International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 11 Nov. 2022, pp: 617-619 www.ijaem.net ISSN: 2395-5252

Accurate and Robust Saliency Detection

Pedditi Satvika[1], Parankusham Vaishnavi[2], Mamindla Mounika[3], Dr. K. Srujan Raju[4]

[1,2,3]*B.Tech Student, Dept.Of CSE, CMR Technical Campus,Medchal,Hyderabad.*
[4] *Professor,Dept. Of CSE, CMR Technical Campus,Medchal,Hyderabad.*

Submitted: 15-11-2022 Accepted: 25-11-2022

ABSTRACT: Saliency Detection is a image processing activity that deals with detection and identification of main point of interest in a picture. Even though this action is very basic to the human eye, it is very complex to the machines.Our project aims to design a simple and powerful deep network architecture for saliency detection i.e., U2-Net. The architecture of the U2 -Net is a two-level nested U-structure. The architecture helps us to train a deep network from scratch without using pre-trained backbones from image classification tasks. It is trained from scratch to achieve competitive performance. U2 -Net architecture allows the network to delve deeper and attains relatively high resolution and, without substantially increasing the memory and estimated cost. This is achieved with the help of nested U-structure. Residual U- Block (RSU) module is present on the bottom level which will extract multi-scale features from the image without significantly downgrading the feature map resolution. U-Net like structure is present on the top level in which every stage is filled with a RSU block.

KEYWORDS:U2- Net, Saliency Detection, RSU, U- Net, architecture.

I. INTRODUCTION

Saliency Detection focuses on segmenting the most prominently attractive objects in an image. It is broadly used in many areas, like visual tracking and image segmentation. Lately, with the evolution of deep convolutional neural networks (CNN), specially with the elevation of Fully Conventional Neural Networks (FCN) in image processing, saliency detection has been enhanced substantially.

Ourmain idea is to propose a simple network architecture, called U2 -Net. U2- Net is a stratifiedstructure of U- Net, with two connected in series. CNN models usually use pre-trainedclassification models by ImageNet for Backbone. But we proposed a model structure optimized for the background separation task, and the performance is comparable to state-of-the-art models.It mostly consists of two paths. First is the encoder path and the other one is the decoder path. Encoder path conquer the context of the image and hence producing the feature maps. Encoder path is just a sheaf of convolution layer and max pooling layer. Decoder path enables formal localization using transposed convolutions. U-net consists of Convolutional layers and do not contain any Dense layer due to which it can accept images of all sizes.

II. LITERATURE SURVEY

There is a usual pattern in the design of most saliency detection networks. That is, it concentrates on making good use of deep features extracted from existing backbones, such as Alexnet , VGG , ResNet , ResNeXt and DenseNet, etc. However, all these backbones wereinitially designed for image classification. They extract features that convey semantic meaning rather than local details and global contrast information that are necessary for saliency detection. It also requires pre-training on ImageNet data which is data-inefficient mainly if the target data follows a different distribution than ImageNet.

The proposed system achieves saliency detection without the help of the pre-trained backbones. It allows the network to delve deeper and attain high resolution, without substantially increasing the storage and computational cost.

III. PROPOSED METHOD

The proposed method uses multi-scale feature extraction which is trained from scratch without using the pre-trained backbone models such as VGG, Alex-Net, Res-Net, etc. It eliminates the process of using patches to withdraw contextual features and instead of it, it follows a technique called pyramid scheme. It is developed by the inspiration from U-Net, using Residual U-block. RSU extracts multi-scale features from the image.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 11 Nov. 2022, pp: 617-619 www.ijaem.net ISSN: 2395-5252

The acquired features are down-sampled so as the contextual features get confined so that at the end, our output does not get affected by the captured fine details of the picture i.e., color, texture, etc.After the features are captured, they are encoded and decoded after up-sampling.

IV. METHODOLOGY

Saliency detection requires both local and global features. A 3×3 filter is good for extracting local features ineach layer. However, it is difficult to extract global information by simply increasing the filter size as it willincrease the number of parameters and computation costsdramatically.

Multi-scale feature extraction aims at designing new modules for extracting both local and global details from features obtained from backbone networks., Almost all of the saliency detection methods tryto make better use of the feature maps generated by the existing image classification backbones. Instead of developing and adding more complex modules and strategies to exploit the capabilities of these backbones, we proposed a novel and simple architecture, which directly extracts multi-scale featuresincrementally, for saliency detection.

V. ARCHITECTURE

This project is totally based on identifying the important or particularly noticeable features in an image. The model is built for saliency detection which is an internal functionality of robotics. The model is built using U2-Net architecture which is based on U-Net machine learning model. The data from the image is collected and preprocessed. Then features are extracted from the image based on pixels and standardization is performed.

