A DUAL CONVOLUTIONAL NEURAL NETWORK WITH U-NET AND MULTI-RESIDUAL NETWORK FOR MEDICAL IMAGE SEGMENTATION

BY

SYED QAMRUN NISA

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

2024

A DUAL CONVOLUTIONAL NEURAL NETWORK WITH U-NET AND MULTI-RESIDUAL NETWORK FOR MEDICAL IMAGE SEGMENTATION

BY

SYED QAMRUN NISA

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy in Computer Science

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

JANUARY 2024

ABSTRACT

Medical image segmentation is a growing field in medical image analysis. The segmentation of diseases, organs, or abnormalities in medical images has become demanding for the analysis of medical images. It is an essential task in medical image analysis for the identification of disease, treatment planning, and observation. Segmentation is a challenging job because of the numerous artifacts present in the medical images. Over the past few years, convolutional neural networks (CNNs) have shown outstanding performance in the segmentation of images. U-Net is a convolutional neural network, that has achieved excellent performance in the segmentation of medical images. Despite the outstanding performance of the U-Net and variants of U-Net, these architectures have certain limitations and are not sufficient for the segmentation of medical images. U-Net has an overfitting problem, a vanishing gradient issue, and it has a small receptive field. This thesis proposes an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for medical image segmentation to improve the segmentation. The proposed architecture consists of two networks. One network is U-Net, which uses pre-trained Resnet_50 as an encoder and the second network is MultiRes U-net. The residual block in the proposed architecture solves the vanishing gradient issue. Atrous spatial pyramid pooling (ASPP) block is used in both networks that increase the receptive field to capture more information. The proposed architecture is evaluated on three medical datasets: gastrointestinal polyp dataset, brain tumor dataset, and dental dataset. The gastrointestinal polyp dataset and brain tumor dataset are publicly available open-access dataset and is used by researchers for the segmentation. The dental dataset is collected from the Oral Radiology Unit, Kulliyyah of Dentistry, International Islamic University Malaysia. Data augmentation is applied to the training set to increase the dataset which will help to reduce the overfitting problem and improve the performance of the segmentation. The result of a proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for medical image segmentation is compared to the existing architectures: FCN, U-Net, and Unet_Resnet. The performance of the proposed architecture is evaluated in terms of Dice Similarity Coefficient (DSC), Intersection over Union (IOU), Precision, and Recall. The experimental results show that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network achieves higher DSC, IOU, Precision, and Recall for the segmentation of the gastrointestinal polyp dataset, dental dataset, and brain tumor dataset and outperforms the existing segmentation architectures.

خلاصة البحث

تعد تجزئة الصور الطبية مجالا متناميا في تحليل الصور الطبية. أصبحت تجزئة الأمراض والأعضاء والشذوذ في هذا المجال أكثر تطلبا من أجل تحليل الصور الطبية. وهي مهمة أساسية في تحليل الصور الطبية للتعرف على الأمراض وتخطيط العلاج والملاحظة. كانت التجزئة وظيفة صعبة بسبب العديد من الآثار الموجودة في الصور الطبية. خلال السنوات القليلة الماضية، أظهرت الشبكات العصبية الإلتفافية (CNNs) أداء متميزا في تجزئة الصور. U-Net هي شبكة عصبية إلتفافية التي قد حصلت على أداء ممتاز في تجزئة الصور الطبية. وعلى الرغم من الأداء المتميز عند U-Net وأنواعها، لكن تلك البنية لها حدود معينة وليست كافية لتدقيق عمل تجزئة الصور الطبية. ويوجد عند U-Net مشكلة إفراط في التدريب وتدرج التلاشي وحقل استقبال صغير. طرحت هذه الرسالة بنية وهي شبكة العصبية الإلتفافية الثنائية (A Dual Convolutional Neural Network) التي تحتوي على U-Net و Multi-Residual Network من أجل تجزئة الصورة الطبية. وتهدف لتحسين التجزئة. تتكون هذه البنية من الشبكتين وهما U-Net التي تستخدم Resnet_50 مدربا مسبقا كالتشفير والشبكة الأخرى هي MultiRes U-Net. الكتلة المتبقية في البنية المقترحة تحلل مشكلة المشقة المتلاشية. يستخدم تجميع الهرم المكاني الافتراضي ASPP في هاتين الشبكتين اللتين رفعتا حقل الاستقبال لتحصيل المزيد من المعلومات. يتم تقسيم البنية المقترحة إلى ثلاث مجموعات من البيانات الطبية وهي: مجموعة بيانات البوليبات المعدية المعوية، ومجموعة بيانات الورم الدماغي ومجموعة بيانات الأسنان. تتوفر مجموعة بيانات البوليبات المعدية المعوية ومجموعة بيانات أورام المخ المفتوحة المصدر والمتاحة للجميع ويستخدمها الباحثون للتجزئة. وتم جمع مجموعة بيانات الأسنان من وحدة أشعة الفم، كلية طب الأسنان، في الجامعة الإسلامية العالمية بماليزيا. تتم تطبيق زيادة البيانات على مجموعة التدريب لإضافة البيانات التي ستساعد في تخفيض مشكلة الإفراط في التدريب وتحسين أداء التجزئة. تقارن نتيجة البنية المقترحة لصور التجزئة الطبية وهي الشبكة العصبية الإلتفافية الثنائية مع U-Net و Multi-Residual Network بالبنية الحالية وهي FCN، U-Net و Unet_Resnet. يقَيّم أداء البنية المقترحة من ناحية معامل دايس للتشابه (DSC)، و التقاطع على الاتحاد (IOU)، والدقة والاسترداد. وتبين نتائج التجربة للبنية المقترحة بأن الشبكة العصبية الإلتفافية الثنائية مع U-Net وMulti-Residual Network تحصل على أفضل DSC، IOU، ودقة والاسترداد لتجزئة مجموعة بيانات البوليبات المعدية المعوية، ومجموعة بيانات الورم الدماغي ومجموعة بيانات الأسنان وتتفوق على بنية التجزئة الحالية.

APPROVAL PAGE

The thesis of Syed Qamrun Nisa has been approved by the following:

____
Amelia Ritahani binti Ismail
Supervisor

____
Mohd. Adli bin Md. Ali
Co-Supervisor

____
Akram M Z M Khedher
Internal Examiner

____
Azlan Mohd Zain
External Examiner

____
Mohammad Naqib Eishan Jan
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Syed Qamrun Nisa

Signature .... Date ....

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

A DUAL CONVOLUTIONAL NEURAL NETWORK WITH U-NET AND MULTI-RESIDUAL NETWORK FOR MEDICAL IMAGE SEGMENTATION

I declare that the copyright holder of this thesis is Syed Qamrun Nisa.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.

3. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Syed Qamrun Nisa

……..…………………….. ………………………..

Signature Date

This thesis is dedicated to my beloved parents; Syed Nazir Ahmad Shah & Syed Firdousa Jan for their endless love, support, and encouragement

ACKNOWLEDGEMENTS

In the name of Allah (S.W.T.), the Most Gracious and Most Merciful, I humbly express my profound gratitude and reverence. All praise and gratitude are extended to the Almighty, the Creator of the universe, the Compassionate and Merciful, the Source of all knowledge and wisdom. Through His benevolence, I have been bestowed with the health, intellect, and capabilities to embark on this sacred academic journey. I also extend my heartfelt appreciation to my beloved Prophet (P.B.U.H) for serving as an everlasting source of inspiration and guidance throughout my life.

I would like to acknowledge the invaluable guidance and encouragement from my dedicated teachers. Their unwavering support has been pivotal in shaping my academic endeavors. Special gratitude goes to my family—Prof Asadullah Shah, Syed Nazir Ahmad Shah, Syed Firdousa Jan, Syed Muzamil, Mahreen Mahajan, and Syed Mirha—for standing by me with their unwavering support, understanding, love, and belief in my abilities.

I am profoundly thankful to my esteemed supervisor, Assoc. Prof. Ts. Dr. Amelia Ritahani binti Ismail, whose expertise, wisdom, and motivation have been indispensable throughout my research. Her willingness to share knowledge has enriched my academic experience.

Throughout my journey of seeking knowledge and exploration, I have been privileged to be immersed in a community of scholars, colleagues, friends, and family whose unified support and encouragement have played a crucial role in advancing my journey. Their unwavering belief in my abilities has consistently provided a foundation of strength and motivation.

Lastly, I offer my gratitude to Allah (S.W.T.) for His continuous blessings and guidance. This thesis stands as a testament to His mercy and my unwavering commitment to the pursuit of knowledge.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
ASD	Average Surface Distance
ASPP	Atrous Spatial Pyramid Pooling
BN	Batch Normalization
CBCT	Cone Beam Computed Tomography
CE-MRI	Contrast Enhanced Magnetic Resonance Imaging
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CT	Computed Tomography
DL	Deep Learning
DNN	Deep Neural Network
DSC	Dice Similarity Coefficient
EM	Electron Microscopy
FCN	Fully Convolutional Network
FN	False Negative
FP	False Positive
GPU	Graphics Processing Unit
IOU	Intersection over Union
JPEG	Joint Photographic Expert Group
MLP	Multi-layer Perceptron
MRI	Magnetic Resonance Imaging
PET	Positron Resonance Imaging
ReLu	Rectified Linear Unit
ResNet	Residual Network
RNN	Recurrent Neural Network
ROI	Region of Interest
SE	Squeeze and Excitation
TN	True Negative
TP	True Positive
US	Ultrasound
VGG	Visual Geometry Group

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE RESEARCH

Image analysis is a technique of obtaining information by measuring objects within an image. It is a procedure of extracting the information from images. The various image processing techniques for image analysis are image pre-processing, image compression, edge detection, and image segmentation (Muthuselvi & Prabhu, 2017). For image analysis, segmentation is one of the most efficient and widely used techniques, which is the procedure of dividing digital images into multiple segments in order to improve the quality of the images (Merjulah & Chandra, 2017). Analysis of medical images is crucial for providing rapid and precise diagnosis and medical treatment (Tran et al., 2021). Medical images have to go through numerous phases prior to the identification of the disease. In medical applications, it is required to process the medical images to identify the details of the medical images. In the initial phase, medical images must be gathered, then the preprocessing has to be performed and after that, the data must be stored in memory. It requires huge computational power and memory (Vardhana et al., 2018). In medical images, segmentation is performed to extract the region of interest for medical image analysis (Merjulah & Chandra, 2017). Image Segmentation in the medical field is one of the significant challenges in the past few years (Tran et al., 2021).

Medical image segmentation, a process that involves determining the pixels of lesions or organs from background medical images such as Magnetic Resonance Imaging (MRI) or Computed Tomography (CT) images, is one of the most crucial tasks in medical image analysis in order to provide critical information about the volumes and shapes of these organs (Hesamian et al., 2019). In healthcare facilities, accurate medical image segmentation is necessary for proper diagnosis (Ashkani Chenarlogh et al., 2022). With the rapidly developing field of artificial intelligence, particularly deep learning (DL), image segmentation approaches that utilize deep learning have produced excellent outcomes in the domain of the segmentation of images. Deep learning provides several advantages over conventional machine learning and computer vision techniques regarding segmentation accuracy as well as speed. Consequently, the application of deep learning to segment the medical images can efficiently assist clinicians in confirming the dimension of cancerous growths and determining the effect prior to and following medical treatment, significantly decreasing the burden on clinicians. Despite the remarkable achievements in the segmentation of medical images in the past few years, the use of deep learning for the segmentation of medical images continues to encounter complications in studies. For instance, the resolution of the medical images is low, the availability of the medical images in the dataset is less and the quality is poor. The inaccurate medical image segmentation findings are not able to satisfy the necessary medical requirements (X. Liu et al., 2021). This research aims to develop a deep neural network architecture capable of segmentation with high accuracy. To achieve this aim, the proposed architecture is inspired by U-Net and is optimized by developing two networks in sequence using U-Net as a baseline architecture and a Multires U-Net. The Multires U-Net employs the Multiresidual block and Res path. Multiresidual block combines the features from DenseNet and ResNet.

This chapter will explain: the problem statement for medical image segmentation, research objectives, research questions, research hypotheses, and research significance, respectively. This chapter will also represent the research scope and limitations, and research contributions. The thesis outline and chapter summary are shown at the end of this chapter.

1.2 PROBLEM STATEMENT

A large amount of medical image datasets prevents manual segmentation in a reasonable time (Pereira et al., 2016). Manual segmentation techniques produce inaccuracy and require a huge amount of time (Abirami & Sheela, 2014). There are lots of medical image segmentation techniques; however, these techniques are sensitive to noise (Norouzi et al., 2014). The fundamental techniques of image segmentation are insufficient for achieving complex object segmentation across the numerous types of medical images (Ashkani Chenarlogh et al., 2022).

Deep learning has achieved significant advancements in recent years for image segmentation (Tran et al., 2021). Medical image segmentation based on deep learning has yet to overcome research challenges, despite the significant advancements it has made in recent years. The accuracy of medical image segmentation is not high, and also the availability of medical datasets is small. Inaccurate medical image segmentation findings are unable to fulfill the actual criteria (X. Liu et al., 2021).

Convolutional Neural Network (CNN) is the most well-known and widely used algorithm in the field of deep learning (Alzubaidi et al., 2021). Nevertheless, CNN needs a huge amount of data during training to produce good results (Hu et al., 2020) which increases the memory requirement (Chakravarty & Sivaswamy, 2019). U-Net is one of the architectures of CNN that is a popular and successful model for the segmentation of images (Lundervold & Lundervold, 2019).

U-Net has shown a remarkable ability in the segmentation of medical images, even with a limited amount of training datasets (Ibtehaz & Rahman, 2020). However, if the characteristic of the image is not prominent, the performance of the U-Net network will be very low (Du et al., 2020). Moreover, for the segmentation of intricate medical images with less clear boundaries, U-Net tends to over-segment or under-segment (Ibtehaz & Rahman, 2020). U-Net skips the segmentation of the region of interest completely for the complex medical images (Ibtehaz & Rahman, 2020; Lou et al., 2021). In addition, U-Net and U-Net variants struggle to segment small region of interest or small organs (Zhu et al., 2018; Lou et al., 2021). Furthermore, U-Net and U-Net variants are limited to a single receptive field which affects the performance of the network for the segmentation task (Su et al., 2021).

1.3 RESEARCH QUESTIONS

The research intends to address the following research questions:

1. What are the issues in medical image segmentation with the U-Net architecture of the Convolutional Neural Network?

2. Why to develop an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for medical image segmentation?

3. How to develop A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for medical image segmentation?

4. How effective is A Dual Convolutional Neural Network with U-Net and Multi-Residual Network compared to other existing segmentation architectures?

1.4 RESEARCH OBJECTIVES

The main research objective is to develop A Dual Convolutional Neural Network with U-Net and Multi-Residual Network architecture for the segmentation of medical images. To achieve these the following research objectives, need to be carried out:

1. To investigate the issues of the U-Net architecture of Convolutional Neural Networks for medical image segmentation.

2. To propose A Dual Convolutional Neural Network with U-Net and Multi-Residual Network architecture for the segmentation of medical images.

3. To evaluate and validate the effectiveness of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network architecture on different medical image datasets for segmentation by comparing the results with other existing segmentation architectures.

1.5 RESEARCH HYPOTHESES

The hypothesis for this research is that the segmentation of medical images using A Dual Convolutional Neural Network with U-Net and Multi-Residual Network can segment medical images accurately and can outperform other convolutional neural network segmentation architectures.

1.6 RESEARCH SIGNIFICANCE

Segmentation of medical images is a crucial step in many applications of the clinical field, such as disease diagnosis and treatment planning. Various researchers have used numerous techniques to improve the segmentation of medical images. However, due to the complexity of medical images, the accurate segmentation of medical images remains a challenging task. The significance of this research is to develop an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the segmentation of medical images. The optimized segmentation architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, will improve the performance of segmentation for medical images.

This research aims to provide a more accurate segmentation for the medical images. A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is proposed for the segmentation of medical images in this research. The proposed architecture is optimized by using the two networks, U-Net and MultiRes U-net. The optimized segmentation architecture for the medical images utilizes the atrous spatial pyramid pooling (ASPP) with different dilation rates, which increases the receptive field in order to acquire more information, also squeeze and excitation attention mechanism is used to select the more useful information while ignoring the irrelevant information. Data augmentation is also applied to the dataset, which helps to improve the performance of the segmentation. Then the optimized segmentation architecture is applied to the different datasets. The performance results are evaluated in terms of the dice similarity coefficient, intersection over union, precision, and recall and are compared with the existing architectures for the segmentation of medical images. Overall, the findings demonstrate that the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, provides an improvement for the segmentation of the medical image, indicating that the result is promising.

1.7 RESEARCH SCOPE AND LIMITATION

This section explains the scope and limitations of the research on the development of an optimized architecture for the segmentation of medical images.

The research is focused on the development of the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, to improve the performance of the segmentation of medical images. The proposed architecture for the medical image segmentation is optimized by the implementation of the Resnet_50, Atrous Spatial Pyramid Pooling, Squeeze and Excitation, multires block, and respath to improve the segmentation performance of the architecture. The optimized architecture is trained and evaluated for the segmentation of different medical datasets. In this research, the evaluation of the performance of the optimized architecture is conducted by using the metrics: dice similarity coefficient, intersection over union, precision, and recall.

The limitation of this research is listed below:

a) This research is limited to the evaluation of the optimized architecture on the limited dataset because of the less availability of the medical image datasets.

b) The computational resources and the Graphics Processing Unit (GPU) memory limit the research.

1.8 RESEARCH CONTRIBUTION

This research proposed an optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the segmentation of medical images using U-Net as a baseline architecture. The proposed segmentation architecture utilizes two networks, the first network is U-Net, and the second network is MultiRes U-Net. The architecture consists of two encoders and two decoders. The first network in the proposed architecture contains a pre-trained Resnet-50 encoder, that has been trained on ImageNet. The second network consists of the multires blocks and the respath. Also, squeeze and excitation block and atrous spatial pyramid pooling (ASPP) are used in both the networks of the proposed architecture.

The experiment is conducted on three different medical image datasets with a variety of imaging modalities including the two publicly available datasets, the gastrointestinal polyp dataset, and the brain tumor dataset. Furthermore, the third dataset, dental images, is gathered from the Oral Radiology Unit. The datasets are pre-processed before it is fed to the proposed architecture for segmentation. The dental dataset is manually cropped and the binary mask for the ground truth is generated (see section 3.2.2.2). Moreover, data augmentation is applied to all three datasets to increase the training dataset.

The performance of the proposed segmentation architecture is evaluated using metrics Dice Similarity Coefficient (DSC), Intersection over Union (IOU), precision, and recall. Comparison is performed with the other segmentation architectures to demonstrate that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, performs better for the segmentation of medical images.

The impact of this research lies in the performance achieved through the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-

Residual Network. The optimized architecture is efficient in identifying fine details and characteristics in medical images by using two networks, which results in more accurate segmentation outcomes, providing medical professionals with better insights into complex anatomical structures. The optimized architecture not only advances the current state of the art in medical image segmentation but also provides insight for future research in the field.

1.9 THESIS ORGANISATION

This thesis is divided into six chapters: introduction, literature review, methodology, experimental design, experimental results, and conclusion and future work. Figure 1.1 illustrates the organization of the thesis.

CHAPTER 2 – LITERATURE REVIEW
Detailed understanding of the segmentation techniques for medical images.

↓

CHAPTER 3 – METHODOLOGY
Covers the conceptual framework for research and conceptual design of the architecture.

↓

CHAPTER 4 – EXPERIMENTAL DESIGN
Precisely describes the proposed architecture for segmentation of medical images.

↓

CHAPTER 5 – EXPERIMENTAL RESULTS
Evaluation and performance analysis of the proposed architecture

↓

CHAPTER 6 – CONCLUSION AND FUTURE WORK
Discusses the findings of research and outline the future work.

Figure 1.1 Organization of the Thesis

Chapter Two provides an overview of the literature on the published segmentation techniques. It discusses the work related to the segmentation of medical images using deep neural networks and also explores work related to segmentation techniques using the optimized algorithm by different researchers.

Chapter Three emphasizes the phases of the proposed methodology for the research. It discusses the conceptual framework and conceptual design that is used to develop an optimized architecture for the segmentation of medical images.

Chapter Four presents the experimental design of the proposed architecture for the segmentation of medical images. Two networks, U-Net, and MultiRes U-Net, from the proposed architecture are explained. The use of atrous spatial pyramid pooling (ASPP), squeeze and excitation block, Multires block, and the res path in the proposed architecture to achieve high performance for the segmentation of medical are explained.

Chapter Five reports the experimental results of the proposed architecture. It includes the evaluation and performance analysis using metrics for the segmentation of different medical datasets using the proposed architecture. Further, it illustrates the comparison results of the proposed architecture to the benchmark architecture for the segmentation of different medical datasets to prove the ability to generate better segmentation results.

Chapter Six concludes the research by remarking on the process of developing A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images. It presents the conclusion of the thesis and possible direction for future work.

1.10 CHAPTER SUMMARY

This chapter provides an overview of segmentation techniques for medical images. Section 1.1 presents the background of the research. Section 1.2 explains the problem statement. Research questions are mentioned in Section 1.3 and research objectives are presented in Section 1.4. Section 1.5 shows the research hypotheses. Research significance is illustrated in Section 1.6. The research scope and limitations of the research are discussed in Section 1.7. Section 1.8 explains the contribution of the research. The organization of the thesis is demonstrated in Section 1.9.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

Chapter 2 provides a depth of insight into the content of the segmentation techniques for medical images using deep neural networks. This chapter illustrates a review of the prior literature works related to deep neural networks for the segmentation of medical images, along with existing gaps. This chapter is categorized into sections that discuss medical images, image analysis, medical image segmentation, deep neural network, convolutional neural network, and U-Net. In section 2.2, the chapter discusses the information regarding medical images. This section presents the medical images used in the published experiments for the segmentation. Section 2.3 explains the techniques for image analysis. Section 2.4 illustrates literature regarding the image segmentation. It gathers the work related to the segmentation of medical images. Deep learning techniques for the segmentation of medical images are presented in section 2.5. Section 2.6 discusses the U-Net and work related to U-Net for the segmentation of medical images. The research gap of the previous research is discussed in Section 2.7. Finally, Section 2.8 presents the chapter summary.

2.2 MEDICAL IMAGES

Medical images play an essential role in monitoring and diagnosing the condition of the patient's health (Mehmood et al., 2019). Medical images such as X-Rays, Ultrasound (US), Magnetic Resonance Imaging (MRI), Computed Tomography (CT) scans, and Positron-emission Tomography (PET) (Anwar et al., 2018) are used for a visual representation of the internal body for clinical studies, diagnosis, and treatment planning (Mittal et al., 2020). Medical images contain rich features with high resolution, a massive amount, and complex features (Bo et al., 2020). Medical images have been used and stored continuously for diagnosis and research reasons (Sharma & Aggarwal, 2019).

In a previous study, Feng et al. (2023) used X-Ray images of sella turcica for the segmentation. Ferl et al. (2022) used CT scans of the lungs for the segmentation of lung tumors. Hashem & Kamil (2022), used an X-Ray of the chest for the segmentation. Chen et al. (2021) used CT images of abdominal organs and MRI images of the cardiac diagnosis for the segmentation. Yousefirizi et al. (2021) used PET images for oncological segmentation. Pasban et al. (2020) used an MRI of an infant's brain for the segmentation. Yang et al. (2020) have used ultrasound for the segmentation of medical images. Singh et al. (2020) used ultrasound for the segmentation of breast tumors. Seeja & Suresh (2019) used dermoscopy images for the segmentation of skin cancer. Zhao et al. (2018) used an MRI of the head and neck for the segmentation of cancerous lymph nodes.

2.3 IMAGES ANALYSIS

Image analysis is a technique of acquiring information by the evaluation of objects in an image. It is the procedure of extracting information from images. Image analysis has become very useful for industrial and research purposes due to its capability for processing digital images and objectively, analyze aspects such as size, color, distance, and the number of particles without disturbing the sample, (Saarinen & Petterson, 2006). The various image processing techniques for image analysis are image pre-processing, image compression, edge detection, and image segmentation (Muthuselvi & Prabhu, 2017). The primary goal of image processing is to improve the quality of the image and extract significant data from the images (Abubakar & Idris, 2020). Image pre-processing is used to remove background noise to enhance data images prior to computational processing. Image compression is a type of data compression that decreases the redundancy of the image and stores or transmits data in a more efficient manner (Muthuselvi & Prabhu, 2017). Edge detection is a technique that utilizes border identification of connected objects or regions, and this is used to locate the discontinuity of the object (Mohamed Y. Abdallah & Alqahtani, 2019). Image segmentation is a technique in which an image is partitioned into several segments, each of which represents some type of information (W. Khan, 2014).

2.4 IMAGES SEGMENTATION

Image segmentation is an important step in image analysis. It is a fundamental stage to analyze the images and extract data from them (S.Kannan et al., 2013). Image segmentation is a challenging and crucial phase in the processing of images for image analysis. Segmentation techniques for images are improving more quickly and precisely (X. Liu et al., 2021). Segmentation of images is a process that divides an image into regions or segments which is used to locate boundaries and objects such as curves, and lines in images (Abirami & Sheela, 2014). Segmentation of images is a procedure of splitting an input image into various segments that have a strong association with the region of interest (ROI) in the provided image (Malhotra et al., 2022). For image segmentation, commonly used segmentation techniques are Threshold-based image segmentation, Edge-based image segmentation, Region-based image segmentation, (Shirly & Ramesh, 2019), and Deep Neural Network-based image segmentation (Malhotra et al., 2022). In the threshold-based image segmentation technique, pixels were assigned to distinct categories based on the range of values in which a certain pixel resides. In the edge-based image segmentation technique, the pixels in an image are classified as non-edge or edge according to the filter applied to an image. In the region-based image segmentation technique, groups of pixels with different values and adjacent pixels with similar values are separated (Malhotra et al., 2022). Deep Neural Network-based techniques have helped to produce better and more efficient image segmentation algorithms in the past few years, and it had attained remarkable accuracy rates on several datasets (Malhotra et al., 2022).

2.4.1 Image Segmentation for Medical Images

The medical images need to pass a number of stages before the disease can be diagnosed. In the beginning, the images must be collected, after that the preprocessing is done and then the data is stored in memory. It needs enormous memory and processing time. Processing of images is essential in medical domains in order to obtain information about the images for analysis (Vardhana et al., 2018). Segmentation is a crucial phase in medical images for analysis and interpretation (Gupta et al., 2018). The domain of the analysis of medical images is growing, and the segmentation of abnormalities, diseases, or organs in medical images has become challenging because of the numerous artifacts present in the images (Malhotra et al., 2022). The intention of image segmentation is to increase the quality of medical images (Merjulah & Chandra, 2017). Segmentation is used to detect and extract the feature areas in medical images(Gupta et al., 2018). Image segmentation extracts features from medical images for better and more accurate medical diagnostics (Shirly & Ramesh, 2019). The goal of the medical image segmentation is to present an image in a meaningful form in order to identify the region of interest (ROI), examine the anatomy, analyze the volume of tissue to determine the tumor and assist in deciding the dose of the medication and treatment planning (Malhotra et al., 2022).

Hsiao et al. (2005) proposed an edge-based segmentation technique for images. The segmentation is performed on different kinds of images. The input image for this algorithm is a grey-level image. This segmentation algorithm for the images uses a region-growing method with a morphological edge detector. Initially, images are enhanced through the morphological closing operations to minimize noise while improving the edge features of the grey-level image, and then the edge detection of the image is performed by using the morphological dilation residue edge detector. Finally, the region growing and region merging techniques are applied for the segmentation of images into various segments. The edge-based image segmentation technique has the drawback that noise can affect the performance and false edges could be present, which adversely impact the results of segmentation (Shirly & Ramesh, 2019).

Wu et al. (2008) suggested a segmentation algorithm that employed a seeded region growing (SRG) algorithm for the automatic organ segmentation of MR images of the abdominal. They have tested this algorithm on 12 sets of 3D MR images of the abdominal. Initially, they extract texture features in the Region of Interest (ROI) for each pixel based on the homogeneity criteria. Co-occurrence and semivariogram are the two texture features that have been analyzed. Co-occurrence is effective for biomedical image processing, whereas a semi-variogram is a commonly used scale for geostatistics dissimilarity. Then the automatic seeded region growing algorithm is applied to the feature spaces. The cost function is decreased with three factors to calculate the seed. The lower value is taken to acquire the threshold prior to the one producing an ‘explosion’. Few modifications are implemented to prevent over-segmentation, and under-segmentation, and accelerate the evaluation. The algorithm of the seeded region growing is then implemented and the extraction of the right kidney is in the experiment.

Taheri et al. (2010) proposed threshold-based image segmentation which utilizes level sets for 3D segmentation of tumors in MR images of the brain. They suggest a level-set method for the segmentation of 3D brain tumors that uses a speed function that will not involve calculating the density function and is achieved by minimum user intervention. They have developed a speed function by using a threshold value. The level set initialization has been used to calculate an initial threshold and then it is modified sequentially during the segmentation process. When the tumor border is reached, the threshold variance reduces due to the severity difference between non-tumor and tumor, and then the procedure ends. The efficiency of this suggested segmentation is superior when the degree of severity is greater between the non-tumor and tumor region.

Li et al. (2011), proposed a region-basic algorithm for the segmentation of MR images. In this proposed algorithm, the segmentation of the image is based on the intensity inhomogeneity of the image. The level set function is determined by the local intensity clustering property. The bias field and image domain partition are represented by the level set function to illustrate the intensity inhomogeneity of the image. This technique minimizes the level-set function in order to concurrently segment the image and calculate the bias field. The calculated bias field is then employed for intensity inhomogeneity correction. The disadvantage of a region-based technique for the segmentation of images is that it is limited to use on closed boundaries, and there is the possibility of under-segmentation and over-segmentation of the region (Shirly & Ramesh, 2019).

Cao et al. (2016), proposed advanced segmentation algorithms for whole-heart CT series images. They improve the two traditional segmentation algorithms; one algorithm is region-based and another algorithm is edge-based segmentation and then they apply these improved segmentation algorithms to the CT images of the whole heart. This study has evaluated the algorithms for segmentation on three features: performance of an algorithm, evaluation of segmentation results of medical images by subjective evaluation, and Jaccard coefficient and Dice coefficient deviation method are used to calculate the performance of segmentation. When these two advanced algorithms for segmentation are compared to the results of manual segmentation, the advanced segmentation algorithms demonstrate excellent performance in the accurate segmentation of the CT sequence images of the whole heart.

2.5 DEEP LEARNING FOR SEGMENTATION OF MEDICAL IMAGES

Recently, great progress has been achieved in the development of efficient and accurate algorithms of machine learning to perform medical image segmentation (Seo et al., 2020). A rapid increase in deep learning based methods has overwhelmed the machine learning community (Ghosh et al., 2019). The capability of deep learning-based models to automatically extract features from input data has significantly increased its application in recent times. Deep learning techniques have generated promising outcomes for several applications, including image classification and segmentation of medical images (Ashkani Chenarlogh et al., 2022). Deep learning models have demonstrated use in numerous image segmentation applications. The achievements and the great performance of the deep learning algorithms are responsible for enormous growth (Malhotra et al., 2022). Segmentation of images is effectively addressed by different types of deep learning methods, including Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs) (Seo et al., 2020). CNN performs effectively on issues pertaining to biomedical images, convolutions play a significant part in CNN's performance in image analysis procedures such as classification and segmentation (Merjulah & Chandra, 2017).

2.5.1 Deep Neural Network

An Artificial Neural Network (ANN) is one of the algorithms of machine learning that is influenced by the human brain that consists of layers with connected nodes. Although, because of low computing power and inadequate learning data, an artificial neural network has suffered from overfitting problems for training deep networks. Due to the limitation of ANN, ANN is expanded to the Deep Neural Network (DNN) by building multi-hidden layers between input and output with linked nodes (Kim et al., 2019). The learning capability and structural complexity of deep neural networks have received great attention for performance enhancement because of the multiple hidden layers and ability to learn better features from image data which hence can respond well to actual-world scenarios (Lau & Hann Lim, 2018). Figure 2.1 shows Artificial Neural Network and Figure 2.2 shows Deep Neural Network.

