Chapter 1

Introduction

This chapter presents the introduction of the thesis. The background of salient object detection is discussed in Section 1.1. The motivation behind choosing the topic of research is explained in Section 1.2. Section 1.3 defines the problem statement and lists the thesis objectives. The contributions to the thesis are mentioned in Section 1.5.

1.1 Background

Humans tend to promptly identify and give attention to the vital region in visual representation. To model, the computer algorithm for such an astonishing ability, Salient object detection or Visual saliency evolved as a fundamental research domain in computer science from psychology, neurobiology, and cognitive science [1], [2], [3].

The formulation and computation of visual saliency are broadly classified into two models, namely, Salient Object Detection(SOD) [4], [5] [6] and Fixation Prediction(FP) [7], [8]. The first recognized model is Fixation Prediction(FP). The Fixation Prediction originated from psychology and cognitive science [9], [10]. The main objective of FP models is to predict the most vital part of the image without detecting the whole object. This model is a pioneer and initial platform in the field of saliency computation. The salient object detection is the second revolution from fixation prediction and recent a to the FP model. The main objective of this model is to detect the entire object with the exact boundary and replace the concept of segmentation and detection.

The most prominent, top-performing Heuristics models (non-deep SOD models [4], [5] based on one or two low-level handcrafted features(e.g., color contrast [11] , background prior [12]). These models gained popularity from 2009 to 2019, before the era of deep learning models. Although, It provides the platform for next-level saliency enhancement using deep learning-based models. These models improved the saliency gradually by using an over-segmentation process based on the regions [13], and super-pixels [14], or object proposals [15] to detect the exact salient objects with clear object boundaries. This model failed in images where the background is complex, cluttered, and similar to salient objects. These models also failed in structural, spatial, and regional similarities.

With the remarkable success and invention of deep learning technologies in computer vision, various deep learning-based SOD models have improved saliency since 2015.

Although, traditional heuristic-based models have various limitations and drawbacks to finding silent object detection accurately. These limitations are the motivational factor applying the Convolutional neural networks (CNNs) [16], [17]based learning method. CNN is successfully used in salient object detection. The characteristics of CNN, like multi-level and multi-scale features, accurately capture the most salient regions without using any prior knowledge like segment-level information. The advantage of using multi-level features is accurately locating the boundaries of the detected salient regions. The various algorithm, in series, is proposed to exploit the high-level semantic features and learn the essential saliency features from large data sets. These CNN-based models continuously refresh the records on almost all existing datasets and become current and mainstream solutions. The arrival of depth-sensing devices with high-resolution depth maps has opened the door to the real-time application of salient object detection in 3D applications. The depth maps have sufficient spatial information, providing the current research stream in saliency computations. The traditional and deep SOD models currently use Depth information from Depth maps to detect and locate the exact salient objects.

1.2 Motivation and Scope

The motivation behind proposing various SODs models is to propose the solutions to the long-prevailing issues of the complex and cluttered background image, where

objects and background are approximately the same. As time passes, more challenging Benchmark datasets like DUT-RGBD, LFSD, ReDWeB-S, and SIP have increased the complexity of simulating prominent objects in real-life situations. It is shown in Fig.1.1, where Initial, salient object detection focused on a mainly single parameter or feature to distinguish and separate the regions. Salient object detection has diverse applications. It is attracted the interest of researchers for the last

FIGURE 1.1: A set of images to demonstrate the challenge to extract salient objects in complex and cluttered backgrounds.

two decades in the field of computer vision. it is an intermediate or integral step in most of the complex vision-related applications like, Object Classification [18], Object Recognition [19], Image/Video Summarization [20], semantic segmentation [21],

Neurobiology of Attention [3], camouflaged object detection [22], and Dermoscopic Segmentation [23]. It has been used extensively in various robotic applications like, object discovery [24], [2] and human-robot interaction [25]. It is also used in various Graphics applications like non-photorealistic rendering [26] image retargeting [27] and image cropping [28] etc. All these applications are improving society by providing easy and exact solutions. The efficiency and accuracy of salient object detection will impact the above-mentioned and other related applications. This is our biggest motivation factors to improve the saliency and proposed more improved models.

1.3 Problem Statement

Salient object detection (SOD) aims to find the most prominent and conspicuous object(s) in a given image, similarly identified and detected by the human visual system. Detecting and localizing the salient object in the complex image is the most challenging task. The proposed framework in this thesis is utilized the probabilistic and deep learning-based approach to solve the challenging problem and improve the performance.

Study, implementation, and proposing new methods and models using probabilistic and Deep Learning-based approaches for detection and localization of complex Salient Objects, their experimental and comparative analysis.