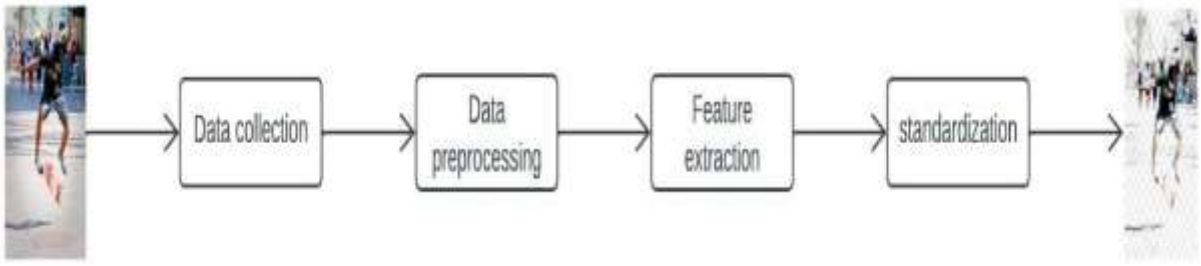
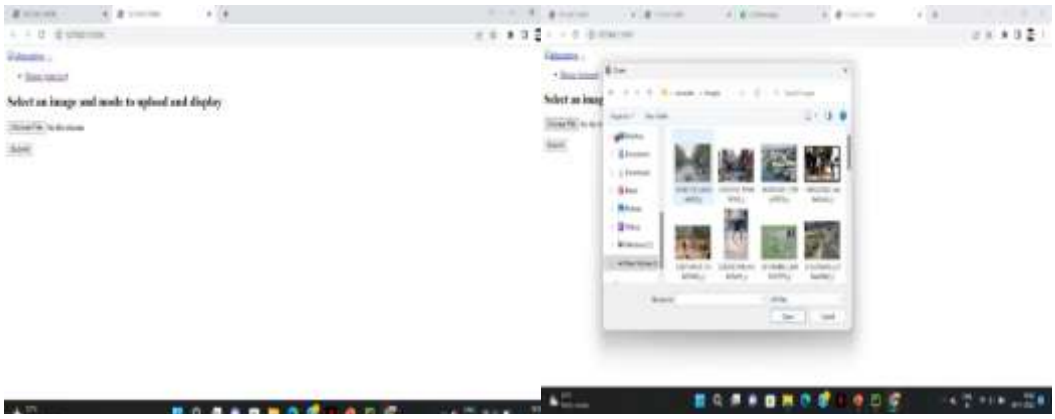


Figure 1: Architecture of accurate and robust saliency detection.

VI. RESULTS



Screenshot 1:User page.Screenshot 2:Selecting a picture

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 11 Nov. 2022, pp: 617-619 www.ijaem.net ISSN: 2395-5252

Screenshot 3:Uploading the picture Screenshot 4: Salient image with transparent Background.

VII. CONCLUSION

The main idea is to propose a simple deep network architecture called U2 - Net, for saliency detection. The main architecture of our U2 -Net is a two-level nested U Structure. The nested U structure with our redesigned RSU blocks enables the network to capture richer local and global information from both shallow and deep layers regardless of the resolution. Compared with those saliency detection models built on the existing backbones, our U2 -Net is built only on the proposed RSU blocks so that it can be trained from scratch and has different model sizes according to the constraints of the target environment. We provided a full size U2 - Net (176.3 MB, 30 FPS) and a smaller size version U2 –Net (4.7 MB, 40 FPS) in this project.

Experimental results on six datasets for saliency detection shows that both models perform very competitively against other 20 state-of-the-art methods, both qualitatively and quantitatively.

Despite the fact that our models attain ambitious results against other state-of-the-art methods, faster and smaller models are required for computation and memory limited devices, such as mobile phones, robots, etc. In the near future, we will explore different techniques and architectures to further improve the speed and reduce the model size. Moreover, training a more accurate and robust model requires a larger and more diverse dataset on prominent objects.

VIII. ACKNOWLEDGEMENT

We thank CMR Technical Campus for supporting this paper titled with “ACCURATE AND ROBUST SALIENCY DETECTION”,which provided good facilities and support to accomplish our work.Sincere thanks to our Chairman,Director,Deans,Head of the Department,Department of Computer Science and Engineering,Guide and Teaching and Non-Teaching faculty members for giving valuable suggestions and guidance in every aspect of our work

REFERENCES

[1]. S. Li, M. Xu, Y. Ren, and Z. Wang, “Closed-form optimization on saliency-guided image compression for hevc-msp,” IEEE Transactions on Multimedia, 2017.

[2]. M. Ding and R.-F. Tong, “Content-aware copying and pasting in images,” The Visual Computer, vol. 26, no. 6-8, pp. 721–729, 2010.

[3]. Y. Fang, Z. Chen, W. Lin, and C. Lin, “Saliency detection in the compressed

[4]. domain for adaptive image retargeting,” IEEE transactions on image processing, vol. 21, no. 9, pp. 3888–3901, 2012.

[5]. H. Wu, G. Li, and X. Luo, “Weighted attentional blocksfor probabilisticobject tracking,” The Visual Computer, vol. 30, no. 2, pp. 229–243, 2014.