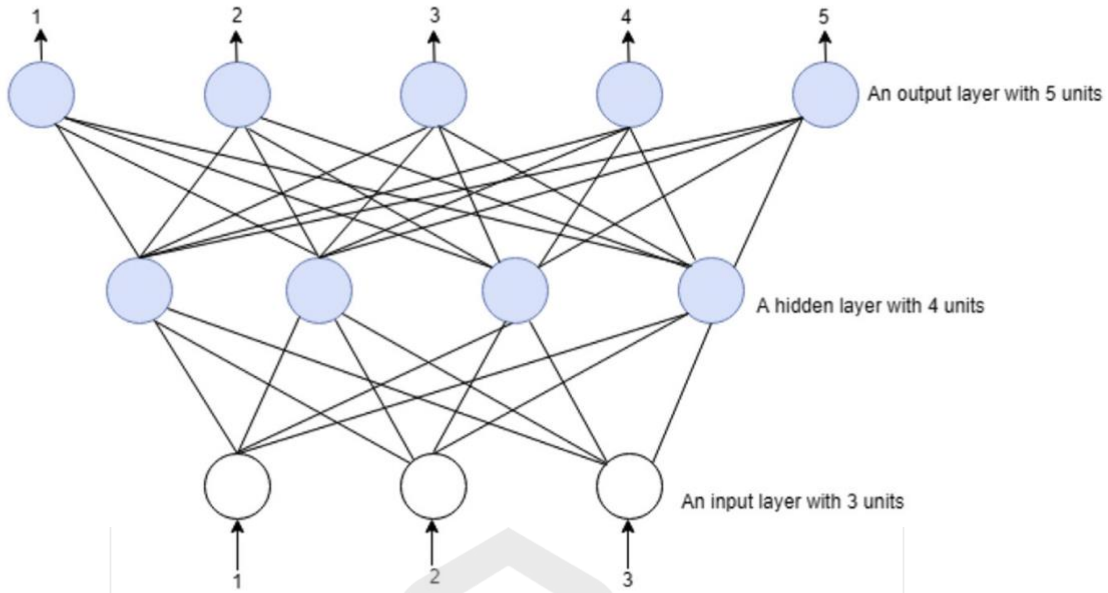


Figure 2.1 Artificial Neural Network

Source: (Abiodun et al., 2018)

Figure 2.1 shows an Artificial Neural Network, which has one input layer, one hidden layer, and one output layer and it also shows all connections between the units in the layers (Abiodun et al., 2018).

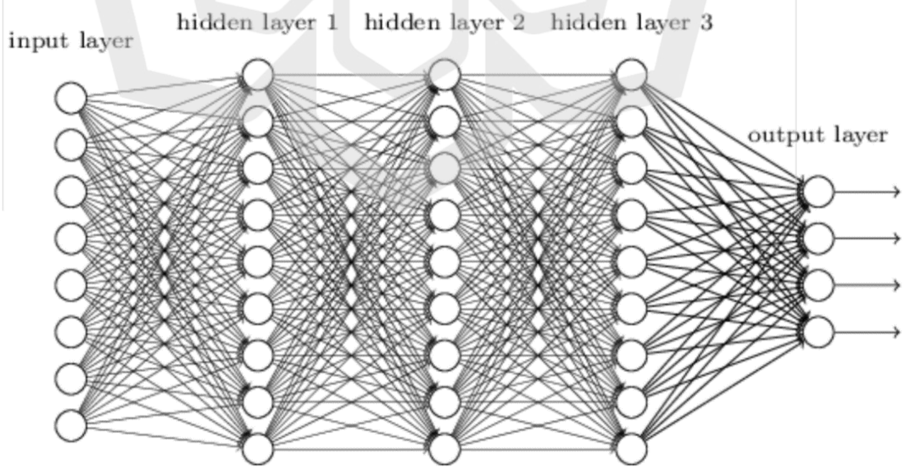


Figure 2.2 Deep Neural Network

Source: (Tripathi et al., 2017)

Figure 2.2 demonstrates the architecture of a Deep Neural Network (DNN) with several hidden layers. The Deep Neural Network in Figure 2.2 comprises an input layer, an output layer, and three hidden layers, where the output from one layer acts as an input for the next layer (Tripathi et al., 2017).

2.5.2 Convolutional Neural Network

A Convolutional Neural Network (CNN) is a branch of a deep neural network, consisting of a sequence of layers performing a specific operation and each intermediate layer receives the results of the previous layer as its input (Hesamian et al., 2019). Convolutional Neural Network layers mainly consist of convolutional layers, pooling layers, and fully connected layers (T. Guo et al., 2017) as shown in Figure 2.3.

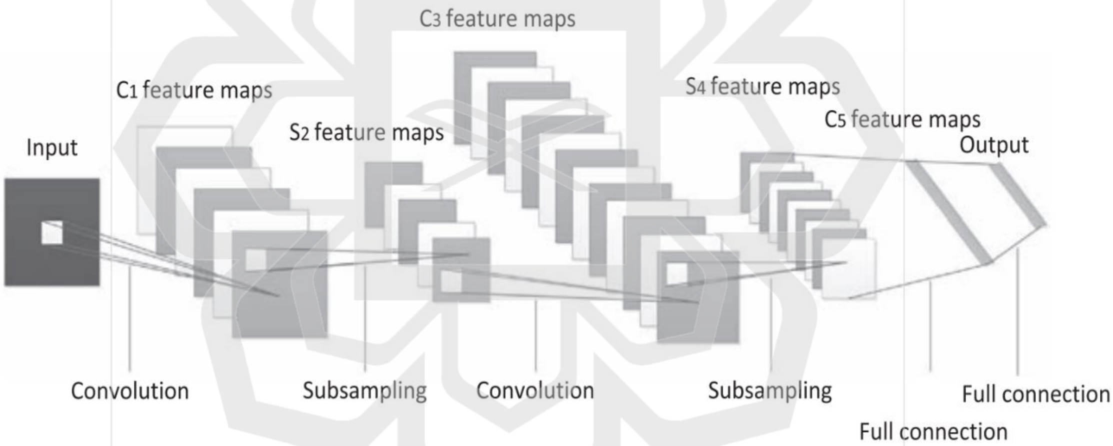


Figure 2.3 Architecture of CNN
Source: (W. Liu et al., 2017)

A convolutional layer (c-layer) is the key component of Convolutional Neural network architecture that performs the extraction of features (Yamashita et al., 2018). The output of the layer is considered the feature map (Pai et al., 2017). A pooling layer is often accompanied by convolutional layers. A pooling layer or subsampling layer (s-layer) typically receives the input data from the feature map produced after the convolution operation (W. Wang, Yang, et al., 2019). A pooling layer reduces the number of parameters and also reduces the number of computations required for training the network (Murphy, 2016). The dimension of the feature map is reduced by the pooling and it increases the effectiveness of the extraction of features (T. Guo et al., 2017). The output of the last convolutional layer or pooling layer feature map gets converted as a one-dimensional (1D) array (flattened) and is connected to one or more fully connected layers where each input is connected with each output by a learnable weight (Yamashita et al., 2018). The last fully connected layer is followed by an output layer (T. Guo et al., 2017).

Convolutional Neural Network has various types of architectures in deep neural networks. Some of the Convolutional Neural Network architectures are AlexNet, VGG, ResNet, FCN, U-Net, and SegNet (Mittal et al., 2020). AlexNet, VGG, and ResNet are considered to be the most common architecture that performs successfully for image classification and image recognition tasks (A. Khan et al., 2020).

A Fully Convolutional Network (FCN) is a segmentation approach that allows pixel-wise prediction based on the ground truth images. FCN extracts crucial feature maps and restores those feature maps to the image labels. The ability of FCN to represent features is the main reason for achieving success in object detection, classification, and image segmentation. FCN will perform well only when a large number of the training dataset is available. In the field of medical imaging, data collection is expensive and is less available because of privacy and regulation concerns (Lai et al., 2019). FCN provides less accurate results without enough training data (Bi et al., 2018) and results in inaccurate segmentation for small organs (Hesamian et al., 2019). So FCN is extended to U-Net to do segmentation on medical images that perform well with comparatively less number of the training dataset (Lai et al., 2019).

SegNet is an efficient novel technique of a CNN architecture for the image segmentation task (Ait Skourt et al., 2018). But U-Net shows a high speed of process during the training phase (Pashaei et al., 2020). SegNet has lower learning performance as compared to U-Net (Song et al., 2019). Kasar et al. (2021), compare the performance of Segnet and U-Net for the segmentation task. The authors demonstrate that the U-Net outperforms the SegNet. Another author, Vianna et al. (2021), performs the segmentation using U-Net and SegNet. The authors present that U-Net performs better than the SegNet for image segmentation. The U-Net architecture and its variants, an encoder-decoder based deep convolutional neural network, are widely employed to deal with the segmentation tasks of medical images (Nie et al., 2022).

2.6 U-NET ARCHITECTURE

U-Net is one of the most popular Convolutional Neural Network architectures for medical image segmentation tasks (Z. Zhang et al., 2020). U-Net network mainly consists of a convolutional layer, upsampling (expensive path), downsampling (contracting path), and activation function (Yang, Feng, et al., 2020). The architecture of U-Net is shown in Figure 2.4.

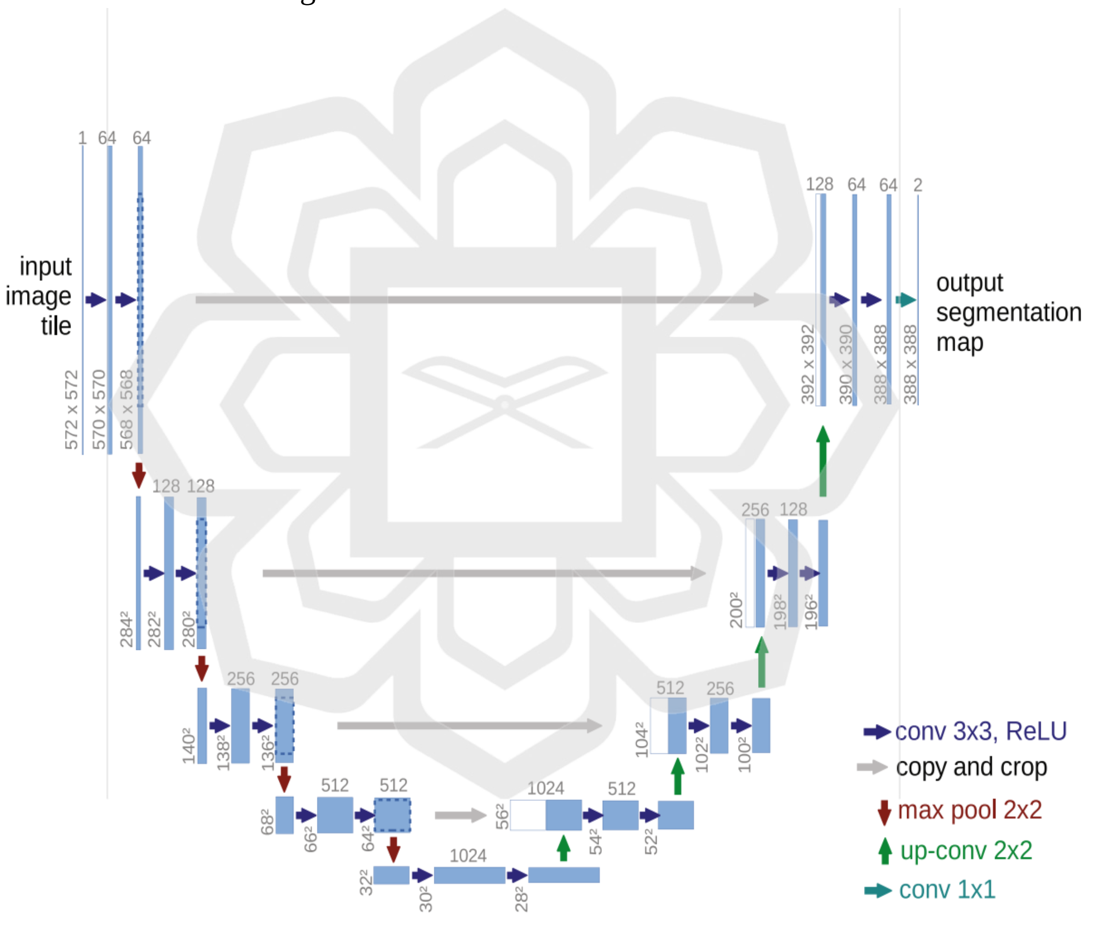


Figure 2.4 Architecture of U-Net

Source: (Ronneberger et al., 2015)

In Figure 2.4 The U-Net architecture comprises a downsampling/ contracting path (left side) and an upsampling/ expansive path (right side). The downsampling consists of two 3x3 convolutional layers by two rectified linear unit (ReLU) activation functions and 2x2 max pooling. At pooling, the number of features doubles. The bottleneck of the network is between downsampling and upsampling. It is made up of two convolutional layers. The upsampling consists of 2x2 convolution (up-convolution) that reduces the number of features, a concatenation with the correspondingly cropped feature map from the downsampling path, and it contains two 3x3 convolution layer followed by two ReLU activation functions. A 1x1 convolution layer is used in the final layer to map each feature to the required number of classes (Ronneberger et al., 2015). This architecture has gained a huge interest in the segmentation of medical images, based on which several variants have been created (Hesamian et al., 2019). The pseudocode for U-Net is shown in Pseudocode 1.

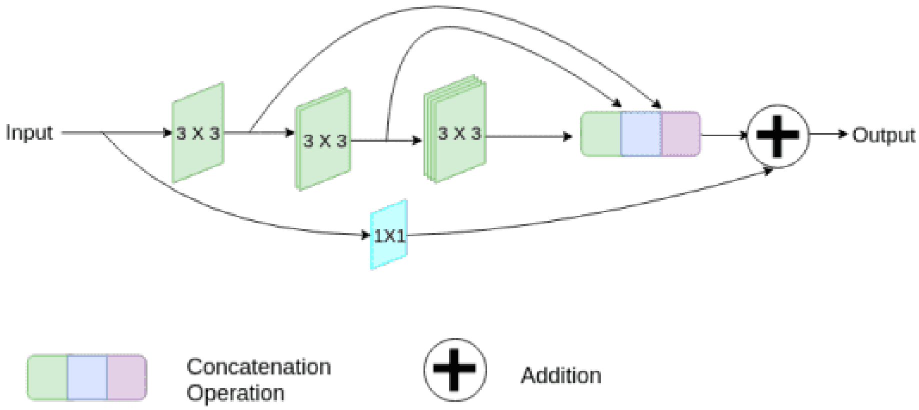
Pseudocode 1: Pseudocode for U-Net Architecture

```
# Import  libraries
import tensorflow as tf

# Define the U-Net model
def  unet_model(input_shape=(height, width, channels)):

    # Encoder (Contracting Path)
    inputs = tf.keras.layers.Input(shape=input_shape)
    conv1 = convolution_block(inputs, 64)
    conv2 = convolution_block(conv1, 128)
    conv3 = convolution_block(conv2, 256)
    conv4 = convolution_block(conv3, 512)

    # Bottleneck
    bottleneck = convolution_block(conv4, 1024)

    # Decoder (Expansive Path)
    upconv4 = upconvolution_block(bottleneck, conv4, 512)
    upconv3 = upconvolution_block(upconv4, conv3, 256)
    upconv2 = upconvolution_block(upconv3, conv2, 128)
    upconv1 = upconvolution_block(upconv2, conv1, 64)

    # Output layer
    outputs = tf.keras.layers.Conv2D(1, 1, activation='sigmoid')(upconv1)

    # Create the model
    model = tf.keras.models.Model(inputs, outputs)
    return model

# Define a convolution block
def convolution_block(x, filters):
    conv = tf.keras.layers.Conv2D(filters, (3, 3), activation='relu', padding='same')(x)
    conv = tf.keras.layers.Conv2D(filters, (3, 3), activation='relu', padding='same')(conv)
    pool = tf.keras.layers.MaxPooling2D((2, 2))(conv)
    return pool

# Define an upconvolution block
def upconvolution_block(x, skip, filters):
    upconv = tf.keras.layers.Conv2DTranspose(filters, (2, 2), strides=(2, 2), padding='same')(x)
    concat = tf.keras.layers.concatenate([upconv, skip], axis=-1)
    conv = convolution_block(concat, filters)
    return conv

# Initialize the U-Net model
model = unet_model()
```

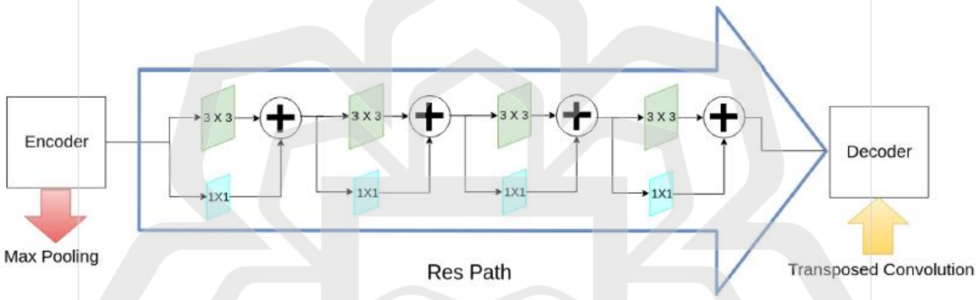
Pseudocode 1 Pseudocode for U-Net Architecture

U-Net was particularly created for image segmentation tasks. U-Net is highly utilized in the medical image industry and has gained widespread adoption as the main tool for medical image segmentation (Siddique et al., 2021). Despite the fact that U-Net has excelled in the segmentation of medical images (H. Wang et al., 2022), U-Net has an overfitting problem (C. Guo et al., 2021; Ren et al., 2021), a vanishing gradient issue (Y. Cao et al., 2020), a semantic gap between encoder and decoder (Amer et al., 2022) and it has a small receptive field (Nie et al., 2022; Li et al., 2021; S. Wang et al., 2020; J. Zhang et al., 2021). In order to prevent an overfitting problem, data augmentation techniques can be applied to the dataset, and data augmentation helps to improve the results of the segmentation (Malhotra et al., 2022; Rice et al., 2020; Altinsoy et al., 2019). Residual blocks help to tackle the issue of vanishing gradient in the network (Ashkani Chenarlogh et al., 2022; Alom et al., 2018; Gudhe et al., 2021; Amer et al., 2022). Respath as a skip connection can solve the semantic gap between the encoder and decoder by transmitting the features of the encoder through a series of convolutional layers (Ibtehaz & Rahman, 2020; Gudhe et al., 2021). Atrous spatial pyramid pooling (ASPP) with various dilation rates expands the receptive field in the network to extract more information at different scales (L. Chen et al., 2018; L. Chen et al., 2017; Kiran et al., 2022; Nie et al., 2022). The larger the receptive field, the higher segmentation performance can be achieved (Nie et al., 2022).

MultiResUNet (Ibtehaz & Rahman, 2020) is one of the variants of U-Net, developed for the segmentation of medical images. In MultiResUNet, Multi-res blocks are used in the encoder and decoder part of the architecture, and the res-path is employed as a skip connection to achieve the segmentation. Residual blocks assist in addressing the network's vanishing gradient problem. Res-path is used to decrease the semantic gap between the encoder and decoder feature maps. Figure 2.5 demonstrates the Multires block and Res-path of MultiResUNet.



(a)



(b)

Figure 2.5 MultiresUNet: (a) Multires block (b) Res-path

Source: (Ibtehaz & Rahman, 2020)

Figure 2.5 shows the Multires block and Res-path in the MultiresUNet. The Multires block in Figure 2.5 (a) uses a series of convolutional blocks. The output of the three convolutional blocks is then concatenated to the spatial features at various scales. Also, a residual connection of 1×1 convolutional layer is added to gather additional spatial features. Figure 2.5 (b) shows the Res Path, there are 3×3 and 1×1 filters. Rather than directly merging the feature maps of an encoder with the features of a decoder, the encoder features are used to pass through a series of convolutional layers to reduce the semantic gap between the encoder and decoder. Additionally, residual connections are also added in the Res path to facilitate learning. The pseudocode for the Multi-Residual block and Res Path is shown in Pseudocode 2 and Pseudocode 3 respectively.

Pseudocode 2: Pseudocode for Multi-Residual Block

```
def multi_residual_block(input_features, filters):
    # Input: input_features - Tensor from the previous layer
    #          filters - Number of filters for convolutional layers
    # First convolution layer with 3x3 kernel size
    conv1 = Conv2D(filters, (3, 3), activation='relu', padding='same')(input_features)
    # Second convolution layer with 3x3 kernel size
    conv2 = Conv2D(filters, (3, 3), activation='relu', padding='same')(conv1)
    # Third convolution layer with 3x3 kernel size
    conv3 = Conv2D(filters, (3, 3), activation='relu', padding='same')(conv2)
    # Residual Path
    residual_path_output = residual_path(input_features, filters)
    # Add the residual connection
    multi_residual_output = add([residual_path_output, conv3])
    return multi_residual_output
```

Pseudocode 2 Pseudocode for Multi-Residual Block

Pseudocode 3: Pseudocode for Res Path

```
def ResPath(input_layer, filters, length):
     # input_layer -- Input_features - Tensor from the previous layer
     # filters  -- Number of filters in convolutional layers
     # length -- Length of ResPath
    # Create a shortcut connection
    shortcut = input_layer
    shortcut = Conv2D(filters, (1, 1), activation=None, padding='same')(shortcut)
    shortcut = BatchNormalization(axis=3)(shortcut)
    # First convolutional block
    output_layer = Conv2D(filters, (3, 3), activation='relu', padding='same')(input_layer)
    output_layer = BatchNormalization(axis=3)(output_layer)
    output_layer = add([shortcut, output_layer])
    output_layer = Activation('relu')(output_layer)
    # Additional convolutional blocks based on the specified length
    for i in range(length-1):
        # Create a shortcut connection
        shortcut = output_layer
        shortcut = Conv2D(filters, (1, 1), activation=None, padding='same')(shortcut)
        shortcut = BatchNormalization(axis=3)(shortcut)
        # Convolutional block
        output_layer = Conv2D(filters, (3, 3), activation='relu', padding='same')(output_layer)
        output_layer = BatchNormalization(axis=3)(output_layer)
        output_layer = add([shortcut, output_layer])
        output_layer = Activation('relu')(output_layer)
    return output_layer
```

Pseudocode 3 Pseudocode for Res Path

Zhou et al. (2018), proposed a Nested U-NET architecture for the segmentation of medical images. They implemented U-Net++ architecture for the medical image segmentation based on nested and dense skip connections. The concept of their architecture is to capture the fine-grained information about the objects in the foreground effectively while the high-resolution feature maps of the encoder network are slowly enriched before combining with the corresponding semantically rich features of the decoder network. The U-Net++ was analyzed on four datasets of medical images, that is colon polyp segmentation, lung nodule segmentation, liver segmentation, and cell nuclei segmentation. This architecture takes the benefit of re-designed skip pathways and deep supervision. The redesigned skip pathways are to reduce the semantic gap in the feature maps of the encoder and the decoder, which results in easier optimization problems to solve by the optimizer. Deep supervision allows more precise segmentation, especially in the case of multiple-scale lesions. Their experiments have shown that U-Net++ has obtained more IOU as compared to U-Net and wide U-Net.

Zafar et al. (2020), proposed an algorithm that combines two architectures which are U-Net and ResNet for the segmentation of skin lesion images. The combination of these two architectures is called Res-Unet. They perform segmentation on datasets of Electron ISIC 2017 and PH2 datasets. Before giving input to the model they preprocess the images using scaling, resizing, data centering, and hair removal techniques. Morphological operations are applied to remove the noise. The design of the network is based on both ResNet and U-net perspectives. The network operates in an encoder-decoder form. The convolutional side, that is, the contracting path is based on the architecture of ResNet, and the deconvolutional side, which is, the expansive path is based on the pipeline of U-net. The architecture is trained for 100 epochs and data augmentation is applied at runtime, which improves the efficiency of the model and produces good segmentation results. However, this model is complex and takes more time to train as compared to classical U-Net. This model achieved an IOU of 0.854 and a dice similarity coefficient of 0.924.

Hoang et al. (2019), proposed a Convolutional Neural Network (CNN) for the segmentation of CT images of the liver using data across multiple health clinics. They examined three Convolutional Neural Network architectures identified as DRIU (Deep retinal image understanding), V-Net, and FCN-CRF (conditional random field) for the segmentation of the liver. The performance of CNNs was evaluated on the basis of the

Hausdorff distance, false-positive rate, Dice score, and mean surface distance. Their result indicates that these three Convolutional Neural network architectures obtain an overall mean dice score of more than 90 percent for contrast-enhanced liver CT images. It showed that CNN-based segmentation for the liver CT images achieves good performance. However, non-contrast-enhanced and low-dose liver CT image segmentation remains an issue.

Habijan et al. (2019), proposed a 3D U-Net architecture of the Convolutional Neural Network (CNN) method for the segmentation of the whole heart. They have used two convolutional neural networks based on 3D U-Net architecture, one architecture is used to locate the bounding box all over the heart, and the other architecture is used for segmentation. The analysis of their proposed method is carried out on 20 2D CT images. The dataset of 20 2D CT is divided into two sets which consist of 15 training images and 5 validation images. The 3D U-Net architecture comprises two parts: one is a contracting encoder part and the other is the consecutive expanding decoder part. The task of the contracting encoder part is to examine the entire image and the task of the decoder part is to generate segmentation in high resolution. Due to the less number of datasets, the data augmentation process is used to maximize the number of data for the training which improves the performance. The dice coefficient is utilized during training for calculating the performance of the model. They achieved a mean dice score of 89% for the segmentation of the whole heart.

W. Wang, Chen, et al. (2019), have developed a model using coordination-guided deep neural networks for the automatic segmentation of pulmonary lobes using chest CT images. They use automatic segmentation of the lung CT images to extract the lung region and then implement a volumetric convolutional neural network (V-net) for the segmentation of the pulmonary lobe. The performance was evaluated on four different datasets by various measuring parameters. Experimental results indicate the high efficiency of the suggested method compared to previous state-of-the-art methods.

Zhou et al. (2019), developed a segmentation architecture for the medical image dataset. The U-net architecture is used as a baseline and has used a nested and dense skip connection for the segmentation. This architecture is evaluated for segmentation on the six different medical images. The segmentation architecture does not explore sufficient information from the full scale. It has a low robustness performance in learning about the location and boundary of the organs (Qian & Zhou, 2022).

Ait Skourt et al. (2018), proposed a deep learning algorithm using U-NET architecture for the segmentation of lung CT images. The dataset they used for the experiment contains thoracic computed tomography (CT) scans of diagnostic and lung cancer with highlighted-up illustrated lesions. Before the experiment was performed, for the ground truth, manual segmentation of lung parenchyma was provided. After manual segmentation, cropping of images is done to eliminate the details that are not for their field of analysis. The performance of architecture was calculated by using the dice coefficient index used for the calculation of similarity. They obtain a dice-coefficient measure of 0.9502 for effective segmentation.

P. Wang et al. (2018), proposed a Convolutional Neural Network (CNN) method for the automatic segmentation of two-dimensional (2D) brain ventricles ultrasound (US) scans. In this method, they use the combination of SegNet and U-Net architecture for the ventricle segmentation. They select randomly test set and training set from the sample data set of 687; 50 random samples for test data and 637 for training data. The model has been trained for 100 epochs to make sure that the loss function converges. The segmentation performance is measured by the Dice coefficient, pixel wise accuracy (Pixel Acc.), and Intersection over Union (IoU) that is, they obtain performance 0.908±0.053, 84.84±0.078, and 92.14±0.063 respectively.

S. Wang et al. (2017), proposed the segmentation of lung nodules using a multi-view convolutional neural network. Their network architecture comprises three convolutional neural network branches, each of which has seven stacked layers and accepts multi-scale nodule patches as input. Their network can segment CT images of lung nodules without any shape hypotheses or modifications with the user-interactive parameter, and it learns differential nodule sensitive features automatically from a huge number of image data. They perform segmentation on 393 nodules and calculate the performance by the Average Surface Distance (ASD) and the Dice similarity coefficient (DSC). They obtain DSC =77.67% and ASD = 0.24.

J. L. Wang et al. (2020), proposed U-net and curriculum learning strategy using a multi-task semi-supervised attention-based model for the segmentation of Intracranial Hemorrhage (ICH) CT scans. Their model worked with both a small labeled dataset as well as a large labeled dataset. This semi-supervised model achieved a Dice coefficient of 0.67 when trained with 80% of the ground truth dataset. The problem with this model was the incapability to segment smaller regions.

Y. Cao et al. (2020), developed an architecture for segmentation that uses the advantages of U-Net and DenseNet. The architecture is built from the contracting path, expanding path, and dense block. The convolutional layers have been replaced by the dense blocks to prevent the problem of vanishing gradient. The architecture is evaluated on the electron microscopy (EM) of the neuron dataset. Data augmentation was also applied to the dataset. The result shows poor performance for the segmentation of images with complex backgrounds (X. Huang et al., 2023).

H. Huang et al. (2020), developed a full-scale connected U-Net architecture for the segmentation of medical images. The architecture takes advantage of deep supervision and full-scale skip connections. It combines the features of multi-scale by skip connection re-designing. The evaluation of the architecture is demonstrated on two datasets, that are, liver and spleen datasets. The experimental results demonstrate that it does not have robust segmentation performance on the small region of interest when the number of datasets is small (Qian & Zhou, 2022).

Cherguif et al. (2019), proposed a deep convolution network based on the U-Net model for the segmentation of brain tumors. They test and evaluate the method on an image dataset produced by Multimodal Brain Tumor Image Segmentation (BRATS) 2017. They evaluate the performance of segmentation for the tumor region using the Dice Similarity Coefficient (DSC). The Dice similarity coefficient achieved for this model is 0.81. This model does not give any information on convergence, so the cross-entropy error gives a regular learning error.

Ibtehaz & Rahman (2020), developed a novel architecture for the segmentation of medical images. Multi-res blocks are applied in the encoder and decoder part and the res-path is used as a skip connection to obtain the segmentation. Res-path is used to decrease the semantic gap between the encoder and decoder feature maps. Various medical datasets with different modalities are used for the segmentation. The datasets have different sizes, which are then resized for the input image. The results are compared with the U-Net architecture. U-Net tends to under-segment and over-segment. The experimental results of the developed architecture show that where the boundaries of medical data are not clear, the segmentation results are influenced (Lou et al., 2021).

Jha et al. (2020), developed an Unet-based architecture for medical image segmentation. Two U-Net architectures are combined, in which the first U-Net employs the pre-trained VGG-19 as an encoder. The architecture concatenates two encoder-decoder networks. Atrous Spatial Pyramid Pooling (ASPP) is used between the encoder and decoder of both U-Net. Squeeze and Excitation block is also used in the network. The architecture is evaluated on four publicly available datasets. Data augmentation has been performed on the dataset to improve performance. The evaluation metrics used to evaluate the architecture are Dice Similarity Coefficient (DSC), Intersection over Union (IOU), Precision, and Recall. This architecture performs significantly better for segmentation when compared to U-Net. The result performance illustrates that the segmentation technique has robust for the cutting of boundary areas of images (C. Huang et al., 2021).

Sathananthavathi & Indumathi (2021), proposed a U-Net based architecture for the segmentation of retinal blood vessels. The encoder part of the architecture is improved by the encoder enhanced atrous. The enhanced encoder is improved by the depth concentration process with the additional layers. In this architecture, all the convolutions are replaced by the atrous convolution that helps to increase the receptive field. It consists of contraction and expansion paths. The contraction and the expansion path are asymmetric. The contraction path has five blocks of an encoder, batch normalization, pooling, and Relu activation function, and each stage contains two atrous convolutions. The encoder part contains two additional layers which improve the extraction of edge information. This architecture is evaluated on the dataset that is publicly available. The dice coefficient, accuracy, specification, and sensitivity metrics are utilized to evaluate the performance of the architecture.

Lou et al. (2021), developed a novel U-Net-like architecture for the segmentation of medical images. Res-path is used in the architecture, and also dual channel model is used for the segmentation. Three different datasets are used for the segmentation. The dataset includes Infrared breast images, electron microscopy of the ventral nerve cord, and endoscopy images of polyps. The architecture is compared with the MultiResNet and classic U-Net. The performance is evaluated by the Tanimoto accuracy metrics.

Liang et al. (2023), proposed an architecture based on U-Net for the segmentation of lung and liver datasets. The dual encoder is employed to increase the depth of the network and enhance the ability to extract more context features. The dual encoder model is enhanced with the Squeeze and Excitation block to extract the channel-level global features. Additionally, full-scale connections have been included, which encourages the combination of high-level semantic information with low-level details. Dilated convolution is also used to extract the feature maps. A large receptive field can be obtained by the use of dilated convolution. The use of dilated convolution improves the effect of segmentation and recognition of small objects. The performance of the architecture is evaluated by dice coefficient, F1 score, precision, and recall.

Jin et al. (2022), proposed an architecture for the segmentation of MR images of the brain tumor. This architecture combines multi dense encoder with U-Net. The encoder part includes a dense block, convolution layers, and transition. In this architecture dense block is used instead of skip connections. It concatenates all the features in an encoder part and passes it to the decoder part. The encoder part and decoder part are symmetric. The size of the feature maps is not changed by the dense block but doubles the channel and the feature maps learned by various layers are concatenated. The output feature is passed to the decoder and transition layers. In the encoder part, the transition layer gets dense block feature maps and the size of the feature maps is reduced. The performance of this architecture for the segmentation is evaluated by IOU metrics.

Ahmad et al. (2021), proposed an architecture based on a multi-scale hierarchy for the segmentation of brain tumor MRI. In this architecture, a hierarchical block is included between the encoder and decoder part for obtaining and combining features to acquire multi-scale information. Densely connected blocks are used in the architecture. Residual- Inception blocks are utilized to get full contextual information. This architecture is implemented on four MICCAI datasets. The performance is measured by the dice similarity coefficient metrics.

