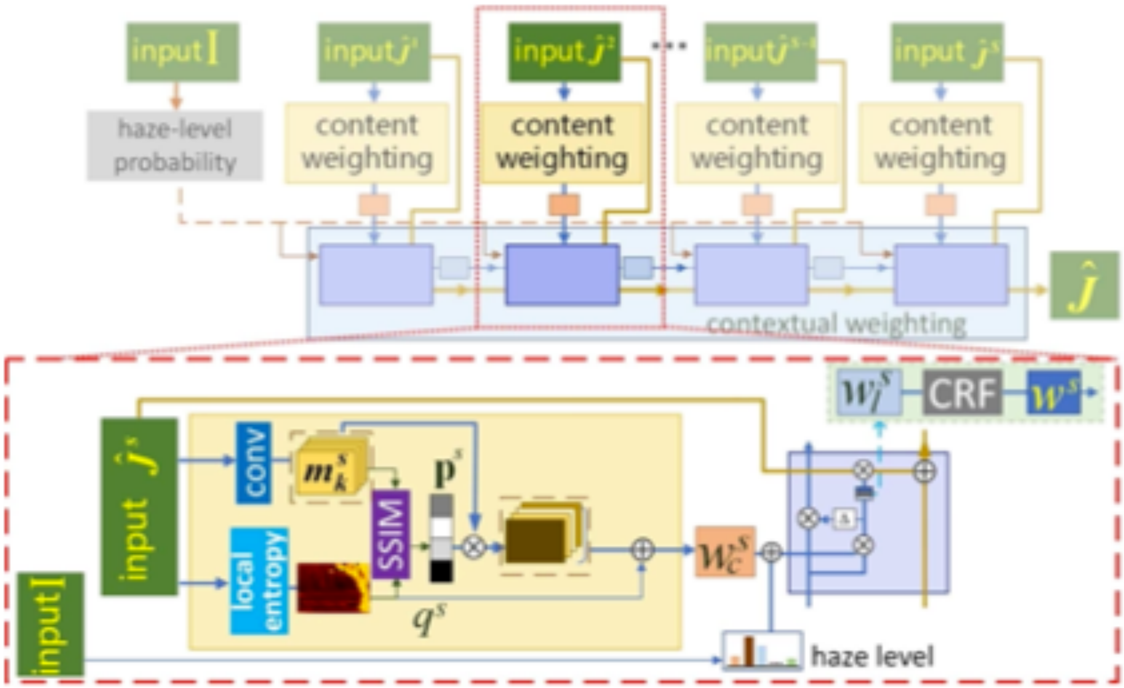


The A-net is implemented as a “residual dense pooling network,” effectively reading in a vast amount of data about the ambient light settings, obtaining a discretized interval value via classification task (i.e. a real value in [Al, Ah]), culminating in a final estimate for the input’s global illuminance. Light is doubly important when dealing with haze, since the haze refracts light, and can scatter the pixel values significantly at distance. By gathering information on the input’s light levels the amount of guesswork is reduced, leading to higher quality results.

The t-net is responsible for making adjustments to the saturation over the haze to evaluate how “deep” the haze is. If there are distant features in an input image, the amount of haze that must be traversed for the light to reach the lens is significant; by fitting these filters over the input image in separately supervised stages, the restoration layer (recipient of t-net output) will be able to produce varied levels of restoration attempts. If the haze is very light, the need for a t-net may not be apparent, though the overall operation time for light, medium, and heavy haze situations, the t-net provides useful information to detect and sharpen the image.

The restoration layer, as noted above, creates multiple restorations based on the information provided by the t-net. Not all of the restoration attempts will be used, there will be cases of over/under-correcting that the I-net will reject, but having them created is more interpretative for the neural networks, while also aiming to produce the highest-quality image possible.

