CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Salient Object Detection from Underwater Image

by

Zeeshan Ur Rehman

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Computing
Department of Computer Science

2019

I dedicate this to my mother and father

CERTIFICATE OF APPROVAL

Salient Object Detection from Underwater Image

by

Zeeshan Ur Rehman

(MCS-153014)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Muhammad Nazir	HITEC, Taxila
(b)	Internal Examiner	Dr. Nadeem Anjum	CUST, Islamabad
(c)	Supervisor	Dr. Mohammad Masroor Ahmed	CUST, Islamabad

Dr. Mohammad Masroor Ahmed
Thesis Supervisor
November, 2019

Dr. Nayyer Masood
Head
Dept. of Computer Science
November, 2019

Dr. Muhammad Abdul Qadir
Dean
Faculty of Computing
November, 2019

Author's Declaration

I, **Zeeshan Ur Rehman** hereby state that my MS thesis titled "**Salient Object Detection from Underwater Image**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Zeeshan Ur Rehman)

Registration No: MCS-153014

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Salient Object Detection from Underwater Image**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been dully acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Zeeshan Ur Rehman)

Registration No: MCS-153014

Acknowledgements

I would first like to thank my thesis supervisor, **Dr. Mohammad Masroor Ahmed**. The door to Dr. Masroor Ahmed office was always open whenever I ran into a trouble spot or had a question about my research or writing. The whole work proposed in this thesis cannot be completed without the professional instructions and support from him. I greatly appreciate his patience and kindness.

My colleague, Khawar and Shabbar always along me with their moral support and technical guidance. This accomplishment would not have been possible without them. Thank you.

Nobody has been more important to me in the pursuit of this project than family members. I would like to thank my mother, whose love and guidance are with me in whatever I pursue. She is the ultimate role models. Most importantly, I wish to thank my loving and supportive wife, Taibba, and my wonderful children, who provide unending inspiration.

(Zeeshan Ur Rehman)
Registration No: MCS-153014

Abstract

Precise analysis of underwater images is a challenging area of research. The challenge is mainly due to the loss of image visual quality. The presence of water influences tonal quality which reduces the contrast. The reduction blurs boundaries and restricts fine demarcation of neighboring regions. This situation makes precise segmentation of underwater images a challenging task. The present study proposes a method for obtaining accurate segmentation results. This method uses gray-world assumption and CLAHE for enhancing the existing visual quality. The enhanced image is processed by saliency identification approach which uses clustering method. Clustering, besides saliency division also separates foreground from background. The proposed approach was tested on 700 plus benchmark images and was compared with two state of the art methods. The proposed approach produced convincing results with an accuracy rate of 94%.

Contents

List of Figures

List of Tables

Abbreviations

ACM	Active Contour Model
AHE	Adaptive Histogram Equalization
BDM	Background Driven Measure
CLAHE	Contrast Limited Adaptive Histogram Equalization
FCM	Fuzzy C-Mean
HE	Histogram Equalization
HVS	Human Visual System
KM	K-Mean
SOD	Salient Object Detection
SODM	Salient Object Detection Measure

Chapter 1

Introduction

Image segmentation is the fundamental tool for the investigation of the marine life. Nonetheless in practice, the light attenuate exponentially while travelling through water and this results in severe degradation of the image quality. Hence the extraction of information via segmentation from an underwater image becomes rather challenging task.

Object Detection is the process to understanding the image fundamentals. In recent arena Salient Object Detection (SOD) is an interesting investigating and widely used domain in computer vision and image processing. The Salient object in an images is defined as any visually distinctive and semantically meaningful object. It is certainly possible that an image does not contain any salient object and an image way very well be full of many such salient objects. The goal SOD is to detect and extract the object from the background in an image. An important term in this field is **Perfect Detection** also known as Ground Truth (GT) which is presenting the salient object in white color and rest of image as black.

Figure 1.1 All the images in the first column are from the MRSA10K(formally named as THUS10000) database (containing 10,000 images) and their corresponding SOD results.The MSRA10K benchmark data set provides the ground truth annotation for all 10,000 MSRA images at a per-pixel bases. The total size of the database is 181MB. All images has clearly distinguishable salient object and the all object areas are also labeled with pixel-wise ground-truth. [1]

When human seeing an image, they will focus on the important and informative region of an image. In order to focus on the foreground object without any background deviation, we are going to the SOD. The fundamental responsibility of SOD process is to locate an interesting and prominent region of the image from the background cues. These cues are calculated using the image features like contrast, color, spatial and texture etc. Images saliency detection has been studied for many years. The estimation of salient features in an image is a valuable resource in image and signal processing. Although current methods tend to offer considerable variation in their methodology to the solution for this problem. Many recent disciplines such as computer vision, neuroscience, Robotics and graphic fields are using SOD model to locate and identify the interesting and important feature from the images. detection[2]. Recently, visual attention modeling has become a vital task and has been adopted widely in many applications such as image retrieval [3], [4], smart video presentation [5], image compression [6], object recognition [7], image re-targeting [8], sensation enhancement[9].

The expected outcome of an segmented image is an object or meaningful parts. [10]. Object or meaningful part is not clear, it can be a standalone object or part of another object. If the object is not defined clearly it could makes the segmenting process a difficult and challenging task. The work in human perception has opened doors and made some useful guidelines in order to develop a segmentation algorithm.

The question of object representation in computers is another challenging task and it further adds to the already challenging process of object segmentation. Human Visual System (HVS) works at a whole other level. When a human sees an object it perceive it as a whole not in a layered manner. In a computer an image is represented on the lower level features of the object, like its color, the texture of it, convexity, the curvature etc. These lower level features have their own local properties, it it makes the capturing of global object information.

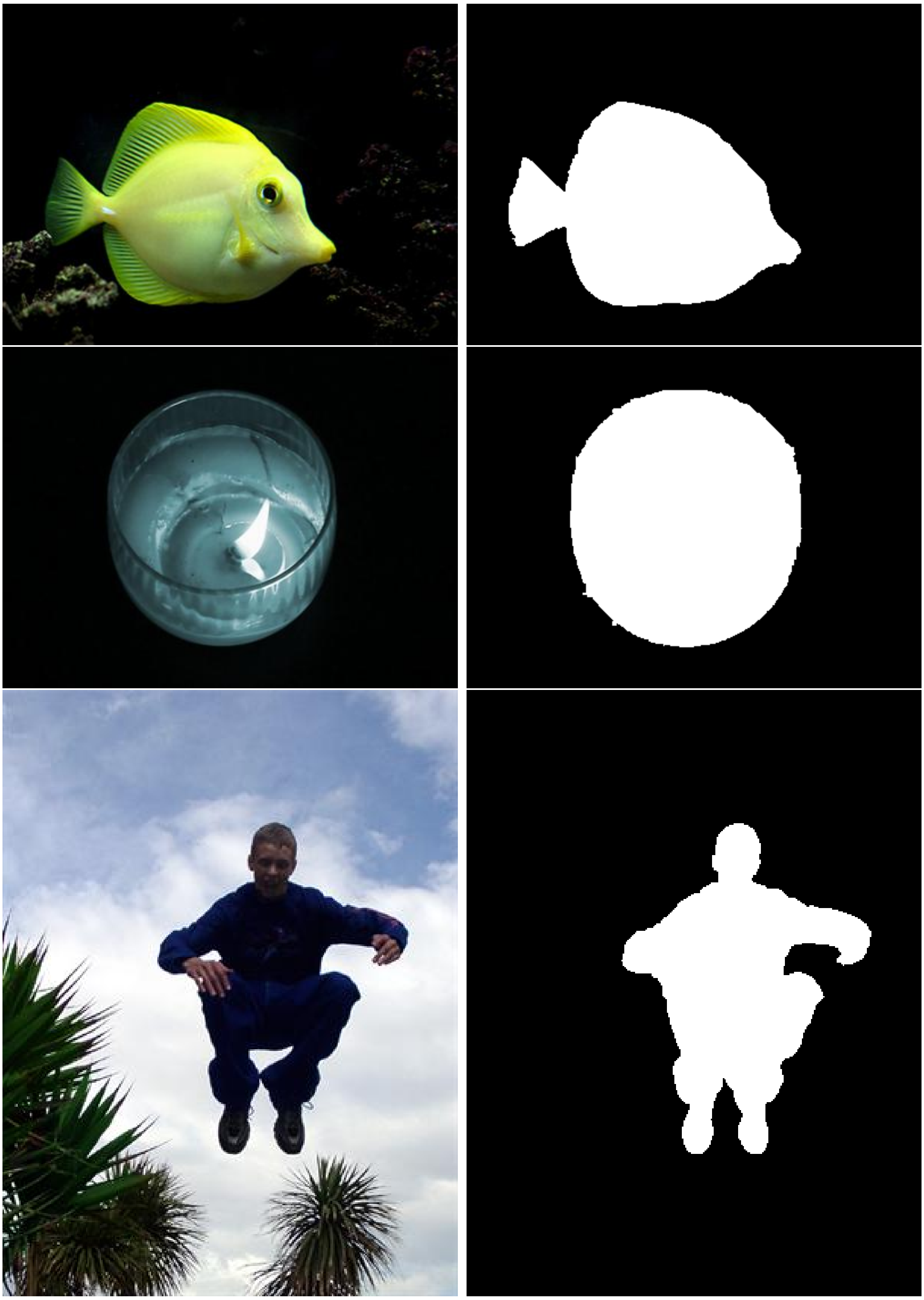

FIGURE 1.1: Sample images with respective SOD Masks

Many saliency detection approaches are used, like[1],[11], [12], [13], [14], here in all these methods they target the lower level priors and cues. In reality, even with these lower level features may not be of much use in promoting the salient foregrounds object to become more prominent from a background which is rather confusing due to lower contrast images as is the case for the underwater images. So its safe to assume that the conventional salient object detection solutions may not be readily applicable to underwater environment. In addition to saliency detection the conventional segmentation procedures may also not work very well.

The light attenuation in water makes it loss the energy rapidly which results in color depletion. Also the presence of suspended particles in water both organic and inorganic adds to light beam randomly reflecting and deflecting before entering the camera sensor and results in an image with lower contrast. Due to these problems of color shifting and degraded contrast make the underwater image segmentation process more challenging and difficut.

The amount of light with in water is always less than the amount of light over the surface of water. Therefore images obtained under water generally have low visual quality. The scarcity of light under water is usually because of two unavoidable facts. One, the light under water is loses its true intensity, and second the chances for scattering of light within water is quite high. The immediate impact of this insufficient amount of light is the color distortion and illumination of the under water scene visibility.

The light scattering and absorption effect the visual appearance of underwater images. These phenomena are caused by turbid medium, light travelling behavior in water, a selective attenuation to different light color elements depending upon their wavelengths, water depth and scene distance from camera. Underwater scene color change is caused by different light components wavelength when travelling in water medium. Also the image contrast degradation is because of light reflection and refraction by suspending particles in water.

In the computer vision applications, one of the more important ways of carrying the information is by interpreting the images. There could be thousands of uses for examining the images and interpreting them, from steering of autonomous ships and submarines to diagnosis of cancer cells, to recolonization of an airport by a drone from far far away. So in essence the image segment ion solutions provides a way to interpret the images and extracting more useful information with could become input for some other process. So the image segmentation provides its users with a tool-set to extract useful information from an image and interpret the extracted information. The aforementioned process is known as image analysis or image survey. These Image segmentation algorithms are routinely being used for image classification and identification in numerous fields like agriculture, forensics and off course medical.

Another area where image segmentation has a prominent role is underwater navigation and automatic target acquisition and recognition. Image segmentation is also extensively employed for the examination of marine life. The main difficulties here lays in the fact the images produces underwater lacks color intensity and have low contrast which in return makes the image segmentation algorithms less efficient due to less information available in the image. Usually the salient object are present in the foreground of an image and the salience detection process is to visually highlight the salient features of the image in observation. Also the area of video and image processing routinely use the saliency detection for a whole range of apps. like the adaptive image compression, content aware image re-sizing and image retrival etc.

Our main focus will be on foreground object extraction by ignoring background and also saliency segmentation which in turn is built upon saliency detection. Saliency detection process means to make the identification of more prominent features in the foreground of an image and making this identification in a fast and accurate manner.

So work steps are: segment an image, identify foreground prominent objects, ignore the background and extract these identified foreground objects from the image with more accuracy and in efficient manner. The main focus of this work is underwater environment and images taken in this environment.

1.1 Background

1.1.1 Characteristics of Underwater Images

Unlike conventional imaging taken above sea in open air, underwater photography shows a strong dominance of bluish and greenish colors. On the other hand, the strong attenuation of light in the water with respect to the air and a greater diffusion of the incident light have the consequence of considerably reducing the visibility. Thus, objects at distant distances from the acquisition system or the observer but also at medium distances, or even relatively short in some cases, are hardly visible and poorly contrasted with respect to their environment. In addition, in the presence of particles suspended in water (sand, plankton, algae, etc.), the incident light is reflected by these particles and forms a kind of *in-homogeneous* mist that adds to the scene observed. This turbidity of the water, most often white, also affects the visibility but also the colour dynamics of the objects contained in the image by tarnishing or veiling them. On the other hand, the formation of an underwater image is highly dependent on the nature of the water in which it was acquired. Natural waters can have very varied constitutions in terms of plants or minerals dissolved or suspended in water. The behaviour of the propagation of light in such a medium is strongly governed by this factor.

1.1.2 Underwater Image Processing

When dealing with underwater image processing, water medium properties and light travelling in water are two important related domains. The knowledge of these domains is necessary in order to process underwater image for any task. [15]. Water medium is different from air due to its physical characteristics and generates light degradation effects in images as compare to taken in air. The scene visual quality is affected with poor contrast and makes image hazy and light attenuation factor in water is the reason for hazy images. The light propagation behaviour in water, the light attenuates and vanishes at approximately 20 meters

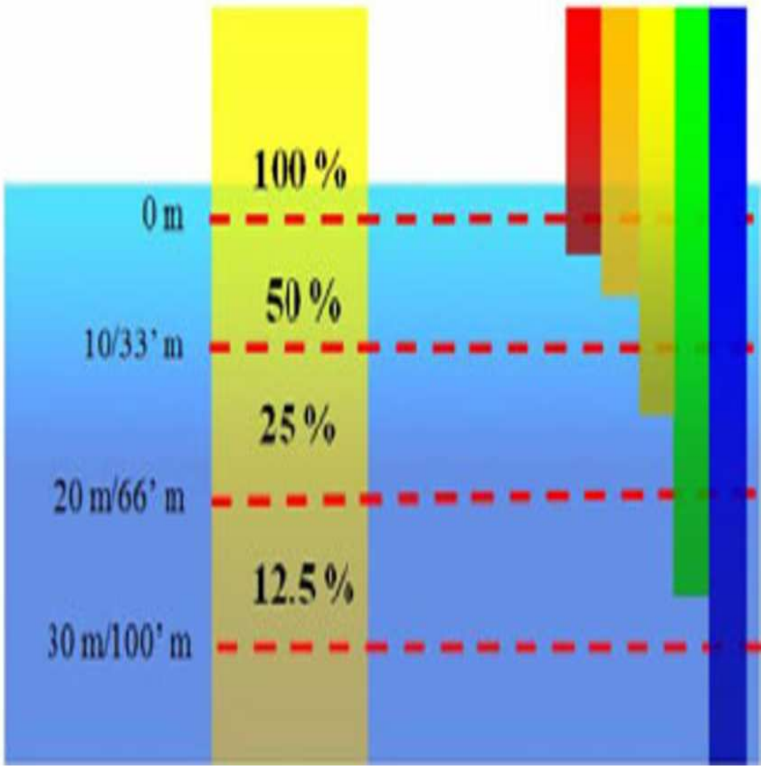

FIGURE 1.2: Light components wavelengths Presentation Model Under Water; Blue light travels larger distance due to shorter wavelength and Red Component of light travels shortest distance due to larger wavelength : image from [16]

distance when water is clear. On the other hand, in turbid water, the light travels less then 5 meters distance show in Figure 1.2. Light attenuation is due to decrease in light intensity (Absorption)and change in reflection path (scattering).

Due to light attenuation, in clear water the visibility goes off after 20m at about twenty meters and in turbid water the visibility vanishes at about 5 meter distance as show in Figure 1.2. Absorption (defuse the light energy) and scattering (alters the light reflection path) are major cases of light attenuation process.

The scattering and absorption effects due to light attenuation process, make the underwater image quality lower and underwater image visual quality is suffered in water environment. Light deviation from its path from an object to camera is called forward scattering and generates blurring effect on an image. In contrast, some part of light reflection from middle without reaching the object is called backward scattering generally limits the contrast of the images.

