Image Dehazing using Deep Learning

MIDTERM REPORT

submitted by

ABU BAKR SIDDIQ & RAZEEM AKHTAR

(170929098 & 170929194)

Under the Guidance of

DR. MUNENDRA SINGH

Assistant Professor

Manipal Institute of Technology

DEPARTMENT OF MECHATRONICS

MANIPAL INSTITUTE OF TECHNOLOGY MANIPAL - 576104

APRIL, 2021

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

DEPARTMENT OF MECHATRONICS ENGINEERING

CERTIFICATE

This is to certify that the project titled *Image Dehazing using Deep Learning* is carried out by Abu Bakr Siddiq (170929098) and Razeem Akhtar (170929194) during Feb-May, 2021 and the Midterm report is submitted to the Department of Mechatronics Engineering as part of requirement of B.Tech. (Mechatronics Engineering) project work evaluation.

Dr. Munendra Singh
Guide
Assistant Professor
Department of Mechatronics
MIT Manipal

ABSTRACT

Many modern-day applications require a clear image to analyse and extract information, but generally the presence of a turbid medium hinders their performance. This has given rise to the field of Image dehazing. The dehazing methods can be extended to applications such as image enhancement, denoising, colour correction etc. The results can be also extended to depth correction models of an image which would aid in improving the performance of autonomous systems and make systems like self-driving cars achieve level 5 automation thereby making it function in most weather conditions.

As a part of our study, we looked at the SOTA models and implemented them on TensorFlow framework. After implementing them we performed a comparative study to identify best features of each model and used those unique features in developing a novel model based on Deep-Retinex model.

From the comparative study we could validate that Feature fusion attention network (FFA-net) was clearly a standout network. We also identified features such as attention blocks, dilations, multi-layer feature extraction and Mean Absolute error clearly improved the performance of dehazenet model when added appropriately.

From our work we have found that dehazenet is a simple dehazing model which if modified by adding the learnings of other models could compete with the best of dehazing models. We also have singled out Deep-Retinex model, as it incorporates concept that can be extended to Image dehazing.

Packages used: *TensorFlow, PyTorch, matplotlib, NumPy, OpenCV.* Software's used: *Jupyter Lab, Google colab, MATLAB.*

Keywords: *Image dehazing, image denoising, Depth correction, Deep-Retinex, Dehazenet*

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

SOTA	State of the Art
CNN	Convolutional Neural Network
PSNR	Peak Signal To Noise Ratio
SSIM	Structural Similarity Index Measure
MS SSIM	Multi-Scale Structural Similarity Index Measure
DCP	Dark Channel Prior
HSV	Hue Saturation Value
RESIDE	Realistic Single Image DEhazing
BLIINDS	BLind Image Integrity Notator
MSE	Mean Square Error
MAE	Mean Absolute Error
LCA	Light Convolutional Autoencoder
GMAN	Generic Model-Agnostic Convolutional Neural Network
GCA	Gated Context Aggregation Network
FFA	Feature Fusion Attention
LOL	Low light paired dataset
NSS	Natural Scene Statistic
MVG	Multi-Variate Gaussian
NIQE	Naturalness Image Quality Evaluator
FADE	Fog Aware Density Evaluator
CEIQ	Contrast-changed Image Quality measure

NOMENCLATURE

$d(x)$	distance from scene point (Depth Map)
$\theta_0, \theta_1, \theta 2$	Coefficients in Color Attention Prior Model
$v(x)$	Pixel value in value channel in HSV
$s(x)$	Pixel value in Saturation channel in HSV
$\epsilon(x)$	Random Variable
P_i	clear image patch (Synthesizer model)
P_j	hazy image patch (Synthesizer model)
t	transmission map
A	Atmospheric light
$I(x)$	Observed Intensity (Hazy Image)
$J(x)$	Scene Radiance (Dehazed Image)
$\beta(x)$	Scattering coefficient

CHAPTER 1

INTRODUCTION

1.1 *General Discussion*

The quality of an image is critical to performance of many algorithms as well as human judgement. Today with the rise of air pollution there has been a rapid increase in particulate matter present around us, which causes increased atmospheric scattering thereby hindering the ability to see clearly. Single image dehazing models propose a way to combat this problem by reducing the haze density to a certain extent such that the depth and aesthetic information in the dehazed image is as close to a clear image retaining the naturalness and structure of image. Haze if entirely removed not only makes the image unnatural but also reduces the depth information that a clear image has, as there is always a need for certain amount of haze so that human eye can perceive depth. Most dehazing assume the atmospheric scattering model as the base model, so we also start with the same model. As described in many works the problem of dehazing is a non-trivial one especially when we try to improve the image quality with only one image which in most cases fails to provide enough information.

1.2 *Present Day Scenario*

The increasing research in image dehazing is well highlighted with the growing research in the field.The models for image dehazing proposed fall under three categories image enhancement, image fusion and image restoration. The most popular model for convention dehazing falls under image restoration which are based on a prior about hazy images, [1] assumes that Dark channel for hazy images is lighter that dark channel for haze free images. This prior based model is one of the ground breaking ones as it formed the basis of future work in the field. There have been other approaches based on polarization and filtering .