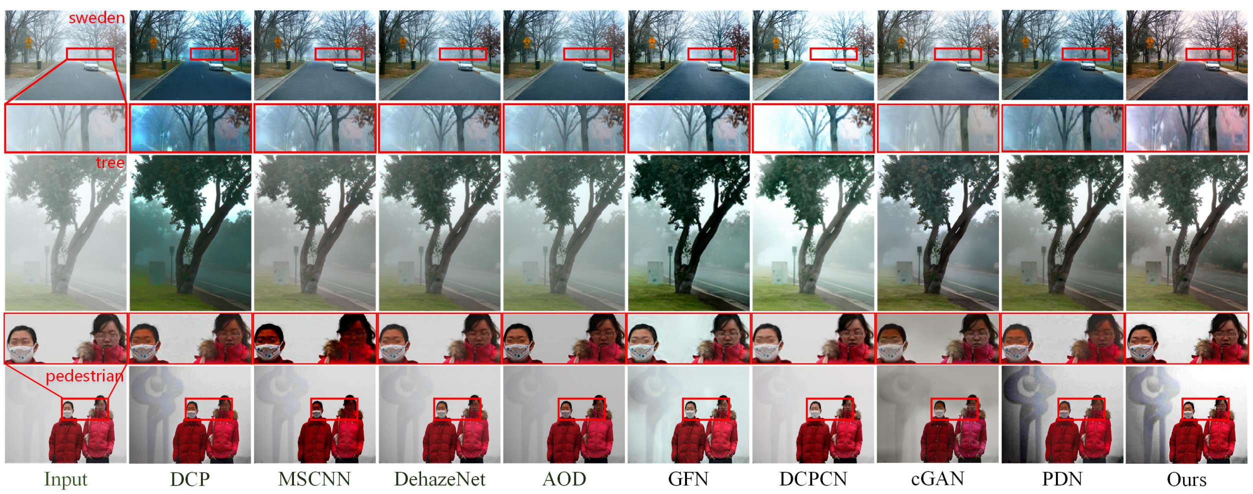
The I-net is where LAP-Net steps into its stride. The intermediary steps of LAP-Net have been summarized, with the I-net being of particular importance due to the images being weighted by the program, then passed into memory and final computation.



Primarily, the I-net handles all of the input from the intermediary steps via content and contextual weightings. The restoration level provides restored images from the t-net's progressive results, passed into the I-net to be weighted individually, then cumulated together for the final result of "ˆJ": aptly named the "ultimate restoration." The I-net's functionality is showcased best in the figure taken from the research paper, shown above.

As a penultimate step, the altered image is passed through a convolution, along with the final weights determined previously. Then, the final picture is available, with greatly reduced haze and, ideally, improved visibility of the scene.

The process of LAP-Net is a very cohesive, all-in-one experience that aims to improve several facets of single-layer dehazing networks. It is difficult to determine a single part of this project that doesn't contribute to the overall execution, and that appears as a success in and of itself. Included below is Figure 5 from the original paper, which show each result of different dehazing networks to allow the distinguishing features of each to stand out.



LAP-Net's success becomes a bit more clear after viewing these images, or at least, that's what happened in my experience. Computer vision research is growing relatively quickly, and it must be kept in mind that chasing these incremental gains are immense undertakings for the researchers involved.

Observing DCPCN (Densely Connected Pyramid Dehazing Network by He Zhang *et. al*), which specializes in edge preservation techniques in hazy images, the variability in results is evident: the objects in the foreground are sufficiently brightened and clear, yet the background suffers from mild overexposure throughout the three images.

cGAN (Conditional Generative Adversarial Network by Runde Li *et. al*) is generative, and performs best when the training set closely matches the input image's level of haze. The results produced are still clearer than some others, though the limitations are seen when haze levels vary between inputs. Single-layer networks have a harder time adjusting to variable haze densities, which can be a detriment in real-world weather conditions, where nothing is guaranteed to look as the dataset expects.

Examining the previously discussed images a bit closer, we can see that the colour saturation in all images as been increased, which makes the resulting photo appear more lifelike than some of the other networks. The objects in this example - trees, signs, people, cars, etc - are all more defined, and at least more visible to the naked eye. Upon zooming in, there are entire pieces of the image that have been made visible by LAP-Net. One example is the area behind the sign in the "tree" image, where two vertical lines of pixels have materialized from behind the haze. This is an especially exciting discovery that seems capable of discovering information as pixels that have been obscured by haze. The shape in the "pedestrian" image is another example: a rough outline in the input image was transformed into a distinct shape with colour and clear structural features (i.e. the beams of metal crossing over each other).

Without doubt, the efforts of the researchers of LAP-Net was met with conclusive success. Computer vision advancements are appearing more and more rapidly in our current epoch, which will hopefully lead to even more discoveries being made about how we interact with computers, and their interactions with the world around us. Developmental comments aside, these gradual improvements lead to higher levels of refinement for industry-leading technology, which we can use today in order to facilitate future growth.

Li, Yunan, et. al "LAP-Net: Level-Aware Progressive Network for Image Dehazing." *LAP-Net (ICCV Permalink)*, https://openaccess.thecvf.com/content_ICCV_2019/papers/Li_LAP-Net_Level-Aware_Progressive_Network_for_Image_Dehazing_ICCV_2019_paper.pdf.