Table 2.1 illustrates the research work performed on deep learning Architectures for medical image segmentation. It presents the architecture used for the segmentation, segmentation part, medical image modality, and evaluation metrics.

Table 2.1 Articles Used Deep Learning Architectures for the Segmentation of Medical Images

Author/Year	**Architecture**	**Segmentation Part/ Image Modality**	**Evaluation Metrics**
(Liang et al., 2023)	Dual encoder model based on U-Net	Lung and liver, CT images	Dice coefficient, F1 score, precision, and recall
(Jin et al., 2022)	Multi dense encoder with U-Net	Brain tumor, MR images	Intersection over Union
(Ahmad et al., 2021)	Multi-scale hierarchical-based architecture	Brain tumor, MR images	Dice similarity coefficient
(Sathananthavathi & Indumathi, 2021)	Encoder enhanced atrous U-Net	Retinal blood vessel images	Dice similarity coefficient, accuracy, specification, and sensitivity
(Lou et al., 2021)	Novel U-Net-like architecture	Infrared breast images, Ventral nerve cord electron microscopy, and Polyps	Tanimoto accuracy

		endoscopy images	
(J. L. Wang et al., 2020)	U-net and curriculum learning strategy	Intracranial Hemorrhage (ICH), CT scans	Dice and jaccard coefficient
(Zafar et al., 2020)	Combines U-Net and ResNet	skin lesion, dermoscopy images	Intersection over Union and Dice similarity coefficient
(Jha et al., 2020)	Combines two U-Net architectures	Polyps, Skin lesions, Nuclei images	Intersection over Union, Dice Similarity Coefficient, Recall, and Precision
(Ibtehaz & Rahman, 2020)	MultiResUNet	Nuclei fluorescence microscopy image, Electron Microscopy, Skin lesion dermoscopy images, Polyps colonoscopy images	Intersection over Union

(H. Huang et al., 2020)	Full-scale connected U-Net	Liver CT scan, Spleen CT scan	Dice Coefficient
(Cherguif et al., 2019)	Deep convolution network based on U-Net	Brain tumor, MR images	Dice similarity coefficient
(W. Wang, Chen, et al., 2019)	Volumetric convolutional neural network (V-net)	Pulmonary lobes, chest CT images	Dice similarity coefficient
(Habijan et al., 2019)	U-Net	Whole heart, cardiac CT angiography	Dice similarity coefficient
(Hoang et al., 2019)	CNN: FCN-CRF, DRIU, and V-net	Liver, CT images	Dice similarity coefficient
(Zhou et al., 2019)		ventral nerve cord electron microscopy, Cell-CT scan, nuclei fluorescence images, Brain MRI, Liver CT scan, Lung nodule CT scan	Dice similarity coefficient and Intersection over Union

(Zhou et al., 2018)	UNet++ (nested U-Net)	colon polyp-microscopy, lung nodule-RGB image, liver-CT scan, and cell nuclei-CT scan	Intersection over Union
(Ait Skourt et al., 2018)	U-Net	Lung, thoracic CT scans	Dice similarity coefficient
(P. Wang et al., 2018)	Based on a Convolutional Neural Network (CNN), combines of U-Net and SegNet	Brain ventricles, ultrasound (US) scans	Intersection over Union (IoU), Dice coefficient, and pixel-wise accuracy
(S. Wang et al., 2017)	Deep Convolutional Neural Networks	Lung Nodule, CT scan	Dice similarity coefficient and Average Surface Distance

Table 2.1 shows the previous studies and the deep learning architectures used for the segmentation of medical images, part of segmentation, and medical image modality, and metrics used to evaluate the architecture are also presented.

2.7 RESEARCH GAPS

Accurate segmentation of medical images has several obstacles, involving the difficulty of representing different objects of interest, and noise present in the medical images.

The fundamental segmentation architectures for medical images are insufficient for complicated target segmentation among several medical image types (Ashkani Chenarlogh et al., 2022). This section discusses the previous Convolutional Neural network architectures for the selected articles and the gaps. The previous research that has been conducted on optimizing Convolutional Neural network segmentation techniques for medical images is summarized in Table 2.2 and also illustrates the advantages and disadvantages.

Table 2.2 Summary of Previous Research with Advantages and Disadvantages

Previous Research	**Segmentation Technique**	**Advantages**	**Disadvantages**
(Xu et al., 2023)	Improved U-Net 3+	Effective segmentation performance and feature extraction with fewer parameters	Less performance on a small dataset
(Q. Zhang et al., 2021)	Improved U-Net	Improves the segmentation quality	The dataset is small, contains only 22 images, and is from a single source
(Ibtehaz & Rahman, 2020)	MultiResUNet	Multires block and res path are utilized to improve the segmentation of medical images	The results of the segmentation get influenced if the boundaries of the medical image are not clear
(H. Huang et al., 2020)	full-scale connected U-Net	The use of full-scale skip connections	It does not have robust

		and deep supervision delivers a more accurate segmentation	segmentation performance on the small region of interest when the number of datasets is small
(Y. Cao et al., 2020)	DenseUNet	Takes advantage of U-Net and DenseNet, which helps to boost the segmentation performance. Data augmentation is also applied to the dataset to improve performance	The result shows poor performance for the segmentation of images with complex backgrounds
(J. L. Wang et al., 2020)	U-Net and curriculum learning strategy	Uses a multi-task attention mechanism to focus on specific features, and separate the image's foreground and background	Incapability to segment small region
(Chi et al., 2020)	U-Net	A dense connection is used to avoid feature vanishing, which improves the efficiency of segmentation. Multi-resolution feature concatenation and	Nodule with low contrast and unclear region is not detected. Nodule on the edge can not be distinguished from the outer region

		multi-scale pooling are used for the classification	
(Zhou et al., 2019)	UNet++	Nested structure and re-designed skip connections show significantly improved performance over U-Net.	Low robustness performance in learning about the location and boundary of the organs
(W. Wang, Chen, et al., 2019)	Volumetric convolutional neural network	Additional feature maps are generated to reduce misclassification	Only used thin slices of CT images
(Hoang et al., 2019)	CNN: FCN-CRF, DRIU, and V-net	Evaluates three CNN models for the CT image segmentation of the liver from multiple site datasets	The result shows that segmentation for low-dose and non-contrast-enhanced is not solved and remains an issue.
(Chunran & Yuanyuan, 2018)	FCN	Accepts any size of input images	Less visible speculation is absent. The low-intensity speculation is difficult to extract after the improvement

The literature review of the selected previous research is presented in Table 2.2 for the segmentation of medical images. The disadvantage of each technique is represented which shows the low performance of the segmentation for the medical images. Previous studies (Xu et al., 2023; Huang et al., 2020; J. L. Wang et al., 2020) demonstrate that the segmentation techniques have low performance when the number of datasets is small. Zhou et al. (2019), show significant segmentation performance but have low robustness performance in learning about the location and boundary of the region of interest. Studies (Ibtehaz & Rahman, 2020; Chi et al., 2020) show that performance gets influenced by unclear boundaries and unclear regions. Y. Cao et al. (2020), improve the segmentation performance for medical images but it shows poor segmentation performance for images with complex backgrounds. Other previous studies presented in Table 2.2 show the gaps in the optimized techniques used for the segmentation of medical images.

2.8 CHAPTER SUMMARY

This chapter presented the literature review on the existing segmentation techniques. The aim to review the literature is to understand the previous studies used for the segmentation of medical images. Section 2.2 provides information regarding the medical images. Section 2.3 and Section 2.4 provide the details of image analysis and image segmentation respectively. Deep learning for the segmentation of the medical images is explained in Section 2.5. U-Net and different variants using U-Net as a baseline architecture are presented in section 2.6, also the architecture, segmentation part, modality used for the medical image, and metrics used to evaluate the architecture are shown in this section. Section 2.7 shows the research gap findings of the selected previous studies, in this section advantages and disadvantages of the optimized segmentation technique for medical images of selected previous studies are presented.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 INTRODUCTION

This chapter describes the methodology that is implemented to conduct the research for this thesis. The general layout of this chapter is presented as follows: section 3.2 explains the research methodology, and section 3.3 presents the research map which illustrates the mapping between research questions, research objectives, and research methodology.

3.2 RESEARCH METHODOLOGY

The intention of this research is to develop A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images. The design of the research methodology is categorized into the following phases: research design, data collection, design of proposed architecture, evaluation and comparison of the algorithm, and documentation. Figure 3.1 briefly emphasizes the phases of how the proposed research is fulfilled.

Five phases are involved in the research methodology. The first phase concentrates on the review of the literature. Data collection and data pre-processing are discussed in the second phase. Architecture development is described in the third phase. The performance analysis of the developed architecture is discussed in the fourth phase. Finally, the research findings are documented and presented in phase five.

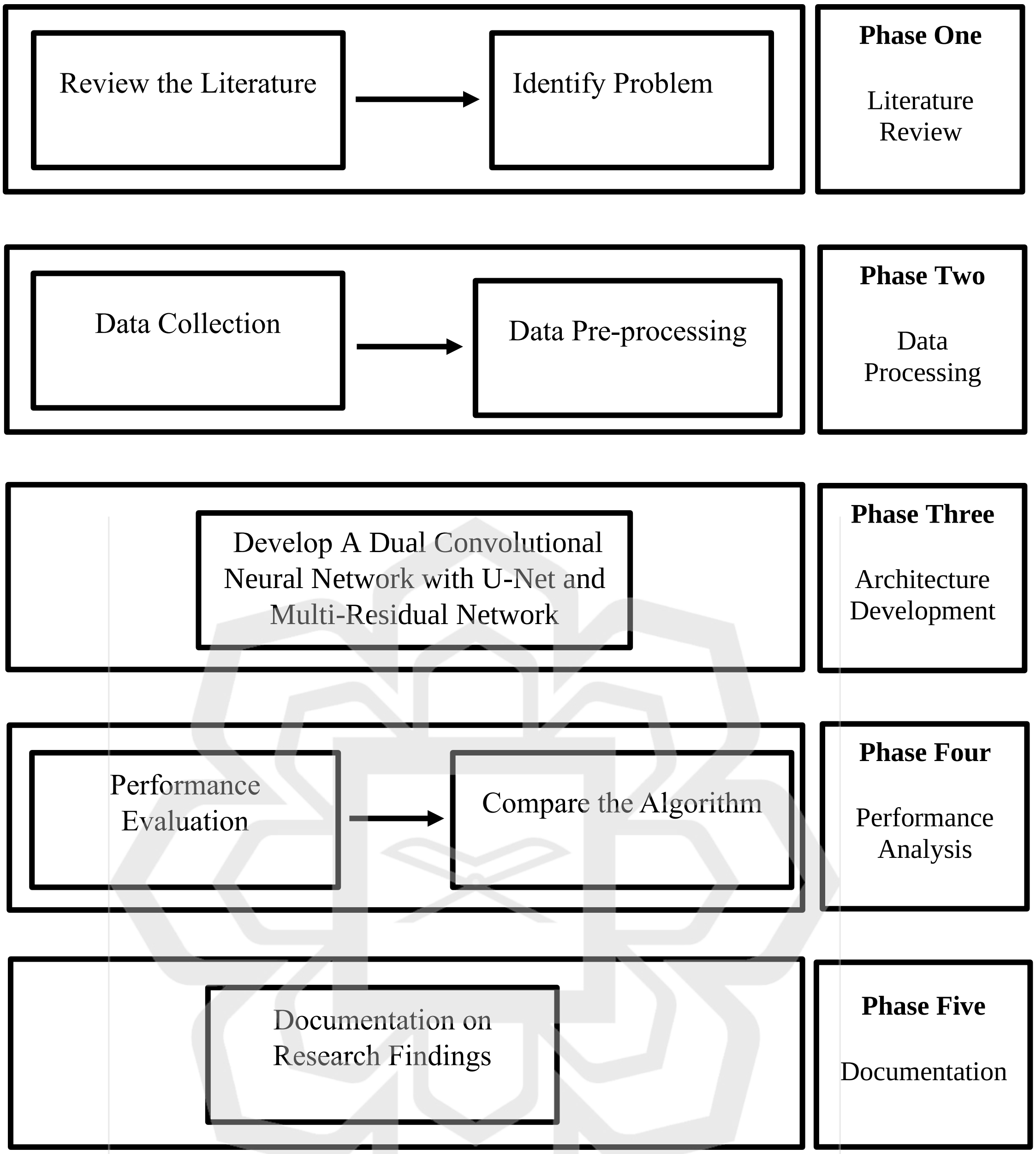


Figure 3.1 Research Methodology Phases

3.2.1 Phase One - Literature Review

The first phase is to review and analyze the earlier studies that are examined using queries 'image segmentation', 'medical images', 'deep learning', and 'convolutional neural network'. A literature review is essential to recognize the problem statement and identify the techniques that have been applied. According to the literature review, a convolutional neural network has been widely employed for the segmentation of medical images. This phase covers the basic knowledge of image segmentation and enables the identification of the numerous techniques and architectures of image segmentation and their advantages and disadvantages that are proposed to optimize the convolutional neural network.

3.2.2 Phase Two – Data Processing

This phase describes the data processing for this research, including data collection and pre-processing. Three different types of medical image datasets are collected: Gastrointestinal Polyp dataset, Brain tumor dataset, and Dental dataset.

3.2.2.1 Data Collection

3.2.2.1.1 Gastrointestinal Polyp Dataset

A gastrointestinal polyp image dataset is collected from the open access 'Kvasir-SEG dataset https://datasets.simula.no/kvasir-seg/'. These images in the dataset were gathered and validated by professional gastroenterologists from the vestre viken health trust in Norway.

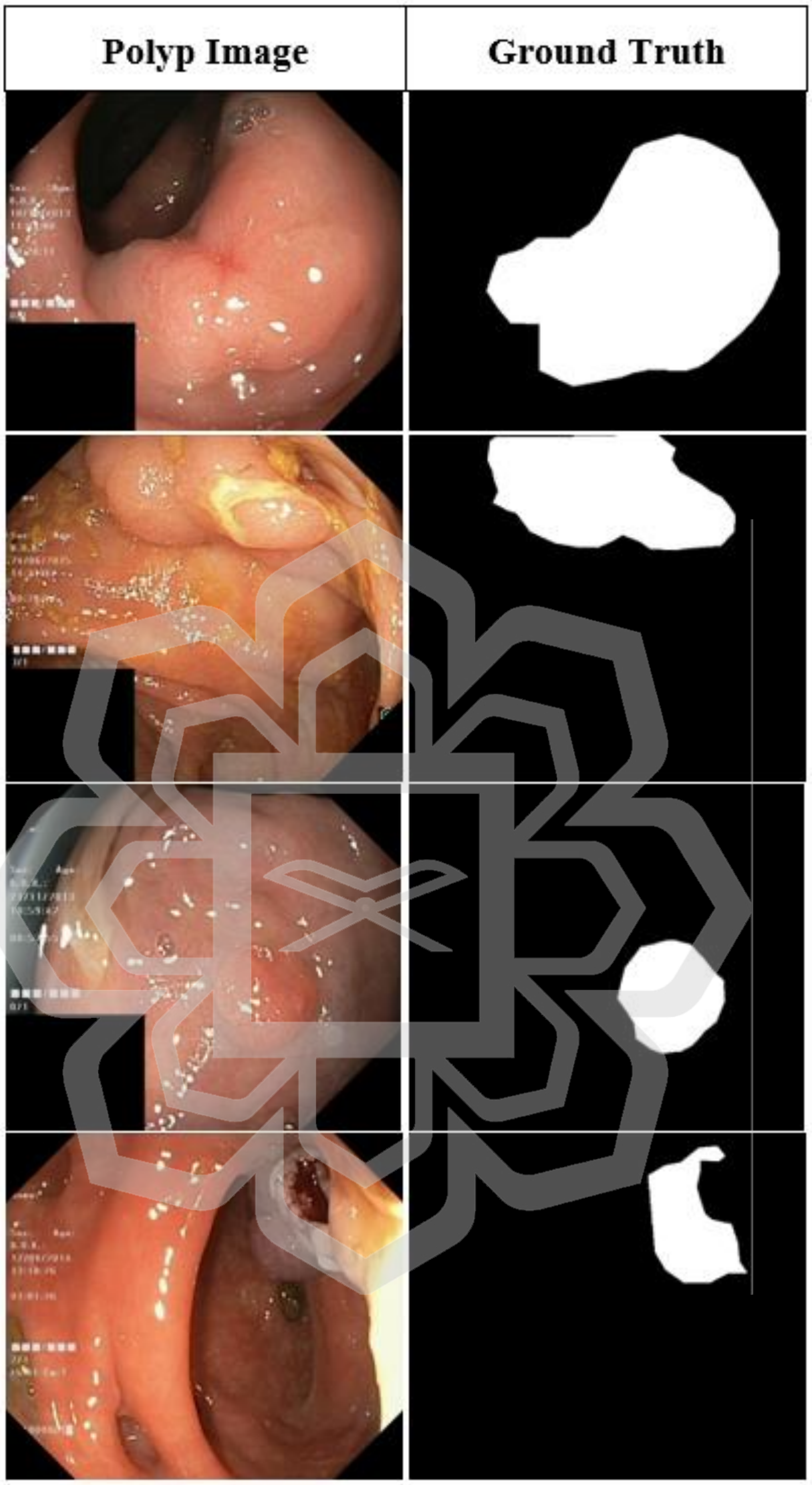


Figure 3.2 Example of Gastrointestinal Polyp Images and Corresponding Ground Truth

The dataset of gastrointestinal polyps from Kvasir-SEG consists of 1000 images and 1000 corresponding ground truth images. The ground truth is the mask of the region of interest (ROI) in the image; the ROI in the gastrointestinal dataset is a polyp in the image. Figure 3.2 depicts an example of a gastrointestinal polyp dataset from the Kvasir-SEG dataset. The resolution of the images in the dataset of Kvasir-SEG ranges from 332x487 to 1920x1072 pixels. As seen in Figure 3.2, the white mask in the foreground represents the region of interest (ROI) pixels indicating polyp tissue, whereas the black in the background does not have any positive pixels.

3.2.2.1.2 Brain Tumor Dataset

The dataset of brain tumor is gathered from https://figshare.com/. It is an open-access dataset, available publicly at https://figshare.com/articles/dataset/brain_tumor_dataset/1512427. The dataset of brain tumors contains contrast-enhanced MRI (CE-MRI), consisting of 3,064 images and 3,064 corresponding ground truths. The dataset has three types of brain tumors from 233 patients, comprising 1426 gliomas, 708 meningiomas, and 903 pituitary tumors. The tumor boundary was marked manually by the radiologists to produce the binary image of the tumor mask. The images in the brain tumor dataset have a resolution of 512x512. An example of a brain tumor dataset containing images and corresponding ground truth is shown in Figure 3.3.

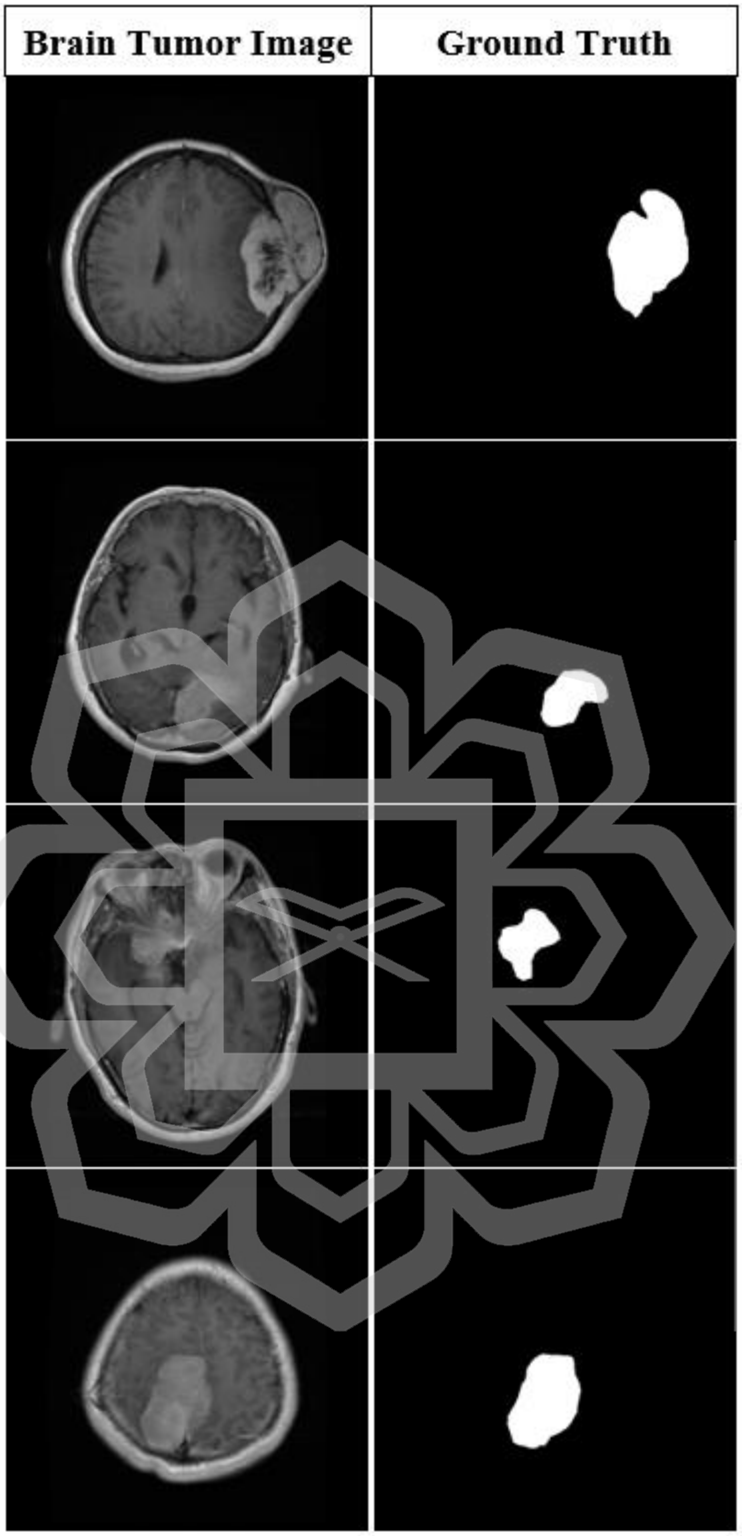


Figure 3.3 Example of Brain Tumor Images and Corresponding Ground Truth

3.2.2.1.3 Dental Dataset

The dental dataset is collected from the Oral Radiology Unit, Kulliyah of Dentistry, International Islamic University Malaysia (The dental dataset was approved by the International Islamic University Malaysia Research Ethics Committee, approval no. (IIUM IREC 2022-152)). The dental dataset contains cone beam computed tomography (CBCT) images of dental radiolucent lesions of 28 patients. The dental dataset includes 409 images and 409 corresponding ground truths. The lesions are marked by the experienced radiologist to generate the binary image for the ground truth. The dimension of the images in the dental dataset varies. An example of the dental dataset is illustrated in Figure 3.4.

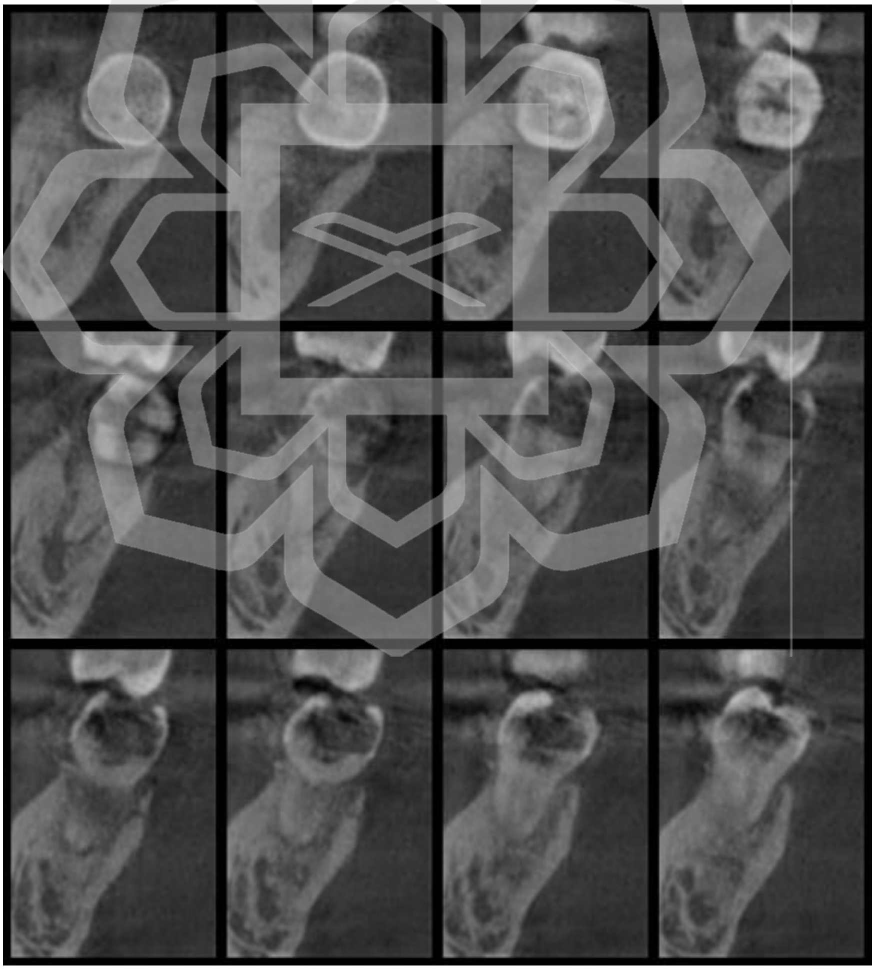


Figure 3.4 (a)

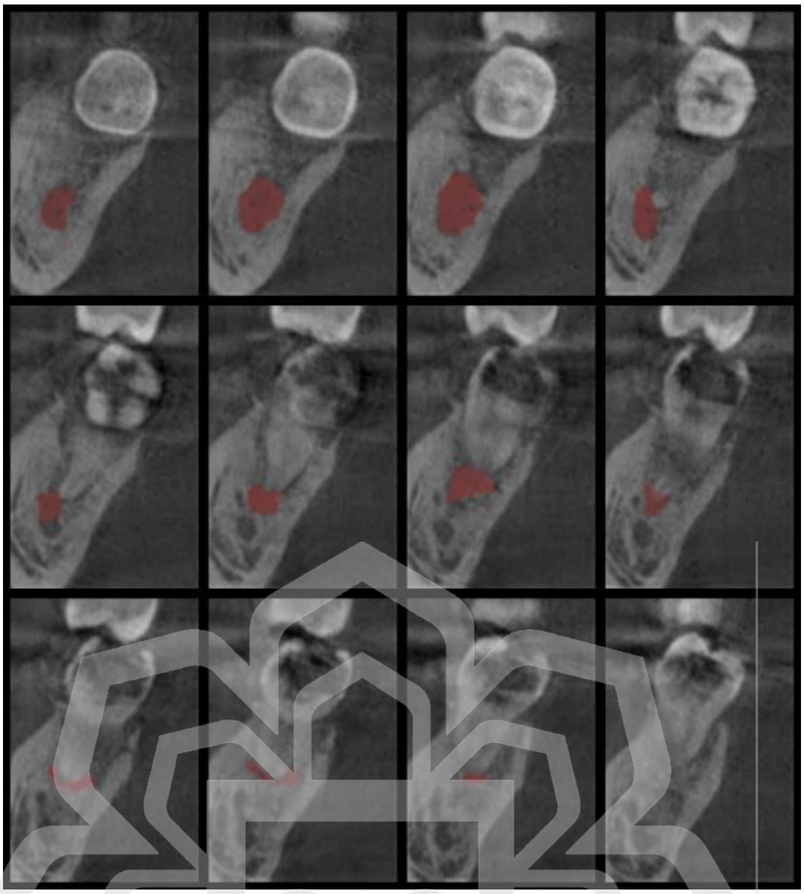


Figure 3.4 (b)

Figure 3.4 Example of Dental Dataset: (a) Images and (b) Corresponding Ground Truth

3.2.2.2 Data Pre-processing

After the collection of the data, pre-processing of the dataset is done. Pre-processing the dataset is crucial before the data is fed to the architecture. The ground truth is checked for each image for all three datasets. As illustrated in Figure 3.4, the dental dataset provided contains many slices of images in one image. So, the dental dataset images are cropped manually to get the individual slice of images and ground truth. After the images are cropped into separate pieces, binary masks are generated for the ground truth. The pseudocode to generate the binary mask from the image is presented in Algorithm 1. Figure 3.5 depicts the example of a cropped image and the corresponding ground truth from the dental dataset.

Algorithm 1: Pseudocode for conversion of image to mask

Input: RGB image, border image

Output: Binary mask image

1. Begin
2. Subtract the border image from the input image.
3. Subtract the border image from the result of step 2.
4. Apply a Gaussian blur.
5. Threshold the image using a value of 0.8.
6. Copy the resulting image to a new variable called mask.
7. Sum the values of the mask across the color channels and compare to a threshold of 760.
8. Threshold the resulting image with a value of 0.5.
9. Return the resulting binary mask image.
10. End

Algorithm 1 Pseudocode for Converting Image to Mask

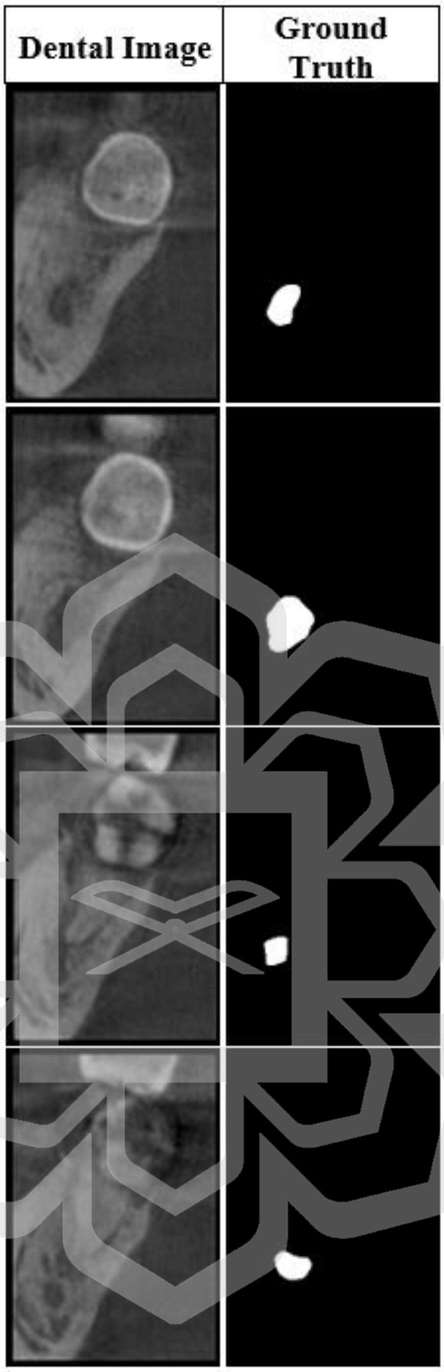


Figure 3.5 Example of Cropped Images and Corresponding Ground Truth of Dental Dataset

After checking the images and the corresponding ground truth, the images in all three datasets are then resized to 256x256 for the computational constraints. The collected datasets are categorized into three sections: training set, validation set, and testing set. The training set is data that is used to train the model. The model learns from the training set of data. The validation set is data employed during the training process to evaluate the performance of the model on a separate set of data that it has not been trained on. The validation set is utilized to evaluate after each epoch of training to tune hyperparameters like the learning rate and prevent overfitting. The testing set is data employed to evaluate the final performance of the model on the unseen dataset after training.

The gastrointestinal polyp dataset is divided into 80% training set (800 images), 10% validation set (100 images), and 10% testing set (100 images). The brain tumor dataset is split into an 80% training set (2,452 images), a 10 % validation set (306 images), and a 10 % testing set (306 images). The dental dataset is also divided into an 80% training set (329 images), a 10% validation set (40 images), and a 10% testing set (40 images). After splitting the datasets into the train, valid, and test sets, data augmentation is implemented on the training set to increase the number of data for training that improves the performance. The data augmentation increases the gastrointestinal polyp training set (800 images) up to 13,600 images. After applying data augmentation, the brain tumor training dataset (2,452 images) is increased to 36,780 images. The dental training dataset (329 images) is also increased to 4,935 images by applying data augmentation. The augmentations used to increase the training set are center crop, crop, random rotation 90-degree, elastic transform, transpose, grid distortion, optical distortion, grayscale, vertical flip, horizontal flip, random brightness contrast, random gamma, huesaturationvalue, RGBshift, random brightness, random contrast. Figure 3.6 presents an example of image augmentation applied to the training dataset.

