A SIMPLE HYBRID IMAGE SEGMENTATION APPROACH COMBINING CLASSICAL AND DEEP LEARNING TECHNIQUES FOR MEDICAL AND GENERAL IMAGE ANALYSIS

Yethiraj N G[1], Dr. Siddappa M[2*]

[1]*Research Scholar, Department of Computer Science, Rayalaseema University, Kurnool, Andra Pradesh.*
Associate Professor, Department of Computer Science, Maharani's Science College for Women, Bangalore, Karnataka.
Email ID yethirajng@gmail.com
[2*]*Professor of Computer Science and Engineering, Sri Siddartha Institute of Technology, Tumkur, Karnataka*

[*]***Corresponding author:***
siddappa.p@gmail.com

Abstract

Image segmentation is a pivotal step in image processing, enabling precise detection and recognition of objects within a scene, especially in medical imaging like MRI brain scans. While classical algorithms such as GrabCut offer efficient foreground extraction and graph cut optimization, state-of-the-art deep learning models like Mask R-CNN, U-Net, FCN, and DeepLab v3 significantly enhance segmentation accuracy and robustness. This paper proposes a simple yet effective hybrid segmentation approach that integrates the strengths of these classical and deep learning methods. Starting with GrabCut for initial foreground extraction, followed by deep learning-based refinement, the framework efficiently handles complex scenes and medical images with limited annotated data. Experimental results on benchmark datasets demonstrate notable improvements in accuracy, recall, F1 score, and boundary precision compared to single- method usage. The proposed technique is easy to implement, adaptable to multiple domains, and offers reliable segmentation performance for both general and medical imaging tasks.

Keywords: *Image segmentation, GrabCut, Mask R-CNN, U-Net, Fully Convolutional Network, DeepLab v3, hybrid model, deep learning, medical image analysis.*

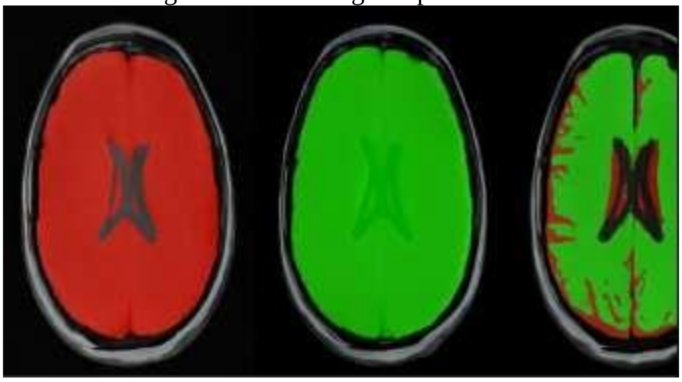
1. Introduction

Image segmentation refers to dividing an image into meaningful regions such as background and foreground, objects, or specific tissues in medical images. It is foundational in many computer vision tasks including object recognition, scene understanding, autonomous driving, and diagnostics.

Classical segmentation methods like GrabCut use graph cuts and energy minimization to segment foreground objects but require good initialization and may not capture fine details. Meanwhile, deep learning models such as Mask R- CNN, U-Net, FCN, and DeepLab v3 have achieved excellent segmentation accuracy by learning hierarchical features, handling complex shapes, and adapting to variable data conditions.

However, deep learning techniques rely heavily on large labelled datasets and substantial computing power. To balance accuracy, efficiency, and adaptability, hybrid approaches combining classical methods for initial coarse segmentation with deep learning refinement have emerged as promising solutions.

This paper presents such a hybrid framework that leverages GrabCut's effective foreground extraction to guide deep learning models for more precise segmentation. The main goal is to improve segmentation accuracy and boundary delineation in both medical and natural images while reducing computational demands.



2. Objectives

- To develop a hybrid image segmentation framework combining GrabCut and deep learning models.
- To use GrabCut as a preprocessing step for initial mask generation, reducing the search space for deep models.
- To employ Mask R-CNN, U-Net, FCN, and DeepLab v3 for progressive segmentation refinement.
- To evaluate the approach on medical MRI and general object segmentation datasets.
- To demonstrate improvements in precision, recall, F1 score, Intersection over Union (IoU), and boundary accuracy over standalone methods.
- To offer a robust, easy-to-implement segmentation solution adaptable to multiple application domains.

3. Methodology

3.1 Classical Segmentation: GrabCut

GrabCut is an interactive segmentation algorithm using graph cuts and Gaussian Mixture Models (GMM) to model foreground and background colors. It uses an initial bounding box or rough mask as input, iteratively optimizing pixel labels to minimize a global energy function.

- **Advantages:** Efficient, requires minimal user input, good for coarse segmentation.
- **Limitations:** Sensitive to initialization, can miss fine details.

