IMAGE REFINEMENT AND RECONSTRUCTION USING MACHINE LEARNING

Project report submitted in partial fulfillment of the requirement

for the degree of Bachelor of Technology

in

Computer Science and Engineering/Information Technology

By
Shivam Kumar Mishra (191394)

Under the supervision of

Dr. Nafis Uddin Khan
to

Department of Computer Science & Engineering and Information Technology
Jaypee University of Information Technology Waknaghat, Solan-173234, Himachal Pradesh

Certificate

Candidate's Declaration

I hereby declare that the work presented in this report entitled **"Image Refinement And Reconstruction Using Machine Learning."** in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering/Information Technology** submitted in the department of Computer Science & Engineering and Information Technology**,** Jaypee University of Information Technology Waknaghat is an authentic record of my own work carried out over a period from July 2022 to May 2023 under the supervision of **Dr.Nafis Uddin Khan,** Assistant Professor (SG) in Department of Electronics And Communication Engineering.

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

Shivam Kumar Mishra (191394)

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

(Supervisor Signature)
Dr. Nafis Uddin Khan
Assistant Professor (SG)
Department of Electronics And Communication Engineering.
Dated:

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

PLAGIARISM VERIFICATION REPORT

Date: ……………………………

Type of Document (Tick): **PhD Thesis** | **M.Tech Dissertation/ Report** | **B.Tech Project Report** | **Paper**

Name: ___________________________ __**Department:** _________________ **Enrolment No** _________

Contact No. ______________________________**E-mail.** ______________________________________

Name of the Supervisor: ____

Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): ________________________

____

____

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

- Total No. of Pages =
- Total No. of Preliminary pages =
- Total No. of pages accommodate bibliography/references =

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at ……………….(%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

(Signature of Guide/Supervisor) **Signature of HOD**

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

Copy Received on	**Excluded**	**Similarity Index (%)**	**Generated Plagiarism Report Details** (Title, Abstract & Chapters)	
	• All Preliminary Pages • Bibliography/Images/Quotes • 14 Words String		Word Counts	
Report Generated on			Character Counts	
		Submission ID	Total Pages Scanned	
			File Size	

Checked by
Name & Signature **Librarian**

……

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com

ACKNOWLEDGEMENT

First, I would like to express my gratitude to Dr. Nafis Uddin Khan, my project's supervisor from the Jaypee University of Information Technology's Department of Electronics and Communication Engineering. Anytime I encountered a problem or had a query concerning my research, the door of our professor's office was always open. He continuously gave me the freedom to make this project our own, but he gave me direction when he felt we needed it. I also like to thank Dr. Himanshu Jindal, who served as the project's second reader and co-supervisor at Jaypee University of Information Technology Department of Computer Science Engineering. I am incredibly appreciative of their input and assistance with this project.

Finally, I must express our sincere gratitude to our parents for their excellent support and never-ending encouragement throughout the years of study as well as during the process of conducting the research for and finishing this project. Without them, this feat would not have been possible. I am grateful.

Author,
Shivam Kumar Mishra

Table Of Contents

List Of Figures

List Of Graphs

ABSTRACT

Today's imaging conditions could result in almost little vision for the human eye. They may frequently be caused by a loss of clarity, similar to influences included in the earth's atmosphere that affect the visuals owing to haze, fog, and other daytime impacts. The impacts on such photos may exist, thus it is important to clarify and improve the relevant information obtained under those circumstances in order to identify the objects as well as other useful information and that is why Image enhancement techniques are becoming increasingly popular and useful in such scenarios.One of the key elements of digital image processing is image refinement, which is crucial for making an image useful for a variety of purposes, including those related digital application. Poor-quality photographs can be improved with the help of image refinement and reconstruction using machine learning. This paper's main objective is to use machine learning and its aligned technologies concepts to attempt to enhance the quality of digital photographs. In this project I am applying different techniques of machine learning and its aligned technologies on different sets of images and checking out which one of them is giving the best results.

Chapter 1

1.1 Introduction

Real-world situations frequently experience poor clarity and low light impacts. High dynamic range photos can more successfully retrieve contextual information. Where nighttime images were taken, they may have low brightness and be plagued by noise as a result of the weak signal.

The project's objective is to increase contrast and color to make

vital information more aesthetically pleasing. These techniques can be used for pixel segmentation and detection analysis. Normally, this low-light picture enhancement is required for a number of purposes, including surveillance, astronomy, and medical imaging.

Sparse representation technique was one of the earliest answers to the problem. It is a simple method that introduces important picture data, thus it is vital to enhance the quality of the reconstruction and offer alternative function operators. Histogram equalization (HE) is a relatively simple technique that often introduces a number of artifacts and significantly reduces the amount of subtleties in the image. It flips the dark input by enhancing the low light image. The illuminant component between frames is then improved using the Dehaze algorithm that is then supplied.When reconstructing the images, artifacts were introduced as a result of component and parameter problems.

In this project, various Histogram equalization methods are used, such as Contrast Limited Adaptive Equalisation (CLAHE) and Adaptive Histogram Equalisation (AHE), followed by some denoising methods to improve the outcomes.

1.1.1 Convolutional Neural Network (CNN)

CNN is a common artificial neural network for image refinement and analysis work. CNN does not follow matrix multiplication procedure. It is based on the idea of convolution, which entails using a filter to extract features from an input image. CNNs learn progressively more sophisticated versions of the input image by layering convolutions, pooling, and nonlinear activation functions.

Convolutional networks consist of multiple layers that contain artificial neurons . The Artificial neurons act as a mathematical function and compute the weighted sum of several inputs and output an activation value

1.1.2 Inception Model

The Google Network also known as Inception Network . The Inception model processes images differently than other models, allowing it to attain high accuracy with fewer parameters. It achieves this by employing a collection of "Inception modules," or convolutional layers of various sizes that capture features at various scales and resolutions.

1.1.3 ResNet-50

ResNet stands for Residual Network. It is an extension of CNN. The number "50" in ResNet-50 refers to the network's total number of layers, which include 48 convolutional, 1 pooling, and 1 fully linked layers. The introduction of skip connections, which allow information to flow straight from one layer to another while avoiding some intermediate levels, is one of the fundamental breakthroughs of the ResNet architecture. The issue of vanishing gradients, which can arise when gradients get too small to be effective for updating the weights of deep layers, is lessened as a result of this.

ResNet-50 has also been applied to other applications such as image captioning, semantic segmentation, and object detection. Due to its high accuracy and adaptability, it has become a popular option for many computer vision applications.

Deep neural network design has advanced significantly with ResNet-50, which has facilitated innovations in a variety of applications involving computer vision. It continues to be a popular option for both researchers and practitioners because of its creative use of skip connections, which has sparked the creation of many additional effective models.

1.1.4 Yolo-v7

Finding and locating things of interest in an image or video is the objective of the computer vision task known as "object detection." The assignment is locating and classifying the objects in an image according to their boundaries and locations.There are two main categories for state-of-the-art techniques

(i) Prioritize inference speed

(ii) Prioritize detection accuracy

YOLO stands for "You Only Look Once" . This architecture advanced for real-time computer vision applications because it processed data much more quickly than other object detectors. Since then, other YOLO versions and variations have been put forth, each of which offers a notable improvement in effectiveness and performance.

A backbone network comes first in the YOLOv5 design, then a detection head but YOLOv7 architecture contains a backbone, a neck, and a head.The task of removing features from the input image frames falls to the backbone network. Convolutional, pooling, and activation layers are combined to create feature maps, which are subsequently sent to the detection head.In order to determine the bounding boxes and class probabilities for the objects in the image or video, the detection head employs the feature maps that the backbone network has generated. It has a number of convolutional layers, a layer that pools global averages after that, and then a number of fully connected layers that result in the output. Utilizing a new anchor-based method to predict object bounding boxes is one of the main upgrades of YOLOv7 over earlier iterations. This method helps to cut down on false positives and enables for more precise object localization.

YOLOv7 is an object detection model that produces its output from a single head. But to help with training, it also has extra heads in the middle layers. YOLOv7 employs a Label Smoothing strategy to enhance the training process, which assigns soft labels based on the ground truth and network prediction outcomes. Instead of directly referencing the ground truth to construct hard labels based on predetermined rules, soft labels are calculated utilizing optimisation approaches that take into account the quality and distribution of the prediction output in addition to the ground truth.

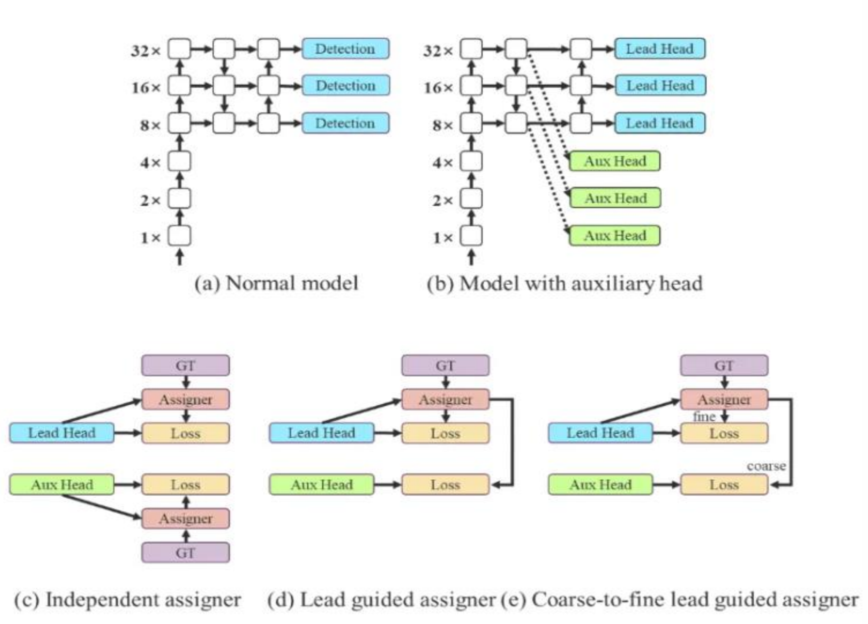


Fig 1

1.2 Problem Statement

This project's main goal is to talk about contrast and the different ways we can use existing technologies like CNN and histogram equalization to considerably increase it. Also covered are various histogram approaches and CNN models also how well each one works with different kinds of images.