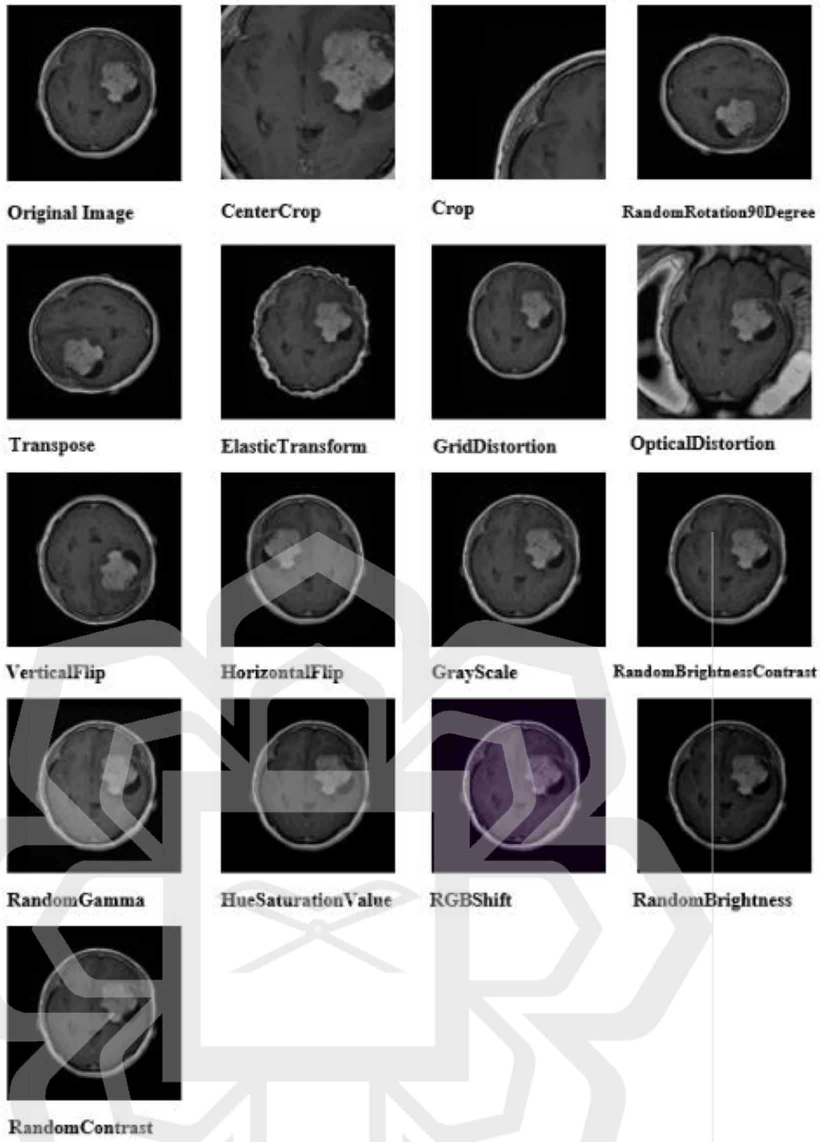

Original Image
CenterCrop
Crop
RandomRotation90Degree
Transpose
ElasticTransform
GridDistortion
OpticalDistortion
VerticalFlip
HorizontalFlip
GrayScale
RandomBrightnessContrast
RandomGamma
HueSaturationValue
RGBShift
RandomBrightness
RandomContrast

Figure 3.6 (a)


Original Image
CenterCrop
Crop
RandomRotation90Degree
Transpose
ElasticTransform
GridDistortion
OpticalDistortion
VerticalFlip
HorizontalFlip
GrayScale
RandomBrightnessContrast
RandomGamma
HueSaturationValue
RGBShift
RandomBrightness
RandomContrast

Figure 3.6 (b)

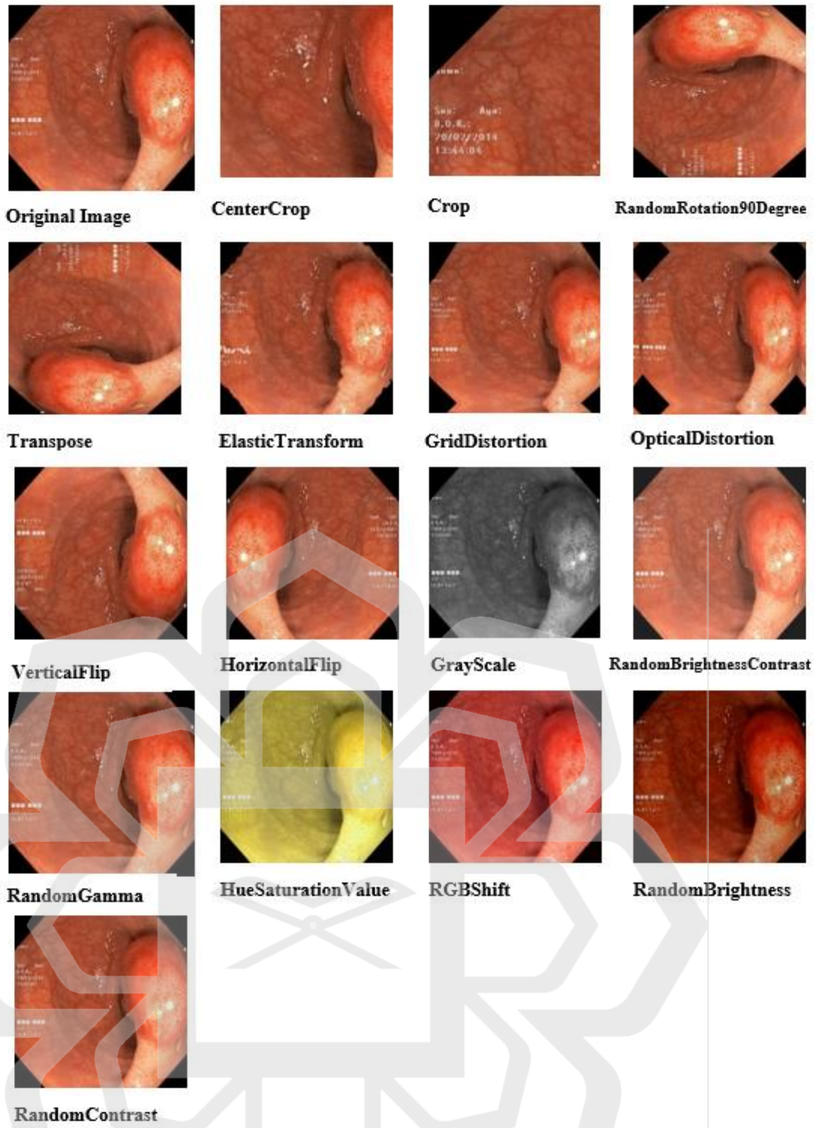


Figure 3.6 (c)

Figure 3.6 Example of Image Augmentation (a) Brain Dataset, (b) Dental Dataset, (c) Gastrointestinal Poly Dataset

The center crop is an augmentation technique used to crop the central area of the image. The crop data augmentation technique is used to crop the images into a particular dimension. In Figure 3.6, the top part of the image is cropped. Random rotation 90 degrees rotates the image randomly by 90 degrees. Transpose augmentation flips an image along its diagonal axis. Elastic transform is an augmentation technique used in image processing to create a new image sample by deforming the image. The grid distortion technique distorts the image by moving grids of pixels around. In grid distortion, a grid of control points is defined on the image, and the position of these control points is randomly shifted. The image is then distorted based on the movement of these control points, resulting in a transformed image that is similar to the original image but with some distortion. The optical distortion augmentation technique distorts certain elements in the image that simulates the effect of the optical element on the image. Vertical flip augmentation flips the image vertically around the x-axis. The horizontal flip augmentation technique flips the image horizontally around the y-axis. The grayscale augmentation technique converts an image to a single channel that is a grayscale image. Randombrightnesscontrast allows changing the contrast and brightness of the image randomly. Randomgamma is a technique that controls the brightness and intensity of the image. Huesaturationvalue is an augmentation technique that represents the color of an image based on the hue, saturation, and value. It allows random value adjustment for hue, saturation, and value from a particular range of values. Rgbshift randomly shifts the value of the red, green, and blue channels of the image. Random brightness and random contrast randomly adjust the brightness and the contrast of the image, respectively.

3.2.3 Phase Three - Architecture Development

This phase presents the architecture development of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images. A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is built using two networks. The first network is U-Net, which utilizes pre-trained Resnet-50 as an encoder part, and Squeeze and Excitation block. In addition, Atrous Spatial Pyramid Pooling (ASPP) is used as a bridge between the encoder and decoder. The decoder part in the first network is built from scratch. The second network is Multi-Res U-net, in which Multi-Residual Network (multi-resnet) block is used in both the encoder part and the decoder part, and a squeeze and excitation block is also used in the second network. In addition, Atrous Spatial Pyramid Pooling (ASPP) is utilized as a bridge between the encoder and decoder in Multi-Res U-net. Figure 3.7 illustrates the flowchart of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images.

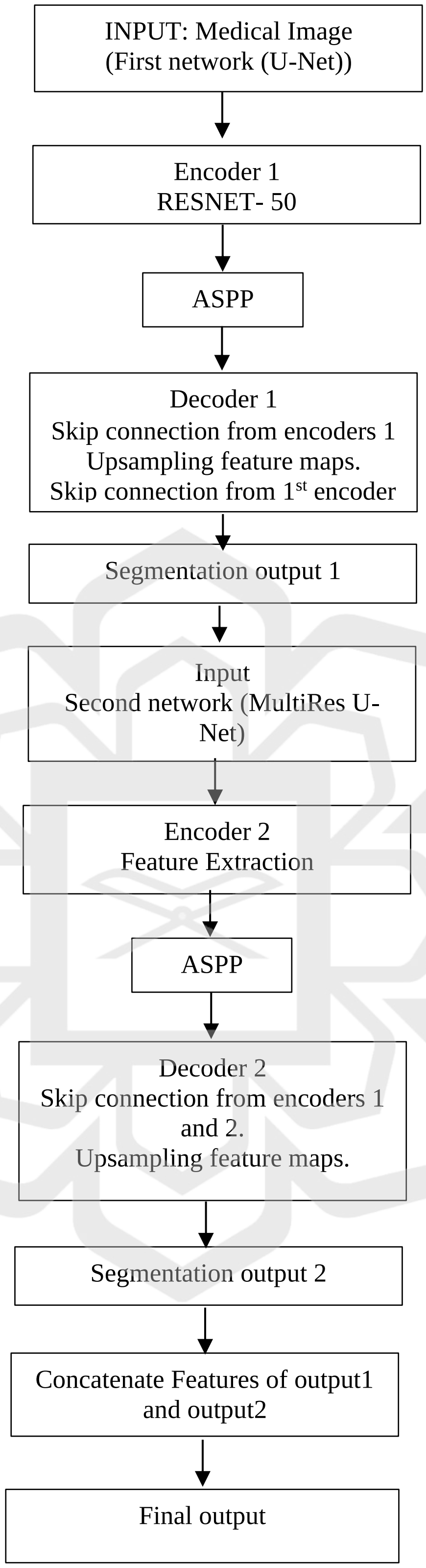


Figure 3.7 Flowchart of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network

In Figure 3.7, the medical image as input is fed to the first network (U-Net) of the architecture, resulting in the generation of the predicted mask (segmentation output 1). The output from the first network serves as the input for the second network (MultiRes U-Net). The second network generates another segmentation mask (segmentation output 2). Finally, output1 and output2 are concatenated to generate the final segmentation output. The development of the architecture of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is explained in detail in Chapter Four.

3.2.4 Phase Four – Performance Analysis

This phase includes performance analysis of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network. This research examines the performance of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, FCN, Unet_Resnet, and U-Net architectures. After performance analysis, a comparison is conducted between the U-Net, FCN, Unet_Resnet, and A Dual Convolutional Neural Network with U-Net and Multi-Residual Network to analyze which algorithm performs best for the segmentation of medical images (see section 5.5).

The performance of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images is evaluated using the metrics of Intersection over Union (IOU), Dice Similarity Coefficient (DSC), Precision, and Recall (see section 5.4), which is explained in detail in Chapter Five.

3.2.5 Phase Five - Documentation

In this phase, the results of the proposed architecture and existing architectures for medical image segmentation, performance analysis, and comparison of results are documented. The achievement and the performance of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, are compared and explained in Chapter Five.

3.3 RESEARCH MAP

This section presents a comprehensive mapping between the formulated research questions, established research objectives, and the selected research methodology. The research map serves as a visual guide, offering a clear depiction of the alignment and interconnection among these critical components. Figure 3.8 below provides a visual representation that explains the connections between research questions, research objectives, and research methodology in this research.

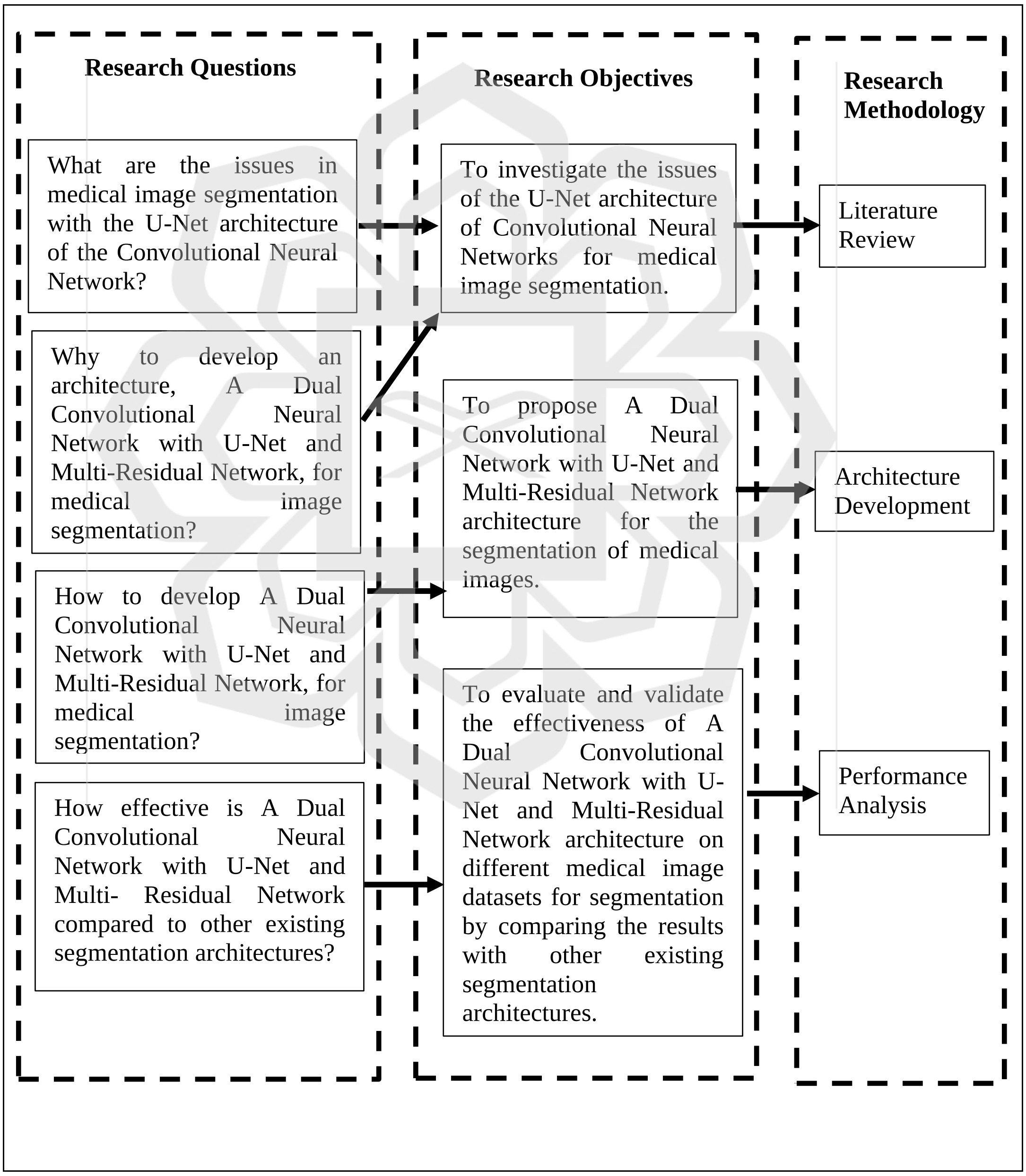


Figure 3.8 Research Map

3.4 CHAPTER SUMMARY

The research methodology for this thesis is represented in this chapter. Each phase of methodology for this research that is utilized to develop A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images is explained. The literature review, data processing, architecture development, performance analysis, and documentation are discussed in this chapter. Furthermore, this chapter illustrates the mapping between research questions, research objectives, and research methodology. In chapter four, the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, will be discussed for the medical image segmentation.

CHAPTER FOUR

EXPERIMENTAL DESIGN

4.1 INTRODUCTION

The experimental design of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the segmentation of medical images is presented in this chapter. This chapter explains all blocks involved in the proposed architecture and illustrates the pseudocode for each block. In this chapter, Section 4.2 explains the design of the proposed architecture. Sub-sections of Section 4.2 present all the blocks used to develop a proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network. The summary of this chapter is provided in Section 4.3.

4.2 DESIGN OF PROPOSED ARCHITECTURE

In this research, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network architecture is proposed for the segmentation of medical images. Figure 4.1 illustrates the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network. The architecture of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network consists of two networks: the first network is the U-Net, and the second network is the MultiRes U-Net.

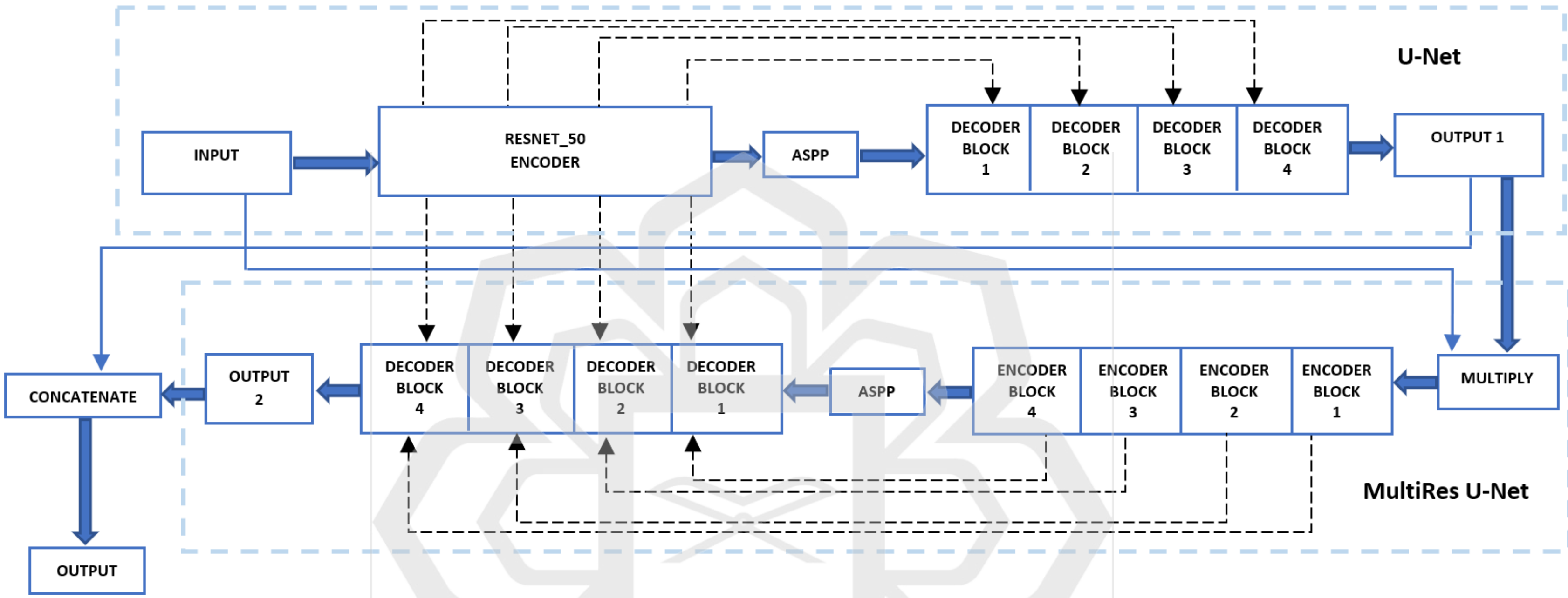


Figure 4.1 Architecture of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network

Figure 4.1 presents the architecture of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, in which the first network U-Net utilizes the pre-trained Resnet_50 as an encoder, and the second network, MultiRes U-Net utilizes the multires block and res path. The U-Net and MultiRes U-Net of the proposed architecture comprises an encoder, decoder, and ASPP blocks. This section explains all the blocks of U-Net and MultiRes U-Net of the proposed architecture of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network.

4.2.1 Encoder-Decoder Block

The encoder in the network serves as a feature extractor that extracts the relevant features from the input image through a sequence of encoder blocks. The decoder takes the extracted features from the encoder and constructs a segmentation mask. The encoder and decoder are connected via a bridge. The encoder receives the input and passes it through the entire encoder network, where the number of filters increases, and the spatial dimension is reduced by half. While in the decoder network, the number of filters decreases, and the feature map is upsampled by factor 2.

4.2.2 Encoder in First Network (U-Net)

In A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, RESNET_50 is utilized as a pre-trained encoder in the first network, U-Net, that has already been trained to learn features from ImageNet and could be applied to another task in order to expedite model development. The pseudocode for the encoder in the first network, U-Net, is shown in Algorithm 2.

Algorithm 2: Pseudocode for Encoder in First Network

Input: Tensor, input tensor to the encoder

Output: Output: tensor, output tensor of the encoder

skip_connections: list, list of tensors representing the outputs of layers in the ResNet50 model

1. skip_connections = empty list
2. model = ResNet50 model with inputs
3. names = list of strings containing the names of the layers ["input_1", "conv1_relu", " conv2_block3_out", "conv3_block4_out"]
4. for name in names do
5. layer_output = output tensor of the layer in model with name=name
6. append layer_output to skip_connections
7. end for
8. output = output tensor of the layer in model with name="conv4_block6_out"
9. return output, skip_connections

Algorithm 2 Pseudocode for Encoder in First Network

Algorithm 2 explains the pseudocode for the encoder in the first network (U-Net), in which pre-trained RESNET_50 is used as an encoder. It returns the output tensor of the layer which is a feature map for the next block and it also returns the skip connection which is used by the decoder block.

4.2.3 Atrous Spatial Pyramid Pooling (ASPP)

The encoder is connected to the decoder via an Atrous Spatial Pyramid Pooling bridge. A Dual Convolutional Neural Network with U-Net and Multi-Residual Network utilizes the advantage of Atrous Spatial Pyramid Pooling (ASPP). In ASPP, parallel atrous convolution layers are used with different dilation rates that increase the receptive field to capture multiscale information. The atrous convolution can store more information about the image to increase the accuracy of segmentation. Atrous convolution is defined in Equation 4-1

$$y[i] = \sum_{k} x[i + r \cdot k]w[k]$$

Equation 4-1

Where $y[i]$ is the output of the pixel whose index is i, k is the location index in kernels, r is the dilation rate that determines the size of convolution kernels, and w is filters with weight. Atrous convolution allows to adjust the field of view of a filter by changing the value of the dilation rate. The structure of ASPP is illustrated in Figure 4.2

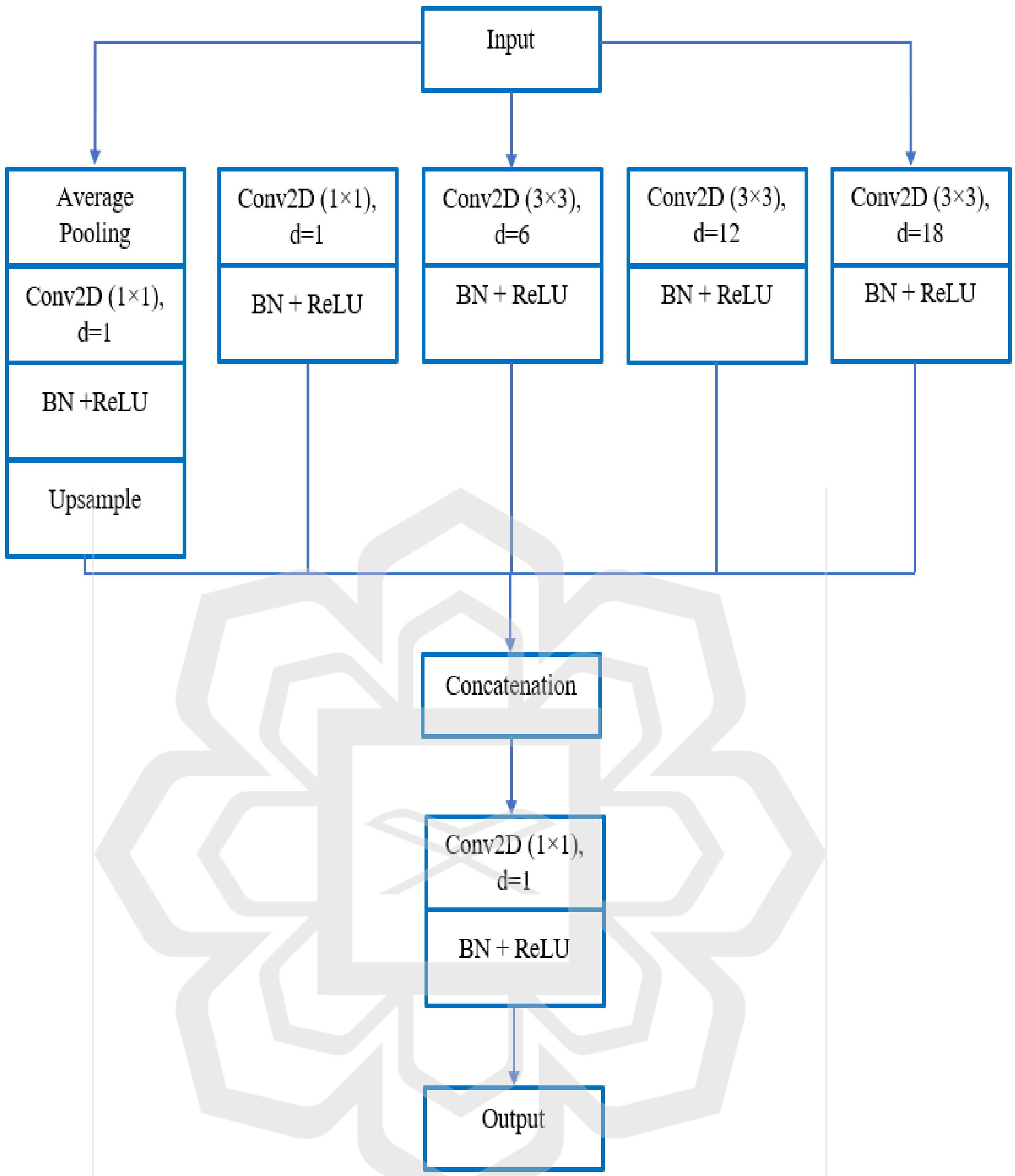


Figure 4.2 Atrous Spatial Pyramid Pooling (ASPP)

The ASPP structure consists of four parallel atrous convolutions that have different atrous rates. It includes one 1×1 convolution, and three 3×3 convolutions with 6, 12, and 18 dilation rates, respectively, followed by the batch normalization (BN) and rectified linear unit (ReLU) activation function. Average pooling is also used. The resulting features from each branch are upsampled bilinearly to the size of the input and then the concatenation is implemented and then it is passed through further 1×1 convolution. ASPP is applied to the feature map generated by the encoder part, and the resultant feature map from the ASPP is given to the decoder part. The pseudocode for ASPP is depicted in Algorithm 3.

Algorithm 3: Pseudocode for Artous Spatial Pyramid Pooling (ASPP)

Input:

x: tensor, input tensor to the ASPP block
filter: int, number of filters in the convolutional layers (default=256)

Output:

y: tensor, output tensor of the ASPP block

1. shape = shape of x
2. y1 = Average Pooling 2D of x with pool_size=(shape[1], shape[2])
3. y1 = Conv2D layer with filter filters, kernel_size=1, and same padding of y1
4. y1 = Batch Normalization and ReLU activation of y1
5. y1 = UpSampling 2D of y1 with size=(shape[1], shape[2]) and bilinear interpolation
6. y2 = Conv2D layer with filter filters, kernel size=1, dilation rate=1, same padding, and No bias of x
7. y2 = Batch Normalization and ReLU activation of y2
8. y3 = Conv2D layer with filter filters, kernel_size=3, dilation_rate=6, same padding, and no bias of x
9. y3 = Batch Normalization and ReLU activation of y3
10. y4 = Conv2D layer with filter filters, kernel_size=3, dilation_rate=12, same padding, and no bias of x
11. y4 = Batch Normalization and ReLU activation of y4
12. y5 = Conv2D layer with filter filters, kernel_size=3, dilation_rate=18, same padding, and no bias of x
13. y5 = Batch Normalization and ReLU activation of y5
14. y = Concatenate tensors y1, y2, y3, y4, and y5 along the channel axis
15. y = Conv2D layer with filter filters, kernel_size=1, dilation_rate=1, same padding, and no bias of y
16. y = Batch Normalization and ReLU activation of y
17. return y

Algorithm 3 Pseudocode for ASPP

Algorithm 3 presents the pseudocode for Atrous Spatial Pyramid Pooling. It acts as a connection between the encoder part and the decoder part. The input of the ASPP is the feature map from the encoder block and the output of the ASPP is the feature map which is fed to the decoder block. The different dilation rates 6, 12, and 18 are used in Aspp as shown in Algorithm 3.

4.2.4 Decoder in First Network (U-Net)

In A Dual Convolutional Neural Network with U-Net and Multi-Residual Network architecture, each block in the decoder conducts a 2×2 bilinear upsampling on the feature maps of the input, which increases the dimension of the input feature maps. The decoder block in the first network, U-Net, is illustrated in Figure 4.3.

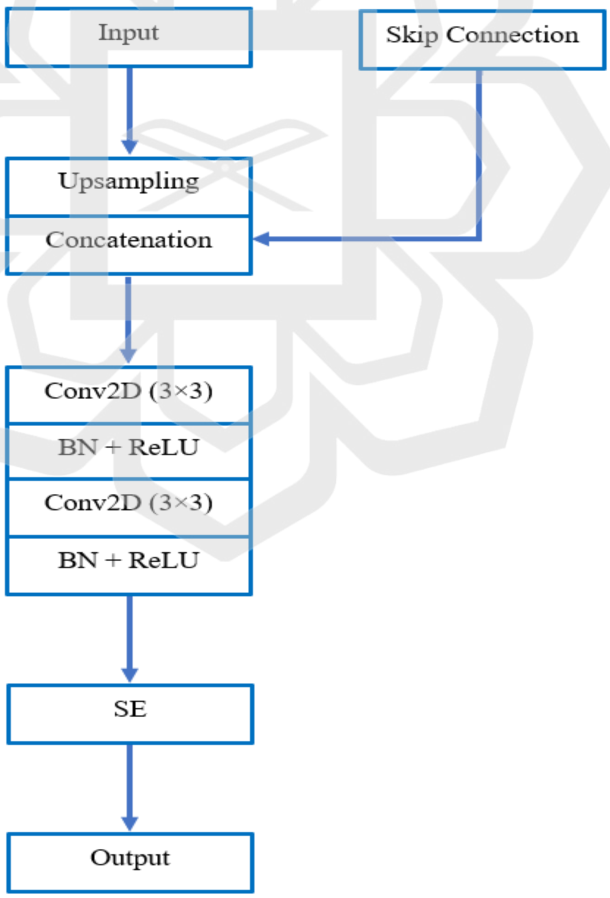


Figure 4.3 Decoder Block in the First Network

The architecture of the decoder in the first network of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is the same as basic U-Net, except squeeze and excitation block is used, and also, bilinear upsampling is used instead of transpose convolution. It utilizes the feature maps of skip connection from the encoder (RESNET_50) of the first network. Skip connection skips some layers and transfers the output of one layer to the subsequent layers as the input (instead of only the next one). Skip connection enables the transfer feature map from the encoder block to the corresponding layers in the decoder block. These skip connections allow the network to retrieve the spatial information lost due to pooling operations. In the decoder of the first network, the decoder block performs upsampling on the input feature maps, and a skip connection from the pre-trained RESNET_50 encoder is used. After the upsampling in the decoder of the first network, concatenate is applied to the appropriate skip connection from the encoder of the first network to the output feature maps.

After concatenation, it is followed by a convolution block which includes two 3×3 convolution operations, and each convolution operation is then followed by batch normalization (BN) and the ReLU activation function. A convolution block is used to extract the features from the input. After that, the squeeze and excitation (SE) block (explained in 4.2.4.1) is used in the decoder. Figure 4.4 shows the convolution block with squeeze and excitation (SE) block. The pseudocode for the convolution block is illustrated in Algorithm 4.