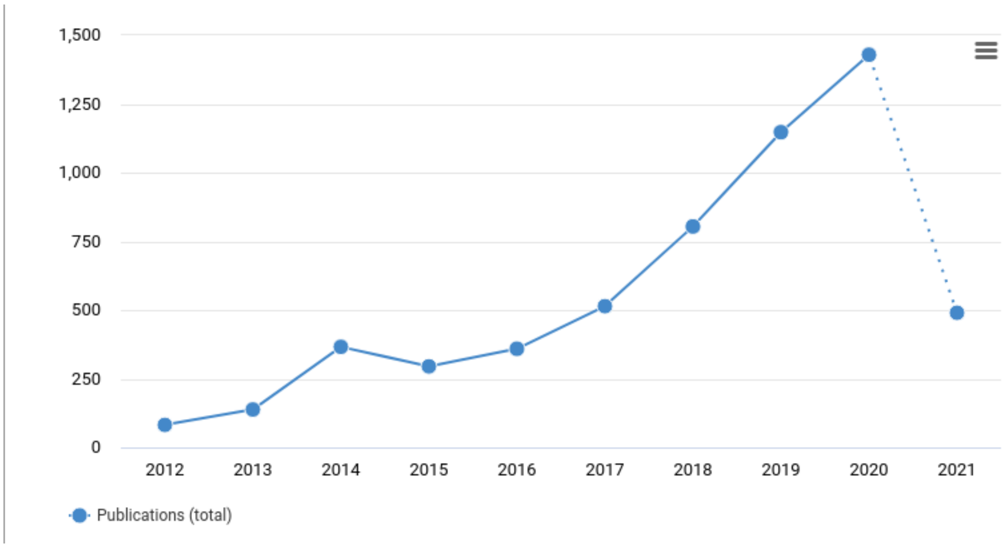

Figure 1.1: Growing Trend in Image dehazing

app.dimensions.ai

With the availability of better resources there was a boom in deep learning models which fall into the image enhancement category. These models could give satisfactory results but failed to consider the features specific to haze and therefore the were overshadowed by image restoration based deep network which considered the properties of haze. Some of the current SOTA models like DehazeNet, gated context attention network and feature fusion attention networks which incorporated haze specific features have given amazing performance with respect to metrics such as PSNR and SSIM.

1.3 *Motivation*

Early in 2020 some of the statistics about air pollution by IQAir caught our attention. According to the report 22 out of 30 most polluted cities are in India. There have been any efforts to combat this problem, but we wanted to develop a solution which would help people in the short term till long term solution take effect.

Figure 1.2: Fog on roads of Delhi

Hindustan Times

The image perfectly sums up the problem we are trying to address. There has already been a lot of work done in this area and the literature review section (2) goes into depth analyzing many popular models and highlighting their shortcoming. Most models assume a homogeneous atmospheric light model and we wanted to test this assumption in our work. We also wanted to strike a balance between computation resource, time and quality of output.

We looked at some of the popular websites for comparing the current models and found that there was no uniform comparison report available for comparison, websites such as www.paperswithcode.com compared models on the basis of PSNR and SSIM which is the most basic criteria. So, we through our work have tried to compare models based on haze specific metrics discussed in the methodology section (6). We also have tried to adopt best feature and metrics from field of image processing, low light enhancement and image dehazing.

1.4 *Significance of results*

The model developed on image dehazing can also be extended to image denoising, image demosaising and image deraining. The algorithm will also have a impact on the field of computer vision in general as this model would improve the accuracy models in the field of image segmentation, object recognition and image enhancement.

