Chapter 7

Conclusion and Future Work

This chapter concludes the models presented in this thesis. It also suggests the future works that can be explored in this area.

7.1 Conclusions

Salient object detection has dominated the research domain for the last two decades, and thousands of papers have been proposed. Excellent success has been achieved by traditional and probabilistic-based methods from 2007 to 2017, followed by deep learning-based methods. This domain has the objective to predict accurate localized and exact salient object that is either directly utilized to solve various prevailing problems in computer vision and image processing or taken as input to further process for complex applications. The success of 3D imaging, robotic applications, automatic vehicle negations, and other profound domain applications depend on the exact localization and prediction of prominent objects. There has been continuous research in this field, yet none of the methods are competent to extract the salient object in all cases and complex scenarios. Some of the ways fail to identify heterogeneous objects in a heterogeneous environment. Another issue and challenge for failure is the similarity of the background to the salient object. The deep learning-based methods achieved tremendous success, although they still do not produce generalized solutions. The main issue of this research domain is the point correspondence problem because the z- parameter(distance) in the actual scene is lost during image accusations. Various approximations based or stereo depth sensing camera-based depth map gives another dimension in this domain to produce more robust solutions. Several methods and models have been proposed in this thesis to counter these issues.

Chapter 1: introduced the field of salient object detection. It explained the motivation behind opting for this research domain. The problem statement and contributions to the thesis are also introduced.

Chapter 2: discusses the theoretical foundation of salient object detection. The historical background and literature survey in this research domain have been adequately investigated to determine the issues and challenges in the past decade. The benchmark databases and the performance metrics are also described in detail.

Chapter 3: proposed two probabilistic models for salient object detection. The first model proposed a novel method for salient object detection using a probabilistic approach using Poisson distribution to generate an initial saliency. This surface is used to integrate the regional depth, color, and spatial-based saliency. Generalization can easily be achieved by using probabilistic distribution. The second model is based on a robust and novel global topographical surface (GTS), which is used as a global reference plane. The proposed global reference plane works efficiently and effectively in integrating regional saliencies. The robustness of GTS makes it suitable for structure, shape, and border preservation in saliency estimation. The probabilistic approaches are applied to produce more robust and generalized solutions.

Chapter 4:used an additional parameter depth to increase the robustness of 3D saliency detection in complex clutter backgrounds. This method prepares an innovative and robust global concave topographical surface (GCS) approach for regional features integration. This surface is designed with the difference in Gaussian-based contours. So, this reference plane is used to minimize the border region's discrepancies. This integration works efficiently and effectively in regional saliencies integration to reduce the interior, exterior, and regional saliency discrepancies. The robustness of GCS increases the preservation of the structure, shape, and border-related information in saliency estimation. These regional saliencies integrations remove the interior saliencies discrepancies. Gaussian weighted background estimation and central saliency integration remove the exterior saliencies discrepancies. The traditional and probabilistic-based models have limitations because they are not based on learning the features.

Chapter 5: proposed a deep learning-based model $CSA - Net$ with 2×3 encoder and decoder to produce a Holistic Feature Space. The Feature space includes all essential features, which are utilized to generate modality-specific and cross-complimentary fusion-based saliencies. An innovative three-stream decoder based on a two-stage encoder efficiently locates the salient object with the exact object border and removes the background. The creative, progressive learning is designed to generate these features into three-stream decoder networks, and the two independent fusion strategy remarkably improves the performance.

Chapter 6: proposed a novel mutual attention-based distinguish window, MADW, to enhance encoded features. A multi-stage mutual attention map-based encoder could be integrated to produce enhanced encoded and deep global localized feature maps to address the challenge of a complex image, low depth map, and complex backgrounds. It is proposed to obtain better feature representation. This attention map is formulated with spatial, channel-wise, mutual, and feature-level attention mechanisms. Furthermore, the deep global localized feature map is the fusion process in the decoder to localize the salient object using the proposed attention map MADW. It is used as a reference plane to distinguish the salient and non-salient regions during the fusion process in SDD. Thus, the boundaries of detected objects could be better preserved.

7.2 Future Work

Despite the continuous research and tremendous success in this field, significant issues exist in this field. The definition and applications of visual saliency are highly subjective and diverse. A more precise understanding of the human attention mechanism and cognizance can clear the existing ambiguity. Various algorithms give gray-scale results, which make them difficult to evaluate. Most existing databases, like MSRA10K, ECSSD, NLPR, NJUD, and DES, are simple, while others are very complex, like PASCAL and LFSD. All the existing 3D datasets have inaccurate, unstructured, and low-depth issues. Most of the 3D datasets have a maximum of 1000 depth maps. Therefore a large-scale data set with good quality depth maps are needed to map for better solutions. Another objective should be incorporated in designing a database that captures nearly all the varieties of images, which can help create better models. Heterogeneity in objects is still not entirely resolved. Studying mechanisms for better separation of background and foreground in cases where they are similar will benefit military and wildlife applications.

Deep learning methods are mainly based on learning and testing. So a large scale and standard training set are needed, making the homogeneous testing environment for better analysis of improvements. The existing models are mainly based on the existing standard backbone and pretend models. Therefore, a composite backbone specific for salient object detection incorporating an attention mechanism similar to the human attention mechanism is needed. Some models must be developed that incorporate the benefits of probabilistic and deep learning models. Such hybrid models can be time-efficient as deep learning methods require much time for training.

The proposed method for salient object detection can be extended and improved using cross-complementary features extracted from better deep learning models. The deep learning models can be developed by utilizing some prior information from traditional or probabilistic models. The salient object detection should be applied directly to formulate medical, image annotation, and robotics applications for better features to be extracted and superior models of multi-label classification and prediction.