TITLE:

Transformer-Based Visual Segmentation

MEMBERS:

1.

Mukhil Charles T

2310110528

[mt555@snu.edu.in](mailto:mt555@snu.edu.in)

2.

Jeevithesh R

2310110489

Jr479@snu.edu.in

3.

Nandan Menon

2310110532

[nm392@snu.edu.in](mailto:nm392@snu.edu.in)

ABSTRACT:

Visual segmentation partitions images, into multiple meaningful segments. Transformers which originated from NLP ,contain significantly advanced vision tasks , which include segmentation. We are focussing on transformer based segmentation methods, covering architectures, challenges faced, and future directions.

Paper(s):

Title: Transformer Based Visual Segmentation

Conference: IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE

INTELLIGENCE(ITPAMI)

Year of publication: 2024

Motivation:

Ever since CNNs have dominated visual segmentation, this paper is a case of when, transformers have outperformed them. Through this, we get to see how transformers simplify model architectures, and work better for segmentation.

Practical applications include: Tumor and organ segmentation.

Object segmentation in real-time(for Augmented Reality based systems)

Scene capturing( understanding) for AI-based robots would be benefitted.

Project Deliverables:

Inputs:

Visual Data: Images, Videos

Labels for supervised learning

Pre-Trained models and Datasets : Mask2Former, COCO

Hyperparameters:

Learning rate, and other parametes.

Outputs:

Segmented image as output (semantic segmentation)

Object masks

Object and background segmentation

Contribution of each member:

Mukhil Charles.T

* Research: study transformers based segmentation techniques, summarize research output, and justify the model selection
* Model Selection and training
* Deployment

Jeevithesh R

* Data Collection & Preprocessing
* Model Inference & Optimization: convert trained models to ONNX
* Result analysis : interpret model performance and suggest improvements/optimization

Nandan Menon

* Performance Evaluation with performance metrics such as iOu, PQ and AP.
* Building a User Interface: Develop a CLI application for testing segmentation results
* Implement an **API (Flask, FastAPI)** for real-time processing.