TOMSK POLYTECHNIC UNIVERSITY
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
Направление подготовки 09.04.04 Программная инженерия
Отделение школы (НОЦ) Информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Rectal Cancer Image Segmentation Method Based on U-Net Network (Метод сегментации изображения рака прямой кишки на основе сети U-Net)

УДК 004.62:004.93'1

Студент

Группа	ФИО	Подпись	Дата
8ПМ0И	Ван Юйцянь		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Губин Е. И.	к.ф.-м.н.		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОСГН ШБИП	Меньшикова Е. В.	к.ф.н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ООД ШБИП	Антоневич О. А.	к.б.н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Савельев А.О.	к.т.н.		

Томск – 2022

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ООП

по направлению 09.04.04 «Программная инженерия»

Код компетенции	Наименование компетенции
Универсальные компетенции	
УК(У)-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий
УК(У)-2	Способен управлять проектом на всех этапах его жизненного цикла
УК(У)-3	Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели
УК(У)-4	Способен применять современные коммуникативные технологии, в том числе на иностранном (-ых) языке (-ах), для академического и профессионального взаимодействия
УК(У)-5	Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия
УК(У)-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки
Общепрофессиональные компетенции	
ОПК(У)-1	Способен самостоятельно приобретать, развивать и применять математические, естественно-научные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
ОПК(У)-2	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач
ОПК(У)-3	Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями
ОПК(У)-4	Способен применять на практике новые научные принципы и методы исследований
ОПК(У)-5	Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
ОПК(У)-6	Способен самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности
ОПК(У)-7	Способен применять при решении профессиональных задач методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях
ОПК(У)-8	Способен осуществлять эффективное управление разработкой программных средств и проектов
Профессиональные компетенции	
ПК(У)-1	Способен к созданию вариантов архитектуры программного средства
ПК(У)-2	Способен разрабатывать и администрировать системы управления базам данных

ПК(У)-3	Способен управлять процессами и проектами по созданию (модификации) информационных ресурсов
ПК(У)-4	Способен проектировать и организовывать учебный процесс по образовательным программам с использованием современных образовательных технологий
ПК(У)-5	Способен осуществлять руководство разработкой комплексных проектов на всех стадиях и этапах выполнения работ

Tomsk – 2022

TOMSK POLYTECHNIC UNIVERSITY | ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
Направление подготовки (специальность) 09.04.04 Программная инженерия
Отделение школы (НОЦ) Информационных технологий

УТВЕРЖДАЮ:
Руководитель ООП
_________ ________ Савельев А.О.
(подпись) (дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
8ПМ0И	Ван Юйцянь

Тема работы:

Rectal Cancer Image Segmentation Method Based on U-Net Network (Метод сегментации изображения рака прямой кишки на основе сети U-Net)	
Утверждена приказом директора (дата, номер)	№ 145-46/с от 25.05.2022

Срок сдачи студентом выполненной работы:	15.06.2022

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе *(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).*	Эта работа посвящена методу кластеризации при работе с многомерными наборами данных, в которых в качестве меры расстояния используется общее евклидово расстояние, и оно не может реально отражать взаимосвязь между точками в многомерных ситуациях.

Перечень подлежащих исследованию, проектированию и разработке вопросов *(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).*	1.Понимать цели и результаты проекта. 2.Понимать типы, характеристики и практические вопросы моделей машинного обучения. 3.Сопоставлять. 4.Следовать указаниям по построению модели машинного обучения.

Консультанты по разделам выпускной квалификационной работы
(с указанием разделов)

Раздел	Консультант
Основная часть	Доцент ОИТ ИШИТР, к.ф.-м.н., доцент Губин Е. И.
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Доцент ОСГН ШБИП, к.ф.н., доцент Меньшикова Е. В.
Социальная ответственность	Доцент ООД ШБИП, к.б.н., доцент Антоневич О. А.

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Губин Е. И.	к.ф.-м.н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ0И	Ван Юйцянь		1

TOMSK POLYTECHNIC UNIVERSITY | ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
Направление подготовки (специальность) 09.04.04 Программная инженерия
Уровень образования магистратура
Отделение школы (НОЦ) Информационных технологий
Период выполнения весенний семестр 2021 /2022 учебного года

Форма представления работы:

Магистерская диссертация

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	15.06.2022

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
10.06.2022	Основная часть	60
10.06.2022	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	20
10.06.2022	Социальная ответственность	20

СОСТАВИЛ:
Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Губин Е. И.	к.ф.-м.н.		

СОГЛАСОВАНО:
Руководитель ООП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Савельев А. О.	к.т.н.		

TASK FOR SECTION
«FINANCIAL MANAGEMENT, RESOURCE EFFICIENCY AND RESOURCE SAVING»

Student:

Group	Full name
8PM9I	Wang Yuqian

School	Information Technology and Robotics	**Division**	Information Technology
Degree	Master	**Educational Program**	09.04.04 Software Engineering

Input data to the section «Financial management, resource efficiency and resource saving»:	
1. *Resource cost of scientific and technical research (STR): material and technical, energetic, financial and human*	- Salary costs – 205088,79 rub. - STR budget – 366315,47 rub.
2. *Expenditure rates and expenditure standards for resources*	- Electricity costs – 5,8 rub per 1 kW
3. *Current tax system, tax rates, charges rates, discounting rates and interest rates*	- Labor tax – 27,1 %; - Overhead costs – 30%;
The list of subjects to study, design and develop:	
1. *Assessment of commercial and innovative potential of STR*	- comparative analysis with other researches in this field;
2. *Development of charter for scientific-research project*	- SWOT-analysis;
3. *Scheduling of STR management process: structure and timeline, budget, risk management*	- calculation of working hours for project; - creation of the time schedule of the project; - calculation of scientific and technical research budget;
4. *Resource efficiency*	- integral indicator of resource efficiency for the developed project.
A list of graphic material *(with list of mandatory blueprints):*	
1. *Competitiveness analysis* 2. *SWOT- analysis* 3. *Gantt chart and budget of scientific research* 4. *Assessment of resource, financial and economic efficiency of STR* 5. *Potential risks*	

Date of issue of the task for the section according to the schedule	

Task issued by adviser:

Position	Full name	Scientific degree, rank	Signature	Date
Associate Professor	Menshikova E.V	PhD		

The task was accepted by the student:

Group	Full name	Signature	Date
8PM9I	Wang Yuqian		

TASK FOR CHAPTER
«SOCIAL RESPONSIBILITY»

Student:

Group	Name
8PM9I	Wang Yuqian

School	Information Technology and Robotics	**Division**	Information Technology
Degree	Master	**Educational Program**	09.04.04 Software Engineering

Topic of FQW:

Rectal Cancer Image Segmentation Method Based on U-Net Network	
Initial data for the chapter «social responsibility»:	
1. Characteristics of the researched object (substance, material, device, algorithm, technique, working area)	– Tumor segmentation based on rectal cancer images – Use U-Net neural network to segment medical images, and set loss function and precision to judge segmentation accuracy – This project was completed in the TPU dormitory at Ycova 21/2.
List of questions to be researched, designed and developed:	
1. Legal and organizational issues of occupational safety **consider special (specific to the projected work area) law norms of labor legislation.** indicate the features of the labor legislation in relation to the specific conditions of the project.	– GOST 12.2.032-78 SSBT. Workplace when performing work while sitting. General ergonomic requirements.
2. Occupational safety: 2.1. Analysis of the identified harmful and dangerous factors: the sourse of factor, the impact on human' s body 2.2 Suggest measures to reduce the impact of identified harmful and dangerous factors	– Increased level of noise – Lack or lack of natural light, insufficient illumination – electromagnetic fields – Increased voltage in an electrical circuit, the closure of which can pass through the human body
3. Environmental Safety: Influence on the atmosphere, hydrosphere, lithosphere	– Hydrosphere: toxins, chemicals and heavy metals: lead, cadmium, lithium, alkaline manganese, and mercury – Lithosphere: glass, metal, plastics, lead, barium, and rare earth metals.
4. Emergency Safety: describe the most likely emergency situation	– Fire
Date issue of the task for the chapter	

Consultant:

Post	Name	Academic degree	Date	Signature
Docent Professor	Antonevich O. A	PhD		

Student:

Group	Name	Date	Signature
8PM9I	Wang Yuqian		

Abstract

Final qualifying work 90 pages, 15 figures, 26 tables, 43 sources.

Keywords: medical image segmentation, deep learning, fullly convolutional network, U-Net, colorectal cancer.

The object of the study is data on the colorectal cancer.

The subject of the research the colorectal cancer tumor segmentation model and make a system interface.

The objective of the project is to develop the system from colorectal cancer tumor segmentation model by using Python programming to help doctors perform rectal cancer tumor segmentation, reduce treatment risk, choose better surgical options and accurately determine tumor size and morphology.

In the research project, to learn how to use python programming for network model building and use Python for rectal cancer tumor region segmentation.

As a result of the study, it is show that to improves the accuracy of rectal cancer tumor segmentation can be improved by using the U-Net network model that allows segmentation of medical images with small data sets. It is concluded that the program calculator is made.

Basic design, technological and technical and operational characteristics: the developed method can determine the size and morphology of rectal tumors.

Scope: the developed technique can be used to improve the accuracy of rectal cancer tumor segmentation and the calculator's interface makes the U-Net model easier to use.

Publications:

XIX Международной научно-практической конференции студентов, аспирантов и молодых ученых.

МОЛОДЕЖЬ И СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ МСИТ-2022

Томск, 21-25 марта 2022г.

Ван Юйцянь(Томский политехнический университет) “Rectal Cancer Image Segmentation Method Based on U-Net Network”

Table Content

List of abbreviations

DSA: Digital Subtraction Angiography

CAD: Computer Aided Design

KNN: Korea New Network

NBNN: Naive-Bayes Nearest-Neightbor

S-KNN: Soft Korea New Network

FCN: Fully Convolutional Networks

BP: Back Propagation

CNN: Convolutional Neural Network

DSC: Dice Similarity Coefficient

ROI: Region Of Interest

Introduction

In recent years, with the vigorous development and application of imaging technology, in more and more medical diagnosis, imaging technology is used to improve the efficiency of doctors' diagnosis and reduce the rate of misdiagnosis. In this case, medical institutions need to process more and more medical images, and the segmentation of medical images is very important. In the past clinical diagnosis, the segmentation is usually performed manually by experienced doctors. This method is time-consuming and labor-intensive, the results are unstable, and there are certain differences in the segmentation results. Therefore, it is particularly important to segment medical images accurately, quickly and efficiently.

The paper uses U-Net network for intelligent segmentation of rectal cancer CT images, incorporates techniques such as image enhancement and batch normalization to alleviate the overfitting phenomenon, and determines the optimal initial learning rate and the number of convolutional kernels through several experiments, and achieves the ideal segmentation of rectal cancer tumors using U-Net network: 85.76%. The experiments show that U-Net works well for medical image segmentation on small data sets, and the similarity of segmentation can be accurately measured using Dice coefficients for data sets with extremely skewed positive and negative samples.

In this work, I have done a literature review to describe the steps to create a rectal cancer tumor network segmentation model using Python programming, including:

- Data preparation
- Variable selection
- Train the model
- Improve segmentation accuracy

However, medical images are more complex, and doctors have different experiences and diagnose diseases differently. This image segmentation model will help doctors make decisions about patients' conditions and control risks. By automatically segmenting the tumor area of rectal cancer, it can help doctors to judge the condition more accurately.

