A

Project report on

Image restoration by using Convolutional Neural Networks

by

Vikas Bhadoria (vikasbadhoria69@gmail.com)

Department of Informatics engineering

Wismar University of Applied Sciences: Technology, Business and Design

Wismar, Germany

ABSTRACT

This project focuses on image enhancement using Deep Learning technique. It has been observed that during the image data collection process or while capturing an image, the image gets blurred or has some noise in it. Specially, in the case of optical communication where a slight change in the signal can disrupt the entire message. Thus image enhancement assumes a very significant role in recovering the original image. In order to meet the intended objectives I first surveyed the related literature to find the current trends in image enhancement.

In this work I first successfully used 'Transfer Learning' technique where I used the already pre trained vgg19 convolutional neural model to recover the image content. The results for this were encouraging and an understanding was developed as to how powerful convolutional neural networks can be when it comes to image restoration. I then created my own deep convolutional neural model by taking inspiration from different award winning models thus trying to mimic their abilities. I carried out training of my model on a huge dataset and the final predictions made by the model were very encouraging if not 100 percent perfect.

This project report consists of 5 chapters. Chapter 1 introduces the image restoration and its relevance. Chapter 2 presents the problem statement and the objectives of this project. Chapter 3 is devoted to the related literature survey. Chapter 4 of this report discusses theory and methodology that I have followed; to test the outcome of the application of deconvolution technique in image enhancement. Chapter 5 highlights phase by outcome of the project and concludes the report.

TABLE OF CONTENTS

Chapter 1: Introduction

Image restoration is one of the challenging issues in image processing. The images acquired from the digital media may not be of desired good quality. The poor image quality can result due to a number of reasons such as uneven illumination, due to noise introduced during the transmission of image data or due to digital to analog transformation. As a consequent, image quality usually needs to be improved. Image restoration plays a very significant role in improving the image quality of such images. Image restoration emphasizes the image content to make them more suitable for human vision, analysis or for use by another machine.

The domain of image processing provides unique functionalities, since it only considers an image as a valuable entity. The project works on the recovery or reconstruction of the distorted image using the concept of Neural Networks. The project is developed in Python using the deep neural framework of Keras & Tensorflow and is divided into four phases.

- Phase 1 is about defining noise parameters and reducing the brightness of dataset images.
- Phase 2 Using transfer learning to deploy Vgg19 model to observe the power of CNN for image restoration.
- Phase 3 Building my own CNN model and training it to remove the noise of any new input image.
- Phase 4 describes how good the model's predictions are.

Chapter 2: Problem Statement and Objectives

To build and train a model that can be used for restorating the content of a distorted image.

The objectives of this project are:

1. To use different techniques for adding noise to the image and using OpenCV library to reduce the brightness so as to create a perfect dataset of distorted images.

2. Using the vgg19 CNN model as a part of transfer learning technique to see how effective CNN is to extract the original content from a distorted image.

3. Building my own CNN model and training it on a huge dataset to efficiently restore distorted images.

Chapter 3: Literature Survey

The basic definition of image processing refers to processing of digital image, i.e removing the noise and any quiet irregularities present in a picture using the computer . The noise or irregularity may creep into the image either during its formation or during transformation etc. For mathematical analysis, a picture could also be defined as a two dimensional function f(x,y) where x and y are spatial (plane) coordinates, and the amplitude of “f()” at any pair of coordinates (x, y) is named the intensity or gray level of the image at that time . When x, y, and therefore the intensity values of f are all finite, discrete quantities, we call the image a digital image. it's vital that a digital image consists of a finite number of elements, each of which features a particular location and value. These elements are called picture elements, image elements, pels or pixels. Pixel is the most widely used term to denote the elements of a digital image.

Various techniques have been developed in Image Processing during the last five decades. Most of the techniques are developed for enhancing images obtained from unmanned spacecrafts, space probes and military reconnaissance flights. Image Processing systems are becoming popular due to easy availability of powerful personal computers, large size memory devices, graphics software etc [1]. Image Processing is used in various applications such as:

Remote sensing, Medical imaging, Forensic study, Textile, Material science, Military, Film industry

Image enhancement

Image enhancement is the way toward altering computerized images with the goal that the outcomes are more reasonable for show or further image investigation. Image enhancement should be possible by evacuating noise, image sharpening, or lighting up an image, making it less demanding to distinguish the key highlights[2]. Image enhancement algorithms include deblurring, filtering, and contrast methods. Deep learning utilizes neural systems to find out helpful portrayals of highlights straightforwardly from image information. Neural systems are pre trained to distinguish and remove different sorts of disruption from images and improve the images.

