Comparative Analysis on Real-time Video Dehazing models

Domain Specific Project

Pratik Patil
Data Science
DA-IICT
Pune, India
202118023@daiict.ac.in

Tanya Jagyasi
Data Science
DA-IICT
Ahmedabad, India
202118039@daiict.ac.in

***Abstract*—The presence of atmospheric haze, fog and smoke makes observed objects less visible and less contrasted, which is problematic for outdoor photography and computer vision applications. Comparing a set of techniques/algorithms to remove haze from pictures and videos, we estimated the global atmospheric light, extract the scene object transmission, along with estimate the dark channel using new haze-free image priors and deep learning models for single image input. Further, developing these models to work in real-time and conducting comparative analysis to determine which model produces the best results in real time.**

***Index Terms*—Video Dehazing, Color Attenuation Prior, Guided Image Filtering, Image Processing, Atmospheric light, Deep Learning, GPU**

I. Introduction

"Images or videos recorded outside in severe weather usually have low visibility. Aerosols scatter in the atmosphere and absorb light before it reaches the observer, and the light that reaches the camera is also mixed with light reflected from other directions, known as the air light." These degraded photographs frequently lack visual appeal and vividness since the color of the seen objects fades and the contrast is reduced as a result of this process.

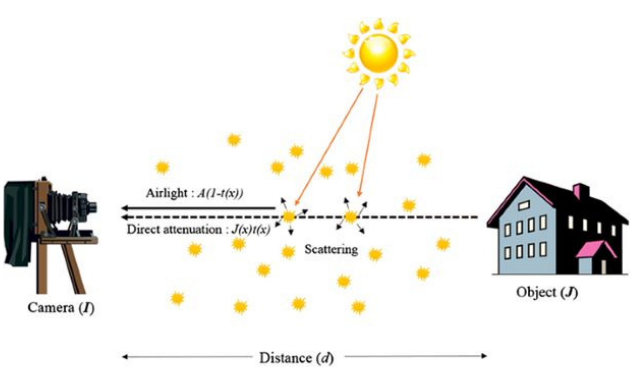


Fig. 1: Iterative Dehazing method [6]

We create a dehazed image from a foggy one first. The depth map is then estimated at the most basic level. The transmission map and atmospheric light are then obtained at the highest resolution using iterative upsampling and guide image filtering to prevent information loss. Additionally, by making transmission values temporally coherent, we extend the single-image dehazing algorithm to real-time video dehazing to lessen flickering artifacts in dehazed videos. The proposed algorithm can be used in real-time applications, according to experimental results, and offers performance that is equivalent to or even superior to that of state-of-the-art algorithms.

II. Background / Literature Review

In this paper, we present some improved methods for real-time image and video dehazing by studying several papers and selecting the appropriate models for single-image dehazing, which were further converted to real-time.

A. Models Used:

1) Real-time Video Dehazing using Color Attenuation Prior (CAP): The model proposes a novel color attenuation prior for single image dehazing in this research [1]. This straightforward and effective prior can aid in the development of a linear model for the hazy image's scene depth. The bridge between the hazy image and it's associated depth map is efficiently generated by learning the parameters of the linear model using a supervised learning method. We can easily remove the haze from a single foggy image using the retrieved depth information. Fig. 8 below depicts an overview of the suggested dehazing process. This dehazing approach has a significantly higher efficiency and dehazing effectiveness than other existing dehazing algorithms.

2) Image Dehazing using Boundary Constraint and Contextual Regularization (BCCR): The model proposed an efficient way for removing hazes from a single image in this research [2]. Exploration of the transmission function's underlying boundary constraint enhances the method greatly. To retrieve the unknown transmission, this constraint, along with a weighted L1-norm based contextual regularization, is posed as an optimization problem. To overcome the optimization problem, an efficient technique based on variable splitting is also proposed. This approach can produce visually appealing outcomes with more accurate color and finer image details and structures. There is a derived boundary limitation on the transmission of an image from the radiance cube from a geometric standpoint of image dehazing. Although the boundary restriction imposes a significantly weaker limitation on the dehazing process, it is remarkably successful for the dehazing of most natural photos when paired with contextual regularization. More broadly, a tighter radiance envelop, not confined to a cubic shape, can be used to provide a more accurate constraint on transmissions. This may help to eliminate the ambiguity between color and depth and avoid numerous erroneous image enhancements.

