# Self Supervised Learning for Endoscopic Image Segmentation

*End Semester Project Review*

**Prankur Shukla**

**Roll No.: EC21B1074**

under the supervision of
**Dr. J. Umarani & Dr. T. Sreenath Reddy**

Department of Electronics & Communication Engineering
Indian Institute of Information Technology
Design and Manufacturing, Kancheepuram, Tamil-Nadu-600127, India

May 8, 2025

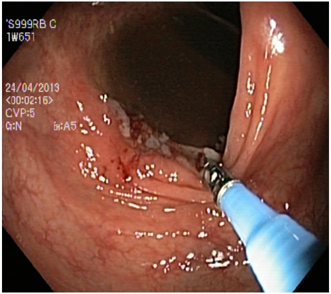
## Seminar Outline

1. Introduction
   - Motivation
   - Objective
2. Selected Literature Survey
3. Proposed Work
   - Work Methodology
   - Experimental Results and Analysis
   - Conclusion & Future Work

# Background

- Medical Image Segmentation plays a vital role in AI-assisted diagnosis, robotic surgeries, and surgical navigation.
- Deep Learning-based segmentation models like U-Net and DeepLabV3+ improve accuracy but rely on large labeled datasets.
- Challenges: Annotation is expensive and time-consuming.
- Precise segmentation of instruments can enhance tasks such as tumor detection, brain segmentation, disease diagnosis, and surgical planning

## Challenges in Endoscopy




Motion artifacts and occlusions.
Variable lighting conditions.
Limited annotated datasets.

# Motivation

- **Without Endoscopic Instrument Segmentation:**
  - **Surgeons face difficulties** in accurately tracking and identifying instruments during minimally invasive procedures.
  - **Increased risk of errors** due to occlusions, motion blur, and poor visibility.
  - **Manual intervention** required for real-time corrections, slowing down surgeries and increasing fatigue.
  - **Limited AI-assisted** guidance, affecting precision in robotic surgeries.

# Objective

- Develop a **Self-Supervised Learning (SSL)** framework for endoscopic instrument segmentation.
- Implement **SimCLR and MoCo** contrastive learning techniques to learn feature representations from unlabeled data.
- Fine-tune **U-Net and DeepLabV3+** models with limited labeled data (50%) and compare their performance with fully supervised models.
- Evaluate segmentation accuracy using **Dice Similarity Coefficient** (DSC), **Intersection-over-Union** (IoU), and Accuracy.
- Demonstrate that SSL-based models can achieve near-supervised performance while reducing annotation costs

# Selected Literature Survey I

- **Early Methods:** Rule-based techniques such as thresholding and edge detection were used for medical image segmentation but struggled with low-contrast images and motion artifacts.
- **Machine Learning Methods:** SVMs and Random Forests were early machine learning techniques but required extensive feature engineering, making them less suitable for complex endoscopic images.
- **Deep Learning Revolution:**
  - U-Net: Widely used, but struggles with small objects like surgical instruments due to limited multi-scale feature extraction.
  - DeepLabV3+: Introduced Atrous Spatial Pyramid Pooling (ASPP), improving accuracy for complex structures like surgical instruments.