1.3 Objectives

The primary goal of this project is to evaluate three distinct techniques utilized for image processing and enhancement. The project implementation involves leveraging Google Colab and Jupyter Notebook.

I. Histogram Equalization.

II. AHE stands for Adaptive Histogram Equalization .

III. CLAHE stands for Contrast Limited Adaptive Equalization

2.DIGITAL IMAGE PROCESSING

2.1 Image And Picture:

Even as already established, humans are typically visual beings that primarily rely on vision to sense their environment. Without only visualizing the things, they are able to quickly scan the space and get a sense of the whole atmosphere. Humans are able to quickly process a large amount of visual information, distinguish between colors, and recognise faces with ease. For our purposes, a single photograph that shows a person, an animal, a landscape, a microphotograph of an electrical component, or the result of a medical imaging operation is referred to as an image. The image won't be a random blur, even if it takes some time to become apparent.

2.2 What Is Image Processing?

An image's properties must be changed in order to improve the graphical details and make it more suitable for autonomous machine perception as well as human interpretation.Digital image processing entails changing the nature of a picture, which is what it is concerned with. Simple, uncluttered visuals are preferred by machines over sharp images, which people like. The edges of an image are improved to make them appear crisper. Figure 1 serves as an example.

As a result, the second image is sharper than the first. The decision to sharpen is made since printing is required for the edge to look great on the printed page.

(a) The original image (b) Result after "sharperning"

Fig 2

2.3 Digital Images:

Consider a black-and-white or colorless image for the time being.The image's brightness at any given time is therefore assumed to depend on the values of the two-dimensional function that these images represent. For the time being, let's assume that the brightness levels in these photos can be categorized as real integers between 0 and 1, with 0 signifying darkness and 1 signifying whiteness.

.

The image will surely have an impact on the scale of x and y, but they are actually scaling from low to high values.

Since these numbers are frequently integers, the brightness level values for the image will range from 0 to 255, or from black to white, and 0 and 1 will each contain values between 1 and 256. Digital images can be visualized as a huge group of unique dots, each with a different level of brightness..

They are known as pixels or the components of an image.
New colours can be created by changing the ratios of the RGB, or red, green, and blue, three primary colours. A colour image is represented using the three-dimensional matrix N M 3, where each layer represents the distribution of the primary color's grey levels.

Multiple points are included in an image, and these points are referred to as "pixels" and are identified by their (x, y) coordinates. The smallest piece of information in an image is represented by a pixel.

The image data contains information on the intensity level, which represents the irradiance that has been detected. The difference in grey levels between two adjacent pixels provides the contrast needed to distinguish between objects and regions. Only if a difference is a certain size can the human eye recognise it as a border.

2.4 Application

Image processing techniques have a wide range of applications, primarily in the sciences and technologies. Below, a summary of a few image processing applications will be covered.

I. Medicine Examining and describing images from various medical applications such as X-rays, MRI scans, and cell analysis such as chromosome karyotypes.
II. Aerial and satellite images of the land can be used in agriculture to determine how much land is used for certain

activities, whether an area is suitable for growing a particular crop, and whether or not fresh or expired produce is being used.

III. Inspection of products on a production line automatically.

IV. Deblurring or Increasing the pace Examples of fingerprint analysis include the use of camera photos for face identification in law enforcement.

2.5 Digital Image Types

2.5.1 Binary Image:

Each pixel requires one bit and can have one of two values. Pixels only come in black or white. As a result, such photographs may be effective in terms of storage. [3] Examples of images that might be represented in binary include text, fingerprints, and others. The figure below has two colors: black for the background and white for the edges.

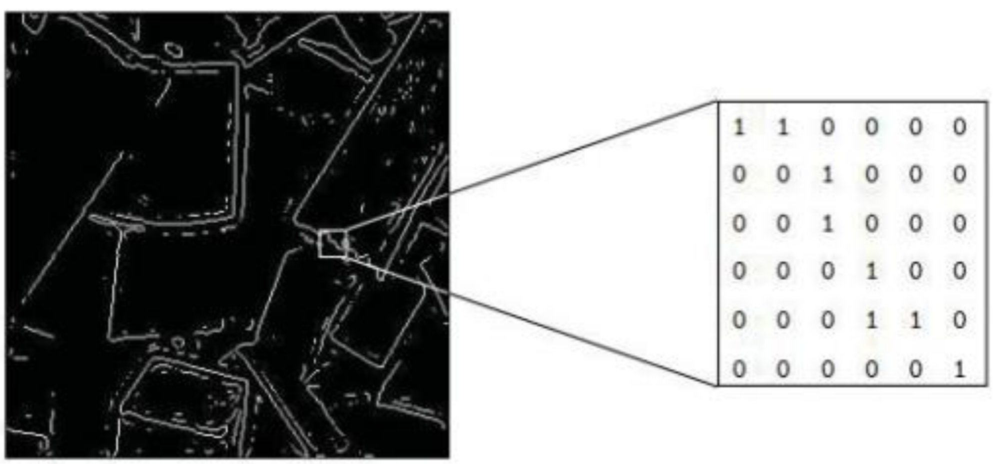


Fig 3

2.5.2 GreyScale Images:

Each pixel in the grayscale image has a shade from 0 (black) to 255. (white).One byte, or exactly eight bits, can be used to represent each pixel in this fashion. This is a typical range for managing image files. Greyscale's standard power is 2. X-rays, printed art, and many other sources may contain such images. Interestingly, most natural objects can be identified by using just 256 distinct shades of gray.

230	229	232	234	235	232	148
237	236	236	234	233	234	152
255	255	255	251	230	236	161
99	90	67	37	94	247	130
222	152	255	129	129	246	132
154	199	255	150	189	241	147
216	132	162	163	170	239	122

Fig 4

2.5.3 RGB(True color):

The number of red, green, and blue planes within each pixel determines the specific hue that it has. There are 16,777,216 distinct colors that might be created if the component ranges from 0 to 255. Images that require a total of 24 bits for each pixel are referred to as 24-bit color images. This image is made up of three matrices, one for each of the colors red, green, and blue in a pixel. This implies that there are three values that correspond to each pixel.

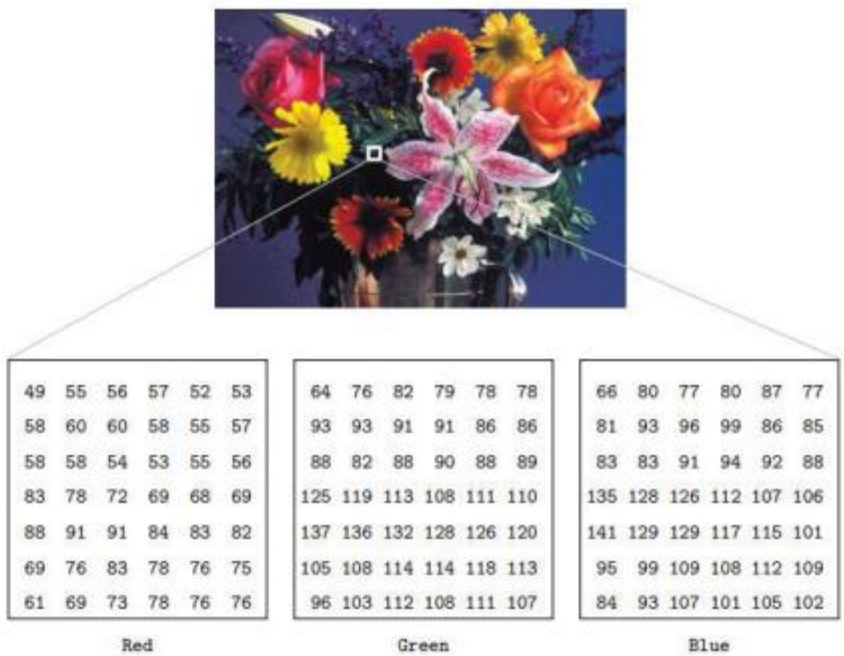


Fig 5

2.5.4 HSV Model :

Based on an RGB image, the HSV color model image Hue, Saturation, and Brightness are referred to as HSV.

I. The term "hue" refers to a characteristic of the visual sense that enables us to perceive a region as similar to one of the perceived hues, which may include red, yellow, green, blue, or a combination thereof."
II. Saturation: The amount of color in a region relative to its brightness.
III. Brightness is the quality of a visual perception that describes how much or how little light an area appears to emit.
IV. The HSV model primarily uses a cylindrical coordinate system. It is primarily founded on three elements.
V. The degree indicates the color of that pixel because H (hue) represents the circumference.
VI. The distance from the cylinder axis is S (saturation).
VII. V (value, often known as brightness) is typically expressed as a number between 0 and 1.
VIII. where a pixel with a value of 0 is black. From all of these, the HSV model can be seen in the figure as a cone.