Another approach to dealing with ambiguity is to add more sound restrictions or create new picture priors, such as employing scene geometry or directly using available depth information into scene transmission estimates.

3) Real-time Video Dehazing using All in One Dehazing (AOD): This research proposed the All in-One Dehazing Network, an efficient end-to-end dehazing convolutional neural network(CNN) model(AOD-Net) [**3**]. While some prior haze removal models described the ”end-to-end” notion, we believe AOD-Net is unique in that it is the first to optimize the end-to-end pipeline from hazy photos to clean images, rather than an intermediate parameter estimation phase. AOD-Net is built around a reformulated atmospheric scattering model. It is trained on blurry synthetic images and evaluated on both synthetic and natural photos. Experiments show that AOD-Net outperforms numerous state-of-the-art approaches.

The proposed AOD-Net is divided into two parts: A K-estimation module that employs five convolutional layers to estimate K(x), and a clean image production module that employs an element-wise multiplication layer and several element-wise addition layers to generate the recovery image (Fig. 2).

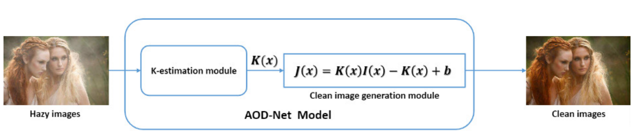


Fig. 2: Structure of AOD-Net

The K-estimation module is the critical component of AOD-Net, being responsible for estimating the depth and relative haze level. As depicted in Fig. 3, the model uses five convolutional layers, and form multi-scale features by

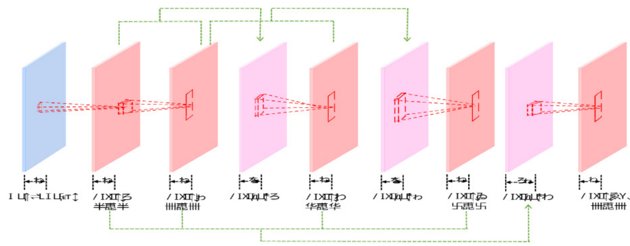


Fig. 3: K-estimation of AOD-Net

fusing varied size filters and used parallel convolutions with varying filter sizes.

4) Image Dehazing using GMAN net: This paper takes a different, more agnostic approach to the dehaze problem, presenting a dehaze neural network that simply focuses on producing a haze-free version of the input image [**4**]. It makes use of current breakthroughs in deep learning to construct an encoder-decoder network architecture shown in figure below (Fig.4) below that has been taught to directly restore the clear image while avoiding the parameter estimation problem entirely. The proposed method has the potential to recognize complicated haze structures in the training data that the atmospheric scattering model did not capture.

The first and second layers are made out of 64-channel convolutional blocks. They are followed by two-step downsampling layers that encode the input image into a 56x56x128 volume. The encoded image is then put into a residual layer composed of four residual blocks, each with a shortcut connection (see Fig. 5). This layer symbolizes the transition from encoding to decoding because it is followed by the deconvolutional layer, which upsamples the residual layer output and reconstructs a fresh 224x224x64 volume for another round of convolutions. The final two layers are made up of convolutional blocks. The upsampled feature maps are converted into an RGB image, which is then added to the input image and thresholded using a ReLU to obtain the haze-free version.

The experimental results validated GMAN’s ability to generate haze-free images and demonstrated that it is capable of overcoming some of the frequent problems of cutting-edge approaches, such as color darkening and excessive edge sharpening. Furthermore, because of GMAN’s generic architecture, it may pave the way for

