PROJECT STAGE - I

On

Retina Gray Scale Image Perception using Deep Learning

Submitted in partial fulfillment of the requirements for the award of degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE ENGINEERING

by

19WH1A0562 A. DHANUHYA

19WH1A05A3 P. HARSHINI REDDY

19WH1A05A7 N. ANJALI

Under the esteemed guidance of

Ms. Shanmuga Sundari

Assistant Professor

Department of Computer Science Engineering

BVRIT HYDERABAD

College of Engineering for Women

(Approved by AICTE, New Delhi and Affiliated to JNTUH, Hyderabad)

Accredited by NBA and NAAC with A Grade

Bachupally, Hyderabad – 500090

2022-2023

DECLARATION

We hereby declare that the work described in this report, entitled "**Retina Gray Scale Image Perception using Deep Learning**" which is submitted by us in partial fulfillmentfor the award of the degree of **Bachelor of Technology** in the department of **Computer Science Engineering** at **BVRIT HYDERABAD College of Engineering for Women**, affiliated to **Jawaharlal Nehru Technological University Hyderabad**, Kukatpally, Hyderabad – 500085 is the result of original work carried out under the guidance of **Ms. Shanmuga Sundari, Assistant Professor, Department of CSE**.

This work has not been submitted for any Degree / Diploma of this or any other institute/university to the best of our knowledge and belief.

A. Dhanuhya
(19WH1A0562)

P. Harshini Reddy
(19WH1A05A3)

N. Anjali
(19WH1A05A7)

Department of Computer Science Engineering

BVRIT HYDERABAD
College of Engineering for Women

(Approved by AICTE, New Delhi and Affiliated to JNTUH, Hyderabad)

Accredited by NBA and NAAC with A Grade

Bachupally, Hyderabad – 500090

CERTIFICATE

This is to certify that the major project entitled "**Retina Gray Scale Image Perception using Deep Learning**" is a bonafide work carried out by **Ms. A. Dhanuhya (19WH1A0562),Ms. P. Harshini Reddy (19WH1A05A3), Ms. N. Anjali (19WH1A05A7)** in partial fulfillmentfor the award of B.Tech. degree in **Computer Science & Engineering** , **BVRIT HYDERABADCollege of Engineering for Women, Bachupally, Hyderabad**, affiliated to **Jawaharlal Nehru Technological University Hyderabad**, under my guidance and supervision.

The results embodied in the project work have not been submitted to any other University or Institute for the award of any degree or diploma.

Internal Guide	**Head of the Department**
Ms. Shanmuga Sundari	**Dr. R. Suneetha Rani**
Assistant Professor,	**Professor and HoD,**
Department of CSE	**Department of CSE**

External Examiner

ACKNOWLEDGEMENTS

The satisfaction that accompanies in successful completion of the task would be incomplete without the mention of the people who made it possible.

We would like to express our sincere thanks to **Dr. K. V. N. Sunitha, Principal, BVRIT HYDERABAD College of Engineering for Women**, for her support by providing the working facilities in the college.

Our sincere thanks and gratitude to **Dr. R.Suneetha Rani, Head, Department of CSE, BVRIT HYDERABAD College of Engineering for Women**, for all timely support and valuable suggestions during the period of our project.

We are extremely thankful to our Internal Guide, **Ms. Shanmuga Sundari, Assistant Professor, CSE, BVRIT HYDERABAD College of Engineering for Women**, for her constant guidance and encouragement throughout the project.

Finally, we would like to thank our Major Project Coordinator, all Faculty and Staff of CSE department who helped us directly or indirectly. Last but not least, we wish to acknowledge our **Parents** and **Friends** for giving moral strength and constant encouragement.

A. Dhanuhya (19WH1A0562)

P. Harshini Reddy (19WH1A05A3)

N. Anjali (19WH1A05A7)

ABSTRACT

Millions of people in the world are affected by ocular fundus diseases such as diabetic retinopathy, age-related macular degeneration, glaucoma, retinal detachment, and fundus tumors, without accurate diagnoses and timely appropriate treatment, these fundus diseases can lead to irreversible blurred vision or even blindness. Manual vessel segmentation is always difficult in a retinal image because of the varying thickness and width of retinal blood vessels. Retinal blood vessel segmentation has become mandatory for any type of retinal disease.

Segmentation of retinal blood vessels is considered an effective technique for diagnosing these diseases. We are aiming to solve the problems of serious segmentation errors of eyes by using network architectures, like semantic pixel-wise segmentation (SegNet) and U-net. The accurate segmentation of retinal blood vessels in the fundus is of great practical significance to help doctors diagnose fundus diseases.

LIST OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

As a vital part of the eye, the retina plays a significant role in the human body. Humans rely on their eyes for 80% of the ways they receive information from the outside world. The quality of human eyesight is significantly influenced by the retina. Clinical medicine currently relies on ophthalmological tools and medical imaging technology to diagnose and treat individuals with retinal injury or detached blindness.

The retina produced by the current imaging technology has significant flaws, which makes it difficult for medical treatment and performing attentive observation and correct judgment. Therefore, some CNN Algorithms can be used. These algorithms will help the doctors by providing a grayscale image of the retina image in which has a clear vision of optic nerves, and blood vessels of the eye which helps in better understanding of the eye for the doctors. Then from the image doctors can treat the patients with better analysis.

In Order to early detection of Eye Disease, Convolutional Neural Network (CNN) approach is proposed using the images (retinal) as input to it. CNNs were applied individually to solve certain problems.CNN is a method of deep learning which has a magnificent record of advantages for image analysis and interpretation. CNN is considered one of the important methods to fix an automatic analysis of retinal images.

1.1 Objective:

The main objective of the project is to help doctors to analyze retinal images and get a clear knowledge of the patient's conditions. This facilitates the work of medical professionals by offering a more accurate understanding of the blood vessels, and optic nerves of an eye

1.2 Problem Statement:

80% of the channels for humans to obtain information from the outside world come from the eyes. The retina is an important part of the human eye and plays a decisive role in the quality of human vision. Grey-scale images of the retina can help patients with damaged or detached blindness achieve the possibility of improving their vision and regaining light.

1.3 Existing system:

The existing system have combined the light field imaging technology with reconstruction algorithms like nearest neighbors interpolation, linear interpolation and optimization algorithm.