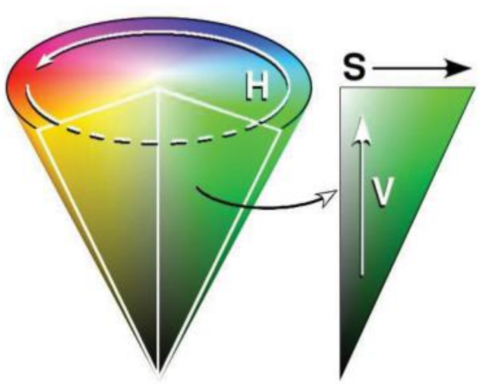


Fig 6

How exactly pixels work

On a digital display or in an image, a pixel is a basic unit that represents a particular colour scheme made up of RGB colours of varying intensities. Both single and multiple integers can be used to represent pixels. Depending on how the image is viewed, pixels can take many forms. For example, grayscale images are depicted in varying tones of black and white, with each pixel representing an integer between 0 and 255. Each pixel in a colour image is made up of three numbers, or RGB, which stand for the relative intensities of red, green, and blue. RGB is the standard interpretation for colour images. RGBA is an expansion of RGB that includes a separate alpha field to express the opacity of the image. Typically, the particular image determines the best interpretation, but

CNN

CNN also known as ConvNet stands for a typical type of neural network used for image recognition applications is the convolutional neural network. Images' dimensionality is decreased via the convolutional layer without any information being lost. CNNs are therefore very useful for these kinds of applications.
Three matrices make up an image that is written in RGB notation. The colour value of each matrix corresponds to a pixel in the image. The first matrix defines the red component, the second matrix defines the green component, and the third matrix defines the blue component. As a result, the RGB components of a 3x3 image are represented by three 3x3 matrices

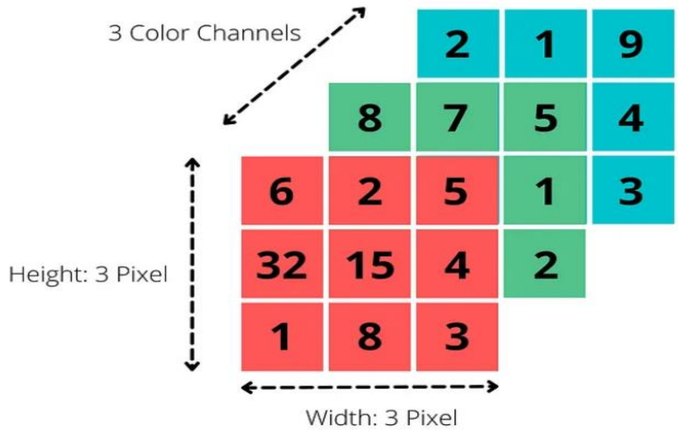


Fig 7

Every pixel in the image must be given into the neural network for it to process it. One example is a 200x200x3 picture, which has 120,000 input neurons because it has 200 pixels on 200 pixels and 3 colour channels. There are a total of 200x200 entries for each matrix due to the representation of each colour channel in a separate 200x200 matrix. As a result, we have three copies of this matrix—one for the red, blue, and green channels. However, as each neuron in the first hidden layer would need 120,000 weights from the input layer, adding more neurons to the hidden layer would quickly increase the number of parameters.

Convolutional neural networks' key operation takes place at the convolution layer, where a filter is employed to choose the size of partial images and the distance between computations. This process successfully decreases the image's dimensionality without distorting key details.

The pooling layer follows and, in terms of processing, it functions similarly to the convolutional layer. Nevertheless, depending on the particular application, we either take the maximum or average value from the output in this layer. By using this method, it is possible to save crucial pixel details needed for the task and remove extraneous ones..

The fully-connected layer, which is the last layer and is employed in conventional neural networks, is comparable to the previous layers. We can use tightly connected layers because the image's dimensions have been drastically decreased. Once more, connections between the multiple sub-images are used to find patterns and complete the classification.

Now that we know the general functions of each layer, we can examine the specific steps involved in classifying an image. In order to achieve this, we attempt to determine whether a dog is present in a 4x4x3 image.

Convolution Layer

In the first stage, a filter with a 2x2 dimension is used for each color channel to reduce the size of a 4x4x3 image. With a step length of 1, the filter travels one pixel at a time while doing calculations. By slightly lowering the image's dimensionality, this method keeps the image's fine details. Our convolutional layer produces a 3x3 matrix by changing the 4x4 matrix into a 2x2 matrix and relocating one column or row for each step. As shown in the image, the component values of the matrix are calculated using the scalar product of the two 2x2 matrices.

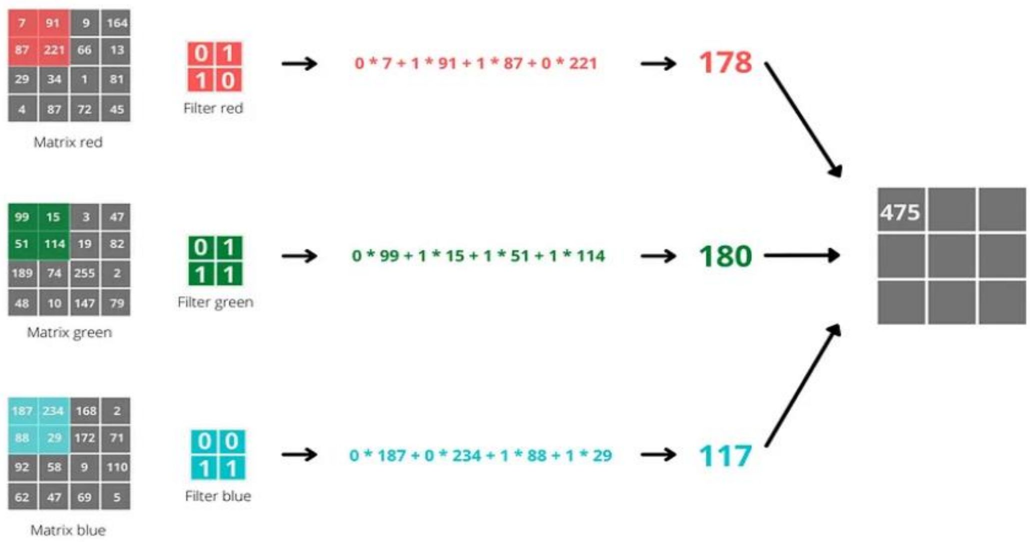


Fig 8

Pooling Layer

The 3x3 convolution layer input matrix is reduced in dimension while maintaining critical image properties by the (Max) Pooling Layer. By dividing the input into all feasible partial matrices of size 2x2, we may produce a 2x2 matrix as the layer's output. The greatest value in each of these fields is then chosen as the value for the field of the output matrix. Instead of using the max-pooling layer, we would compute the average of the four fields if we were to use the average pooling layer.

In addition to assisting in lowering the dimensionality of the input matrix, the pooling layer also removes any extraneous data or background noise from the image that does not aid in categorization. For instance, it might not initially matter if a dog is in front of a house or a forest.

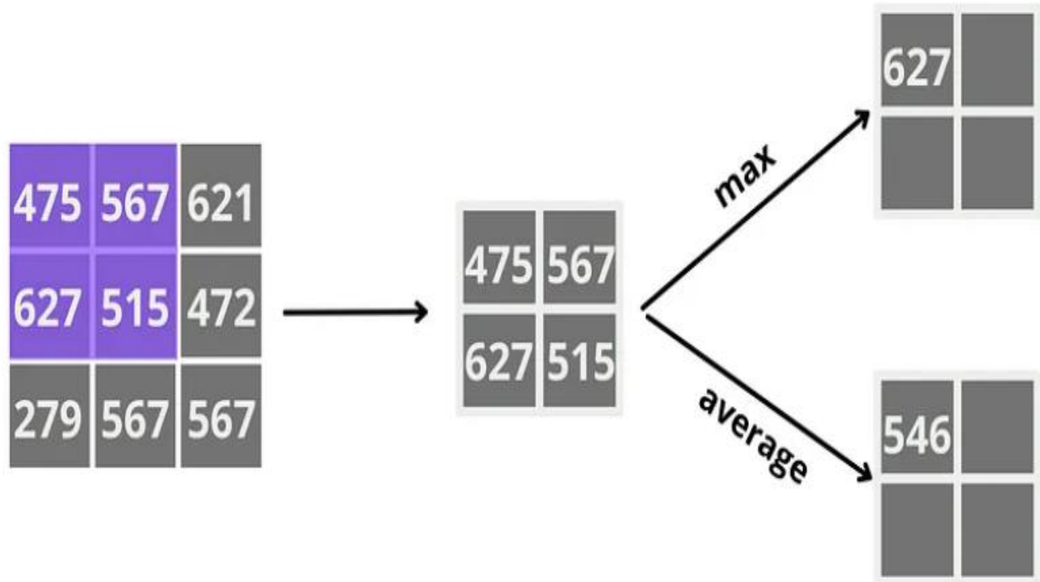


Fig 9

Fully-Connected Layer

This layer's goal is to accomplish the same thing as the entire image's original processing. The number of dimensions and required training resources are greatly decreased as a neuron is built for each element in the smaller 2x2 matrix and connected to every neuron in the following layer.

In conclusion, the last layer aids in locating the crucial aspects of the image required for classification. Before moving on to the fully-connected layer in the case of larger images, it is feasible to stack numerous convolutional and pooling layers. Because of the reduced dimensionality as a result, less training is necessary.

3. IMAGE INTENSIFICATION:

3.1 Introduction

The primary goal of picture enhancement is to improve the input. Image in order for the enhancement's results to be appropriate for the intended use.Based on factors like contrast, edges, boundaries, etc.The image enhances the desired result. The amount of the original information won't be increased by enhancement.

Instead, it will broaden the dynamic range of the selected features, not the range of the data.Quantifying the standard for image enhancement is the main challenge.Enhancement strategies necessitate participatory processes because they are about improvement.

to achieve the desired results.

There are two basic ways to process a picture, known as the spatial domain and frequency domain, depending on the processing domain. The term "spatial domain processing" describes an image processing method in which the pixels of the image are directly changed in the main image plane. Contrarily, frequency domain processing is based on modifying the Fourier transform-determined spatial frequency spectrum of the image.

3.1.1 Spatial domain enhancement methods:

The group of pixels that make up the image are referred to as the spatial domain technique, and it works directly on these pixels. The following formula is used to represent the spatial image processing function:

$$g(x, y) = T [f (x, y)].$$

In this case, the input image data are f(x, y), the output image data are g(x, y), and T is the operation performed on the set of input photos, such as doing pixel-by-pixel sums or averaging a number of images for noise reduction. In order to prevent the snowball effect of the altered grey levels, the resultant is saved independently in this case rather than modifying the pixel values.

3.1.2 Frequency domain enhancement methods:

The convolution theorem is the foundation of the frequency domain method.An image that has been processed is created by convolution of an image with location f(x, y).Operation with invariant h (x, y).

$$f(x, y) = hg(x, y)\ (x, y).$$

The following frequency domain relation is obtained by the convolution theorem:Here, 2D convolution is used to perform DFT in the frequency domain.

$$G(u, v) \text{ equals } F(u, v)\ H\ (u, v).$$

Where G, H, and F are the corresponding Fourier transformations of g, h, and f. H(u, v) (u, v) represents the process's transfer function.