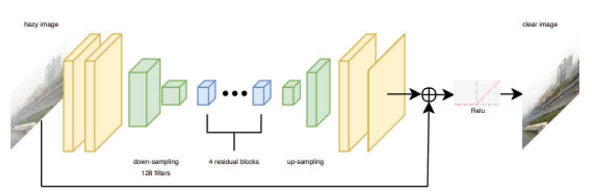


Fig. 4: Structure and details of GMAN

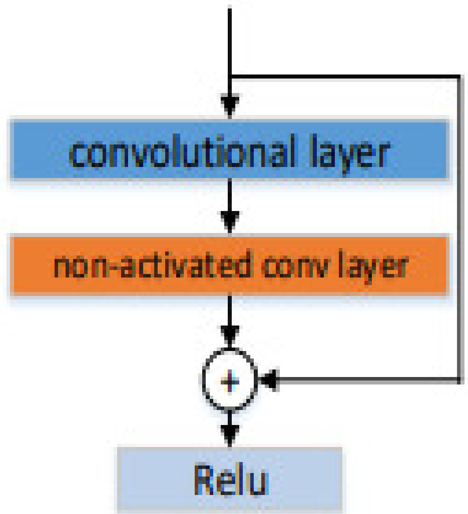


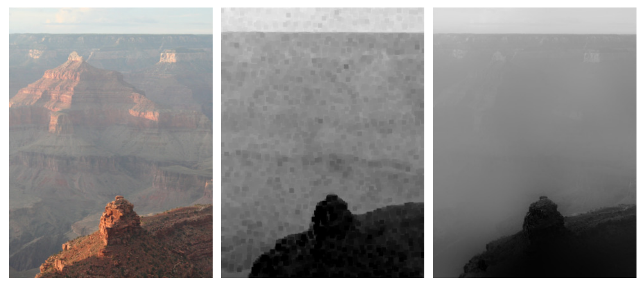
Fig. 5: A residual block used in the middle layer of the proposed GMAN. Relu is used as the activation function after the addition operator of every block

future research on general-purpose image restoration.

5) A Novel Image Dehazing and Assessment Method (Using DCP): The model estimate transmission map and the unknown value K associated with the additive component in the traditional picture degradation model and use Guided Filtering in this research [**5**]. The 'Airlight Co-efficient' is the name given to this constant. While various approaches have been developed to estimate the global airlight constant A, this model estimates the value of K, which is the proportionality constant and whose value is dependent on the particular type and form of the scattering function. The method's advantage is that the result of this alteration is better color contrast when compared to other cutting-edge algorithms. Furthermore, the strategy works for dense haze as well. The estimation of airlight co-efficient method suggested in this research yields some good results, as evidenced by quantitative and qualitative analyses, but we also discovered that in some circumstances, this method underperformed state-of-the-art methods. As a result, other researchers and developers will struggle to discover better estimates for the airlight co-efficient. One advantage of our outcomes is their high visibility and brightness. Another advantage of this research is the metrics based on haze theory, which quantifies how much haze is removed and should be used to evaluate haze removal algorithms.

B. Important Functionalities:

1) Estimation of Depth Information (Depth Map): Dept map calculates and represents the distance of the subject from the camera for each pixel in the image. It can be said to be the relationship between three factors: screen depth d, brightness of image v, and saturation s [**1**]. For instance, white objects in an image usually have high brightness and low saturation and hence, image processing models consider scene objects with white color to be distant from the camera (see Fig. 6(b)). But as this may lead to misclassification, which may result in inaccurate estimation of depth information. To avoid this, we can use a guided image filter to smooth the image and refine the depth map. Figure shows the final restored depth map of the hazy image (Fig. 6(c)).



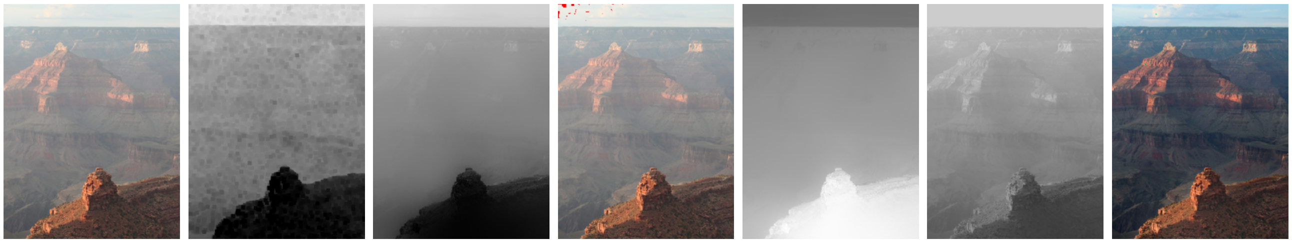
(a) (b) (c)

Fig. 6: (a) Input Image (b) Original Depth Map (c) Refined Depth Map

2) Estimation of Atmospheric Light: Atmospheric light is a source of illumination in the natural environment in the image. Atmospheric light can be easily calculated from the depth map [**1**]. In the estimated depth map, bright regions stand for distant places in the image. We can choose a very small percent (generally less than 0.5 percent(%)) of these brightest pixels in the depth map, and select the pixel with highest intensity in the corresponding hazy image I as the atmospheric light A.
Fig. 7(b) illustrates the estimated atmospheric light with the brightest pixels in red color (top-left corner)