1.4 Proposed System:

On the basis of the literature review and subsequent findings, a deep learning-based model/method for retinal image segmentation is proposed. The model employs a modified version of the popular U-NET architecture and Seg-Net Architecture along with proposed pre-processing algorithms in order to segment the input images.

2. LITERATURE SURVEY

Title: Light-Field Imaging Reconstruction Using Deep Learning Enabling Intelligent Autonomous Transportation System

Authors: Davi Ribeiro Militani, Demóstenes Zegarra Rodríguez, Juan Casavílca Silva, Lunchakorn Wuttisittikulkij , Muhammad Saadi , Renata Lopes Rosa, and Sattam Al Otaibi

Summary:

Plenoptic cameras, also referred to as light-field (LF) cameras, enable the recording of the 4D LF distribution of target scenes. Commercial plenoptic cameras, intelligent transportation systems, and other sectors and applications have all been investigated using THE light-field (LF) imaging. Because surface flaws and parallax problems must be taken into account when building accurate maps for the Intelligent Autonomous Transport System (IATS), light-field (LF) camera images are essential. To produce high-quality photos, a learning-based architecture that directly imitates the LF distribution is provided. The performance of the suggested framework is assessed using a variety of image quality assessment techniques, including PSNR, SSIM, IWSSIM, FSIM, GFM, MDFM, and HDR-VDP. Deep learning algorithms like CNN, GAN , SRCNN. In this work, LF image reconstruction framework that efficiently learns the picture distribution, reconstructs high-quality and densely sampled LF images, and extracts the spatioangular information of a scene. The suggested model additionally improves the spatial and angular resolutions by simulating the local LF distribution.

Title: Harnessing Multi-View Perspective of Light Fields for Low-Light Imaging

Authors: Kranthi Kumar Rachavarapu, Kaushik Mitra and Mohit Lamba

Summary:

The main goal of this effort was to improve LF recorded under dim lighting.Light Field (LF) offers special benefits like depth estimation and post-capture focussing, but low light situations severely restrict these capabilities. For Low-Light Light Field (L3F) restoration, a deep neural network called L3Fnet that not only improves the visual quality of each LF view but also maintains the epipolar geometry across views. L3F-wild dataset, which contains LF captured late at night with virtually zero lux values, to further examine the effectiveness of low-light restoration techniques. LF Processing Algorithm is used . completing trials on the L3F-20, L3F-50, and L3F-100 datasets, the effectiveness of L3Fnet on LFs at various levels of low light was demonstrated. Results on the L3F-wild dataset demonstrated that L3Fnet could now automatically adjust to various light levels as a result of this preprocessing module.

Title: Raindrop Removal With Light Field Image Using Image Inpainting

Authors: Hang Su , Tao Yang , Nathan Crombez , Yassine Ruichek, Tomas Krajnik , Xiaofei Chang and Zhi Yan

Summary:

In this study, an image inpainting technique for removing raindrops from light field images. First thing is to identify raindrop regions using the depth map created from the light field image, and these regions are then expressed as a binary mask. By shifting the focus to the far distance, the original raindrops image is made better. precision of the visual localisation percentages for the original sub-aperture photographs, improved images, and images with raindrops removed are 80%, 82%, and 86%, respectively.To assess the suggested picture restoration technique, image quality analysis is carried out. For the perception tasks in adverse weather, it's crucial to take advantage of both the rich ray information in light field images and the most recent advancements in deep learning-based image processing. Additional applications of the restored images include object detection and visual. Making raindrop removal using light field image an end-to-end is intriguing paradigm going forward. Additionally, merging the light field dataset with the EU long-term dataset for long-term vehicle autonomy is a focus of future work.

Title: A Double-Deep Spatio-Angular Learning Framework for Light Field-Based Face Recognition

Authors: Alireza Sepas-Moghaddam, Fernando Pereira, Kamal Nasrollahi , Mohammad A. Haque, Paulo Lobato Correia and Thomas B. Moeslund

Summary:

In this paper, a framework for light field-based face identification using double-deep spatio-angular learning is proposed. Using two deep networks in succession, intra-view/spatial and inter-view/angular data The long short-term memory (LSTM) recurrent network, whose inputs are VGG-Face descriptions, generated using a VGG-16 convolutional neural network, is a component of the proposed double-deep learning system (CNN). The LSTM network then evaluates a series of VGG-Face spatial descriptions. Utilizing the IST-EURECOM light field face database, a wide range of experiments tackling complex recognition problems have been carried out. Although each SA image's location within the multi-view array and the order in which the SA images are to be scanned are known, there is still some additional information regarding the inter-view angular information/dependencies, such as parallax, that could be further utilised during the learning process to improve the recognition accuracy and/or convergence speed. Future studies will examine how to further use the additional angle information by extending the LSTM for spatio-angular visual recognition challenges.

Title: Deep Light Field Super-Resolution Using Frequency Domain Analysis and Semantic Prior

Authors: Gangyi Jiang, Mei Yu, Yeyao Chen , Yo-Sung Ho and Zhidi Jiang

Summary:

We have been widely concerned by the limited size of the imaging and angular resolutions. To this end, this paper proposes a new. enhance the spatial and angular resolutions of LFI. a frequency restoration process with new cascaded 2D and 3D convolutional neural networks. The designed network to enhance its representation ability and domain transformation. The field of LF research's "hot spot" has been LFSR.

Generally speaking, LFSR can be classified into two categories: angular SR, which synthesizes fresh views to improve angular resolution, and spatial SR, which improves the spatial resolution of LF images. Firstly, bidirectional optical flows and spatial interpolation are utilized to reconstruct the intermediate. the spatial and angular information frequency components. In addition, semantic priors containing the network to further improve its expression ability. The recovered results are transformed by inverse DCT to reconstruct the high-quality LFI. Unfortunately, effectively determining disparity from low-resolution (LR) LFI is difficult and expensive. CNN was employed to handle LF data because of its widespread performance in image restoration tasks and its ease of usage. Since LFI has a 4D structure, adopting 4D convolution for modelling makes sense, but it produces a large network architecture and necessitates a large amount of training data. In order to create a 2D CNN for modelling, the spatial-angular properties of LFI were examined. This led to the development of bidirectional recurrent CNN, spatial-angular interactive network, and spatial-angular separable convolution Demonstrates that the proposed method can effectively generate LFI with high spatial-angular resolution. performance in both subjective and objective aspects.