1.5 *Project Work Schedule*

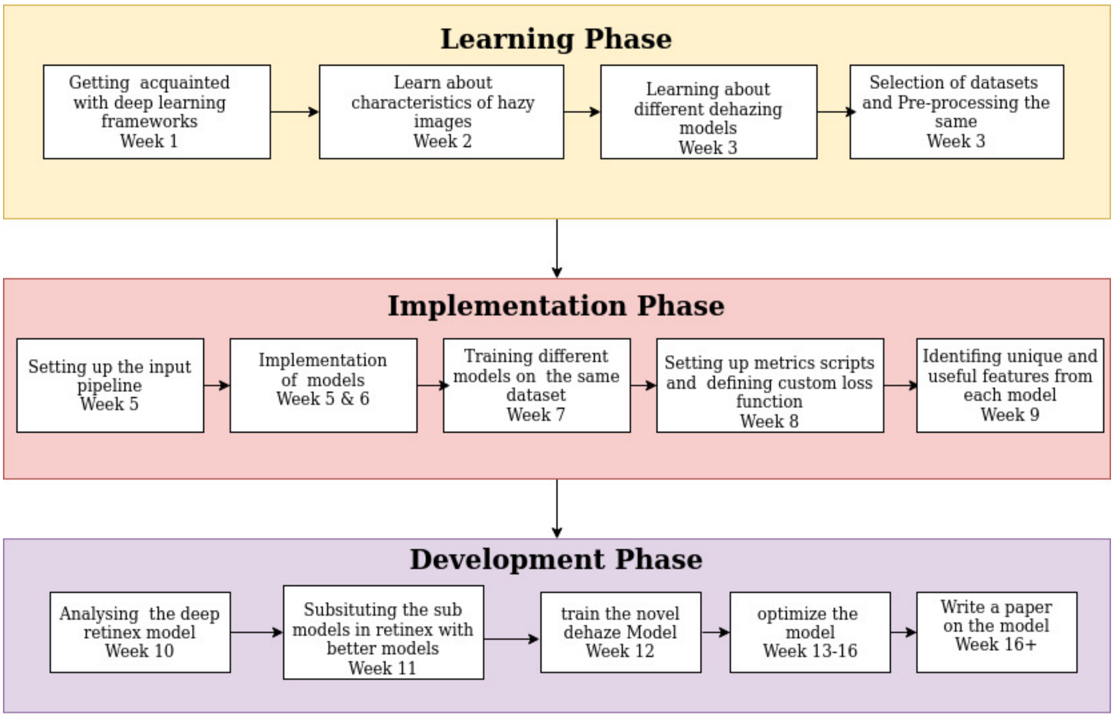

Figure 1.3: Project schedule

Currently we are in development phase (week 10).

CHAPTER 2

LITERATURE REVIEW

The problem of single image dehazing is a non-trivial one and many approaches have been used in solving it. Most works start off with the atmosphere scattering model (5).The problem encountered in single image is that its combination of three factors namely transmission map, atmospheric light,input image and to get the dehazed image there is not enough information available. Many conventional approaches get around this by coming up with suitable priors.

2.1 *Conventional approaches*

He et al. [1] proposes a dark channel prior, which consists of minimum intensities over 3 channels over a certain neighborhood. The assumption is that the dark channel for hazy images is brighter than the dark channel for non-haze images because the haze component is generally white, so it adds equal intensities to all the color channels. He et al. [1] also assumes that the hazing is homogeneous and uses 0.1% of brightest pixels in dark channel and compares them and the pixel having highest value in input image is considered as atmospheric light. The paper also suggests using soft matting to refine the transmission map and improve the results, the soft matting function increases complexity and the time required. In further works it has been shown by He et al. [2] that the guided filter if used instead of soft matting, improves the performance of the DCP algorithm.

Zhu et al. [3] is based on the color attention prior based approach. The paper observes that the density of haze increases with the increase in depth, and makes the assumption that depth is positively correlated to the concentration of depth which in turn depends upon the difference between value and saturation of the pixel in Hue Saturation Value (HSV) space. The paper proposes a linear model for modelling the scene depth

$$d(x) = \theta_0 + \theta_1 v(x) + \theta_2 s(x) + \varepsilon(x) \tag{2.1}$$

Where d(x) is depth map of a pixel, v(x) and s(x) are pixel values of value and saturation in HSV format. $\theta_0, \theta_1, \theta_2$ are the coefficients of the model. For dehazed image recovery the scattering coefficient is assumed constant and atmospheric light is assumed in th same way as done in [1].

The above prior based methods are successful to a certain degree in recovering the dehaze image but both suffer from some notable drawbacks. [1] generally over enhances the contrast of image, it also suffers from halo effects. [2] suffers from color distortion and background noise. Both the approaches are based on prior which may generally be useful but may not always hold true. This shortcoming pave way for approaches based on deep learning models.

2.2 *Deep learning approaches*

The major advantage with deep learning approach is that the model can itself come up with suitable priors based on the training methods and datasets.

2.2.1 *Dataset*

Datasets play a critical part in the performance of deep learning models.

For this reason, we looked at Tang et al. [4]. The paper's aim is to identify the best features combination for dehazing, the paper also looks at features such as multi-scale dark channel, multi-scale local max contrast, hue disparity and multi-scale local max saturation. The paper observers that dark channel is positively correlated to amount of haze and color saturation is negatively correlated. The most interesting part of the work is that it proposes a way to develop a synthetic hazy image as for a scene it is very difficult to have a haze free and hazy pair. The synthesizer assumes that depth is locally constant. The synthesizer takes a non hazy patch (P_i), a transmission t and atmospheric light A, and outputs the hazy image patch (P_j) according to the equation:

$$P_i = tP_j + (1 - t)A \tag{2.2}$$

Here the transmission map is assumed as a random value between 0.3 and 0.7 thereby helping us generate image with variable amount of haze. For dehazing the paper proposes firstly a method for robust preprocessing based on adaptive atmospheric light, and then this is passed throughout the random forest regressor which is trained on the generated dataset. The results of the paper further go on to show that the synthetic dataset is very effective in training the model.

In deep learning models that are generally trained for hazing of images there is generally a limitation on the quantity of data, which affects the model's performance. B Li. et al. [5] aims at providing a benchmark for single image dehazing. The paper introduces Realistic single image dehazing dataset (RESIDE), which is a large-scale synthetic dataset which originally had 15,000 images and was later updated and the RESIDE: V0 has around 500,000 images focused at both objective and subject evaluations. B Li. et al. [5] further observers that the conventional metrics like PSNR, SSIM fail to satisfy the needs of a dehazing model when testing, so the paper compares some of the dehazing algorithms based on PSNR, SSIM and some of the popular No reference metrics such as the blind image integrity notator (BLIINDS) which is based on discrete cosine transform. The paper finally suggests the use of perceptual loss for improving the training process.

