INDIAN INSTITUTE OF ENGINEERING SCIENCE AND TECHNOLOGY, SHIBPUR

A Term
Paper On

A Deep Reinforcement Learning Based Dynamic Approach For Cancer Detection with Stage

NAME: ANWESH KUILA

M.TECH, 1ST YEAR (2ND SEMESTER)

ROLL NO: 2024ITM002

Under the supervision of
Dr. Santi Prasad Maity

DEPARTMENT OF INFORMATION TECHNOLLOGY

CERTIFICATE

This is to certify that the work presented in this term paper entitled "A Deep Reinforcement Learning Based Dynamic Approach For Cancer Detection with Stage", submitted by Anwesh Kuila, having the examination roll number 2024ITM002, has been carried out under my supervision for the partial fulfilment of the degree of Master of Technology in Information Technology during the session 2024-26 in the Department of Information Technology, Indian Institute of Engineering Science and Technology, Shibpur.

(Dr. Santi Prasad Maity)
Professor

Department of Information Technology
Indian Institute of Engineering Science and Technology, Shibpur

(Dr. Tuhina Samanta)
Head of the Department

Department of Information Technology
Indian Institute of Engineering Science and Technology, Shibpur

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere thankfulness and deep regard to Dr. Santi Prasad Maity, for his impeccable guidance, valuable feedback, and constant encouragement throughout the duration of the project. His valuable suggestions were of tremendous help in preparing the term paper. Working under him was an extremely knowledgeable experience for me.

I would like to convey my gratitude to the department of Information Technology, our respected and honourable head of the department of Information Technology, Dr. Tuhina Samanta, and to our Central Library for proving the necessary resources. I would also like to give my sincere gratitude to all my friends and seniors for their valuable cooperation. Without their help this study would have been incomplete.

Date:

Anwesh Kuila

Roll No: 2024ITM002

Department of Information Technology,
Indian Institute of Engineering Science
and Technology, Shibpur

ABSTRACT

Accurate detection and staging of cancer from medical imaging is crucial for effective treatment planning and prognosis. Traditional deep learning methods, particularly convolutional neural networks (CNNs), have shown remarkable success in static tumour segmentation tasks. However, they often fail to adapt dynamically to changing tumour characteristics across time-series data. In this study, we discuss the way to develop a novel framework that combines CNN-based medical image segmentation with Deep Reinforcement Learning (DRL) to detect cancerous regions and monitor their progression over time. The system learns a policy to refine segmentation results and classify cancer stages by interacting with sequential medical images as an evolving environment. Using spatial features from CNN-extracted tumour masks, a DRL agent is trained to make informed decisions about the tumour's development, leveraging temporal consistency across imaging sessions. This approach should enable the model to dynamically adapt to morphological changes in tumour shape, size, and location, which are often poorly captured in static models. Furthermore, we also discuss our further work which should provide better performance.

Index Terms - Reinforcement learning, deep convolutional neural networks, image segmentation, semantic segmentation.

Contents

Chapter 1

Introduction

1.1 Introduction to the problem

1.1.1 Objective

To develop an intelligent and adaptive medical image segmentation framework that combines Convolutional Neural Networks (CNNs) for feature extraction and spatial segmentation, with Reinforcement Learning (RL) for to track, stage, and predict cancer progression using sequential imaging data.
If cancer found the model should be capable of dynamically adjusting its analysis strategy based on changes observed across time-series medical images, ultimately supporting real-time clinical decision-making for oncology.

1.1.2 Motivation

Cancer diagnosis is not a one-time task - continuous monitoring of the disease's evolution is critical for staging, treatment planning, and prognosis. While CNNs are effective for static image segmentation, they are limited in capturing temporal changes. Reinforcement Learning offers the capacity to learn progression patterns and adapt decisions (e.g., classify cancer stage, assess spread) based on feedback from sequential inputs, making it a strong complement to CNN-based architectures.