Title: CapsField: Light Field-Based Face and Expression Recognition in the Wild Using Capsule Routing

Authors: Ali Etemad, Alireza Sepas-Moghaddam, Fernando Pereira and Paulo Lobato Correia

Summary:

Such as biometrics and affective computing are recognition in the wild in this regard, this research suggests that learning hierarchical links between capsules, created from each LF image, be preserved and made public. A detailed performance assessment research employing the new to perform personal identity, and expression recognition advances after the first automatic facial recognition. Despite recent developments in face and expression analysis, some conditions still cannot be accurately assessed. Systems for analysis cameras with lenslet light fields(LF), have recently come into prominence as they are able to simultaneously capture the intensity of light rays coming from multiple directions in space. The visual scene is "seen" by an LF camera. A multi-view is formed by the collection of 2D SA pictures that have been rendered face image analysis. The performance of the recognition solutions presents the biggest hurdle in this situation. This work takes the recognition paradigm into account. It suggested a solution by CapsField: CapsField is a brand-new deep learning approach that is suggested for face and expression identification in the real world. CapsField blends convolutional and capsule networks to take advantage of both the spatial and angular data present in LF images. Recently, a few solutions based on capsule networks have been put forth in the context of facial image analysis, including age and gender classification and expression identification, demonstrating how well-suited capsule networks are to handle variations in face pose. CapsField takes things a step further by using dynamic routing between capsules for the first time to harness the angle variances between the various views acquired in LF photos. Additionally, CapsField ignores the spurious dimensions and gives higher weights to the more significant features to strengthen its robustness.

Title: Light Field Image Super-Resolution via Mutual Attention Guidance

Authors: Yao Lu and Zijian Wang

Summary:

Progress in light field image super-resolution has been significantly accelerated by deep learning-based techniques. However, the majority of them disregard lining up various light field sub-aperture elements.image prior to aggregation, producing less than ideal super-resolution outcomes. For sub-aperture feature aggregation, we want to provide an effective feature alignment technique. In order to do this, we create a technique for mutual attention for sub-aperture feature alignment and suggest a mutual attention guidance block (MAG).With the help of the centre attention guidance module (CAG) and the surrounding attention guidance module, MAG achieves the mutual attention mechanism between the centre feature and surrounding feature (SAG). To create bidirectional center-view, SAG aligns the refined surrounding-view feature with the original surrounding-view feature after CAG aligns the center-view feature with the surrounding-view feature.

Title: AIFNet: All-in-Focus Image Restoration Network Using a Light Field-Based Dataset

Authors: Bin Chen, Jizhou Li, Lingyan Ruan , and Miu-Ling Lam

Summary:

The performance of the image such as object recognition and image segmentation degrades when images are blur. Image converges out of image plane. This region is called circle of confusion (COC). However, defocus blur is undesirable for most computer vision and image processing tasks. Competitive results have been obtained in addressing single image motion deblurring by using DOF and LFDOF. However, the performance of existing deblurring techniques is still unsatisfactory.

So in order to restore the defocussed image in this paper they have implemented a convolutional neural network(CNN) architecture AIFNet for removing spatially varying defocus blur from a single defocused image for multiple image datasets they have used recognition techniques and Light Field Synthetic aperture. AIFNet consists of three modules: defocus map estimation, deblurring, and domain adaptation. The effects and performance of various network components are extensively evaluated. We also compare our methods with existing solutions using several publicly available image datasets. We leverage the photographic features of light fields. superior deblurring results. We fully analyze the contributions of. our method against existing techniques using different testing datasets. our AIFNet presents state-of-the-art performance. illustrate the benefits of all-in-focus image restoration.

Title: Quality Prediction on Deep Generative Images
Authors: Alan C. Bovik, Dae Yeol Lee, HyunsukKo , and Seunghyun Cho

Summary:

Deep neural networks have been used for several tasks, including the creation of images. For applications like image reduction, generative adversarial networks (GANs) in particular are capable of creating incredibly lifelike images. Similar to normal compression, monitoring and managing the encoding process requires the ability to automatically evaluate the perceptual quality of generated images. On GAN-generated content, traditional image quality algorithms are ineffectual, particularly in textured regions and at high compressions. Here, we provide a new image quality predictor for generative images that is "naturalness"-based. Utilizing a multi-stage parallel boosting technique based on structural similarity features and assessments of statistical similarity, we have developed a novel GAN picture quality predictor. According to our experimental findings, our suggested GAN IQA model provideshigher-qualityy predictions on both generative image datasets and conventional image quality datasets. The SSQP model, which we suggested, was developed utilising two groupings of attributes that stand for structural and statistical similarity. In order to understand the nonlinear relationship between the proposed characteristics and the subjective scores, we additionally used a multi-stage parallel boosting method. We developed a database of generated images with high quality made up of GAN generative images, and we performed an empirical analysis on it. The experimental results show that SSQP performs better than existing FR models on the new database by a wide margin. On three conventional image quality datasets, it also achieved similar prediction accuracies as current DNN-based IQA models.

Title: An Off-Axis Flight Vision Display System Design Using Machine Learning

Authors: Jianlin Zhao, Shan Mao, Zhenbo Ren

Summary:

An advanced optoelectronic tool known as the AFLIGHT simulator can recreate the procedures involved in takeoff and landing as well as the surrounding environment on the ground.for pilots. This study investigates a deep neural network for an off-axis flight visual display system design, focusing primarily on a free-form surface for take-off and landing pilot training. The ZPL-macro programming in the reverse-engineered ZEMAX software is used to propose the surface type and its related expression, together with optical requirements for a flight vision display system. The network has great accuracy in predicting the free-form surface shape from the original structures and parameters. Finally, we explore these data for neural network learning and training. This approach and neural network may provide follow-up research for complex and high-quality optical system design based on artificial intelligence methodology, greatly cut design time, increase design accuracy, and lessen the dependence of optical engineers on experience.