The scientific community continues to be lacking a reference dataset to judge objectively and quantitatively the performance of projected dehazing ways. The few datasets that are presently thought-about, each for assessment and training of learning-based dehazing techniques, completely deem artificial hazy pictures, to deal with this limitation, the paper Ancuti et al. [6] introduces the primary out of doors scenes information (named O-HAZE) which has both real hazy and corresponding haze-free pictures. The hazy pictures are captured in presence of real haze, generated by haze machines, and it contains forty-five different outdoor images portraying a similar visual content recorded in haze-free and hazy conditions, below similar illumination parameters.

2.2.2 *Deep Learning Networks*

We have looked at some of the most popular as well as distinct models.

Pavan et al. [7] has proposed a light convolution network which emphasis on factors such as time and computational efficiency for real-time purposes and faster network. The paper boasts an encoder-decoder network with the modern CNN with no atmospheric scattering assumptions. With the trade-off between network complexity-image quality in the LCA neural network, it can be used in any light processing device. With ideal elimination of haze with the help of a light convolutional neural network for quicker results on the current standard datasets, this gives a proportional result with respect to image quality compared to current revolutionary and modern dehazing models with significantly less computational resources. The paper suggests the mean squared error (MSE) as its loss function which is optimal and helps with the variance between the high contrast pixels and also helps with faster and smoother convergence to the local minima with the gradient descent. Although there are contrast problems in the paper with no means to reconstruct the color fidelity and intensity. The downside of this network is its inaccurate representation and reconstruction on low-light and thick hazed images and the shortcomings and flaws are overcome with the help of other models which will be discussed briefly in the next section.

B. Cai et al. [8] proposes a dehazenet which offers good trade between computational complexity and dehazing output, and hence making this model a viable and optimal starting point for dehazing application. The paper suggests a deep learning architecture based on Convolutional Neural Networks (CNN) with parallel networks with the help of concatenation of layers for feature extraction with lesser trainable parameter. The paper basically aims at generating a transmission map through 4 steps namely feature extraction, multi scale mapping with kernel of size 3,5,7 and then finding the local extremum and finally applying nonlinear regression. The paper is delineated to train the model to learn and assume the prior atmospheric assumptions. The paper proposes a nonlinear activation function called Bilateral Rectified Linear Unit (BReLU), which helps the model to incorporate the best features of Rectified linear Unit(ReLU) and maintain the transmission map in range [0,1]. The paper achieves prime performance over current models yet keeping it optimal and ease of access. The paper [8] limits itself to the atmospheric scattering model, as it can also be learned in a deeper neural network. The network due to its simplicity fails to perform as well as the other state of the art models but with around 800 citations it has been a base for many of the newer deep learning models.

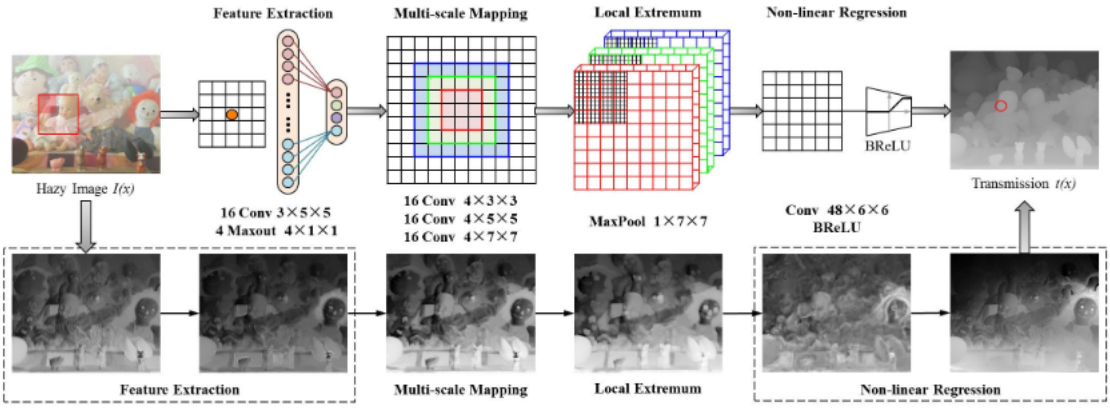

Figure 2.1: Dehazenet Architecture

B. Cai et al. [8]

Liu et al. [9] proposes a Generic Model-Agnostic Convolutional Neural Network (GMAN) for Single Image Dehazing. The paper does not estimate any atmospheric scattering model and only depends on trainable parameters and hence it learns complex hazy features and structures. The paper has numerous advantages over the Gated Fusion Network in terms of optimization, input size versatility, efficiency and complexity of the network. The paper suggests that the network is tailored for reconstructing the hazy image which is more generic and could be used for broader application. It also uses image by upscaling and downscaling by a factor of 2. As it is a new type of network there is not any concrete papers and research on this network.

Chen et al. [10] proposes a Gated Context Aggregation (GCA) Network for Image Dehazing and Deraining. Unlike the traditional computer vision papers which leverages traditional mathematical or image priors as the restoration constraints, e.g., DCP, the paper uses an gated context aggregation network to directly restore the final haze-free image. The paper proposes the modern smoothed dilation technique with dilated convolution in the network which takes up few to little trainable parameters and also uses a gated sub-network which blends features from multiple scales of the same image which helps the network remove the hazy elements from the image and also artifacts. The paper surpasses previous state-of-the-art dehazing methods by a large margin both quantitatively and qualitatively. The paper boasts the dilated convolutional layer, which braces the exponential expansion of the responsive field without loss of resolution and maintaining fidelity. The paper is computationally heavy, can be improved with custom loss function and can be optimized and made efficient for ease of access.