1 Project introduction

1.1 Project background and research purpose and significance

In recent years, with the booming development and application of MRI, positron emission tomography, imaging spectroscopy, ultrasound imaging, X-ray, DSA imaging and other image generation techniques, the use of image generation techniques has been successfully used in more and more medical diagnoses. More and more medical diagnoses, image generation technologies are successfully used to improve the efficiency of doctors' diagnoses and reduce the rate of physician misdiagnosis. Image generation technology allows doctors to have a very intuitive understanding of the current status of the patient, the trend of the condition and trends in the patient's condition, which helps doctors to analyze and judge the case later. This can significantly increase the probability of cure, which is of great significance for medical diagnosis and treatment. With the widespread use of image generation technology, medical institutions need to process more and more medical images. The manual processing and analysis of medical images by doctors alone is not only time-consuming and labor-intensive, but also time-consuming because of the difference in doctors' level and physical fatigue. This is not only time consuming, but also due to the difference in the level of doctors and physical fatigue, which can make the processed results unstable or even wrong. This makes medical institutions in urgent need of automated medical image processing technology to help alleviate the burden of medical doctors and provide stable and effective results. This makes it urgent for medical institutions to have automated medical image processing technology to help reduce the burden of medical doctors and to process medical images in a stable and effective way.

Along with the accumulation of medical image images, researchers have been meticulously processing and analyzing the huge amount of medical image images one after another, from which the association between some diseases and their medical image images is automatically learned by algorithms, and certain achievements have been made in the automated medical image image processing[1-3]. With the widespread use of image generation technology in medical institutions, image processing related technologies will become more and more important. Breakthroughs in automated medical image processing technology are crucial for improving the efficiency of medical diagnosis. Because of the above reasons, the research and application of automated medical image processing-related technologies are of critical importance, which makes the full exploitation of the potential value of medical images and the improvement of the efficiency of clinical applications a top priority for academics and medical institutions.

Image segmentation is a crucial task in medical image processing, and many other related tasks in medical image processing require prior image segmentation, making it a very important part of medical image processing. Since the beginning of the twentieth century, along with the flourishing development of computer-related and artificial intelligence technologies, CAD technology has been used to analyze medical images to assist doctors in diagnosis. In the process of diagnosing diseases using CAD technology, the computer is not disturbed by emotions and works for long hours without fatigue, which allows objective and efficient medical diagnosis and significantly avoids many diagnostic errors. Medical image segmentation is the first and foremost issue in medical image processing and is one of the key steps that can raise the processing of images to a higher level and help to analyze and process medical images in greater depth by understanding the content of images from a semantic perspective. Therefore, the research and application of automatic algorithms for medical image segmentation is crucial. It is beneficial for the professional physicians to use the segmentation results for qualitative and even quantitative analysis and processing of the condition, to compare the patient's physical condition before and after the medical treatment in detail, and to adjust the treatment plan in time according to the patient's physical condition[4]. In conclusion, medical image segmentation is an important part of CAD technology, which is the key to medical image processing technology.

Medical image segmentation generally refers to extraction in some target areas in the entire image through some way, such as nucleus, organs or tissues. The results of the division of medical images usually have no intersection in each area, and each area that is divided has a certain similarities within it[5]. Medical image segmentation is sometimes the first step in the diagnosis and treatment of certain diseases, for example, for early detection of diseases associated with the liver, estimation of the volume of the human liver and creation of a 3D model of the liver, it is necessary to first perform an accurate segmentation of the liver in the image[6]. In the previous clinical diagnosis, the general image segmentation is a senior doctor with rich experience and professional knowledge to artificially divide the artificial division. The error, because doctors may be disturbed by factors such as emotions or fatigue, and each doctor's judgment on the results of the segmentation will also have some differences[7]. In order to effectively auxiliary doctors and improve the effect of medical image segmentation, it has extremely critical importance to the research and application of automated medical image and picture processing technology, which makes the potential value of medical image pictures and improves the efficiency of clinical application. At present, academic and medical institutions are emphasized and explored.

Compared with the images of natural scenes, such as houses, roads, trees, etc., medical images have great special features. While images of natural scenes are usually separated from each other clearly and the outlines of the target objects are clear, medical images, due to their own complexity, have blurred separation between their components and the boundaries between them are not clear enough. This phenomenon is partly due to the variability and complexity of the human body itself, but also due to the objective limitations of medical imaging equipment and immaturity of technology. Therefore, the segmentation of medical images is more challenging than the semantic segmentation of images of natural scenes. When segmenting medical images, it is important to fully consider the special characteristics of medical images compared with natural scenes, so that a better segmentation result can be obtained. However, many past studies did not carefully analyze the differences and special features between the two, and applied the segmentation algorithm for natural scene images to medical images directly without any improvement of the algorithm, resulting in mediocre segmentation results. In conclusion, segmentation for medical images is very challenging and has attracted an increasing number of researchers from home and abroad to solve this problem.

1.2 Current status of domestic research and analysis

It is because of the great social value and difficulties of segmentation of medical images that more and more researchers at home and abroad have been studying and exploring this problem. Researchers abroad have started to study medical image processing technology for a long time, but because of the immaturity of image generation technology at that time, and because of the great complexity and diversity of medical images, even experts may make mistakes in analysis. Along with the booming development and wide application of related technologies, a huge amount of medical images to be processed and analyzed manually or by algorithms have gradually accumulated, which makes medical institutions increasingly hope that medical image related technologies, represented by medical image segmentation, can be applied to automated medical image processing as soon as possible. In recent years, the research on medical image processing technology represented by medical image segmentation has been gradually increasing in China, and more and more related research projects are carried out in various universities and research institutes.

Before deep learning received wide attention from academia and industry, traditional machine learning algorithms based on manual extraction of features in medical image processing, such as medical image segmentation, have made great breakthroughs and developments, and many researchers at home and abroad have proposed numerous traditional algorithms to solve the problem of medical image segmentation. However, these traditional algorithms have obvious limitations, and there are still great difficulties and challenges in applying these traditional algorithms for accurate segmentation of medical images, and newer and more efficient algorithms need to be explored to solve them. The processing of medical images is more difficult compared to images of natural scenes, with diverse variations, chaotic distribution of parts in the image, inconspicuous contrast, lack of obvious edges between parts, and large differences in physical features such as the shape of the target segmentation object[8-9].

Along with the gradual and intensive research on medical image segmentation by domestic and foreign researchers, a number of results based on traditional algorithms have been achieved in this research direction at present. These include algorithms that use different regions of the image for segmentation[10], algorithms that use a predetermined threshold for segmentation[11], algorithms that use boundary information in the image for segmentation[12], and so on. The studied segmentation algorithms can be broadly classified into three categories. The first is the algorithms that use the discontinuity of boundary information for segmentation, such as fitting surfaces, differential operator parallelism, and deformation models; the second is the algorithms that use the similarity of different regions of the image for segmentation, such as region growing algorithms, thresholding methods, classification and clustering algorithms, and statistically related algorithms; and finally, the algorithms that combine the above two methods[13-15].

The methods of medical image segmentation have evolved roughly from manual segmentation in the beginning to semi-automatic segmentation to the recent fully automatic segmentation[16]. In the beginning, it mainly relied on professional doctors to draw the edges of the target segmentation manually, which produced better results, but it relied heavily on the experience and expertise of doctors, and it was less stable, difficult to reproduce the same results, and time-consuming, and it seemed impossible for doctors to segment manually in the face of the accumulated huge amount of medical images to be processed. With the booming development and wide application of computer-related and artificial intelligence technologies, the process of medical image segmentation can be combined with some advanced methods to achieve semi-automatic segmentation of medical images, which can significantly improve the efficiency of image segmentation and make the segmentation results less subjective to professional doctors. However, this method also requires a certain level of experience and expertise from the operator. For example, researcher Worth used his experience and expertise in anatomy to semi-automatically segment MR images of the human brain, and this method achieved relatively good results[17]. Along with the flourishing and widespread use of traditional machine learning techniques, researchers have devised algorithms that can automate image segmentation[18].

The algorithm is completely automatic and does not require the researcher to perform the related operations, but it still has some problems. In using the properties of K-nearest neighbor algorithm and neighborhood approximation, Grabowski, Bieniecki et al. proposed an algorithm to segment images by using neighborhood templates to speed up the finding[19], thus solving the shortcomings of KNN finding inefficiency to some extent, and its final segmentation effect is better than using K-nearest neighbor algorithm alone.Boiman, Shechtman Coupé, Manjón et al. proposed the Naive Bayes Nearest-Neighbor (NBNN) algorithm based on plain Bayes[20], and their experiments on several datasets proved its superiority over the previous algorithms.Coupé, Manjón et al. sliced medical 3D images into overlapping block pixels, then used a series of methods to reduce the number of block pixels, and finally used KNN to search for the most similar block pixels and complete the segmentation by voting on the labels of the pixel points[21]. This method reduces the number of KNN retrievals although it is a segmentation of grayscale images, while using a small number of CT image samples to create a large number of block pixel image samples, which enhances the classification ability of KNN, and the method using block pixels does not require any alignment, and finally, it is analyzed that the use of smaller block pixels is prone to the problem of segmentation holes.Rousseau, Habas used block pixels and KNN algorithm to segment medical images of the human brainwith good results.Konukoglu, Glocker et al[22]. used random forest to approximate the neighborhood structure of KNN[23], which exponentially accelerated the KNN retrieval.WazarKar, Keshavamurthy et al. proposed an S-KNN (Soft KNN) model[24], which has some similarity with the image block-based segmentation method.

In recent years, with the rapid development of the science and technology and the Internet industry, artificial intelligence technology has increasingly applied in all aspects of scientific research, engineering technology and living services, such as recommended systems, driverless, intelligent Q & A, computing advertising, computing advertising , Quantitative investment, voice recognition, face recognition, etc. Among them, the most impressive is deep learning technology.

Deep learning technology has been booming in recent years, and it is one of the newest and most promising subfields of machine learning technology. Deep learning technology can be thought of as deepening and transforming the traditional artificial neural networks in machine learning to process unstructured data such as text and images, which are difficult to extract features manually. The original intention of deep learning is to imitate the mechanism of human brain cognition and learning, and it uses a deeper and more complex network structure than traditional machine learning to extract high level abstract features from raw data. Deep learning is developed from the artificial neural network in traditional machine learning, and its core feature is to use multi-layer artificial neural network to extract high level abstract features. The main advantage of deep learning over traditional machine learning is that it can automatically extract high-level features from raw data instead of researchers manually designing features that may be useful based on their expertise and deep understanding of the problem. Because of its good nonlinear fitting ability and significant breakthroughs in image classification, recommender systems, target detection and tracking, computational advertising, image segmentation, quantitative investment, speech recognition, etc., deep learning techniques have become the focus of research in various universities, research institutes, and companies nowadays. Deep learning can learn high level abstract features from low level features such as pixels of images, which in turn can be applied to many fields where traditional AI techniques are not competent. Since deep learning has gained widespread attention and application in academia and industry, convolutional neural networks have gradually made unexpected breakthroughs in advanced computer vision research such as image classification, target detection, and target tracking. These exciting results have motivated researchers to continuously try to apply convolutional neural networks to more computer vision tasks, one of which is the semantic segmentation of images for pixel-by-pixel classification. The advantage of convolutional neural networks in deep learning over traditional rule-based or machine learning segmentation methods is that they can automatically derive useful high-level features from the original features for different tasks and purposes, enable end-to-end, pixel-to-pixel semantic image segmentation without the need for manual expertise in feature design, and enable network applications with appropriate fine-tuning of the network structure or parameters. The network structure or parameters can be fine-tuned to make the network applicable to other scenarios. In recent years, there has been an increasing number of studies related to semantic segmentation of images using deep learning.Yadav, Patra et al. proposed a deep convolutional neural network based on the color line model[25], which can segment rural roads and is well suited for autonomous driving techniques.Liu, Li et al. optimized Markov random fields using deep learning[26] and achieved 77.5% on the PASCAL VOC-2012 dataset. 2012 dataset to achieve a Mean Intersection over Union (MIoU) ratio of 77.5%; Chen, Papandreou et al. proposed a deep learning system for semantic segmentation of color images called DeepLab[27] to achieve a MIoU of 79.7% on the same dataset. Mishra, Daescu et al. on the other hand applied deep learning convolutional neural networks to segmentation of medical images[28], which outperformed Otsu segmentation results in the segmentation task of color skin damage images. Deep learning is usually better for segmentation, but it requires a large number of labeled samples for training and GPU for acceleration.