Title: A Deep Learning for Unsupervised Anomaly Localization in Industrial Images

Authors: Chandranath Adak, Shaohua Yan , Xian Tao , Xin Zhang and Xinyi Gong

Summary:

As we know, deep learning-based visual inspection has been highly successful with the help of supervised learning methods. However, in real industrial scenarios, the scarcity of defect samples, the cost of annotation, and the lack of knowledge of defects may render supervised-based methods ineffective. In recent years, unsupervised anomaly localization algorithms have become more widely used in industrial inspection tasks. The researchers in this field by comprehensively surveying recent achievements in unsupervised anomaly localization in industrial images using deep learning. The survey reviews covering different aspects of anomaly localization, mainly covering various concepts, challenges, taxonomies, benchmark datasets, and quantitative performance comparisons of the methods reviewed. In reviewing the achievements to date, this paper provides detailed predictions and analysis of several future research directions. This review provides detailed technical information for researchers interested in industrial anomaly localization and who wish to apply it to the localization of anomalies in other fields.

Title: The Recognition Framework of Deep Kernel Learning for Enclosed Remote Sensing Objects

Novel stock prediction for Indian stock market using ensemble techniques

Authors: DAZHENG FENG, JIE CHEN , LONG SUN , and MENGDAO XING

Summary:

Remote sensing image target recognition is used in various fields, such as ships, tanks, airplanes, and vehicles, which are closed targets. The features of these targets include target outlines that are obvious and target discriminant features that are significantly different from the surrounding environment, and the targets are characterized as small and dense. Therefore, the recognition of these types of targets is a popular topic. We proposed a recognition framework consisting of a remote sensing image target recognition method based on deep saliency kernel learning analysis, which uses a target region extraction method based on the visual saliency mechanism and implements a nonlinear deep kernel learning saliency feature analysis method to realize target extraction and recognition. Experimental results show that a 95.9% recognition rate is achieved for SAR remote sensing target recognition on the public MSTAR data set, a 96% recognition rate on the UC Merced Land Use data set, and an 85% recognition rate on a self-built visible light remote sensing image data set. The recognition framework can be used for video recognition.

Title: Deep Light Field Spatial Super-Resolution Using Heterogenous Imaging

Authors: Gangyi Jiang, Haiyong Xu, Mei Yu, Yeyao Chen, Yo-Sung Ho

Summary:

Light field (LF) imaging expands traditional imaging techniques by simultaneously capturing the intensity and direction information of light rays, and promotes many visual applications. However, owing to the inherent trade-off between the spatial and angular dimensions, LF images acquired by LF cameras usually suffer from low spatial resolution. Many current approaches increase the spatial resolution by exploring the four-dimensional (4D) structure of the LF images, but they have difficulties in recovering fine textures at a large upscaling factor. To address this challenge, this paper proposes a new deep learning-based LF spatial super-resolution method using heterogeneous imaging (LFSSR-HI). The designed heterogeneous imaging system uses an extra high-resolution (HR) traditional camera to capture the abundant spatial information in addition to the LF camera imaging, where the auxiliary information from the HR camera is utilized to super-resolve the LF image. Specifically, an LF feature alignment module is constructed to learn the correspondence between the 4D LF image and the 2D HR image to realize information alignment. Subsequently, a multi-level spatial-angular feature enhancement module is designed to gradually embed the aligned HR information into the rough LF features. Finally, the enhanced LF features are reconstructed into a super-resolved LF image using a simple feature decoder. To improve the flexibility of the proposed method, a pyramid reconstruction strategy is leveraged to generate multi-scale super-resolution results in one forward inference. The experimental results show that the proposed LFSSR-HI method achieves significant advantages over the state-of-the-art methods

in both qualitative and quantitative comparisons. Furthermore, the proposed method preserves more accurate angular consistency.

Title: Unrolling Graph Total Variation for Light Field Image Denoising

Authors: Gene Cheung, Huy Vu , Kazuya Kodama , RinoYoshida , Takayuki Hamamoto

Summary:

A light field (LF) image is composed of multiple sub-aperture images (SAIs) from slightly offset viewpoints. To denoise a noise-corrupted LF image, leveraging recent development in deep algorithm unfolding, we pursue a hybrid graph-model-based / data-driven approach. Specifically, we first connect each pixel in a target patch of an SAI to neighbouring pixels within the patch, and to pixels in co-located "similar" patches in adjacent SAIs. Given graph connectivity, we formulate a maximum a posteriori (MAP) problem using graph total variation (GTV) as signal prior. We then unroll the iterations of a corresponding optimization algorithm into a sequence of neural layers. In each unrolled layer, we learn relevant features per pixel from data using a convolutional neural net (CNN) in a supervised manner, so that edge weights can be computed as functions of feature distances. Each neural layer can be interpreted as a graph low-pass filter for a 4D LF image patch. Experiments show that our proposal outperformed two model-based and two deep-learning-based implementations in numerical and visual comparisons.

Title: Camera -Based Light Emitter Localization Using Correlation of Optical Pilot Sequences

Authors: Ashwin Ashok, Kristin J. Dana , Macro Grusteser , MD Rashed Rahman , Narayan B. Madayam , Shubham Jain and T. V. Sethuraman

Summary:

Visual identification of objects using cameras requires precise detection, localization, and recognition of the objects in the field-of-view. The visual identification problem is very challenging when the objects look identical and features between distinct objects are indistinguishable, even with state-of-the-art computer vision techniques. The problem becomes significantly more challenging when the objects themselves do not carry rich geometric and photometric features, for example, in visual identification and tracking of light emitting diodes (LED) for visible light communication (VLC) applications. In this paper, we present a camera based visual identification solution where objects or regions of interest are tagged with an actively transmitting LED. Motivated by the concept of pilot symbols, typically used for synchronization and channel estimation in radio communication systems, the LED actively transmits unique pilot symbols which are detected by the camera across a series of image frames using our proposed spatio-temporal correlation based algorithm. We setup the visual identity as a problem of localization of the LED on the camera image, which involves identifying and the unique ID corresponding to the LED. In this paper, we present the algorithm and trace-based evaluation of the identification accuracy under real-world conditions including indoor, outdoor, static, and mobile scenarios.