Y. Lee et al. [11] proposes the popular U-net based architecture for dehazing that employs squeeze and excitation blocks. The paper incorporates the dilation principle to increase the receptive field in the encoder part of the network. The paper uses squeeze blocks to provide channel attention and nearest neighbour upscaling in the decode network to improve output image quality. The model uses a sum of mean square error and perceptual loss as loss function for training. When implementing we found the model to exhibit colour distortion and dullness (like a light image).

Ayush et al. [12] proposes a pyramid convolution which dehaze's in multiple layers of scales. The Uses a U-Net block which extracts new and complex features from the input image. The paper proposes a deep learning method which makes it as a generalized model resulting in it to be more versatile compared to other models. The whole of the algorithm is an extensive experimentation. It uses a combination of many losses. This model uses a pyramid convolution technique. The downside is reconstruction of the colors.

Qin et al. [13] proposes an end-to-end feature fusion attention network. The surpasses all the previous methods by a very high margin. It supersedes the previous algorithm with its powerful advantage of superior result in thick haze regions. Includes a FA module which mixes both channel attention and pixel attention mechanism which helps it focus more on the thick haze part. Local residual learning and feature attention (FA) each have a separate block which helps the model to process the image in the thin haze region through many skip connections. Attention-based feature fusion (FFA) structure, it can process thin hazed regions by running it through multiple deep layers which helps it outperform every other feature fusion method. But one of the key drawbacks of the model is that it is quite computationally heavy.

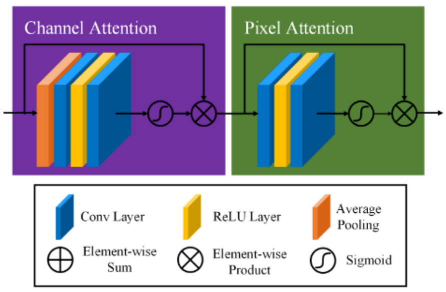

Figure 2.2: Feature Attention Blocks

Qin et al. [13]

C. Wei et al. [14] proposes a deep retinex decomposition model for low light enhancement. The model decomposes a image into reflectance and illuminance based on the retinex theory, and reconstructs the image and later uses image enhancement operations. C. Wei et al. [14] uses a data driven approach, and built a large dataset with paired low and normal light images named Low light paired dataset (LOL). The model takes input as both low light image and normal image during the training stage and only low light image during the testing stage. The paper uses a combination of 3 losses while training namely reconstruction loss, invariable reflectance loss and illumination smoothness loss. The results of the paper are quite good therefore we used this model skeleton when developing our model as described in methodology section (6).

2.2.3 *Losses*

Most of the networks we looked at in the above section, used either Mean square error (L2) or Mean absolute error (L1) as their loss function for computation of gradients for back propagation.

H. Zhao et al. [15] looks into loss functions which are specific to the tasks of image restoration for applications such as denoising, deblurring, super resolution and de-mosaicking. The paper looks into 4 different loss function candidates. Firstly, at Mean square error (MSE) which is one of the most common loss functions. MSE correlates poorly to image quality and human perception. It then looks at Mean absolute error (MAE) whose performance is better than MSE but still sub optimal. Then it looks at multi scale structural similarity (MS SSIM) , as SSIM with sigma value for the gaussian filter if used fails to preserve the local structure. The finally proposes a weighted combination of both MAE and MS SSIM.

2.2.4 *Metrics*

Metrics play a huge part in deciding how well a model performs, most of the models described earlier look at PSNR and SSIM as a metric to decide the quality of results. We in our work have tried to go one step further and have studied No reference image quality evaluators.

Mittal et al. [16] proposes a no reference opinion unaware, distortion unaware image quality assessment model. The model proposed by the paper derives quality aware around 36 quality aware features from natural scene statistics on a local patch. The optimal size of the local patch can vary from 32x32 to 160x160 as observed by the paper. The paper uses a corpus of 125 images selected from Berkley image segmentation dataset. A simple model of NSS is obtained by fitting the features to a multi-variate Gaussian (MVG). The difference between MVG of test image and that of the natural images gives the value of the Naturalness image quality index.

K. Choi et al. [17] propose a Fog aware density evaluator which is dependent on natural scene statistic like NQIE, but it only considers fog specific features. The paper considers a total of 12 feature like sharpness, colorfulness, dark channel prior etc. and fits a Multi variate Gaussian to the features which are derived on a local patch. The Mahalanobis-like distance between the haze free images and hazy images gives the density of haze. The paper further shows that this metric is linearly related to mean score of humans thereby making it a very import and suitable metric in the field of dehazing.