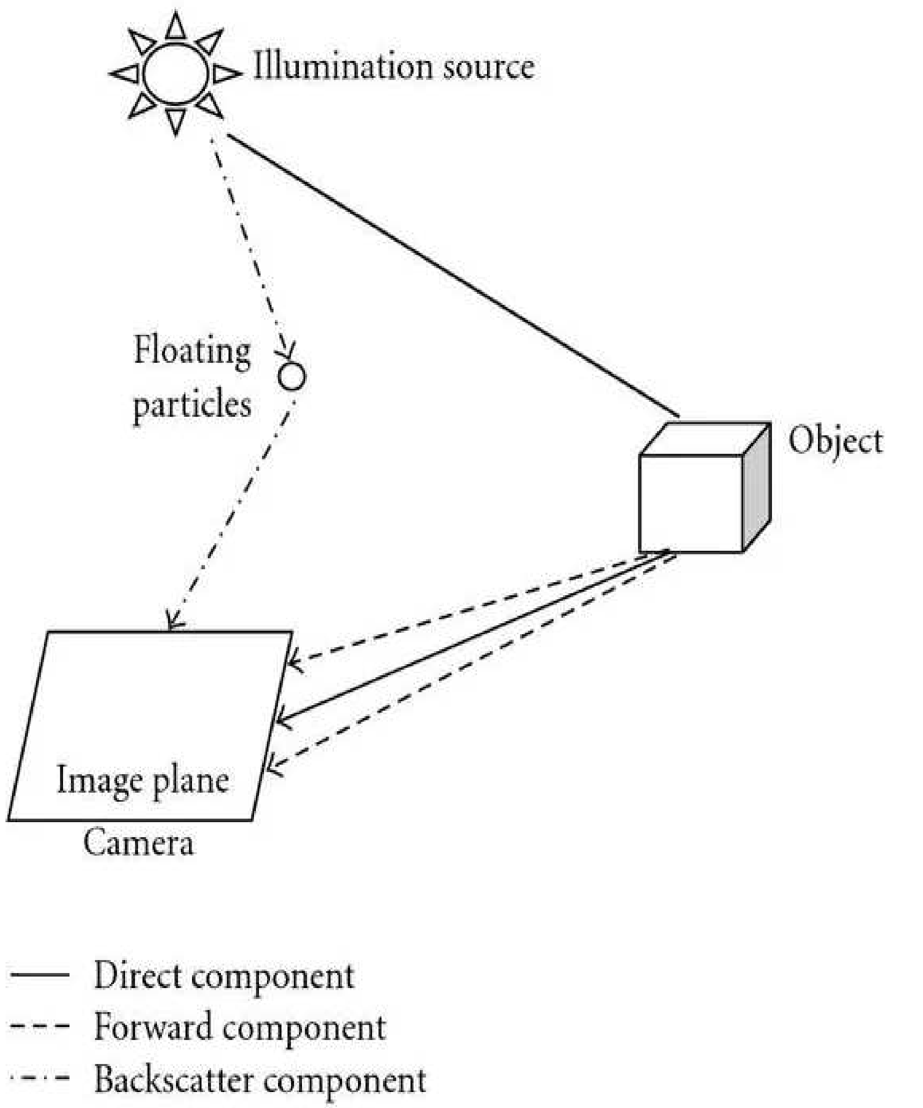

FIGURE 1.3: Underwater optical model with it's three components: Direct component - display as solid line, forward scatter component - display as dashed line and backward scatter component - display as dash-dot line : image from [15]

The light intensity get lower as light travels in water and the light color components gradually disappears depending upon their wavelengths. Due the shortest wavelength, the blue color penetrated underwater more as compare to other components of the light, making the underwater images bluish due to blue color light component dominance. The underwater images suffer from poor quality and suffer from these problems:un-even lighting, poor visibility, low contrast,blurring,disappearing of color components (result into greenish and bluish appearance) and noise. Due to these problems , any computer vision application targeting underwater images needs to rectify one or more issues from these.

There are two different strategies to process an image. one is image restoration and other is image enhancement. The details are:

1. In image restoration, a degradation model is suggested to recover the poorly constructed image quality. The original image is required to recover the quality of an image. These methods are reliable but they need many environment properties values as parameters for their model. These parameters can be light attenuation and water turbidity diffusion coefficients. These parameters vary time to time and have variable values. And underwater depth of scene is also an important parameters and is required for this inverse modeling to restore the image quality.

2. No prior information is required for image enhancement method. This technique uses qualitative subjective criteria in order to enhance image quality. No physical model is required to produce visually pleasant images from the original images. This kind of schemes are not very complex and simpler in implementation. These are relatively have better performance as well.

1.1.2.1 Underwater Light Travelling

Light propagation properties in the water are discussed in this section. When light travels in water medium, two phenomenon are produced, absorption and scattering. Intensity of light goes low when light travels in the any medium and this whole process is called absorption. The light intensity is linearly dependent upon the medium's index of refraction. The index of refraction for water is 1.33 and for air, refractive index has value 1. The deflection of light from a straight line path when propagating, is called scattering. Due to the turbidity nature of water, the suspending particles in water is the root cause of deflection.

Hence, the amount of light with in water is always less than the amount of light over the surface of water. Therefore images obtained under water generally have

Energy Loss of light while travelling in ocean								
	Violet		Blue-green			Yellow	Orange	Red
λ (μm)	0.3	0.45	0.45	0.51	0.53	0.59	0.62	0.7
open seas clearest	17%	3%	2%	3%	5%	8.5%	30%	41%
open sea turbid waters	57%	16%	11%	10%	13%	19%	36%	55%
near coastal sea, average		64%	38%	30%	29%	31%	46%	75%

FIGURE 1.4: Energy Loss of light while travelling in ocean

low visual quality. The scarcity of light under water is usually because of two unavoidable facts. One, the light under water is loses its true intensity, and second the chances for scattering of light within water is quite high.

The relationship between decay of light intensity and light travelling medium is described by Lambert-Beer empirical law as: the decay of light intensity is related to the properties of the material (through which the light is travelling) via an exponential dependence.

The irradiance E at position r can be modeled as:

$$E_d = E_o e^{-c_\lambda r} \tag{1.1}$$

where c is the total attenuation coefficient of the medium. Light decay due to the absorption and scattering per unit length of travel is measured by this coefficient. Subscript denotes that c depends on the wavelength. This equation also explains

TABLE 1.1: Different water Attenuation Coefficients

Deep Ocean Water	$0.05m^{-1}$
Coastal Water	$0.2m^{-1}$
Bay water	$0.33m^{-1}$

why white balance may only be used to correct registered colors if all registered points are approximately equidistant to the camera. In this case the shift in color is similar for all the points and therefore white balance may succeed. However, it is important to remember, that applying white balance to the image will not restore true colors - it will make the image look more naturally, but it is based on some general assumption (e.g. so called grey world assumption) which may not be true for a given picture a therefore one should not expect to get an accurate color reconstruction.

1.1.2.2 Image Restoration

Image restoration is process which operates on a poor quality image and estimates the clear, original image. Image restoration operation tries to recover the original image f(x,y) from the noisy image g(x,y) using available explicit knowledge about the degradation function h(x,y) and the noise properties n(x,y)

$$g(x, y) = f(x, y) * h(x, y) + n(x, y) \tag{1.2}$$

where * expresses the convolution operator. The degradation function comprising of the system response from the imaging system itself and the effects of water medium.

1.1.2.3 Color Correcting and Contrast Enhancement

For these methods no prior information of the environment is required. Image enhancement techniques are usually less complex and these enhancement techniques take less execution time as compare to image restoration techniques. The color channels from propagating light disappear as light propagates in water. The disappearing of color channel from travelling light in water depend on their wavelength. Because red channel has longer wavelength it disappears first of all at the depth of 3m. The orange channel disappears at the depth of 5m. Yellow color mostly disappears at the depth of 10m. In the last, green and purple disappear at further depth. The blue component have shortest wavelength and hence disappears after travelling the longest distance in the water. Due to their wavelengths the blue and green colors are more dominant in underwater images. The light source variations also plays its role to affect the color. Overall the underwater images are suffered from a strong and non uniform color cast.

1.1.3 Image Segmentation

To analyze the image and extract useful information from the image, a processing technique called Image segmentation is used. Segmentation is process to to partition an image into several different regions. These regions are classified based on common pixel characteristics. The image pixel characteristics can be color, intensity , gray level, texture etc. The ultimate objective of the segmentation steps are to yield as maximum as possible information in the region of interest in an image, which helps to locate an object in image scene with more accuracy. The main objective of image segmentation is to accurately locate the foreground objects and identify the background region in an image.

If I represents an image, then the partitioning of image I into into smaller parts $I_1, I_2, I_3......I_n$ is called image segmentation and represents as :

$$I = I_1 \cup I_2 \cup I_3....... \cup I_n \tag{1.3}$$

There are two main general groups of Image segmentation techniques:

1. Image Segmentation Techniques based on Layers
2. Image Segmentation Techniques based on Blocks

Based on Blocks, image segmentation is mostly used and use various features and information found in an image. These feature and information can be color information, pixel information or texture information.

Further based on discontinuity and similarity, the block based image segmentation is generalized into three main categories :

- Region Based (Discontinuities)
- Edge Based (Similarity)
- Hybrid

The saliency detection process is challenging because it tries to performs the same task as of Human Visual System(HVS) and HVS is a very complex system to replicate. It is widely acknowledged that the visual extent of an object extends beyond the object itself[17]. Based on HVS, the saliency detection models can be categorized as:

1. Fixation Prediction Models - Human Attention Prior
2. Salient object detection models
3. Convolutional Neural Networks (CNNs) Based Models

Fixation Prediction (FP) tries to predict the fixation points similar to HVS.This aims to find the fixation locations as human viewers would focus on at first glance. Contrary to SOD models, the fixation prediction models normally establish a prediction where humans look at first glance.

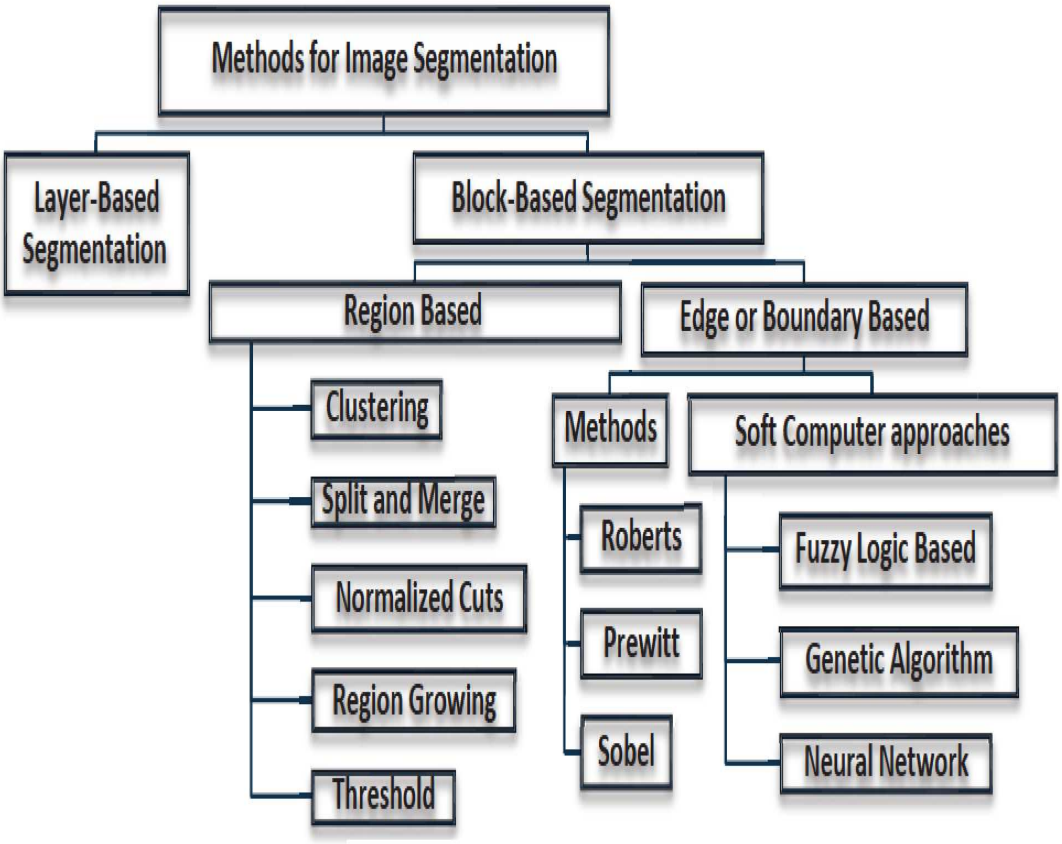

Figure 1.5: Image Segmentation Methods

Different from FP, the main objective of SOD models are to locate and highlight the image most prominent details and then use some segmentation technique(s) to segment this prominent information. In more recent efforts, SOD models based on deep learning, capture more attention and made some substantial improvements.

In recent times CNNs attracted great attentions in many disciplines due to accuracy of results, CNNs based models are providing. In object detection CNN based models also gains success and is widely using in this domain.

Ability to detect prominent features from scenes by any vision model plays an important role to study and analysis the filtered out information. This Visual saliency mechanism serves as a identifier to filter out the irrelevant details and also extracts only the required features and details with respect to image context.

In recent times, in many computer vision applications are using SOD and it provides fast solutions to several complex processes. The SOD process work is compromised of two steps. The first step locates and detects the most salient details in a scene. The second step segments the whole extent of details(detected in first step) from the scene. The end result of this processing is, usually a map where each pixel intensity represents either that pixel belonging to the salient object or not.

With slight difference the Salient Object Detection is treated same as segmentation problem . By ignoring background details, the SOD models segment only the salient foreground features. On the other hand, the image is partitioned into regions based on pixel properties in the traditional segmentation algorithms.

To predict salient locations, some early methods were mainly based on local contrast and global contrast with the assumption that salient regions in each image should be distinguish from others objects in surroundings or from the entire image as a whole. Later on, new approaches started to use other factors as well like background cues, high-level cues and depth information . In recent time, progress in deep neural networks, attracts it for saliency detection. Remarkable results and performance is achieved by using neural networks.

The importance of saliency detection methods are in many fields and their applications are: object detection and recognition, image compression, video monitoring and analysis etc.

One would expect a segmentation algorithm to decompose an image into the objects or semantic/meaningful parts.[10]

1.1.4 Salient Object Detection(SOD)

In Computer Vision, SOD or salient object segmentation(SOS) terms are used interchangeably to detect prominent features of an image. The Salient Object Detection process compromises of two steps[18]:

1. Detecting the most prominent foreground feature
2. Accurately segment the detected feature.

Generally, to detect salient object precisely, the criteria is defined as:

1. Accurate detection: The precision should be high and recall should be low
2. High Resolution: to accurately locate salient objects and retain original image information, the saliency maps should have high or full resolution.
3. Computationally Efficient: the SOD methods and models have low execution time and should detect salient regions in minimum time.

Based of prior knowledge, SOD is categorized into two kind of approaches:

1. Bottom Up Approach
2. Top Down Approach

A supervised learning is required for top down approach. The bottom up techniques are based on prior calculations and calculate these priors based on image contrast, spatial distance, color and textures to generate the saliency maps.[19]

1.2 Motivation

Underwater Image segmentation methods usually include exploring the underwater optical model and compensating the bad effects caused by water and particles while segmenting and exploring the image features. The information extraction from underwater images is a crucial and plays a decisive role in ocean engineering and scientific underwater research, such as monitoring sea life, ocean rescue and accessing the geological environment. The light absorption and scattering limit the visibility of the underwater objects when they are captured. Many Salient

Object Detection algorithms which perform well and produce good results for images captured in a normal environment, reduce their accuracy in an underwater environment substantially. On the other hand, we have the number of methods which are widely using in single underwater image enhancement. So, to detect prominent features from a single image without any parameter and prior information, the aim is to use a state of the art image enhancement technique before performing cluster-based image segmentation.

1.3 Objective

The target of this work is developing an improved SOD model for underwater images to identify prominent features and extract these foreground objects in order to analyze and investigate ocean life.

The task includes the following objective:

- Improving underwater image quality in order to segment underwater images with accuracy and clear edges.
- Studying and analyzing the principles and mechanisms of underwater objects which suffer from light absorption and scattering, and exploring the underwater issues,
- Data analysis of underwater segmented images and Salient Object Detection
- Performance comparison of proposed solution with other underwater image segmentation and Salient Object Detection techniques and solutions
- Analyzing the advantages and limitations after data analysis

Chapter 2

Literature Review

2.1 Introduction

This chapter presents a comprehensive review of literature focusing on various techniques that are used to enhance visual quality of under water image. Also focus is to present a review of models used to detect salinet objects from a scene. There is a wide variety of technologies as well as techniques and models are used for underwater image restoration and locate the prominent foreground features of an image.

