

碩士學位論文

Cross-Modal 지식 증류와 Hybrid CNN-Transformer를  
활용한 향상된 의료 이미지 분할 기법

Advancing Medical Image Segmentation Using Cross-Modal  
Knowledge Distillation and Hybrid CNN-Transformer

韓國交通大學校 大學院

소프트웨어 學科

Saeed Ahmad

June, 2024

# Advancing Medical Image Segmentation Using Cross-Modal Knowledge Distillation and Hybrid CNN-Transformer

Saeed Ahmad

碩士學位論文

Cross-Modal 지식 증류와 Hybrid CNN-Transformer를  
활용한 향상된 의료 이미지 분할 기법

Advancing Medical Image Segmentation Using Cross-Modal Knowledge  
Distillation and Hybrid CNN-Transformer

指導教授 郭 政 桓

June, 2024

韓國交通大學校 大學院

소프트웨어 學科

Saeed Ahmad

Cross-Modal 지식 증류와 Hybrid CNN-Transformer를  
활용한 향상된 의료 이미지 분할 기법

Advancing Medical Image Segmentation Using Cross-Modal Knowledge  
Distillation and Hybrid CNN-Transformer

指導教授 郭 政 桓

이 論文을 工學碩士 學位論文으로 提出함

June, 2024

韓國交通大學校 大學院

소프트웨어 學科

Saeed Ahmad

Saeed Ahmad 工學碩士學位論文으로 提出함

June, 2024

審査委員長 오염덕 (印)

審査委員 우성희 (印)

審査委員 곽정환 (印)

韓國交通大學校 大學院

# Abstract

## Advancing Medical Image Segmentation Using Cross-Modal Knowledge Distillation and Hybrid CNN-Transformer

Medical image segmentation is a critical and laborious task that involves precisely delineating anatomical structures, organs, and abnormalities within medical imaging data. This process plays a fundamental role in diagnosis, treatment planning, and quantitative analysis of various medical conditions. Manual analysis of these images requires the expertise of the radiologist to detect the abnormality and segment the imagery. However, this approach is susceptible to human error and is inherently limited in scalability. Consequently, there is a compelling need for automated analysis of medical images, which would expedite the process and extend its accessibility and efficiency to a broader population. This research focuses on contributing to the development of efficient automatic brain tumor segmentation algorithms. It proposes a configured and optimized hybrid residual attention UNet (COHRA-UNet) for multimodal brain tumor segmentation and a novel multi-teacher cross-modal knowledge distillation (MTCM-KD) framework for unimodal segmentation.

The first study comprehensively examines multiple hyperparameters, including activation functions, normalization techniques, upsampling methods, and loss functions. We also conducted a detailed exploration of various attention mechanisms, encompassing spatial, channel, and transformer-based self-attention in diverse configurations, to evaluate their effects on the UNet architecture and identify their optimal settings for improving brain tumor

segmentation performance. Additionally, we examined different hybridizations of the transformer-based CNN architecture. The proposed COHRA-UNet integrates these optimal parameters and configurations, demonstrating superior performance for the given task and surpassing previous state-of-the-art methods.

In our second study, we developed the MTCM-KD framework to tackle the issue of unavailable MRI modalities in clinical settings. This framework is specifically designed to segment the  $T_{1ce}$  MRI sequence, which is frequently used in clinical settings and shares structural similarities with the  $T_1$  MRI scan and provides adequate information for effective segmentation. The framework integrates two distinct knowledge distillation (KD) strategies: 1) a performance-oriented response-based KD and 2) a cooperative deep supervision fusion learning (CDSFL). The former strategy adjusts confidence weights to each teacher model based on their performance, ensuring that the KD process is guided by performance. Moreover, the CDSFL module bolsters the learning capabilities of the multi-teacher models through facilitated mutual learning. Additionally, the combined knowledge from these models is distilled into the student model to enhance its deep learning supervision. Comprehensive testing on the BraTS datasets has shown that our framework delivers promising results in unimodal segmentation for both  $T_{1ce}$  and  $T_1$  modalities, surpassing earlier state-of-the-art techniques.