When using traditional shallow machine learning algorithms to solve a pain point in production life, when encountering some unstructured data such as text, images, and audio, it is a very challenging problem to use what kind of features can efficiently represent these unstructured data. Unstructured data such as image data, the original image consists of pixels without semantic features, and the pixel-level features cannot make the machine understand the content of the image. Traditional machine learning to recognize an object often requires manual construction of some features. For example, to recognize a picture of a cat, researchers need to abstract some features of the cat, such as the cat's mouth, cat's tail, etc., based on expertise or long experience, and then combine these features in some form to recognize the image of the cat. Before the widespread use of deep learning, traditional machine learning used manually constructed features to learn, but in order to solve different problems, different features had to be constructed manually. The effectiveness of machine learning algorithms is seriously affected by the quality of the manually selected or constructed features. The manually selected or constructed features are often obtained by researchers with professional knowledge and long experience. Feature selection and construction consumes most of the time and effort in solving a machine learning problem. Because of the time and effort required to construct features manually, it is particularly important for algorithms to learn some high-level features from raw data, especially unstructured data, autonomously, and deep learning can achieve this goal and is therefore widely used and rapidly developing.

2 Medical image segmentation based on convolutional neural network

2.1 Introduction

In recent years, along with the booming development of science and technology and Internet industry, artificial intelligence technology has been increasingly applied to various aspects of scientific research, engineering technology and life services, such as recommendation systems, driverless, intelligent Q&A, computational advertising, quantitative investment, voice recognition, face recognition, etc. Among them, the most impressive is deep learning technology.

2.2 Introduction to deep learning techniques

Deep learning techniques have made a booming development in recent years, and it is one of the newest and most promising subfields of machine learning techniques. Deep learning technology can be thought of as deepening and transforming the artificial neural networks in traditional machine learning to process unstructured data such as text and images, which are difficult to extract features manually. The original intention of deep learning is to imitate the mechanism of human brain cognition and learning, and it uses a deeper and more complex network structure than traditional machine learning to extract high-level abstract features from raw data. Ultimately, deep learning is an extension of the artificial neural network in machine learning, and its core feature is the use of multi-layer artificial neural networks to extract high-level abstract features. The main advantage of deep learning over traditional machine learning is that it automatically extracts high-level features from raw data rather than researchers manually designing features that may be useful based on their expertise and deep understanding of the problem. Because of its good nonlinear fitting ability and significant breakthroughs in image classification, recommendation systems, target detection and tracking, computational advertising, image segmentation, quantitative investment, speech recognition, etc., deep learning techniques have become the focus of research in various universities, research institutes, and companies today. Deep learning can learn high level abstract features from low level features such as pixels of images, which in turn can be applied to many fields where traditional AI techniques are not competent.

When using machine learning algorithms to solve a real-world problem, it is a very challenging problem to represent unstructured data, such as text, images, and audio, efficiently with what kind of features. Unstructured data such as image data, the original image consists of pixels without semantic features, and the pixel-level features cannot make the machine understand the content of the image, the pixels are just numbers to the machine, and the pixels are discrete and cannot represent the local connection. For example, to recognize a picture of a cat, researchers need to abstract some features of the cat, such as the cat's mouth and tail, based on professional knowledge or long experience, and then use these features to combine to recognize the cat's image. Before deep learning was widely used, traditional machine learning used manually constructed features to learn, but different features needed to be constructed manually when facing different challenges. The quality of the manually selected or constructed features seriously affects the effectiveness of machine learning algorithms. Manual selection or construction of features often requires researchers' expertise and long experience to obtain them. Feature selection and construction consumes most of the time and effort in solving a machine learning problem. Because of the time and effort required to construct features manually, it is especially important for algorithms to learn some high-level features from raw data, especially unstructured data, autonomously, and deep learning can achieve this effect, which has been widely used and rapidly developed.

The famous backpropagation (BP) algorithm was proposed in the 1980s[29], as shown in Figure 1, which broke the bottleneck of artificial neural networks at that time and brought an opportunity for the development of traditional artificial neural networks, thus stimulating the enthusiasm for artificial neural networks. However, Marvin Minsky found through careful research that the typical perceptron in traditional artificial neural networks could not solve XOR or other linearly indistinguishable problems[30], and the applicability of shallow learning was very small, which led to the concern of researchers, and the research on traditional artificial neural networks came to a standstill, and the first peak period of artificial neural networks ended.

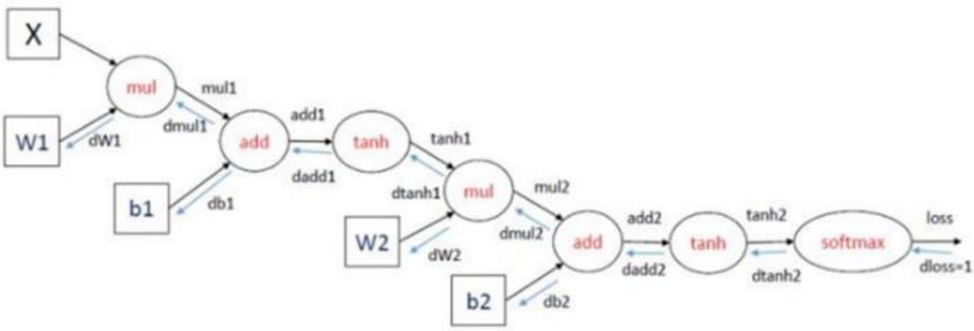

Figure 1 Diagram of BP algorithm

Until the early twentieth century, Professor Hinton, the flagship of the deep learning field, presented research results that delighted countless researchers[31], and deep learning received widespread attention and research in the industry again, enabling rapid development and application, making people's research on deep learning reach a peak once again.

Deep learning is based on artificial neural networks in traditional machine learning algorithms, the most representative model of which is the convolutional neural network (CNN). In the 1960s, Wiesel and Hubel studied the cells responsible for visual perception by inserting electrodes through a hole in the back of the cat's skull and discovered a special cell that could perceive visual information with a certain directional selectivity. In 1998, Le Cun, inspired by the cat study, simulated the visual cognition principle of "direction-selective cells" and invented a multilayer convolutional neural network[32], as shown in Figure 2. In 1998, Le Cun invented a multilayer convolutional neural network (CNN), named LeNet[33]. The model was then applied to small-scale problems such as MNIST, a handwritten digit recognition database, and as shown in Figure 3, the performance of the Le Net network is a significant improvement over traditional machine learning models. Convolutional neural networks (CNNs) are usually able to use image pixels directly as input, eliminating the need for manual feature selection and feature construction, automatically learning abstract features when training with large amounts of image data, and using weight sharing so that each convolutional kernel uses the same set of parameters for convolution, thus significantly reducing the size of parameters in the network. The famous Image Net competition provides a huge amount of free labeled image data to participants from all over the world in computer vision competitions, which provides ample learning materials and an arena for researchers from all over the world to give full play to their ingenuity, and greatly promotes the development of deep learning technology in computer vision. In October 2012, Alex and his teacher Hinton applied the deeper convolutional neural network AlexNet on top of the Image Net image classification competition with the network structure shown in Figure 4 and achieved an accuracy rate that shocked the industry, beating traditional machine learning models based on manual feature extraction by nearly 10% to win the image classification category[34]. AlexNet has achieved a major breakthrough in image classification without processing the original image and without researchers manually extracting any abstract features. AlexNet's excellent results have led to numerous researchers' enthusiasm for deep learning and computer vision, and have led to a number of significant research advances that have led to the application of deep learning to more and more aspects of computer vision. Kaming He et al. team from Microsoft Asia Research proposed Resnet with 152 layers in the ILSVRC 2015 competition, and its prediction of the five categories with the highest probability of being correct actually reached 96%, which exceeded the level of recognition of relevant objects in the dataset by ordinary people and attracted wide attention from the community[35-37].

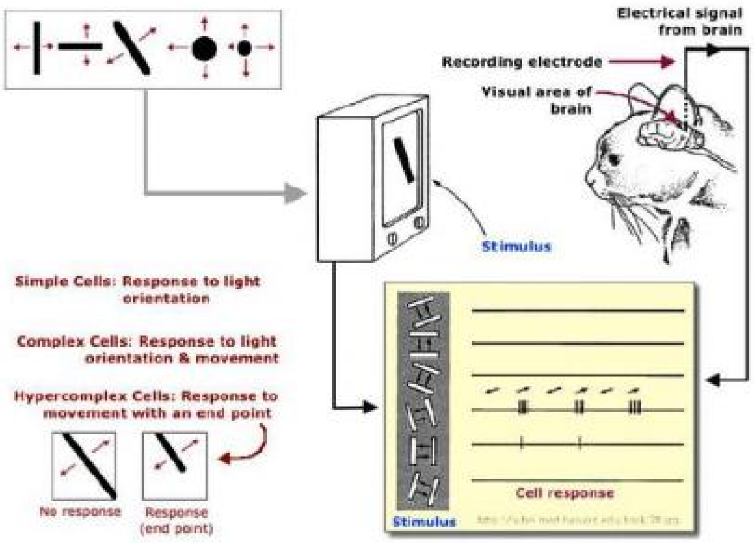

Figure 2 Diagram of directional selective cells of cat

Figure 3 Hand-written numeral recognition using LeNet

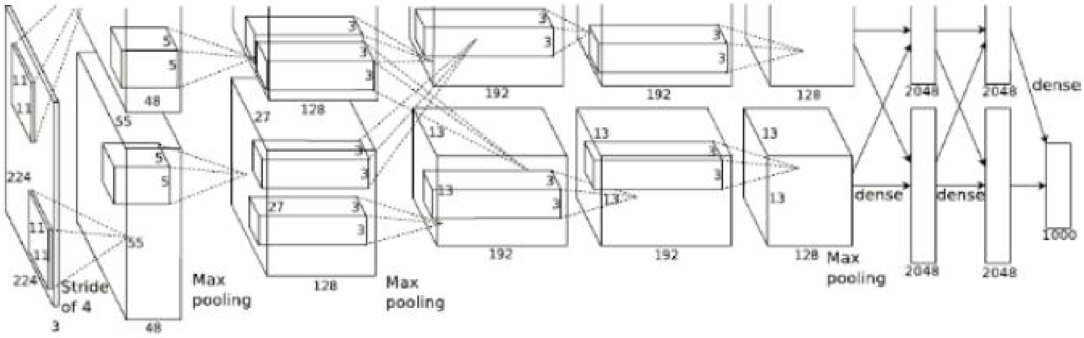

Figure 4 Architecture of AlexNet

2.3 Basic concepts of convolutional neural network (CNN) algorithms

Convolutional neural networks have a significant structural difference from networks with only fully connected layers, where adjacent layers are partially connected and the latter are fully connected. Since images often contain many pixel points, feeding them into a network with only fully connected layers would result in a large parameter size, which would require a lot of computational resources and be prone to overfitting, a problem that convolutional neural networks can effectively solve.