Identifying, locating and extracting prominent features from real world environments is referred as salient object detection. In recent times salient object detection has become more interesting and attracted domain in computer vision. Number of models are presented and experimented and these models are applied in many application in different domains. Still a deep learning of this domain is lacking.

Human visual ability is capable to isolate distinctive items, so called salient objects, from scene regions quickly and with less efforts at first glance. These isolated regions are then analyzed and processed to extract more finer details for the extraction information from the image.

The acquiring of information and understating of real wold scenes is studying by cognitive scientists from last many years and now in recent times this study is adopted with much interest by computer vision experts as well. The main reason to adopt this in computer vision is because it locates and isolate the objects or regions from a scene, a very handy tool to understand a scene which itself a very complex computer vision problem.

2.1.1 Theory Behind Salient Object Detection (SOD)

SOD or salient object segmentation is generally considered in computer vision as a two stage process:

1. Identifying the most prominent regions and objects
2. Extracting the identified objects and regions using segmentation accurately

In first stage this is no limitation on objects, more then one object can be also the detected. However most of present proposed methods, target to detect most prominent object or region, but their final saliency maps can be used to detect and locate other salient regions in a scene.

The second stage of SOD can be categorized as the classic segmentation problems in computer vision. The only differentiation is, the SOD process accuracy is measured based on the most salient object or region.

Generally , to decide about the accuracy of the detected salient object good or not, the following three criteria is used to measure its accuracy:

1. Accurate detection: The precision should be high (missing real salient regions should be low) and recall (falsely marking the background should be low) should be low
2. High Resolution: to accurately locate salient objects and retain original image information, the saliency maps should have high or full resolution.

3. Computationally Efficient: the SOD methods and models have low execution time and should detect salient regions in minimum time. The model should detect the salient region in minimum time

We can use this criteria to determine the accuracy of salient object detection model and based on this accuracy, these models can be compared with each others.

2.1.2 Locating SOD

SOD models normally target to locate the most prominent regions only in an image and ignoring the background of the image and segment the whole extent of detected regions.

On the other hand, Fixation prediction models, normally target to define a prediction same like human visual system. The HVS typically predict a minimum numbers fixation points

The final output of both SOD and Fixation Prediction is continuous-valued saliency map.The higher value in the final saliency map describes as the corresponding image pixel is looked more at first glance. Both these models are used to locate the salient regions.

There is a strong relationship between SOD and fixation points. It is also observed that humans seeing results match with each other when they have asked to locate the most prominent object in an image [20] Fig. 2.1.

The results of different models are show in Figure 2.2. We can observer the different between different produced outputs.

2.1.3 SOD History

Itti et al. [13] has presented initial saliency models. This presentation opened the new horizons across multiple fields like computer vision, cognitive psychology and neuroscience.

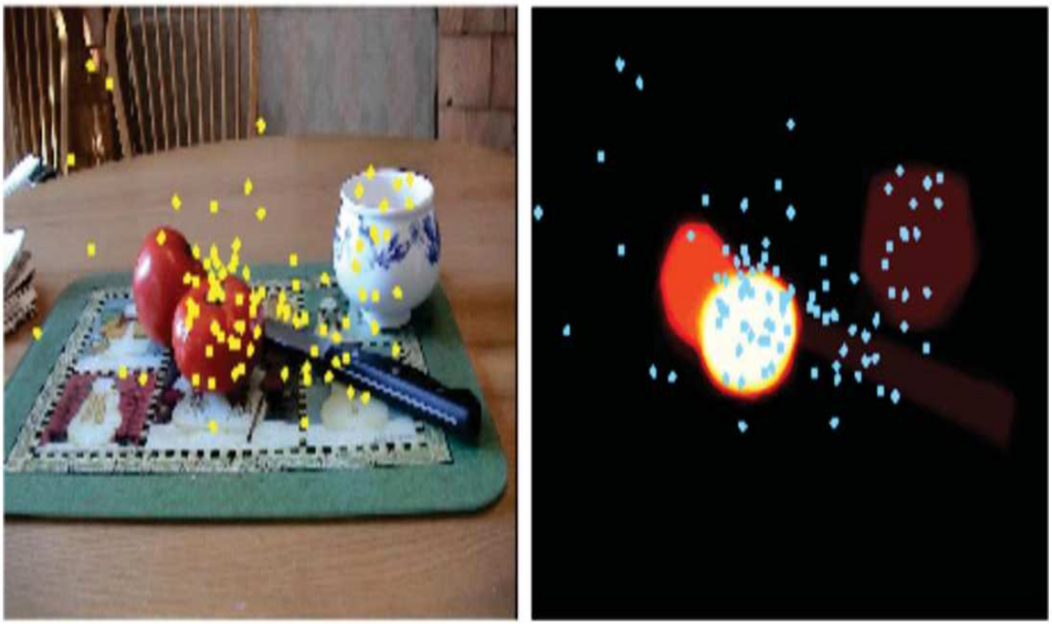

FIGURE 2.1: Image from Borji et al. s experiment [20] with detected salient regions. The dots are depicting a 3-second fixations points.

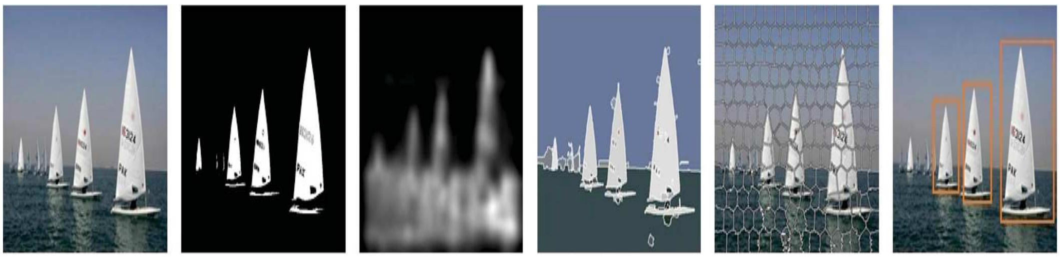

FIGURE 2.2: The Final saliency outputs - outcome of different object detection models. Left to right: input image, SOD [2], fixation prediction model [13], region based image segmentation [21], superpixel based image segmentation [22], and object proposals (true positives) [1]

Another wave is generated and considered as second arena, when salient object detection problem is presented as binary segmentation problem.

The third wave is generated in recent years, when convolutional neural networks (CNNs) become popular and also used for salient object detection.

Although SOD and segmentation methods have made great improvements in last few years, an efficient and robust SOD model that can generate high quality results for nearly all images is still lacking. This particular work is focused on models presented in second wave.

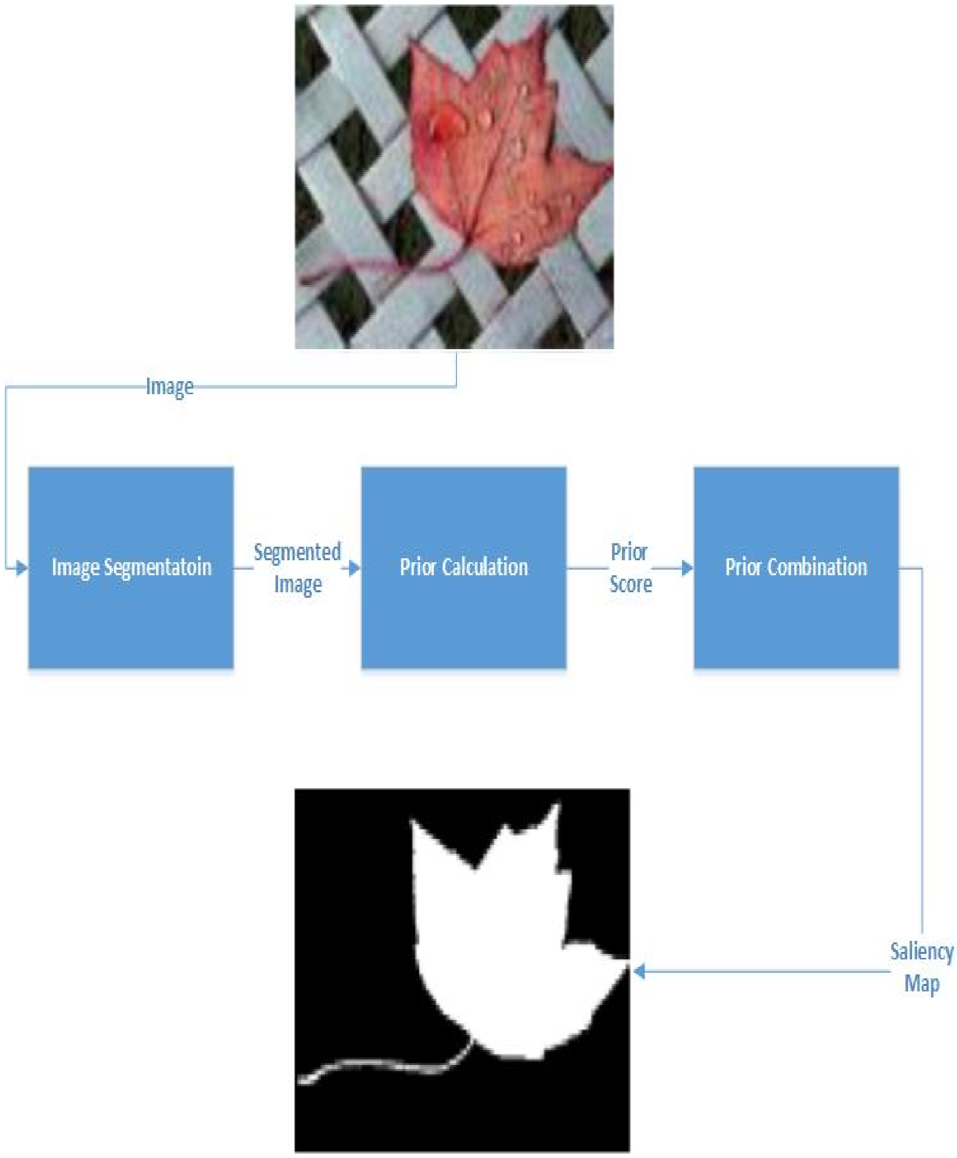

FIGURE 2.3: Salient Object Detection, Base Model.

2.1.4 Salient Objects Detection (SOD) Base Model

Many approaches have been proposed and using to detect salient feature accurately. However, in general few stages are common in all approaches, as illustrated in Fig 2.3

Image segmentation is the first step, partitions the image into multiple segments.

The image segments from first step are input for second step and in second step different priors are calculated for input segments. This step computes various scores for different image characteristics and these characteristics can be color, contrast, spatiality or texture.

The third and last step, prior combination, is the process to combine the scores generated in step two and generates the final saliency map. Each of these stages is explored further in the sections that follow.

2.1.4.1 Segmentation

First step is Image segmentation, at this important stage, the image is divided into regions. How to divide the image into how many regions, is the critical at this stage.

Pixel level saliency is used consistently in last few years [13], [23]. No context or indication is derived from the pixel itself and in which object pixel lies.So now a days, the group level saliency is based on pixel groups.

There are two general categories for the image segmentation process: super-pixel based [24], [25], [26] and cluster based [27], [28]. In Super-pixels segmentation, continuity is the advantage and clusters are not grouped spatially.

2.1.4.2 Cue Calculation

The most common cue is color contrast [27], [29], [30], [1]. Regions are compare globally with other regions for color contrast calculations and other methods use color contrast related to neighboring regions.

Spatial cue is another prior, based on pixel location [31], [32]. The idea is that human visual system model (HVS model) focuses on objects closer to the image centers. Center bias is mostly used in methods use.

Color distribution is more recent addition [31], [33] defined based on spatial variation of the color. Compact color distribution is normally characteristics of salient region.

2.1.4.3 Cues Combination

when cues are calculated, to form a final saliency map, these calculated cues are combined. We can use multiplication, average, summation or weighted average to combine these cues. Multiplication is preferred where precision is desired more then recall and the mostly adopted way; it can be either linear [27] or nonlinear [33].

2.2 Underwater Image Enhancement and SOD Methods

The saliency detection process, a two step saliency detection approach is used. The clustered-based method is used to extract a highly cohesive global constraint. After extracting highly cohesive global constraint, the bottom-up sapatial and contrast saliency cues are generated [27]. segmentation is done using a cluster based method and linear multiplication is performed for combination. The multiplication is used in order depress the noises instead of summation. They also apply their method the sequence of multiple images, to detect saliency accross the images (co-saliency). The co-saliency is not in domain of our work.

In this paper[34], they proposed a two level combined model for segmenting an underwater image. To segment and extract underwater image features, cluster based co-saliency object detection and local statistical active contour models are pipe lined.The cluster based co-salience detection algorithm processes sequence of underwater images and detect common features among input images. After co-saliency is established in first step, the region based local statistical active contour model, use saliency information and extract the detected objects with more accuracy.

Another interesting Saliency Object Detection approach is proposed in [35]. They discussed a mechenism for saliency based on cellular automata. They generated the map based on color and spatial contrast, by focusing on background. The closely similar neighbour regions are explored and get connected. They use superpixel segmentation for image segmentation step. Bayesian algorithm is used to combine multiple saliency maps.

Li et al. [36] create two maps from the background perspective and from the foreground salient features. Two saliency measures, the salient object driven measure (SODM) and the background driven measure (BDM) are generated. Color contrast cues are used in SODM to isolate salient features. On the other hand, boundary information is used by BDM to remove the background pixels. The combination of two calculated cues generates the final saliency, with the aim remove the background and highlighting the foreground (salient pixels).

This paper [37] class3 fuzzy C-means clustering and CLAHE image contrast enhancement are combined in their proposed method to process an underwater image for salient object detection. Thresholding is also used with class 3 fuzzy Cmeans clustering. Before image segmentation, the image contrast of underwater image is enhances using CLAHE enhancement method. The idea is to prominent the features to improve the segmentation. A distributed pseudorandom numbers generator function is used to initialize the fuzzy membership function.

Bazeille et al. [38] An underwater image algorithm is proposed to process the underwater images. It improves the image quality by reducing the underwater perturbations. This method contains many steps and each step perform independently. These steps make illumination correct, sharpen edges, remove noise and correct the colors. No prior information and parameter is required by this algorithm. This method is applied to detect edges. In Figure 2.2 images before and after applying this algorithm are shown [38].

Iqbal et al. [39] present this work for underwater image enhancement method. They perform different processing on different underwater image color spaces and generate an integrated color model. The have used Slide stretching for their method. In RGB color space perform the contrast stretching to equalize the color

FIGURE 2.4: Images (a) before and (b) after applying Bazeille et al.' processing. Image from : [38]

contrast. Secondly in HSI color space , saturation and intensity stretching is applied. This step correct the colors and enhance contrast as well. In Figure 2.3 two sample images are shown with corresponsing images after applying Iqbal et al'.

To process single underwater image without prior information is a challenging task and some complex methods are proposed. Ancuti et al. [40] built-up a fusion strategy to deal with under water images. This method takes a sequence of inputs derived from the initial image. The method takes two inputs as the candidates for fusion. The two input are derivation of original image, one is white balance of original image and another is histogram equalization out of original image, then four weight maps are generated to compute the normalized weight map, and follows by the Gaussian pyramid-based fusion process. The draw back of this method is, it over-correct some of the images especially bluish images. It add red channel into areas where it is not required.

In literature review we have discussed different papers and techniques they have

Figure 2.5: Original images with their corresponding images after applying Iqbal et al'. technique. Image from: [39].
Input images in left column, the processed images in right column.

presented. These papers and publications are from the domains of image segmentation, salient object detection, underwater environment and its properties, the affect of ocean properties on images taken in water. We also studied, how different techniques are used in order to identify salient object from underwater images and what are their strengths and weaknesses.

Table 2.1: Image Enhancement and SOD related work

Year	Development	Author	Reference
2013	Cluster-based Co-saliency Detection	Huazhu Fu,	[27]
2017	Underwater Image Segmentation with Co-Saliency Detection and Local Statistical Active Contour Model	Yue Zhu1,	[34]
2015	Saliency detection via cellular automata	Qin, Yao and Lu,	[35]
2014	Saliency detection via foreground rendering and background exclusion	Li, Yijun and Fu,	[34]
2014	Segmentation of underwater objects using clahe enhancement and thresholding with 3-class fuzzy c-means clustering	Singhet al.,	[34]
2017	TWO-STEP Approach for Single Underwater Image Enhancement	Xueyang Fu,	[41]
2007	Underwater Image Enhancement Using an Integrated Colour Model.	Iqbal et al.,	[39]
2006	Automatic underwater image pre-processing	Bazeille et al.,	[38]
2011	Fusion-based restoration of the underwater images	Ancuti et al.,	[40]
2019	Salient object detection: A survey	Borji et al.,	[18]

Chapter 3

Underwater Image Salient Object Detection using Image Enhancement and Clustering

3.1 Introduction

In this chapter, a simple yet effective two-step Salient Object Detection technique is presented. The first step enhance underwater image visual quality and second step detects the salient features of the underwater image.