Methods in image enhancement

There are certain methods for Image Enhancement some of them are listed below [3]: **Histogram matching**: Histogram matching is the change of an image with the goal that its histogram coordinates a predefined histogram. The surely understood histogram equalization strategy is an exceptional case in which the predetermined histogram is consistently appropriated.

Contrast-limited adaptive histogram equalization (CLAHE): It is used to enhance the contrast of the grayscale image assumed as I by transforming. CLAHE works on small regions in the image called tiles, rather than the whole images[3]. Contrast of every tile is enhanced, therefore, the histogram of the output region approximately matches the histogram predefined the „Distribution‟ parameter.

Wiener filter: Wiener filter is a filter used to create a gauge of a coveted or target arbitrary process by linear time-invariant (LTI) filtering of an observed noisy process, accepting known stationary signal and noise spectra, and added substance noise. The Wiener filter limits the mean square error between the evaluated random process and the desired procedure.

Median filter: The median filter is a nonlinear computerized filtering method, regularly used to expel noise from an image. Such noise reduction is a common pre-processing step to enhance the results of later processing for example, edge recognition on an image. Median filtering is broadly utilized as a part of digital image processing in light of the fact that, under specific conditions, it preserves edges while removing noise.

Linear contrast adjustment: In this the contrast adjustment block changes the contrast of an image by linearly scaling the pixel values amongst lower and upper limits. Pixel values that are below or above this range are saturated to the lower or upper limit value, individually.

Deep neural network: Execute image processing undertakings, for example, removing noise from images and constructing high-resolution images from low-resolutions images, utilizing convolutional neural networks. Deep learning utilizes neural networks to learn valuable portrayals of highlights straightforwardly from information. For instance, you can utilize a pertained neural network to recognize the images and remove various types of noise from images.

Types of noise:

Gaussian noise: It is statistical noise having a probability density function (PDF) equal to that of the normal distribution, which is also known as the Gaussian distribution[4]. In other words, the values that the noise can take on are Gaussian-distributed.

Salt and pepper noise: is a form of noise sometimes seen on images. It is also known as impulse noise[4]. This noise can be caused by sharp and sudden disturbances in the image signal. It presents itself as sparsely occurring white and black pixels.

Poisson noise: The appearance of this noise is seen due to the statistical nature of electromagnetic waves such as x-rays, visible lights and gamma rays. The x-ray and gamma ray sources emitted a number of photons per unit time [4]. These rays are injected into the patient‟s body from its source, in medical x rays and gamma rays imaging systems. These sources are having random fluctuations of photons. The resulting gathered image has spatial and temporal randomness. This noise is also called quantum (photon) noise or shot noise.

Deep learning

Deep Learning, as a branch of Machine Learning, utilizes calculations to process information and impersonate the reasoning procedure, or to create deliberations [5]. Deep learning (DL) utilizes layers of calculations to process information, comprehend human discourse, and outwardly perceive objects. Data is gone through each layer, with the yield of the previous layer giving contribution to the following layer. The main layer in a system is known as the info layer, while the latter is called a yield layer. Every one of the layers between the two are mentioned as hidden layers as appeared in the Figure 1. Each layer is commonly a straightforward, uniform calculation containing one sort of actuation work. Learning can be directed, unsupervised or hybrid.

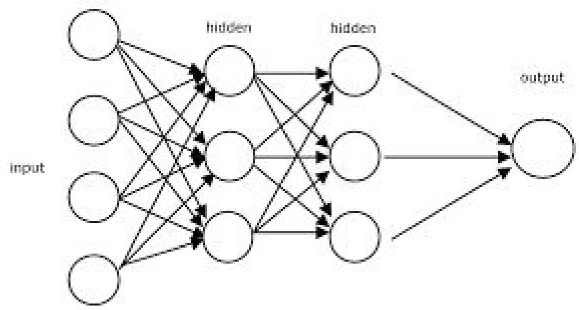

Fig 1: A feed-forward neural network

Convolution Neural Network(CNN)

CNN is a much better approach to work with images in Deep Learning. In CNN the goal is to extract and learn the specific features of an image. Images are arrays of pixels. These pixels are going to be processed by convolution filters also known as kernels (feature detector). For example a 3x3 Kernel can be used and applied on the image. This can be done by sliding the kernel at every location of the image. The amount by which the kernel is shifted is known as stride. The area where the operation takes place is known as the receptive field. Inside the receptive field every cell is multiplied by every cell in the kernel, the sum of all numbers is taken and the result is divided by the total number of cells. The final result is then shown on the feature map. By doing this the specific features of the image are captured. The values of weights in kernel are learned by the CNN during training through a gradient descent algorithm the values in kernel are changed to best learn the image.Its weight sharing ability reduces the number of parameters thus reducing the complexity of the model.