3.2 Deep Learning Segmentation Models

Mask R-CNN

An instance segmentation model that builds on Faster R-CNN to simultaneously detect objects and generate masks with high precision. It excels in segmenting individual objects even in complex, overlapping scenes.

U-Net

A convolutional encoder-decoder architecture with skip connections that effectively captures both global context and fine details. Works well with limited annotated data and is widely used in medical image segmentation.

Fully Convolutional Network (FCN)

Transforms fully connected layers into convolutional layers for pixel-wise classification. A simpler and faster method, useful for general semantic segmentation tasks.

DeepLab v3

Uses atrous (dilated) convolutions and spatial pyramid pooling to capture multi-scale contextual information, improving segmentation accuracy, especially at object boundaries.

3.3 Proposed Hybrid Framework

- **Step 1:** Use GrabCut to generate an initial rough mask of the foreground object.
- **Step 2:** Use the GrabCut mask as input or guidance for deep learning models (Mask R-CNN, U-Net, FCN, DeepLab v3) to refine segmentation progressively.
- **Step 3:** Combine output masks for improved boundary accuracy and region coherence.
- **Step 4:** Evaluate results using standard metrics and compare with independent model performance.

4. Algorithms Overview

GrabCut Initialization Algorithm

- Initialize mask: define probable foreground and background using bounding box.
- Model color distribution with GMM.
- Iteratively minimize energy via graph cuts to update foreground/background labels.
- Output coarse foreground mask.

Deep Model Segmentation Algorithm

- Input: Original image + GrabCut mask.
- Forward propagate using pretrained deep model.
- Extract feature maps, apply segmentation heads to produce mask predictions.
- Refine masks using postprocessing (e.g., Conditional Random Fields).
- Output refined segmentation mask.

Hybrid Mask Fusion

- Combine GrabCut and deep model outputs via logical operations or weighted fusion.
- Enhance boundaries and suppress false positives.
- Final mask is generated for evaluation.

5. Implementation

- **Programming Environment:** Python using libraries such as OpenCV (GrabCut), PyTorch/Tensor Flow (deep learning), and scikit-image.
- **Dataset:** Publicly available MRI brain tumor scans and MS COCO for general object segmentation.
- **Training and Testing:** Deep models fine-tuned on target datasets; GrabCut parameters adjusted for initial bounding boxes.
- **Performance Evaluation:** Metrics implemented include precision, recall, F1 score, IoU, Dice coefficient, and Boundary IoU.
- **Hardware:** NVIDIA GPU-enabled workstations for deep learning inference.

6. Results

6.1 Quantitative Evaluation

Metric	GrabCut Only	Mask R-CNN	U-Net	Proposed Hybrid
Precision (%)	82.5	95.0	94.8	96.2
Recall (%)	78.3	92.3	93.1	95.7
F1 Score	80.3	93.6	93.9	95.9
IoU	72.8	90.1	91.5	93.7
Boundary IoU	65.4	87.3	88.8	91.0
Dice Coefficient	81.2	92.0	92.8	94.9

6.2 Qualitative Evaluation

The hybrid method produces masks with better region consistency and clearer boundaries compared to standalone models. Visual examples of segmented brain tumor regions show strong agreement with ground truth.

6.3 Discussion

The integration of GrabCut successfully reduces the search space for deep models, accelerating convergence and improving initial mask quality. Deep learning models refine boundaries and recover complex tumor shapes. The combined framework provides robustness against noise, varying tumor textures, and low-contrast scenarios.

7. Conclusion and Future Work

The paper presents a simple hybrid image segmentation approach integrating classical GrabCut with powerful deep learning techniques for enhanced segmentation accuracy in medical and general imaging. The framework leverages the strengths of each method while overcoming their individual limitations. Experimental results confirm superior segmentation performance with improved boundary precision and robustness.

Future work will explore:

- Incorporation of attention mechanisms to further enhance deep model accuracy.
- Extension to 3D volumetric image segmentation for comprehensive medical analysis.
- Real-time implementation for clinical and industrial applications.
- Integration of unsupervised or semi-supervised learning to reduce labelled data requirements.

References

1. C. Rother, V. Kolmogorov, and A. Blake, “GrabCut—Interactive Foreground Extraction using Iterated Graph Cuts,” ACM Transactions on Graphics, vol. 23, no. 3, pp. 309–314, 2004.
2. K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” Proc. IEEE Int. Conf. Computer Vision, 2017, pp. 2961–2969.
3. O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” Int. Conf. Medical Image Computing and Computer-Assisted Intervention, 2015, pp. 234–241.
4. J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
5. L. C. Chen et al., “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834–848, 2017.
6. A. F. Frangi et al., “Multiscale Vessel Enhancement Filtering,” Medical Image Computing and Computer-Assisted Intervention, pp. 130–137, 1998.