A typical convolutional neural network (CNN) is a deep neural network that can be used to process two-dimensional grid-type data, such as image data. Convolutional neural networks (CNNs) have yielded satisfactory results in a very large number of real-world scenarios. The difference from the original fully-connected layer-only network is that the network structure of CNNs contains a feature extractor capable of extracting high-level abstract features, which consists of convolutional and pooling operations in turn. In the Convolution Layer, each node is only partially connected to the adjacent network layers. Each layer generally contains several feature maps, and each feature map consists of a number of nodes arranged in a square pattern, and the nodes of the same feature plane The nodes in the same feature plane use the same weight parameters, i.e., shared weights. In the training process, the convolutional kernel is usually initialized randomly using a matrix of very small values, and the appropriate weights are finally trained through continuous learning. The shared weights of each kernel reduce the connectivity between adjacent layers and reduce the possibility of overfitting the convolutional neural network on the training set. The pooling layer, also known as the down sampling layer, generally consists of max pooling and mean pooling, with the former being more commonly used in practice. The former is more commonly used in practice. The convolutional and pooling layers of a convolutional neural network can greatly reduce the complexity of the network and significantly reduce the parameters of the network.

2.4 Basic components of convolutional neural network (CNN)

A typical convolutional neural network (CNN) is usually composed of several layers of neurons, and generally contains five basic components: Fully Connected Layer, Activation Layer, Convolution Layer, Pooling Layer, and Softmax Layer. In this network, the convolution operation and the pooling operation alternate, and the activation function generally uses the Relu function. When training the convolutional neural network, the parameters of the convolutional neural network can be updated layer by layer using the back propagation algorithm (BP) by using the gradient descent (SGD), Adam and other optimization methods. These basic components are briefly described in the following section.

2.4.1 Convolution layer

The convolutional layer, as shown in Figure 5, is an important part of the convolutional neural network, which limits the local perceptual field of nodes to a certain range to extract local features of data such as images, and can automatically extract higher-level abstract features from the feature map through learning. The shallow convolutional layer starts with pixel-level features and often extracts low-level image features, such as image corners, textures, edges and lines etc. These low-level image features are then transformed nonlinearly by the deeper convolutional layers using convolutional kernels to extract higher-level, more complex features such as shapes and contours, as shown in Figure 6. The convolutional neural network uses mainly smaller random numbers to initialize the parameters to initialize the weights of each convolutional kernel in the convolutional layers, and the inputs between layers are the original image pixel matrix and the feature maps formed after convolution or pooling. The convolutional layer is computed by using a fixed size convolutional kernel with pre-defined dimensions, sliding on top of the output feature map of the previous layer in pre-defined steps, performing inner product operations followed by other non-linear operations, and finally transforming the feature map of the previous layer into the feature map of the next layer, as shown in Figure 7. The size of the feature map is influenced by the size of the convolution kernel, the step of sliding, the amount of padding and the size of the input feature map, and the number of channels of the output feature map is determined by the number of convolution kernels.

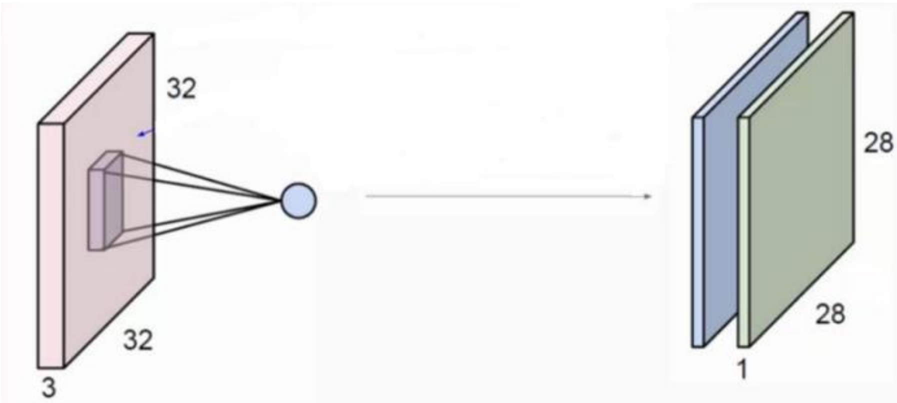

Figure 5 Convolution layer

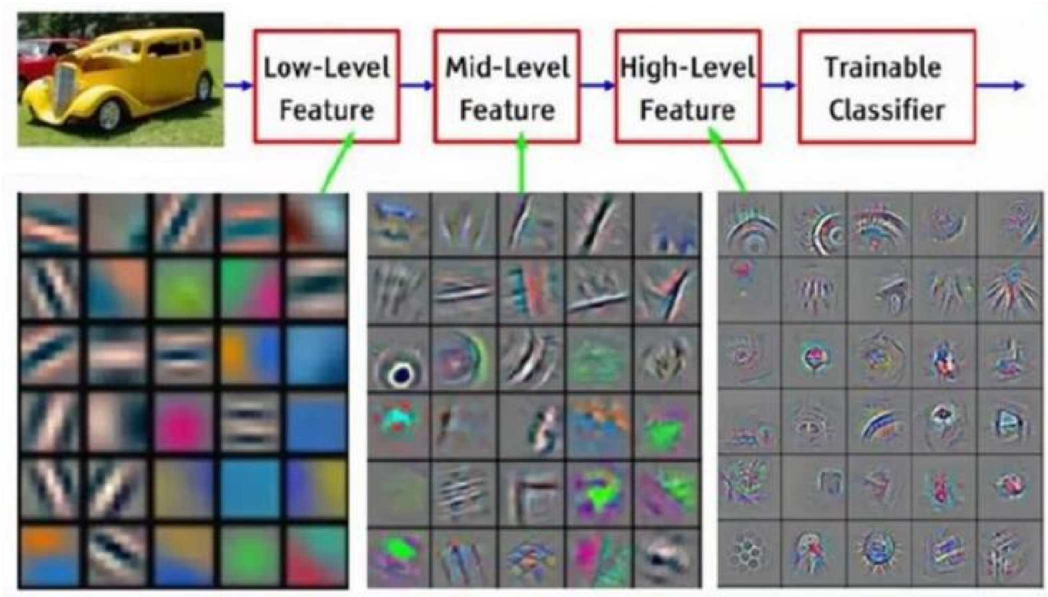

Figure 6 Diagram of convolution layer extract features step by step

Local perceptual field and weight sharing are the main parts that make convolutional neural networks different from the previous fully connected neural networks. Each convolutional kernel usually performs only local operations with the feature maps of the previous layer, and the number of output feature map channels is usually the same as the number of convolutional kernels. Usually the size of the convolutional kernel determines the range it can perceive, and the parameters are usually shared among neurons, which significantly reduces the size of the convolutional layer parameters.

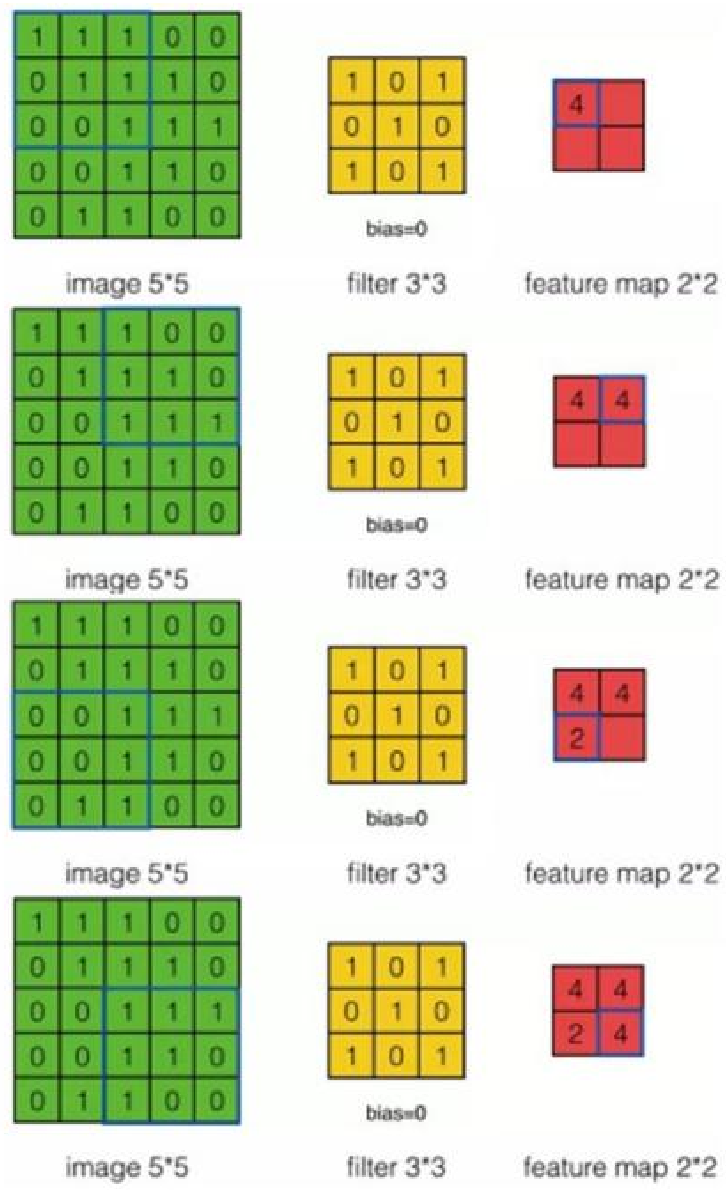

Figure 7 Convolution operation

2.4.2 Pooling Layer

The pooling layer is also often referred to as the downsampling layer. After the image has been slidingly convolved in the convolution layer, high-latitude features with local correlation are obtained, and if all the extracted features are used to make predictions, not only are they prone to overfitting, but also the computational effort will be relatively large. To cope with this difficulty and to achieve the goal of representing large-size feature maps with small-size feature maps, the resulting feature maps can be downsampled, a process called pooling, as shown in Figure 8, using a convolution kernel size of 2 × 2 and a sliding step of 2 to perform sliding, taking the largest element in each rectangle or the average of each element, so that the pooling operation is completed The original 4×4 size feature map is transformed into a 2×2 size feature map by this operation, and the elements on the four feature maps are turned into one element. The pooled features have much lower dimensionality than the convolutional ones, and significantly suppress the overfitting of the model, thus enhancing the generalization ability and reducing the computational complexity of the model. The commonly used pooling methods are maximum pooling and average pooling. If the input image of the model is shifted by a relatively small amount in a certain direction, the final predicted output of the model will not change much, which is called shift invariance, but the shift of the original image must be very small.

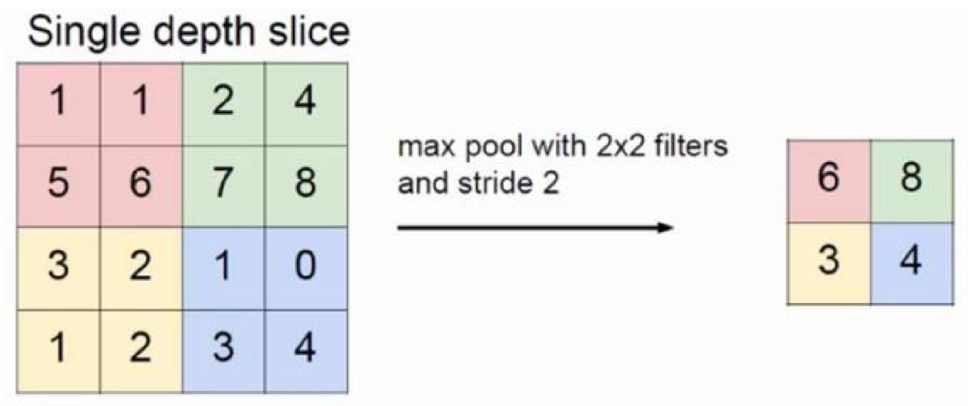

Figure 8 Pooling operation

2.4.3 Fully-connected Layer

The fully connected layer is globally connected, that is, the nodes of all neighboring network layers are connected to each other, and this connection is different from the local connection of the convolutional layer.