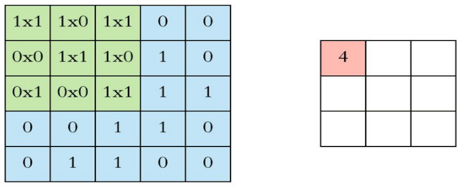

Fig 2: Convolution process.

Layers in CNN:

☐ Convolution layer: All the feature extraction takes place here. The feature maps from each filter are stacked together. Each feature map is different as each filter has different weights. After convolution the Relu activation function is applied to introduce non-linearity. The vanishing gradient problem is solved by using Relu.

☐ Pooling layer: Acts to shrink the image stack by reducing the dimensionality of each feature map thereby also reducing the complexity of the model. It's done to reduce the complexity and to reduce overfitting.

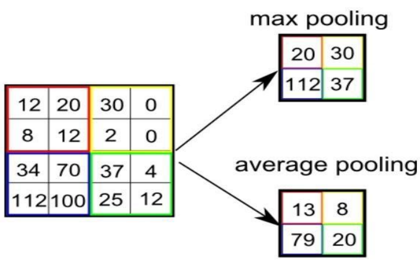

Fig 3: Pooling in CNN

Gradient Descent: Gradient Descent is an optimization algorithm used to minimize some function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient. In Deep learning, gradient descent is used to update the parameters of the model.

Chapter 4: Theory and Methodology

Theory:

Images received from the dataset are added with noise and reduced brightness in order to mimic the distorted images that a camera sensor may capture in bad conditions, low light situations or image distortion during network communication. These images should be recovered by building and training a CNN which can output a clear restored image.

Dataset:

The image dataset consists of 8,750 images and most of the images I have scrapped from the internet. The dataset also consists of pictures from my personal life. All these images are then resized to 500x500 before using in the model.

Methodology:

The project work was divided into the following 4 phases:

Phase 1: In order to test the theory there has to be a dataset which consists of clear images and the corresponding distorted images. The idea is to use supervised learning techniques to train the model and for that the real images will act as labels. In order to create artificial corresponding distorted images, noise has been added to the real images and then by using the OpenCV library brightness of images are reduced.

The noise added here is “Salt & Pepper” noise as different approaches were tried along with gaussian noise, poisson noise, speckle but the best results were with “Salt & Pepper”.

Phase 2: In phase 2 ‘transfer learning’ technique has been used as a part of research to observe the effects of CNN for image restoration. The award winning VGG19 neural model has been used. This model consists of 16 convolutional layers and 3 fully connected layers hence the name Vgg19. The pre-trained parameters of this model are kept constant and only the feature extraction part of the model has been used for this project’s phase 2.

Content Loss:

Content loss definition is actually quite simple. The network is passed with both the desired content image and the base input image that needs to be transformed. This will return the intermediate layer outputs from the model. Then simply take the euclidean distance between the two intermediate representations of those images. More formally, content loss is a function that describes the distance of content from our noise image 'x' and the clear image 'p'.

Backpropagation is then performed in the usual way such that it minimizes this content loss. Thus keep changing the noise image until it generates a similar response in a certain layer (defined in content_layer) as the original real image.

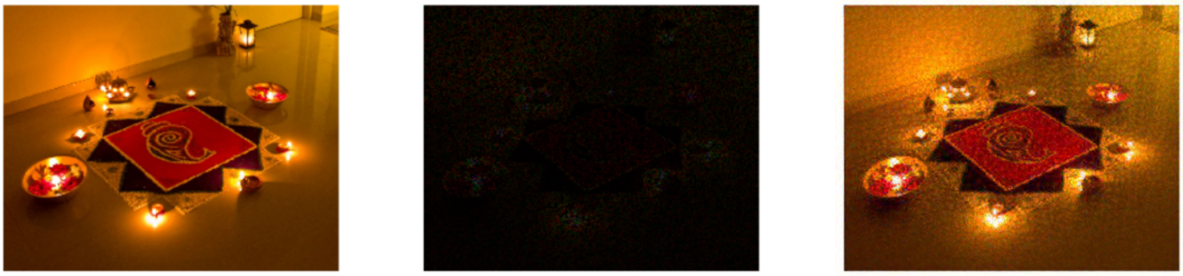

Fig 4: Output of vgg19. Original Image(left), Noise Image(centre), Restored Image(right)

The output of this phase shows how powerful CNN can be to restore images with noise and captured in low light intensity. The results motivates me to build a CNN and train it to output such restored images when a noisy image is fed as an input.