Y Liu et al. [18] proposes a novel no reference image quality assessment algorithm for contrast distorted images like hazy images. Contrast distortion is a general type of distortion found in many images and these hinders the quality of the images, so to quantify the quality of the image with the amount of contrast distortion present in an image. The paper proposes spatial characteristics, image histogram, visual perception characteristics and chrominance to quantify the quality. All the above four mentioned properties are extracted and then passed through a Support Vector Regression (SVR) to integrate a regression line to extract information. The downside of this method is its time complexity as it runs twice as slow as PSNR and SSIM.

CHAPTER 3

PROBLEM DEFINITION

As stated earlier, we have taken up the project to combat the visibility issues faced by people and machines. When we look at the available approaches and there seems to a clear lack of ranking system, this has prompted us to take up the work. We further looked at developing a model which although being domain specific could also be extended to many computer vision algorithms like image deraining, image denoising, processing before object detection and segmentation.

CHAPTER 4

OBJECTIVES

- To compare various SOTA dehazing algorithms and compare them using common performance metrics.
- To develop a novel model by taking inspiration from deep retinex model.
- To Extend the application of the model from single image dehazing to sequence of image dehazing (Video dehazing).
- Analyse different areas of application for the algorithm and suggest the most suitable area of use.
- To present the new architecture, and a suitable metrics a research article.

CHAPTER 5

THEORETICAL BACKGROUND

5.1 *Atmospheric scattering model*

The haze present in a image is due to two features:

1. Direct Attenuation: results from decay in the medium
2. Air light: occurs due to the scattering of light in the surrounds

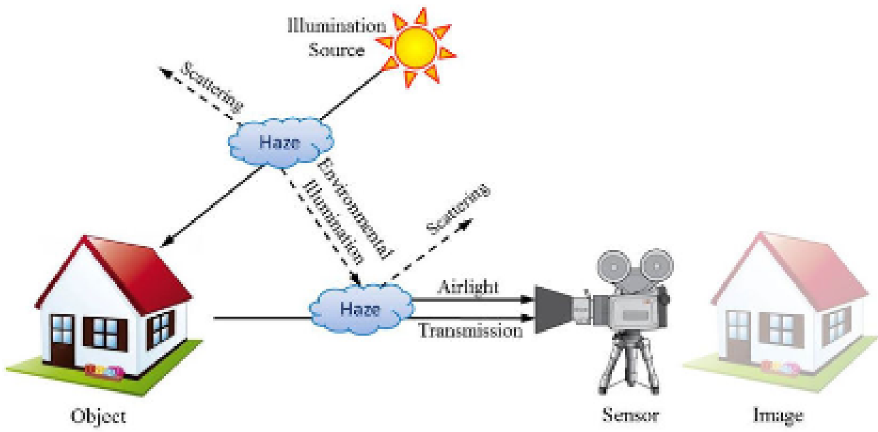

Figure 5.1: Atmosphere Scattering Model

B. Cai [8]

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{5.1}$$

Where I the hazy image captured, J is the dehazed image (scene Radiance), A is the global atmospheric light, and t is the transmission of the medium i.e. the portion of the light that is unscattered.

$$t(x) = e^{-}\beta.d(x) \tag{5.2}$$

Where β is the scattering coefficient of the atmosphere and d(x) is distance from the scene point It indicates that the scene radiance is attenuated exponentially with the scene depth. The model assumes homogeneous scattering of light and a constant scattering coefficient.

5.2 *Convolutional neural networks*

A Convolutional Neural Network (ConvNet/CNN) is a type Deep Learning method for images to extract features and to classify or to reconstruct them. Traditionally, the initially layers try to find simple features and patterns. As the layers go on, it learns more complex features that comprehends the images as we do. The Pooling layer is similar to conv layer but it helps to reduce the dimensionality of an image to reduce the computation. By dimensionality reduction, the computing power needed to process the data is reduced. Pooling is classified into two types: average pooling and max pooling. Max Pooling gives the max value from the Kernel-covered part of the image. Average Pooling, on the other hand, gives the average of all the values from the part of the image covered by the Kernel.

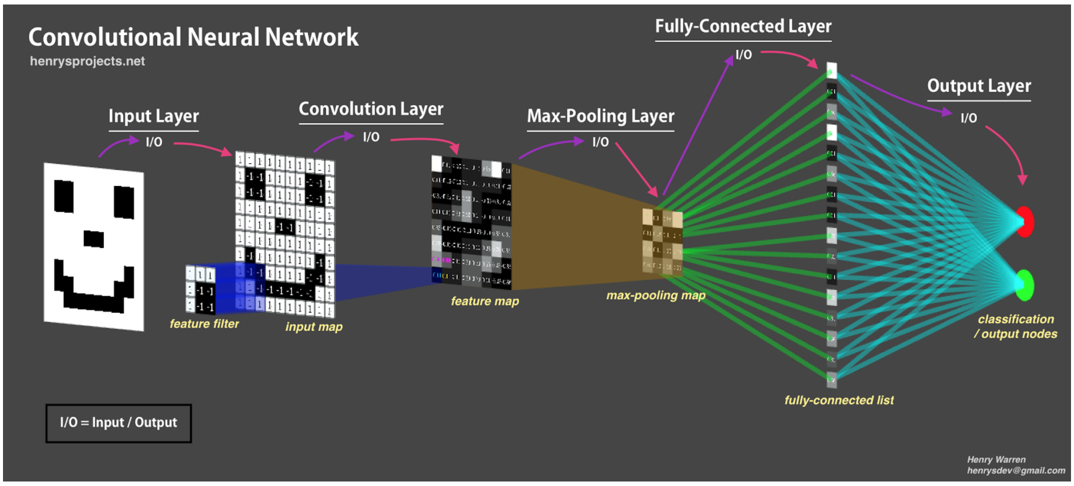

Figure 5.2: CNN

Henry Projects

Because of dimensionality reduction this makes process for learning an image effective. CNN outperforms other conventional ML algorithms in terms of performance, such as SVM, and XGBoost classifier. The CNN trend was set by ALEXNET with only 8 layers.