3. REQUIREMENTS

3.1 Software Requirements:

- Windows 10
- Python 3.8
- Google Colab
- Jupyter

3.2 Hardware Requirements:

- Intel core i5 processor
- RAM 8GB

4. PROPOSED DESIGN AND METHODOLOGY

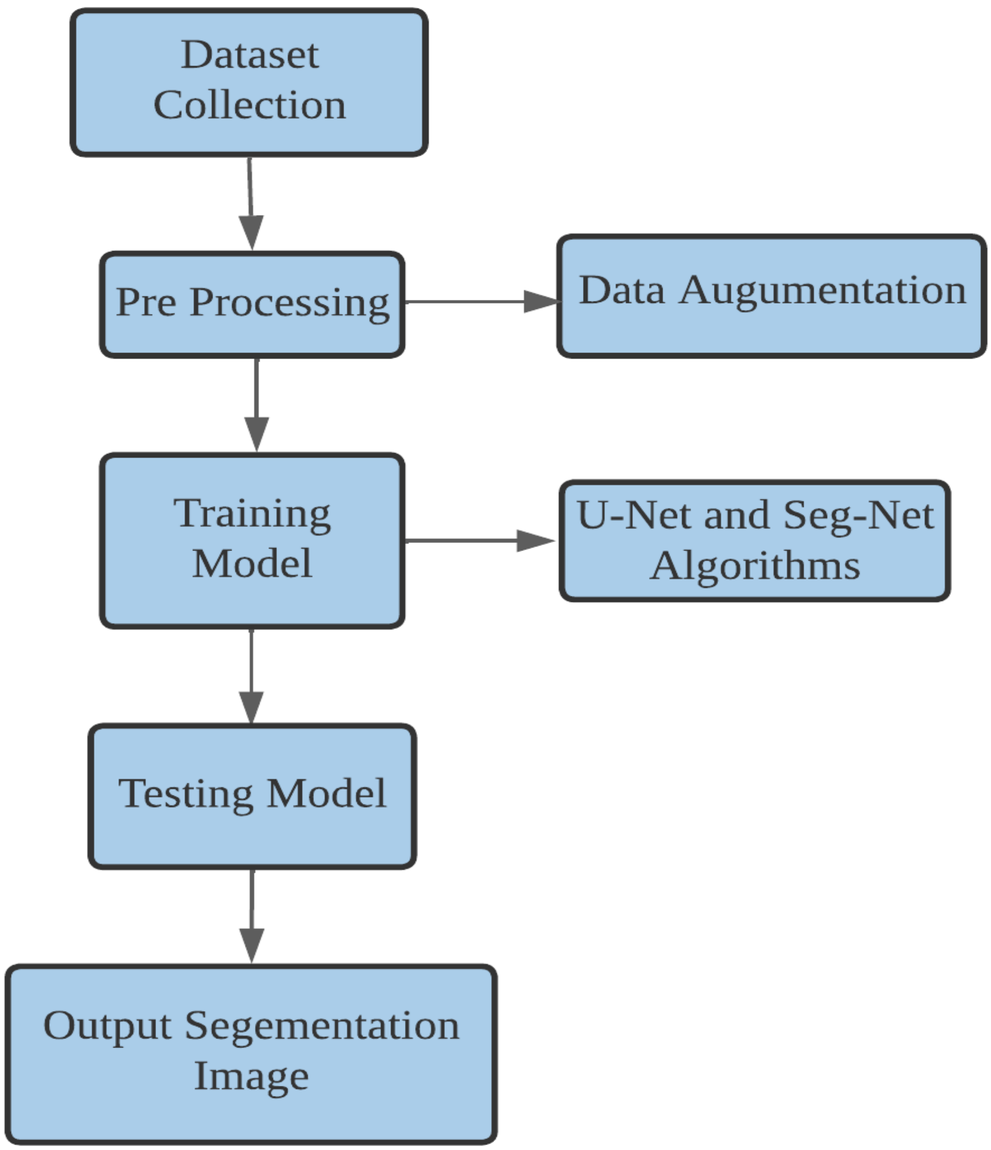


Figure1: Architecture of Implementing the project

4.1 Dataset

Proper dataset is required during the training and the testing phase. The dataset for the experiment is downloaded from the Kaggle which contains 40 different Retinal images. The images are split into training set and testing set. Each having the images classified into – 1st manual, images, mask. The images are in TIF format.