My proposed process is built upon some previous image enhancement and salient object detection state-of-the-art methods. The block representation of proposed method is shown in figure 3.1.

Other then the Salient Object Detection three steps (image segmentation, cues calculation, and cues combination), my method added image enhancement features of color correction and color contrast enhancement. A prepossessing step is added before salient object detection three steps. The prepossessing step is further divided into two sub steps, underwater image color correction and the adjustment of underwater image contrast.

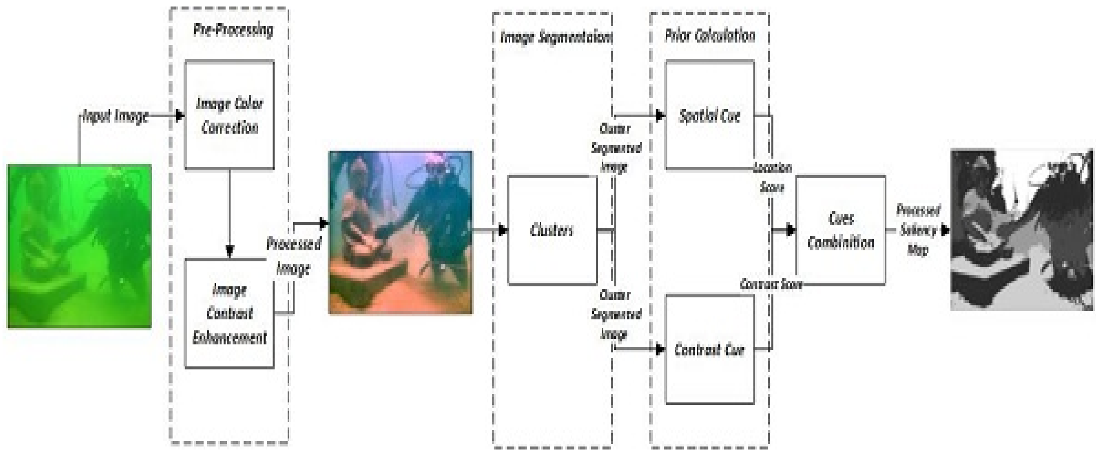

FIGURE 3.1: Block Diagram for Underwater Image Salient Object Detection

3.2 Proposed Methodology

The proposed methodology flow is shown in figure 3.2. The methodology is developed after reviewing literature and studying the available techniques and models. During my literature review I have compared existing models for image color correction, image contrast enhancement, image segmentation and salient object detection. The proposed methodology consists of following steps:

1. The first prepossessing step is underwater image color correction. Different color correcting models are studied and after comparing their results, the Gray World Assumption [41] is picked for my image color correcting step.

2. The second step of prepossessing is, to adjust image contrast. CLAHE [42] is used for this step.

3. The objective of this image segmentation step is to partition the under water image into manageable regions and this prepare the image for next step of prior calculations. For image segmentation K-mean(Hard clustering), Fuzzy C-mean clustering(Soft Clustering) and Active Contour Model are compared. After comparing their results, K-mean found better in terms of accuracy and efficiency. Hence I have used K-mean for my proposed solution for image segmentation.

4. Because ore proposed solution does not rely on any prior knowledge. So to detect salient object a step is required to calculate some prior knowledge. This knowledge is calculated normally using input image features like image textures, image colors, image contrast, image pixel locality etc. I have used contrast and spatial features of input image to calculate the priors.

5. To combine saliency maps, calculated as contrast and spatial priors, linear multiplication or linear addition is used. Linear multiplication is desired where accuracy matters and linear addition is desired where recall have more weight-age over precision. In our case, accuracy or precision is more desirable over recall. So I have used linear multiplication to combine the priors in order to get more accurate salincy map.

6. In last step I have compared my proposed solution with avaialable existing models.

3.3 Flow Chart

The implementation flow chart is shown in figure 3.3. The six steps are implemented. The six steps are:

1. Underwater image color correction
2. Underwater image contrast adjustment
3. Enhanced image K-mean clustering
4. Calculate contrast prior for each cluster
5. Calculate spatial prior for each cluster
6. Combine contrast and spatial priors using linear multiplication
7. Object Segmentation

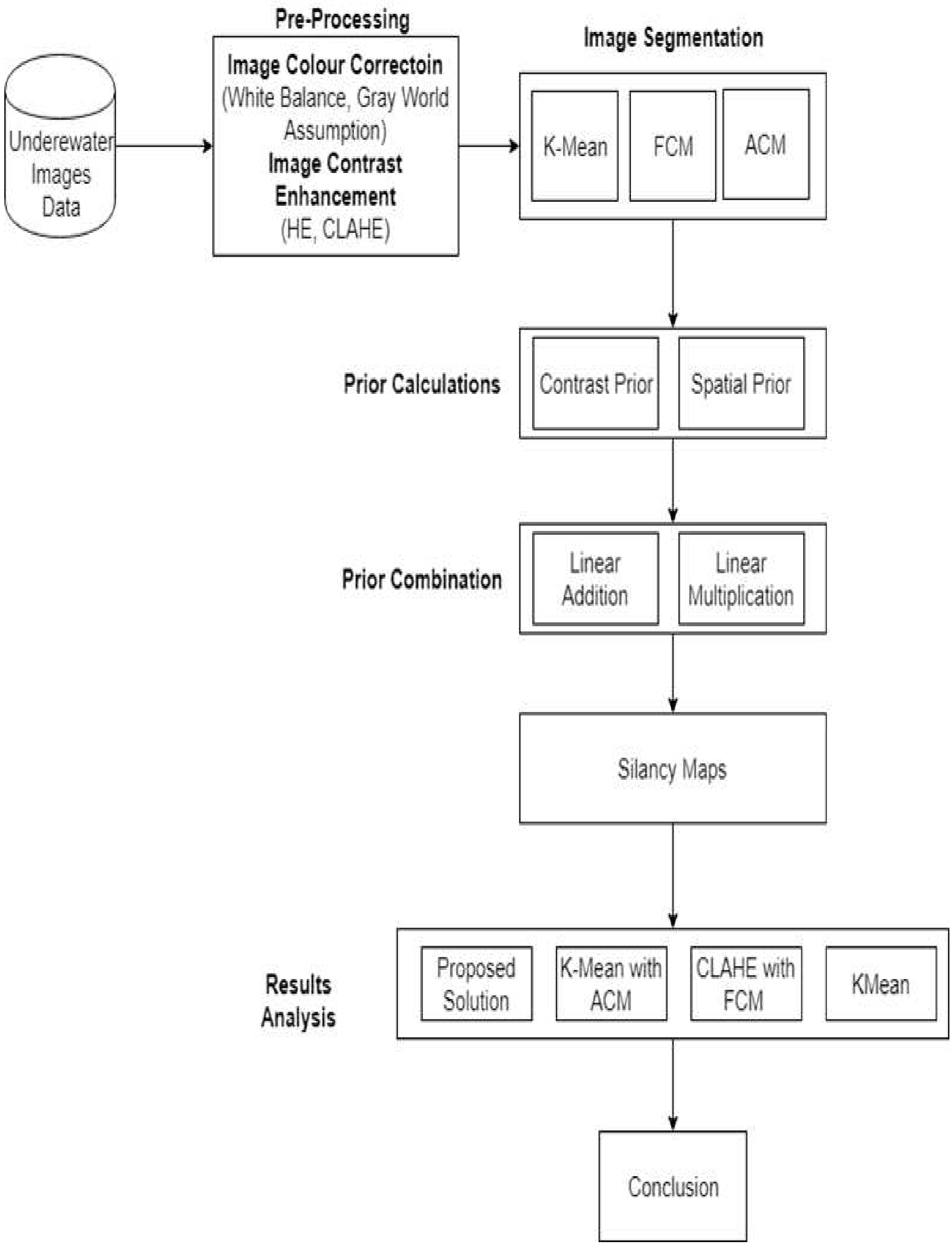

FIGURE 3.2: Proposed Methodology.

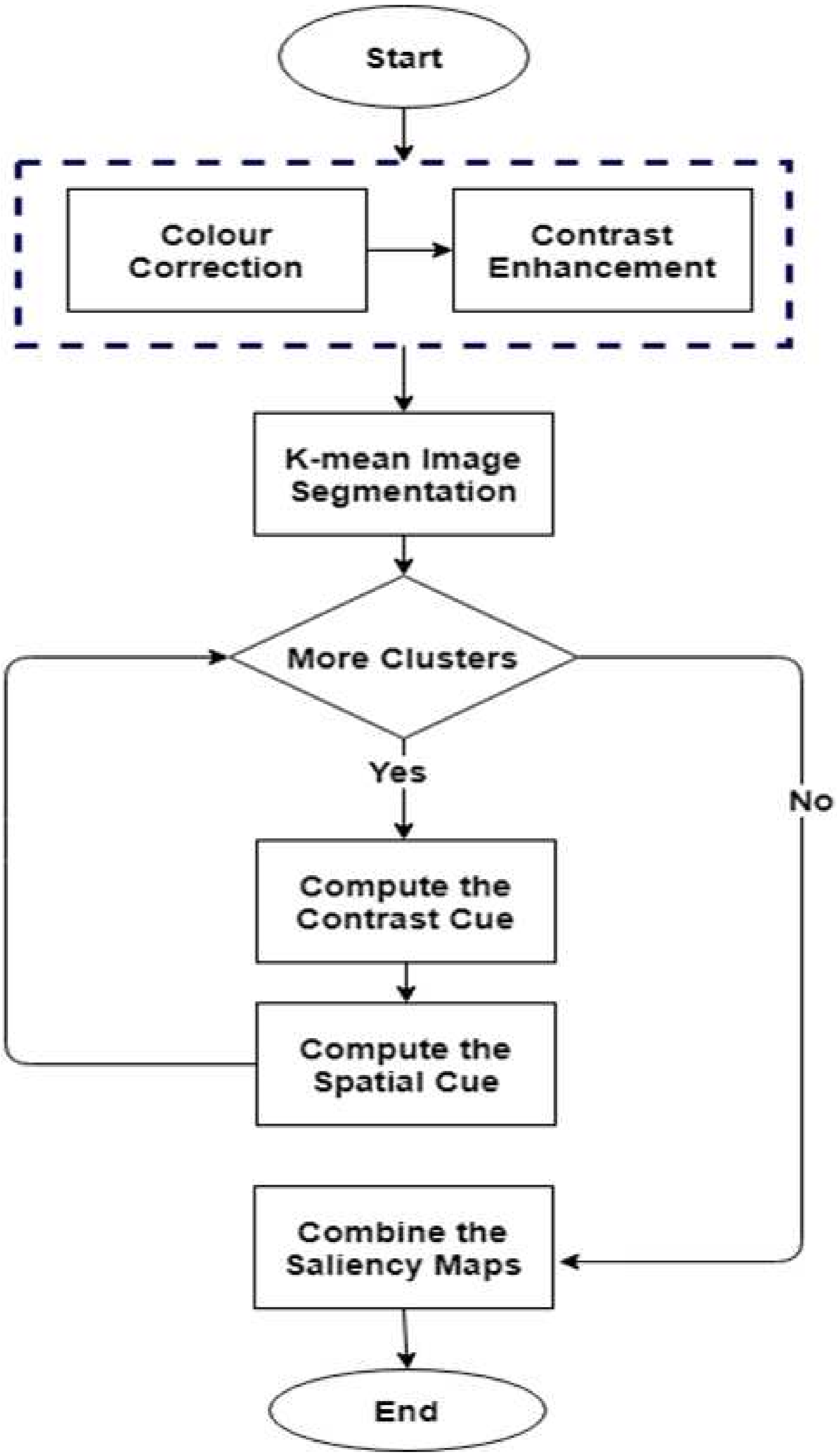

FIGURE 3.3: Process Flowchart.

3.4 Pre-Processing

The objective of pre-processing is to improve edges and boundaries of an image that to get better results from saliency cues calculations processes. This pre-processing helps in improving the performance of the segmentation under the noisy and non uniform aquatic environments.

This first step deals with the two major issues of underwater image, the color distortion and low contrast. For the color distortion problem, a color correcting strategy based on a piece-wise linear transformation is proposed [41] after inspiring with gray world assumption[43]. For low contrast,contrast-limited adaptive histogram equalization (CLAHE)[42] is used.

In some cases, the white areas in the scene appear coloured in the pictures. Generally invisible to the human eye, which is able to do the correction automatically, it causes extremely saturated colours after the application of a restoration method. This problem is particularly sensitive in underwater conditions during the restoration.

The Gray World is one of the oldest methods used to perform a white balance. It considers that the average intensities of the three RGB channels must be equal. The principle is to keep the green channel unchanged and to multiply the red and blue channels, respectively by the gains

The gray-world hypothesis assumes that the average color in a scene is achromatic[43]. Based on this gray-world hypothesis, this color correcting approach adopts a piece-wise linear transformation to stretch the images mean value toward 128.The pre-processing first step results for three sample images are shown in Figure 3.2.

A piece-wise linear transformation function is applied for stretching the mean of the image towards 128. S variable is the input image. For each color channel $C \in r, g, b$ the max, the min and the mean is computed. The mathematical expression of it as follows,

FIGURE 3.4: Results of Applying Color Correction Algorithm. First Row: Original images. Second Row: Color Corrected images.

$$S_{CR}^{C} = \begin{cases} \left(S^{C} - S_{mean}^{C}\right)\frac{S_{min-128}^{C}}{S_{min}^{C}-S_{mean}^{C}} + 128, & S_{mean}^{C} \leq 128, \\ \left(S^{C} - S_{mean}^{C}\right)\frac{S_{max-128}^{C}}{S_{max}^{C}-S_{mean}^{C}} + 128, & S_{mean}^{C} > 128 \end{cases} \tag{3.1}$$

S_{CR}^{C} Shows the color corrected image while $S^{C}, S_{min}^{C}, S_{max}^{C}$ are the mean, maximum and minimum of the channel. The mean decides the direction to stretch.

$$S_{CR}^{C} = \begin{cases} S^{C} - \lambda\left(S_{mean}^{C} - 128\right), P^{c} > 0.7, \\ \left(S^{C} - S_{mean}^{C}\right)\frac{S_{min-128}^{C}}{S_{min}^{C}-S_{mean}^{C}} + 128, & S_{mean}^{C} \leq 128, \\ \left(S^{C} - S_{mean}^{C}\right)\frac{S_{max-128}^{C}}{S_{max}^{C}-S_{mean}^{C}} + 128, & S_{mean}^{C} > 128 \end{cases} \tag{3.2}$$

here λ behaves as a controlling parameter, its a positive number and manipulates the shifting range and P^{c} describes the probability of a pixel values which are equal to or less than 40. It handles the over-correction rather effectively.

The color correction effect histogram is show in Figure 3.4. The histogram of input image in first row, shows that the green channel have large values and red and blue channels have low values in this greenish image. When apply gray world piece-wise linear transformation, all channels are stretch towards middle. The resultant image and its histogram is shown in second row of Fig 3.4.

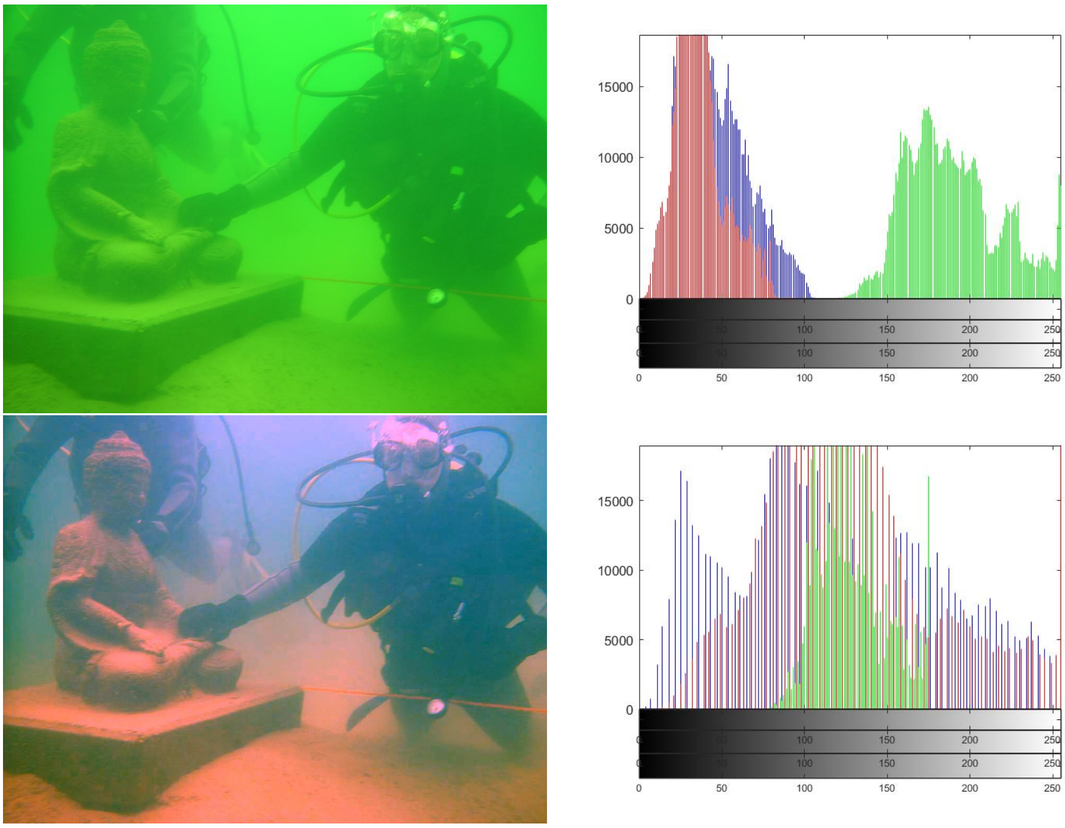

FIGURE 3.5: Histograms After and Before Applying Color Correction Algorithm. First Row: Original images. Second Row: Color Corrected images.

