



Two Simple Unfolded Residual Networks for Single Image Dehazing

Bartomeu Garau¹^a, Joan Duran^{1,2}^b and Catalina Sbert^{1,2}^c

¹*Institute of Applied Computing and Community Code, Universitat de les Illes Balears (UIB), Edifici Complexo d'R+D, Cra. de Valldemossa km 7.4, E-07122 Palma, Spain*

²*Dept. of Mathematics and Computer Science, UIB, Cra. de Valldemossa km 7.5, E-07122 Palma, Spain*
b.garau@uib.cat, {joan.duran, catalina.sbert}@uib.es

Keywords: Image Dehazing, Deep Learning, Unfolding, Residual Network, Channel Attention, Variational Methods.

Abstract: Haze is an environmental factor that impairs visibility for outdoor imaging systems, presenting challenges for computer vision tasks. In this paper, we propose two novel approaches that combine the classical dark channel prior with variational formulations to construct an energy functional for single-image dehazing. The proposed functional is minimized using a proximal gradient descent scheme, which is unfolded into two different networks: one built with residual blocks and the other with residual channel attention blocks. Both methods provide straightforward yet effective solutions for dehazing, achieving competitive results with simple and interpretable architectures.

1 INTRODUCTION

The rapidly increasing use and demand for efficient outdoor imaging systems have brought issues like dehazing to the forefront of image processing. Outdoor images are often affected by atmospheric conditions such as haze, smoke, rain or snow. In particular, haze reduces visibility, giving scenes a gray tone and lowering contrast. Tackling these issues is highly relevant for a wide range of applications, including surveillance, autonomous systems, and remote sensing.

Haze is an environmental phenomenon, caused by the scattering of light as it travels through the atmosphere, where airborne particles distort the light. Moreover, the degradation depends on both the depth of the scene and the haze density. This makes dehazing a particularly challenging problem.


Various strategies for image dehazing have been explored in the literature (Wang and Yuan, 2017; Guo et al., 2022; Jackson et al., 2024). Some methods address the problem as an enhancement task, using techniques such as histogram equalization (Jun and Rong, 2013; Thanh et al., 2019) or the Retinex theory (Zhou and Zhou, 2013; Galdran et al., 2018). Other methods leverage the physical principles underlying hazy scenes (McCartney, 1977). The resulting models can be approached in different ways, including di-


rect computation (Tan, 2008; He et al., 2011) or variational techniques (Fang et al., 2014; Galdran et al., 2015; Liu et al., 2022).


With the rapid growth of artificial intelligence, numerous dehazing methods involving deep learning networks have emerged (Cai et al., 2016; Qin et al., 2019; Lei et al., 2024). Some of these methods include unfolding architectures (Yang and Sun, 2018; Fang et al., 2024), which combine the strengths of model-based and data-driven learning approaches.

In this paper, we propose two simple model-based deep unfolded approaches to variational image dehazing. Our proposals are based on the dark channel prior (He et al., 2011) to estimate the main components of a hazy image, specifically the transmission map and the atmospheric light of the scene. We introduce a simple variational formulation to obtain the haze-free image as the minimizer of an energy functional. The minimization of this energy is performed using a proximal gradient descent algorithm, in which the proximal operators are replaced by residual networks.

The rest of the paper is organized as follows. In Section 2, we review the related work on image dehazing. Section 3 introduces the two proposed models and, in Section 4, we discuss their implementations and compare them with state-of-the-art approaches. Section 5 conducts an ablation study to justify the configurations of our architectures. Finally, conclusions are drawn in Section 6.

^a <https://orcid.org/0009-0008-3439-8316>

^b <https://orcid.org/0000-0003-0043-1663>

^c <https://orcid.org/0000-0003-1219-4474>

2 RELATED WORK

Physical principles can be used to describe the atmospheric scattering that generates haze. In this context, McCartney (McCartney, 1977) posits that a hazy image is formed through the combined effects of light attenuation and air-light scattering. This leads to the following expression:

$$\mathbf{I}(x) = \mathbf{J}(x)t(x) + \mathbf{A}(1 - t(x)), \quad (1)$$

where \mathbf{I} is the hazy image, \mathbf{J} is the haze-free image, t is the transmission map (the proportion of the clear image that reaches the camera), and \mathbf{A} is the atmospheric light of the scene. Usually, the transmission is related to the depth map of the scene. Estimating \mathbf{J} , \mathbf{A} , and t from \mathbf{I} is a highly ill-posed inverse problem.

To solve the decomposition problem arising from (1), one either requires additional information or must rely on some prior assumptions. Tan et al. (Tan, 2008) assume that hazy images exhibit lower contrast and that the variation of the air light is a smooth function of the distance. Fattal et al. (Fattal, 2008) separate surface shading from the transmission map. He et al. (He et al., 2011) introduce the dark channel prior, which states that in most local patches of haze-free outdoor images there are pixels with very low intensities in at least one color channel.