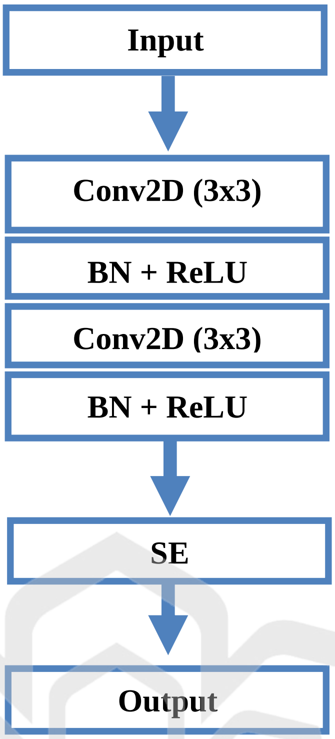


Figure 4.4 Convolution Block

Algorithm 4: Pseudocode for Convolution Block (conv-Block)

Input:

x: tensor, input tensor to the Conv-Block

filters: int, number of filters in each convolutional layer

Output:

x: tensor, output tensor of the Conv-Block

1. x = Convolutional layer with filters, 3x3 kernel size, and "same" padding of x
2. x = Batch Normalization layer of x
3. x = ReLU activation layer of x
4. x = Convolutional layer with filters, 3x3 kernel size, and "same" padding of x
5. x = Batch Normalization layer of x
6. x = ReLU activation layer of x
7. x = Squeeze-and-Excite Block of x **Algorithm 6 section 4.2.4.1**
8. return x

Algorithm 4 Pseudocode for Convolutional Block

The pseudocode of the convolutional block is shown in Algorithm 4, which includes two 3×3 convolutional layers, each convolutional layer is then followed by the batch normalization (BN) and ReLU activation function, and after that, the Squeeze and Excitation block follows.

The pseudocode for the decoder of the first network (U-Net) is illustrated in Algorithm 5.

Algorithm 5: Pseudocode for Decoder in First Network

Input:

inputs: tensor, input tensor to the Decoder

skip_connections: list of tensors, list of skip connections from the Encoder

Output:

x: tensor, output tensor of the Decoder

```
num_filters = list of integers [256, 128, 64, 32]
skip_connections.reverse()
x = inputs
for i = 0 to length of num_filters - 1 do
        x = UpSampling2D with scale (2, 2) and interpolation "bilinear" applied to x
        x = Concatenate of x and skip_connections[i]
        x = conv_block(x, num_filters[i])  Algorithm 4
return x
```

Algorithm 5 Pseudocode for Decoder

In Algorithm 5, the pseudocode for the decoder in the first network is demonstrated. The algorithm shows that the input for the decoder block is the tensor from the previous block and it has skip connections from the encoder block of the first network (U-Net).

4.2.4.1 Squeeze and Excitation (SE) Block

Squeeze and Excitation (SE) is a channel-wise attention mechanism that calibrates each channel independently, which improves the performance of the network. The squeeze and excitation attention mechanism adds a parameter to each channel that rescales them independently, so it becomes more sensitive towards significant features while ignoring the irrelevant features. Figure 4.5 shows the squeeze and excitation block.

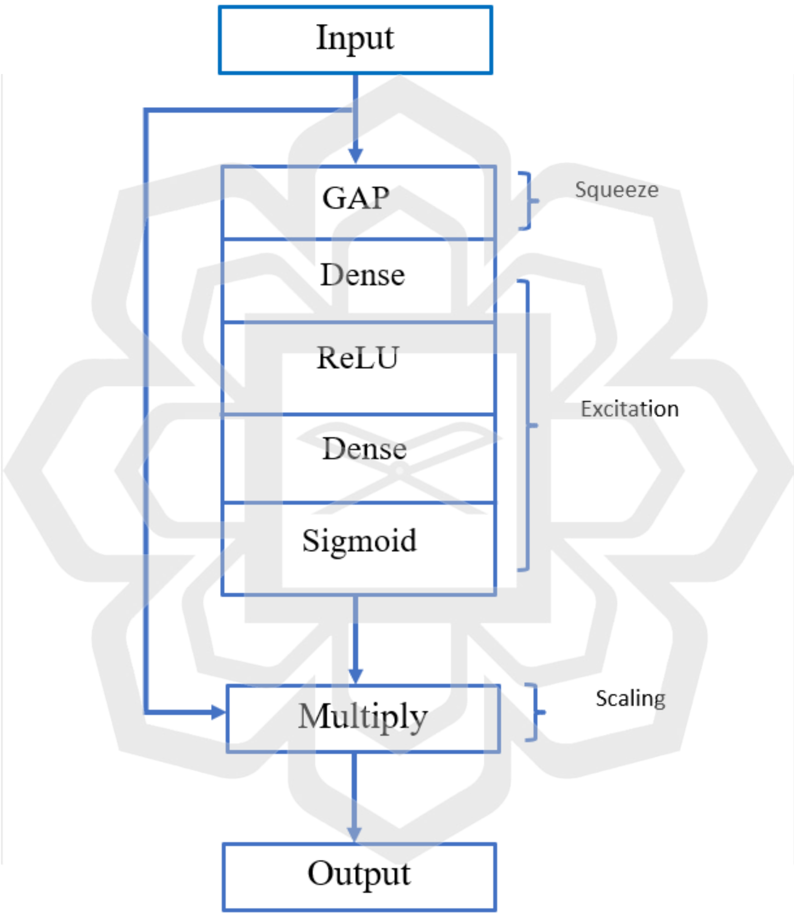


Figure 4.5 Squeeze and Excitation Block

The squeeze and excitation block are comprised of three components: Squeeze, Excitation, and Scaling.

Squeeze operation extracts the global information of the feature map from each channel. The feature map is the convolution layer's output of size:

$$B \times H \times W \times C$$

Where B is the batch size, H is the height of the feature map, W is the width of the feature map, and C is the number of channels in the feature map.

In Squeeze, the pooling operation is used to decrease the spatial dimension of the feature map. Global Average Pooling (GAP) is used in squeeze operation to reduce the feature map B×H×W×C to B×1×1×C.

Now the feature map is decomposed to a smaller dimension (B×1×1×C). The size H×W of each channel is reduced to a singular vector. In the Excitation operation, a bottleneck structure fully connected Multi-layer Perceptron (MLA) is used. Multi-layer perceptron generates the weights that are used to scale each channel of the feature map.

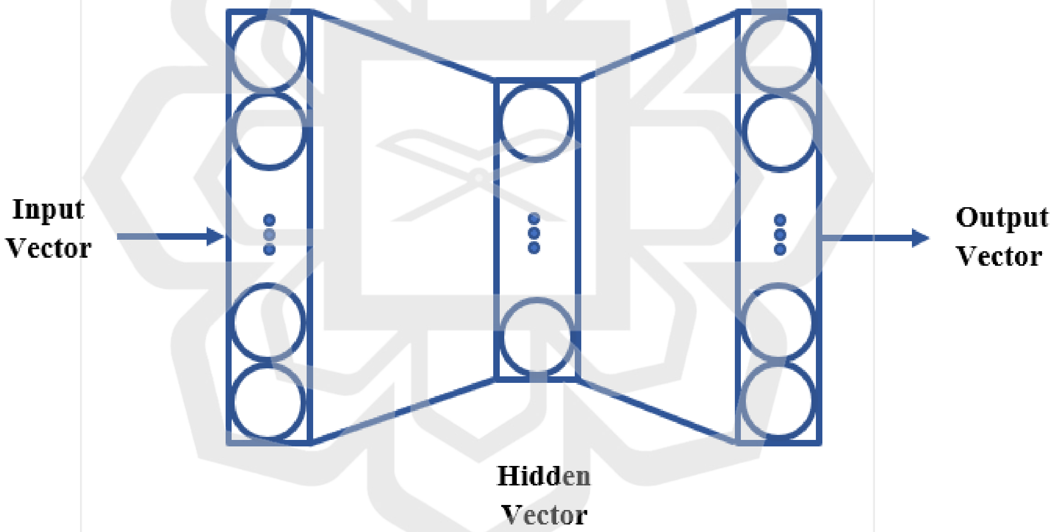


Figure 4.6 Multi-layer Perceptron (MLP)

Figure 4.6 shows the Multi-layer Perceptron (MLP), which includes input, output, and hidden layers. The hidden layer is utilized to reduce the number of features determined by the reduction factor r. The dimensionality of the feature map at each layer of the Multi-Layer Perceptron (MLP) is:

i. The shape of the input is (B×1×1×C), which is reduced to B×C. Thus, in the input layer, there is C number of neurons.

ii. The hidden layer reduces the number of neurons by the reduction factor r. As a result, the hidden layer contains a C/r number of neurons.

iii. The number of neurons is increased back in the output layer to the same dimension as the input, that is, C number of neurons.

Basically, the MLP receives (B×1×1×C) as an input and returns the same dimension as an output (B×1×1×C).

In scaling operation, after obtaining the excited tensor (B×1×1×C) from the excited operation, a sigmoid activation function is then applied to the excited tensor. An activation function sigmoid converts the value of the tensor in the range of 0 and 1. After that, the element-wise multiplication is carried out between the input feature map and the output of the sigmoid activation function. If the value is close to 0, that means the channel is less important, and the feature channel's value will be reduced. On the other hand, if the value is close to 1, then the feature channel is crucial. The pseudocode for squeeze and excitation is depicted in Algorithm 6.

Algorithm 6: Pseudocode for Squeeze-and-Excite Block (SE-Block)

Input:

inputs: tensor, input tensor to the SE-Block

ratio: int, reduction ratio for the SE-Block (default=8)

Output:

x: tensor, output tensor of the SE-Block

1. init = inputs
2. channel_axis = -1
3. filters = shape of init along channel_axis
4. se_shape = (1, 1, filters)
5. se = Global Average Pooling 2D of init
6. se = Reshape se_shape of se
7. se = Dense layer with filters//ratio units and ReLU activation, with no bias
8. se = Dense layer with filters units and sigmoid activation, with no bias
9. x = Multiply element-wise of inputs and se
10. return x

Algorithm 6 Pseudocode for Squeeze and Excitation

4.2.5 Encoder in Second Network (MultiRes U-Net)

The encoder in the second network, MultiRes U-Net, in A Dual Convolutional Neural Network with U-Net and Multi-Residual Network consists of a Multires block and Res path.

4.2.5.1 Multires Block

Multires block contains the features from the DenseNet and ResNet. In addition, squeeze and excitation block is also added in the Multires block in the second network. The Multires block is illustrated in Figure 4.7

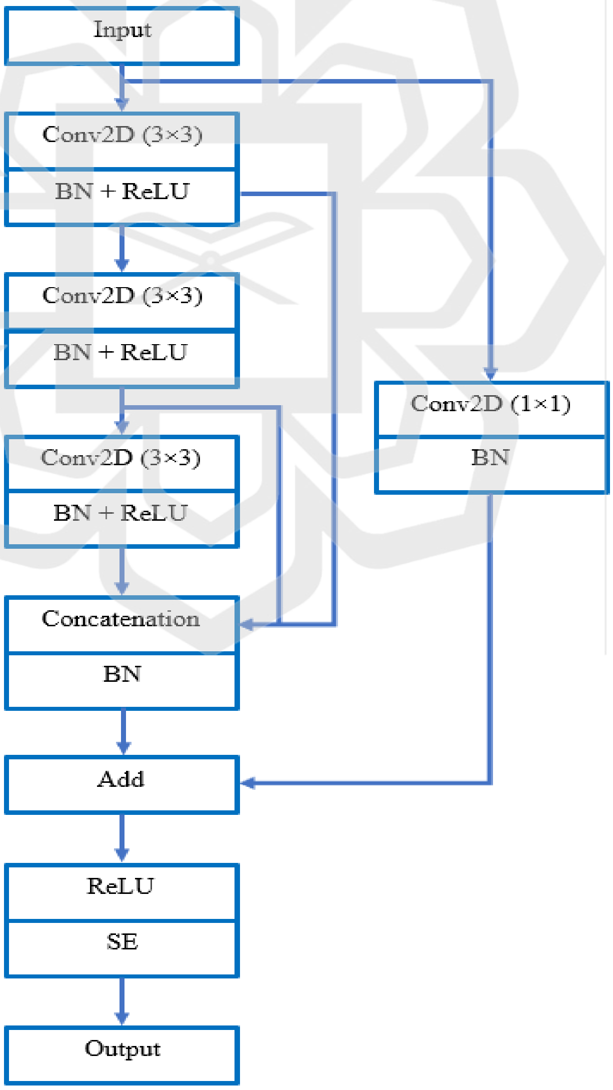


Figure 4.7 Multires Block

Multires block consists of a sequence of three 3×3 convolutional layers. These three convolutional layers are followed by batch normalization (BN) and the ReLU activation function. The output of these three 3×3 convolutional layers is concatenated, which is the feature of DenseNet. Multires block also contains shortcut connection or residual connection, which is the feature of ResNet. It performs an element-wise addition of the input and the concatenation feature. It is then followed by the ReLU activation function and squeeze and excitation block. The pseudocode for the Multires block is shown in Algorithm 7.

Algorithm 7: Pseudocode for Multires Block

Input:
x: tensor, input tensor to the Multires-Block
num_filters: int, number of filters for the Multires-Block
alpha: float, scaling factor for the number of filters (default=1.67)

Output:
x: tensor, output tensor of the MultiRes Block

1. W = num_filters * alpha
2. x0 = x
3. x1 = Conv2D layer with int(W*0.167) filters and 3x3 kernel applied to x0
4. x2 = Conv2D layer with int(W*0.333) filters and 3x3 kernel applied to x1
5. x3 = Conv2D layer with int(W*0.5) filters and 3x3 kernel applied to x2
6. xc = Concatenate x1, x2, and x3 along channel axis
7. xc = BatchNormalization layer applied to xc
8. nf = int(W*0.167) + int(W*0.333) + int(W*0.5)
9. sc = Conv2D layer with nf filters and 1x1 kernel applied to x0, with no activation function
10. x = Element-wise addition of xc and sc
11. x = ReLU activation function applied to x
12. x = BatchNormalization layer applied to x
13. x = Squeeze-and-Excite Block applied to x
14. return x

Algorithm 7 Pseudocode for Multires Block

Algorithm 7 shows the pseudocode for Multires Block. The three 3×3 convolutional layers are followed by batch normalization and the ReLU activation function. The output of these three 3×3 convolutional layers is concatenated. Shortcut connection or residual connection in Algorithm 7 uses the 1×1 convolutional layer, having no activation function. It performs an elementwise addition of the input and the concatenation feature. It is followed later by the ReLU activation function, as well as the squeeze and excitation block.

4.2.5.2 Res Path

In Multires U-Net, the second network of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, the skip connection is changed to res path. Res path is used in between the encoder and the decoder of the second network. Res path reduces the semantic gap between the encoder and the decoder features. The features of the encoder are passed through a sequence of res path, and then the res path's output is concatenated with the features of the decoder.

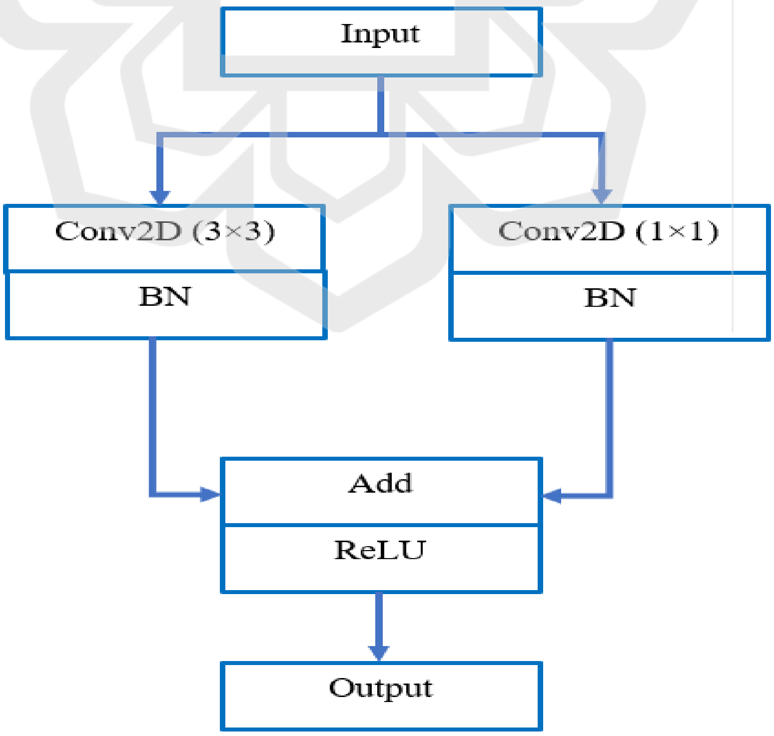


Figure 4.8 Res Path

Figure 4.8 shows the Res path. The Res path takes the input that is features from the encoder and passes it through the convolutional layer of 3×3 filters followed by batch normalization and is accompanied by the residual connection of 1×1 convolution, which is also followed by batch normalization. The features of these 3×3 convolution and 1×1 convolution are added, and then it is followed later by the ReLU activation function. The pseudocode for the Res path is illustrated in Algorithm 8.

Algorithm 8: Pseudocode for Res Path Block

Input:

x: tensor, input tensor to the residual path function

num_filters: int, number of filters to be used in convolution layers

length: int, number of convolution layers in the residual path function

Output:

x: tensor, output tensor of the ResPath Block

```
for i in range length do
        x0 = x
        x1 = convolution layer with num_filters filters, kernel size of 3, no activation
        sc = convolution layer with num_filters filters, kernel size of 1, no activation
        x = element-wise addition of x1 and sc
        x = ReLU activation function applied to x
        x = Batch Normalization applied to x
end for
return x
```

Algorithm 8 Pseudocode for Res Path

In Algorithm 8 pseudocode for Res Path is presented, which has two convolutional layers: 3×3 convolutional layer and also 1×1 convolutional with no activation function, followed by batch normalization. An elementwise addiction is performed between the two of these convolutional layers. After that, the activation function ReLU is applied in the Res Path.

The encoder of the second network, Multires U-Net, of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is depicted in Figure 4.9.

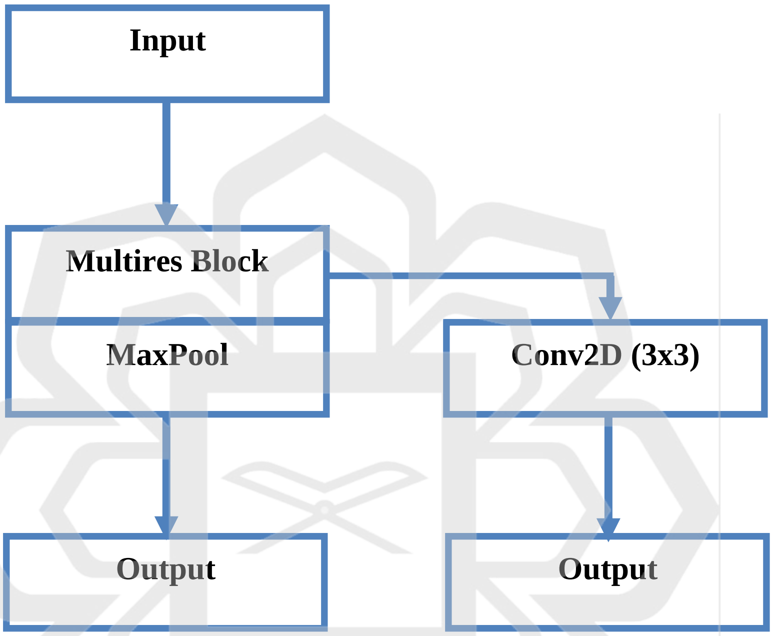


Figure 4.9 Encoder Block in Second Network

The encoder in the second network of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is built from scratch, the input is passed through the Multires block, and the output of the Multires block is then followed by the Maxpooling with a pooling size of 2×2, where the spatial dimension (height and width) of the feature maps is decreased by half. The output from the Multires block also passes through the res path, which serves as a skip connection for the associated decoder block. The pseudocode for the encoder in the second network is illustrated in Algorithm 9.

Algorithm 9: Pseudocode for Encoder in Second Network

Input:

inputs: tensor, input tensor for the encoder

Output:

x: tensor, output tensor of the encoder

skip_connections: list of tensors, skip connections for the decoder

```
1. num_filters = list of integers, the number of filters for each convolutional layer [32, 64,
        128, 256]
2. skip_connections = empty list
3. x = inputs
4. for i in range length of num_filters:
5.        x = call multires_block with input x and number of filters num_filters[i]
6.        s = call res_path with input x, number of filters num_filters[i], and index of
            previous num_filters[i-1]
7.        append s to skip_connections
8.        x = MaxPool2D layer with (2,2) pool size applied to x
9. return x, skip_connections
```

Algorithm 9 Pseudocode for Encoder in Second Network

The pseudocode of the encoder in the second network illustrated in Algorithm 9 uses the input feature map from the previous block (which is the feature map of the elementwise multiplication between the output block and input from the first network). The output of the encoder is the feature map to the ASPP and the skip connection which is used by the decoder in the second network.

The encoder of the second network is connected to the decoder of the second network through an Atrous Spatial Pyramid Pooling bridge (ASPP) explained in section 4.2.3.

4.2.6 Decoder in Second Network (MultiRes U-Net)

In the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, the decoder block in the second network, MultiRes U-Net, performs a 2×2 bi-linear up-sampling on the input feature maps, which increases the dimension of the feature maps of the input. Furthermore, the decoder in MultiRes U-Net of the proposed architecture uses skip connection from both encoders, skip connection 1 from the first network's encoder and skip connection 2 from the second network's encoder, which retains the spatial resolution and improves the quality of the feature maps of the output. The decoder in the second network is shown in Figure 4.10.

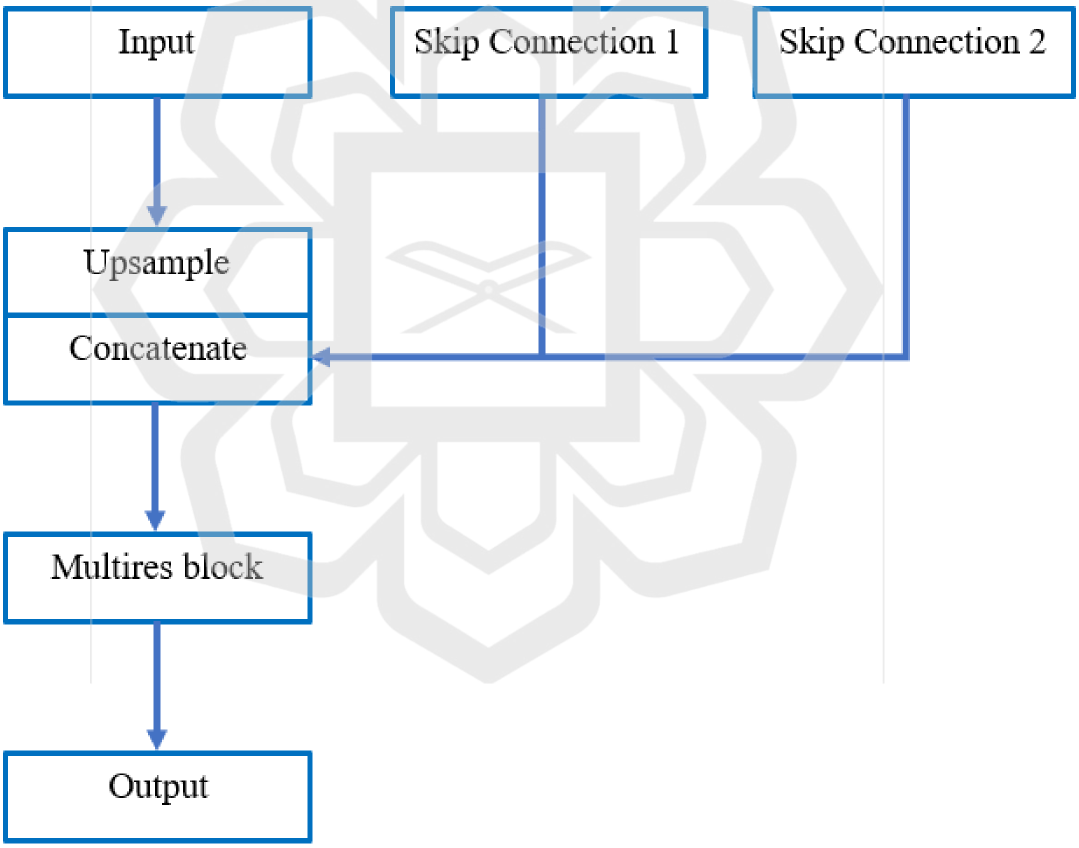


Figure 4.10 Decoder in the Second Network

After the concatenation of the upsampled feature map with the output of the two skip connections, skip connection 1 and skip connection 2, in the decoder of the second network of the proposed architecture, it is then followed by the Multires block (mentioned in section 4.2.5.1). In the decoder of the second network, squeeze and excitation block (discussed in section 4.2.4.1) is also used. Finally, the output block consists of a 1×1 convolution layer, which is then followed by a sigmoid activation function, that is used to generate the mask for the network. Algorithm 10 illustrates the pseudocode for the output block, and Algorithm 11 presents the pseudocode for the decoder.

Algorithm 10: Pseudocode for Output Block

Input:

inputs: tensor, input tensor to the Output Block

Output:

x: tensor, output tensor of the Output Block

1. x = Convolutional layer with 1 filter, 1x1 kernel size and inputs
2. x = Sigmoid activation function of x
3. return x

Algorithm 10 Pseudocode for Output Block

In Algorithm 10 Pseudocode for Output Block is shown, which uses the 1×1 convolution layer and then followed by the sigmoid function.

Algorithm 11: Pseudocode for Decoder in Second Network

Input:

inputs: tensor, input tensor to the decoder

skip_1: list of tensors, skip connections from the first network's Encoder

skip_2: list of tensors, skip connections from the second network's Encoder

Output:

x: tensor, output tensor of the decoder

1. num_filters = list of integers [256, 128, 64, 32]
2. skip_2.reverse()
3. x = inputs
4. for each f in num_filters do the following
5. x = Upsampling2D with bilinear interpolation of size (2,2) applied to x
6. x = Concatenate along the channel axis of x, skip_1[i] and skip_2[i]
7. x = multires_block with f filters applied to x
8. return x

Algorithm 11 Pseudocode for Decoder in Second Network

In Algorithm 11, the pseudocode of the decoder block in the second network is presented, it conducts a 2×2 bilinear upsampling on the input feature maps. And, also it uses skip connection from both encoders, skip connection 1 from the first network's encoder, and skip connection 2 from the second network's encoder.

The proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, shown in Figure 4.1, has two networks: the first network is U-Net which has pre-trained resnet-50 as an encoder, and the second network is MultiRes U-net. Squeeze and Excitation block is employed in the encoder part of the second network and in the decoder part of both networks, U-Net and MultiRes U-net, it minimizes the redundant information and transmits the curial information. Furthermore, Atrous Spatial Pyramid Pooling (ASPP) serves as a bridge between the encoder block and decoder block in both networks, it helps to extract high-quality feature maps that lead to superior performance. The pseudocode of the method for the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, is illustrated in Algorithm 12.

Algorithm 12: Pseudocode of the Method for the Proposed Architecture

```
Input:
    input_shape: tuple, shape of the input tensor
Output:
    model: a Keras model object
inputs = Input layer with input_shape
x, skip_1 = encoder from first network(inputs)
x = ASPP layer with 64 filters
x = decoder1(x, skip_1)
output1 = output_block layer with x as input        (output from first network)
x = inputs element-wise multiplied by output1
x, skip_2 = encoder from second network(x)
x = ASPP layer with 64 filters
x = decoder2(x, skip_1, skip_2)
output2 = output_block layer with x as input        (output from second network)
final output = Concatenate layer with output1 and output2 as inputs
model = Model with inputs and outputs as arguments
return model
```

Algorithm 12 Pseudocode of the Method for the Proposed Architecture

In the first network of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, an input image is fed to U-Net with Resnet_50 as an encoder that produces a predicted mask (output1). After that, the multiplication is applied to the predicted mask (output1) and the input image, which serves as an input for the MultiRes U-net (second network) that generates another predicted mask (output2). Finally, output1 and output2 are concatenated to get the final output that produces a better mask.

4.3 CHAPTER SUMMARY

This chapter illustrates the proposed architecture for the segmentation of medical images. Section 4.2 shows the design of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network. This chapter presents all the blocks used in the proposed architecture and the pseudocode. The proposed architecture consists of two networks, first network, U-Net, which uses pre-trained RESNET_50 as an encoder, is explained in section 4.2.2. ASPP in section 4.2.3 is used as a bridge between the encoder and the decoder in the first network. The decoder for the first network is explained in section 4.2.4. The squeeze and excitation block specified in section 4.2.4.1 is implemented in the decoder. In the proposed architecture’s second network, MultiRes U-Net, the encoder mentioned in section 4.2.5 consists of the Multires block explained in section 4.2.5.1 and the res path described in section 4.2.5.2. ASPP is used between the encoder and decoder of the second network. The Decoder block for the second network is explained in section 4.2.6. Squeeze and Excitation block is also applied in the second network's encoder and decoder block.

CHAPTER FIVE

ANALYSIS OF RESULTS

5.1 INTRODUCTION

In this chapter, the experimental results of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, are explained here to show that the predicted segmentation mask achieved from the proposed architecture is more accurate. The performance of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network (explained in Chapter Four) for the segmentation of medical image datasets mentioned in Chapter Three is evaluated in terms of Intersection Over Union (IOU), Dice Similarity Coefficient (DSC), Precision, and Recall. The performance is then compared to the existing architectures: FCN, U-Net, and Unet_Resnet for the segmentation. This chapter is divided into five sections. 5.2 Section discusses the experimental environment of the dataset used to perform the segmentation. 5.3 Section explains the evaluation metrics. The experimental results of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of different medical image datasets are stated in Section 5.4. Section 5.5 states the comparison results between the proposed architecture and the other existing architecture. Lastly, section 5.6 provides an overview of Chapter 5 of this thesis.

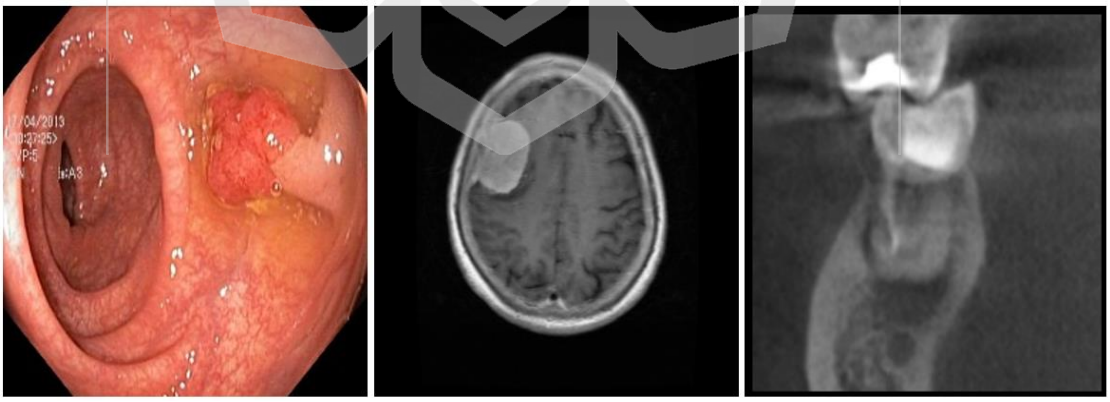
5.2 EXPERIMENTAL ENVIRONMENT

As stated in Chapter Three, three different datasets of medical images: Gastrointestinal Polyp dataset, Brain tumor dataset, and Dental dataset are collected for the segmentation (see section 3.2.2). The datasets contain the original images and the corresponding ground truth. These three datasets of medical images used for the experiment have different modalities and sizes. The input images are resized to reduce the computation time. The modality and body parts of the medical dataset, number of images in the dataset, size of the images, and input size of the image are illustrated in Table 5.1.

Table 5.1 Summary of Medical Image Dataset Used in the Experiment

Body Part	**Modality**	**No. of Images**	**Size**	**Input Size (Resize)**
Gastrointestinal	Colonoscopy	1000	varies	256*256
Brain	CE-MRI	3064	512*512	256*256
Dental	CBCT	409	varies	256*256

All three datasets used in the experiment are divided into a training set, testing set, and validation set. Data augmentation is applied to a training set of all three medical datasets to increase the data for training that improves the performance. Figure 5.1 shows the sample of all the datasets used. All the input datasets of medical images are in JPG format with the “.jpg’ extension.



(c) (b)

Figure 5.1 Original Images (a) Gastrointestinal Colonoscopy (b) Brain MRI and (c) Dental CBCT

5.3 EVALUATION METRICS

The performance of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is evaluated based on the Intersection over Union (IOU), Dice Similarity Coefficient (DSC), Recall, and Precision. All these metrics are evaluated for the segmentation of all three different medical datasets: the Gastrointestinal Polyp dataset, the Dental dataset, and the Brain tumor dataset.

Dice Similarity Coefficient (DSC) is utilized to measure the similarity between the ground truth mask and the predicted mask. Dice Similarity Coefficient (DSC) calculates similar pixels in the predicted mask and the ground truth and multiplies it by 2, and then it is divided by the total number of pixels present in the predicted mask and ground truth. The expression of DSC is described in Equation 5-1

$$DSC = \frac{2 \times |X \cap Y|}{|X| + |Y|}$$

Equation 5-1

Where X denotes the predicted mask, Y denotes the ground truth, and $\cap$ is the intersection. $|X \cap Y|$ represents common pixels between the predicted mask and ground truth, $|X|$ and $|Y|$ represents all the pixels in the predicted mask and ground truth respectively.