3.2 Histogram Processing:

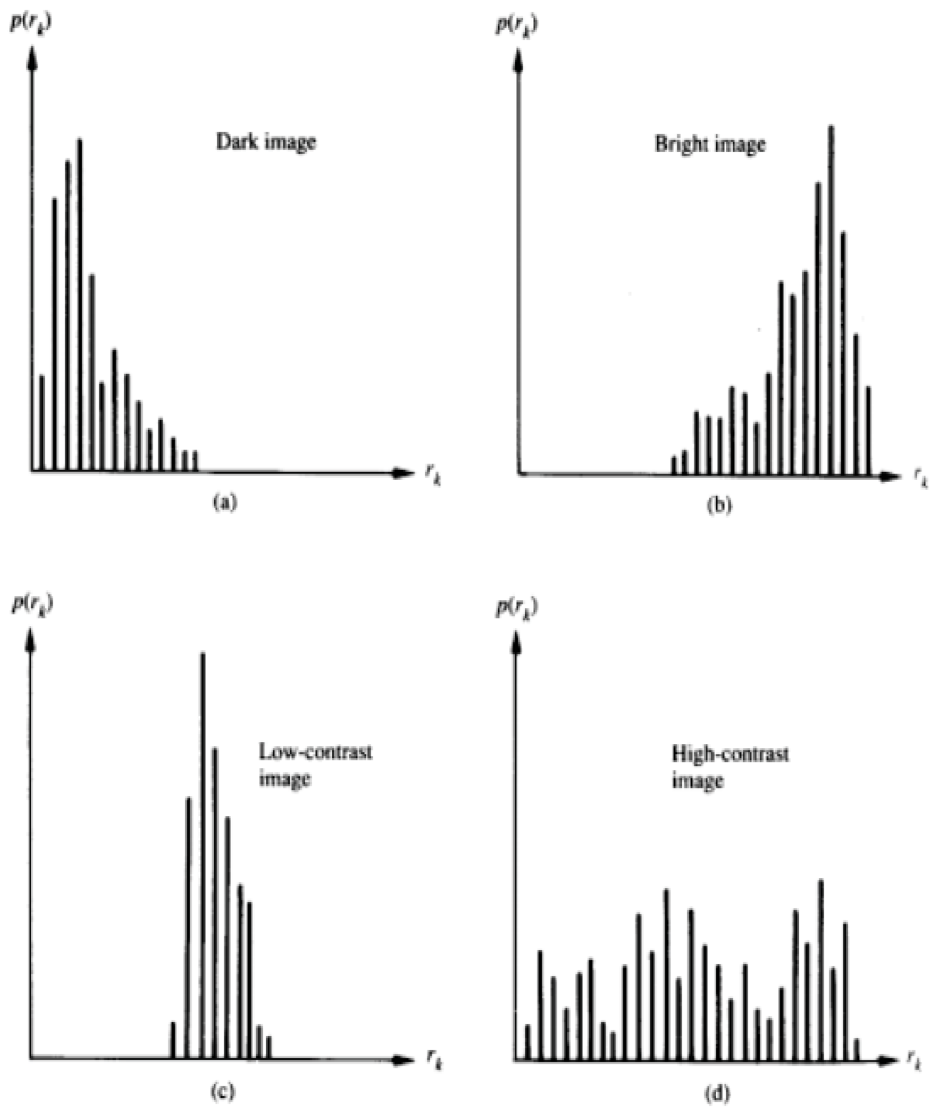
A digital image's histogram, which is a discrete function with grey levels between [0, L-1],

$$P(r_k) = n_k/n$$

The number of gray-level pixels in the original raw image is shown below. There are n total pixels in the image, and k is the grey level, which is defined with k=0, 1...L-1.

P provides the expected probability of grey scale occurrence ( ). The information regarding the potential for contrast enhancement is found using the image's histogram.

If the histogram is narrow and indicates a limited range of values, then the contrast of the image will be low.



3.2.1 Histogram equalization:

I. Histogram equalization's primary goal is to make a histogram that, after being related, is uniform by relating an input image to an output image.

II. Here, r will stand in for the image's gray levels, and s, the increased output, will be represented by the transformation s=T. (r).

III. To improve the dynamic range of pixels and obtain uniform density in gray levels, an image can be transformed using a function that is equivalent to the cumulative distribution of the probability density of the image pixels, denoted as r.

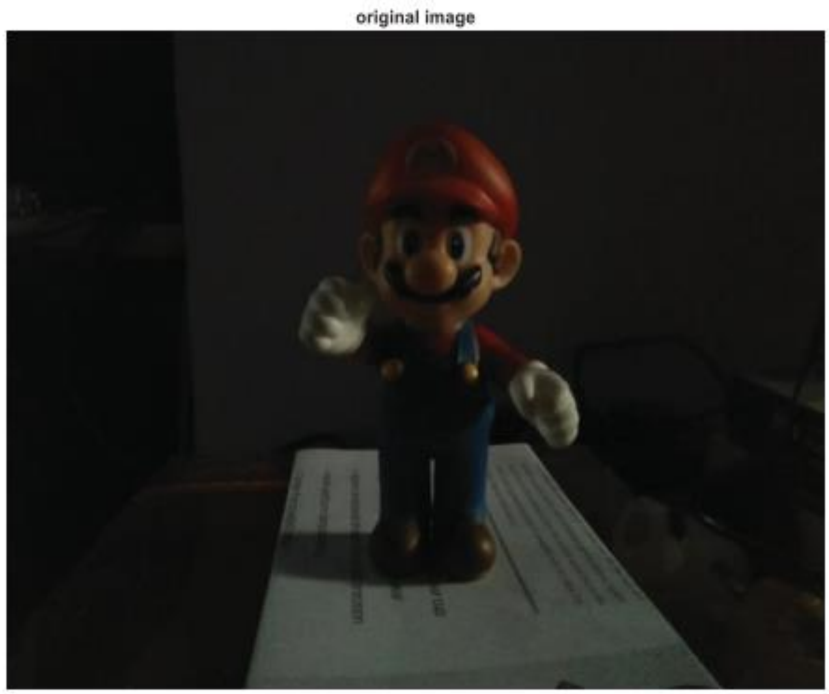
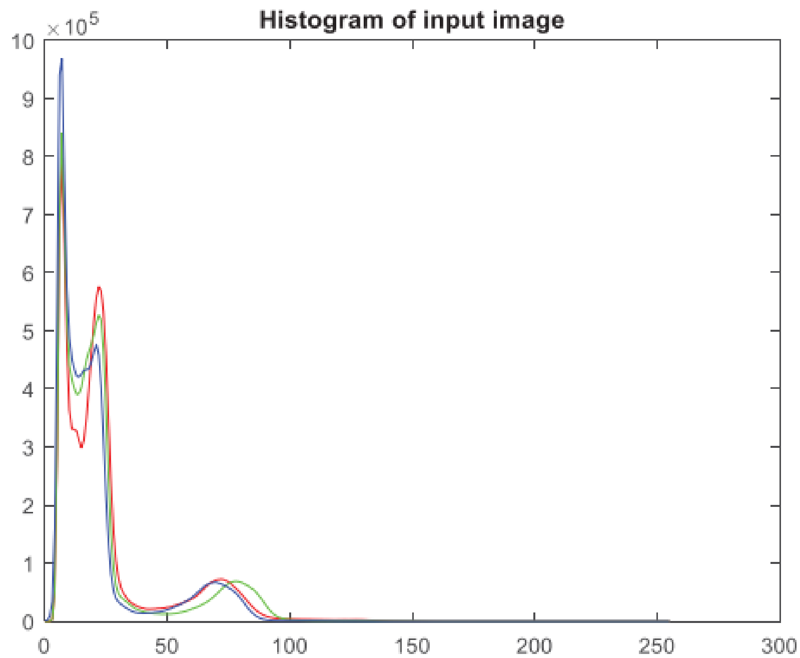
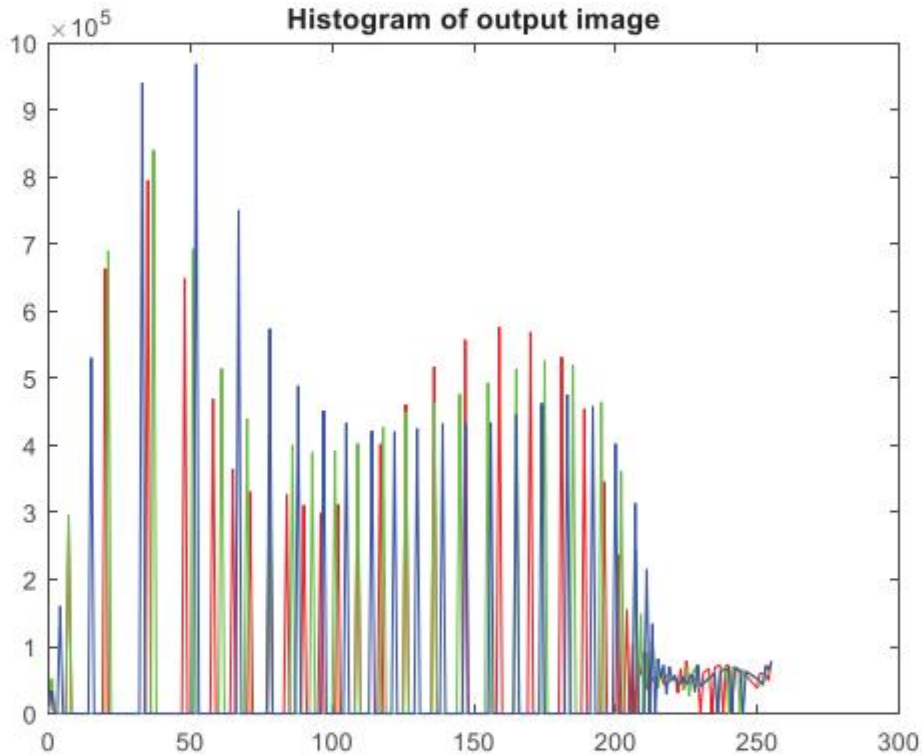


Fig 10


Histogram of input image
× 10^5
10
9
8
7
6
5
4
3
2
1
0
0
50
100
150
200
250
300


output image

Fig 11



Histogram equalization is a technique that enhances the contrast of an image by distributing intensity values throughout the entire range. However, this method is limited to images with even illumination, as it increases the intensity of brighter areas and decreases the intensity of darker areas. The result is an image with a wider dynamic range, but it may not be suitable for images with uneven lighting. The primary objective of histogram equalization is to uniformly distribute the image's contrast over the entire available dynamic range.