1.2 Introduction to CNN

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed to process and analyze images. Their architecture mimics the visual perception mechanism of the human brain, enabling them to automatically learn features from pixel data. This makes CNNs particularly well-suited for medical image analysis, where tasks such as tumor detection, tissue classification, and organ segmentation require fine-grained, localized understanding of complex image structures.A CNN consists of stacked layers-convolutional layers for feature extraction, pooling layers for downsampling, and fully connected layers for classification.
In the context of cancer detection, CNNs have demonstrated state-of-the-art performance in identifying and segmenting malignant regions from images. By stacking

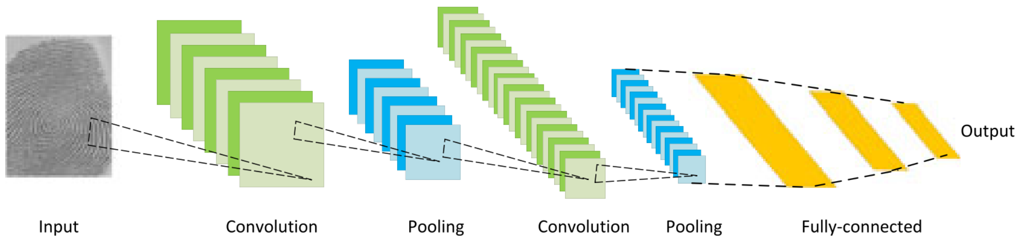

Figure 1.1: Architecture of CNNs. From [1]

convolutional layers with nonlinear activations and pooling operations, CNNs can extract deep semantic features based on edges, textures, color gradients, contours, etc. That distinguishes cancerous tissues from healthy ones.

1.3 Introduction to reinforcement learning

RL explicitly considers the problem of a goal-directed agent interacting with an uncertain environment. The learner is not told beforehand which actions to take; instead, it should discover which actions yield the most reward by trying them. The agent's objective is to maximize its reward and thus determine the best course of actions, or policy, to achieve that objective.

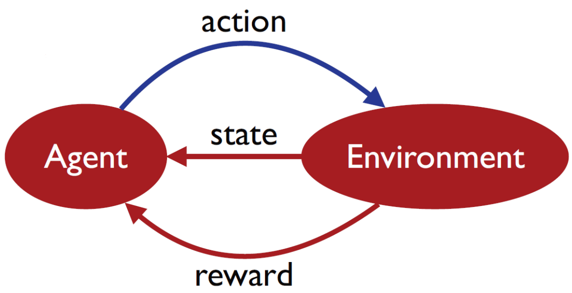

Figure 1.2: Representation of the general scenario of reinforcement learning.

1.4 Why Reinforcement Learning?

In reinforcement learning, guidance is provided not by a supervisor but through a reward-penalty system, which naturally encourages exploration over exploitation. This makes it particularly effective for problems involving dynamic and abstract environments, where the relationship between input features and the desired output is weak or unclear.

Our problem aligns well with these characteristics. The environment - in this case the cancerous region in medical images - evolves over time, and the model is expected to predict the stage of cancer as it progresses. Moreover, due to the high variability and lack of clearly defined size or shape of cancerous regions, it is difficult to rely on a conventional supervisory signal for training. Thus, reinforcement learning offers a fitting solution for such a complex and uncertain setting.

Chapter 2

Related works

2.1 Different types of segmentation techniques

Over the years, a wide variety of segmentation techniques have been developed. These techniques can be roughly classified as shown in the figure 2.1.