2.5 Convolutional neural network (CNN) based colorectal cancer image segmentation

Convolutional neural networks (CNN) can achieve very good results in images because of their deep network structure, which can continuously abstract high-level features from the original pixel values of images, and their ability to learn the feature representation of images from the original data. The shallower convolutional layers in a convolutional neural network have a smaller field of perception and can learn some local, detailed features, while the deeper convolutional layers have a larger field of perception and can learn higher-level, more abstract features. The higher-level abstract features are often less sensitive to small-scale changes in image translation, rotation, and scaling, which can effectively improve the robustness and stability of convolutional neural networks for computer vision tasks such as image classification. However, high level abstract features lose many detailed features and are less accurate for determining edges and contours in images. In contrast, image semantic segmentation requires a label for each pixel, which is a spatially (pixel) intensive computer vision task, and thus the original convolutional neural network can greatly affect the accuracy of the image segmentation task.

Image segmentation can be treated as a pixel-level classification, where each pixel in an image is assigned to a class. The general approach to image segmentation using a primitive convolutional neural network is that in order to classify each pixel in an image into a class, a predefined size image block, such as a 15×15 or 17×17 image block, is intercepted around each pixel point in the image as the training data for the convolutional neural network. Then, using the gold standard for image segmentation previously labeled manually, the category of the image block is usually the category to which the most central pixel belongs, and these blocks are then fed as training data, making it possible to finally output the category of that central pixel. Once the class of each pixel in the image is finally determined, the whole image can be segmented. The image is segmented by inputting a block of images centered on it pixel by pixel, and then classifying the block of images, the result of which is the class of the pixel, and finally post-processing to determine the region to which each object belongs.

The image segmentation based on traditional convolutional neural networks has the following major problems: first, it leads to a large storage overhead, because during training, many image blocks need to be intercepted in the original image, which requires more storage space to be allocated; second, it is computationally inefficient, because for adjacent pixels, the intercepted image blocks usually contain a very large overlap, which will lead to many pixels in the computation of convolutional Third, the size of the image block also limits the size of the perceptual field of the deep convolution layer, so that it cannot exceed the size of the intercepted image block, and the intercepted image block is usually much smaller than the whole image, which eventually makes it possible to extract only some local features within the size of the image block, thus making it difficult to achieve accurate segmentation and limiting the overall algorithm recognition performance. In general, the image segmentation based on convolutional neural network can roughly segment the general outline of the object, but it is poor at segmenting the edges of the object, and it is easy to over-segment or under-segment.

3 Medical image segmentation based on improved U-Net

3.1 Introduction

The original U-Net is used for colorectal cancer tumor segmentation, in which a simple convolutional neural network with fewer layers and a simpler structure is used to construct the encoder on the left side of the U-net after removing the fully connected layers and randomly initializing the parameters in the network. However, the convolutional neural network with fewer layers and simpler structure is relatively ineffective in image classification tasks, because its fewer layers and simpler structure do not sufficiently extract some abstract high-level features of the image and cannot maximize the effect of the encoder. Moreover, the training of random initialization parameters is slow, and the final result may not be satisfactory. Since the pooling layer in U-net loses the spatial detail information in the image during downsampling, this paper uses a more complex convolutional neural network and uses a multi-channel convolutional layer with a convolutional kernel size of 1×1 to replace the fully connected layer to build the encoder, and uses its pre-trained parameters for the image classification task to initialize the encoder, and will apply the null convolution to the encoder to further improve the accuracy of the image segmentation task. The segmentation results for the original U-net are less than satisfactory, still coarse compared to the manually labeled results, poorly segmented for smaller tumors, and with small over- and under-segmentation. To a certain extent, this result is due to the different sizes of tumor edges and the small area of the tumor in the whole image. Therefore, in this paper, the loss function of the original U-net is improved so that it can solve the previous segmentation problems.

3.2 U-Net Model

In the field of medical image segmentation, U-Net has been widely used as a very effective model in various applications, and it appeared almost simultaneously with the FCN model. The entire structure of U-Net is a Ushaped network, as shown in Figure 9.

U-Net is mainly divided into three parts: contraction path, jump connection, and expansion path. The systolic path encodes the original map through a convolutional layer, and then uses a pooling layer for downsampling, doubling the number of feature map channels each time it is encoded through the convolutional layer. The jump connection mainly sends the features of the systolic path directly to the corresponding expansion path, fusing the low-dimensional features with the high-dimensional features. The expansion path will decode the features of the systolic path through the deconvolution layer, then perform upsampling, and finally output the segmented mask. When upsampling is performed, there are two modes, one is to interpolate the original map directly, and the other is obtained by convolving the original map with a learnable deconvolution layer.

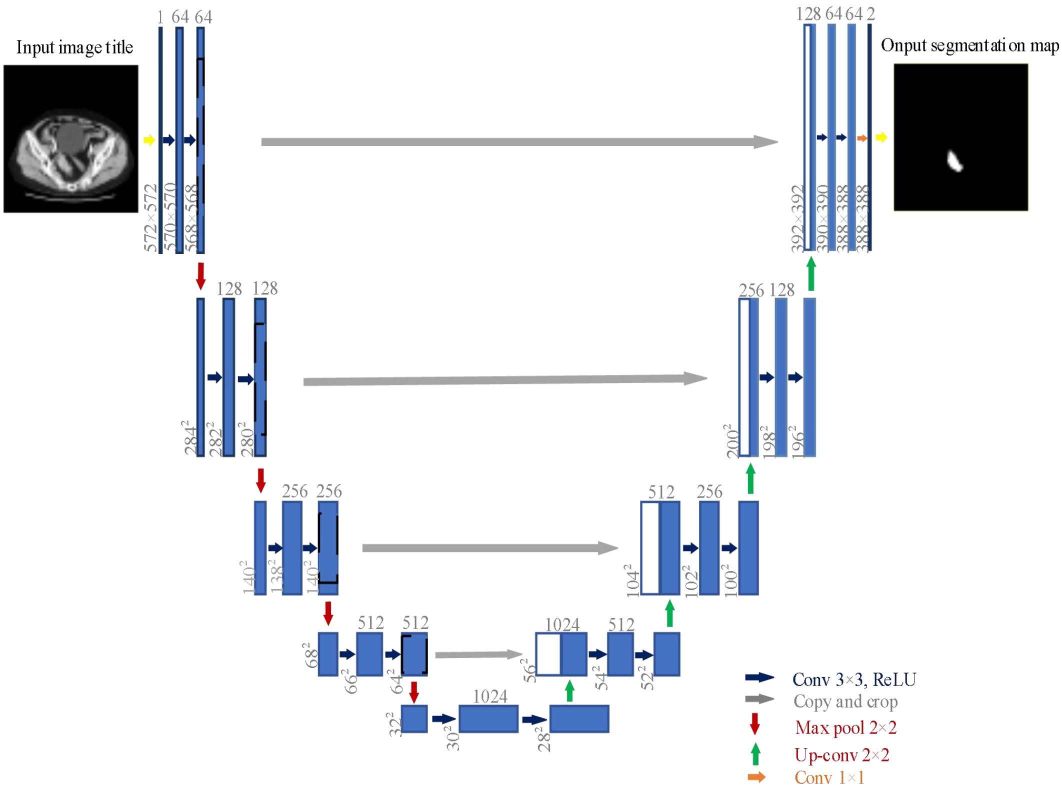

Figure 9 U-Net model

The symmetric structure of U-Net's decoding and coding with multiple downsampling to fuse multi-scale features has a great influence on the subsequent segmentation networks, and many networks have borrowed the structure of U-Net.

3.3 Medical image segmentation evaluation metrics

When evaluating the model in this study, the following evaluation indicators were used:

1) Loss function: the losses commonly used for image segmentation can be broadly classified into two types, one is directly on the pixel-based level and the other is on the image-based level. Hybrid loss functions have performed well in many studies, and therefore, this paper also samples both pixel-based and image-based hybrid loss functions[38-39]. The most common among the pixel-based loss functions is the cross-entropy function, which in the dichotomous case is formulated as follows:

$$L_{cross\_entropy} = \frac{1}{N}\sum_i L_i = \frac{1}{N}\sum_i -[y_i \cdot \log(p_i) + (1-y_i) \cdot \log(1-\mathrm{p_i})]$$

For each pixel point in the image, its corresponding loss is calculated and then accumulated. However, the cross-entropy function has a much smaller loss for the pixel points in the minority class than in the majority class when the classes are not balanced, resulting in a poor segmentation effect in the minority class. In medical images, the segmentation target is mostly small compared to the background, so the cross-entropy function is not applied directly to segmentation, but the weights between different categories are adjusted according to the differences between the segmentation categories. In order to adjust the adjust the loss between different classes, some researchers solve this problem by adjusting the weight of direct loss of unbalanced classes and adjusting the relative ratio of loss between difficult samples and simple samples, so that the model can learn for difficult samples, and this loss function is called Focal loss, which can be defined as:

$$L_{FL} = \frac{1}{N}\sum_i L_i = \frac{1}{N}\sum_i -[\alpha \cdot {y_i}^{\gamma} \cdot \log(p_i) + (1-\alpha) \cdot (1-y_i)^{\gamma} \cdot \log(1-\mathrm{p_i})]$$

Focal loss adjusts the relative contribution of difficult samples to the loss by the gamma parameter, while alpha is used to adjust the loss of pixels between unbalanced categories[40]. When gamma=0,alpha=1, it is equivalent to the cross-entropy loss function.Focal loss can adaptively increase the penalty on difficult samples, so that the model gradually focuses on learning difficult samples to achieve better results. Unlike the pixel-based loss function, the image-level loss function is based on the whole prediction probability map, such as the common Dice loss, which originates from the Dice Coefficient (Dice Similarity Coefficient), which directly calculates the overlap between the prediction map and the mask, and the larger the Dice Coefficient, the more the overlap, and it is one of the most commonly used in medical image segmentation It is one of the most commonly used evaluation criteria in medical image segmentation[41].

The Dice coefficient is calculated by ignoring the background and calculating the foreground directly, which is very useful for tasks with large backgrounds and small targets in medical images. It is expressed as a loss function as follows:

$$L_{Dice} = 1 - 2 \times \frac{\sum_{d=1}^{D}\sum_{h=1}^{H}\sum_{w=1}^{W}(p_{h,w,d}g_{h,w,d})+\varepsilon}{\sum_{d=1}^{D}\sum_{h=1}^{H}\sum_{w=1}^{W}(p_{h,w,d}+g_{h,w,d})+\varepsilon}$$

Among them, D, H, W correspond to the depth, width and length of the image respectively, and ε is the smoothing term, preventing the denominator from being 0.