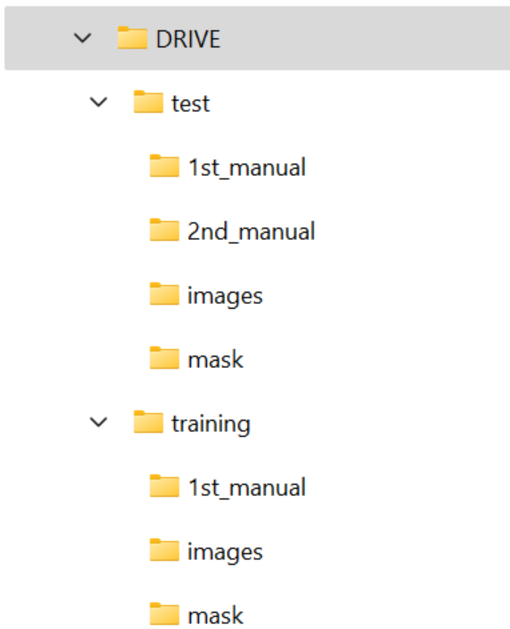


Figure 2: Dataset

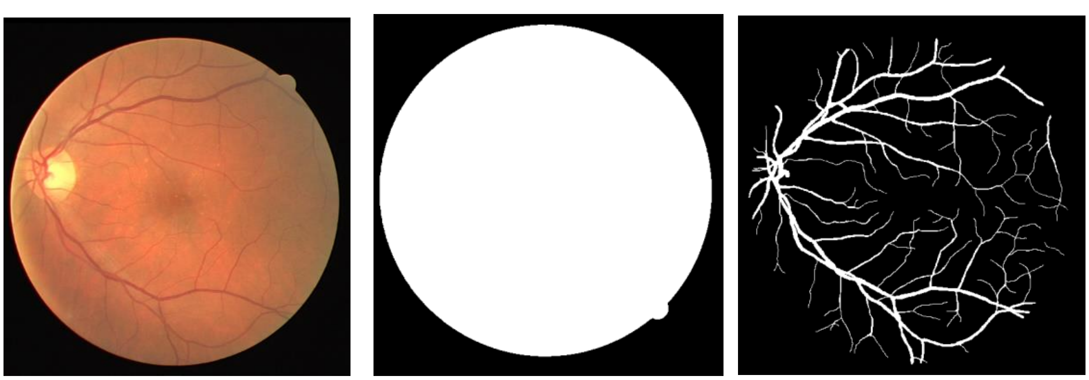


Figure 3: Sample Dataset

4.2 PROPOSED MODELS

4.2.1 U-Net

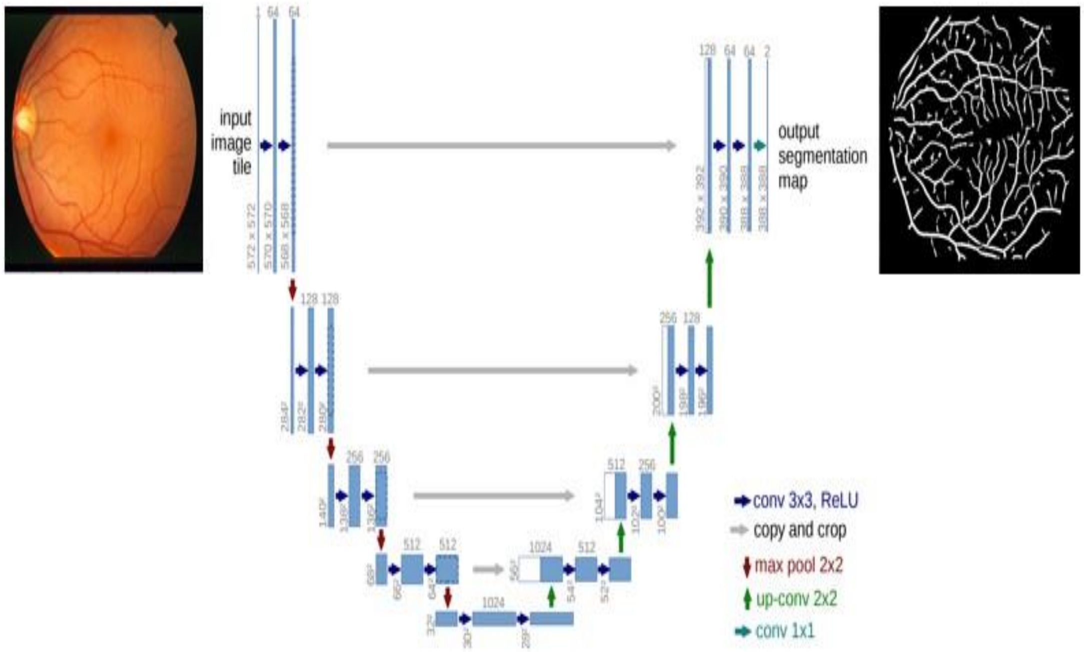


Figure 4: U-Net architecture for retinal image segmentation

U-Net is an extension of Fully Connected Network where it results in accurate segmentations even though the dataset is relatively small. U-Net does not need a huge amount of training data and that makes it ideal for Biomedical image analysis because of the relative scarcity of images in the field of Biomedicine. The U-NET model classifies the pixels of an input image into either a vessel pixel or a non-vessel pixel and therefore extracts the retinal blood vessels from the input image. The original U-NET architecture consists of two main components. The left half is the first component of the model that is called the contraction path or the encoder path while the right half is the second component of the model is called the expansion path or the decoder path. The encoder comprises of convolution layers.

4.2.2 Seg-Net

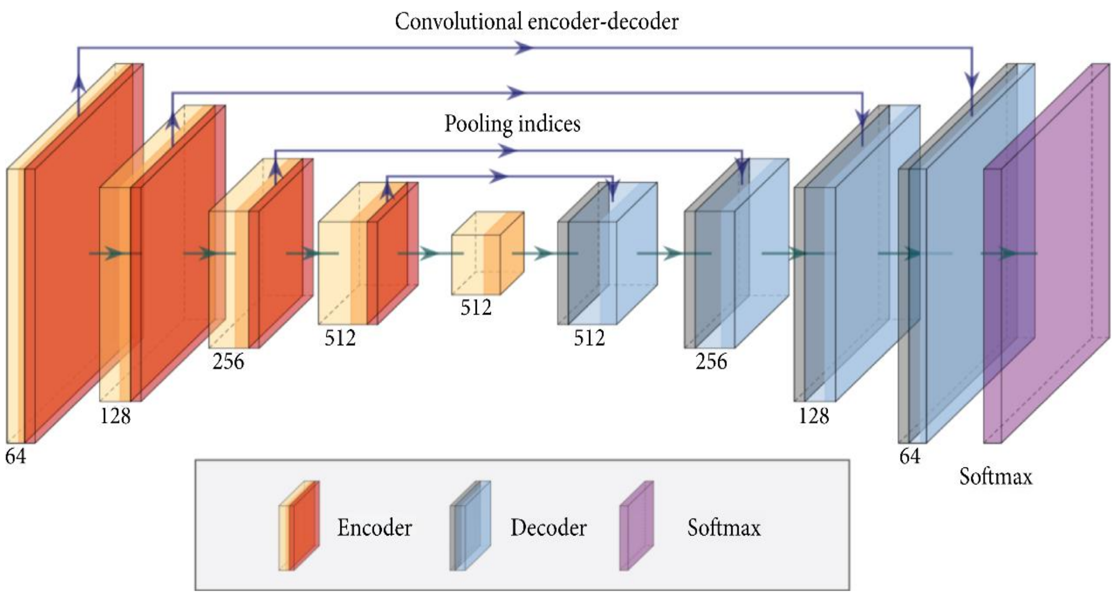


Figure 5: Architecture of Seg Net.