Intersection Over Union (IOU), which is also referred to as the Jaccard index, measures the overlap between the predicted mask and the ground truth. Intersection Over Union (IOU) is the region of overlap between the predicted mask and the ground truth and is divided by the region of union between the predicted mask and the ground truth. The expression of Intersection Over Union (IOU) is illustrated in Equation 5-2

$$IOU = \frac{|X \cap Y|}{|X \cup Y|} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|}$$

Equation 5-2

In Equation 5-2, X denotes the predicted mask, Y represents the ground truth, $\cup$ is the union, and $\cap$ is the intersection.

Precision and recall metrics are represented in the following terms:

- True positive (*TP*) represents that both the ground truth and the predicted mask are true.
- True Negative (*TN*) represents that both the ground truth and the predicted mask are false.
- False positive (*FP*) represents that the ground truth is false, but the predicted mask is true.
- False negative (*FN*) represents that the ground truth is true, but the predicted mask is false.

The confusion matrix to interpret the precision and recall is demonstrated in Table 5.2

Table 5.2 Confusion Matrix

	GROUND TRUTH		
		POSITIVE	NEGATIVE
PREDICTED	POSITIVE	TRUE POSITIVE	FALSE POSITIVE
	NEGATIVE	FALSE NEGATIVE	TRUE NEGATIVE

Precision is a metric that measures the positive predictions that are truly positive; that is, out of all the positive predictions, how many percent are true positive predictions? Precision is calculated by using the formula presented in Equation 5-3

$$PRECISION = \frac{TP}{TP + FP} = \frac{TRUE\ POSITIVE}{PREDICTIONS}$$

Equation 5-3

A recall is employed to measure the true positive rate; that is, out of all actual positives, how many percentages are true positive predictions? A recall is calculated by using the formula presented in Equation 5-4

$$RECALL = \frac{TP}{TP + FN} = \frac{TRUE\ POSITIVE}{GROUND\ TRUTH}$$

Equation 5-4

The loss function is employed to measure the loss between the actual and the predicted segmentation mask. The Dice Coefficient Loss is often used for the imbalance dataset, which is frequent in the medical field (Ashkani Chenarlogh et al., 2022). In this research, Dice Coefficient Loss function is used to calculate the loss. The dice coefficient loss function is one minus the estimated value of the dice similarity coefficient (DSC). The dice coefficient loss is calculated by the formula defined in Equation 5-5

$$Dice\ Coefficient\ Loss = 1 - DSC$$

Equation 5-5

In Equation 5-5, the Dice coefficient loss is equal to one minus the Dice similarity coefficient (calculated by Equation 5.5). When the value of the Dice Similarity Coefficient (DSC) increases, the value of the Dice Coefficient loss decreases.

5.4 EXPERIMENTAL RESULTS

All the architectures: FCN, U-Net, U-Net-Resnet50, and A Dual Convolutional Neural Network with U-Net and Multi-Residual Network are implemented using the Python 3.9 software framework with Tensor flow 2.5 and CUDA 11.2 is installed. The experiment is conducted on the system with an Intel core i7-10750H CPU processor, 16

GB RAM, and an NVIDIA GeForce RTX 2060 GPU with 6 GB of RAM. The system is running on Windows 11 OS. An Adam optimizer is utilized, and the value of the learning rate (lr) is set to 1e-4. All the architectures used for the experiment are trained for 50 epochs.

5.4.1 Proposed Architecture Results

In this section, the results of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network are shown. The proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, is conducted on the three medical image datasets: Gastrointestinal Polyp, Brain tumor, and Dental (see Chapter Three, Section 3.2.2.1). All dataset is partitioned into 80% training, 10% testing, and 10% validation mentioned in Chapter Three Section 3.2.2.2. All the evaluation metrics (see section 5.3) range from 0-1.

5.4.1.1 Results on Gastrointestinal Polyp

A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is performed for the segmentation on a gastrointestinal polyp dataset containing 1000 images and 1000 corresponding ground truth. The gastrointestinal polyp dataset is divided into 800 images for training, 100 images for testing, and 100 images for validation. Data augmentation is performed to the training set (800 images) which increased the training set up to 13,600 images. The results of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network on the gastrointestinal polyp dataset are depicted in Table 5.3

Table 5.3 Results on Gastrointestinal Polyp Dataset

	DSC	**IOU**	**PRECISION**	**RECALL**
A Dual Convolutional Neural Network with U-Net and Multi-Residual Network	0.89184	0.83082	0.90826	0.91026

In Table 5.3, the Dice Similarity Coefficient (DSC), Intersection Over Union (IOU), precision, and recall are illustrated. For Gastrointestinal polyp segmentation, the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, achieved DSC of 0.89184, IOU of 0.83082, 0.90826 precision, and 0.91026 recall. Figure 5.2 shows the snippet cropped from the Python 3.9 software output, demonstrating the results of the DSC, IOU, Precision, and Recall metrics for the gastrointestinal polyp dataset:

Figure 5.2 Snippet Cropped from the Python 3.9 Output for Gastrointestinal Polyp Dataset

The proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the segmentation of the dataset of gastrointestinal polyp images, is depicted in Figure 5.3.

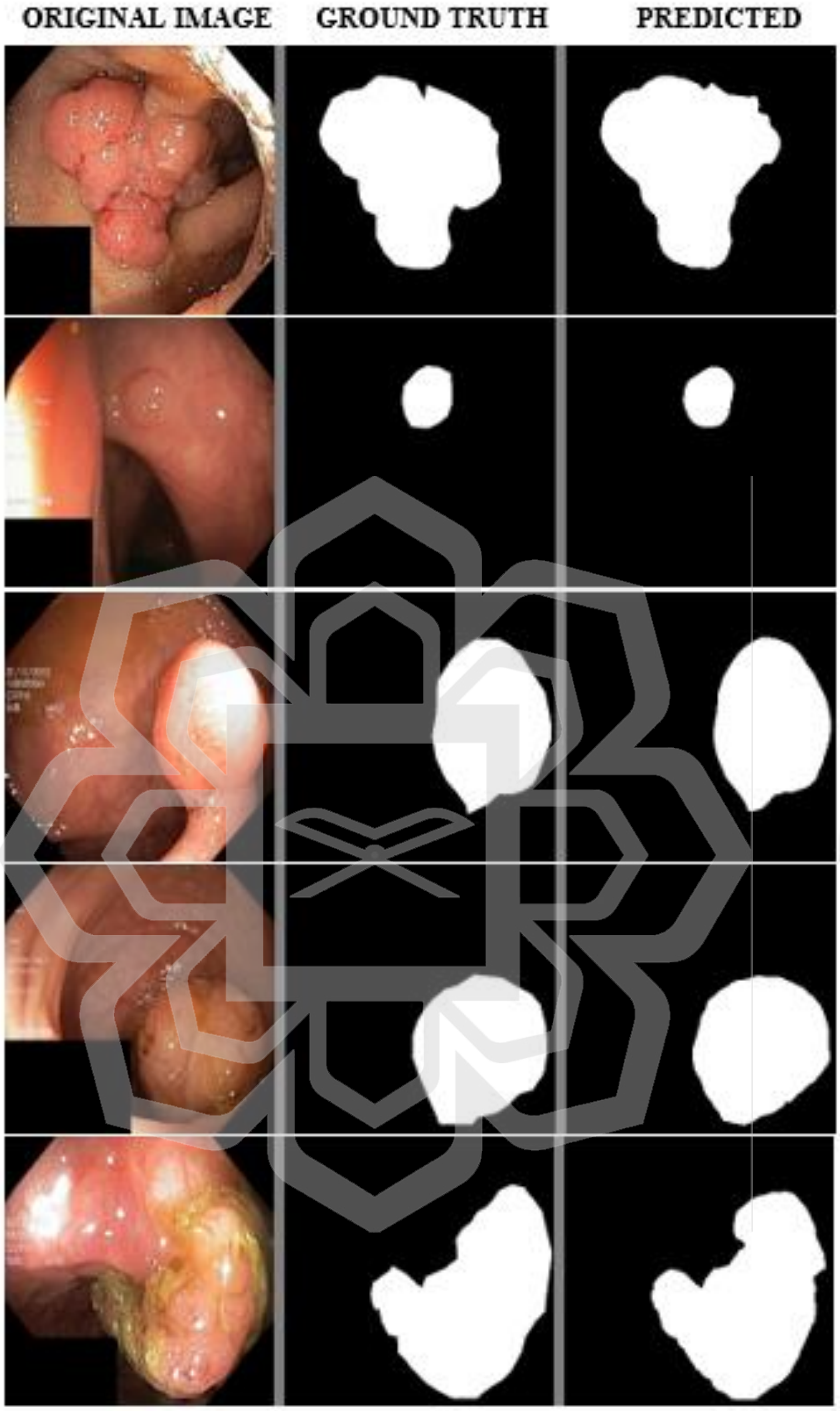


Figure 5.3 Gastrointestinal Polyp Segmentation

Figure 5.3 shows the original image of the gastrointestinal polyp, ground truth, and the predicted mask which is a segmented mask produced by the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network.

5.4.1.2 Results on Brain Tumor

A Dual Convolutional Neural Network with U-Net and Multi-Residual Network performs segmentation on Brain tumor dataset. The brain tumor dataset contains 3,064 images and corresponding 3064 ground truth masks. The brain tumor dataset is divided into 2,452 images for training, 306 images for testing, and 306 images for validation. Data augmentation is implemented to the training set (2,452 images) which increased the training set up to 36,780 images. Table 5.4 provides the results of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network on the brain tumor dataset.

Table 5.4 Results on Brain Tumor Dataset

	DSC	**IOU**	**PRECISION**	**RECALL**
A Dual Convolutional Neural Network with U-Net and Multi-Residual Network	0.84675	0.76817	0.86380	0.85769

For the segmentation of the brain tumor dataset, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network achieves a DSC of 0.84675, IOU of 0.76817, precision of 0.86380, and recall of 0.85769. Figure 5.4 shows the snippet cropped from the Python 3.9 software output, demonstrating the results of the DSC, IOU, Precision, and Recall metrics for the brain tumor dataset:

```
test
X: 306 - Y: 306
 34%|                                                              | 103/306 [00:16<00:30,  6.70it/s]
ics\_classification.py:1248: UndefinedMetricWarning: Precision is ill-defined and being set to 0.0 due to no predicted
  _warn_prf(average, modifier, msg_start, len(result))
100%|█████████████████████████████████████████████████████████████| 306/306 [00:53<00:00,  5.75it/s]
DSC: 0.84675
IOU: 0.76817
Recall: 0.85769
Precision: 0.86380
```

Figure 5.4 Snippet Cropped from the Python 3.9 Output for Brain Tumor Dataset

The segmentation results of brain tumor images performed by the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, are depicted in Figure 5.5.

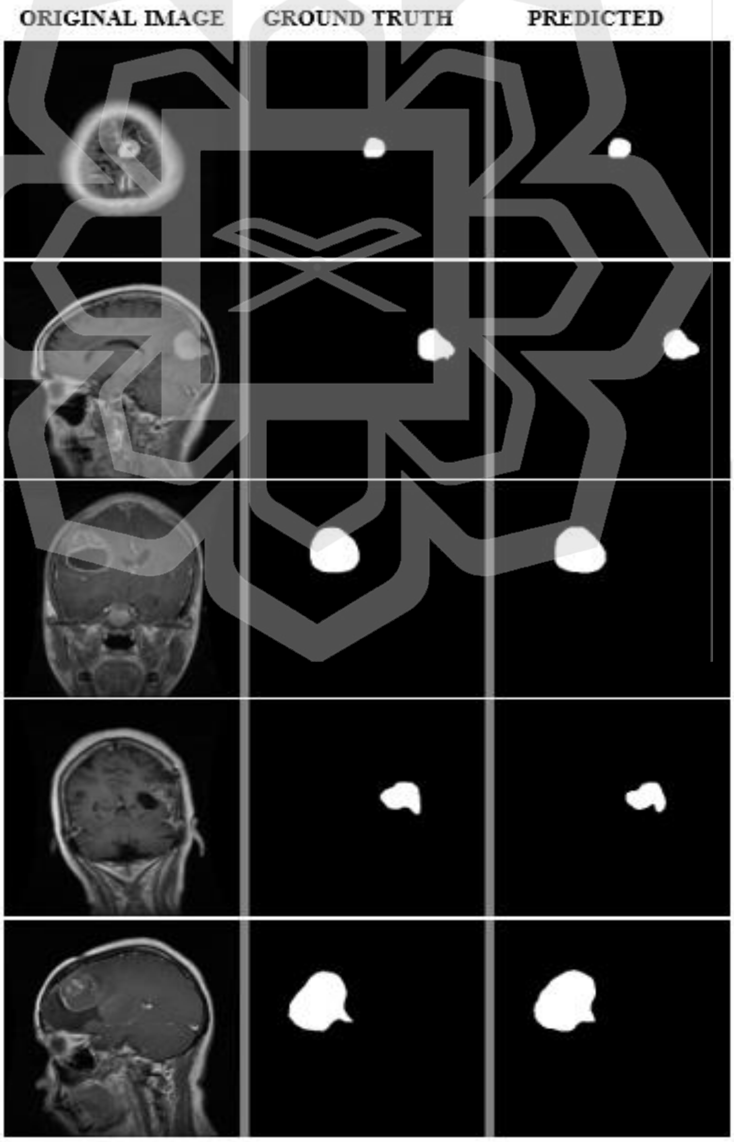


Figure 5.5 Brain Tumor Segmentation

The original image of the brain tumor, the corresponding ground truth mask, and the predicted mask that is achieved by the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the segmentation is shown in Figure 5.5.

5.4.1.3 Results on Dental Dataset

The dental dataset contains 409 images and corresponding 409 ground truth masks. The dental dataset is divided into 329 images for training, 40 images for testing, and 40 images for validation. Data augmentation is performed on the training set (329 images) which increased the training set up to 4,935 images. The results using metrics IOU, DSC, precision, and recall of the dental dataset achieved by A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for segmentation of the dental dataset are provided in Table 5.5.

Table 5.5 Results on Dental Dataset

	DSC	**IOU**	**PRECISION**	**RECALL**
A Dual Convolutional Neural Network with U-Net and Multi-Residual Network	0.66693	0.58323	0.72291	0.65424

A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for segmentation of the dental dataset achieves 0.66693, 0.58323, 0.72291, and 0.65424 DSC, IOU, precision, and recall respectively. Figure 5.6 shows the snippet cropped from the Python 3.9 software output, demonstrating the results of the DSC, IOU, Precision, and Recall metrics for dental dataset:

```
test
X: 40 - Y: 40
100%|████████████████████████████████████████| 40/40 [00:04<00:00,  8.49it/s]
DSC: 0.66693
IOU: 0.58323
Recall: 0.65424
Precision: 0.72291
```

Figure 5.6 Snippet Cropped from the Python 3.9 Output for Dental Dataset

The results of the segmentation of dental dataset images performed by the proposed architecture are shown in Figure 5.7.

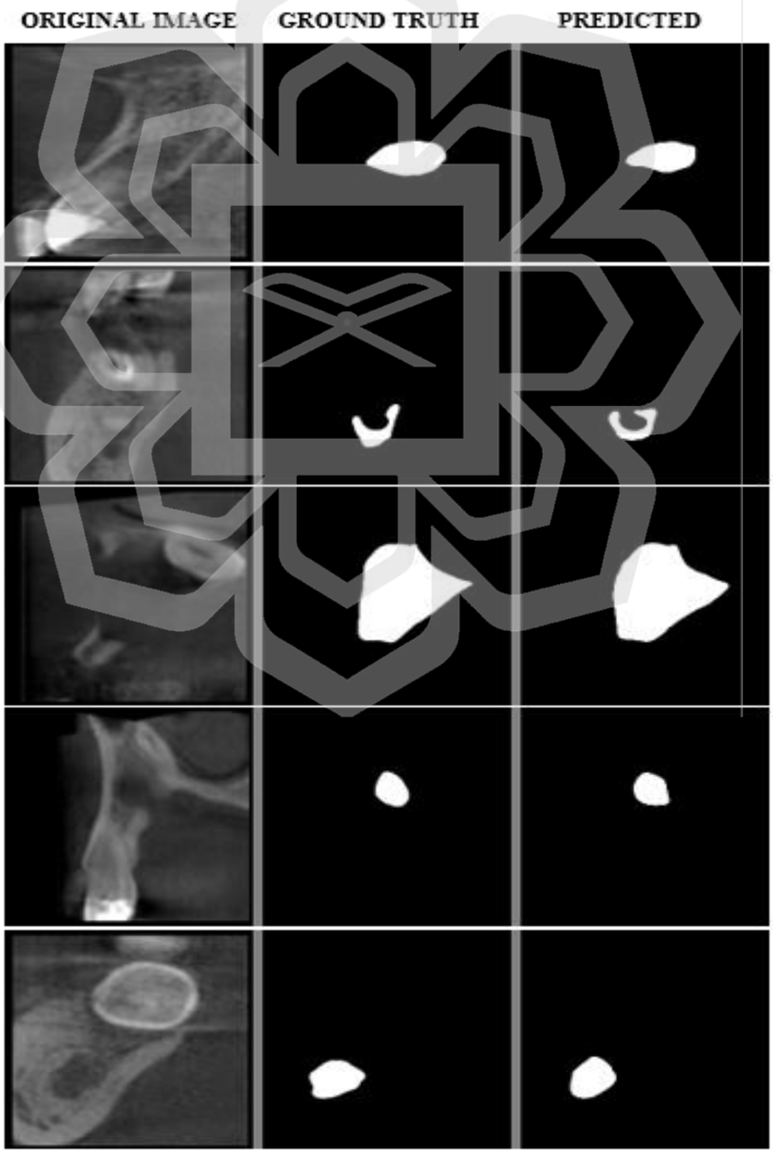


Figure 5.7 Dental Segmentation

Figure 5.7 shows the original image of the dental dataset, the corresponding ground truth, and the predicted mask that is achieved by the proposed architecture for segmentation, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network.

5.5 COMPARISON OF PERFORMANCE RESULTS

This section shows the comparison of performance results of the segmentation of the medical images dataset (see Chapter Three) generated by the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network (see Chapter Four), with the other existing segmentation architectures including FCN, U-Net, and Unet_Resnet. To evaluate the results for the comparison of segmentation performance are measured by Dice Similarity Coefficient (DSC) Equation 5-1 and Intersection Over Union (IOU) Equation 5-2, Precision Equation 5-3 and Recall Equation 5-4 (see section 5.3).

5.5.1 Comparison Analysis of Gastrointestinal Polyp Segmentation

This section shows the segmentation results of the gastrointestinal polyp dataset. Table 5.6 compares the segmentation results of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, FCN, U-Net, and Unet_Resnet in terms of the metrics mentioned in Equation 5-1, Equation 5-2, Equation 5-3, and Equation 5-4 (section 5.3).

Table 5.6 Comparison Results of Test Set of Gastrointestinal Polyp

Architecture	**DSC**	**IOU**	**Precision**	**Recall**
FCN	0.81474	0.71895	0.82747	0.86093
U-Net	0.82841	0.74813	0.82963	0.86816
Unet_Resnet	0.84494	0.77359	0.87904	0.84668
A Dual Convolutional Neural Network with U-Net and Multi-Residual Network	**0.89184**	**0.83082**	**0.90826**	**0.91026**

The experimental results of the test set in Table 5.6, for the segmentation of the gastrointestinal polyp dataset, show that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, achieves the highest Dice Similarity Coefficient (DSC) of 0.89184, Intersection Over Union (IOU) of 0.83082, Precision of 0.90826, and Recall of 0.91026. Furthermore, the result shows that A Dual Convolutional Neural Network with U-Net and Multi-Residual Network outperforms the existing segmentation architecture: FCN, U-Net, and Unet_Resnet in terms of the following metrics DSC, IOU, Precision, and Recall. FCN obtained a DSC of 0.81474, IOU of 0.71895, Precision of 0.82747, and Recall of 0.86093. U-Net obtained a DSC of 0.82841, IOU of 0.74813, Precision of 0.82963, and Recall of 0.86816. Unet_Resnet achieved a DSC of 0.84494, IOU of 0.77359, Precision of 0.87904, and Recall of 0.84668. The visual results of the comparison for the segmentation of gastrointestinal polyps from the testing set are shown in Figure 5.8.

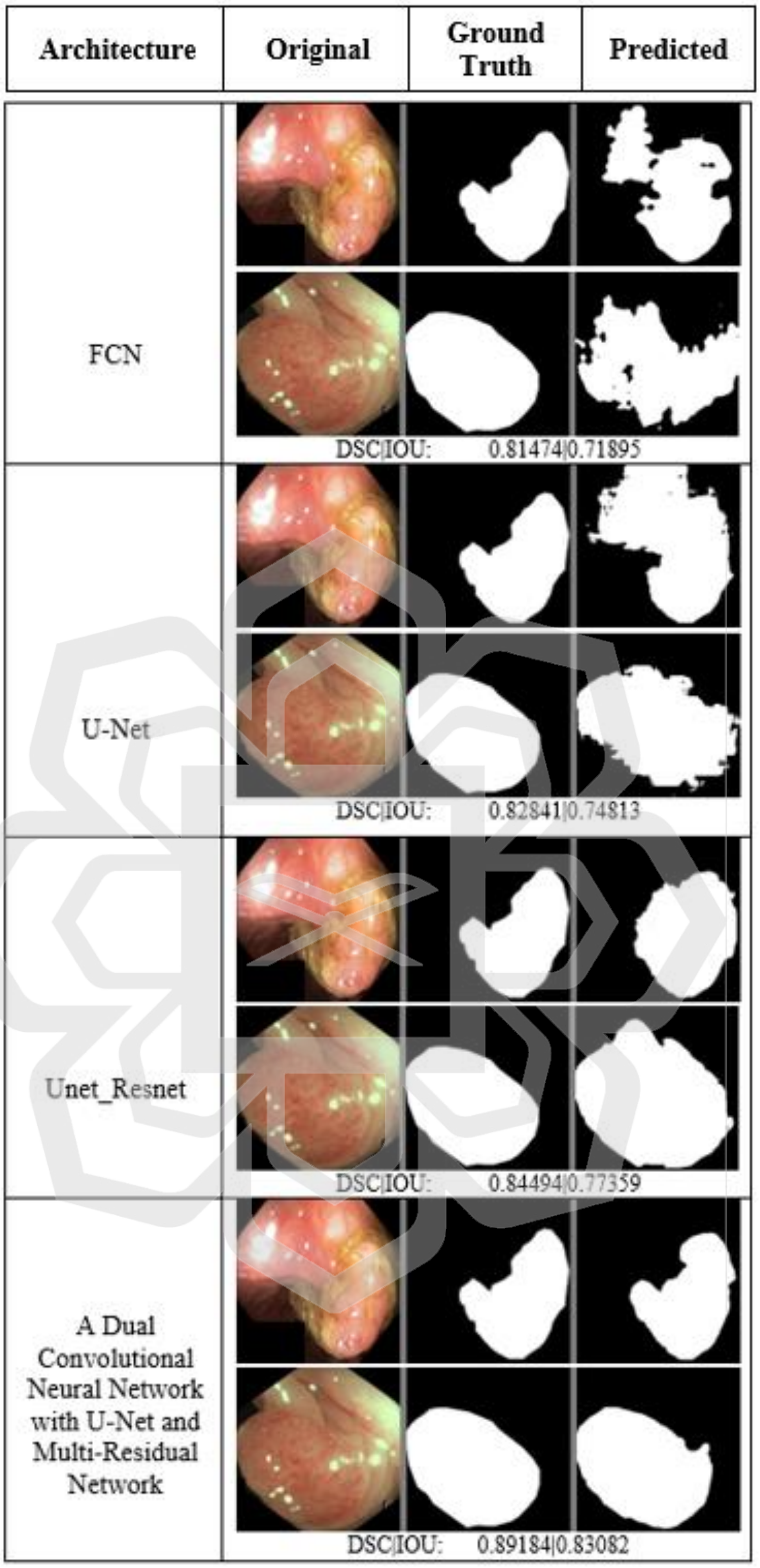


Figure 5.8 Comparison Results of the Segmentation on the Test Images of Gastrointestinal Polyps

In Figure 5.8, it represents the original image of gastrointestinal polyps, the ground truth corresponding to the original image, and the predicted mask which is a segmented result by FCN, U-Net, Unet_Resnet, and A Dual Convolutional Neural Network with U-Net and Multi-Residual Network. For polyp images with clear boundaries, it can be observed that A Dual Convolutional Neural Network with U-Net and Multi-Residual Network performs better segmentation than other architectures. For vague boundaries in polyps, existing architectures FCN, U-Net, and Unet_Resnet, struggle to perform segmentation. In both cases, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network provides better segmentation results. Figure 5.9 illustrates the graphical depiction of the performance of the training set on each epoch of the gastrointestinal polyps for all the architectures used in this research.

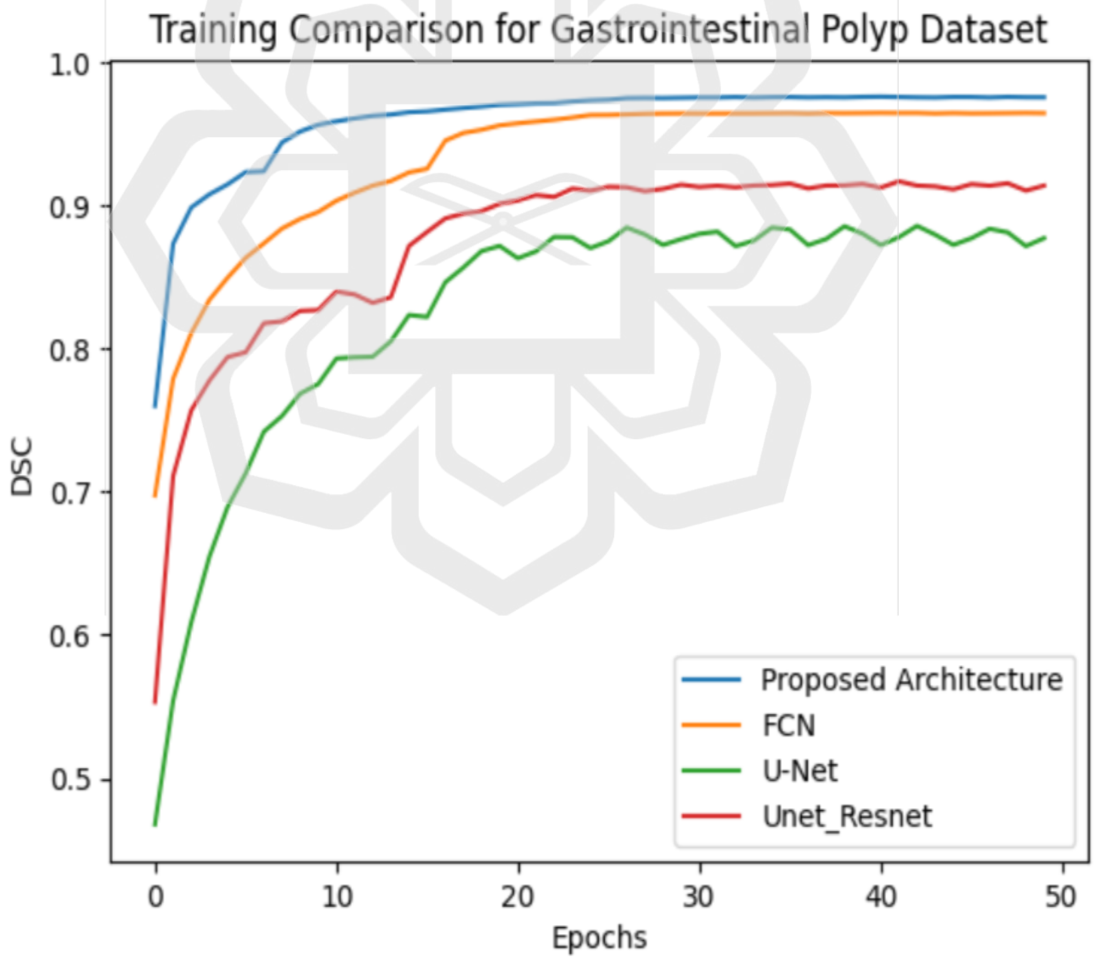


Figure 5.9 Training Performance of the Dice Similarity Coefficient on Gastrointestinal Polyps

Figure 5.9 represents that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, provides better performance of Dice Similarity Coefficient during the training phase when compared to FCN, U-Net, and Unet_Resnet. The dice coefficient loss (Equation 5-5) during the training is graphically demonstrated in Figure 5.10.

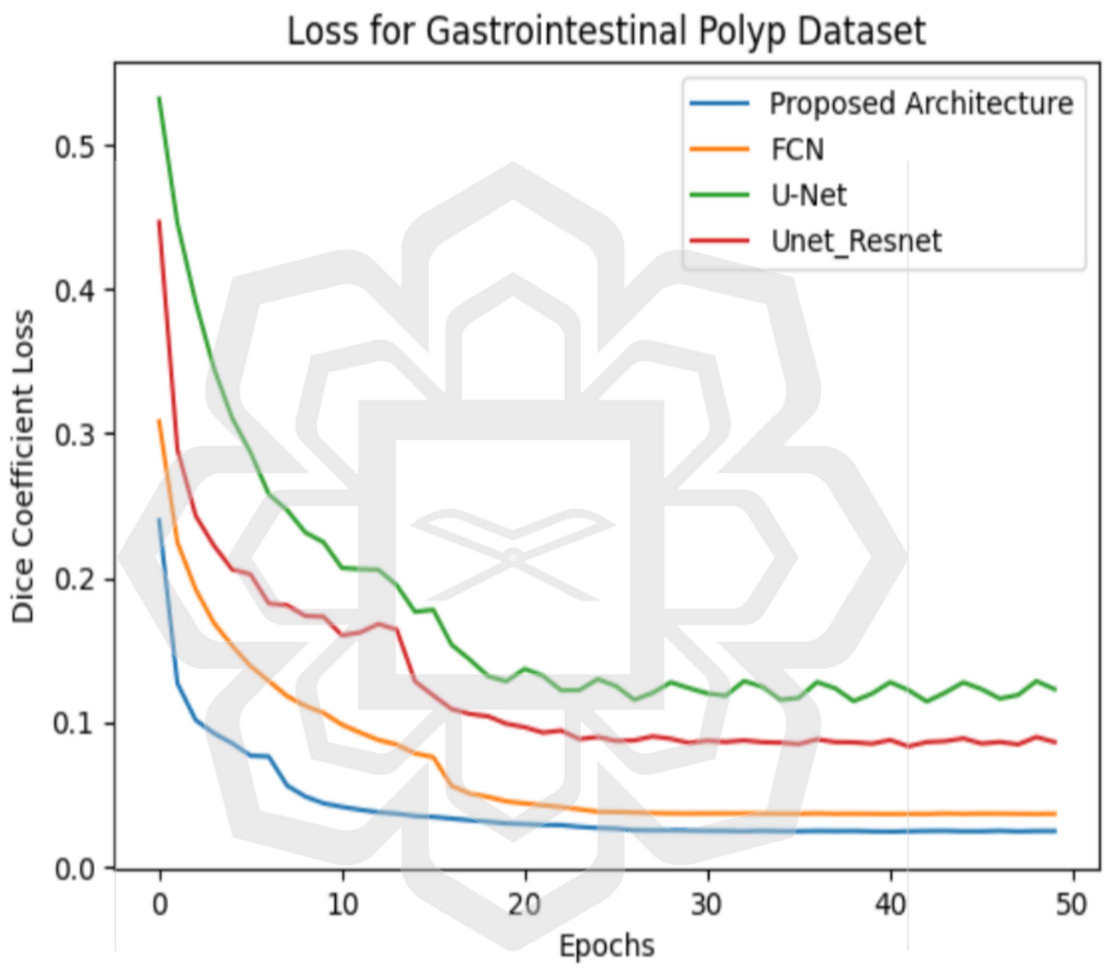


Figure 5.10 Training Dice Coefficient Loss for Gastrointestinal Polyps

In Figure 5.10, It can be observed that dice coefficient loss decreases exponentially with respect to the number of epochs during training for the gastrointestinal polyp dataset.

5.5.2 Comparison Analysis of Brain Tumor Segmentation

The comparison results of the segmentation on the brain tumor dataset by A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, FCN, U-Net, and Unet_Resnet measured by the Dice Similarity Coefficient (DSC) (mentioned in Equation 5-1), Intersection Over Union (IOU) (mentioned in Equation 5-2), Precision (mentioned in Equation 5-3), and Recall (mentioned in Equation 5-4) metrics are shown in Table 5.7.