2) Recall rate (Recall), for the prediction result P and the real label G, the higher the recall rate, the better the result. The recall rate is defined as:

$$Recall = \frac{TP}{TP+FN} = \frac{|P \cap G|}{|G|}$$

3) Precision, for the prediction result P and the true label G, the higher the precision, the better the result, and its false positive rate is defined as:

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{|P \cap G|}{|P|}$$

3.4 Improvement of loss function

In the localization stage, because the predicted mask is downsampled several times, it contains fewer pixel points in the foreground, and Dice loss is calculated by means of ratio, when there are fewer pixel points its will its value will be unstable, there are large fluctuations, and the model parameters cannot be updated effectively, resulting in difficulty in convergence of the model. In this study, the design of cross-entropy loss function with weights is borrowed in the localization stage[42], and the ratio of front and back view pixels is used to replace the weight parameter alpha in Focal Loss that needs to be adjusted manually.

$$L_{FL} = \frac{1}{N}\sum_i L_i = \frac{1}{N}\sum_i -[\alpha \cdot {y_i}^{\gamma} \cdot \log(p_i) + (1-\alpha)\cdot(1-y_i)^{\gamma} \cdot \log(1-\mathrm{p_i})]$$

$$\alpha = \frac{N}{P+N}$$

In the equation, for one of the categories, P is the number of pixels in its category and N is the number of pixels not in its category. In the segmentation stage, this paper combines Dice Loss and Focal loss to obtain the convergence speed of Dice loss in unbalanced samples while Focal loss is able to learn further for difficult samples, constituting a loss function of the following form.

$$L_{region} = L_{Dice} + \lambda \cdot L_{FL}$$

Many studies have shown that boundary information has a large impact on segmentation performance, and strengthening the weight of the boundary is beneficial for the model to obtain better performance[43]. In this paper, the tumor edge is obtained by expanding and eroding the original tumor label and adding it as a prediction target to the loss function, and the edge contour is extracted as shown in Figure 10:

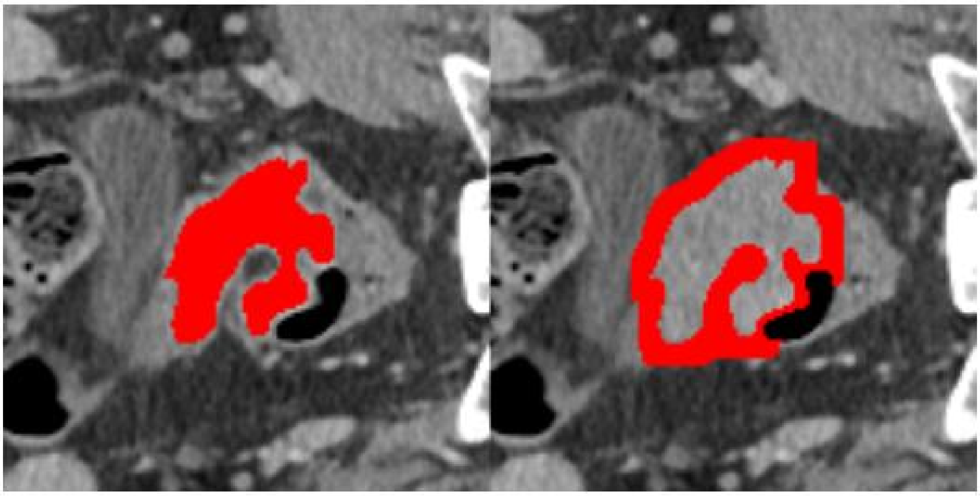

Figure 10 Contour schematic

In summary, the loss function used in the segmentation phase of this paper can eventually be expressed in the following form.

$$L = L_{region} + L_{Contour}$$

3.5 Experimental

The experiments in this paper were conducted on a computer with a 64-bit operating system, using python 3.6 to program the deep learning and deep multitasking learning frameworks and MATLAB®2016a for image cropping. Deep learning was developed on tensorflow2.0.

Tensorflow2.0 is developed on tensorflow1.0 version. tensorflow1.0 version has been widely used, but its counterpart pytorch has dynamic framework, while tensorflow1.0 only has static graph mode, such that tensorflow1.0's Therefore, Google has launched a new version of tensorflow, tensorflow2.0, which is a major upgrade to tensorflow1.0. tensorflow2.0 changes the default execution mode to eager, and reorganizes the The default execution mode of tensorflow2.0 has been changed to eager, and the relatively confusing modules of tensorflow1.0 have been reorganized. Moreover, tensorflow2.0 integrates the powerful and comprehensive features of tensorflow1.x with the easy-to-learn features of keras, and can directly call all the functions of keras.

3.5.1 Experimental data

There are 107 cases with 3029 CT images and corresponding mask images in this dataset, and the data samples are 512×512 pixels. And there are 1745 CT image sequences for each case. Among them, there are only 860 CT images with tumor area and the tumor area accounts for a small percentage of CT images, which is a positive The data set with negative sample imbalance. Therefore, when dividing the dataset, in order to make the positive and negative sample distributions of the training set, validation set and test set as similar as possible, the number of cases is used as the division object, and the number of training set, validation set and test set are divided by 8:1:1, respectively.

Through the mask map, it can be observed that the tumor locations are all within a certain area of the image, therefore, this paper uses the OpenCV library to calculate the maximum and minimum pixel values of the tumor area. Statistically, a pixel range of 126~382 in length and 202~458 in width can contain all tumor features. Therefore, the dataset is finally cropped into 256×256 size images, i.e., irrelevant features are filtered, while saving resources and reducing computation.

3.5.2 Image pre-processing

In today's rapid development of deep learning technology, the requirements of deep learning for data have also increased, and the scale and quality of data are crucial to the performance of the algorithm. However, the proportion of various samples collected is often unbalanced, and if the unbalanced data is used for training, it will often produce unsatisfactory prediction results, and the results tend to be more biased for categories with a large proportion of data. For example, in the dichotomous classification problem, the ratio of positive and negative samples is expected to be close to 1:1, but in different scenarios, the data ratio is often not what we expect, such as 1:2, 1:10, 1:100, or even more exaggerated, for example, in the task of classifying spam emails, normal emails often account for the majority and spam emails only for a very small number, so that the ratio of samples in the two categories will severely imbalanced, and the classification network will be biased towards the large ratio; in fact, people are more concerned about the few classified samples, so that the network predicted in this way loses its original meaning.

Data imbalance is divided into intra-class imbalance and inter-class imbalance. Intra-class imbalance refers to the fact that the data itself is complex and has a diverse distribution, where one class usually consists of several aggregated subclasses and the sample size of these subclasses is significantly unbalanced. The between-class imbalance refers to the apparently excessive difference between the amount of data in one category and the other categories in the data set, when the model parameters will tend to fit the category with a larger number, reducing the model effect.

Since the divided training set has only 1820 CT images, it is far from sufficient for deep learning. Therefore, the sample size of the training set needs to be expanded by image enhancement technique. In this paper, the sample size of the training set is expanded to 6 times of the original size using horizontal flip, vertical flip, 90-degree rotation, and the combination of flip and rotation. Image enhancement has a great improvement on the training of model performance, which can both reduce the occurrence of overfitting phenomenon and improve the generalization ability of the model. The effect is shown in Figure 11.

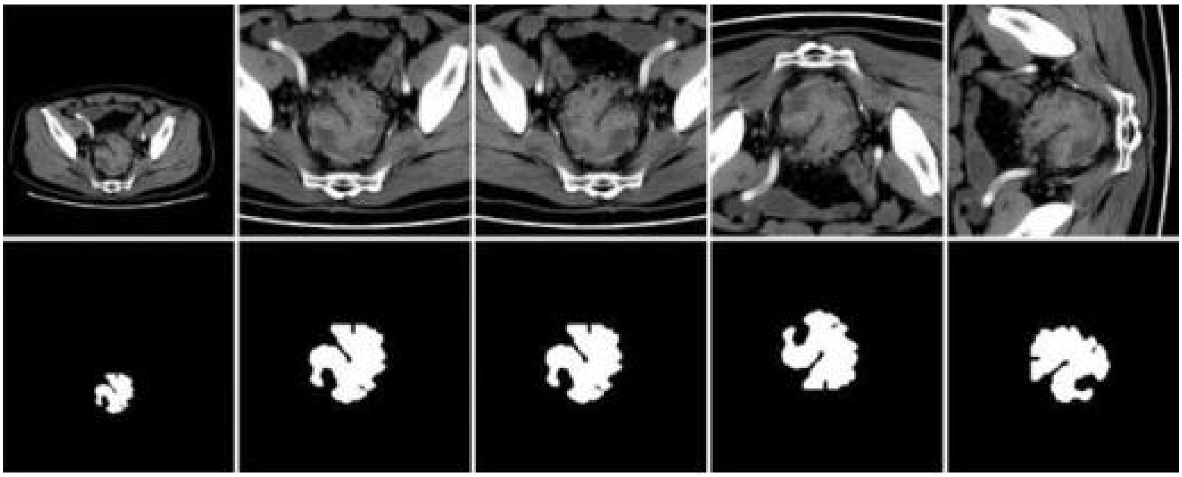

Figure 11 Image enhancement

To compare the impact of U-Net on the model performance, the performance of the original training set and the expanded training set were compared on the U-Net network separately. Tables 1 show that the increase in the amount of U-Net data improved the performance of the network.

Table 1 Comparison of performance indicators before and after data augmentation

Training set	P(%)	R(%)	L(%)
Original training set	83.26	78.90	82.42
Post-expansion training set	**85.76**	**84.85**	**86.98**

3.5.3 Data normalization

The specific meaning of data normalization is the normalization of the pixels in the image. If the data not normalized by the image pixels is used as the input of the neural network, it will often cause problems such as slow convergence of the network, and the Z-score method is used to normalize the dataset. The Z-score formula is as follows:

$$x^* = \frac{x-\mu}{\sigma}$$

where μ represents the mean of this experimental data and σ represents the variance of this experimental data.

3.5.4 Intercepting the POI region

Then, the ROI region (Region Of Interest) is extracted by first finding the center of the region of interest, and then using this center as the origin and extending it by 64 pixels at the top, bottom, left and right, and finally truncating a 128x128 pixel region as the input for deep learning and multi-task learning, as shown in Figure 12.

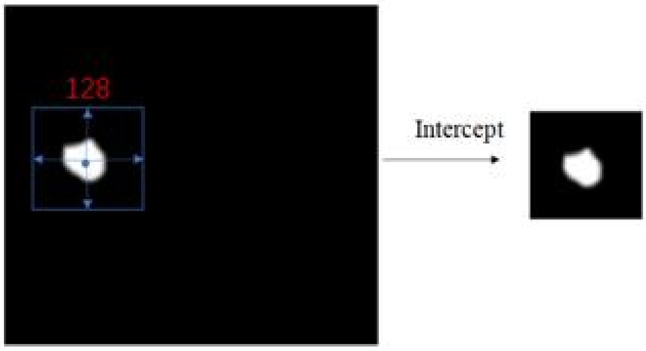

Figure 12 Intercept ROI area

4 Experiment results

The experimental configurations for all experiments in this paper are shown in Table 2：

Table 2 Experimental configuration

Experimental Environment	Concrete Arrangement
Operating System	Windows 10
CPU	Intel Core i7-9750H, 2.6GHz
GPU	NVIDIA GTX 1660Ti 6g
RAM	8G
Disk	512GB
Programming Languages	Python 3.6
Deep Learning Framework	Tensorflow/Keras
Other	CUDA 10.2、cuDNN 7.6

4.1 Hyperparameters

In this paper, we use the Adam algorithm to optimize the loss function, which combines the features of both AdaGrad and RMSProp to adjust the learning rate for each gradient without decaying the learning rate to so small that the model fails to learn later.

Table 3 Comparison of performance indicators of different optimizers

Optimizer	P(%)	R(%)	L(%)
SGDM	44.29	97.30	62.88
AdaGrad	43.56	91.15	82.52
DiffGrad	85.13	83.68	86.54
Adam	**85.76**	**84.85**	**86.98**

After several experiments, the initial learning rate set in this paper is 1×10-5. Although a small learning rate requires more training times, the model results will be better; while a larger learning rate makes the learning curve oscillate significantly, and may even lead to the gradient updating in the wrong direction.