# Contents

Abstract	i
Table of Contents	iii
List of Tables	iv
List of Figures	v
1 Introduction	1
2 Related Work	2
3 Proposed Methodology	3
4 Experiments Evaluation and Analysis	4
5 Conclusion and Future Work	5
Abstract	6



## List of Tables

## List of Figures

# I. Introduction

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

## II. Related Work

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

### III. Proposed Methodology

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

## IV. Experiments Evaluation and Analysis

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

## V. Conclusion and Future Work

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

## Abstract

# Cross-Modal 지식 증류와 Hybrid CNN-Transformer를 활용한 향상된 의료 이미지 분할 기법

의료 영상 분할은 의료 영상 데이터 내의 해부학적 구조, 기관 및 이상을 정확하게 묘사하는 중요한 작업으로, 진단, 치료 계획 및 다양한 질병의 정량적 분석에 근본적인 역할을 한다. 이러한 이미지를 수동으로 분석하려면 방사선 전문의의 높은 전문 지식이 필요하지만, 이 방법은 인간의 오류에 취약하고 확장성에 한계가 있다. 따라서 의료 영상의 자동 분석에 대한 요구가 커지고 있으며, 이는 분석 과정을 가속화하고 접근성과 효율성을 증대시킬 수 있다. 본 연구는 효율적인 자동 뇌종양 분할 알고리즘 개발에 중점을 둔다. 이를 위해 다중 모드 뇌종양 분할을 위한 구성 및 최적화된 하이브리드 잔여 주의 UNet(COHRA-UNet)과 단일 모드 분할을 위한 새로운 다중 교사 교차 모달 지식 증류(MTCM-KD) 프레임워크를 제안한다.

첫 번째 연구에서는 활성화 함수, 정규화 기술, 업샘플링 방법, 손실 함수 등 다양한 하이퍼파라미터를 종합적으로 조사하였다. 또한, 공간, 채널 및 트랜스포머 기반 self-attention을 포함한 다양한 주의 메커니즘의 영향을 평가하고, UNet 아키텍처의 뇌종양 분할 성능을 향상시키기 위한 최적의 설정을 탐색하였다. 이 과정에서 트랜스포머 기반 CNN 아키텍처의 다양한 하이브리드화를 검토하였다. 제안된 COHRA-UNet은 이러한 최적의 매개변수와 구성을 통합하여 우수한 성능을 발휘하며, 기존의 최신 방법론들을 능가하는 결과를 보였다.

두 번째 연구에서는 임상 환경에서 사용 가능한 MRI 양식이 제한적인 문제를 해결하기 위해 MTCM-KD 프레임워크를 개발하였다. 이 프레임워크는 임상에서 자주 사용되는  $T_{1ce}$  MRI 시퀀스를 효과적으로 분할할 수 있도록 설계되었으며,



$T_1$  MRI 스캔과 구조적 유사성을 공유하여 충분한 정보를 제공한다. 프레임워크는 성과 중심 응답 기반 KD와 협력적 심층 감독 융합 학습(CDSFL)이라는 두 가지 고유한 지식 증류(KD) 전략을 통합한다. 성과 중심 응답 기반 KD 전략은 각 교사 모델의 성과에 따라 신뢰도 가중치를 조정하여 KD 프로세스를 성과에 맞게 조정한다. 또한, CDSFL 모듈은 상호 학습을 촉진하여 다중 교사 모델의 학습 능력을 강화한다. 이러한 모델들의 결합된 지식은 학생 모델에 증류되어 딥러닝 감독을 향상시킨다. BraTS 데이터 세트를 통한 포괄적인 테스트 결과, 우리 프레임워크는  $T_{1ce}$  및  $T_1$  양식의 단일 모드 분할에서 기존의 최신 기술을 능가하는 유망한 결과를 보였다.