Table 5.7 Comparison Results of Test Set of Brain Tumor

Architecture	**Dice Similarity Coefficient (DSC)**	**Intersection Over Union (IOU)**	**Precision**	**Recall**
FCN	0.77542	0.68518	0.81098	0.79645
U-Net	0.80697	0.72836	0.82537	0.82586
Unet_Resnet	0.82916	0.75020	0.84265	0.84660
A Dual Convolutional Neural Network with U-Net and Multi-Residual Network	**0.84675**	**0.76817**	**0.86380**	**0.85769**

Table 5.7 illustrates the comparison results of the segmentation on the brain tumor test set. The evaluation results demonstrate that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, achieves a DSC of 0.84675, IOU of 0.76817, Precision of 0.86380, and Recall of 0.85769. All the evaluation metrics, DSC, IOU, Precision, and Recall for the proposed architecture are higher when compared to other architectures, FCN, U-Net, and Unet_Resnet. FCN achieved a DSC of 0.77542, IOU of 0.68518, Precision of 0.81098, and Recall of 0.79645. U-Net achieved a DSC of 0.80697, IOU of 0.72836, Precision of 0.82537, and Recall of 0.82586. Unet_Resnet achieved a DSC of 0.82916, IOU of 0.75020, Precision of 0.84265, and Recall of 0.84660. The visual analysis of the segmentation results is depicted in Figure 5.11.

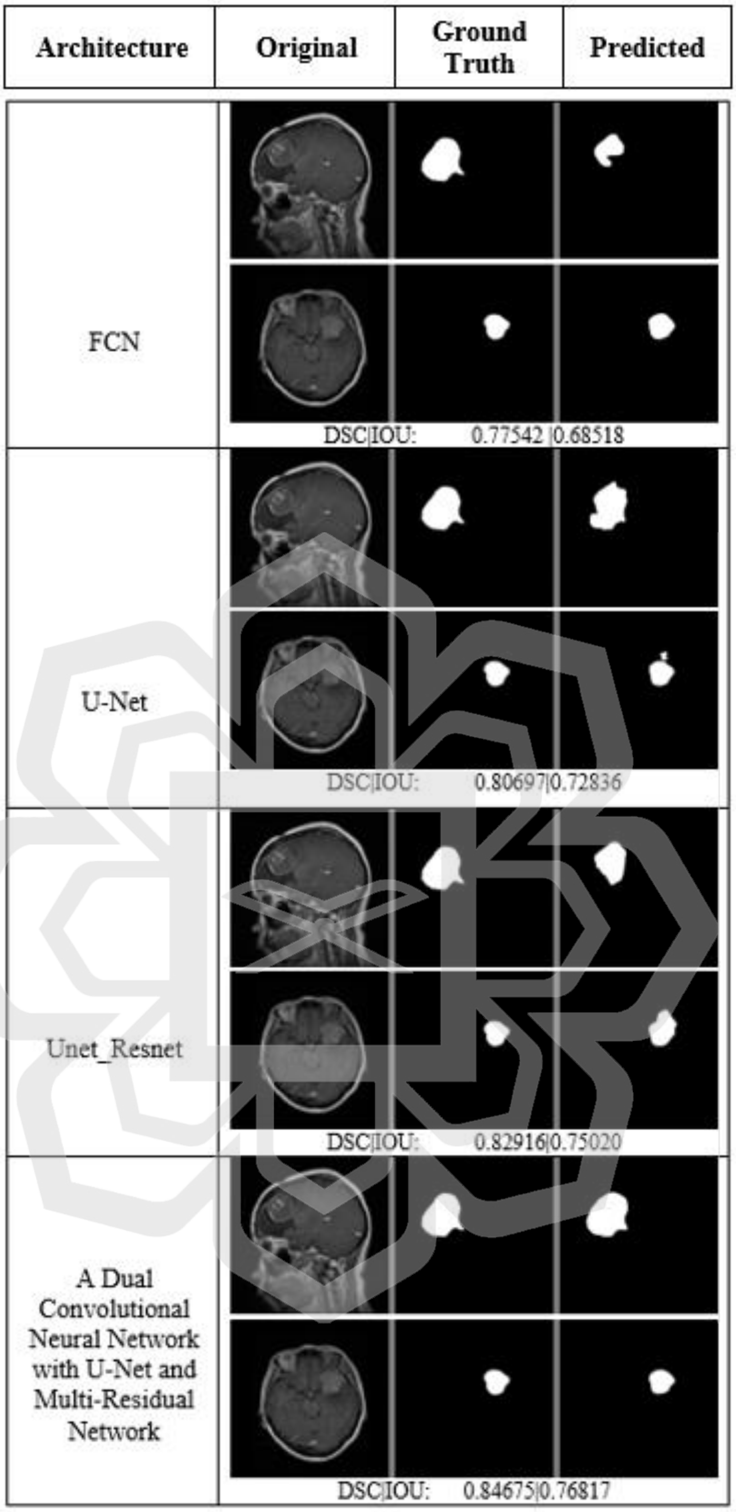


Figure 5.11 Comparison Results of the Segmentation on the Test Images of Brain Tumor Dataset

Figure 5.11 shows that FCN seems to under-segment or less accurately segment the tumor, U-Net produces irregularities along the boundaries and also makes a false prediction, and Unet_Resnet also predicts segmentation less accurately and over-segment the tumor. The proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, performs reasonably well on the brain tumor dataset and generates better segmentation masks as compared to the other architectures. Figure 5.12 illustrates the training performance of the brain tumor dataset for all the architectures used in this research.

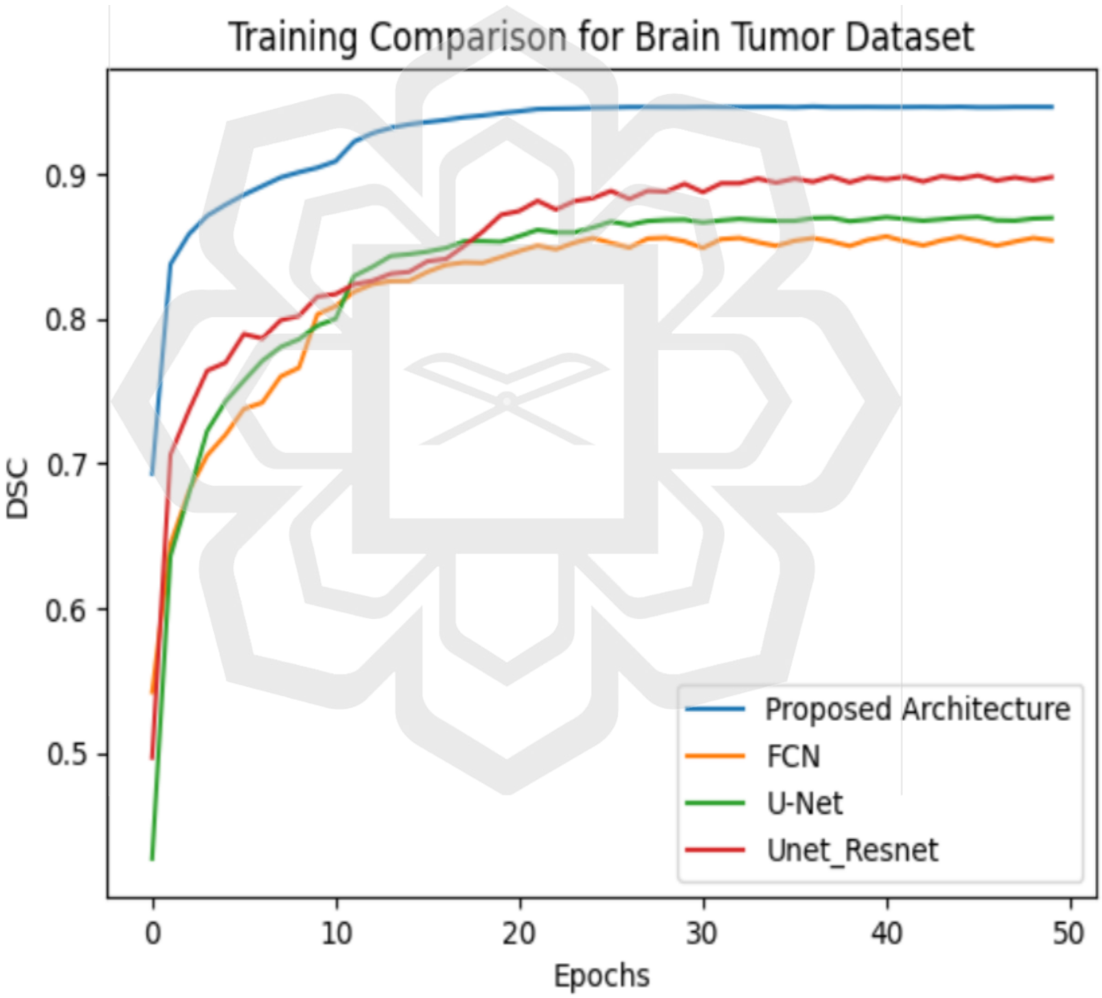


Figure 5.12 Training Performance of the Dice Similarity Coefficient on Brain Tumor Dataset

The training performance on the brain tumor dataset by Dice Similarity Coefficient (DSC) shown in Figure 5.9 represents that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, provides better performance during the training phase as compared to FCN, U-Net, and Unet_Resnet.

The dice coefficient loss (Equation 5-5) during the training for the brain tumor dataset is graphically depicted in Figure 5.13.

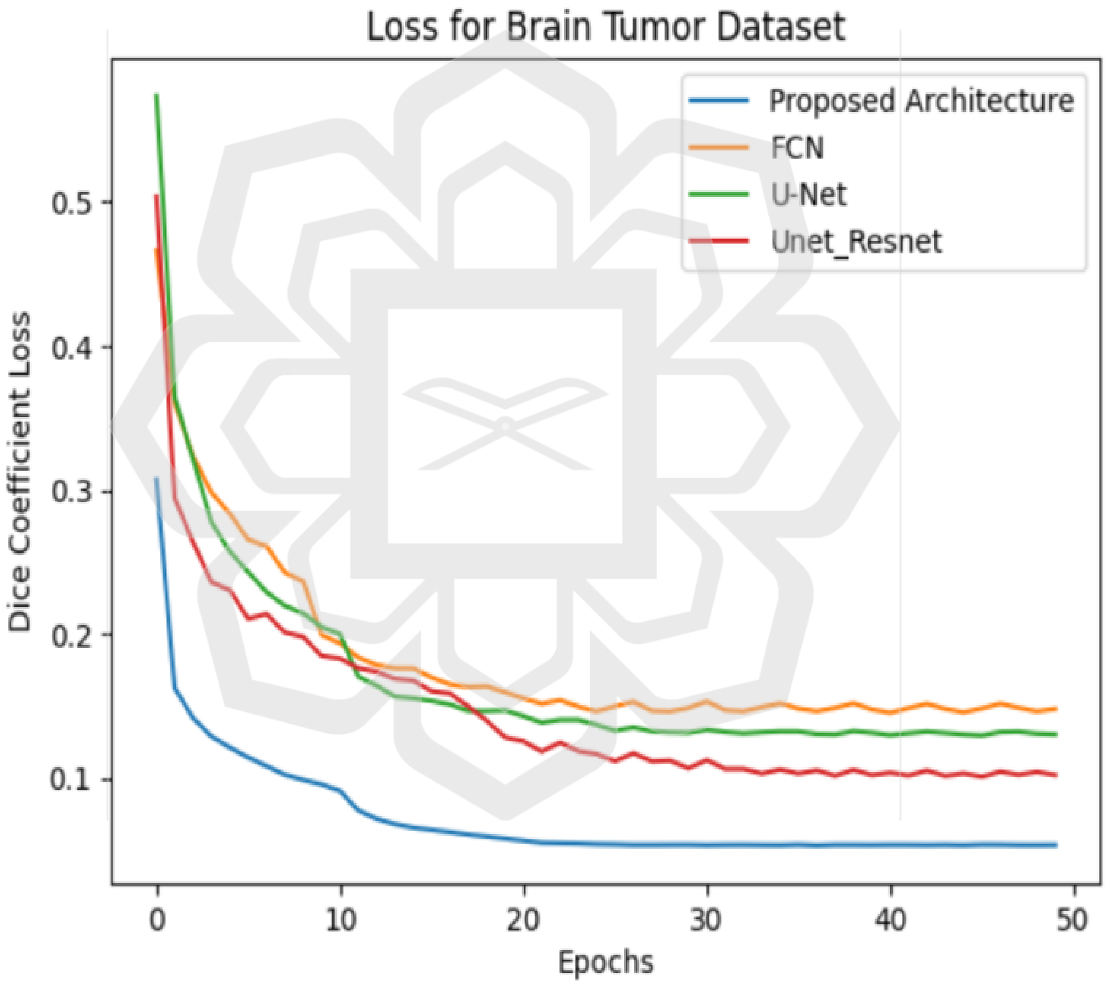


Figure 5.13 Training Dice Coefficient Loss for Brain Tumor

It can be seen in Figure 5.13 that proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, performs well, as the dice coefficient loss decreases more than the other architectures.

5.5.3 Comparison Analysis of Dental Dataset Segmentation

Table 5.8 represents the comparison results of segmentation on the dental dataset by A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, FCN, U-Net, and Unet_Resnet measured by the Dice Similarity Coefficient (DSC) (explained in Equation 5-1), Intersection Over Union (IOU) (explained in Equation 5-2), Precision (explained in Equation 5-3), and Recall (explained in Equation 5-4) metrics.

Table 5.8 Comparison Results of the Test Set of Dental Dataset

Architecture	**Dice Similarity Coefficient (DSC)**	**Intersection Over Union (IOU)**	**Precision**	**Recall**
FCN	0.56370	0.47588	0.62608	0.55854
U-Net	0.61079	0.53468	0.61264	0.63357
Unet_Resnet	0.58044	0.51460	0.61278	0.59357
A Dual Convolutional Neural Network with U-Net and Multi-Residual Network	**0.66693**	**0.58323**	**0.72291**	**0.65424**

The segmentation results of the test set in Table 5.8 on the dental dataset show that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, achieves DSC of 0.66693, IOU of 0.58323, Precision of 0.72291, and Recall of 0.65424. The segmentation result represents that A Dual Convolutional Neural Network with U-Net and Multi-Residual Network outperforms the existing segmentation architecture, FCN, U-Net, and Unet_Resnet, in terms of DSC, IOU, Precision, and Recall. FCN achieved a DSC of 0.56370, IOU of 0.47588, Precision of 0.62608, and Recall of 0.55854. U-Net achieved a DSC of 0.61079, IOU of 0.53468, Precision of 0.61264, and Recall of 0.63357. Unet_Resnet achieved a DSC of 0.58044, IOU of 0.51460, Precision of 0.61278, and Recall of 0.59357. Figure 5.14 demonstrates the visual results of the comparison for the segmentation of brain tumor from the testing set.

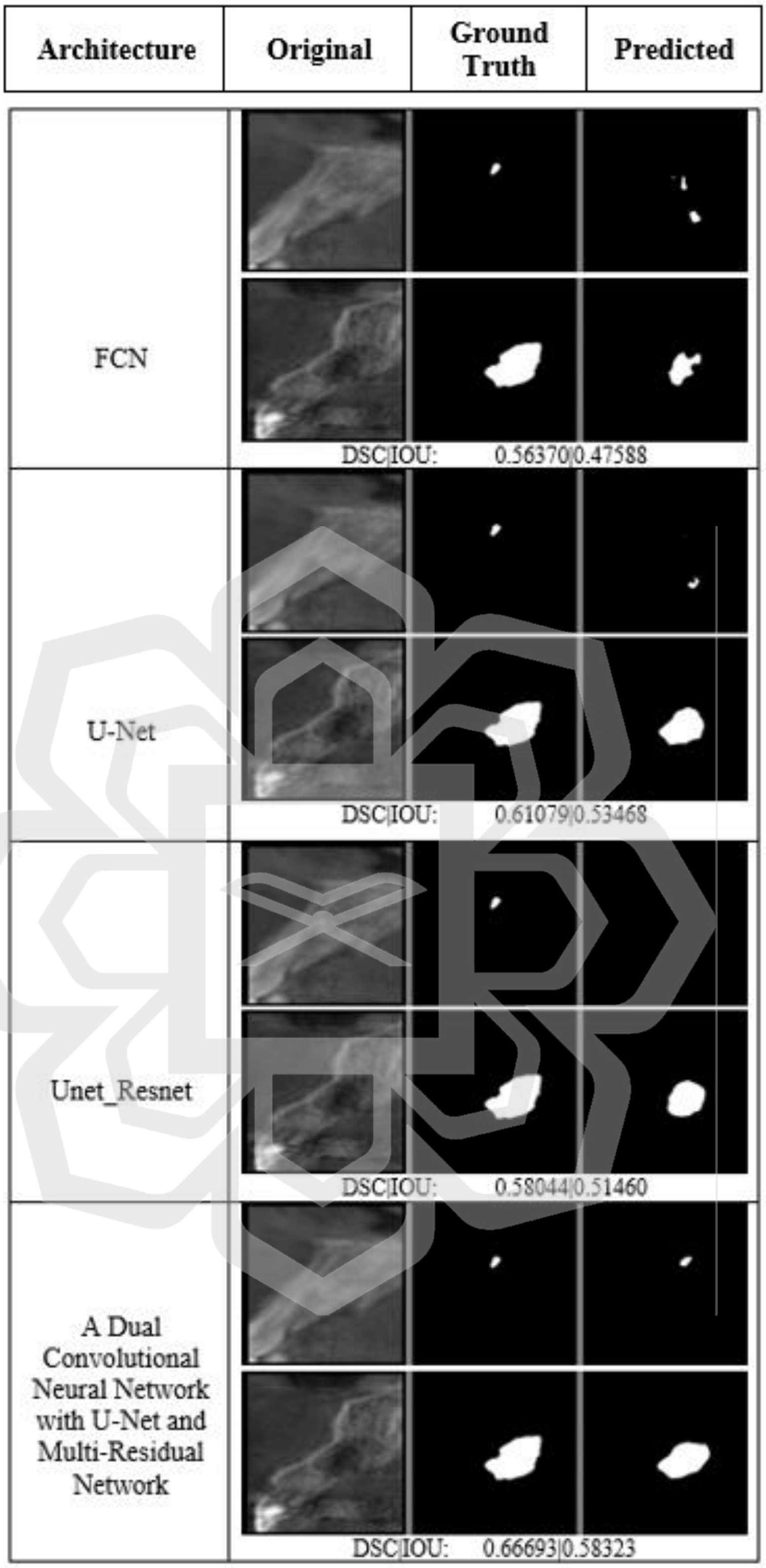


Figure 5.14 Comparison Results of the Segmentation on the Test Images of the Dental Dataset

In Figure 5.14, for more challenging images, FCN, U-Net, and Unet_Resnet seem to be struggling with the segmentation in the dental dataset. For the segmentation of small regions in the dental data set, FCN and U-Net made the wrong predictions, Unet_Resnet missed the prediction completely. On the other hand, the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, has performed considerably better. For irregular boundaries, proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, manages to segment better than the FCN, U-Net, and Unet_Resnet. Figure 5.15 illustrates the graphical representation of the performance of the training dental dataset.

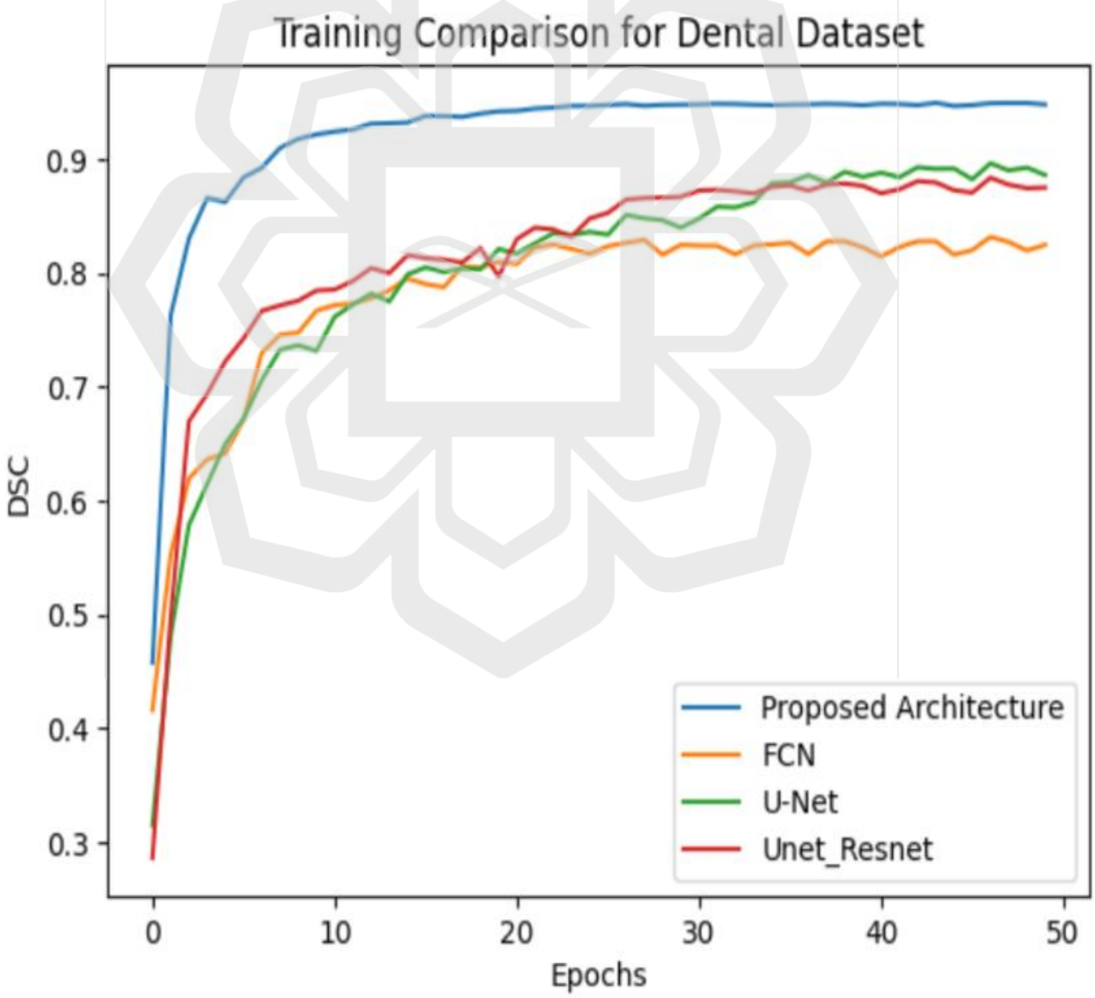


Figure 5.15 Training Performance of the Dice Similarity Coefficient on Dental Dataset

Figure 5.15 illustrates the performance of the training set on each epoch of the dental dataset for all the architectures used in this research. The results show that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, obtains better performance during the training phase as compared to FCN, U-Net, and Unet_Resnet. The dice coefficient loss (Equation 5-5) during the training for the dental dataset is graphically demonstrated in Figure 5.16. The dice coefficient loss decreases gradually for the proposed architecture compared to the FCN, U-Net, and Unet_Resnet.

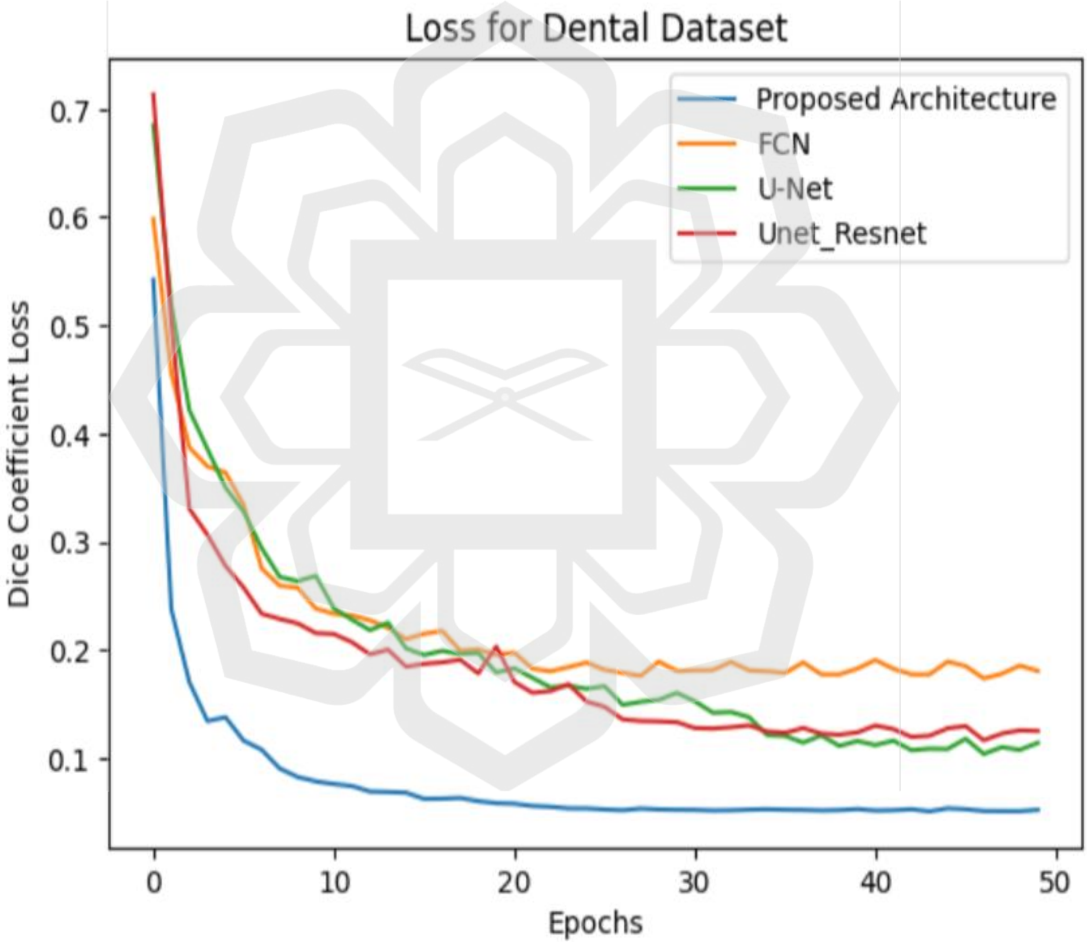


Figure 5.16 Training Dice Coefficient Loss for Dental Dataset

5.6 CHAPTER SUMMARY

This research proposed an architecture that uses a baseline U-Net, multires blocks, res path, ASPP, and squeeze and excitation for the segmentation of medical images. The performance of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation is examined in this chapter. The performance of the segmentation of the medical image dataset using architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, is calculated in terms of metrics: DSC, IOU, precision, and recall (section 5.3). In this research, three medical datasets are used for the segmentation. The performance metrics are then compared to the existing architectures: FCN, U-Net, and Unet_Resnet. It is observed in section 5.5 that the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, outperforms the existing segmentation architecture in terms of DSC, IOU, precision, and recall.

CHAPTER SIX

CONCLUSION AND FUTURE WORK

6.1 INTRODUCTION

This chapter summarizes and concludes the process of the research work for the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, to improve the segmentation of medical images. Furthermore, it presents an overview of the findings discussed in previous chapters. The importance and the contribution of this research are emphasized in this chapter along with the potential for future work.

6.2 CONCLUSION

Segmentation is a significant step in the analysis of medical images. The study's findings demonstrate that the inaccurate results of segmentation for medical images lack the ability to fulfill the actual medical requirements. Segmentation techniques for medical images have been studied for this research. U-Net has become the most popular and effective architecture of convolutional neural networks for the segmentation task of medical images. Therefore, the aim of this chapter is to show the optimized architecture for the segmentation of medical images that utilizes the two networks, U-Net and MultiRes U-Net, and takes advantage of Atrous Spatial Pyramid Pooling (ASPP), and Squeeze and Excitation block. In this research, three different medical image datasets are used for the segmentation. The segmentation performance of the optimized architecture is evaluated by Intersection over Union (IOU), Dice Similarity Coefficient (DSC), Precision, and Recall. The performance of the optimized architecture is compared to the other existing segmentation architectures for medical images: U-Net, FCN, and Unet_Resnet.

This research significantly contributes to the current state of the art in medical image segmentation by addressing a notable gap in the existing segmentation architectures. While many current approaches struggle with the segmentation of medical images, this research introduces an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, that overcomes these challenges.

The experimental results demonstrate a higher performance compared to existing segmentation architecture, showcasing the potential of this approach in real-world medical applications. This research aligns with the current trends in medical image segmentation, as evident from recent studies. In the future, this research will provide a foundation for the segmentation of medical images and has the potential to make a lasting impact on medical diagnostics and treatment planning.

The main aim of the research is to illustrate the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, that improves the performance of the segmentation of medical images. Chapter Two of this research provides a comprehensive literature review of the published segmentation techniques. The study gives an overview of the work related to medical image segmentation using deep neural networks. In addition, work related to segmentation techniques using optimized architectures is explored. It also criticized and investigated the segmentation techniques to identify the advantages and disadvantages of architectures of Convolutional Neural Networks for the segmentation of medical images of previous articles. The major aim of these segmentation techniques is to enhance the performance of the segmentation of medical images. Dice Similarity coefficient, Intersection over Union, Precision, and recall are the metrics used to evaluate the performance of the segmentation architecture. Various segmentation techniques were used by many researchers to optimize the architecture to increase the performance of segmentation for medical images.

Chapter Three highlights the stages of the methodology proposed for the optimized architecture for the segmentation of medical images. Conceptual framework and conceptual design are illustrated that is implemented for developing an optimized segmentation architecture for medical images. The conceptual framework for the optimized segmentation architecture was covered in Section 3.2. The processing of the datasets for the segmentation was explained in Section 3.2.2. The conceptual design for the optimized segmentation architecture for medical images was explained in Section 3.2.3.

Chapter Four of this research presented the experimental design of the proposed architecture for the segmentation of medical images. Section 4.2 demonstrated the design of the proposed segmentation architecture. The proposed segmentation architecture consists of two networks, the first network is U-Net which has pre-trained Resnet_50 as an encoder (see Section 4.2.2), and the second network is Multires U-Net, which consists of MultiRes block and Respath explained in Section 4.2.5.1 and Section 4.2.5.2, respectively. Algorithm 3 Atrous Spatial Pyramid Pooling (see Section 4.2.3), Algorithm 6 Squeeze and Excitation (SE) block (Section 4.2.4.1), Algorithm 7 Multires block (Section 4.2.5.1), and Algorithm 8 Respath (Section 4.2.5.2) were utilized to optimize the proposed architecture to improve the segmentation for the medical images. Algorithm 12 in Section 4.2.6 illustrated the method for the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, that improved the performance of the segmentation for the medical images.

Chapter Five demonstrated the experimental results of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network. The metrics employed to evaluate the performance of the proposed architecture for the segmentation are Dice Similarity Coefficient (DSC), Intersection over Union (IOU), Precision, and Recall presented in Section 5.3. The results of the experiment of the proposed architecture for the segmentation of three different medical datasets using the evaluation metrics are shown in Section 5.4. The evaluation of the performance of the proposed architecture and the comparison to the three benchmark architectures for the segmentation of medical images are presented in Section 5.5. The segmentation results of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the gastrointestinal poly (Section 5.5.1), brain tumor (Section 5.5.2), and dental (Section 5.5.3) are compared to existing architectures, which shows that the proposed architecture outperforms the existing architecture for segmentation of medical images. The experimental results prove the ability of the performance of the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, in terms of evaluation metrics and generate a better segmentation result.

This research is intended to accomplish the objectives stated in Section 1.4 of Chapter 1. The objectives achieved by the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the segmentation of medical images, are discussed in Table 6.1. The questions that were presented at the start of this research for developing an optimized segmentation architecture for medical image segmentation are answered in Table 6.2.

Table 6.1 Research Objectives Achieved

Objective	**Achievements**
O1: To investigate the issues of the U-Net architecture of Convolutional Neural Networks for medical image segmentation.	The previous studies for the segmentation of medical images are investigated in Chapter 2. The issues of the U-Net architecture of the Convolutional Neural Network for the segmentation of medical images are analyzed and discussed in Chapter 2 Section 2.6. The previously optimized Convolutional Neural Network segmentation architectures utilizing U-Net and their advantages and disadvantages for medical image segmentation are discussed in Section 2.7.
O2: To propose an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images.	To improve the performance of the segmentation for medical images, an optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, was proposed. The optimized architecture utilizes the U-Net with Resnet_50 as an encoder, MultiRes U-net having MultiRes block and Respath, Atrous Spatial Pyramid Pooling (ASPP) block, and Squeeze and Excitation (see Chapter 4, Section 4.2). In addition, data augmentation is applied to the training dataset (see Section 3.2.2.2) to improve the performance of the optimized

	segmentation architecture for medical images.
O3: To evaluate and validate the effectiveness of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network architecture on different medical image datasets for segmentation by comparing the results with other existing segmentation architectures.	To evaluate the performance of A Dual Convolutional Neural Network with U-Net and Multi- Residual Network for medical image segmentation, the experiment is employed on three different medical datasets: gastrointestinal polyp dataset, brain tumor dataset, and dental dataset (see Section 3.2.2.1). Intersection over Union (IOU), Dice Similarity Coefficient (DSC), Precision, and Recall metrics (see Section 5.3) are employed to evaluate the performance of the optimized segmentation architecture. The performance of an optimized segmentation architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for three different medical image datasets, is evaluated in Section 5.4.1. The analysis of the performance of the experimental results on the three different medical image datasets is conducted and evaluated results are compared to the other existing segmentation architecture: FCN, U-Net, and Unet_Resnet (see Section 5.5). The results prove that the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi- Residual Network for the

	segmentation of medical images improves the segmentation results and achieves better performance compared to the existing segmentation architectures in terms of Intersection over Union (IOU), Dice Similarity Coefficient (DSC), Precision, and Recall.