Contrast limited adaptive histogram equalization (CLAHE)[42] is used to enhance the contrast and minimize the noise problem. This differs from standard histogram equalization in the respect that this method operates on small tiles of an image and computes several histograms, each corresponding to a different region of the image and use them to redistribute the lightness values of the image. CLAHE[42] improves local contrast of an image.The pre-processing second step results for three sample images are shown in Figure 3.5.

3.5 Image Segmentation

The aim of image segmentation is to partition the image into manageable regions that prepare the image for prior calculations. My method adopts a cluster-based technique to preserve edge integrity and object coherence. A cluster-based method

Figure 3.6: Results of Applying CLAHE Process. First Row: Original images. Second Row: Color Corrected images. Third Row: Contrast Adjusted images

to detect saliency in an image[27]. The cluster-based idea is inspired by the global-contrast methods on the single image.

3.5.1 Cluster-based Segmentation

After image color correction and contrast enhancement, clustering is the next step. The cluster based idea is derived from global contrast based phenomenon[1]. K-mean clustering is used and is defined as:

$$L = \{C_K\}_{k=1}^{k_c} \tag{3.3}$$

K-means is used to separate pixels based on the Lab color space. After different experiments I have divided image into 6 clusters.

3.6 Cue Calculation

At this stage we have six clusters after partitioning the image into appropriate regions. So six cluster is input for this cue calculation process. To assign salient scores to each regions is the main objective of this cue calculation process. The higher score indicates more salient (e.g. prominent feature) and a lower score indicates less salient (e.g. less prominent features and background).

3.6.1 Cluster-Based Saliency Cues

Two cluster-based cues are introduced to measure the cluster-level saliency. These two cues are contrast and spatial cues, which are used in the single image saliency detection. The main property of cluster-based method is that the visually prominent cues detect on cluster-level rather than the individual pixel-level.

Notations: The pixel is denoted by $\{p_i^j\}_{i=1}^{N_j}$ with index i in the image I^j, where the N_j denotes the j^{th} image lattice. $\{z_i^j\}_{i=1}^{N_j}$ denotes the normalized location of the pixel p_i^j in the image I^j. Given M images $\{I^j\}_{i=1}^{M}$, we obtain K clusters $\{C^k\}_{k=1}^{K}$. The clusters are denoted by a set of D-dimensional vectors $\{\mu^k\}_{k=1}^{K}$, in which μ^k denoted the prototype $\big($cluster centers$\big)$ associated with the cluster C^k. And the function $b : {}^2 \rightarrow \{1...K\}$ associates the pixel p_i^j and the cluster index $b\big(p_i^j\big)$.

1. Contrast Cue: Contrast cue defines the unique prominent feature on the single or multiple images. Contrast is one of the most widely used cues for measuring saliency in single image saliency detection algorithms [1], [3], [7], since the contrast operator simulates the human visual receptive fields. The contrast cue $w^2\big(k\big)$ of cluster C^k is defined using its feature contrast to all other clusters:

$$w^c\big(k\big) = \sum_{i=1, i\neq 1}^{K} \left(\frac{n^i}{N} \parallel \mu^k - \mu^i \parallel_2 \right) \tag{3.4}$$

2. Spatial Cue: In human visual system, the regions near the image center draw more attention than the other regions [32][34]. When the distance between the object and the image center increases, the attention gain is depreciating. This scenario is known as central bias rule in single image saliency detection.

Contrast cue is expert in discriminating the most salient object. Spatial cue is good at handling the textured background around the image boundaries. Our final single image saliency map joints two cues and obtains a satisfactory saliency map. [27] extends this concept to the cluster-based method, which measures a global spatial distribution of the cluster.

The spatial cue $w^s(k)$ of cluster C^k is defined as:

$$w^s(k) = \frac{1}{n^k} \sum_{j=1}^{M} \sum_{i=1}^{N_j} \left[\mathcal{N}(\parallel z_i^j - o^j \parallel^2 | 0, \sigma^2) \cdot \sigma\left[b(p_i^j) - C^k \right] \right] \tag{3.5}$$

where $\sigma(\cdot)$ is the kronecker delta function, o^j denotes the center of the image I^j, and Gaussian kernel $\mathcal{N}(\cdot)$ computes the Euclidean distance between pixel z_i^j and the image center o^j, the variance σ^2 in the normalized radius of images. And the normalized coefficient n^k is the pixel number of cluster C^k. Different from the single image model, our spatial cue w^2 represents the location prior on the cluster-level, which is a global central bias on the multiple images.

3.6.2 Cues Combination

So far, two bottom-up cues in our cluster-based method are introduced. Each cue, if used independently, has its advantages and, of course, disadvantages. After computing these two cues, I combine the cues with using point-wise multiplication. A common fusion is formulated as a linear summation [1], [7] or point-wise multiplication [35] of static salient features. The multiplication is better in depressing the noises than summation. And summation is better in getting higher recall. For saliency detection, however, the precision is more important than recall [31]. In

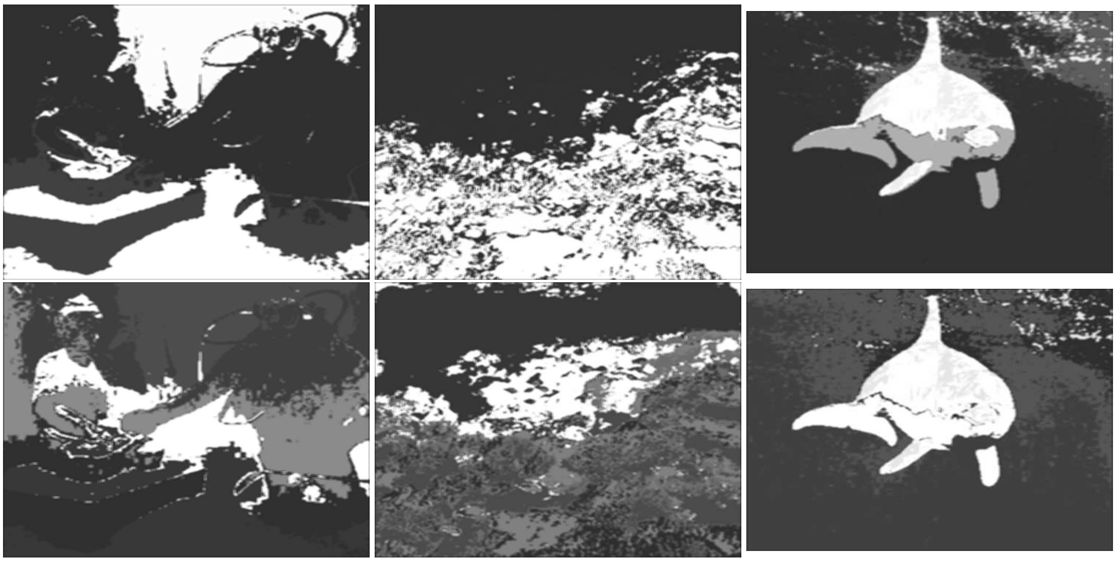

Figure 3.7: Prior Saliency Maps. First Row: Contrast Cues. Second Row: Saliency Cues

our work, we also prefer a precise, rather than a large, saliency map. Therefore, we employ the multiplication operation to integrate the saliency cues in order to get more accurate results.

Before combining saliency cues, we normalize each cue map to standard Gaussian using the distribution of scores across all clusters. Then the cluster-level co-saliency probability p(k) of cluster k is defined as:

$$p(C^k) = \prod_i w_i(k) \tag{3.6}$$

$$p(x|\, C^k) = \mathcal{N}(\mid \| v_x - \mu^k \|_2 \mid, 0, \sigma_k^2) \tag{3.7}$$

where ν_x denotes the feature vector of pixel x, and the variance σ_k of Gaussian uses the variance of cluster C^k. Hence, the marginal saliency probability $p(x)$ is obtained by summing the joint saliency $p(C^k)p(x|\, C^k)$ over all the clusters:

$$p(x) = \sum_{k=1}^{K} p(x, C^k) = \sum_{k=1}^{K} p(x|\, C^k)p(C^k) \tag{3.8}$$

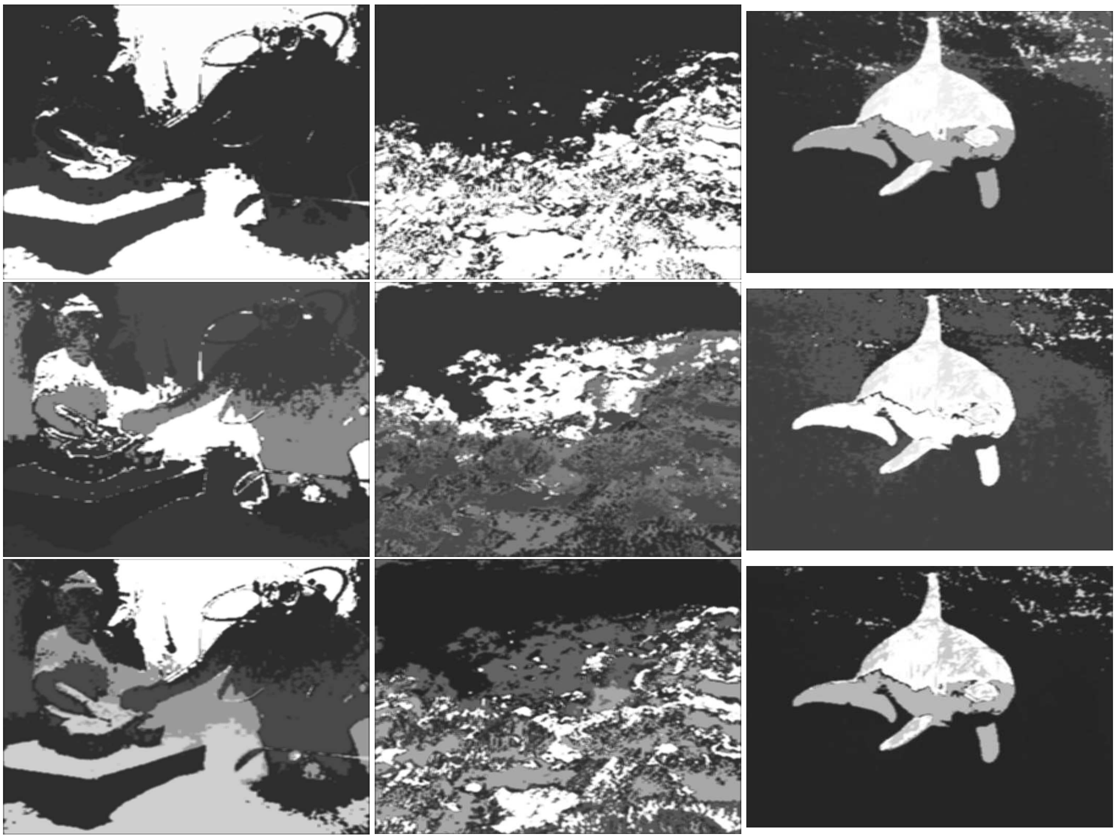

FIGURE 3.8: Prior Saliency Maps. First Row: Contrast Cues. Second Row: Saliency Cues. Third Row : Combined Saliency Cues

3.6.3 Algorithms

The saliency detection method and its steps are summarized in Algorithm 1.

Algorithm 1: Cluster-based saliency Detection

Input: Input pre-processed image, Intra cluster number K1
Clustering image into K1 Clusters;
foreach *each cluster* **do**
  Computing the contrast cue using Eq. (3.3) and spatial cue using Eq. (3.4);
end
Combining two saliency cues using Eq. (3.5);
foreach *each cluster* **do**
  Obtaining the final single saliency map using Eq. (3.7);
end
Output: The saliency map

The proposed method and its steps are summarized in Algorithm 2.

Algorithm 2: Proposed solution

Input: Input under water image, Intra cluster number K1

`/* Input Image color correction */`

Apply gray world assumption linear transformation on input image to correct the image color using Eq. (3.1)

`/* Adjust color contrast */`

Apply CLAHE on color corrected image to enhance the color contrast

`/* Cluster-based saliency Detection */`

Clustering image into K1 Clusters;

foreach *each cluster* **do**

  Computing the contrast cue using Eq. (3.3) and spatial cue using Eq. (3.4);

end

Combining two saliency cues using Eq. (3.5);

foreach *each cluster* **do**

  Obtaining the final single saliency map using Eq. (3.7);

end

Output: The saliency map

3.6.4 Final Results

Methods for single image without prior knowledge have some disadvantages, to accommodate the underwater image complexity some relatively complicated methods are proposed. Ancuti et al. [15] propose a fusion-based image enhance method to deal with underwater images. This method uses two inputs as the resources of fusion, one is white balance result and another is histogram equalization result, then four weight maps are generated to compute the normalized weight map, and follows by the Gaussian pyramid-based fusion process. Because of over-saturation and color aberration for some test images, the author modifies and proposes an improved fusion algorithm [16], which obtains excellent performance. This method works really well for greenish images but makes blue image over correct.

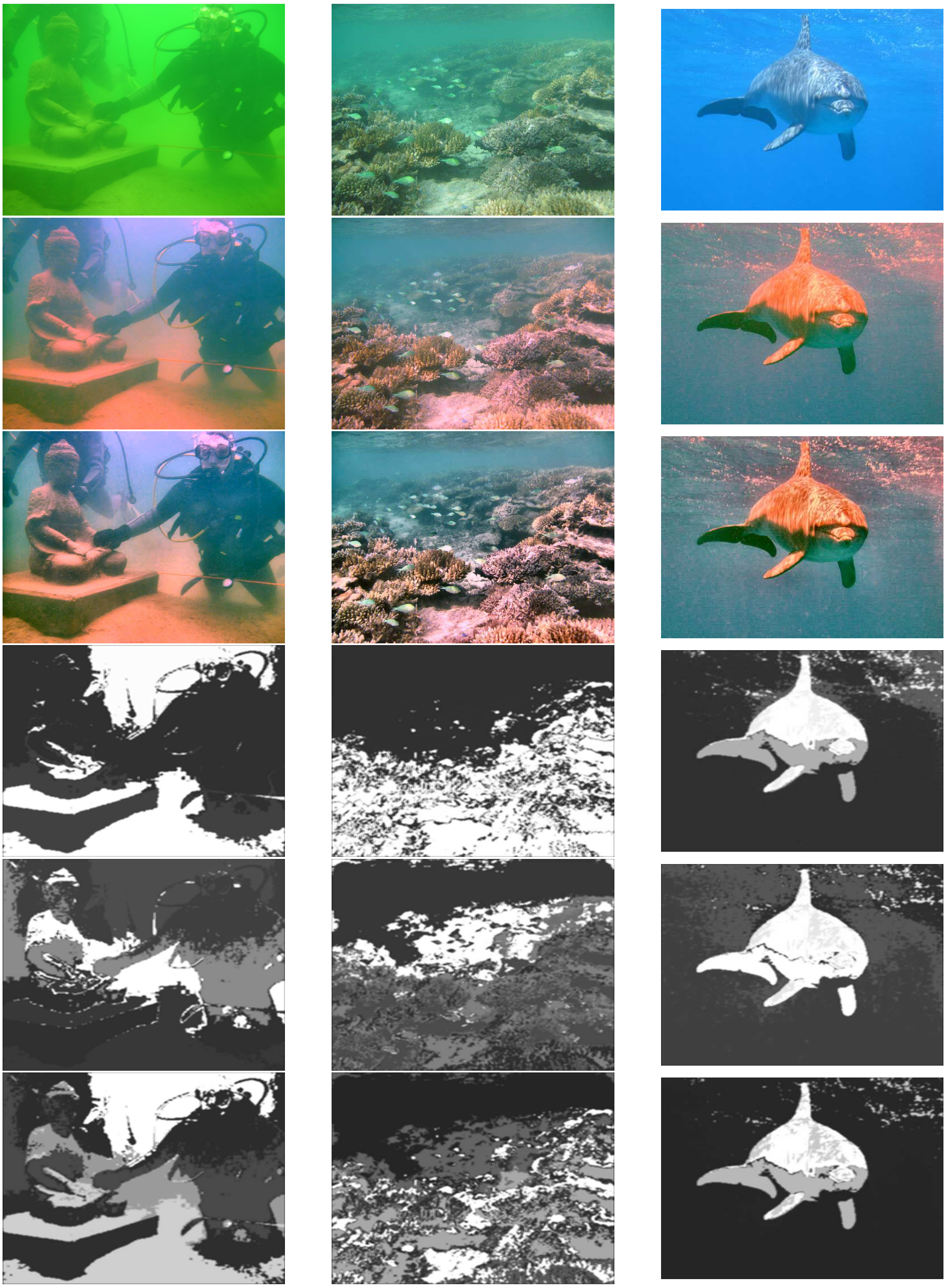

FIGURE 3.9: Final Results. First Row: Original images. Second Row: Color Corrected images. Third Row: Contrast Adjusted images. Fourth Row:Contrast Cues. Fifth Row: Saliency Cues. Sixth Row : Combined Saliency Cues

Chapter 4

Results and Analysis

4.1 Introduction

This chapter presents the results and analysis of the an improved model to detect the salient features of underwater image.I here perform various experiments with different variations to show the effectiveness of my method.In order to show the effectiveness and usefulness of each stage of the method, I first define four variant systems, in sections 4.2.1 through 4.2.4, by removing a step from the proposed model. The complete model is shown in section 4.2.4. Further the results of the variant models are visually compared with the results of the full system in section 4.2.4 and quantitatively compared in section 4.2.4.