(a) (b)

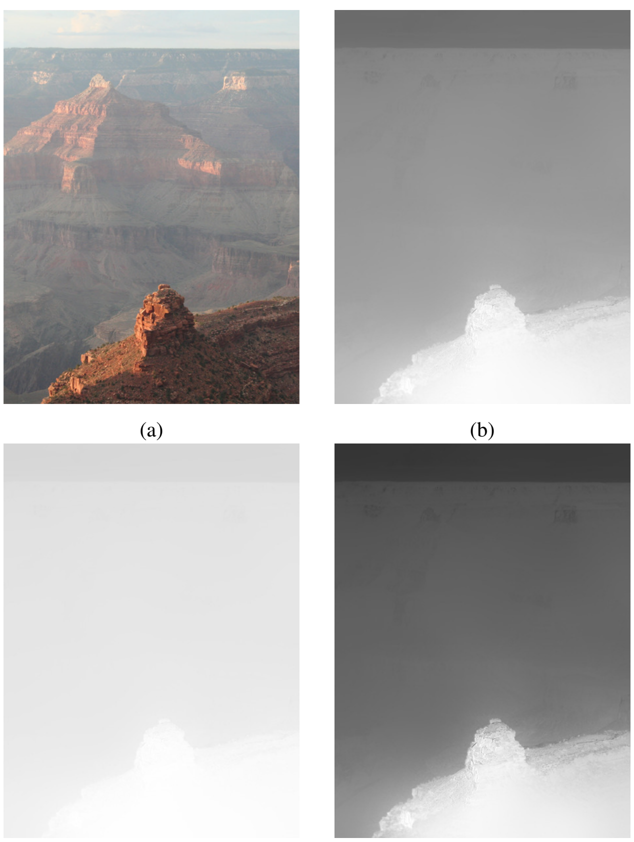
Fig. 7: (a) Input Image (b) Estimation of Atmospheric Light



(a) Input Image (b) (c) (d) (e) (f) (g) Output

Fig. 8: Single Image Dehazing using various filters
where, (b) Depth Map, (c) Refined Depth Map, (d) Estimated Atmospheric Light, (e) Transmission Map, (f) K Map

3) Scene Radiance Recovery (Transmission map): After knowing the depth of scene and it's atmospheric light, we can easily estimate the medium transmission and recover the scene radiance in the image [**1**]. Note that the scattering coefficient β, which can be regarded as a constant in homogeneous regions, represents the ability of a unit volume of atmosphere to scatter light in all directions. In other words, β determines the intensity of dehazing indirectly.

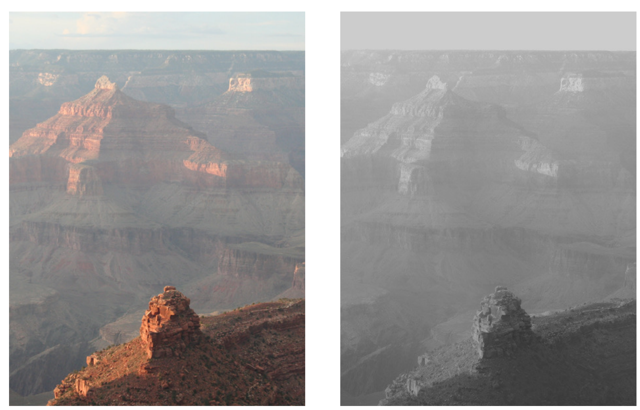


(a) (b) (c) (d)

Fig. 9: (a) Input Image (b) Transmission Map ($\beta = 1$) (c) Transmission Map ($\beta = 0.2$) (d) Transmission Map ($\beta = 1.8$)

As can be seen, on one hand, a small β leads to small transmission, and the corresponding result remains still hazy in the distant regions (see Figure 9(c)). On the other hand, a too large β may result in overestimation of the transmission (see Figure 9(d)). Therefore, a moderate β is required when dealing with the images with dense-haze regions.
In most cases, β = 1.0 is more than enough (Fig. 9(b)).