In the variational framework, the dehazed image is obtained as the minimizer of an energy functional that incorporates both data-fidelity terms, which measure the deviation from prescribed constraints involving the hazy image, and regularization terms, which assess the smoothness of the solution. Fang et al. (Fang et al., 2014) pioneered a variational formulation for image dehazing, where the energy functional enforces total variation to regularize the depth map and weighted total variation for the dehazed image. The authors rely on the dark channel prior to estimate an initial transmission. Since then, several variants have been proposed. For example, Lei et al. (Jin et al., 2024) apply total generalized variation to the depth, while Liu et al. (Liu et al., 2018) introduce nonlocal regularization to refine the transmission map, suppress unwanted artifacts, and preserve image details. In (Stipetić and Lončarić, 2022), the authors propose a smooth variational formulation of the dark channel prior that reaches a minimum when the reconstructed image satisfies the prior.

Other variational approaches avoid (1) and exploit alternative formation models. In this context, Galdran et al. (Galdran et al., 2015) propose an energy that maximizes the average contrast of the image, which is further studied in (Galdran et al., 2017) and applied for image fusion. On the other hand, Liu et al. (Liu et al., 2022) decompose the hazy image as a linear

combination of structure, detail, noise and glow, and use different regularization terms for each of these components.

Recently, the growing popularity of deep learning architectures has led to an increase in dehazing methods. This trend began with (Cai et al., 2016), which improved the estimation of the transmission map using an end-to-end convolutional neural network (CNN). In this framework, some architectures lack physical basis and rely on artificially generated pairs of hazy and ground-truth images (Qu et al., 2019; Qin et al., 2019). However, there are also dehazing networks based on the models and priors discussed previously, such as histogram correction (Chi et al., 2020), the Retinex theory (Li et al., 2021; Lei et al., 2024), and the dark channel prior (Zhang and Patel, 2018; Golts et al., 2020).

The use of formation models makes variational methods robust to distortions, but their performance is limited by rigid priors. Conversely, data-driven learning approaches can easily learn natural priors, but are less flexible and interpretable. Deep unfolding networks combine the strengths of both. The general idea involves unfolding the steps of the optimization algorithm into a deep learning framework. These networks can be based on transformers (Song et al., 2023), pyramid structures (Xiao et al., 2024) or classical optimization algorithms (Yang and Sun, 2018; Fang et al., 2024). In (Yang and Sun, 2018), the authors introduce an energy functional with a novel dark channel regularization term and subsequently unfold a proximal point algorithm into deep CNN structures. More recent architectures like (Fang et al., 2024) use (1) without assuming the dark channel prior. However, the resulting algorithm is unfolded into a cooperative network that increases in complexity.

3 PROPOSED MODELS

Based on the haze formation model (1), we want to recover \mathbf{J} , \mathbf{A} , and t from a single hazy image \mathbf{I} . To address the ill-posed nature of such a problem, we will use the dark channel prior (He et al., 2011) to estimate t and \mathbf{A} , that is,

$$\mathbf{J}_{dark} = \min_{c \in \{R, G, B\}} \left(\min_{y \in w(x)} \mathbf{J}_c(y) \right) \rightarrow 0,$$

where $w(x)$ is a patch of pixels centered at x . Then, we will estimate a rough transmission map as

$$\tilde{t}_0(x) = 1 - \nu \min_{c \in \{R, G, B\}} \left(\min_{y \in w(x)} \left(\frac{\mathbf{I}_c(y)}{\mathbf{A}_c} \right) \right), \quad (2)$$

where ν is a constant set empirically to $\nu = 0.95$. To compute (2), we will first estimate \mathbf{A} by taking the