4.2 System Model Variants

This section illustrates the prepossessing steps(underwater image Color correction and underwater image contrast enhancement), image segmentation, prior calculations and prior combination stages of the model. These result of each stage is compared with each others. Both subjective and quantitative comparisons are used to analyze and compare the results.

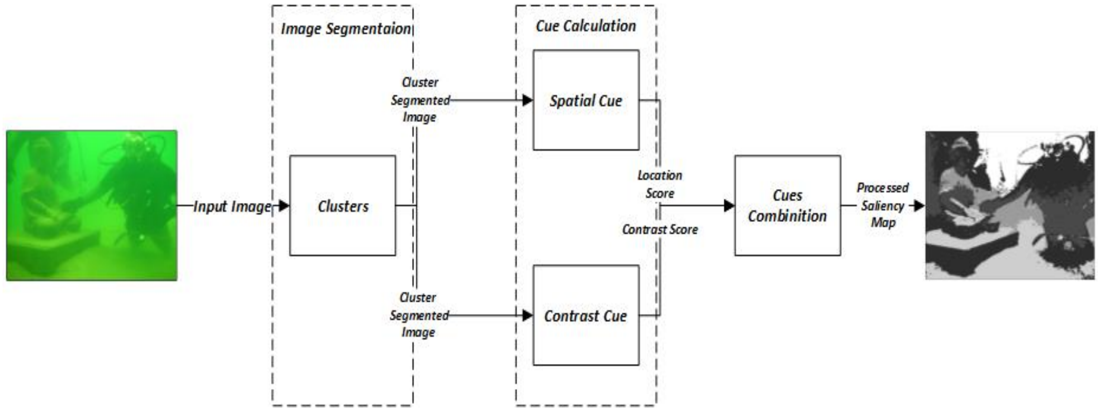

FIGURE 4.1: Block Diagram for Underwater Image Salient Object Detection

4.2.1 Baseline Model

The block diagram of the first variant system is demonstrated by Figure 4.1. The both pre-processing steps are removed. This baseline works very similar to the system in [27]. Sample map results for the initial system are shown in the fourth column of Figure 4.2.

4.2.2 Baseline With Image Color Correction

The block diagram of the second variant system is demonstrated by Figure 4.3. The image contrast enhamcement step from pre-processing stage is removed. This system variant is applying color correction before performing segmentation. Sample map results for the initial system are shown in the fifth column of Figure.

4.2.3 Baseline With Image Contrast Enhancement

The block diagram of the third variant system is demonstrated by Figure 4.11. The image color correction step from pre-processing stage is removed and only color contrast step remains in the system. This system variant is only enhancing image contrast before performing segmentation. Sample map results for the initial system are shown in Figure 4.16.

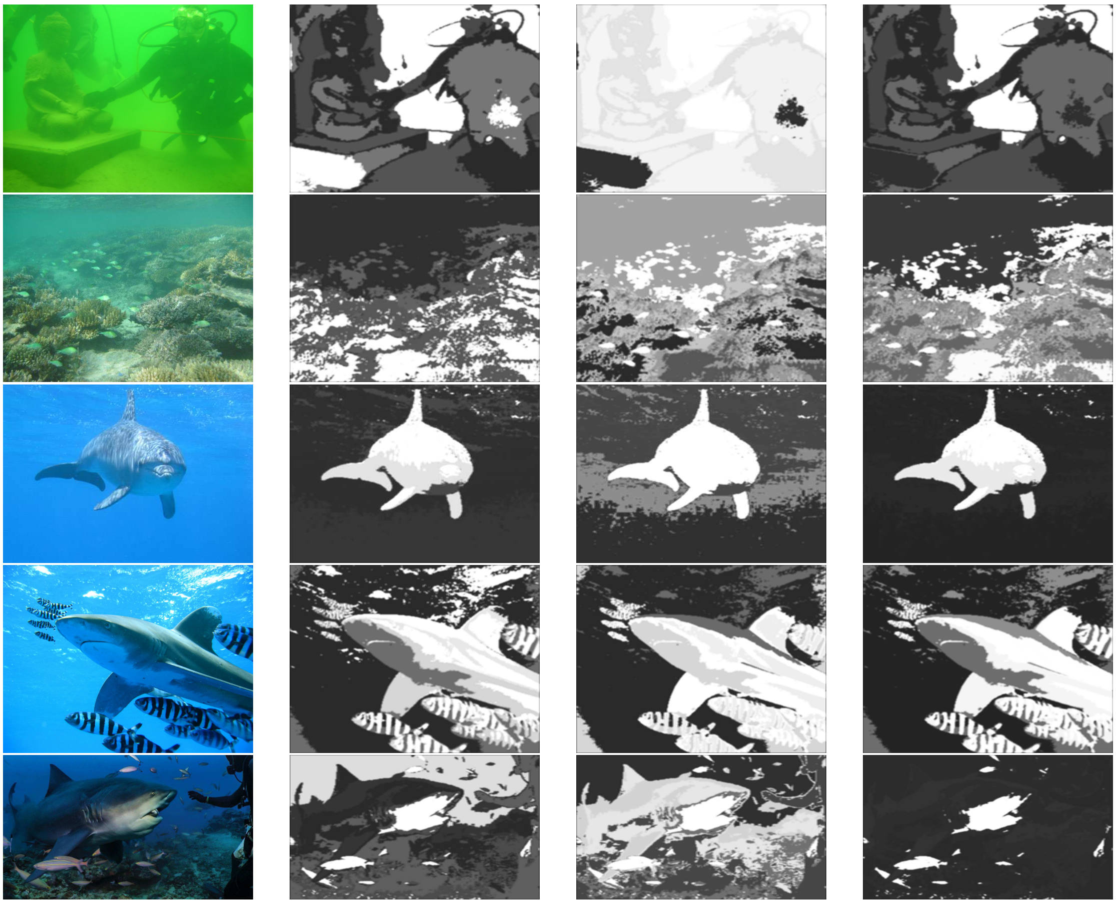

FIGURE 4.2: Results with no pre-processing (a): Original images. (b): Contrast Cue. (c): Saliency Cue. (d): Saliency Map

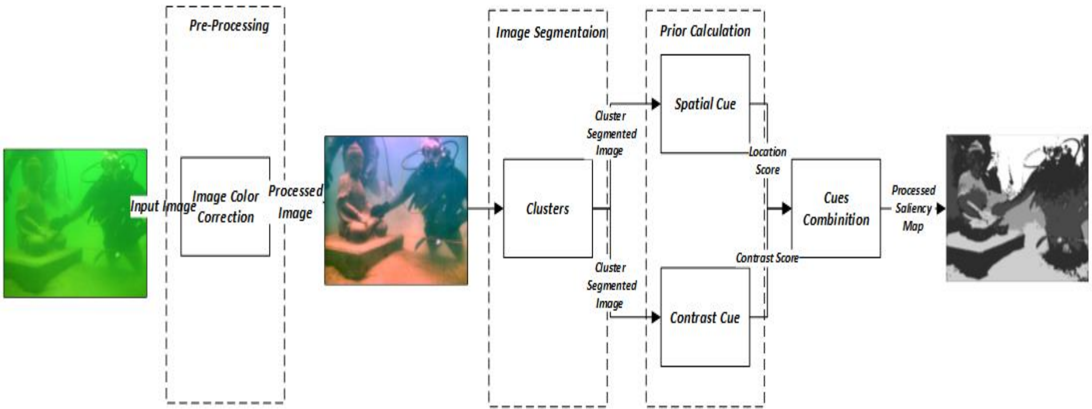

FIGURE 4.3: Block Diagram for Underwater Image Salient Object Detection

FIGURE 4.4: (a)

FIGURE 4.5: (b)

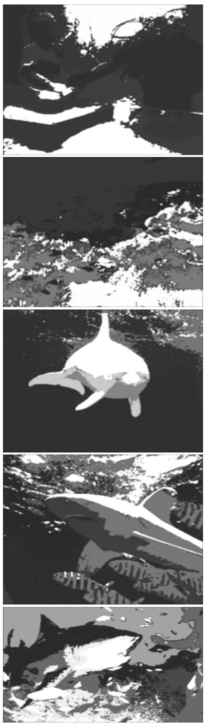

FIGURE 4.6: (c)

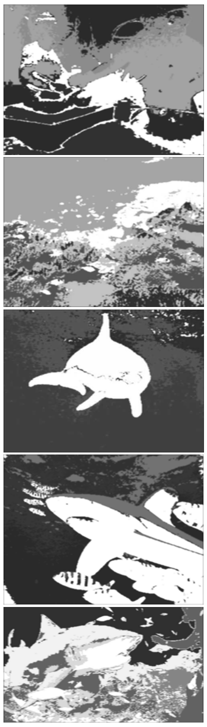

FIGURE 4.7: (d)

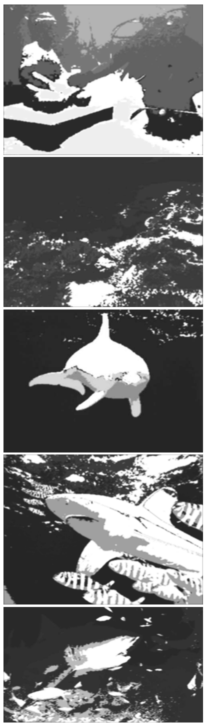

FIGURE 4.8: (e)

FIGURE 4.9: Results of Applying Color Correction Algorithm. (a): Original images. (b):Image with color correction (c): Contrast Cue. (d): Saliency Cue. (e): Saliency Map

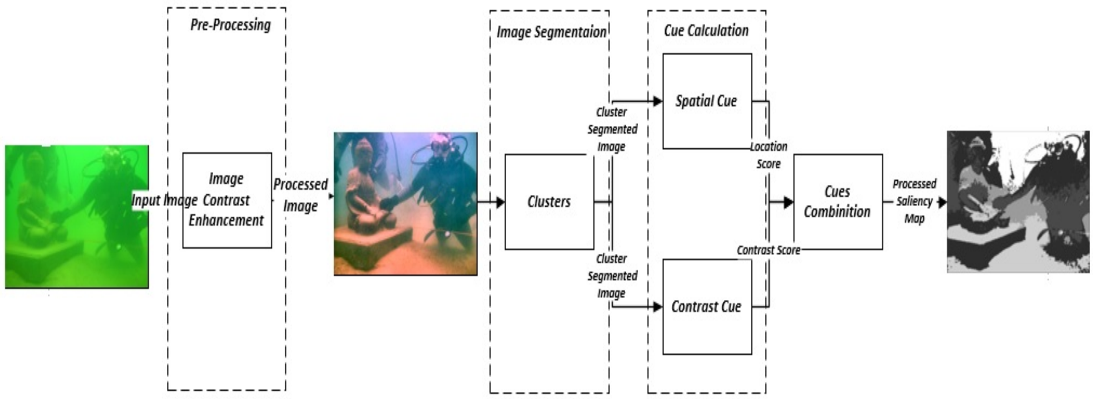

FIGURE 4.10: Block Diagram for Underwater Image Salient Object Detection

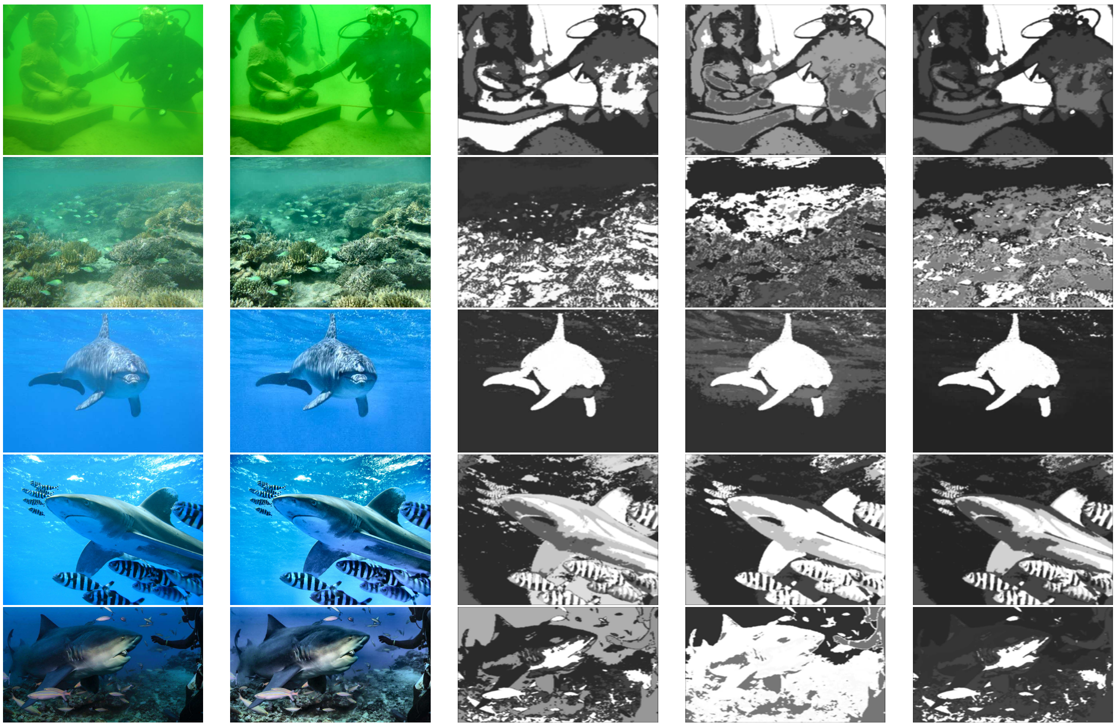

FIGURE 4.11: (a) FIGURE 4.12: (b) FIGURE 4.13: (c) FIGURE 4.14: (d) FIGURE 4.15: (e)

FIGURE 4.16: Results of Applying Color Correction Algorithm. (a): Original images. (b):Image with color correction (c): Contrast Cue. (d): Saliency Cue. (e): Saliency Map

4.2.4 Complete Model

Figure 4.18 presents the results of the variant models and and the full system for the underwater image. The outcome of the different stages are varying with images in out test dataset. For image1 in first row, color correct step result is introducing some extra features but in final result, the information and features are more prominent. The contrast adjustment step result is very similar to result without any pre-processing but contributing in end result. This image has only two objects.

Image2(Row 2) in Figure 4.18, the final result is better then result without pre-processing. Both color correct and contrast adjustment results are contributing well in final result.This image has some patterns of underwater life.