The goal of the histogram equalization technique in image processing is to modify an image's probability density function (PDF) and make it more uniform. Digital images are discrete functions, thus the likelihood of the data can be used to estimate the PDF on the histogram. We can use the likelihood of the data, for instance, to approximate the PDF on the histogram in a picture x with intensity levels ranging from 0 (black) to L1 (white). The objective is to spread out the PDF more evenly, from the lowest to the highest pixel value.

$$pdf(x) = p(r_k) = \frac{total\,pixels\,with\,intensity\,r_k}{total\,pixels\,I\,image\,x}$$

We can then derive the cumulative density function (cdf) from this pdf as follows:

$$cdf(x) = \sum_{k=0}^{L-1} p(r_k)$$

The probability of a pixel's intensity is represented by p(rk). The output of a pixel after histogram equalization is equivalent to the cumulative distribution function (cdf) of the image, or its numerical value..

$$p(s_k) = \sum_{k=0}^{L-1} p(r_k)$$

P(sk) must be multiplied by L1 and rounded to the nearest integer to obtain the value of the pixel.

4. ENHANCEMENT IN IMAGE PROCESSING

4.1 Introduction

Image enhancement enhances or improves the image's visual quality. The most straightforward and appealing aspect of digital image processing is image enhancement. The methods used for enhancement bring out specific features of interest in an image and accentuate their details. For a better-looking image, boost the contrast and brightness. Image enhancement is used to make digital images better suited for display or other picture analysis, such as noise reduction, sharpening, or image brightness. Frequency domain and spatial domain are the two basic divisions of the image enhancement process. In the frequency domain, techniques work with the image's frequency transform, whereas in the spatial domain, they can work directly with the pixels.

4.2 Contrast enhancement

By consistently enhancing the brightness differences throughout the image's whole dynamic range, the technique of tonal augmentation improves the contrast of an image. Contrary to tonal enhancement, which only modifies the brightness differences in a few regions—such as the shadow, midtone, or highlight areas—to balance the brightness differences in other regions, this technique adjusts the brightness differences in all regions.

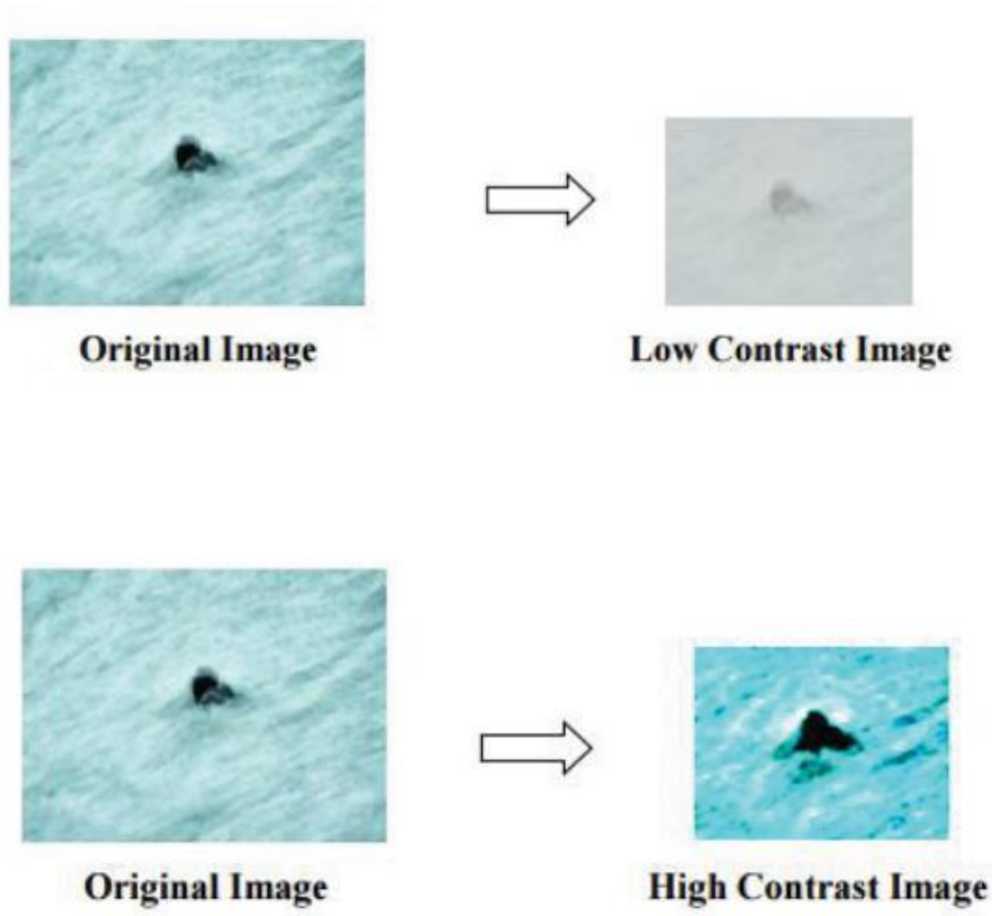


Fig 12

4.3 Some Proposed Algorithm:

4.3.1 Contrast Stretching

This technique increases the dynamic range of the grey levels in the image, which is beneficial for pictures with low contrast because of things like bad lighting, an insufficient dynamic range in the image sensor, or a poor choice of aperture when taking the picture.

4.3.2 Gray Level Slicing

The main goal is to keep the backdrop and grayscale tonalities while exhibiting a specified range of grey levels by brightening the chosen grayscale range. This is accomplished by creating a binary image by displaying high values for each grey that falls inside the relevant range and low values for all other grayscales.

4.3.3 Bit Plane Slicing

Each pixel in a digital image can be represented by 8 bits by splitting the image into eight 1-bit planes. The relevance of each bit in the image may be understood by analyzing these bit planes, which can then help calculate the right number of bits to quantize each pixel. Decomposition like this can help with image compression.

.

4.3.4 Histogram Equalization

One technique for enhancing contrast in a picture by using the histogram is known as histogram equalization [7]. Histogram equalization's primary purpose is to improve the histogram's distribution of intensities and boost the image's overall contrast. This makes it possible for regions with less local contrast to acquire more contrast. This technique works well for the image's foreground and background when they are both bright and dark

Histogram equalization is a method for improving an image's contrast by equally distributing the range of intensity values. For pictures with poorly lit backgrounds, it is ineffective. Because the method only increases the number of pixels in the bright areas of the image and decreases the number of pixels in the dark areas, the resulting image has a wider dynamic range.

Two techniques are used to perform histogram equalization.

1. Histogram Equalization on RGB image:

   The histogram equalization algorithm is used to independently equalize the independent R, G, and B planes in a color image. The RGB space's individual channels were each processed and equalized separately. Following that, all of the RGB components are combined to create a better image than the original.

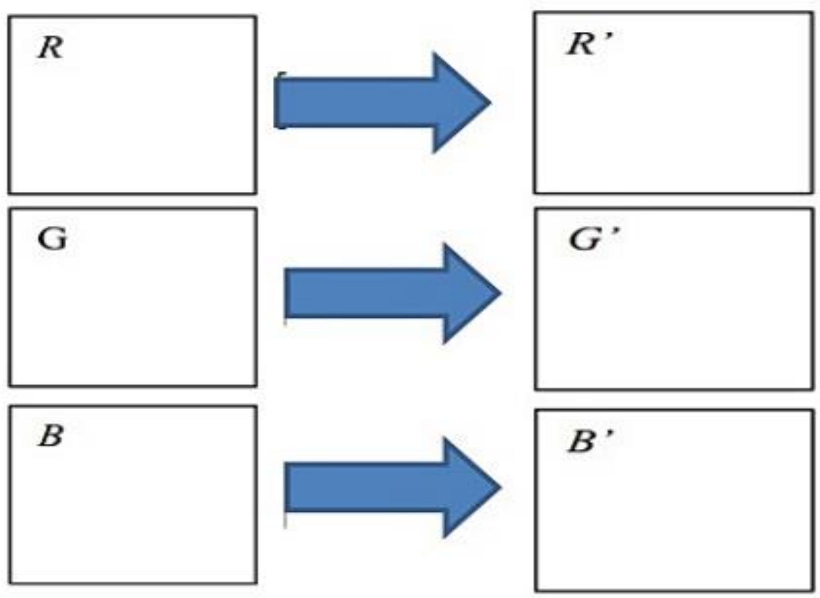


Fig 13

The probability density function (pdf) of an image is used in the histogram equalisation approach, which essentially converts the image's pdf into a uniform pdf that spans from the lowest pixel value ((0) to the highest pixel value (m-1)). As a result, the image's dynamic range, which is between 0 and 1, is divided uniformly. If the pdf is continuous, this can be done rapidly. A digital image's pdf, however, is discrete. Let's say we have an image called x that ranges from 0 (black) to m-1 in terms of dynamic range.

