Data-Driven Active Snakes with Reinforcement Learning for Image Segmentation

Rishi Upadhyay[1], Eshanika Ray[1], Nikhil Isukapalli[1], and Jakob Reinwald[1]

[1]**Department of Computer Science, University of California, Los Angeles**

ABSTRACT

Automated image segmentation is crucial in various applications, including medical imaging. Traditional active contour models, or "snakes," are powerful tools for image segmentation that evolve contours towards object boundaries by minimizing an energy function. However, classical snakes often suffer from limitations such as local minima traps, initialization sensitivity, and tedious manual parameter tuning. While deep learning has significantly advanced computer vision, it typically requires large, manually annotated datasets, which are time-intensive and intricate to create, especially in medical contexts.
This project proposes a novel approach to address these limitations by developing an active snake controlled by a neural network trained using Reinforcement Learning (RL). This data-driven approach aims to move beyond fixed, hand-crafted energy functions and enable the snake to learn adaptive, context-aware deformation strategies. We also explore imitation learning, training our RL snake to emulate a traditional snake to accelerate basic movement learning. This framework has the potential to provide more robust and intuitive segmentation tools that can leverage global image context in addition to local features. We intend to apply these techniques to instance segmentation tasks on traditional images, with potential extensions to medical images.

Keywords: Active Contours, Snakes, Reinforcement Learning, Image Segmentation, Deep Learning, Weak Supervision

INTRODUCTION

Automated image segmentation holds significant clinical utility in various domains, particularly in medical imaging. However, progress in this area often faces challenges, including the demanding need for high-quality ground-truth annotations. Manual annotation, while possible, is elaborate and time-intensive, thereby limiting large-scale studies. Deep learning methods have generated considerable recent success in computer vision, but require a significant amount of labeled samples, which presents a hurdle.

Active contour models, often referred to as "snakes," represent a classic yet powerful approach to image segmentation, dating back to 1988. These deformable splines evolve towards object boundaries by minimizing an energy function, balancing internal forces (for smoothness) and external forces (attraction to edges). While providing direct, smooth object boundaries, classical snakes are prone to limitations such as trapping in local minima, sensitivity to initial conditions, and the need for meticulous manual parameter tuning.

To overcome the data annotation bottleneck for deep learning and the inherent limitations of traditional active contour models, this project introduces a novel learning framework. We propose to develop an active snake controlled by a neural network trained using Reinforcement Learning (RL). This approach aims to leverage RL's adaptive decision-making capabilities to enable the snake to learn complex deformation strategies directly from experience, moving beyond fixed energy functions. The RL agent, embodied by a neural network, will output dynamic offsets for the snake's control points, guided by reward signals derived from traditional snake energy functions. This novel integration promises more robust and intuitive segmentation tools capable of utilizing both local features and global image context. Furthermore, we will explore imitation learning as a strategy to bootstrap the RL agent's learning by mimicking the behavior of a traditional snake, initially focusing on instance segmentation on general images.

RELATED WORK

The evolution of image segmentation methods, particularly those involving deformable models, has seen a significant shift from classical variational approaches to modern deep learning paradigms.

Classical Active Contours

The concept of active contours, or ”snakes,” was pioneered by Kass et al. (1988). These methods model segmentation as an energy minimization problem, where a deformable spline (the snake) evolves under the influence of internal forces (promoting smoothness and continuity) and external forces (attracting the snake to image features like edges). Mathematically, the energy function for a snake is defined as:

$$E_{\text{snake}} = \int_0^1 (E_{\text{internal}}(v(s)) + E_{\text{image}}(v(s)) + E_{\text{constraint}}(v(s)))ds$$

where $v(s)$ represents the contour, E_{internal} ensures smoothness, E_{image} attracts the snake to image gradients, and $E_{\text{constraint}}$ can incorporate user-defined forces or other high-level constraints. Pioneering methods primarily utilized deformable models such as snakes or level set methods to segment structures in 2D and 3D medical images. Despite their foundational importance and ability to yield smooth object boundaries, classical snakes are known to suffer from several limitations: susceptibility to local minima, sensitivity to initial contour placement, and the requirement for meticulous, manual tuning of parameters (e.g., weights for internal and external forces).

Modern Active Contours

Specific advancements in active contours include Deep Active Contours by Rupprecht (2016), which use CNNs to predict vector fields to guide contour evolution Rupprecht et al. (2016). More recently, Deep ContourFlow (2024) focuses on unsupervised learning approaches ??? (2024). The overarching trend has been moving from hand-crafted energy functions to learned features. But as we’ll see, there’s still a gap in adaptive decision-making that our reinforcement learning approach aims to fill.

DeepSnake

DeepSnake, proposed by Peng et al. (2020), bridges classical and modern active contour approaches. This method employs a neural network to iteratively deform an initial contour until it matches the object boundary. Essentially, it ”implements classic snake algorithms with a learning-based approach”. A key innovation in DeepSnake is the use of ”circular convolution for cycle-graph structure”, which effectively exploits the cyclic nature of closed contours for better feature learning. The pipeline operates in two stages: initial contour proposal followed by iterative contour deformation. Circular convolution processes features along the contour, predicting point movements, and this process repeats until convergence. DeepSnake demonstrates a performance of ”32.3 fps on 512x512 images”.

However, DeepSnake has fundamental limitations that motivate further research. It employs a ”fixed deformation strategy”, meaning the CNN applies the same learned transformation pattern regardless of the specific object or image. Consequently, there is ”no adaptation mechanism” for the system to adjust its strategy based on image complexity or challenging boundaries. Furthermore, DeepSnake’s decision-making involves limited global reasoning, primarily relying on local, feature-based decisions along the contour.

Reinforcement Learning in Image Segmentation

Reinforcement Learning (RL) trains a neural network agent to operate in an environment using reward signals to maximize cumulative reward. Early applications of RL in medical image segmentation, such as by Sahba et al. (2006), formulated segmentation as a control problem, utilizing Q-learning for optimal thresholds and post-processing parameter optimization. However, a crucial gap remains in the literature: no one has directly applied RL to control snake evolution, with most existing RL work focusing on parameter selection or region proposal. This project aims to bridge the gap in direct RL control over snake evolution by integrating DeepSnake’s contour representation with RL.

METHODOLOGY

RL Environment Design

Observation Space

Action Space

Reward Design

Episode Termination

Neural Network Architecture

Imitation Learning

Hybrid Model

EXPERIMENTS

Black-and-white Shapes

Real-world Image Datasets

Medical Images

RESULTS

DISCUSSION

CONCLUSION

ACKNOWLEDGMENTS

REFERENCES

??? (2024). Deep contourflow: Unsupervised learning of contour evolution for image segmentation. *???* TODO.

Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331.

Peng, Z., Han, M., Liang, R., Hu, X., Gao, P., Huang, W., Cai, J.-F., and Tan, T. (2020). Deepsnake: A neural network for contour deformation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5687–5696.

Rupprecht, C., Lensch, H. P., and Cremers, D. (2016). Deep active contours. In *German Conference on Pattern Recognition (GCPR)*, pages 155–166. Springer.

Sahba, M., Hamarneh, G., and Abugharbieh, R. (2006). Reinforcement learning for medical image segmentation. *IEEE transactions on medical imaging*, 25(6):706–719.