4) K map or Haze map (Refined Transmission map): The K map or haze map describes the portion of the light that is not scattered and reaches the camera [**5**]. Since the map is a continuous function of depth, it thus reflects the depth information in scene (Fig. 10).



(a) (b)

Fig. 10: (a) Input Image. (b) K Map or Haze Map

III. Comparitive Analysis

The comparative analysis has been performed on the above mentioned models. This analysis is based on various aspects of the output images, like, image quality, image clarity, dehazing capability and color contrast.
As we enhance further, transforming a single image dehazing model to real-time can be tricky. Depending on the time complexity for a single image, straight from reading it, doing some processing and up to displaying or saving the output image, the time complexity and optimization of the model plays a key role. Using the Open CV (cv2) and

Image Utilities (imutils) library in Python, we captured an input video sequence of hazy environment, thus processing on each and every frame to get a clear and dehazed output video sequence in real-time. As all models don't follow the same time complexity and are sophisticated in various ways, we identified three models that can be used for real-time video dehazing, namely:

1) Real-time Video Dehazing using Color Attenuation Prior (CAP)
2) Real-time Video Dehazing using All in One Dehazing (AOD)
3) A Novel Image Dehazing and Assessment Method (Using DCP)

Following are some examples on the comparative analysis:

A. *Single Image Dehazing:*

We conducted a comparison analysis to dehaze a single image using five different models: 1) Color attenuation prior (CAP), 2) Image dehazer function (BCCR), 3) All-in-one dehazing (AOD), 4) Gman-net, and 5) Dark channel prior (DCP), and the results are shown below (Fig.11 and Fig.12).

(a) Hazy Input

(b) CAP

(c) BCCR

(d) AOD Net

(e) GMAN Net

(f) DCP

Fig. 11: Comparative Analysis 1

(a) Hazy Input

(b) CAP

(c) BCCR

(d) AOD Net

(e) GMAN Net

(f) DCP

Fig. 12: Comparative Analysis 2

Here we used two images to conduct a comparative research, one image taken from Internet (Fig. 11) and another one is the actual image taken from a camera (Fig. 12). We determined that the fifth model(DCP), which employs the dark channel, outperforms all others. It dehazes the image appropriately, but it also enhances the image and produces a good outcome, whereas other models, such as gman-net, provide a blurry/grainy effect after dehazing. Although models such as AOD and BCCR produce good results, BCCR requires more processing time to dehaze the image. Also, when discussing CAP, it does not dehaze dense haze in the image. CAP can readily dehaze close-up photos but give comparatively weak results for far-away images.

B. *Real-time Dehazing:*

We got the followiing results after applying these algorithms in real-time: (Ref. Fig. 13, 14 & 15)

These three methods were tested in different times depending on the availablity of resources, but in the same environment by creating artificial smoke (using incense sticks).

As seen in these figures (Fig. 13, 14 & 15) all the three methods give remarkable results and dehazed output sequence can be seen clearly in the display window. The white smoke / haze (in natural environment) can be