Hyper-parameters related to CNN

- Number of Hidden Layers and units: Many layers with regularization will result in better performance. Less would result in under fitting.
- Network Weight Initialization: It is ideal to use different distribution for weight initialization for different types of activation function in a layer.
- Dropout: It is a type of regularization which will drop specified percent of neuron in a particular layer at random so as to increase the dependency of each neuron to learn new features.
- Activation function: These are used to bring out the non-linearity in the model with the help of specific continuous and discrete mathematical function. E.g., ReLu, Sigmoid, tanh etc.
- Learning Rate: Bigger learning rate takes huge steps during back propagation for converging to minima and smaller takes smaller step. Ideally, we would like to use bigger initially and slowly reducing the learning rate so as to converge at the minima with less hinderance. This is called Decaying Learning Rate.
- Momentum: This parameter helps the model to take the step in the forward direction, so that the model doesn't oscillate between a specific point.
- Batch size: Number of subsamples to train before the model back propagates.
- Number of epochs: Number of times the entire training process is repeated.

CHAPTER 6

METHODOLOGY

6.1 *Detailed Methodology*

We started our working trying to implement all the aforementioned papers and compare them using Full reference as well as no reference metrics. We also considered factors like training time, number of blocks and the ability of the model to be extended. After working through all the popular model, we identified Dehazenet as a good base model because of the compactness and seamless reconstruction of the output. We also try to embed some of the good feature from GCA-net and FFA net into our improved Dehazenet. We then looked at different field whose results could be extended to the area of dehazing and found the area of log light enhancement particularly interesting so we picked out some of the low light models and tried to extend it to dehazing with the help of a suitable dataset while training. On the basis of the conceptual uniqueness and ability to be extendible we picked out Deep-retinex and are working on developing a network with similar architecture for dehazing problems.

6.2 *Assumptions*

We though our work wanted to focus on visibility outdoors as we were motivated to take up the problem due to the visibility loss through rise in air pollution. So, we trained all the models on part of the RESIDE dataset which had only outdoor images. We also wanted to break the limitation of current models so we have been looking at implementing a heterogenous atmospheric light assumption which goes one step ahead then the homogeneous atmospheric light assumption. We also looked at the atmospheric scattering model and its inability to capture nonlinear mapping between the terms, so we decided to use a nonlinear mapping between the terms.

Table 6.1: Comparison Study of dehazing models

Metrics	DCP	LCA	DehazeNet	GMAN	U-net	GCA	FFA
PSNR	12.21	17.07	17.23	14.8	19.28	20.13	20.67
SSIM	0.61	0.65	0.66	0.64	0.73	0.77	0.79
FADE	0.38	0.95	0.50	0.65	0.68	0.91	1.24
NQIE	2.85	4.42	2.93	2.70	3.71	2.7	2.67
CEIQ	3.19	3.27	3.33	3.22	3.4	3.22	3.42

6.3 *Experimental Work*

As a part of the work carried out, we choose different dehazing networks and coded them from scratch with the help of the research papers and supporting material online. We then picked out a dataset of 13,000 images and picked out a set of 42 images each with distinct characteristics to form the test set. We then wrote an image quality evaluator script on Matlab which would output the average of all the metrics. Then we trained all the models on a subset of the dataset which contained around 2000 images for around 30 epochs. The results validated the fact that FFA is truly a standout net, but we were particularly interested in dehazenet which as discussed earlier is light and gives a good result. So next we took the weights trained weights of dehazenet implementation in pytorch and reshaped them according to TensorFlow and there was a slight increase in performance. We also implemented features such as dilation, multi scale feature extraction and attention feature on dehazenet and analysed the results. As a part of our development phase, we looked at various network and the Deep-retinex network [14] particularly stood out as it was one of the best networks in the field of low light enhancement. We further took inspiration from Deep-Retinex [14], which has 3 levels: decomposition, adjustment and reconstruction. We extended the dehazenet base model for decomposition to extract the transmission map from it and also attain atmospheric light scattering coefficients from the same pipeline with modules to calculate the coefficients.

6.4 *Tools*

The tools used for the current project are: NumPy, Matplotlib, TensorFlow

NumPy is numerical computational library which is fast and compact with the help of its vectorization and broadcasting. Matplotlib is visualizing tool to display the image outputs and inputs. TensorFlow is an open-source library which is used for machine learning and deep learning. TensorFlow was developed by the Google rain team for internal Google use. We chose TensorFlow over PyTorch is because of the flexibility with Google Colab. In TensorFlow framework a model can be defined in 3 ways: sequential, functional API and model subclassing. We mostly defined our models using sub-classing as it gives us lower-level control. Most models implemented in PyTorch can be converted into TensorFlow model subclass. Google colab lets us run programs on a jupyter notebook kind interface online. It also lets up access GPU and TPU to speed up training.