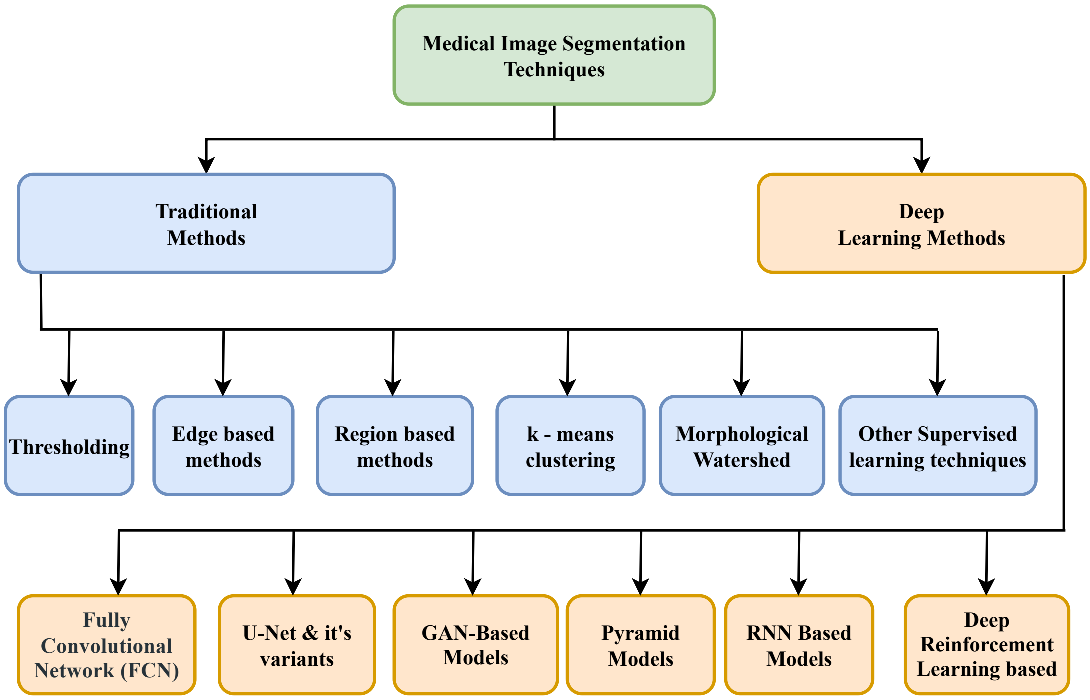

Figure 2.1: Different image segmentation techniques.

2.2 Deep Learning based models

Medical image segmentation using deep learning enables precise identification and delineation of anatomical structures or abnormalities, such as tumors, within medical scans like MRI, CT, or X-rays. By leveraging deep neural networks, particularly

CNNs, this approach enhances diagnostic accuracy and supports clinical decision-making. Key architectural trends include:

- **Fully Convolutional Networks (FCNs):** FCNs, introduced by [2] replace fully connected layers in traditional CNNs with convolutional layers to output a spatial segmentation map of the same size as the input image. This allows pixel-wise predictions for semantic segmentation. They modify existing CNN architectures like VGG16 or GoogLeNet by removing fully connected layers and adding upsampling layers. Skip connections are used to fuse features from earlier layers with deeper layers, improving segmentation accuracy.
  Drawbacks: Computationally expensive for real-time use, lacks global context, and struggles with 3D image generalization.

- **U-Net and its variants:** Models like U-Net [3] and SegNet [4] use an encoder to extract features and a decoder to up-sample for pixel-wise predictions.
  Drawbacks: Loss of fine-grained details due to resolution reduction during encoding.

- **GAN based models:** Generative Adversarial Networks (GANs) consist of a generator (produces segmentation masks) and a discriminator (distinguishes real from fake masks). [5] use adversarial training to refine semantic segmentation by training the generator to fool the discriminator. The generator learns to produce realistic segmentation maps by minimizing the discriminator's ability to detect fake outputs.
  Drawbacks: Training instability, mode collapse, and high computational requirements.

- **Multi-scale and pyramid models:** Paper [6] presents Feature Pyramid Networks (FPNs) and Pyramid Scene Parsing Networks (PSPNs) handle objects at different scales using hierarchical feature extraction.

- **Recurrent Neural Network (RNN) Based Models:** RNNs model sequential dependencies among pixels to capture global context. ReSeg [7] builds on ReNet, stacking RNN layers atop VGG16 features to encode spatial dependencies in four directions (horizontal/vertical, forward/backward), followed by up-sampling for segmentation.
  Drawbacks: Slower due to sequential processing, not easily parallelizable, and struggles with 2D image structures.

2.3 Deep Reinforcement Learning based models

In [8] presents RLSegNet. A U-shaped network designed for brain tumor segmentation on 2D MRI slices (BRATS 2015), integrating reinforcement learning (RL) to improve accuracy and efficiency over traditional 3D CNNs. It includes:

- Feature Extraction Network (FEN): A DenseNet-based encoder that extracts deep, multi-scale features using dense connections, bottlenecks, and transition layers.

- Mask Prediction Network (MPNet): Uses a Dueling DQN with GRU to predict the current mask from the previous frame via affine transformations (translation, scaling). RL defines state, action, reward (IoU), and termination for each step.

- Up-Sampling with Cascade Attention: A decoder with deconvolution layers, refined by attention from MPNet's mask to focus on tumor regions. A Middle Supervision Block (MSB) enhances gradient flow. The loss combines MSE, cross-entropy, and MSB loss.