Phase 3: In phase 3 the idea is to build a CNN model which can be trained on my dataset consisting of real and noisy images. The basic concept behind this training will be the same as in phase 2. The model will be provided with real images as labels and noisy images which will be altered as the model performs gradient descent to minimize the error.

The model built in this phase for training is a 16 layer CNN which can be used for feature extraction of the images. This model is built using inspiration from ResNet model architecture, google's teachable machine CNN model's architecture and Vgg19's architecture.

The architecture and the idea behind all these models were carefully studied and then implemented to build this model.

Fig 5: Structure of the CNN model built.

The model is then trained on my dataset. For training the model the optimizer used is 'Adam', loss function used is 'Mean Squared Error', activation function used is 'Relu'. Since training with images is quite memory expensive, the model is trained in Google Colab which provides free GPU service to train such intensive models.

Phase 4: In this phase the model built and trained is put to test on noisy low light images which it has never seen before. The test images go through the same process of artificial noise addition and brightness reduction. This time only the noisy image will be an input to the model without any labels. The prediction of the model can be seen below.

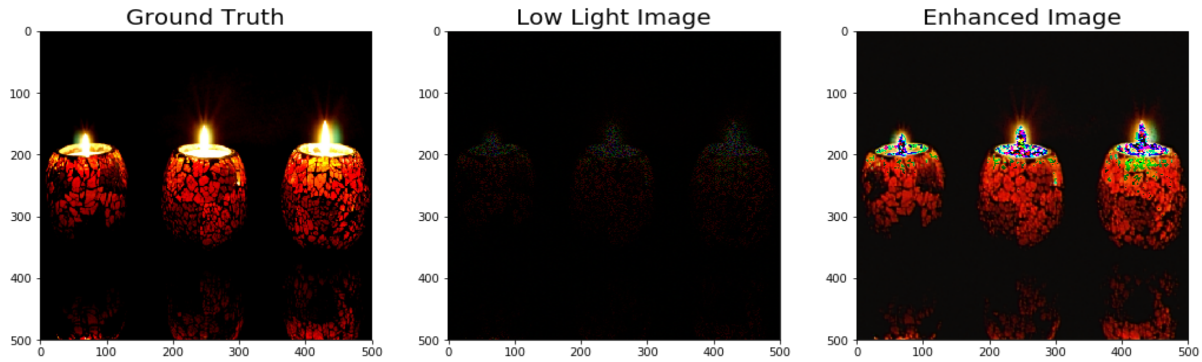

Fig 6 : Model's output prediction.

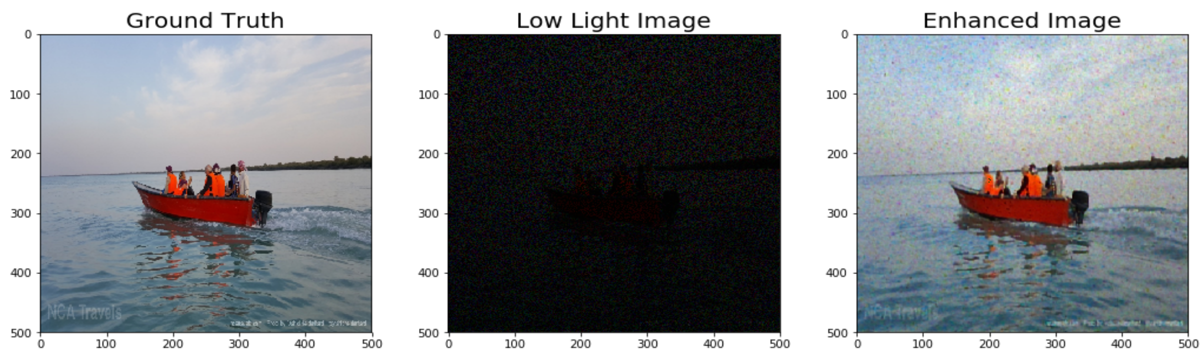

Fig 7 : Model's output prediction.