2. Histogram Equalization using V Component from HSV Space:

To apply this method, the image must first be converted to the hue, saturation, and value (HSV) color space. Saturation measures the degree of color purity while hue corresponds to the dominant wavelength of the color stimulus. When combined, hue and saturation are referred to as chromaticity coordinates in the polar system. The approach involves using histogram equalization on the value (V) component of the HSV color space. After equalization, the new values are combined with H and S in the V plane. Finally, the equalized image is compared to the original input image.

$$\text{HSV} \rightarrow \text{H, S, V}$$

$$\text{V} \rightarrow \text{V (equalize)}$$

$$\text{HSV (equalize)} \rightarrow \text{H, S, V (equalize)}$$

Chapter 2

Literature Review

2.1 IMAGE CONTRAST ENHANCEMENT USING HISTOGRAM EQUALIZATION[1]

Without adding visual artifacts that lower an image's visual quality and give it an unnatural appearance, the contrast of the image can be increased. In comparison to previous contrast enhancement algorithms, the experimental results demonstrate the algorithm's efficacy. The obtained photographs have an appealing aesthetic, are free of artifacts, and appear natural. The absence of flickering in this paper is a positive aspect. This is mostly because the input (conditional) histogram, which does not greatly alter within the same scene, is used by the algorithm as the main source of data. A wide range of photos can be used using this technique. Additionally, it provides a degree of controllability and adaptability that enables various degrees of contrast augmentation to no contrast enhancement.

2.2 IMAGE ENHANCEMENT USING BACKGROUND BRIGHTNESS PRESERVING HISTOGRAM EQUALIZATION[2]

Histogram equalization technique, which is a popular method for enhancing image contrast, but it has a tendency to excessively increase the brightness of the background. To address this issue, they introduced a method called brightness-preserving bi-histogram equalization (BBHE) that maintains the original brightness of the image by dividing it into two parts based on the input mean. These sub-images are then separately equalized and merged to generate the final image.

2.3 NONLINEAR TRANSFER FUNCTION-BASED LOCAL APPROACH FOR COLOR IMAGE ENHANCEMENT[3]

Developed a technique for improving color photographs based on pixel neighborhood and nonlinear transfer function while maintaining details.

In order to avoid affecting the color balance between the different components of the HSV color image, this method only modifies the V (luminance) component. The V component is divided into smaller, overlapping blocks, and a non-linear transfer function is applied to each pixel within these blocks to enhance the brightness. Then, each pixel is further adjusted to improve the image contrast based on the values of the surrounding pixels. Meanwhile, the H and S components are left unchanged..

2.4 CONTRAST ENHANCEMENT OF DARK IMAGES USING STOCHASTIC RESONANCE[4]

Suggested a method based on stochastic resonance that is used to improve very low-contrast photographs. An equation for the ideal threshold has been developed using this method. The low-contrast image has been progressively upgraded with increasing standard deviation Gaussian noise until the quality is at its highest.

2.5 ON THE IMAGE ENHANCEMENT HISTOGRAM PROCESSING[5]

Histogram equalization is a useful technique for enhancing the image's contrast. A source of high-enhancement, an image's natural appearance is not retained. In order to get over this restriction, multi-HE technique is used to maintain both brightness and natural occurrence of the digital image. According to the mean and median threshold values, the recommended approach separates the histogram of an input digital picture into a number of sub-histograms. The narrow segments are determined by the number of segments, vital range of each slab, and intensity level. After detecting the small sections and leaving the bigger areas alone, it is scaled to the complete dynamic range. A separate histogram equalization application is made for each slab. In order to prevent intensity saturation and an unsettling bin distribution, the normalization is eventually constructed for an equalized histogram. As the number of segments rises, the contrast of the picture and brightness preservation both improve. The natural occurrence of the digital image cannot be retained due to the information loss.

2.6 CONVOLUTIONAL NURAL NETWORK(CNN)

The capacity of CNNs to accommodate spatial data, such as photographs, gives them a significant edge over other forms of neural networks. Whatever their position, orientation, or scale, they can learn to recognise patterns in the image. This is made possible by the convolutional layers' shared weights, which enable the network to identify the same feature across several regions of the image.

It is crucial to remember that training CNNs can be computationally demanding and necessitates a substantial amount of labeled data. Additionally, if the images have complex backgrounds or occlusions, they might not perform well on images that are significantly different from the training data.

The fundamental features of a digital image are edges that run horizontally or vertically, and the first layer of CNN extracts these layers. T The first layer acts as a frontline filter and is used as the input for the following layer, which extracts more complicated features including corners, curves, and multiple edges. This continues through CNN's final tier. The classification layer assigns ratings to each image in the 0–1 range, and the model uses these ratings to assign the image to one of several classes depending on the activation map that is present in the final convolution layer..Different models of CNN are

2.6 INCEPTION MODEL

The inception module of the network comprises 27 convolutional layers with diverse filter sizes and is followed by a pooling layer. This allows the network to gather both detailed and high-level features from the input image. The "dimensionality reduction" method, which lowers the input's dimensionality before convolutional layers are applied, is another method used by the Inception model. This aids in lowering the network's computing expense while preserving accuracy.

The Hebbian principle states that while building a new layer in a deep learning model, it is crucial to take into account the knowledge obtained from earlier layers.

In artificial neural networks, the Hebbian principle has also been employed as a learning principle to modify the strength of connections between the virtual neurons. As a result, machine learning models with Hebbian roots, like the self-organizing map and the Hopfield network, have been created.

Architecture of Inception model

The Inception model's combination layers can be further divided into three categories:

i) Pooling layers, which downsample the input and capture spatial information.

ii) 1x1 convolutional layers, which reduce the input's dimensionality and aid in improving computational efficiency.

iii) convolutional layers of greater sizes, which capture more complex characteristics.

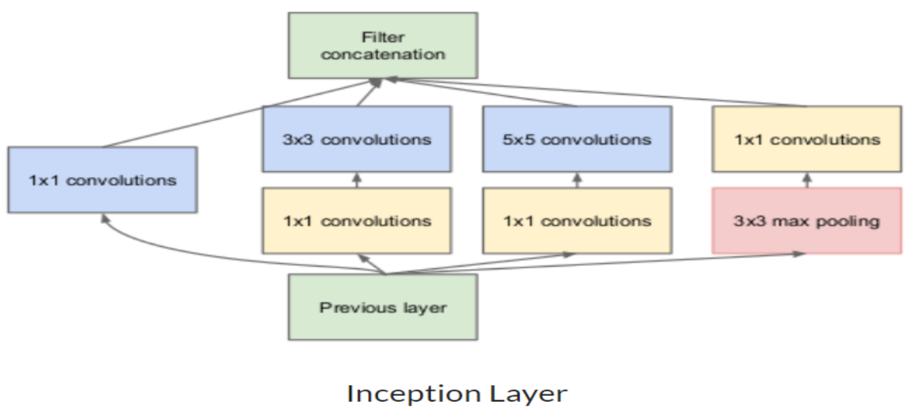
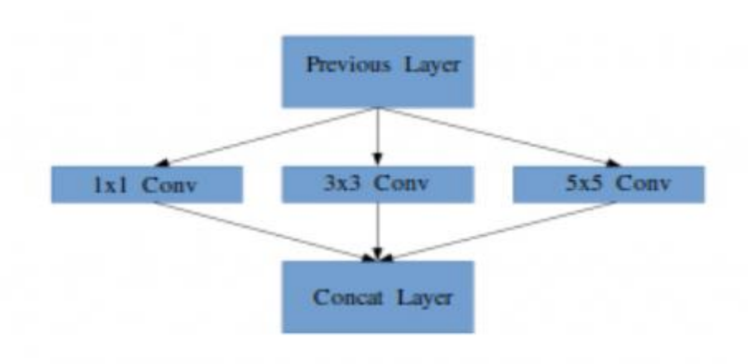


Fig 14

Architecture Of Inception Model

convolution
max pool
convolution
max pool
inception (3a)
inception (3b)
max pool
inception (4a)
inception (4b)
inception (4c)
inception (4d)
inception (4e)
max pool
inception (5a)
inception (5b)
avg pool
dropout (40%)
linear
softmax

Fig 15



Idea of an Inception module

Fig 16

Envision a neural network layer that has mastered the ability to recognise particular facial traits. The next layer of the network would probably concentrate on the complete face in the image to detect various things that were there. The layer needs the proper filter sizes for recognising different objects in order to complete its assignment.

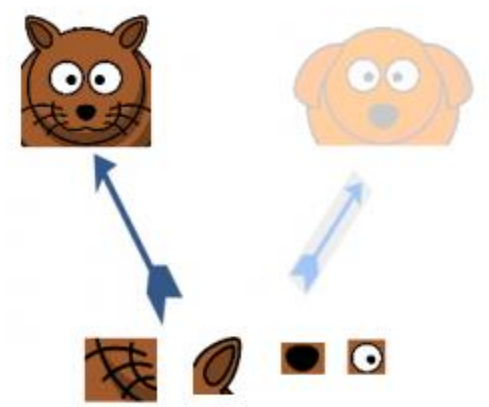


Fig 17

The inception layer emerges at this point. It enables the internal layers to select the appropriate filter size for learning the necessary information. Therefore, the layer functions appropriately to recognise the face. It would probably be necessary to use a larger filter size for the first image while using a smaller one for the second.

2.7 ResNet-50 ARCHITECTURE

ResNet-34, a version of the original ResNet design, included 34 weighted layers. An innovative method for increasing the number of convolutional layers in a CNN while avoiding the issue of vanishing gradients was devised through the use of shortcut connections. This method creates a residual network rather than a regular one by including shortcut links that "jump over" multiple levels.t.

The ResNet architecture is guided by two key design principles.

(i) Each layer has the same number of filters, irrespective of the size of the output feature map.

(ii) the number of filters in each layer is doubled to maintain the time complexity of the layer, even if the size of the feature map is reduced by half.