Image3(Row 3) in Figure 4.18, the result with only contrast adjustment only seems the best map among four. In final result, we can observe more prominent patterns

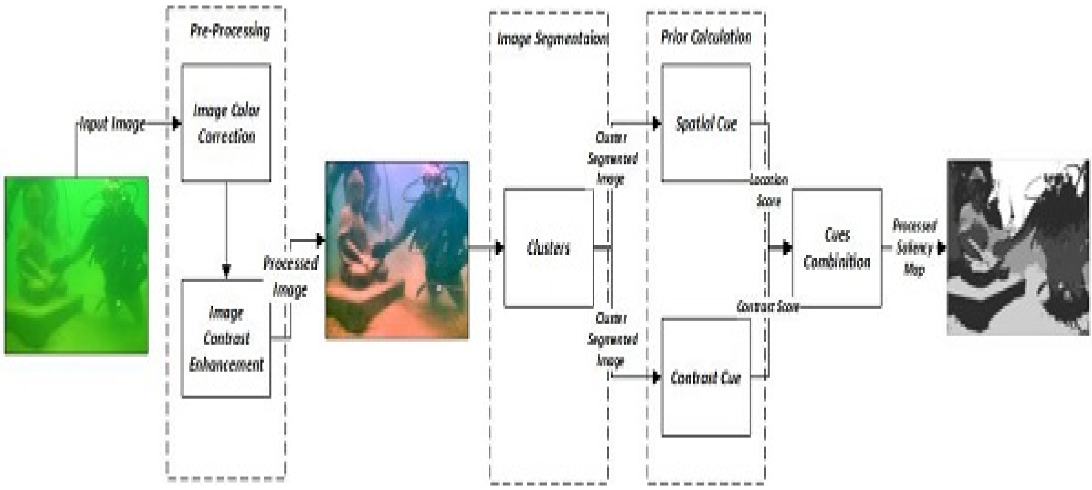

FIGURE 4.17: Block Diagram for Underwater Image Salient Object Detection

within fish object. This image contains only single object.

Image4(Row 4) in Figure 4.18, have multiple overlapping objects. The result without pre-processing and with only color correct look similar.

Image5(Row 5) in Figure 4.18, portrays the more complex details of aquatic life. and fianl results is much better then result without pre-processing.

4.3 Quantative Results

To compare our three variants with proposed model, a quantitative comparison is performed. For quantitative analysis,Precision/Recall (PR) curves and F-number are used to analyze and evaluate the results of all variants. Receiver Operating Characteristic (ROC) curves measure similar features to PR curves and are common in other research fields. However, PR curves are more commonly employed in saliency detection.

Here we discuss the most common metrics used for semantic segmentation. For reference, a general analysis of accuracy metrics for classification tasks can be found in [44]. The accuracy, or the ratio of the correctly classified elements over all available elements can be calculated as follows:

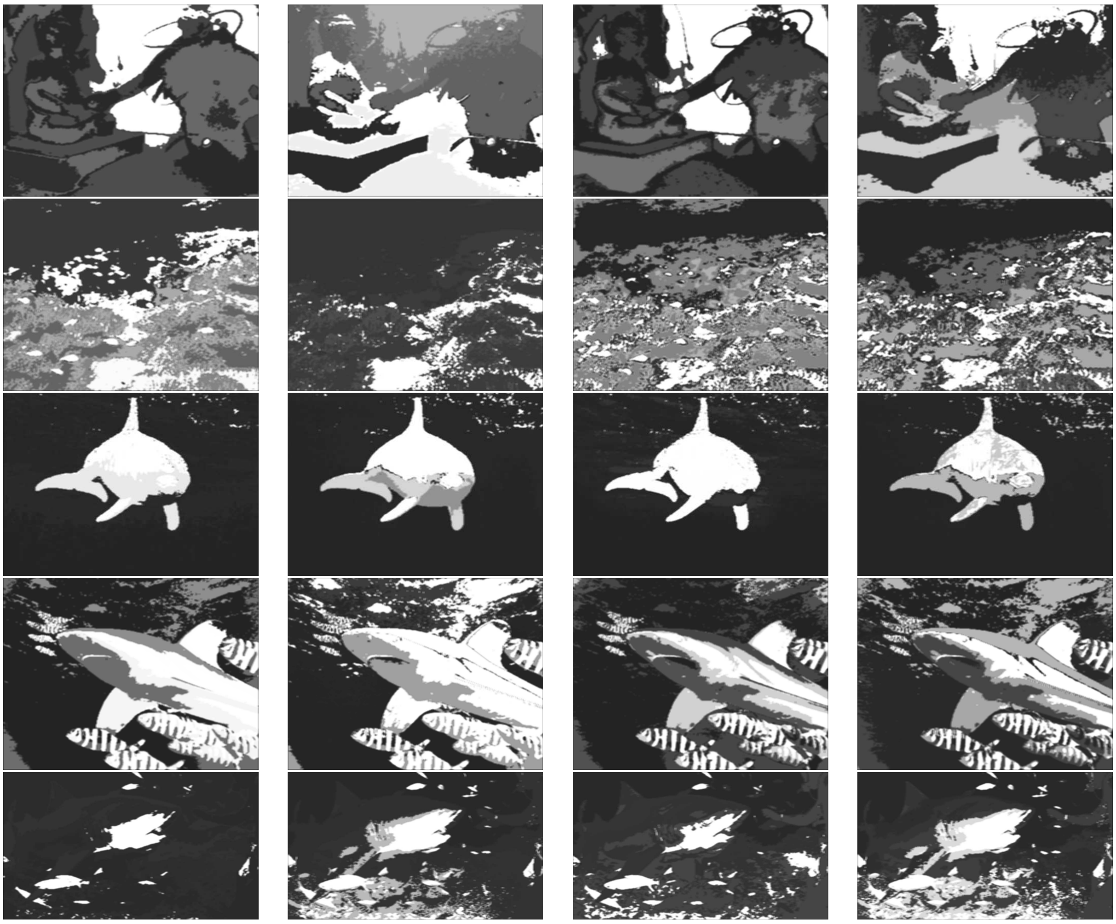

FIGURE 4.18: Final Saliency Maps (a): Saliency Maps without any pr . (b): Color Corrected Saliency Map. (c): Contrast Enhanced Saliency Map. (d): Saliency Maps with both pre-processing steps

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

The precision, or positive predictive value (PPV), is the relation between true positives and all elements classified as positives:

$$Precision = \frac{TP}{TP + FP} \tag{4.2}$$

Table 4.1: Class confusion matrix and notation

		Predicted class	
		Positive	Negative
True Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

The Recall, or true positive value (TPV), is the relation between true positives and all positive elements:

$$Recall = \frac{TP}{TP + FN} \tag{4.3}$$

The F-number, to compare the outputs.This quantity combines precision and recall into a single quantity for a system.The F-measure is a widely used metric to evaluate classification results, which consists of the harmonic mean of precision and recall metrics:

$$F = \frac{(\beta^2 + 1)TP}{(\beta^2 + 1)TP + \beta^2 FN + FP} \tag{4.4}$$

where β is scaling between the precision and recall.To create an F1 number that weights precision and recall evenly, Considering $\beta = 1$, leads to the widely used F1-measure:

$$F_1 = \frac{2TP}{2TP + FN + FP} \tag{4.5}$$

However, a recent trend in computer vision research is setting $\beta^2 = .3$ so as to weigh precision above recall.
It depends, what kind of output is required, Sometimes recall is more desirable over precision. For β^2 equal to or greater then .5 is used when recall have more weigh over the precision.

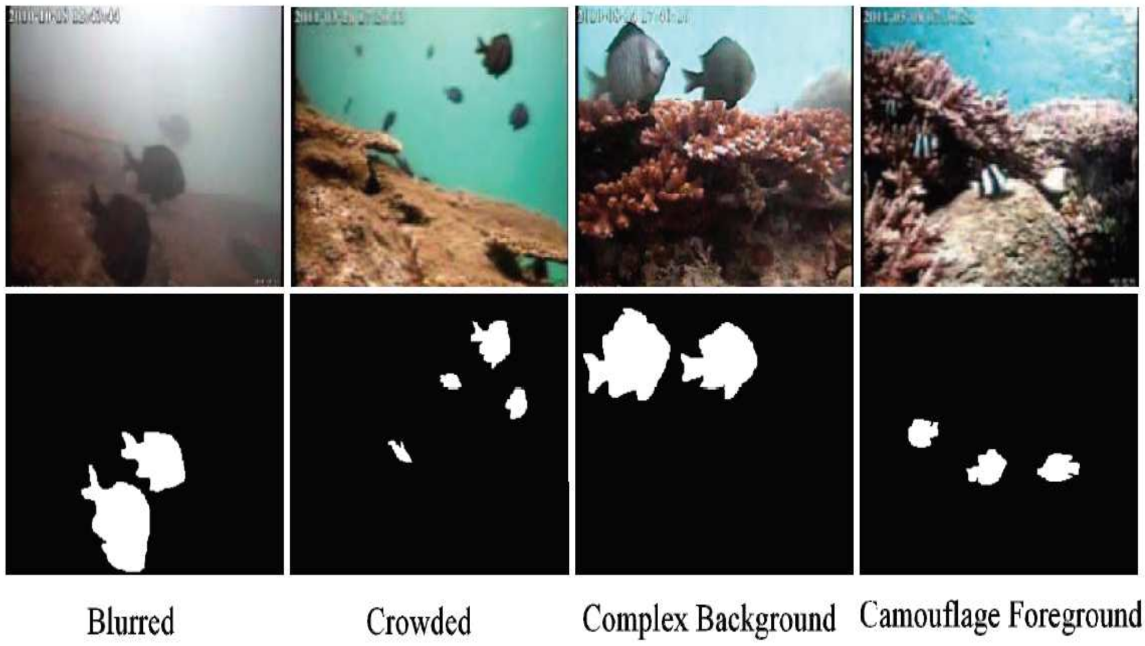

FIGURE 4.19: Fish4Knowledge - Ground truth samples, original images and ground truth Maps from four different categories.

4.4 DataSet

Fish4Knowledge Ground-Truth datasets, the underwater ground truth database having 780 images extracted from seventeen underwater videos and is classified into four different categories:

- Blurred
- Crowded
- Complex Background
- Camouflage

4.4.1 Results Comparison

In this subsection, we use the accuracy measurement to evaluate the accuracy of proposed solution for underwater objects visibility enhancing and detecting methods as well as comparison with the results produced by the state-of-art models.

TABLE 4.2: Precision, Recall and F-measure on four underwater data sets

	Recall	**Precision**	**F_/beta**
Blurred	0.6337	0.9205	0.8337
Crowde	0.6160	0.8659	0.7627
Complex Background	0.3455	0.7627	0.5941
Camouflage Foreground	0.4343	0.8694	0.7106

For this we compare the performance of our method to several state-of-art methods in recent years, where most of them also achieves excellent performance for the underwater imaging issue.

The K means clustering algorithm partitions an image into mutually exclusive clusters with aiming to minimize the distance measure between centroid of the cluster and the pixel.

Fuzzy C means is the also called soft K means and is the variant of K means.

Active Contour Model Computes the internal energy and the external energy

The experimental results are for the K means, Fuzzy C means and Active Contour Model algorithms on images. The algorithms are evaluated using the following measures:

1. Accuracy : It is the proportion of true results (both true positives and true negatives)
2. Peak Signal to Noise ratio (PSNR) : It is a measure of reconstruction quality.
3. Elapsed time : It is the total time taken for the completion of the program and calculates in seconds

Table.4.4. The images were processed using these three model under evaluation. The overall efficiency of the algorithms with user interaction, not only depends on the efficiency of the algorithm but also on the users expertise to give the input. For example, the ACM model outputs are influenced by the shape of initial contour and the speed with which the user specifies the contour. The overall results show that K meams provides better results for experimental images with comparatively minimum computational time.

Images captured in the different underwater environments often show various attenuation and degradation degree, which causes that an underwater imaging method may performs well for several certain underwater environment conditions, but weak for other conditions.

To show the effectiveness of the proposed model, I have performed many experiments in underwater image dataset. In section 4.2 I define several variations of my full system and show comparison results of all the systems.

The proposed method is compare with Yue et al.[34] and Singh et al. [26]. The Comparison results are shown in Figures 4.20, 4.21 and 4.22.

In figure 4.20 under water images with complex background (bottom of images) are chosen.

Images captured in the different underwater environments often show various attenuation and degradation degree, which causes that an underwater imaging method may performs well for several certain underwater environment conditions, but weak for other conditions. We choose a set of images from Fish4Knowledge dataset for this comparison because this dataset provides ground truth images as well. The images and their respective ground truth data is compiled from under water videos.

Yue et al.[34] model is using K-mean clustering with prior calculations and performing post processing using Active Contour Model segmentation.

Singh et al. [26] is using Fuzzy c-mean thresholding for underwater image segmentation after applying applying CLAHE as prepossessing step.

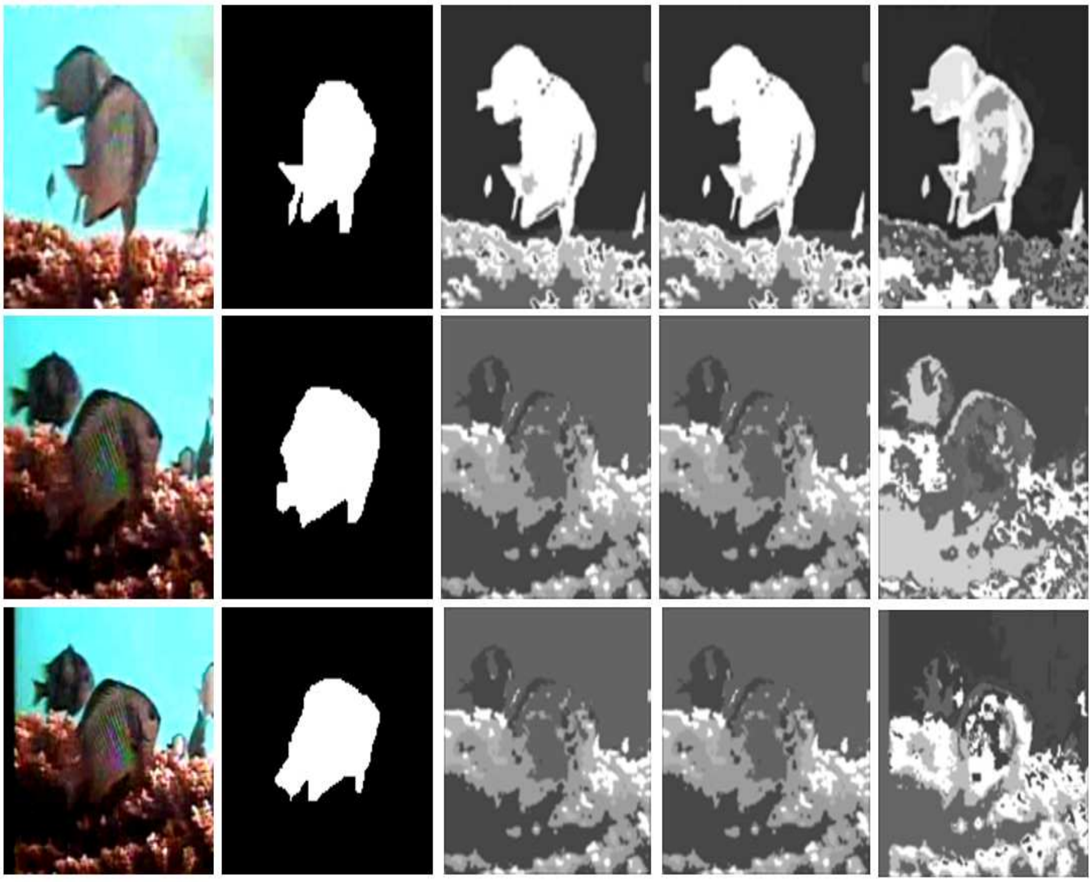

FIGURE 4.20: Results comparison -1 Images with complex background . Left to right: input image, Groud Truth image, Yue et al. [34] saliency Map, Singh [37] Saliency Map, Proposed Solution Saliency Map

The outputs from Yue et al.[34], Singh et al. [26] and proposed solution are compared with respective ground truth images. I have calculated the accuracy using equation 4.1 and error. The calculated accuracy and error values of each model are listed in table 4.3.

In figure 4.21 under water images with plane background are used for experiments.

In figure 4.22 under water images have more then one object and some overlapping features as well.

Table 4.3, the underwater images from Fish4Knowledge dataset were processed using these three models. We can observe that the accuracy of proposed model is higher then other two models.