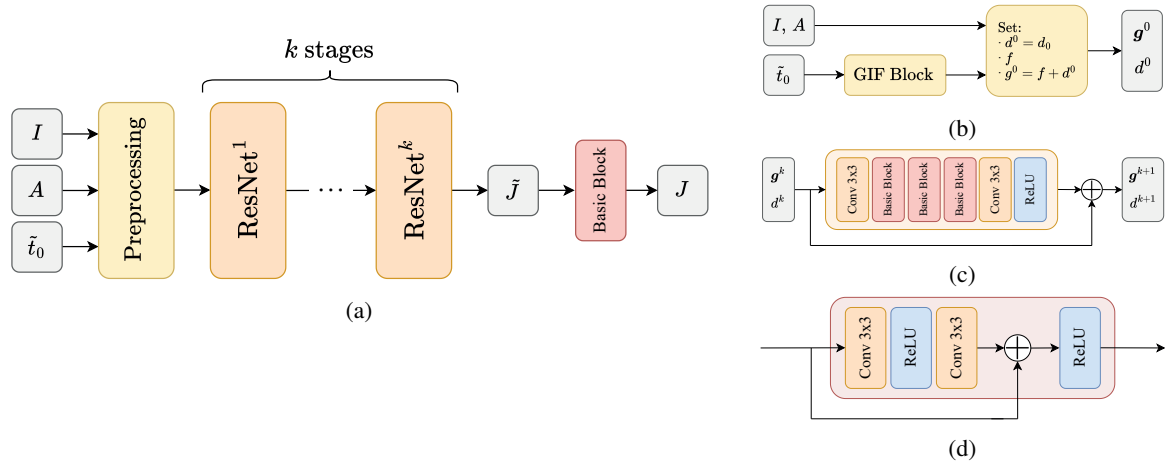


Figure 1: (a) Overall architecture of the unfolded formulation of (7). (b) Preprocessing block. (c) Residual Network (ResNet) architecture. (d) Basic block residual architecture.

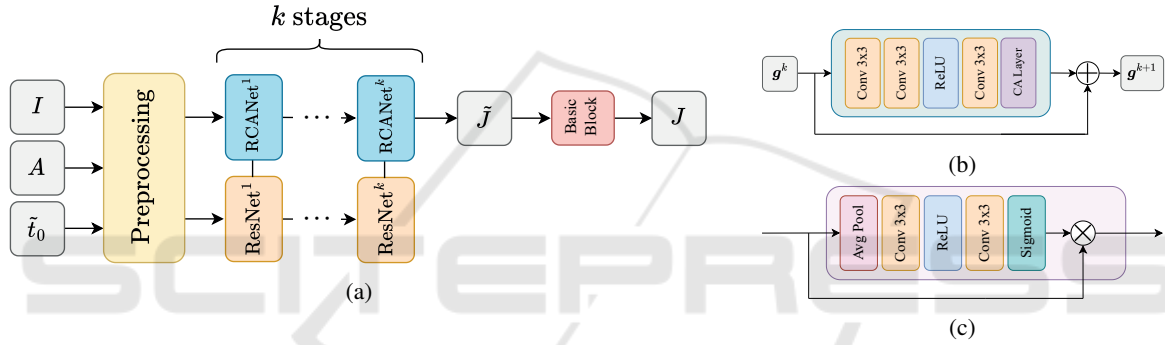


Figure 2: (a) Overall architecture of the unfolded formulation of (8). The preprocessing block, the ResNet block, the basic block are the same ones featured in Figures 1b, 1c, 1d, respectively. (b) Residual channel attention architecture for g . (c) Channel attention (CA) layer.

mean of the top 0.1% brightest pixels of J_{dark} on each channel, as done in (He et al., 2011). Once we have \tilde{t}_0 , we will apply a guided filter (He et al., 2013) to obtain the initial transmission map t_0 .

Following (Fang et al., 2014), we can rewrite each channel of (1) as

$$\mathbf{A}_c - \mathbf{I}_c = t(\mathbf{A}_c - \mathbf{J}_c).$$

By linearizing the model, we get

$$\log(\mathbf{A}_c - \mathbf{I}_c) = \log t + \log(\mathbf{A}_c - \mathbf{J}_c).$$

When the atmosphere is homogeneous, the transmission can be approximated by

$$t(x) = e^{-\eta d(x)}, \quad (3)$$

where $\eta > 0$ describes the scattering of the medium and d is the depth map of the scene. Using (3), and setting $f_c = \frac{1}{\eta} \log(\mathbf{A}_c - \mathbf{I}_c)$ and $g_c = \frac{1}{\eta} \log(\mathbf{A}_c - \mathbf{J}_c)$, we end up with $g_c = f_c + d$ or, in vectorial form,

$$\mathbf{g} = \mathbf{f} + \mathbf{d}, \quad (4)$$

where we denote $\mathbf{g} = \{g_R, g_G, g_B\}$, $\mathbf{f} = \{f_R, f_G, f_B\}$, and $\mathbf{d} = \{d, d, d\}$.

3.1 Variational Formulation

We will estimate d and \mathbf{J} from (4) as the minimizers of an energy functional of the form

$$E(\mathbf{g}, d) := R(\mathbf{g}, d) + F(\mathbf{g}, d),$$

where R and F consist of the regularization and fidelity terms, respectively. On the one hand, we choose different regularizers for \mathbf{g} and d :

$$R(\mathbf{g}, d) := R_1(\mathbf{g}) + \lambda R_2(d),$$

where $\lambda > 0$ is a trade-off parameter. On the other hand, we consider the following fidelity terms:

$$F(\mathbf{g}, d) := \frac{\alpha}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2} \|d - d_0\|_2^2,$$

where $d_0 = -\log t_0$ and $\alpha, \gamma > 0$. Therefore, we aim to solve the following minimization problem:

$$\min_{\mathbf{g}, d} \{R_1(\mathbf{g}) + \lambda R_2(d) + F(\mathbf{g}, d)\}. \quad (5)$$

