Video Salient Object Detection Via

Fully Convolution Networks

Nitika Aggarwal A20427230

Karan Bhatiya A20424290

Paper Details:

Title: - Video Salient Object Detection Via Fully Convolution Networks

Author – Wenguan Wang, Jianbing Shen, Ling Shao **Publication Year -** February 2017

Problem Statement:

Detecting and segmenting salient objects in video, often referred to as video salient object detection, which has attracted a lot of interest in Computer Vision. While many models have been proposed and implemented but the deep understanding of achievements is still lacking. However, this project proposes a deep learning model to effectively detect salient regions in video. We can make use of this model in various tasks such as object detection, video summarization, etc.

Approach:

As described in this paper, our main aim is to detect the most salient object and segmenting the accurate region of the object in video. We all know that Convolution Neural Networks have been utilized in many computer vision regions such as object detection, semantic segmentation. Here too, we used CNN for another computer vision task which is **Video Salient Object Detection**.

We are implementing the video saliency object detection with the help of popular Caffe library and open source framework for training and testing the CNN model. We will build two networks which are as follows:

1. Static Video Saliency Network

2. Dynamic Video Saliency Network

For both the network we will be using stochastic gradient descent (SGD) and a polynomial learning policy.

Team Member's responsibility:

Each of us will implement one CNN network model. Please find the below stages assigned:

Nitika Aggarwal – Static Video Saliency Network.

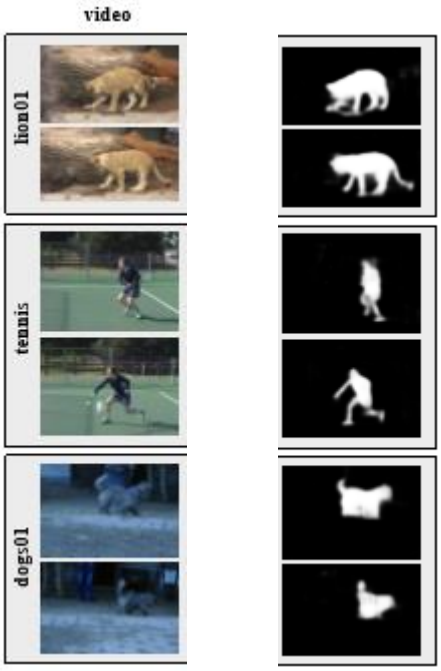
Karan Bhatiya – Dynamic Video Saliency Network.

Finally, both of us together will test and compare our results with the results of paper in OpenCV.

Data:

The performance of our project is reported using **Freiburg-Berkeley Motion Segmentation (FBMS)** dataset and **Densely Annotated Video Segmentation (DAVIS)** dataset. For **training** the CNN model we use two large image saliency datasets: **MSRA10K** and **DUT-OMRON**.

The below example shows the result of the video salient object detection obtain by the above paper.



References:

1. Video salient object detection via fully convolutional networks, IEEE Trans. on Image Processing, 27(1):38-49, 2018 By Wenguan Wang and Jianbing Shen and Ling Shao

2. https://arxiv.org/pdf/1702.00871.pdf

3. https://arxiv.org/pdf/1411.5878.pdf