# Selected Literature Survey II

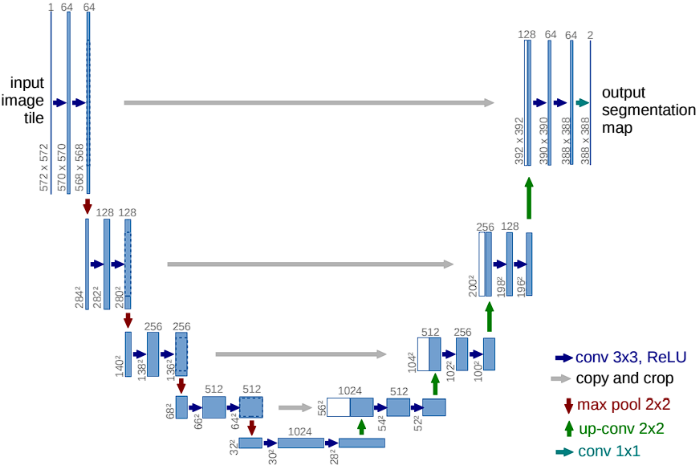



Figure: U-Net Architecture

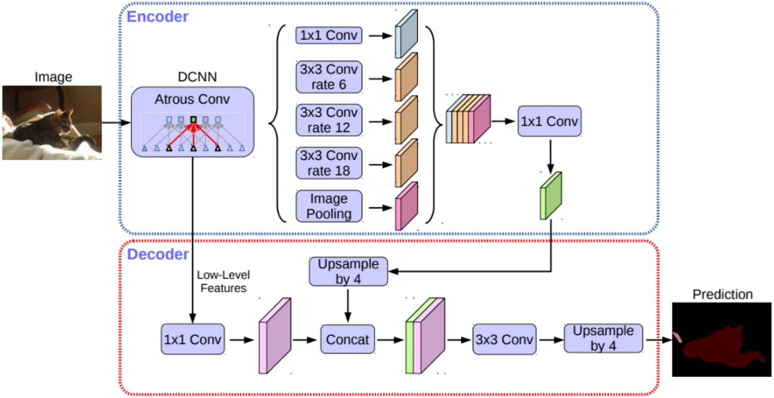



Figure: DeeplabV3+ Architecture

# Contrastive Learning-Based SSL

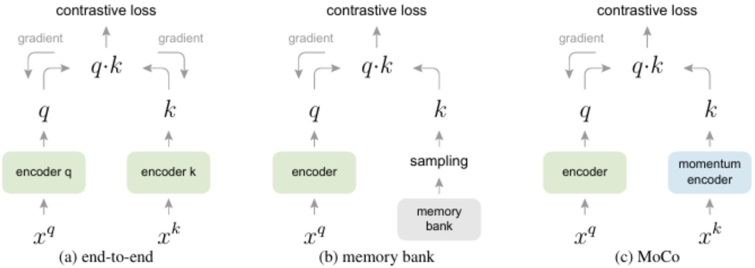



Figure: SimCLR & MoCo

# Methodology I

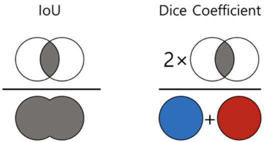
- **Dataset:**
  - Kvasir-Instrument dataset, containing 590 annotated endoscopic images of surgical tools.
  - Focus on instrument segmentation in endoscopic images, which includes challenges like motion artifacts and occlusions.
- **Proposed Deep Learning Model Structure:**
- U-Net & DeepLabV3+ are evaluated under two training paradigms:
  - Fully supervised learning (100% labeled data)
  - SSL pretraining followed by fine-tuning on 50% labeled data.

**Segmentation Metrics:**

- Dice Score (DSC)
- Intersection over Union (IoU)

# Methodology II

- Accuracy




**Goal:**

- Evaluate if SSL-pretrained models can match the segmentation performance of fully supervised models while reducing annotation dependency.

# Methodology III

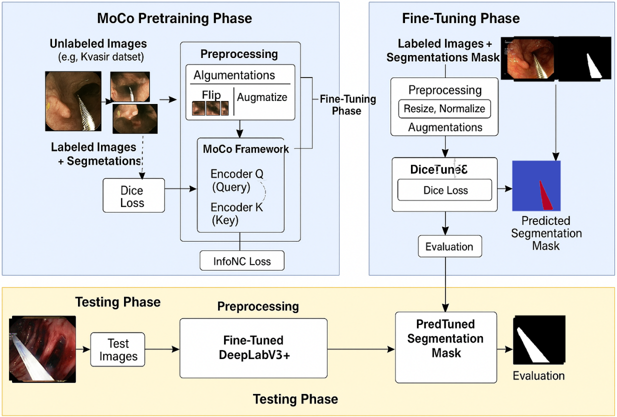



Figure: Workflow Diagram of model architecture

# Experimental Results & Analysis I

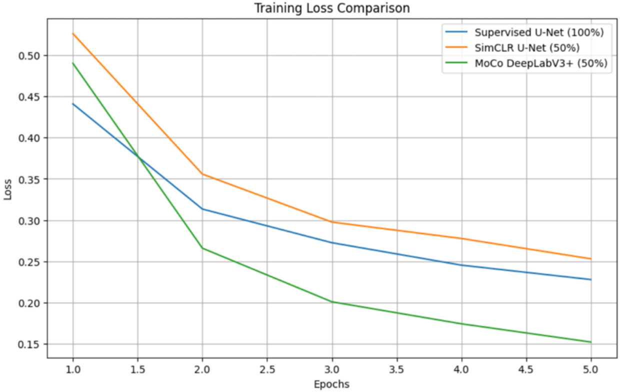



Figure: Training Loss and Convergence Analysis

Table: Segmentation Performance of Different Models

| Model (Pretraining) | DSC Score | IoU Score |
|---|---|---|
| U-Net (No SSL) | 0.82 | 0.75 |
| DeepLabV3+ (No SSL) | 0.89 | 0.80 |
| **DeepLabV3+ (MoCo + Fine-Tuning)** | **0.82** | **0.76** |

Table: Comparison with State-of-the-Art Methods

| Method | DSC Score | IoU Score |
|---|---|---|
| dJha et al. (2020) (U-Net) | 0.91 | 0.85 |
| Keprate et al. (2021) (U-Net) | 0.80 | 0.73 |
| **Ours (DeepLabV3+ + MoCo)** | **0.82** | **0.76** |


Figure: Qualitative results: Ground truth vs. predicted segmentation masks.

# Conclusion & Future Work I

- Study explores the effictiveness of SSL techniques in endoscopic image segmentation.
- Result obtained showed that even with 50% of labelled data SSL is acheving alomost similar accuracy to that of supervised models.
- Future research should focus on scaling SSL techniques to larger, multi-center datasets to improve model generalization across diverse surgical environments.
- Domain adaptation techniques should be explored to ensure robust segmentation across different surgical settings without re quiring extensive labeled data.

# References

Ronneberger, O., Fischer, P., Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A. L. (2017). *DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs*. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Chen, T., Kornblith, S., Norouzi, M., Hinton, G. (2020). *A Simple Framework for Contrastive Learning of Visual Representations (SimCLR)*. International Conference on Machine Learning (ICML).

He, K., Fan, H., Wu, Y., Xie, S., Girshick, R. (2020). *Momentum Contrast for Unsupervised Visual Representation Learning*. IEEE Conference (CVPR).

Jha, D., Smedsrud, P. H., Riegler, M. A., Halvorsen, P., de Lange, T., Johansen, D., Johansen, H. D. (2020). *Kvasir-Instrument: Diagnostic and Therapeutic Tool Segmentation Dataset in Gastrointestinal Endoscopy*. International Journal of Computer Assisted Radiology and Surgery.

Thank You