Table 4 Parameter setting

Learning Rate	Learning rate decay		Early termination	Step length
	factor	patience		
10^-5	0.5	5	10	2

If the Batch is set too small, the model will take more time to train and the effect of batch normalization will not be obvious; if the Batch is set too large, the training speed will be accelerated, but overfitting will occur more easily.

4.2 Model results

In this paper, we build a U-Net network using Tensorflow as the back-end Keras framework structure and train the accelerated model on the server using GPU pairs. The total number of model parameters is 31507 and the time spent for one training session is 44 seconds, finally after the 45th training session. Meanwhile, in order to preserve the integrity of the tumor region during training, thresholding was performed only during testing, which resulted in the best processing. Through the final 200 iterations of training, the training set reached 85.98% similarity and the testing set reached 85.76% similarity, i.e., the segmentation accuracy reached 85.76%, as shown in Figure 13.

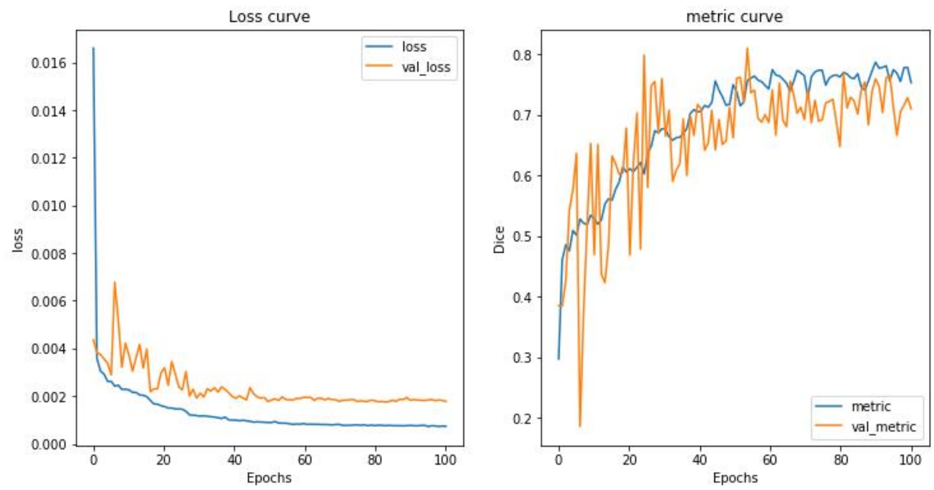

Figure 13 During the training process, the U-Net framework is computed for training loss and Dice for rectal cancer dataset

Table 5 shows the test results of the two algorithms on the colorectal cancer dataset.

Table 5 Test results

Network Structure	P(%)	R(%)	L(%)	Time/(epoch/s)
FCN	76.27	65.38	83.65	33.25
U-Net	**85.76**	**84.85**	**86.98**	**29.56**

Table 5 shows the segmentation effects of the two network models, and the experimental comparison with the two different network structures, FCN and U-Net, shows that the experimental effect of U-Net in colorectal cancer segmentation is significant. From the experimental results, we can see that the performance of FCN algorithm in segmentation is worse than U-Net method, which is due to the fact that FCN cannot make full use of the features between different layers, and the only way to restore the image to its original size is by upsampling, which lacks the connection between pixel location and classification in space, making the experimental results of FCN unsatisfactory. In contrast, U-Net is used as the main network, and jump connections are added to the network, thus more features are obtained. However, there are only a few shallow features in the coding layer and there is a large amount of redundant information, which will affect the image for effective segmentation. By comparing the analysis with other methods, the segmentation method proposed in this chapter has better results.

4.3 Analysis of results

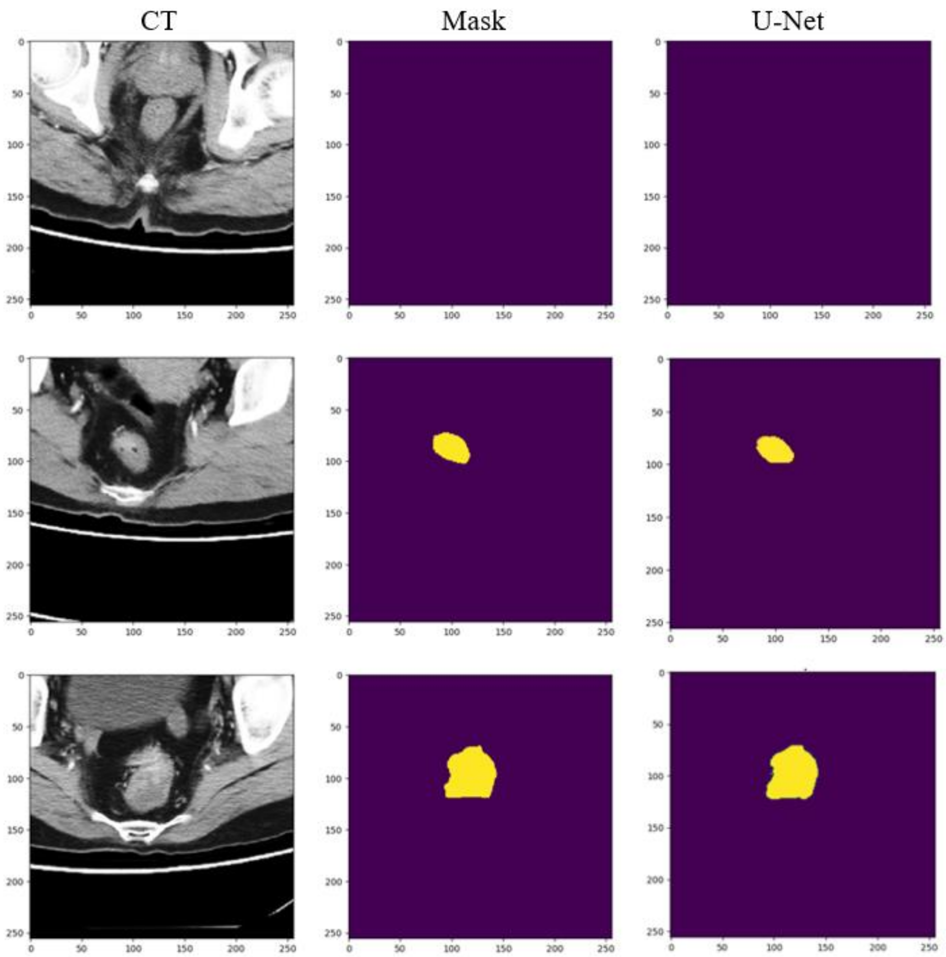

Figure 14 Segmentation results

In order to observe the visualization effect, three CT images in the test set were randomly selected for display, and the results are shown in Figure 14, where the yellow part is the segmented tumor location and the purple part represents the non-tumor location. From the visualization results, it can be seen that the U-Net network can segment the tumor location well, indicating that the model has good segmentation performance.

5 Conclusion

1. In this paper, the classical U-Net image segmentation algorithm is used as the basis and validated against the background of colorectal cancer dataset to segment the rectal tumor region.

2. In this paper, we propose an automatic tumor segmentation model based on U-Net neural network model for tumor segmentation of rectal cancer patients, and validate and analyze the similarity (Dice coefficient) of the segmentation. In this study, the segmentation of rectal cancer tumors using the U-Net network achieved the ideal state: 85.76%. This effectively solved a series of problems caused by doctors segmenting tumor locks with the naked eye and achieved automated and efficient segmentation.

3. I have presented an academic paper “A method for image segmentation of rectal cancer based on U-Net network” at the XIX Международная научно-практическая конференция студентов, аспирантов и молодых ученых (МОЛОДЕЖЬ И СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ).

6 Financial management, resource efficiency and resource saving

The purpose of this section discusses the issues of competitiveness, resource efficiency and resource saving, as well as financial costs regarding the object of study of Master's thesis. Competitiveness analysis is carried out for this purpose. SWOT analysis helps to identify strengths, weaknesses, opportunities and threats associated with the project, and give an idea of working with them in each particular case. For the development of the project requires funds that go to the salaries of project participants and the necessary equipment, a complete list is given in the relevant section. The calculation of the resource efficiency indicator helps to make a final assessment of the technical decision on individual criteria and in general.

This research work aims to develop a neural network architecture for automatic segmentation of biomedical images, with a particular focus on the task of automatically localizing rectal cancer tumor regions from computed tomography data. The development of a software system for processing CT images based on the algorithms described in this paper can detect, identify, and localize rectal cancer tumor regions, a development that has relevance and needs in the radiology departments of medical institutions. The application of this method to the problem of radioisotope diagnostics will qualitatively improve the accuracy of the conclusions drawn by experts on the research findings and significantly reduce the research time. Furthermore, the developed techniques for automatic segmentation of biomedical images can be the basis for developing training software for trainees in the field of radiology research.

6.1 Competitiveness analysis of technical solutions

In order to analyze the consumers of the study results, it is necessary to consider the target market and to segment it. The target market represents the segments that may be implemented and commercially developed in the future. Segmentation is the division of consumers into homogeneous groups, each of which may require a specific product.

Since the main consumers of the development are healthcare institutions, the healthcare profile of the professional institution and the type of tomography study used in that profile were selected as the main segmentation criteria. Segmentation was based on a weighted assessment between the prevalence and applicability of the tomography type (CT) in a given profile and the readiness of that profile for the introduction of automated biomedical image annotation techniques. The segmentation results are shown in Table 6.

Table 6 Map of market segmentation by medical care profile

Medical Type	**Research Type**			
	CT	MRI	X-ray	Ultrasound
Gastroenterology				
Hematology				
Genetics				
Gynecology				
Dermatology				
Cardiology				
Neurology				
Oncology				
Otolaryngology				
Ophthalmology				
Psychiatry				
Pulmonology				
Rheumatology				
Dentistry				
Traumatology and Orthopedics				
Urology				
Surgery				
Endocrinology				

A list of the main competing technology options is shown in Table 7.

Table 7 List of the main competitive technical solutions

No.	Product Name	Manufacturer Country
1	Zebra «AI1»	Israel
2	Arterys «Lung AI»	USA
3	InferVision «InferRead»	China
4	DeepWise «Dr. Wise»	China
5	Items proposed in the current study	Russia

Select the following criteria for competitive analysis:

1. segmentation speed - the speed of image processing.

2. segmentation quality - an assessment of the quality of the information extracted by the researcher after processing the biomedical image. Thus, the product has the lowest score, which results in a numerical assessment of the probability of the presence of the desired pathological sign, and the highest score, which is the precise location of the pathological category by automatic determination.

3. energy efficiency - assessment of the need to run equipment that requires more computing resources and therefore electricity, a parameter that indirectly affects the need to purchase additional expensive equipment.

4. universality - to determine the degree of possibility of using the product within the medical profile and the flexibility of the product with respect to new tasks, for example, retraining the model to use new, as yet unstudied, pathologies or changing the principles of determining the distribution of signs in pathology.

5. development potential - reflecting the potential to improve the algorithm without major changes to its architecture.

6. accessibility - an assessment of the degree of openness and transparency of development. It depends on the availability of documentation, the openness of the source code, and the readiness of the manufacturing company to cooperate.

The analysis of the competing technical solutions was determined by equation. The results of the competitive analysis are shown in Table 8. Analysis of competitive technical solutions is determined by the formula:

$$C = \sum W_i \cdot P_i,,$$

C - the competitiveness of research or a competitor;

Wi – criterion weight;

Pi – point of i-th criteria.