The SegNet architecture includes the encoder layer, decoder layer, and softmax layer. In the encoder layer, there are four convolutions and pooling layers. Each convolution used to extract features is followed by a batch normalization for accelerating learning speed, a rectified linear unit (ReLU), and a maximum pooling operation (step size is 2) for down sampling. In each down sampling, the number of characteristic channels is doubled. In the decoder layer, there are four upsampling layers and four convolutions. After each upsampling is a convolution, and each convolution is followed by batch standardization and ReLU. The last layer of the architecture is the SoftMax layer which classifies each pixel using convolution. This core trainable segmentation engine consists of an encoder network, a corresponding decoder network followed by a pixel-wise classification layer. The architecture of the encoder network is topologically identical to the 13 convolutional layers in the VGG16 network. The role of the decoder network is to map the low-resolution encoder feature maps to full input resolution feature maps for pixel-wise classification. The novelty of SegNet lies is in the manner in which the decoder upsamples its lower resolution input feature map(s). Specifically, the decoder uses pooling indices computed in the max-pooling step of the corresponding encoder to perform non-linear upsampling. This eliminates the need for learning to upsample. The upsampled maps are sparse and are then convolved with trainable filters to produce dense feature maps. Hence, it is designed to be efficient both in terms of memory and computational time during inference.

5. IMPLEMENTATION AND RESULTS

```
from tensorflow import keras

from os import listdir
from os.path import isfile, join
import matplotlib.pyplot as plt
import pandas as pd
from glob import glob
from PIL import Image

import numpy as np
import os
%matplotlib inline

from tqdm import tqdm
import seaborn as sns
import keras.utils as image
import cv2

import glob
import matplotlib.image as mpimg
images = []

for img_path in glob.glob('DRIVE/training/images_tiff//*'):
images.append(mpimg.imread(img_path , format='jpeg'))

plt.figure(figsize=(30,30))
columns = 5
for i, image in enumerate(images):
        plt.subplot(len(images) / columns + 1, columns, i + 1)
        plt.imshow(image)
        plt.xticks([])
        plt.yticks([])

for img_path in glob.glob('DRIVE/training/images_tiff//*'):
        images.append(mpimg.imread(img_path , format='jpeg'))

plt.figure(figsize=(30,30))
columns = 5
for i, image in enumerate(images):
        plt.subplot(len(images) / columns + 1, columns, i + 1)
```

```
    plt.imshow(image)
    plt.xticks([])
    plt.yticks([])

for i in glob.glob('DRIVE/training/images_tiff/*'):
    print(i.split('/')[-1][7:])

import cv2, os
base_path = "DRIVE/training/images_tiff/"
new_path = "DRIVE/trained/images/"
for infile in os.listdir(base_path):
print ("file : " + infile)
read = cv2.imread(base_path + infile)
outfile = infile.split('.')[0] + '.png'
cv2.imwrite(new_path+outfile,read,[int(cv2.IMWRITE_JPEG_QUALITY), 200])

from PIL import Image

# get image
filepath = "DRIVE/trained/images/21_training.png"
img = Image.open(filepath)

# get width and height
width = img.width
height = img.height

# display width and height
print("The height of the image is: ", height)
print("The width of the image is: ", width)

from PIL import Image
import glob

manual_path = glob.glob("DRIVE/training/1st_manual_gif/*")
for gif in manual_path:
img = Image.open(gif)
name = gif.split("\\")[-1]
 name = name.split('.')[0]
img.save("DRIVE/trained/1st_manual/"+name+".png",'png', optimize=True, quality=70)

manual_path = glob.glob("DRIVE/training/mask_gif/*")
for gif in manual_path:
```

```
        img = Image.open(gif)
        name = gif.split("\\")[-1]
        name = name.split('.')[0]
img.save("DRIVE/trained/mask/"+name+".png",'png', optimize=True, quality=70)

dataset_path = 'DRIVE/trained/images/'
train_path = os.path.join(dataset_path, '*')
train_path = glob.glob(train_path)
train_path[4]

from skimage.io import imread
image = imread(train_path[4])

i, (im1) = plt.subplots(1)
i.set_figwidth(15)
im1.imshow(image)

i, (im1, im2, im3, im4) = plt.subplots(1, 4, sharey=True)
i.set_figwidth(20)

im1.imshow(image)  #Original image
im2.imshow(image[:, : , 0]) #Red
im3.imshow(image[:, : , 1]) #Green
im4.imshow(image[:, : , 2]) #Blue
i.suptitle('Original & RGB image channels')

datagen = keras.preprocessing.image.ImageDataGenerator(rotation_range=20, fill_mode='nearest')

dir_It = datagen.flow_from_directory(
    "DRIVE/trained/",
    batch_size=1,
    save_to_dir="DRIVE/images_processed",
    save_prefix="",
    save_format='png',
)

print(len(dir_It))

for _ in range(len(dir_It)):
    img, label = dir_It.next()
    print(img.shape)   #  (1,256,256,3)
    plt.imshow(img[0])
```

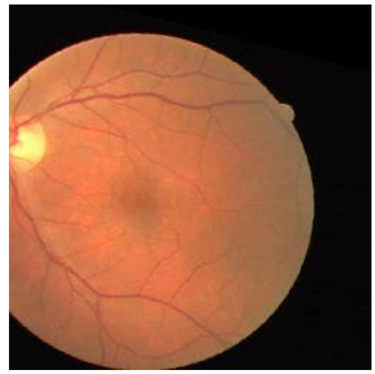
```
    plt.show()
print(len(dir_It))

datagen = keras.preprocessing.image.ImageDataGenerator(
    rescale=1./255,
    rotation_range=180,
    width_shift_range=0.2,
    height_shift_range=0.2,
)

dir_It = datagen.flow_from_directory(
    "DRIVE/trained/",
    batch_size=1,
    save_to_dir="DRIVE/images_processed",
    save_prefix="",
    save_format='png',
)

for _ in range(len(dir_It)):
    img, label = dir_It.next()
    print(img.shape)   #  (1,256,256,3)
    plt.imshow(img[0])
    plt.show()

datagen = keras.preprocessing.image.ImageDataGenerator(brightness_range=[0.5,2.0])

dir_It = datagen.flow_from_directory(
    "DRIVE/trained/",
    batch_size=1,
    save_to_dir="DRIVE/images_processed",
    save_prefix="",
    save_format='png',
)

for _ in range(len(dir_It)):
    img, label = dir_It.next()
    print(img.shape)   #  (1,256,256,3)
    plt.imshow(img[0])
    plt.show()

datagen = keras.preprocessing.image.ImageDataGenerator(featurewise_center =True,
      featurewise_std_normalization = True)
```

```
dir_It = datagen.flow_from_directory(
    "DRIVE/trained/",
    batch_size=1,
    save_to_dir="DRIVE/images_processed",
    save_prefix="",
    save_format='png',
)

for _ in range(len(dir_It)):
    img, label = dir_It.next()
    print(img.shape)   #  (1,256,256,3)
    plt.imshow(img[0])
    plt.show()

data_path ='DRIVE/'

class_data= ['images_processed']
len_class_data = len(class_data)

print(len_class_data)

image_count = {}
train_data = []

      for i , class_data in tqdm(enumerate(class_data)):
    class_folder = os.path.join(data_path,class_data)
    label = class_data
    image_count[class_data] = []

    for path in os.listdir(os.path.join(class_folder)):
        image_count[class_data].append(class_data)
        train_data.append(['{}/{}'.format(class_data, path), i, class_data])
         for key, value in image_count.items():
    print('{0} -> {1}'.format(key, len(value)))
df = pd.DataFrame(train_data, columns=['file', 'id', 'label'])
df.shape
df.head()
df = pd.DataFrame(train_data, columns=['file', 'id', 'label'])
df.shape
df.head()
df['file'][np.random.randint(160)]
```