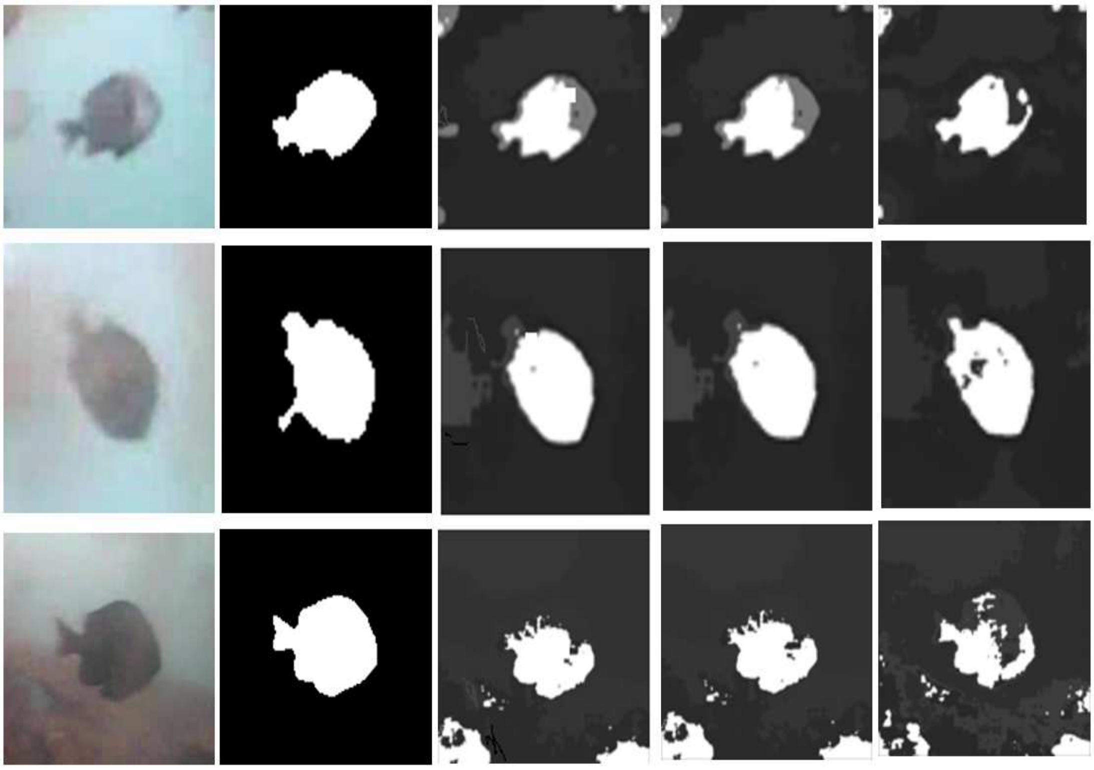

FIGURE 4.21: Results comparison -2 Objects with simple background . Left to right: input image, Groud Truth image, Yue et al. [34] saliency Map, Singh [37] Saliency Map, Proposed Solution Saliency Map

TABLE 4.3: Comparison of proposed solution with state-of-art [37] (Using CLAHE and Fuzzy C-mean) and [34] (Using K-mean and Active Contour Model for segmentation) techniques

	Accuracy	**Error**
Proposed Solution	94.24083	5.759167
Singh et al.[37]	84.13273	15.86272
Zhu et al.[34]	85.19857	14.80143

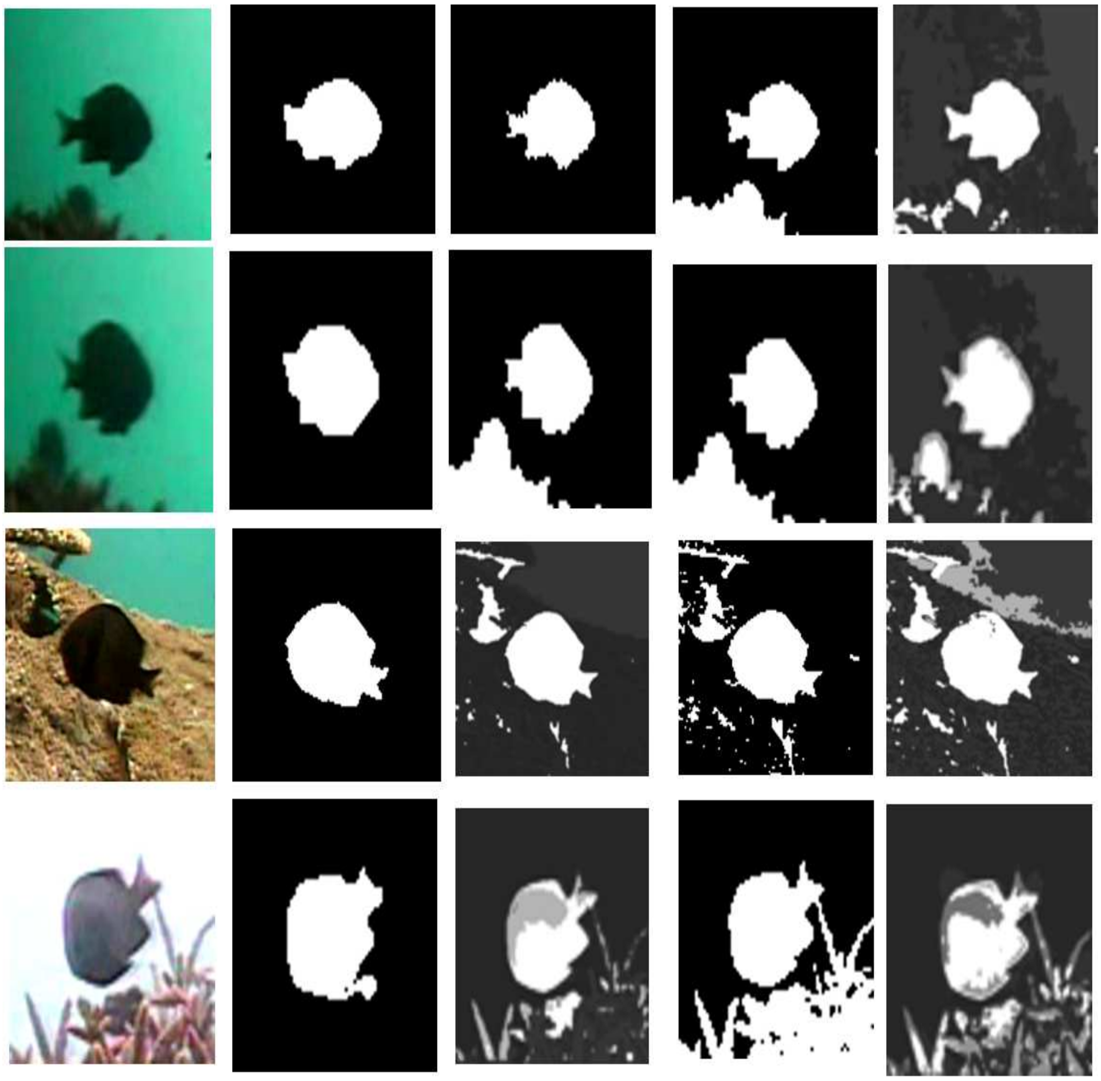

FIGURE 4.22: Results comparison -3 Objects with complex background and foreground. Left to right: input image, Groud Truth image, Yue et al. [34] saliency Map, Singh [37] Saliency Map, Proposed Solution Saliency Map

TABLE 4.4: K-mean, Fuzzy C-mean and Active Contour Model Clustering algorithm Comparison - Measure parameters are Accuracy and Efficiency

	Accuracy	**PNSR**	**Elapsed Time (sec)**
K-Means	90.99	63.86	0.20
Fuzzy C-Mean	83.41	60.56	0.22
Active Contour Model	39.84	56.10	0.35

Chapter 5

Conclusion and Future Work

In this work, underwater environment issues, challenges in underwater images processing and the SOD literature with respect to closely related areas is extensively reviewed. Identifying and extracting salient objects becomes very useful area. Objects in images automatically capture more attention at first glance than background items, such as cars, trees, and players. Therefore, if we can detect salient or important objects first, we can perform detailed reasoning and scene understanding in the next stage. Compared to traditional special-purpose object detectors, saliency models are general, typically fast, and do not need heavy annotation. These properties allow processing of a large number of images at low cost.

In this thesis, several state of the art image segmentation and underwater image enhancement methods are reviewed. In this work underwater images issues and challenges to detect objects in underwater images are discussed. A method composing of image enhancement and salient object detection steps is proposed. After comparing the performance of proposed method with state of art method using distorted images captured in the different water environments. Meanwhile, subjective visual evaluation index, several image segmentation quality metrics are utilized to evaluate and analyze the the proposed methods gains and drawbacks. In general a approach is presented to extract salient feature of underwater images that includes not only the three general steps of image segmentation, prior calculation, and prior combination, but also includes an image enhancement step comprising of color correct and contrast enhancement as pre-processing. The subjective, quantitative and qualitative experimental results show the specific features of poorly formed images, captured in the water.

5.1 Future Work

Future work includes exploring alternate pre-processing techniques, relying less on center bias, and incorporating color distribution. The major drawback to the k-means method is that its results are slightly random. Thus, a method less reliant on k-means would help provide more consistent results. Though humans often focus more towards the center of an image, there is no requirement that a salient object be in the center. Therefore a method less reliant on the center would be helpful for more applications. One way to use center bias less is to incorporate color distribution. Consequently, the next step would be to add the color distribution prior to the method discussed in this thesis. We can also enhance this technique for videos and co-saliency detection.

References

[1] Ming-Ming Cheng, Niloy J Mitra, Xiaolei Huang, Philip HS Torr, and Shi-Min Hu. Global contrast based salient region detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3):569–582, 2014.

[2] Federico Perazzi, Philipp Krähenbühl, Yael Pritch, and Alexander Hornung. Saliency filters: Contrast based filtering for salient region detection. In *2012 IEEE conference on computer vision and pattern recognition*, pages 733–740. IEEE, 2012.

[3] Yuan Gao, Miaojing Shi, Dacheng Tao, and Chao Xu. Database saliency for fast image retrieval. *IEEE Transactions on Multimedia*, 17(3):359–369, 2015.

[4] Xingxing Wei, Zhiqiang Tao, Changqing Zhang, and Xiaochun Cao. Structured saliency fusion based on dempster–shafer theory. *IEEE Signal Processing Letters*, 22(9):1345–1349, 2015.

[5] Meng Wang, Richang Hong, Xiao-Tong Yuan, Shuicheng Yan, and Tat-Seng Chua. Movie2comics: Towards a lively video content presentation. *IEEE Transactions on Multimedia*, 14(3):858–870, 2012.

[6] Sunhyoung Han and Nuno Vasconcelos. Image compression using object-based regions of interest. In *2006 International Conference on Image Processing*, pages 3097–3100. IEEE, 2006.

[7] Ueli Rutishauser, Dirk Walther, Christof Koch, and Pietro Perona. Is bottom-up attention useful for object recognition? In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, volume 2, pages II–II. IEEE, 2004.

[8] Roberto Gallea, Edoardo Ardizzone, and Roberto Pirrone. Physical metaphor for streaming media retargeting. *IEEE Transactions on Multimedia*, 16(4):971–979, 2014.

[9] Jianjun Lei, Cuicui Zhang, Yuming Fang, Zhouye Gu, Nam Ling, and Chunping Hou. Depth sensation enhancement for multiple virtual view rendering. *IEEE Transactions on Multimedia*, 17(4):457–469, 2015.

[10] Hongyuan Zhu, Fanman Meng, Jianfei Cai, and Shijian Lu. Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation. *Journal of Visual Communication and Image Representation*, 34:12–27, 2016.

[11] Xiu Li, Jing Hao, Hongwei Qin, and Liansheng Chen. Real-time fish localization with binarized normed gradients. In *OCEANS 2015-MTS/IEEE Washington*, pages 1–5. IEEE, 2015.

[12] Radhakrishna Achanta, Sheila Hemami, Francisco Estrada, and Sabine Süsstrunk. Frequency-tuned salient region detection. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, number CONF, pages 1597–1604, 2009.

[13] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (11):1254–1259, 1998.

[14] Yun Zhai and Mubarak Shah. Visual attention detection in video sequences using spatiotemporal cues. In *Proceedings of the 14th ACM international conference on Multimedia*, pages 815–824. ACM, 2006.

[15] Raimondo Schettini and Silvia Corchs. Underwater image processing: state of the art of restoration and image enhancement methods. *EURASIP Journal on Advances in Signal Processing*, 2010(1):746052, 2010.

[16] John Y Chiang, Ying-Ching Chen, and Yung-Fu Chen. Underwater image enhancement: using wavelength compensation and image dehazing (wcid). In

International Conference on Advanced Concepts for Intelligent Vision Systems, pages 372–383. Springer, 2011.

[17] Jasper RR Uijlings, Arnold WM Smeulders, and Remko JH Scha. The visual extent of an object. *International journal of computer vision*, 96(1):46–63, 2012.

[18] Ali Borji, Ming-Ming Cheng, Qibin Hou, Huaizu Jiang, and Jia Li. Salient object detection: A survey. *Computational Visual Media*, pages 1–34, 2014.

[19] Tao Xi, Wei Zhao, Han Wang, and Weisi Lin. Salient object detection with spatiotemporal background priors for video. *IEEE Transactions on Image Processing*, 26(7):3425–3436, 2016.

[20] Ali Borji, Dicky N Sihite, and Laurent Itti. What stands out in a scene? a study of human explicit saliency judgment. *Vision research*, 91:62–77, 2013.

[21] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (5):603–619, 2002.

[22] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282, 2012.

[23] Xiaodi Hou and Liqing Zhang. Saliency detection: A spectral residual approach. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. Ieee, 2007.

[24] Zhixiang Ren, Shenghua Gao, Liang-Tien Chia, and Ivor Wai-Hung Tsang. Region-based saliency detection and its application in object recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 24(5):769–779, 2013.

[25] Li Zhou and Zhaohui Yang. Salient region detection based on spatial and background priors. In *2014 IEEE International Conference on Information and Automation (ICIA)*, pages 262–266. IEEE, 2014.

[26] Anurag Singh, Chee-Hung Henry Chu, and Michael A Pratt. Multiresolution superpixels for visual saliency detection. In *2014 IEEE symposium on computational intelligence for multimedia, signal and vision processing (CIMSIVP)*, pages 1–8. IEEE, 2014.

[27] Huazhu Fu, Xiaochun Cao, and Zhuowen Tu. Cluster-based co-saliency detection. *IEEE Transactions on Image Processing*, 22(10):3766–3778, 2013.

[28] Lei Zhou, Yi Jun Li, Yi Peng Song, Yu Qiao, and Jie Yang. Saliency driven clustering for salient object detection. In *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 5372–5376. IEEE, 2014.

[29] Ruchira Manke, Suresh C Raikwar, and Anand Singh Jalal. A robust approach for salient region detection. In *2014 9th International Conference on Industrial and Information Systems (ICIIS)*, pages 1–4. IEEE, 2014.

[30] Na Tong, Huchuan Lu, Lihe Zhang, and Xiang Ruan. Saliency detection with multi-scale superpixels. *IEEE Signal Processing Letters*, 21(9):1035–1039, 2014.

[31] Rajkumar Kannan, Gheorghita Ghinea, and Sridhar Swaminathan. Salient region detection using patch level and region level image abstractions. *IEEE Signal Processing Letters*, 22(6):686–690, 2014.

[32] Pandivalavan Manipoonchelvi and Karuppiah Muneeswaran. Region-based saliency detection. *IET image processing*, 8(9):519–527, 2014.

[33] Lingfu Kong, Liangliang Duan, Wenji Yang, and Yan Dou. Salient region detection: an integration approach based on image pyramid and region property. *IET Computer Vision*, 9(1):85–97, 2014.

[34] Yue Zhu, Baochen Hao, Baohua Jiang, Rui Nian, Bo He, Xinmin Ren, and Amaury Lendasse. Underwater image segmentation with co-saliency detection and local statistical active contour model. In *OCEANS 2017-Aberdeen*, pages 1–5. IEEE, 2017.

[35] Yao Qin, Huchuan Lu, Yiqun Xu, and He Wang. Saliency detection via cellular automata. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 110–119, 2015.

[36] Yijun Li, Keren Fu, Lei Zhou, Yu Qiao, and Jie Yang. Saliency detection via foreground rendering and background exclusion. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 3263–3267. IEEE, 2014.

[37] Shivendra Singh, Manish Soni, and Ravi Shankar Mishra. Segmentation of underwater objects using clahe enhancement and thresholding with 3-class fuzzy c-means clustering. *International Journal of Emerging Technology and Advanced Engineering*, 4(4):798–805, 2014.

[38] Stephane Bazeille, Isabelle Quidu, Luc Jaulin, and Jean-Philippe Malkasse. Automatic underwater image pre-processing. 2006.

[39] Kashif Iqbal, Rosalina Abdul Salam, Azam Osman, and Abdullah Zawawi Talib. Underwater image enhancement using an integrated colour model. *IAENG International Journal of Computer Science*, 34(2), 2007.

[40] Codruta Orniana Ancuti, Cosmin Ancuti, Tom Haber, and Philippe Bekaert. Fusion-based restoration of the underwater images. In *2011 18th IEEE International Conference on Image Processing*, pages 1557–1560. IEEE, 2011.

[41] Xueyang Fu, Zhiwen Fan, Mei Ling, Yue Huang, and Xinghao Ding. Two-step approach for single underwater image enhancement. In *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pages 789–794. IEEE, 2017.

[42] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and

Karel Zuiderveld. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3):355–368, 1987.

[43] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980.

[44] Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4):427–437, 2009.