Since F is differentiable, if we assume R_1 and R_2 to be proper, convex and lower semicontinuous functionals, we can solve (5) using the proximal gradient descent algorithm (Chambolle and Pock, 2016). Therefore, the sequence of iterates $\{(\mathbf{g}^k, d^k)\}$ converging to the solution of the minimization problem (5) is given by

$$\begin{cases} \mathbf{g}^{k+1} = \text{prox}_{\tau R_1}(\mathbf{G}^k), \\ d^{k+1} = \text{prox}_{\sigma \lambda R_2}(D^k), \end{cases} \quad (6)$$

where $\tau, \sigma > 0$ are the step-size parameters, and

$$\begin{aligned} \mathbf{G}^k &:= \mathbf{g}^k - \tau \nabla_{\mathbf{g}} F(\mathbf{g}^k, d^k) = (1 - \tau \alpha) \mathbf{g}^k + \tau \alpha (\mathbf{f} + d^k), \\ D^k &:= d^k - \sigma \lambda \nabla_d F(\mathbf{g}^k, d^k) \\ &= (1 - 3\sigma \alpha - \sigma \gamma) d^k + \alpha \sigma \sum_c (g_c^k - f_c) + \sigma \gamma d_0. \end{aligned}$$

3.2 Unfolded Formulation

If we consider R_1 and R_2 to be two generic regularizers that are proper, convex and lower semicontinuous, we can unfold (6) and replace the proximal operators by learning-based networks. Therefore, (6) becomes

$$\begin{cases} \mathbf{g}^{k+1} = \text{ResNet}^k(\mathbf{G}^k), \\ d^{k+1} = \text{ResNet}^k(D^k). \end{cases} \quad (7)$$

The hyperparameters $\lambda, \alpha, \gamma, \tau$ and σ are learned throughout the training phase and shared across all stages (that is, the number of iterations of the optimization algorithm). However, the residual networks do not share weights between stages.

The overall structure of the network is illustrated in Figure 1a. Figure 1b displays the initialization stage, where \tilde{t}_0 is filtered with the Guided Image Filtering Block (GIF Block) introduced in (Yang and Sun, 2018). This block is fixed and not learned during the training phase. In this way, we obtain the transmission map t_0 , which is used to set the initialization variables d^0 and \mathbf{g}^0 using the relations (3)-(4). Figure 1c shows each stage of the residual network used to compute (7), built with the basic blocks depicted in Figure 1d. From now on, this network will be referred to as URNet (Unfolded Residual Network).

We will now propose an alternative architecture. The nonlocal theory for image processing is used to capture self-similarities across different patches of an image to smooth them out. Since \mathbf{g} contains the haze-free image, we want to regularize patches with the same amount of haze in a similar way. Now, as channel attention modules mimic the behaviour of nonlocal regularization terms (Pereira-Sánchez et al., 2024), we propose to substitute the ResNet used to compute \mathbf{g}^k by a residual channel attention network

(RCANet). Thus, we can now unfold (6) as

$$\begin{cases} \mathbf{g}^{k+1} = \text{RCANet}^k(\mathbf{G}^k), \\ d^{k+1} = \text{ResNet}^k(D^k). \end{cases} \quad (8)$$

Again, the residual networks do not share weights between stages and the hyperparameters are randomly generated and learned during the training phase. The structure of the new network can be seen in Figure 2a. The preprocessing block, the basic blocks are the same as in Figures 1b and 1d, respectively. In Figure 2b, we can see the residual channel attention network (RCANet) used to compute \mathbf{g}^k . The residual network used to compute d^k has the same structure as the one presented in Figure 1c. Figure 2c shows the last layer of the RCANet, the channel attention layer. From now on, this network will be referred to as URCANet (Unfolded Residual Channel Attention Network).

4 EXPERIMENTAL RESULTS

For the performance evaluation, we will use the RESIDE-Standard dataset (Li et al., 2019). We have selected the SOTS-outdoor set, which comprises 500 pairs of outdoor hazy images and their corresponding ground truths. These pairs have been divided into 70% for training, 15% for validation, and 15% for testing.

We compare our dehazing models with various landmark and state-of-the-art methods. Specifically, we compare with He’s dark channel prior (DCP) (He et al., 2011), since we use their estimations for t_0 and \mathbf{A} ; Fang et al.’s variational model (Fang et al., 2014), since our variational framework is based on it; two physical-model-based networks, DehazeNet (Cai et al., 2016) and AODNet (Li et al., 2017); and two straightforward hazy-to-clear networks, FFA-Net (Qin et al., 2019) and ConvIR (Cui et al., 2024). AODNet, FFA-Net and ConvIR have been downloaded from their respective GitHub repositories, while the other models have been implemented in Pytorch from scratch. All methods, including ours, have been trained during 1000 epochs using an ADAM optimizer with a learning rate of 10^{-5} . For more details about the configurations of the two proposed unfolded networks, we refer to the ablation study in Section 5.