CHAPTER 7

RESULTS AND CONCLUSIONS

We have compared various models for dehazing which were developed from scratch in TensorFlow with the help of research paper and supporting works.

The models were trained on a common dataset as mentioned before. It's clearly evident that FFA net outperforms all other networks. FFA network though quite efficient is quite computationally heavy to train, so we have considered Dehazenet as our base model for further work. Dehazenet model performs quite well from a human vision stand point but has some glaring drawback such as

- The transmission map is only considered for enhancements.
- Atmospheric light is assumed homogeneous and is estimated by conventional methods.
- Restricts itself to the atmospheric scattering model.

We are currently working on substituting the Decom-net in Deep-retinex with Dehazenet based on current findings.

CHAPTER 8

WORK PROPOSED

After completing our analysis of the many models used to solve the dehazing problem current we will be trying to incorporate the learning from those models to our novel model (Development Phase). Drawing inspiration from the Deep-retinex model we plan to implement a similar architecture substituting blocks such as Decom and image enhancement with models specific to dehazing. We thought of taking the lowlight enhancement model which decomposes the image into illuminance and reflectance and converting it to a model which outputs dehazing specific image decomposition like transmission map and atmospheric light. We have further plan to incorporate better image enhancement networks so as to improve the performance of the model.

REFERENCES

[1] K. He, J. Sun and X. Tang, ”Single Image Haze Removal Using Dark Channel Prior,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2341-2353, Dec. 2011, doi: 10.1109/TPAMI.2010.168.

[2] K. He, J. Sun, and X. Tang. Guided image filtering. In ECCV, pages 1–14. 2010.

[3] 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.

[4] K. Tang, J. Yang and J. Wang, ”Investigating Haze-Relevant Features in a Learning Framework for Image Dehazing,” 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 2995-3002, doi: 10.1109/CVPR.2014.383.

[5] B. Li et al., ”Benchmarking Single-Image Dehazing and Beyond,” in IEEE Transactions on Image Processing, vol. 28, no. 1, pp. 492-505, Jan. 2019, doi: 10.1109/TIP.2018.2867951.

[6] C. O. Ancuti, C. Ancuti, R. Timofte and C. De Vleeschouwer, ”O-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Outdoor Images,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 2018, pp. 867-8678, doi: 10.1109/CVPRW.2018.00119.

[7] A. Pavan Bennur, Adithya Gaggar, Mohit S S, Shylaja. (2020). LCA-Net: Light Convolutional Autoencoder for Image Dehazing.

[8] B. Cai, X. Xu, K. Jia, C. Qing and D. Tao, ”DehazeNet: An End-to-End System for Single Image Haze Removal,” in IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187-5198, Nov. 2016, doi: 10.1109/TIP.2016.2598681.

[9] 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.

[10] Chen, D, He, M, Fan, Q, Liao, J, Zhang, L, Hou, D, Yuan, L Hua, G 2019, Gated context aggregation network for image dehazing and deraining. in Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019., 8658661, Proceedings - IEEE Winter Conference on Applications of Computer Vision, WACV , IEEE, pp. 1375-1383, 19th IEEE Winter Conference on Applications of Computer Vision (WACV 2019), Waikoloa Village, United States, 7/01/19.

[11] Y. -W. Lee, L. -K. Wong and J. See, ”Image Dehazing With Contextualized Attentive U-NET,” 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 2020, pp. 1068-1072, doi: 10.1109/ICIP40778.2020.9190725.

[12] Ayush Singh, Ajay Bhave, Dilip K. Prasad, "Single image dehazing for a variety of haze scenarios using back projected pyramid network," arXiv:2008.06713 [eess.IV]

[13] Qin, Xu & Wang, Zhilin & Bai, Yuanchao & Xie, Xiaodong & Jia, Huizhu. (2019). FFA-Net: Feature Fusion Attention Network for Single Image Dehazing.

[14] Wang, Wenjing & Yang, Wenhan & Liu, Jiaying. (2018). Deep Retinex Decomposition for Low-Light Enhancement.

[15] H. Zhao, O. Gallo, I. Frosio and J. Kautz, "Loss Functions for Image Restoration With Neural Networks," in IEEE Transactions on Computational Imaging, vol. 3, no. 1, pp. 47-57, March 2017, doi: 10.1109/TCI.2016.2644865.

[16] Mittal, Anish Moorthy, Anush & Bovik, Alan & Chen, Chang-Wen & Chatzimisios, Periklis & Dagiuklas, Tasos & Atzori, Luigi. (2015). No-Reference Approaches to Image and Video Quality Assessment. 10.1002/9781118736135.ch5.

[17] L. K. Choi, J. You and A. C. Bovik, "Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging," in IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3888-3901, Nov. 2015, doi: 10.1109/TIP.2015.2456502.

[18] Y. Liu and X. Li, "No-Reference Quality Assessment for Contrast-Distorted Images," in IEEE Access, vol. 8, pp. 84105-84115, 2020, doi: 10.1109/ACCESS.2020.2991842.