

## Lab- 11

### CSET340- Advanced Computer Vision and Video analytics

#### Task-1:- Visual saliency detection, Unsupervised image segmentation, graph cut.

**Visual saliency detection** – Take an image and a short video stream and detect the salient region in them using any of the saliency techniques/algorithms (preferably OpenCV algos).

<https://pyimagesearch.com/2018/07/16/opencv-saliency-detection/>

**Unsupervised image segmentation** - Take any landscape image and perform k-means clustering on it for image segmentation.

<https://www.v7labs.com/blog/image-segmentation-guide>

<https://nayakpplaban.medium.com/image-segmentation-using-opencv-39013013920a>

**Graph Cut:-** Take an image and perform graph based image segmentation using graph cut algorithm. Also, show the intermediate graph formed with nodes and edges and edge weights.

<https://github.com/Sri-Sai-Charan/Graph-Cut-Segmentation>

#### Task-2:- Synthetic Human Action Video Generation using GANs on UCF101 human activity dataset.

- Choose two action classes (e.g., "JumpingJack" and "PushUps").
- Pick 5 short clips (3–5 seconds) from each class (total: 10 videos).
- Resize videos to consistent resolution, e.g., 64x64 or 128x128.
- Convert videos into frame tensors (e.g., shape:  $T \times H \times W \times C$ ).
- Normalize pixel values to [0, 1] or [-1, 1].
- Align frame counts using trimming or interpolation (e.g., 16–32 frames).
- Store them in an organized folder structure with sub-folder for each activity type.
- Model any of the following GAN for Video Generation i.e. **general GAN**, **TGAN** (Temporal GAN), **Video-GAN** or **styleGAN**

<https://github.com/bogireddytejareddy/gan-video-generation/blob/master/README.md>

<https://github.com/NVlabs/long-video-gan>

<https://github.com/amunozgarza/tsb-gan>

#### Output:-

- 1 new fake video (16–32 frames) generated per action class.
- Final format: .gif, .avi, or sequence of images.
- Min Epochs = 50 epochs.

**UCF101** is a widely-used human action recognition dataset with **13,320 videos** from **101 action categories**, such as "Basketball", "Skiing", "Clapping", etc.

Link to dataset:- <https://www.crcv.ucf.edu/data/UCF101.php>

<https://www.kaggle.com/datasets/pevogam/ucf101>

<https://www.kaggle.com/datasets/matthewjansen/ucf101-action-recognition>

Each video typically contains a single human performing one action.

Format: **.avi** videos at ~25 fps, various resolutions.

**Note:-** Submit the notebook file on LMS and Git link.

- Marks will be deduced for late submission.