Since ground truths are available, the metrics used for objective evaluation are *Peak Signal-to-Noise Ratio* (PSNR), *Structural Similarity Index Measure* (SSIM), and *Spectral Angle Mapper* (SAM).

Table 1 displays the average PSNR, SSIM, and SAM values of each method on the testing set. The proposed models yield the best results in terms of

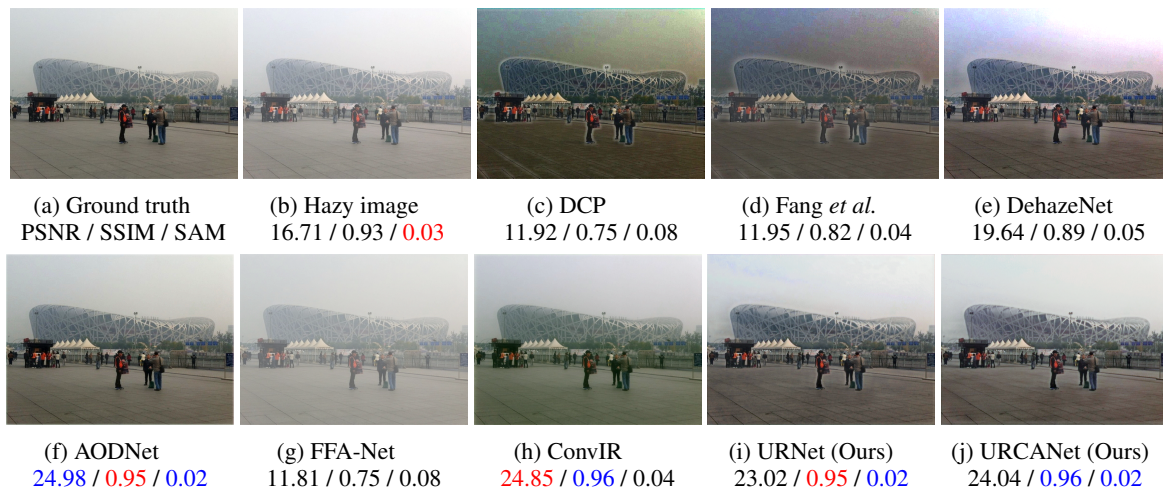


Figure 3: Visual comparison of dehazing methods on an image of the testing set. DCP and Fang et al.’s effectively remove the haze, but tend to darken the image and produce artifacts around the edges. DehazeNet and ConvIR yield a clear output but fail to correctly balance the colors of the scene. AODNet and our models provide the best visual results, with ours offering superior color recovery.

Table 1: Quantitative comparison of various dehazing methods on the testing set of SOTS-outdoor dataset. We highlight in blue the best result, and red the second best.

Method	PSNR \uparrow	SSIM \uparrow	SAM \downarrow
DCP	11.863	0.785	0.075
Fang et al.	15.746	0.824	0.076
DehazeNet	16.382	0.864	0.067
AODNet	21.955	0.874	0.059
FFA-Net	11.221	0.689	0.087
ConvIR	22.370	0.903	0.113
URNet (Ours)	22.045	0.906	0.058
URCANet (Ours)	21.921	0.912	0.058

SSIM and SAM, while our URNet ranks second best in terms of PSNR, just behind ConvIR. However, as illustrated in Figures 3 and 4, the dehazed images provided by ConvIR are undersaturated. We also observe that DCP and Fang et al.’s methods tend to darken the images and introduce artifacts around the edges. Among the deep learning models, all except FFA-Net effectively remove the haze. However, our two methods tend to improve color recovery. DehazeNet often oversaturates the images, returning warmer colors. Conversely, FFA-Net struggles significantly, introducing additional haze. This issue is likely due to being trained on a small dataset and suffering from overfitting. This highlights the importance of robust training protocols in achieving effective methods.

Finally, we also test the quality of the estimations on real life images from the LIVE Image Defogging Database (Choi et al., 2015), as shown in Figure 5. We can see that the results inherit the qualities and problems of the synthetic image testing. He’s, DCP and Dehazenet remove effectively the haze, but oversaturate the sky regions. Fang et al. removes the haze

but generates a strong halo around the edges. Both AODNet and ConvIR remove the haze, but inherit the color of the hazy scene. Our networks combine the strengths of the DCP and neural networks, resulting in a fully dehazed image with a better color balance, even though it also saturates the sky because of the violation of the DCP.