Drawbacks: RLSegNet faces several limitations that impact its robustness and general applicability. Its accuracy on enhancing tumor regions is relatively low, as the model struggles to handle complex and irregular tumor shapes due to the restricted set of reinforcement learning actions. The approach heavily depends on the segmentation of the previous frame, making it prone to error propagation when earlier predictions are inaccurate. Additionally, the use of only basic affine transformations limits its ability to capture non-linear deformations in tumor boundaries. lastly, Its reliance on IoU-based rewards makes it sensitive to noise and artifacts present in MRI scans.

The paper [9] presents a deep reinforcement learning (DRL)-based approach to correct inaccurate bounding boxes in object tracking, addressing common issues like oversized, partial, or misaligned detections. By formulating the task as a Markov Decision Process, a Deep Q-Network (DQN) agent is trained to adjust bounding boxes through discrete actions such as movement and scaling. The model leverages VGG16 features and a memory of past actions, along with a custom reward strategy, to improve regression accuracy. Tested on the MOT 2015 dataset, the method outperforms conventional regression techniques in tracking accuracy metrics like MOTA and MT. textbfDrawbacks: It requires online fine-tuning for new targets, increasing computational demand, and can get stuck in local optima, occasionally needing random actions to recover. The fixed action space may limit flexibility in handling complex object deformations.

2.4 Challenges

Accurate cancer detection and staging through medical image segmentation face several complex challenges, especially when dealing with longitudinal data. These challenges arise from biological variability and data limitations.

- **High Variability in Tumor Appearance:** Tumors exhibit significant variation in size, shape, texture, and location across different patients. These characteristics can also change over time, making consistent detection and segmentation difficult.

- **Scarcity of Labeled Temporal Datasets:** There is a lack of annotated medical imaging datasets that span multiple time points and cancer stages. Labeling such data requires expert radiologists and is time-consuming.

- **Precise Tumor Tracking Across Time:** Ensuring accurate localization and segmentation of tumors in scans taken at different times or from slightly different angles poses a major challenge. Variations in imaging protocols and patient positioning add to the complexity.
- **Dynamic Cancer Staging:** Beyond static segmentation, the model must detect and interpret tumor progression, such as growth and local invasion. This requires temporal understanding and the ability to compare features across multiple scans.

Chapter 3

Future Work

3.1 3D conversion

In the future, we plan to enhance our approach by transforming 2D images into 3D representations through the incorporation of depth information. Utilizing 3D convolution enables more effective exploitation of spatial features, often leading to improved segmentation accuracy compared to traditional 2D convolution methods. However, this advantage comes with significant trade-offs-3D convolutional networks introduce a much larger number of parameters, resulting in increased computational complexity and longer processing times for segmentation.

3.2 Dynamic

Implementing a dynamic cancer detection model that can process input images over time to assess cancer progression and staging is challenging due to the lack of appropriately labeled datasets. In the future, we aim to address these challenges and develop a dynamic algorithm to achieve this goal.

3.3 Quantum CNN

With the emergence of quantum computing, we aim to transition from classical Convolutional Neural Networks (CNNs) to Quantum Convolutional Neural Networks (QCNNs), which offer several compelling advantages-

- QCNNs can efficiently handle high-dimensional data through quantum states, allowing for complex pattern representation with fewer parameters.
- By exploiting superposition and entanglement, quantum systems enable exponential parallelism, processing multiple states simultaneously and potentially accelerating computation. Offering a significant reduction in time complexity, processing an image with n pixels in $O(n)$ time, as opposed to the $O(n^2)$ time required by classical CNNs.
- Quantum operations naturally capture global correlations and entangled features, leading to enhanced feature extraction compared to classical methods.

Bibliography

[1] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[2] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.

[4] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[5] Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. Semantic segmentation using adversarial networks. *arXiv preprint arXiv:1611.08408*, 2016.

[6] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890, 2017.

[7] Francesco Visin, Marco Ciccone, Adriana Romero, Kyle Kastner, Kyunghyun Cho, Yoshua Bengio, Matteo Matteucci, and Aaron Courville. Reseg: A recurrent neural network-based model for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 41–48, 2016.

[8] Yi Ding, Xue Qin, Mingfeng Zhang, Ji Geng, Dajiang Chen, Fuhu Deng, and Chunhe Song. Rlsegnet: An medical image segmentation network based on reinforcement learning. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 20(4):2565–2576, 2022.

[9] Yifan Jiang, Hyunhak Shin, and Hanseok Ko. Precise regression for bounding box correction for improved tracking based on deep reinforcement learning. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1643–1647. IEEE, 2018.