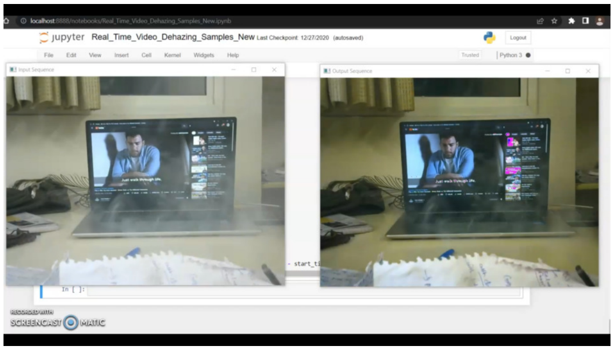


Fig. 13: Real-time dehazing using CAP

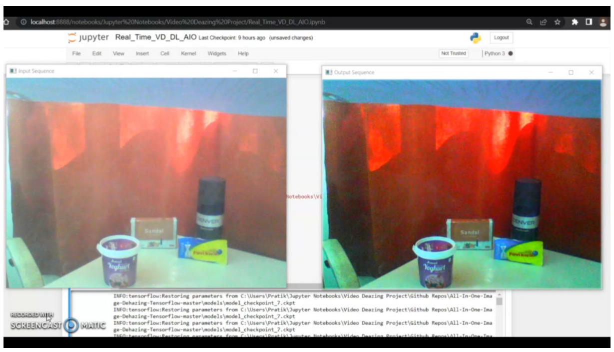


Fig. 14: Real-time dehazing using AOD

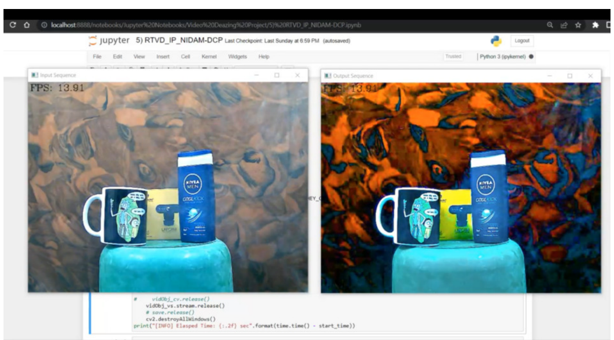


Fig. 15: Real-time dehazing using DCP

cleaned easily. The Color Attenuation Prior model cleans the image with respect to the scene depth from each frame, similarly, the All in One Dehazing algorithm uses the power of tensors to clean the extra white haze and light from the images. Also, as seen previously while dehazing single images, the Dark Channel Prior also gives good results while keeping a high contrast with the surrounding.

C. Performance Comparison:

Here, (Fig. 16 and Fig. 17) we can see the performance of each model on the basis of processing time and frames per second (FPS) of each model.

Model	Single Image	Recorded Video			
	Processing Time	Video Duration	Processing Time	No. of Frames	Processing time per Frame
CAP	5 sec	24 sec	7 min 34 sec	732	620 ms
BCCR	3.5 sec	24 sec	15 min 20 sec	732	1.25 sec
AOD	170 ms	24 sec	1 min 45 sec	732	140 ms
GMAN	1.5 sec	24 sec	6 min 4 sec	732	500 ms
DCP	400 ms	24 sec	3 min 58 sec	732	320 ms

Fig. 16: Model Comparison for Single image and Videos

Model	Real-time Dehazing
	FPS
CAP	5 - 7 FPS
AOD	7 - 9 FPS
DCP	9 - 11 FPS

Fig. 17: Model Comparison for Real-time

IV. Conclusion

On concluding we can say that, after performing tests we discovered that the potential of three models(CAP, AOD, DCP) in generating haze-free images is better than other models in almost all scenarios, whether it is for single image dehazing, static video dehazing, or real-time video dehazing and have a low latency along with good frame rate when used in real-time, and it is also shown that it is capable of overcoming some of the common pitfalls of state-of-the-art models. Furthermore, when comparing performance based on processing time, these three models have less processing time and produce better outcomes than the other two. As a result, these three models perform better in both qualitative and quantitative evaluations.

References

[1] Q. Zhu, J. Mai and L. Shao, *"A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior"* in IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3522-3533, Nov. 2015, doi: 10.1109/TIP.2015.2446191.

[2] G. Meng, Y. Wang, J. Duan, S. Xiang and C. Pan, *"Efficient Image Dehazing with Boundary Constraint and Contextual Regularization"* 2013 IEEE International Conference on Computer Vision, 2013, pp. 617-624, doi: 10.1109/ICCV.2013.82.

[3] B. Li, X. Peng, Z. Wang, J. Xu and D. Feng, *"AOD-Net: All-in-One Dehazing Network"* 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4780-4788, doi: 10.1109/ICCV.2017.511.

[4] S. B. Sami, A. Muqeet and H. Tariq, *"A Novel Image Dehazing and Assessment Method"* 2019 8th International Conference on Information and Communication Technologies (ICICT), 2019, pp. 72-77, doi: 10.1109/ICICT47744.2019.9001939.

[5] Z. Liu, B. Xiao, M. Alrabeiah, K. Wang and J. Chen, *"Single Image Dehazing with a Generic Model-Agnostic Convolutional Neural Network"* in IEEE Signal Processing Letters, vol. 26, no. 6, pp. 833-837, June 2019, doi: 10.1109/LSP.2019.2910403.

[6] Sungmin Lee, Seokmin Yun, Ju-Hun Nam, Chee Sun Won and Seung-Won Jung *"zing algorithms"* Lee et al. EURASIP Journal on Image and Video Processing (2016) 2016:4 DOI 10.1186/s13640-016-0104-y