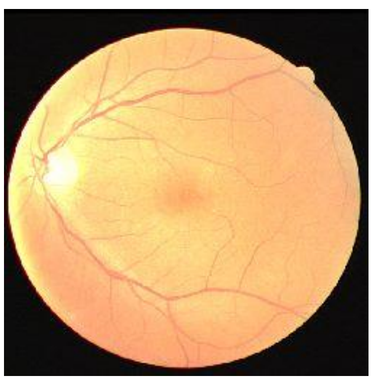
Output

Retina Image

rotation range 20



rotation range 180



brightened image

horizontal flip

Figure 6: Sample Output

6. EXTENSION PLAN (FOR PHASE – II)

- To further increase the images by Data Augumentation
- Implementing the proposed models on the augumented dataset

7. DATASET AND REFERENCES

Dataset - https://www.kaggle.com/datasets/andrewmvd/drive-digital-retinal-images-for-vessel-extraction

Base Paper - LIPENG SI, GUANGYI JIANG, XIUHUA HU, AND BAOLONG LIU, "Retina 3D Perception Reconstruction Algorithm Based on Visual Light Field Image", Oct 2020.

[1] Davi Ribeiro Militani, Demóstenes Zegarra Rodríguez, Juan Casavílca Silva, Lunchakorn Wuttisittikulkij , Muhammad Saadi , Renata Lopes Rosa, and Sattam Al Otaibi ,"Light-Field Imaging Reconstruction Using Deep Learning Enabling Intelligent Autonomous Transportation System", IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 23, NO. 2, FEBRUARY 2022.

[2] Kranthi Kumar Rachavarapu, Kaushik Mitra and Mohit Lamba, "Harnessing Multi-View Perspective of Light Fields for Low-Light Imaging", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 30, 2021.

[3] Hang Su , Tao Yang , Nathan Crombez , Yassine Ruichek, Tomas Krajnik , Xiaofei Chang and Zhi Yan , "Raindrop Removal With Light Field Image Using Image Inpainting", date of current version April 6, 2020.

[4] Alireza Sepas-Moghaddam, Fernando Pereira, Kamal Nasrollahi , Mohammad A. Haque, Paulo Lobato Correia and Thomas B. Moeslund ,"A Double-Deep Spatio-Angular Learning Framework for Light Field-Based Face Recognition", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 30, NO. 12, DECEMBER 2020.

[5] Gangyi Jiang, Mei Yu, Yeyao Chen , Yo-Sung Ho and Zhidi Jiang, "Deep Light Field Super-Resolution Using Frequency Domain Analysis and Semantic Prior", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 24, 2022.

[6] Gangyi Jiang, Mei Yu, Yeyao Chen , Yo-Sung Ho and Zhidi Jiang, "Deep Light Field Super-Resolution Using Frequency Domain Analysis and Semantic Prior", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 24, 2022.

[7] Yao Lu and Zijian Wang, "Light Field Image Super-Resolution via Mutual Attention Guidance", date of publication September 14, 2021.

[8] Bin Chen, Jizhou Li, Lingyan Ruan , and Miu-Ling Lam "AIFNet: All-in-Focus Image Restoration Network Using a Light Field-Based Dataset", IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING, VOL. 7, 2021.

[9] Alan C. Bovik, Dae Yeol Lee, HyunsukKo , and Seunghyun Cho, Fellow, IEEE "Quality Prediction on Deep Generative Images" IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 29, 2020

[10] Jianlin Zhao, Shan Mao, Zhenbo Ren, "An Off-Axis Flight Vision Display System Design Using Machine Learning" IEEE PHOTONICS JOURNAL, VOL. 14, NO. 2, APRIL

[11] Chandranath Adak, Shaohua Yan , Xian Tao , Xin Zhang and Xinyi Gong "Deep Learning for Unsupervised Anomaly Localization in Industrial Images: A Survey", IEEE TRANSACTIONS ON, VOL. 71, 2022.

[12] DAZHENG FENG, JIE CHEN , LONG SUN , and MENGDAO XING "The Recognition Framework of Deep Kernel Learning for Enclosed Remote Sensing Objects", date of current version July 13, 2021.

[13] Gangyi Jiang, Haiyong Xu, Mei Yu, Yeyao Chen, Yo-Sung Ho, "Deep Light Field Spatial Super-Resolution Using Heterogeneous Imaging",2022.

[14] Gene Cheung, Huy Vu , Kazuya Kodama , RinoYoshida , Takayuki Hamamoto "UNROLLING GRAPH TOTAL VARIATION FOR LIGHT FIELD IMAGE DENOISING",2022.

[15] Ashwin Ashok, Kristin J. Dana , Macro Grusteser , MD Rashed Rahman , Narayan B. Madayam , Shubham Jain and T. V. Sethuraman , "Camera-Based Light Emitter Localization Using Correlation of Optical Pilot Sequences ",date of current version March 9, 2022.
ustering, and dimensionalin.