Table 6.2 Research Questions with Answers

Research Questions	**Answers from the Research**
Q1: What are the issues in medical image segmentation with the U-Net architecture of the Convolutional Neural Network?	The issues of the U-Net architecture of the Convolutional Neural Network for the segmentation of medical images are covered in Section 2.6. In Chapter 2, the literature review presents that U-Net architecture has an overfitting problem, a vanishing gradient issue, a semantic gap between the encoder and a decoder, and U-Net has a small receptive field. Also, U-Net tends to over-segment and under-segment.
Q2: Why to develop an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for medical image segmentation?	Previous Convolutional Neural Network architectures for the segmentation of medical images show less performance in terms of metrics. In Chapter 2, Section 2.7 discusses the gaps in previous research that uses Convolutional Neural Network techniques for the segmentation of medical images. To enhance the

	efficacy of medical image segmentation, the development of an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network is imperative. The development of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network architecture represents a significant advancement in the field of medical image segmentation, the experimental results are shown in Chapter 5.
Q3: How to develop an architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for medical image segmentation?	The proposed segmentation architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for medical image segmentation is developed by utilizing the two networks: the U-Net with Resnet_50 as an encoder, MultiRes U-net having MultiRes block and Respath. In addition, it takes advantage of the Atrous Spatial Pyramid Pooling (ASPP) block, and Squeeze and Excitation block (see Chapter 4).
Q4: How effective is A Dual Convolutional Neural Network with U-Net and Multi-Residual Network compared to other existing segmentation architectures?	Chapter 5, Section 5.4.1 demonstrates the segmentation performance of the experimental results of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for three different medical image datasets: gastrointestinal polyp dataset, brain tumor dataset, and dental dataset (see Section 3.2.2.1). The performance is evaluated in terms of the Intersection

	over Union (IOU), Dice Similarity Coefficient (DSC), Precision, and Recall (see Section 5.3). The segmentation results with the evaluation metrics achieved and the predicted mask produced by the proposed architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the gastrointestinal polyp dataset, brain tumor dataset, and the dental dataset is presented in Section 5.4.1.1, Section 5.4.1.2, and Section 5.4.1.3 respectively. Section 5.5 analyzes the performance of the experimental results on the three different medical image datasets and evaluated performance results are compared to the other existing segmentation architecture: FCN, U-Net, and Unet_Resnet. The results prove that the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for the segmentation of medical images improves the segmentation results and achieves better performance compared to the existing segmentation architectures in terms of Intersection over Union (IOU), Dice Similarity Coefficient (DSC), Precision, and Recall. The performance results of optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, for the segmentation of the gastrointestinal

	polyp dataset, brain tumor dataset, and dental dataset, and comparison results that prove that optimized segmentation architecture outperforms the existing segmentation architectures is demonstrated in Section 5.5.1, Section 5.5.2, Section 5.5.3 respectively.

6.3 FUTURE WORK

This thesis illustrates how the optimized architecture, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network, is developed for the segmentation of medical images. The optimized architecture for segmentation in this thesis has the ability to provide promising results when compared to the existing segmentation architectures. The optimized segmentation architecture addressed some of the key concerns regarding enhancing the performance of segmentation architecture for medical images, which provides the possibility of further research in these domains. A Dual Convolutional Neural Network with U-Net and Multi-Residual Network for segmentation of medical images uses the two networks which leads it to have more parameters that make it complex and slower for training the data. This complexity of A Dual Convolutional Neural Network with U-Net and Multi-Residual Network demands substantial computational power and GPU resources. The possible future work to improve the segmentation architecture should focus on simplifying the architecture by reducing the parameters while maintaining the ability of the architecture to generate better segmentation results. Another future work possibility is to focus on architecture to enhance the performance of segmentation by training the architecture on small medical image datasets. Another suggestion for future work is to use architecture to perform segmentation on various modalities of different organ and pattern medical image datasets. Another possible future work can be to use the other pre-trained encoders in the segmentation architecture to see the performance of the segmentation. Also, A Dual Convolutional Neural Network with U-Net and Multi-Residual Network in the future can be improved to apply to 3D medical images and may be compared to other segmentation architectures.

REFERENCES

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. E., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11), e00938. https://doi.org/10.1016/j.heliyon.2018.e00938

Abirami, M. S., & Sheela, D. T. (2014). Analysis of Image Segmentation Techniques. *International Journal of Recent Advancement in Engineering & Research*, *3*(3), 4. https://doi.org/10.24128/ijraer.2017.mn89cd

Abubakar, M. M., & Idris, B. (2020). Image Processing Techniques in Computer Vision: An Important Step for Product Visual Inspection. *International Journal Of Science for Global Sustainability*, *November*.

Ahmad, P., Jin, H., Alroobaea, R., Qamar, S., Zheng, R., Alnajjar, F., & Aboudi, F. (2021). MH UNet: A multi-scale hierarchical based architecture for medical image segmentation. *IEEE Access*, *9*, 148384–148408. https://doi.org/10.1109/ACCESS.2021.3122543

Ait Skourt, B., El Hassani, A., & Majda, A. (2018). Lung CT image segmentation using deep neural networks. *Procedia Computer Science*, *127*, 109–113. https://doi.org/10.1016/j.procs.2018.01.104

Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., & Asari, V. K. (2018). *Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation*. *12085 LNAI*, 207–219. https://doi.org/10.48550/arXiv.1802.06955

Altinsoy, E., Yilmaz, C., Wen, J., Wu, L., Yang, J., & Zhu, Y. (2019). Raw G-Band Chromosome Image Segmentation Using U-Net Based Neural Network. *Artificial Intelligence and Soft Computing: 18th International Conference*, *1*(17), 117–126. https://doi.org/10.1007/978-3-030-20915-5

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. In *Journal of Big Data* (Vol. 8, Issue 1). Springer International Publishing. https://doi.org/10.1186/s40537-021-00444-8

Amer, A., Lambrou, T., & Ye, X. Y. (2022). applied sciences MDA-Unet : A Multi-Scale Dilated Attention U-Net for Medical Image Segmentation. *Applied Sciences*, *12*(7).

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical Image Analysis using Convolutional Neural Networks: A Review. *Journal of Medical Systems*, *42*(11). https://doi.org/10.1007/s10916-018-1088-1

Ashkani Chenarlogh, V., Shabanzadeh, A., Ghelich Oghli, M., Sirjani, N., Farzin Moghadam, S., Akhavan, A., Arabi, H., Shiri, I., Shabanzadeh, Z., Sanei Taheri, M., & Kazem Tarzamni, M. (2022). Clinical target segmentation using a novel deep neural network: double attention Res-U-Net. *Scientific Reports*, *12*(1), 1–17. https://doi.org/10.1038/s41598-022-10429-z

Bi, L., Feng, D., & Kim, J. (2018). Dual-Path Adversarial Learning for Fully Convolutional Network (FCN)-Based Medical Image Segmentation. *Visual Computer*, *34*(6–8), 1043–1052. https://doi.org/10.1007/s00371-018-1519-5

Bo, W., Ying, W., & Lijie, C. (2020). Fuzzy clustering recognition algorithm of medical image with multi-resolution feature. *Concurrency and Computation: Practice and Experience*, *32*(1), 1–12. https://doi.org/10.1002/cpe.4886

Cao, W., Li, J., Liu, J., & Zhang, P. (2016). Two improved segmentation algorithms for whole cardiac CT sequence images. *Proceedings - 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2016*, *32*, 346–351. https://doi.org/10.1109/CISP-BMEI.2016.7852734

Cao, Y., Liu, S., Peng, Y., & Li, J. (2020). *DenseUNet : densely connected UNet for electron microscopy image segmentation*. 1–8. https://doi.org/10.1049/iet-ipr.2019.1527

Chakravarty, A., & Sivaswamy, J. (2019). RACE-Net: A Recurrent Neural Network for Biomedical Image Segmentation. *IEEE Journal of Biomedical and Health Informatics*, *23*(3), 1151–1162. https://doi.org/10.1109/JBHI.2018.2852635

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A. L., & Zhou, Y. (2021). *TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation*. 1–13. http://arxiv.org/abs/2102.04306

Chen, L., Papandreou, G., Member, S., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). *DeepLab : Semantic Image Segmentation with Deep Convolutional Nets , Atrous Convolution , and Fully Connected CRFs*. *40*(4), 834–848.

Chen, L., Papandreou, G., Schroff, F., & Adam, H. (2017). *Rethinking Atrous Convolution for Semantic Image Segmentation*. https://doi.org/10.48550/arXiv.1706.05587

Cherguif, H., Riffi, J., Mahraz, A. M., Yahyaouy, A., & Tairi, H. (2019). Brain tumor segmentation based on deep learning's feature representation. *International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS)*, *7*(12). https://doi.org/10.1109/ISACS48493.2019.9068878

Chi, J., Zhang, S., Yu, X., Wu, C., & Jiang, Y. (2020). *A Novel Pulmonary Nodule Detection Model Based on Multi-Step Cascaded Networks*. *1*, 1–26. https://doi.org/10.3390/s20154301

Chunran, Y., & Yuanyuan, W. (2018). Automatic Detection and Segmentation of Lung Nodule on CT Images. *11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2018)*, 1–6. https://doi.org/10.1109/CISP-BMEI.2018.8633101

Du, G., Cao, X., Liang, J., Chen, X., & Zhan, Y. (2020). Medical image segmentation based on U-Net: A review. *Journal of Imaging Science and Technology*, *64*(2), 1–12. https://doi.org/10.2352/J.ImagingSci.Technol.2020.64.2.020508

Feng, Q., Liu, S., Peng, J. xiang, Yan, T., Zhu, H., Zheng, Z. jun, & Feng, H. chao. (2023). Deep learning-based automatic sella turcica segmentation and morphology measurement in X-ray images. *BMC Medical Imaging*, *23*(1), 1–11.

https://doi.org/10.1186/s12880-023-00998-4

Ferl, G. Z., Barck, K. H., Patil, J., Jemaa, S., Malamut, E. J., Lima, A., Long, J. E., Cheng, J. H., Junttila, M. R., & Carano, R. A. D. (2022). Automated segmentation of lungs and lung tumors in mouse micro-CT scans. *IScience*, *25*(12), 105712. https://doi.org/10.1016/j.isci.2022.105712

Ghosh, S., Das, N., Das, I., & Maulik, U. (2019). Understanding deep learning techniques for image segmentation. *ACM Computing Surveys*, *52*(4). https://doi.org/10.1145/3329784

Gudhe, N. R., Behravan, H., Sudah, M., Okuma, H., & Vanninen, R. (2021). Multi - level dilated residual network for biomedical image segmentation. *Scientific Reports*, 1–18. https://doi.org/10.1038/s41598-021-93169-w

Guo, C., Szemenyei, M., Yi, Y., Wang, W., Chen, B., & Fan, C. (2021). SA-UNet : Spatial Attention U-Net for Retinal Vessel Segmentation. *International Conference on Pattern Recognition (ICPR)*, 1236–1242. https://doi.org/10.1109/ICPR48806.2021.9413346

Guo, T., Dong, J., Li, H., & Gao, Y. (2017). Simple convolutional neural network on image classification. *2017 IEEE 2nd International Conference on Big Data Analysis, ICBDA 2017*, 721–724. https://doi.org/10.1109/ICBDA.2017.8078730

Gupta, K. K., Dhanda, N., & Kumar, U. (2018). A comparative study of medical image segmentation techniques for brain tumor detection. *2018 4th International Conference on Computing Communication and Automation, ICCCA 2018*, 1–4. https://doi.org/10.1109/CCAA.2018.8777561

Habijan, M., Leventic, H., Galic, I., & Babin, D. (2019). Whole Heart Segmentation from CT images Using 3D U-Net architecture. *International Conference on Systems, Signals, and Image Processing*, *2019-June*, 121–126. https://doi.org/10.1109/IWSSIP.2019.8787253

Hashem, S. A., & Kamil, M. Y. (2022). Segmentation of Chest X-Ray Images Using U-Net Model. *Mendel*, *28*(2), 49–53. https://doi.org/10.13164/mendel.2022.2.049

Hesamian, M. H., Jia, W., He, X., & Kennedy, P. (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *Journal of Digital Imaging*, *32*(4), 582–596. https://doi.org/10.1007/s10278-019-00227-x

Hoang, H. S., Phuong Pham, C., Franklin, D., Van Walsum, T., & Ha Luu, M. (2019). An Evaluation of CNN-based Liver Segmentation Methods using Multi-types of CT Abdominal Images from Multiple Medical Centers. *Proceedings - 2019 19th International Symposium on Communications and Information Technologies, ISCIT 2019*, 20–25. https://doi.org/10.1109/ISCIT.2019.8905166

Hsiao, Y. T., Chuang, C. L., Jiang, J. A., & Chien, C. C. (2005). A contour based image segmentation algorithm using morphological edge detection. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, *3*, 2962–2967. https://doi.org/10.1109/icsmc.2005.1571600

Hu, Q., Luís, L. F., Holanda, G. B., Alves, S. S. A., Francisco, F. H., Han, T., & Rebouças Filho, P. P. (2020). An effective approach for CT lung segmentation using mask region-based convolutional neural networks. *Artificial Intelligence in Medicine*, *103*. https://doi.org/10.1016/j.artmed.2020.101792

Huang, C., Wu, H., & Lin, Y. (2021). HarDNet-MSEG : A Simple Encoder-Decoder Polyp Segmentation Neural Network that Achieves over 0.9 Mean Dice and 86 FPS. *ArXiv E-Prints*, 1–13. https://doi.org/10.48550/arXiv.2101.07172

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., & Iwamoto, Y. (2020). *UNET 3 + : A FULL-SCALE CONNECTED UNET FOR MEDICAL IMAGE SEGMENTATION*. *iii*, 1055–1059.

Huang, X., Chen, J., Chen, M., Wan, Y., & Chen, L. (2023). FRE-Net : Full-region enhanced network for nuclei segmentation in histopathology images. *Biocybernetics and Biomedical Engineering*, *43*(1), 386–401. https://doi.org/10.1016/j.bbe.2023.02.002

Ibtehaz, N., & Rahman, M. S. (2020). MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*, *121*, 74–87. https://doi.org/10.1016/j.neunet.2019.08.025

Jha, D., Riegler, M. A., Johansen, D., Halvorsen, P., & Johansen, H. D. (2020). DoubleU-Net: A deep convolutional neural network for medical image segmentation. *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, *2020-July*(1), 558–564. https://doi.org/10.1109/CBMS49503.2020.00111

Jin, T., Chen, K., Yamane, S., & Kuroda, Y. (2022). *M-DenseUNet: Multi Dense Encoder Connected UNet for Biomedical Image Segmentation*. 919–921. https://doi.org/10.1109/gcce56475.2022.10014305

Kasar, P. E., Jadhav, Shivajirao, M., & Kansal, V. (2021). MRI Modality-based Brain Tumor Segmentation Using Deep Neural Networks. *Research Square*, *September*, 1–15. https://doi.org/10.21203/rs.3.rs-496162/v2

Khan, A., Sohail, A., Zahoora, U., & Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. In *Artificial Intelligence Review* (Vol. 53, Issue 8). Springer Netherlands. https://doi.org/10.1007/s10462-020-09825-6

Khan, W. (2014). Image Segmentation Techniques: A Survey. *Journal of Image and Graphics*, *July*, 166–170. https://doi.org/10.12720/joig.1.4.166-170

Kim, M., Yun, J., Cho, Y., Shin, K., Jang, R., Bae, H., & Kim, N. (2019). Deep Learning in Medical Imaging. *Neurospine 16(4)*, *16*(4), 49–56. https://doi.org/10.14245/ns.1938396.198

Kiran, I., Raza, B., Ijaz, A., & Khan, M. A. (2022). DenseRes-Unet : Segmentation of overlapped / clustered nuclei from multi organ histopathology images. *Computers in Biology and Medicine*, *143*(October 2021), 105267. https://doi.org/10.1016/j.compbiomed.2022.105267

Lai, J., Zhu, H., & Ling, X. (2019). Segmentation of Brain MR Images by Using Fully Convolutional Network and Gaussian Mixture Model with Spatial Constraints. *Mathematical Problems in Engineering*, *2019*. https://doi.org/10.1155/2019/4625371

Lau, M. M., & Hann Lim, K. (2018). Review of Adaptive Activation Function in Deep

Neural Network. *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 686–690. https://doi.org/10.1109/iecbes.2018.8626714

Li, C., Huang, R., Ding, Z., Gatenby, J. C., Metaxas, D. N., & Gore, J. C. (2011). A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Transactions on Image Processing*, *20*(7), 2007–2016. https://doi.org/10.1109/TIP.2011.2146190

Li, C., Zhang, J., Zhao, X., Kulwa, F., Li, Z., Xu, H., & Li, H. (2021). *MRFU-Net: AMultiple Receptive Field U-Net for Environmental Microorganism Image Segmentation*.

Liang, B., Tang, C., Zhang, W., Xu, M., & Wu, T. (2023). N-Net: an UNet architecture with dual encoder for medical image segmentation. *Signal, Image and Video Processing*. https://doi.org/10.1007/s11760-023-02528-9

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*(December 2016), 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

Liu, X., Song, L., Liu, S., & Zhang, Y. (2021). A review of deep-learning-based medical image segmentation methods. *Sustainability (Switzerland)*, *13*(3), 1–29. https://doi.org/10.3390/su13031224

Lou, A., Guan, S., & Loew, M. (2021). *DC-UNet : Rethinking the U-Net Architecture with Dual Channel Efficient CNN for Medical Images Segmentation*. *February*. https://doi.org/10.1117/12.2582338

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift Fur Medizinische Physik*, *29*(2), 102–127. https://doi.org/10.1016/j.zemedi.2018.11.002

Malhotra, P., Gupta, S., Koundal, D., Zaguia, A., & Enbeyle, W. (2022). Deep Neural Networks for Medical Image Segmentation. *Journal of Healthcare Engineering*, *2022*. https://doi.org/10.1155/2022/9580991

Mehmood, A., Khan, I. R., Dawood, H., & Dawood, H. (2019). Enhancement of CT images for visualization. *ACM SIGGRAPH 2019 Posters, SIGGRAPH 2019*, 3–4. https://doi.org/10.1145/3306214.3338602

Merjulah, R., & Chandra, J. (2017). Segmentation technique for medical image processing: A survey. *Proceedings of the International Conference on Inventive Computing and Informatics, ICICI 2017*, *Icici*, 1055–1061. https://doi.org/10.1109/ICICI.2017.8365301

Mittal, M., Arora, M., Pandey, T., & Goyal, L. M. (2020). *Image Segmentation Using Deep Learning Techniques in Medical Images*. https://doi.org/10.1007/978-981-15-1100-4_3

Mohamed Y. Abdallah, Y., & Alqahtani, T. (2019). Research in Medical Imaging Using Image Processing Techniques. *Medical Imaging - Principles and Applications [Working Title]*. https://doi.org/10.5772/intechopen.84360

Murphy, J. (2016). An Overview of Convolutional Neural Network Architectures for Deep Learning. *Microway Inc*, 1–22.

Muthuselvi, S., & Prabhu, P. (2017). DIGITAL IMAGE PROCESSING TECHNIQUES – A SURVEY. *International Multidisciplinary Research Journal*, *5*(11), 1–7.

Nie, X., Zhou, X., Tong, T., Lin, X., Wang, L., Zheng, H., Li, J., Xue, E., Chen, S., Zheng, M., Chen, C., & Du, M. (2022). N-Net: A novel dense fully convolutional neural network for thyroid nodule segmentation. *Frontiers in Neuroscience*, *16*. https://doi.org/10.3389/fnins.2022.872601

Norouzi, A., Rahim, M. S. M., Altameem, A., Saba, T., Rad, A. E., Rehman, A., & Uddin, M. (2014). Medical image segmentation methods, algorithms, and applications. *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, *31*(3), 199–213. https://doi.org/10.1080/02564602.2014.906861

Pai, A., Teng, Y. C., Blair, J., Kallenberg, M., Dam, E. B., Sommer, S., Igel, C., & Nielsen, M. (2017). Characterization of Errors in Deep Learning-Based Brain MRI Segmentation. In *Deep Learning for Medical Image Analysis* (1st ed.).

Elsevier Inc. https://doi.org/10.1016/B978-0-12-810408-8.00013-4

Pasban, S., Mohamadzadeh, S., Zeraatkar-Moghaddam, J., & Shafiei, A. K. (2020). Infant brain segmentation based on a combination of vgg-16 and u-net deep networks. *IET Image Processing*, *14*(17), 4756–4765. https://doi.org/10.1049/iet-ipr.2020.0469

Pashaei, M., Kamangir, H., Starek, M. J., & Tissot, P. (2020). Review and evaluation of deep learning architectures for efficient land cover mapping with UAS hyper-spatial imagery: A case study over a wetland. *Remote Sensing*, *12*(6). https://doi.org/10.3390/rs12060959

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, *35*(5). https://doi.org/10.1109/TMI.2016.2538465

Qian, L., & Zhou, X. Z. (2022). UNet #: A UNet-like Redesigning Skip Connections for Medical Image Segmentation. *Eprint ArXiv:2205.11759*, arXiv-2205. https://doi.org/10.48550/arXiv.2205.11759

Ren, J., Tong, L., Li, Y., Yuan, L., & Si, Y. (2021). *IMPROVED UNET COMBINING DROPOUT AND ACNET FOR REMOTE SENSING IMAGE CHANGE DETECTION*. 4380–4383.

Rice, L., Wong, E., & Kolter, J. Z. (2020). Overfitting in adversarially robust deep learning. *International Conference on Machine Learning*, 8093–8104.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9351*(Cvd), 12–20. https://doi.org/10.1007/978-3-319-24574-4

S.Kannan, Gurusamy, V., & G.Nalini. (2013). Review on Image Segmentation Techniques. *Pattern Recognition*, *26*(9), 1277–1294.

Saarinen, T., & Petterson, G. (2006). Image Analysis Techniques. *Tracking Environmental Change Using Lake Sediments*, *2*, 23–39.

https://doi.org/10.1007/0-306-47670-3_3

Sathananthavathi, V., & Indumathi, G. (2021). Encoder Enhanced Atrous (EEA) Unet architecture for Retinal Blood vessel segmentation. *Cognitive Systems Research*, *67*, 84–95. https://doi.org/10.1016/j.cogsys.2021.01.003

Seeja, R. D., & Suresh, A. (2019). Melanoma segmentation and classification using deep learning. *International Journal of Innovative Technology and Exploring Engineering*, *8*(12), 2667–2672. https://doi.org/10.35940/ijitee.L2516.1081219

Seo, H., Khuzani, M. B., Vasudevan, V., Huang, C., Ren, H., Xiao, R., Jia, X., & Xing, L. (2020). Machine Learning Techniques for Biomedical Image Segmentation: An Overview of Technical Aspects and Introduction to State-of-Art Applications. *Physiology & Behavior*, *176*(5), 139–148. https://doi.org/10.1002/mp.13649.Machine

Sharma, S., & Aggarwal, A. (2019). Content-based retrieval of biomedical images using orthogonal Fourier-Mellin moments. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, *7*(3), 286–296. https://doi.org/10.1080/21681163.2018.1493619

Shirly, S., & Ramesh, K. (2019). Review on 2D and 3D MRI Image Segmentation Techniques. *Current Medical Imaging Formerly Current Medical Imaging Reviews*, *15*(2), 150–160. https://doi.org/10.2174/1573405613666171123160609

Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, 82031–82057. https://doi.org/10.1109/ACCESS.2021.3086020

Singh, V. K., Abdel-Nasser, M., Akram, F., Rashwan, H. A., Sarker, M. M. K., Pandey, N., Romani, S., & Puig, D. (2020). Breast tumor segmentation in ultrasound images using contextual-information-aware deep adversarial learning framework. *Expert Systems with Applications*, *162*(February), 113870. https://doi.org/10.1016/j.eswa.2020.113870

Song, H. S., Yoon, H. S., Lee, S., Hong, C. K., & Yi, B. J. (2019). Surgical navigation system for transsphenoidal pituitary surgery applying U-net-based automatic segmentation and bendable devices. *Applied Sciences (Switzerland)*, *9*(24). https://doi.org/10.3390/app9245540

Su, R., Zhang, D., Liu, J., & Cheng, C. (2021). MSU-Net: Multi-Scale U-Net for 2D Medical Image Segmentation. *Frontiers in Genetics*, *12*(February), 1–14. https://doi.org/10.3389/fgene.2021.639930

Taheri, S., Ong, S. H., & Chong, V. F. H. (2010). Level-set segmentation of brain tumors using a threshold-based speed function. *Image and Vision Computing*, *28*(1), 26–37. https://doi.org/10.1016/j.imavis.2009.04.005

Tran, S. T., Cheng, C. H., Nguyen, T. T., Le, M. H., & Liu, D. G. (2021). Tmd-unet: Triple-unet with multi-scale input features and dense skip connection for medical image segmentation. *Healthcare (Switzerland)*, *9*(1), 1–19. https://doi.org/10.3390/healthcare9010054

Tripathi, S., Acharya, S., Sharma, R. D., Mittal, S., & Bhattacharya, S. (2017). Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Data. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*(2). https://doi.org/10.1609/aaai.v31i2.19105

Vardhana, M., Arunkumar, N., Lasrado, S., Abdulhay, E., & Ramirez-Gonzalez, G. (2018). Convolutional neural network for bio-medical image segmentation with hardware acceleration. *Cognitive Systems Research*, *50*, 10–14. https://doi.org/10.1016/j.cogsys.2018.03.005

Vianna, P., Farias, R., & de Albuquerque Pereira, W. C. (2021). U-Net and SegNet performances on lesion segmentation of breast ultrasonography images. *Research on Biomedical Engineering*, *37*(2), 171–179. https://doi.org/10.1007/s42600-021-00137-4

Wang, H., Xie, S., Lin, L., Iwamoto, Y., Chen, X. H. Y., & Tong, R. (2022). *MIXED TRANSFORMER U-NET FOR MEDICAL IMAGE SEGMENTATION College of Computer Science and Technology , Zhejiang University , China College of Information Science and Engineering , Ritsumeikan University , Japan Artificial Intelligence Research Center , Yamag*. 2390–2394.

Wang, J. L., Farooq, H., Zhuang, H., & Ibrahim, A. K. (2020). Segmentation of intracranial hemorrhage using semi-Supervised multi-task attention-based U-Net. *Applied Sciences (Switzerland)*, *10*(9), 1–12. https://doi.org/10.3390/app10093297

Wang, P., Cuccolo, N. G., Tyagi, R., Hacihaliloglu, I., & Patel, V. M. (2018). *Automatic Real-Time Cnn-Based Neonatal Brain Ventricles Segmentation Department of Electrical and Computer Engineering , Rutgers University , Piscataway , NJ Department of Surgery , Rutgers Robert Wood Johnson Medical School , New Brunswick , NJ Departmen*. *Isbi*, 716–719.

Wang, S., Hu, S., Cheah, E., Wang, X., Chen, L., Baikpour, M., Ozturk, A., Li, Q., Chou, H., Lehman, C. D., Kumar, V., & Samir, A. (2020). U-Net Using Stacked Dilated Convolutions for Medical Image Segmentation. *ArXiv Preprint ArXiv:2004.03466*.

Wang, S., Zhou, M., Gevaert, O., Tang, Z., Dong, D., Liu, Z., & Tian, J. (2017). A multi-view deep convolutional neural networks for lung nodule segmentation. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 1752–1755. https://doi.org/10.1109/EMBC.2017.8037182

Wang, W., Chen, J., Zhao, J., Chi, Y., Xie, X., Zhang, L., & Hua, X. (2019). Automated Segmentation Of Pulmonary Lobes Using Coordination-Guided Deep Neural Networks. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, *Isbi*, 1353–1357.

Wang, W., Yang, Y., Wang, X., & Wang, W. (2019). Development of convolutional neural network and its application in image classification: a survey. *Optical Engineering*, *58*(04), 1. https://doi.org/10.1117/1.oe.58.4.040901

Wu, J., Poehlman, S., Noseworthy, M. D., & Kamath, M. V. (2008). Texture feature based automated seeded region growing in Abdominal MRI segmentation. *BioMedical Engineering and Informatics: New Development and the Future - Proceedings of the 1st International Conference on BioMedical Engineering and Informatics, BMEI 2008*, *2*, 263–267. https://doi.org/10.1109/BMEI.2008.352

Xu, Y., Hou, S., Wang, X., Li, D., & Lu, L. (2023). *A Medical Image Segmentation Method Based on Improved UNet 3 + Network*.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, *9*(4), 611–629. https://doi.org/10.1007/s13244-018-0639-9

Yang, Y., Feng, C., & Wang, R. (2020). Automatic segmentation model combining U-Net and level set method for medical images. *Expert Systems with Applications*, *153*, 113419. https://doi.org/10.1016/j.eswa.2020.113419

Yang, Y., Yang, P., & Zhang, B. (2020). Automatic segmentation in fetal ultrasound images based on improved U-net. *Journal of Physics: Conference Series*, *1693*(1), 6–12. https://doi.org/10.1088/1742-6596/1693/1/012183

Yousefirizi, F., Jha, A. K., Brosch-Lenz, J., Saboury, B., & Rahmim, A. (2021). Toward High-Throughput Artificial Intelligence-Based Segmentation in Oncological PET Imaging. *PET Clinics*, *16*(4), 577–596. https://doi.org/10.1016/j.cpet.2021.06.001

Zafar, K., Gilani, S. O., Waris, A., Ahmed, A., Jamil, M., Khan, M. N., & Kashif, A. S. (2020). Skin lesion segmentation from dermoscopic images using convolutional neural network. *Sensors (Switzerland)*, *20*(6), 1–14. https://doi.org/10.3390/s20061601

Zhang, J., Li, C., Kosov, S., Grzegorzek, M., Shirahama, K., Jiang, T., Sun, C., Li, Z., & Li, H. (2021). LCU-Net : A novel low-cost U-Net for environmental microorganism image segmentation R. *Pattern Recognition*, *115*, 107885. https://doi.org/10.1016/j.patcog.2021.107885

Zhang, Q., Du, Y., Wei, Z., Liu, H., Yang, X., & Zhao, D. (2021). *Spine Medical Image Segmentation Based on Deep Learning*. *2021*.

Zhang, Z., Wu, C., Coleman, S., & Kerr, D. (2020). DENSE-INception U-net for medical image segmentation. *Computer Methods and Programs in Biomedicine*, *192*, 105395. https://doi.org/10.1016/j.cmpb.2020.105395

Zhao, B., Soraghan, J., Di-Caterina, G., Petropoulakis, L., Grose, D., & Doshi, T. (2018). Automatic 3D segmentation of MRI data for detection of head and neck cancerous lymph nodes. *Signal Processing - Algorithms, Architectures, Arrangements, and Applications Conference Proceedings, SPA, 2018-Septe*, 298–303. https://doi.org/10.23919/SPA.2018.8563420

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture. In *Deep learning in medical image analysis and multimodal learning for clinical decision support: Vol. 11045 LNCS*. Springer International Publishing. https://doi.org/10.1007/978-3-030-00889-5

Zhou, Z., Siddiquee, M. R., Tajbakhsh, N., & Liang, J. (2019). *UNet ++ : Redesigning Skip Connections to Exploit Multiscale Features in*. *39*(6), 1856–1867.

Zhu, W., Huang, Y., Tang, H., Qian, Z., Du, N., Fan, W., & Xie, X. (2018). AnatomyNet: Deep 3D Squeeze-and-excitation U-Nets for fast and fully automated whole-volume anatomical segmentation. *BioRxiv*, 392969. https://doi.org/10.1101/392969

PUBLICATIONS

1) Nisa, Syed Qamrun, and Amelia Ritahani Ismail. "Dual U-Net with Resnet Encoder for Segmentation of Medical Images." *International Journal of Advanced Computer Science and Applications* 13.12 (2022), DOI: 10.14569/IJACSA.2022.0131265.

2) Nisa, Syed Qamrun, and Amelia Ritahani Ismail. "Comparative Performance Analysis of Deep Convolutional Neural Network for Gastrointestinal Polyp Image Segmentation." International Journal of Innovative Science, Engineering & Technology, Vol. 8 Issue 4, April 2021 (IJISET 2021).

3) Nisa, Syed Qamrun, Amelia Ritahani Ismail, M.A.B MD Ali, and Mohammad Shadab Khan. "Medical Image Analysis using Deep Learning: A Review." *2020 IEEE 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*. IEEE, 2020, DOI: 10.1109/ICETAS51660.2020.9484287.