Special characteristics of ResNet-50

The basic building blocks of ResNet50 is bottleneck design. The bottleneck blocks, which are made up of three layers—a 1x1 convolutional layer, a 3x3 convolutional layer, and another 1x1 convolutional layer—are the foundation of ResNet-50. While the middle 3x3 convolutional layer learns the key features, the first and third layers use 1x1 convolutional filters to first reduce and then restore the dimensions of the input. Additionally, batch normalization layers and ReLU activation functions are included in each bottleneck block.The components of the 50-layer ResNet architecture are listed below:

I. a convolution of a 7x7 kernel with 64 additional kernels and a stride of size 2
II. A stride of two sizes with a maximum pooling layer..
III. There are nine more layers total—three 3x3,64 kernel convolution levels, one 1x1,64 kernel layer, and one 1x1,256 kernel layer. Repeating these three layers three times.
IV. 12 more layers with iterations of 4, 1x1, 128 kernels and 1x1, 512 kernels each were added.
V. 18 more layers with 1×1,256 cores
VI. 9 more layers with 1×1,512 cores

Advantage Of CNN

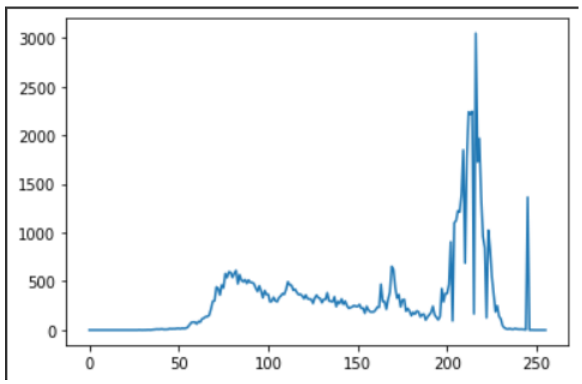
Deep learning is a type of machine learning that involves neural networks consisting of at least three layers. Neural networks with multiple layers generally produce more precise results compared to those with a single layer. Convolutional neural networks (CNNs) or recurrent neural networks (RNNs) are commonly used in deep learning applications. CNNs are a popular neural network architecture for computer vision tasks such as image recognition and classification. They excel at processing large amounts of data and can generate accurate predictions by learning the object's features through multiple iterations. This approach eliminates the need for manual feature engineering, such as feature extraction..Transfer learning refers to the process of reusing or adapting pre-trained models for new tasks or datasets. In particular, with CNNs, developers can modify and fine-tune existing networks with trained weights, or use pre-trained models as a starting point for building new models. This approach saves time and computational resources while also improving the accuracy and generalization capabilities of the model.

Chapter 3

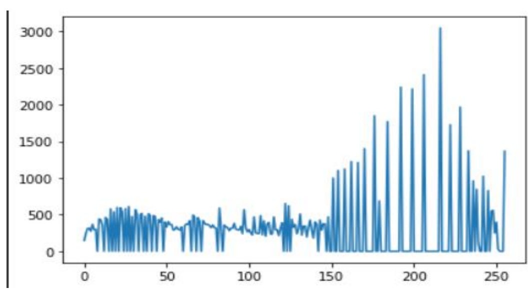
3.1 Analysis

Here we are using a collection of images and drawing their respective Histogram and Adaptive Histogram Graph. Respective graphs are shown below.

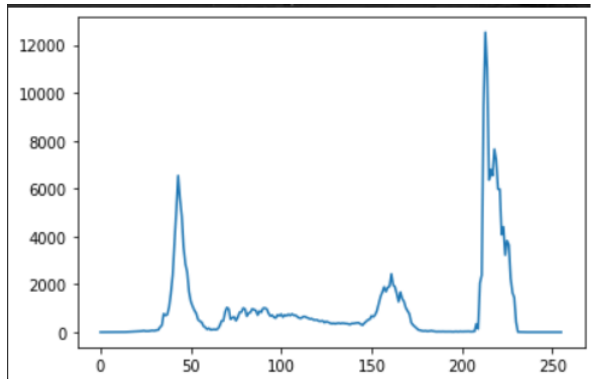
Before Applying AHE



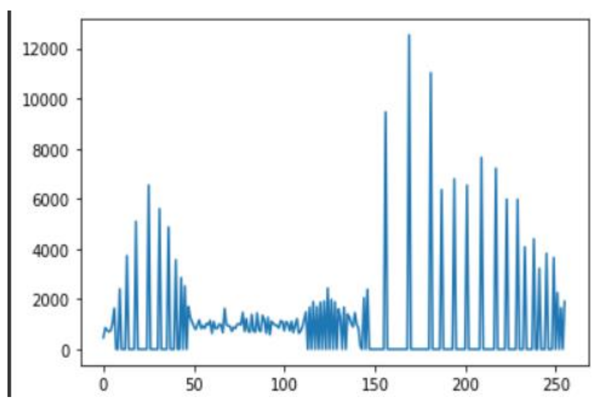
After Applying AHE



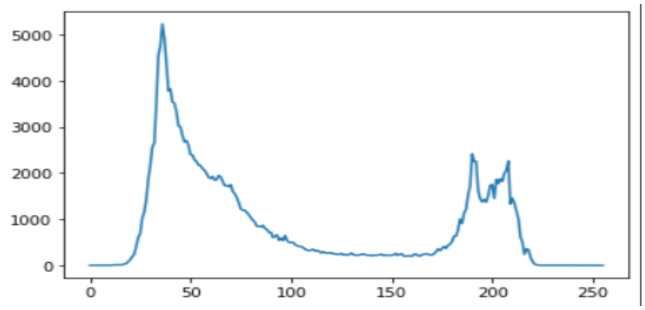
Before Applying AHE



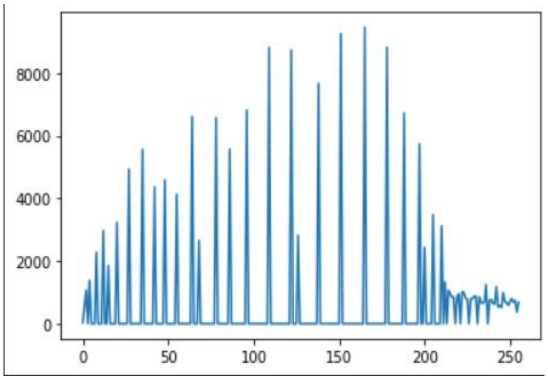
After Applying AHE



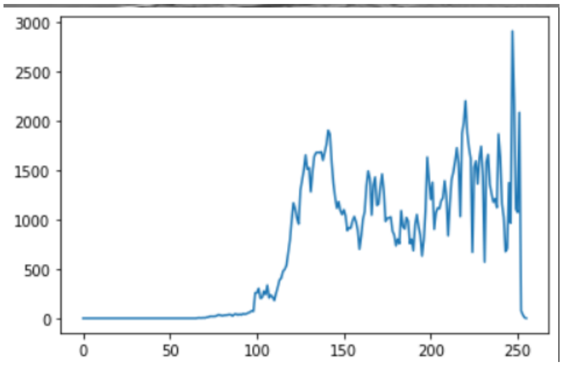
Before Applying AHE



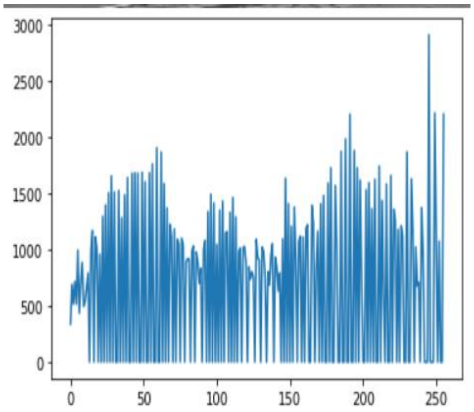
After Applying AHE



Before Applying AHE



After Applying AHE



3.2 Algorithm Used

3.2.1 Histogram Equalization

Histogram equalization is a method of image processing that uses the histogram of the image to change the contrast.

This method is commonly used to enhance the contrast of images when the useful data is represented by similar contrast values. By doing so, the intensities on the histogram can be distributed more evenly, allowing areas with less local contrast to gain more contrast. Histogram equalization effectively distributes the most common intensity values, resulting in this effect. This technique is particularly useful for images in which both the foreground and background have either low or high brightness levels.

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j)$$

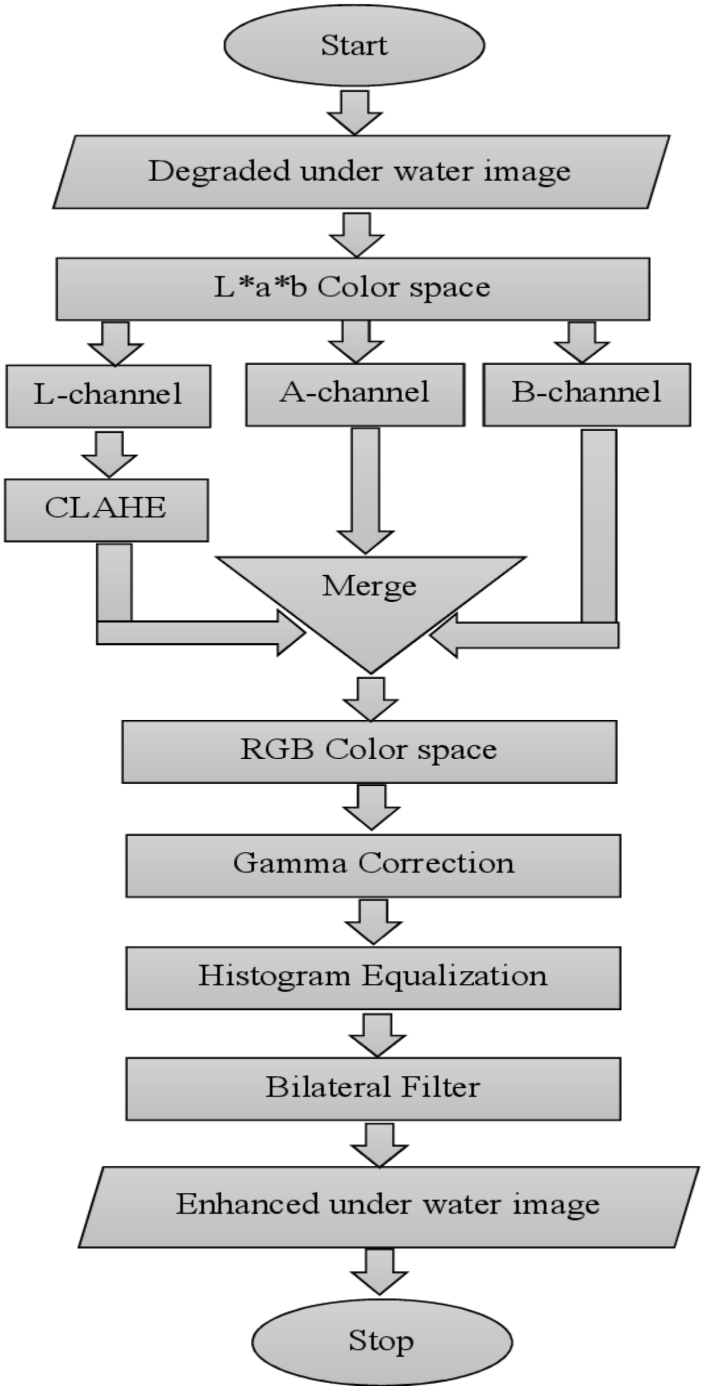
3.2.2 OpenCV

OpenCV is a widely-used open-source library for machine learning, computer vision, and image processing. It has a major role in contemporary systems that require real-time operations. OpenCV can be used to detect people, objects, and handwriting in images and videos. By combining it with other libraries like NumPy, Python can effectively handle the array structure of OpenCV for analysis. To detect and understand visual patterns, we utilize vector space and perform mathematical operations on these features.