5 ABLATION STUDY

To optimize the configuration of our model and validate each component’s contribution to performance, we conduct several ablation studies. These focus on the network structure, hyperparameter impact, the necessity of preprocessing and postprocessing blocks, and the choice of loss function.

Concerning the structure of the URNet, we have chosen to use a ResNet because the proximal operator of a proper, lower semicontinuous and convex function R can also be defined as a resolvent operator:

$$\text{prox}_{\tau R}(\cdot) = (Id + \tau \partial R)^{-1}(\cdot). \quad (9)$$

Then, we have trained the model varying the number of blocks of the ResNet architecture, stages and features chosen. After this study, we have settled for 3 stages, 64 features and 3 basic blocks. With this setting, our model has 1.3M parameters. With the same configuration and stages, we have then computed \mathbf{g} with residual channel attention blocks instead of residual blocks and trained the URCANet. With this setting, the model has 2M parameters.

After the main structure of the network has been chosen, we study possible pre and postprocessing



Figure 4: Visual comparison of dehazing methods on an image of the testing set. DehazeNet and AODNet produce a haze-free images with a warmer tone, while ConvIR results in undersaturated colors. In contrast, our models provide a better color balance, with URNet being the most faithful and URCANet exhibiting a slightly cooler tone.

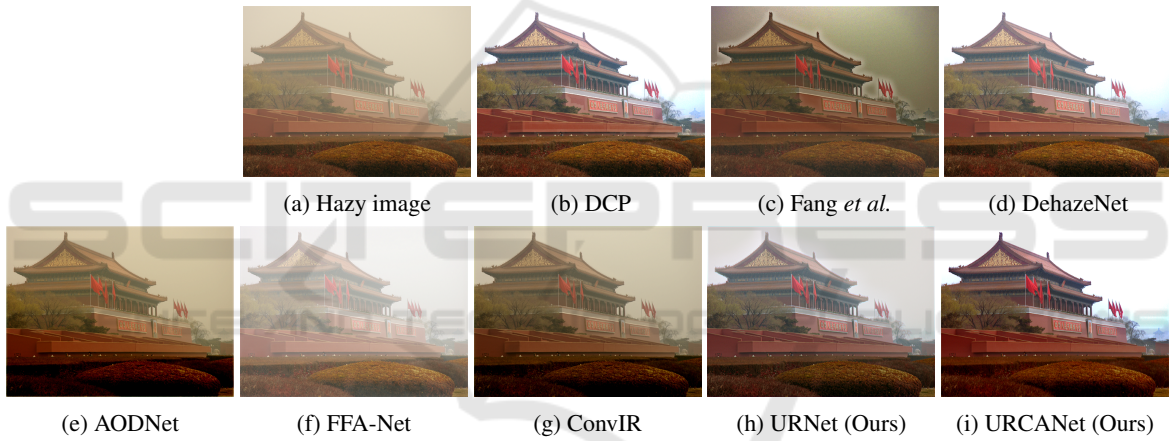


Figure 5: Visual comparison of dehazing methods on a real life image. DCP and DehazeNet remove effectively the haze, but oversaturate the sky, as the dark channel prior hypothesis are violated. Fang *et al.* removes the haze but generates a strong halo around the building. AODNet and ConvIR both remove the haze, but not completely. Both URNet and URCANet remove the haze and the results have a better color balance, even though it saturates the sky because of the violation of the DCP.

blocks. For the preprocessing, we mainly compare the guided filter used by He *et al.* in (He *et al.*, 2011) and the GIF Block in (Yang and Sun, 2018) to refine the transmission map. We have realised that He’s guided filter does not effectively refine the borders of the image, as can be seen in Figure 6, so we have chosen the GIF block to filter t_0 . For the postprocessing, we have studied how different blocks affect the output of the algorithm. The candidates tested are: a residual block (RESB), the same one depicted in Figure 1d, as a denoising tool; a channel attention block (CAB) (Woo *et al.*, 2018) to focus on recovering correctly the colors; a spatial attention block (SAB) (Woo *et al.*, 2018) to address any possible problem resulting on d ’s estimation; and no block at all. Among them, the residual block yielded the best results in terms of PSNR and

SSIM, as can be seen in Table 2.

Table 2: Comparison of PSNR values obtained from different postprocessing blocks during the first 100 epochs.