The main objective of the research work is to design and develop efficient algorithms for complex salient object detection to detect and localize the precise and complete salient objects from clutter backgrounds by integrating various regional, spatial, color, and depth saliency.

1.4 The Research Objectives of Thesis

The robust models for accurately detecting and localizing the complex salient object in a challenging and cluttered background have been designed in this Thesis. The results analysis have been performed on globally available benchmark datasets with the possible evaluation parameters. The probabilistic model-based (probabilistic contrast PC) global topographical reference surface is proposed to achieve this objective. Further, these models are enhanced by proposing Edge Enhanced global reference surface and a global concave reference surface. This reference surface is used to integrate various handcrafted features to predict exact salient object detection. The capabilities of deep learning-based models are utilized to predict the more accurate and robust salient object detection. The objectives of the thesis are:

1. To perform an extensive study of the existing state-of-the-art methods for salient object detection.

2. To understand the issues and challenges and identify the research gaps in existing state-of-the-art methods models.

3. To propose a probabilistic model for complex salient object detection, implementation, and performance evaluation over benchmarks dataset.

4. To utilized the probabilistic model as global topographical reference surface for various regional saliencies integration for complex RGBD(3D salient object detection.

5. To propose deep learning-based approaches for complex salient object detection, its implementation, and performance evaluation over benchmarks dataset to address the limitation of the existing methods.

6. To propose a deep learning-based framework that utilizes the attention mechanism similar to the human for complex salient object detection.

1.5 Contributions to the Thesis

The main contributions to the thesis are as follows:

- The extensive literature survey of the salient object detection is described and analyzed. It includes the top-performing state-of-the-art methods of salient object detection together with limitations and challenges.

- The extensive literature survey of the benchmark Datasets and performance metrics used in this research domain is adequately described.

- Proposed and evaluated a probabilistic model for salient object detection using probabilistic contrast (PC). The various regional, color, and spatial saliency are computed integrated into Probabilistic Contrast. The regional saliencies are computed superpixel-wise.

- Proposed and evaluated the global topographical saliency(GTS) using global contrast and iterative Laplacian of Gaussian-based initial reference surface. This global topographical saliency is used as a reference plane for integrating regional saliencies. The regional saliencies are incorporated into the boundary enhanced global topographical saliency to improve the substantial information about the object. A boundary-based Gaussian weighted background suppression model is used to remove the background and edge effects.

- Proposed and evaluated the global concave reference surface (GCS) for RGBD saliency by addition of DoG-based contour and improved probabilistic contrast. This surface integrates spatial, regional, color, and depth saliencies to minimize the interior saliency discrepancy.

- Proposed and evaluated a CNN-based deep learning model of three independent networks dedicated to color, depth, and cross-modality to extract multi-stage raw saliencies. The stage-wise non-complementary, Complementary features are integrated into the late fusion strategy to predict the final salient object.

- Proposed and evaluated a CNN-based deep learning model with a composite backbone network(Architectural innovations) using an attention mechanism to produce enhanced encoded features, purified deep localized features. These improved complementary features fuse in a self-learning-based dense decoder to produce an exact salient object in a challenging and complex image.

1.6 Outline of the Thesis

The outline of the thesis is as follows:

Chapter 1 introduces the topic of the thesis. It describes the motivation and problem statement of the thesis. Finally, the thesis objectives and contributions to the thesis are summarized.

Chapter 2 presents the detailed literature survey of state-of-the-art methods to identify the limitations and challenges of existing models to establish the direction of this research work.

Chapter 3 presents a novel probabilistic contrast and an edge-enhanced global topographical surface to provide an initial reference surface containing the salient object. The various super-pixel-based regional saliencies are integrated into this global reference surface to pop out the complete salient object.

Chapter 4 presents a global topographical and Concave reference surface of complex salient object detection. The reference surface is used for various regional saliencies integration. The impact of the reference surface has been evaluated.

Chapter 5 presents a CNN-based deep learning based model, "CSA-Net", with three-stream network to produce raw saliencies and Nonlocal Network based deep localized features. These raw saliencies/side-outputs are combined to produce the final salient object. The detailed comparison with state-of-the-art methods is shown.

Chapter 6 presents a deep CNN composite backbone to extract enhanced encoded features. An attention mechanism based on improved deep localized features is produced. A self learning-based dense decoder is proposed to integrate these features.The extensive results analysis and stage-wise validation with state-of-the-art methods are shown.

Chapter 7 concludes the thesis and summarizes main findings of the work done. This chapter also proposes some possible future perspectives of the thesis.