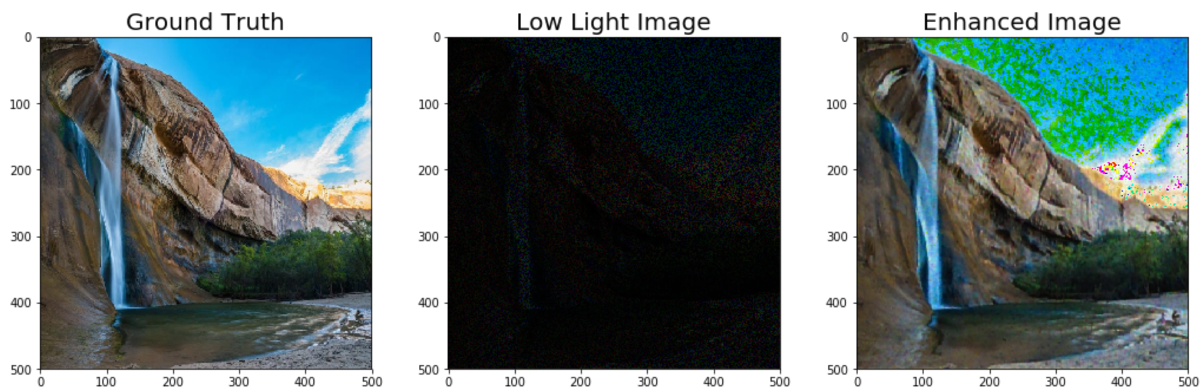

Fig 8 : Model's output prediction.

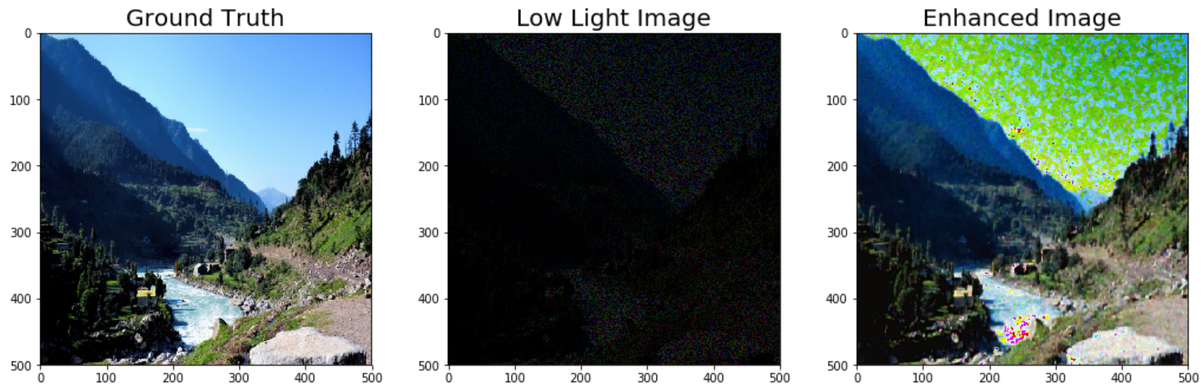

Fig 9: Model's output prediction.

Project status:

Sr.no	Project Phase	Status	Problems
1	Phase 1	complete	--------
2	Phase 2	complete	--------
3	Phase 3	complete	--------
4	Phase 4	partially pending	The model is producing inappropriate results in case of enhancing a very bright section of an image.

Chapter 5: Conclusions

In this project in order to meet the intended objectives I first created a dataset with real images, noise was then added artificially to the images. After trying out different types of noise I found "Salt & Pepper" more appropriate. Vgg19 model was then used to observe the effects of CNN for image restoration and enhancement and the results were very encouraging. In phase 3 I then built my own model by taking inspirations from architecture of other award winning models. I then trained the model on my dataset in google colab. The predictions of the model are very encouraging and can be improved further. The overall results can be concluded in 3 points:

1. The model works pretty good with low light images.

2. The noise used is 'Salt and Pepper' to train the model with a realistic low light image although it is not tested for other type of noises.

3. The model is producing inappropriate results in case of enhancing a very bright section of an image. The reason might be due to presence of salt pepper noise and might produce better results in case of real-time low light image.

References

[1] K.M.M. Rao "OVERVIEW OF IMAGE PROCESSING" Readings in Image Processing Fundamentals Of Digital Image Processing - Anil K.Jain, Prentice-Hall,1989.

[2] Cheng HD, Shi XJ. A simple and effective histogram equalization approach to image enhancement. Digital Signal Processing. 2004; 14(2):158-70.

[3] Lore KG, Akintayo A, Sarkar S. LLNet: a deep autoencoder approach to natural low-light image enhancement. Pattern Recognition. 2017; 61:650-62.

[4] Boyat AK, Joshi BK. A review paper: noise models in digital image processing. Signal & Image Processing: An International Journal. 2015; 6(2): 63-75.

[5] Chen XW, Lin X. Big data deep learning: challenges and perspectives. IEEE Access. 2014; 2:514-25.