	RESB	CAB	SAB	No block
PSNR \uparrow	21.97	20.83	20.87	20.86
SSIM \uparrow	0.902	0.902	0.901	0.902
SAM \downarrow	0.075	0.067	0.067	0.066

Last, we discuss the loss function. Let J be the recovered image and GT the ground truth. Our first idea was to use either the L^1 or the MSE, that is,

$$L^1(GT, J) = \|GT - J\|_1, \quad (10)$$

or

$$MSE(J, GT) = \|GT - J\|_2^2, \quad (11)$$

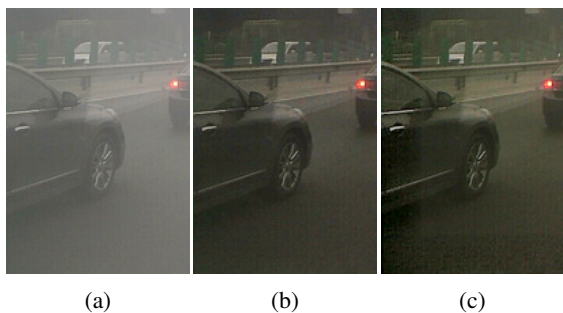


Figure 6: (a) Hazy image (b) Output using a GIF Block (c) Output using He's guided filter. Here we see how the borders on (c) are not fully refined, with a darker frame appearing on the borders of the image.

respectively. We found that the L^1 norm was a better choice, as it did not smooth the edges like the MSE and preserved color more effectively. However, neither of them recovered correctly the edges, causing the appearance of halos around the objects. Then, we

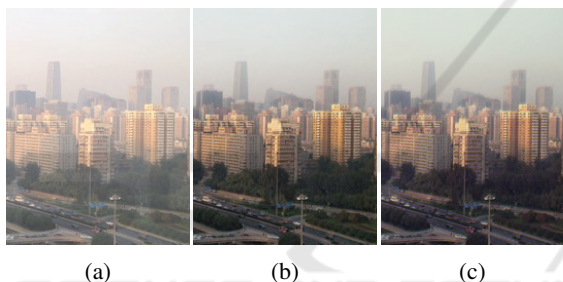


Figure 7: (a) Hazy image (b) Output using (12) as loss (c) Output using MSE instead of L^1 in (12). We see how (b) recovers a sharper image with a correct color balance, while (c) is a bit undersaturated.

added a weighted sum of the loss at each stage of the unfolded algorithm. Again, L^1 performs better than MSE (see Figure 7). In the end, the final loss is set to

$$\mathcal{L}(\mathbf{GT}, \mathbf{J}, \{\mathbf{J}_i\}_{i=1}^{N-1}) = L^1(\mathbf{GT}, \mathbf{J}) + \frac{\omega}{N} \sum_{i=1}^{N-1} L^1(\mathbf{GT}, \mathbf{J}_i), \quad (12)$$

where N is the total number of stages and ω is a constant. After different tests, we set $\omega = 0.3$ as it balances edge preservation and overall image quality.

6 CONCLUSIONS

In this paper, we have proposed two simple unfolded residual networks for single-image dehazing. In both cases, we have designed an energy functional to be minimized via proximal gradient descent. On one hand, this gives us a solid mathematical foundation

and a clear interpretation of all the variables involved in the problem. However, the derivation of this functional involved imposing some restrictive priors on the fidelity terms, such as d being close to d_0 . This could compromise the results when such hypothesis are violated or, for instance, if t_0 is not accurately estimated. However, the unfolding process addresses some of these problems, which can be seen comparing Fang's classical variational model with ours in Figures 3-5.

The results demonstrate that laying a robust mathematical framework not only aids in understanding the modeling process but also facilitates the development of efficient, interpretable neural networks that perform comparably to state-of-the-art methods. Although many models prioritize performance over interpretability, our approach shows that sometimes taking a step back to lay a solid foundation can result in simpler and more effective solutions.

ACKNOWLEDGMENTS

This work is part of the MoMaLIP project PID2021-125711OB-I00 funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGeneration EU/PRTR.

REFERENCES

- Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D. (2016). Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198.
- Chambolle, A. and Pock, T. (2016). An introduction to continuous optimization for imaging. *Acta Numerica*, 25:161–319.
- Chi, J., Li, M., Meng, Z., Fan, Y., Zeng, X., and Jing, M. (2020). Single image dehazing using a novel histogram transformation network. In *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5.
- Choi, L. K., You, J., and Bovik, A. C. (2015). Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Transactions on Image Processing*, 24(11):3888–3901.
- Cui, Y., Ren, W., Cao, X., and Knoll, A. (2024). Revitalizing convolutional network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–16.
- Fang, C., He, C., Xiao, F., Zhang, Y., Tang, L., Zhang, Y., Li, K., and Li, X. (2024). Real-world image dehazing with coherence-based label generator and cooperative unfolding network. *arXiv preprint arXiv:2406.07966*.