3.2.3 CLAHE Histogram Equalization

The CLAHE is an enhanced version of AHE that solves the problem of contrast over-amplification. Rather than processing the entire image at once, CLAHE works on small, discrete sections called tiles. To prevent artificial boundaries, the neighboring tiles are merged using bilinear interpolation. This technique is effective in enhancing the contrast of images..

On color photographs, we can also apply CLAHE;
Typically, the luminance channel is used for this, and when only the luminance channel is adjusted, the outcomes for an HSV image are significantly better than those for a BGR image where all channels are adjusted.


Start
Degraded under water image
L*a*b Color space
L-channel
A-channel
B-channel
CLAHE
Merge
RGB Color space
Gamma Correction
Histogram Equalization
Bilateral Filter
Enhanced under water image
Stop

Chapter 4

Output at several stages are shown below

Image 1

Original

Equalized

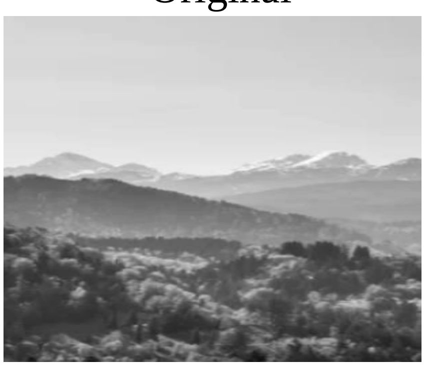
Image 2

Original

Equalized

Image 3

Original



Equalized

Image 4

Original

Equalized

Image 5

Original

Equalized

Image 6

Original

Equalized

Image 7

Original

Equalized

Image 8

Original

Equalized

Image 9

Original

Equalized

Chapter 5

5.1.Conclusion

This project proposes and applies methods for the enhancement of full colour photos that strike a compromise between the demands of appearance enhancement and maintaining an image's original appearance. Results have demonstrated our algorithm's efficacy in enhancing the contrast and vibrancy of the source photos. With our own proposed histogram and three ways of histogram equalization with CNN models on low light images, we have demonstrated in this paper that we can get superior images through histogram specification.

The approaches used in this thesis are used for the enhancement of both RGB and grayscale images, balancing the requirements of the developed enhancement and maintaining the integrity of the original image. The outcomes demonstrate how well the algorithm worked to enhance the original image's color and contrast.

Finally all the results are shown in the Chapter 3 and Chapter 4 where we have compared the different techniques on different images and how efficient they are based on the given image.

The Inception model has some limitations despite its impressive success in image classification. Understanding how the network makes predictions can be challenging since it can be difficult to interpret the features the network learns. The model's high processing demands could also make scaling up training and implementation difficult.

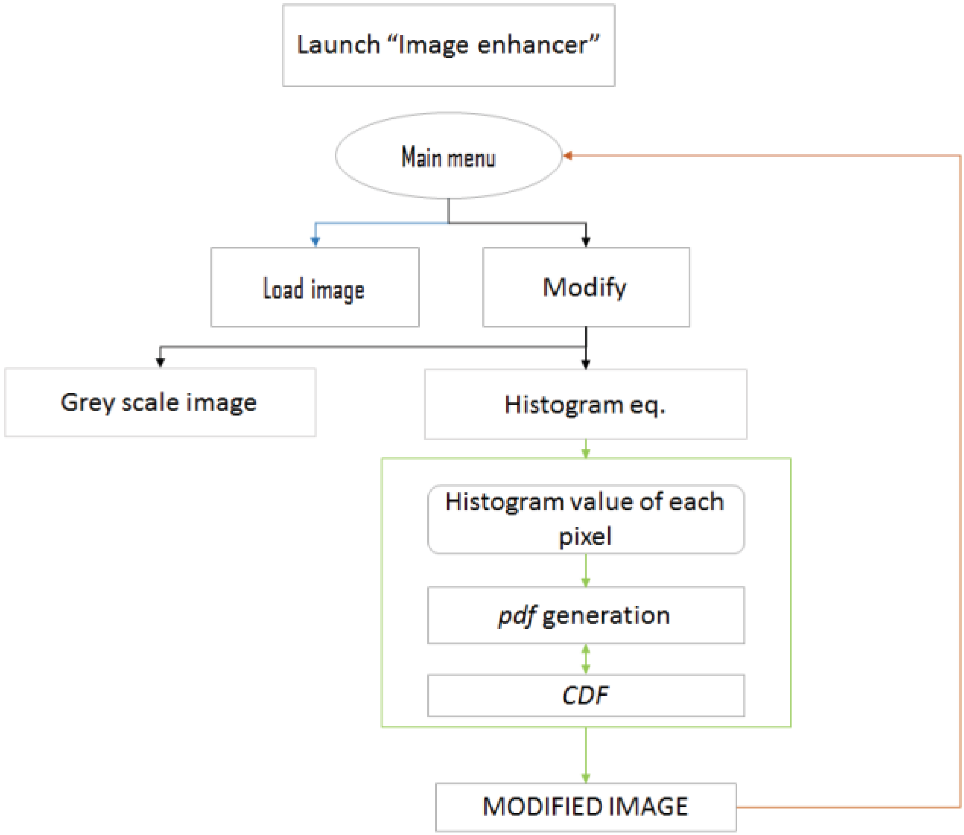
The computational complexity of ResNet-50, which could make it difficult to deploy and train the model on low-end hardware, is one possible drawback.

An updated anchor-based object detection method that outperforms previous iterations is introduced in YOLOv7. With this technique, object localization can be done more precisely, and false positives are less likely to occur.

5.2. Future Scope

Although all these techniques work perfectly fine and give us favorable results we can use android studio to make it work on the everyday smartphone as we believe that it will increase the user experience and will come much handy especially given in the smartphones.

A general flowchart of how our image enhancer will work is given below.



5.3 Application

The main idea behind this project is to increase the contrast of a very low contrast image while preserving all the necessary details and also reducing the noise and graininess.There are three types of histogram techniques used in this project but there are still a few more left like Gamma correction which can be very helpful if the input image is medical image modalities.

5.3.1 Magnetic Resonance Imaging

Nuclear magnetic resonance spectroscopy shares the same fundamentals as magnetic resonance imaging (MRI), a type of medical imaging. MRI is the most reliable and non-invasive technique for disease clinical diagnosis.. The MRI scanner takes
photographs of the body's tissues, organs, and other structures using magnetic and radio waves. An MRI scan produces images that are superior at portraying fine details. As a result, an MRI scanner can be used to take photos of all the body's tissue. Bones, which have the fewest hydrogen atoms, become dark, and fatty tissue, which has the most hydrogen atoms, is noticeably brighter.

5.3.2 Computed Tomography

The method known as computed tomography (CT) uses X-rays along with computer algorithms to create images of the body's tissues. In the field of medical imaging, the CT scan is one of the crucial diagnostic tools used to visualise the inside anatomy of the human body. CT scan provides high-quality images with good contrast between different soft tissues of the body, making it an ideal imaging modality for the brain, muscles, and cancers. Recent technological advancements in CT machines have been made to enhance the image contrast, particularly for diagnostic purposes.

5.3.3 Image Classification

Convolutional neural networks (CNNs) are widely used in business because they allow computers to automatically classify and comprehend visual input, which has many uses in a variety of industries.

References

[1] J.A. Stark, “Image Contrast Enhancement Using histogram equalization”, IEEE Trans. Image Process, vol. 9, no. 5, pp. 889-896, May 2015

[2] T.L. Tans , K.S. Sim and C.P. Tso,”Image enhancement using background brightness preserving histogram equalization”, 2nd International Conference on Intelligent Computing , Communication & Convergence, 2020, pp- 1389-1387.

[3] Deepak Ghimire and Joonwhoan Lee,”Nonlinear Transfer Function-Based Local Approach for Color Image Enhancement”, IEEE Transactions on Consumer Electronics,2019,10.1109/TCE.2019.5955233.

[4] R.K. Jha, P.K. Biswas, B.N. Chatterji “Contrast enhancement of dark images using stochastic resonance” IET Image Process., 2012, Vol. 6, Iss. 3, pp. 230–237; doi: 10.1049/iet-ipr.2010.0392

[5] G.Senthamarai and K. Santhi,"Dynamic multi histogram equalization for image contrast enhancement with improved brightness preservation," Electronics and Communication Systems (ICECS), 2019 2nd International Conference on, Coimbatore, 2019, pp.1205-1209.

[6] Ziad M. Abood, Kadhim K. Kadhim, (2019), Assessment the Quality of Medical Images (CT & MRI) by Using Wavelet Transformation (WT),

International Journal of Emerging Research in Management & Technology, vol. 4, Issue 7.

[7] A. McAndrew, An Introduction to Digital Image Processing. Boston, MA, United States: Course Technology Press, 2014.

[8] Chien-Yao Wang1 , Alexey Bochkovskiy, and Hong-Yuan Mark Liao1 1 Institute of Information Science, Academia Sinica, Taiwan ,YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detector,2022