- Fang, F., Li, F., and Zeng, T. (2014). Single image dehazing and denoising: A fast variational approach. *SIAM Journal on Imaging Sciences*, 7(2):969–996.
- Fattal, R. (2008). Single image dehazing. *ACM Trans. Graph.*, 27(3):1–9.
- Galdran, A., Bria, A., Alvarez-Gila, A., Vazquez-Corral, J., and Bertalmío, M. (2018). On the duality between retinex and image dehazing. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8212–8221.
- Galdran, A., Vazquez-Corral, J., Pardo, D., and Bertalmío, M. (2015). Enhanced variational image dehazing. *SIAM Journal on Imaging Sciences*, 8(3):1519–1546.
- Galdran, A., Vazquez-Corral, J., Pardo, D., and Bertalmío, M. (2017). Fusion-based variational image dehazing. *IEEE Signal Processing Letters*, 24(2):151–155.
- Golts, A., Freedman, D., and Elad, M. (2020). Unsupervised single image dehazing using dark channel prior loss. *IEEE Transactions on Image Processing*, 29:2692–2701.
- Guo, X., Yang, Y., Wang, C., and Ma, J. (2022). Image dehazing via enhancement, restoration, and fusion: A survey. *Information Fusion*, 86-87:146–170.
- He, K., Sun, J., and Tang, X. (2011). Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353.
- He, K., Sun, J., and Tang, X. (2013). Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6):1397–1409.
- Jackson, J., Agyekum, K. O., kwabena Sarpong, Ukwuoma, C., Patamia, R., and Qin, Z. (2024). Hazy to hazy free: A comprehensive survey of multi-image, single-image, and cnn-based algorithms for dehazing. *Computer Science Review*, 54:100669.
- Jin, Z., Ma, Y., Min, L., and Zheng, M. (2024). Variational image dehazing with a novel underwater dark channel prior. *Inverse Problems and Imaging*.
- Jun, W. and Rong, Z. (2013). Image defogging algorithm of single color image based on wavelet transform and histogram equalization. *Applied Mathematical Sciences*, 7:3913–3921.
- Lei, L., Cai, Z.-F., and Fan, Y.-L. (2024). Single image dehazing enhancement based on retinal mechanism. *Multimedia Tools and Applications*, 83(21):61083–61101.
- Li, B., Peng, X., Wang, Z., Xu, J., and Feng, D. (2017). Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE international conference on computer vision*, pages 4770–4778.
- Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., and Wang, Z. (2019). Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505.
- Li, P., Tian, J., Tang, Y., Wang, G., and Wu, C. (2021). Deep retinex network for single image dehazing. *IEEE Transactions on Image Processing*, 30:1100–1115.
- Liu, Q., Gao, X., He, L., and Lu, W. (2018). Single image dehazing with depth-aware non-local total variation regularization. *IEEE Transactions on Image Processing*, 27(10):5178–5191.
- Liu, Y., Yan, Z., Wu, A., Ye, T., and Li, Y. (2022). Nighttime image dehazing based on variational decomposition model. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 639–648.
- McCartney, E. J. (1977). Optics of the atmosphere: Scattering by molecules and particles. *Physics Bulletin*, 28(11):521.
- Pereira-Sánchez, I., Sans, E., Navarro, J., and Duran, J. (2024). Multi-head attention residual unfolded network for model-based pansharpener. *arXiv preprint arXiv:2409.02675*.
- Qin, X., Wang, Z., Bai, Y., Xie, X., and Jia, H. (2019). Ffanet: Feature fusion attention network for single image dehazing. *CoRR*, abs/1911.07559.
- Qu, Y., Chen, Y., Huang, J., and Xie, Y. (2019). Enhanced pix2pix dehazing network. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8152–8160.
- Song, Y., He, Z., Qian, H., and Du, X. (2023). Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 32:1927–1941.
- Stipetić, V. and Lončarić, S. (2022). Variational formulation of dark channel prior for single image dehazing. *J. Math. Imaging Vis.*, 64(8):845–854.
- Tan, R. T. (2008). Visibility in bad weather from a single image. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8.
- Thanh, L. T., Thanh, D. N. H., Hue, N. M., and Prasath, V. B. S. (2019). Single image dehazing based on adaptive histogram equalization and linearization of gamma correction. In *2019 25th Asia-Pacific Conference on Communications (APCC)*, pages 36–40.
- Wang, W. and Yuan, X. (2017). Recent advances in image dehazing. *IEEE/CAA Journal of Automatica Sinica*, 4(3):410–436.
- Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). Cbam: Convolutional block attention module.
- Xiao, B., Zheng, Z., Zhuang, Y., Lyu, C., and Jia, X. (2024). Single uhd image dehazing via interpretable pyramid network. *Signal Processing*, 214:109225.
- Yang, D. and Sun, J. (2018). Proximal dehaze-net: A prior learning-based deep network for single image dehazing. In Ferrari, V., Hebert, M., Sminchisescu, C., and Weiss, Y., editors, *Computer Vision – ECCV 2018*, pages 729–746, Cham. Springer International Publishing.
- Zhang, H. and Patel, V. M. (2018). Densely connected pyramid dehazing network. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3194–3203.
- Zhou, J. and Zhou, F. (2013). Single image dehazing motivated by retinex theory. In *2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation (IMSNA)*, pages 243–247.