Table 8 Evaluation card for comparison of competitive technical solutions

Evaluation criteria	**Criterion weight**	**Points**					**Competitiveness** Taking into account weight coefficients				
		P1	P_2	P_3	P_4	P_5	C_1	C_2	C_3	C_4	C_5
1	2	3	4	5	6	7	8	9	10	11	12
Technical criteria for evaluating resource efficiency											
1. Splitting speed	0.06	5	2	3	4	4	0.30	0.12	0.18	0.24	0.24
2. Dividing quality	0.26	3	2	4	5	3	0.18	0.12	0.24	0.30	0.18
3. Energy efficiency	0.04	5	3	2	3	3	0.30	0.18	0.12	0.18	0.18
4. Multi-functionality	0.08	3	2	3	4	3	0.18	0.12	0.18	0.24	0.18
5. Development potential	0.24	3	2	3	5	5	0.18	0.12	0.18	0.30	0.30
Economic criteria for performance evaluation											
1. Availability	0.20	5	6	3	2	5	0.30	0.30	0.18	0.12	0.30
2. Price	0.12	5	3	1	3	5	0.30	0.24	0.06	0.18	0.30
Total	**1.00**	**29**	**20**	**19**	**26**	**28**	**1.74**	**1.20**	**1.14**	**1.56**	**1.68**

Based on the results of the competitive analysis, it can be concluded that the proposed method is competitive compared to the other considered methods, mainly due to its flexibility and potential to solve a wide range of problems, as well as its availability for the Russian market. However, at this stage of its development, the proposed project lost out to the easily available and inexpensive Israeli product from Zebra-Med.

It is also worth noting that the results demonstrated by the Chinese company

DeepWise set the bar high for the technology under development. But this company is strongly focused on the Chinese market and does not currently demonstrate a willingness to cooperate internationally, a fact that reduces the competitiveness of its product.

6.2 SWOT analysis

The results of the SWOT analysis are taken into account when developing the work structure as part of the project development. Given the characteristics of the industry under study, it was necessary to conduct a comprehensive analysis of the project. During the SWOT analysis, the final matrix was prepared (Table 9).

Table 9 Matrix of SWOT-analysis

	Strengths: S1. High precision of dividing S2. Fast dividing speed S3. Wide range of applications S4. Open source S5. Lack of competitors in the Russian Federation	**Weaknesses:** W1. Loses out to some foreign competitors in terms of accuracy W2. Highly hardware dependent W3. Requires large amount of training data
Opportunities: O1. Further network training to improve the quality of segmentation O2. Rapid development of machine learning O3. Growth in hardware performance O4. Creation of a database of training examples	*1. Due to the development of machine learning and the emergence of new methods and algorithms, it is possible to introduce new optimizations and heuristics in this development to improve the quality of segmentation* *2. With the increase of computer hardware performance, the replacement of components leads to the increase of segmentation speed* *3. Open source code and wide range of applications will allow you to add new features capabilities* *4. The advantages of the project will attract potential customers and organize fruitful cooperation with medical institutions of the Russian Federation*	*Strategy which based on weaknesses and opportunities:* *1. Using more efficient hardware will increase the competitiveness of development* *2. Creating a database with a large number of training examples is a very laborious process and requires the involvement of many specialists* *3. As the hardware productivity increases, the price of equipment decreases, and in turn, refusal to support the project with updated hardware may lead to a decrease in the competitiveness of the project*

Threats: T1. Emergence of competitively developed analogues in the market T1. Untimely funding	*Strategy which based on strengths and threats:* *1. The high accuracy and speed of segmentation and the wide range of applications will avoid the emergence of competing analogs and the loss of developmental relevance*	*Strategy which based on weaknesses and threats:* *1. Weaknesses of the project may lead to the emergence of competing analogues for development* *2. Lack of appropriate funding for the preparation of training datasets by experts makes the project impossible to develop*

Based on the obtained data, it can be concluded that the main drawback of this development is the need for more training examples with annotations, the solution of which requires the participation of expensive labeling specialists (radiologists) in the project. In addition, the high computational complexity of the task of training experimental neural network models requires the purchase of modern computing equipment.

6.3 Project Initiation

Project stakeholders, individuals and organizations that are actively involved in the project or whose interests are affected by the execution or completion of the project are listed in Table 10.

Table 10 Stakeholders of the project

Stakeholders of the project	Stakeholder expectations
НИ ТПУ, ИШИТР	Acquisition of a neural network algorithm for locating areas of lung infiltration based on computed tomography data for further implementation and presentation at a conference.
Medical Center for Enteropathy, Affiliated Hospital of Henan University of Science and Technology	High radiologists work efficiently in preparing medical reports on the presence of lesions, which are tumors in the rectal region.
Medical research groups and institutions	This development will allow a deeper study of the signs and principles of the formation of various pathologies through intelligent neural network analysis of biomedical images.

Companies engaged in the development and production of medical devices	By achieving very high quality in segmentation and pathology detection, this development can be the basis for forming intelligent reports received from the equipment performing radiological examinations.
Student	Master's Thesis Defense.

Table 11 Purpose and results of the project

Purpose of project:	Automatically segment the CT images of colorectal cancer, visualize the shape, size and location of the tumor, and assist doctors in diagnosing rectal cancer.
Expected results of the project:	Create a visualized tumor image and compare it with the mask map annotated by the doctor to provide more detailed tumor information.
Criteria for acceptance of the project result:	Ensure that the accuracy of rectal cancer tumor segmentation is higher than 85%.
Requirements for the project result:	1. The project must be completed by May 31, 2022 of the year.
	2. The results obtained must meet the acceptance criteria for the project result.

The organizational structure of the project

It is necessary to solve some questions: who will be part of the working group of this project, determine the role of each participant in this project, and prescribe the functions of the participants and their number of labor hours in the project.

Table 12 Participant of the project

№	Participant	Role in the project	Functions	Labor time, hours
1	Supervisor	Head of project	1. Drafting and approval of terms of reference. 2. Choice of research direction. 3. Schedule of project work. 4. Results evaluation and iterative changes in algorithm structure. 5. Evaluation of the validity of the results obtained. 6. Evaluation of the feasibility of further research on this topic.	150 hours

2	Master's student	Executor	1. Selection and study of subject material. 2. Study of existing solutions in the field 3. Selection of research direction. 4. Theoretical development of the algorithmic structure of the problem solving. 5. Construction of the trial version of the algorithm. 6. Evaluation of results and iterative changes in the structure of the algorithm. 7. Comparison of the results of the final version of the algorithm with other solutions in the field. 8. Evaluation of the validity of the obtained results. 9. Evaluation of the feasibility of further research on this topic.	534 hours

Project limitations

Project limitations are all factors that can be as a restriction on the degree of freedom of the project team members.

Table 13 Project limitations

Factors	**Limitations / Assumptions**
3.1. Project's budget	115000 RUB
3.1.1. Source of financing	TPU
3.2. Project timeline:	01/12/2021 to 31/05/2022
3.2.1. Date of approval of plan of project	01/10/2021
3.2.2. Completion date	31/05/2022

Project Schedule

As part of planning a science project, you need to build a project timeline and a Gantt Chart.

Table 14 Project Schedule

Job title	Duration, working days	Start date	Date of completion	Participants
Drawing up and approval of the terms of reference	4 days(without 2 weekend)	01/12/2021	06/12/2021	Supervisor
Selecting and studying materials on the topic	21 days(without 10 weekend, 9 holiday and 1 exam day)	01/12/2021	09/01/2022	Student
Study of already existing solutions in the field	7 days(without 4 weekend and 3 exam day)	10/01/2022	25/01/2022	Student
Choosing the direction of research	5 days(without 2 weekendand)/ 3days(without 2 weekendand)	17/01/2022/ 26/01/2022	23/01/2022/ 30/01/2022	Supervisor/ Student
Calendar of work on the topic	5 days(without 2 weekend)	24/01/2022	30/01/2022	Supervisor
Theoretical development of the structure of the algorithm that solves the set task	8 days(without 2 weekendand)	31/01/2022	09/02/2021	Student
Construction of trial versions of the algorithm	27 days(without 10 weekend and without 4 holiday)	10/02/2021	21/03/2021	Student
Evaluation of the results and iterative modification of the structure algorithm	11 days(without 5 weekendand and without 5 holiday)/ 35dayss(without 15 weekendand and without 5 holiday)	25/04/2021/ 22/03/2022	15/05/2021/ 15/05/2022	Supervisor/ Student
Comparison of the results of the final version of the algorithm with other solutions in this area	5 days(without 2 weekendand)	16/05/2021	22/05/2021	Student
Evaluating the effectiveness of the results obtained	5 days(without 2 weekendand)/ 5days(without 2 weekendand)	18/05/2021/ 23/05/2022	24/05/2021/ 29/05/2022	Supervisor/ Student
Assessing the feasibility of conducting further research on this topic	4 days(without 2 weekendand)/ 2days	25/05/2021/ 30/05/2022	30/05/2021/ 31/05/2022	Supervisor/ Student

The total estimated duration of the project is 113 working days. The total labor intensity of the supervisor and engineer is 34 and 113 days, respectively. The total duration of the work phases, including the possibility of executing the planned work in parallel, totaled 118 calendar days. The chronological values of the work phase durations are summarized in the linear schedule of the project implementation, as shown in Table 15.

6.4 Scientific and technical research budget

The amount of costs associated with the implementation of this work is the basis for the formation of the project budget. This budget will be presented as the lower limit of project costs when forming a contract with the customer.

To form the final cost value, all calculated costs for individual items related to the manager and the student are summed.

In the process of budgeting, the following grouping of costs by items is used:

- material costs of scientific and technical research;
- costs of special equipment for scientific work (Depreciation of equipment used for design);
- basic salary;
- additional salary;
- labor tax;
- overhead.

Table 15 Gantt chart

№	Participants	Tc, days	Duration of the project																	
			December			January			February			March			April			May		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1	Supervisor	4																		
2	Student	21																		
3	Student	7																		
4	Supervisor/ Student	5/ 3																		
5	Supervisor	5																		
6	Student	8																		
7	Student	27																		
8	Supervisor/ Student	11/ 35																		
9	Student	5																		
10	Supervisor/ Student	5/ 5																		
11	Supervisor/ Student	4/ 2																		

6.5 Calculation xof material costs

The calculation of material costs is carried out according to the formula:

$$C_m = (1+k_T)\cdot\sum_{i=1}^{m} P_i \cdot N_{consi} ,$$

where m – the number of types of material resources consumed in the performance of scientific research;

N_{consi} – the amount of material resources of the i-th species planned to be used when performing scientific research (units, kg, m, m^2, etc.);

P_i – the acquisition price of a unit of the i-th type of material resources consumed (rub./units, rub./kg, rub./m, rub./m^2, etc.);

k_T – coefficient taking into account transportation costs.

Prices for material resources can be set according to data posted on relevant websites on the Internet by manufacturers (or supplier organizations).

Energy costs are calculated by the formula:

$$C = P_{el} \cdot P \cdot F_{eq} ,$$

where P_{el} – power rates (5.8 rubles per 1 kWh for Tomsk);

P – power of equipment, kW;

F_{eq} – equipment usage time, hours.

Table 16 Material costs

Name	Unit	Amount	Price per unit, rub.	Material costs, rub.
Video card Asus GeForce GTX 1660 Ti, 6 Gb	Pcs	1	38680,00	38680,00
Zalman Wattbit power supply, 600W	Pcs	1	3050,00	3550,00
Silicon Power DDR4 RAM, 8 Gb	Pcs	1	3100,00	6200,00
A4 printer paper, 500 liters	Dozen	1	274,00	274,00
Total				48704,00

Costs of special equipment

This point includes the costs associated with the acquirement of special equipment (instruments, stands, devices and mechanisms) necessary to carry out work on a specific topic.

Table 17 a Costs of special equipment (+software)

№	equipment identification	Quantity of equipment	Price per unit, rub.	Total cost of equipment, rub.
1.	Laser printer	1	11000	11000

OR

Calculation of the depreciation. Depreciation is not charged if an equipment cost is less than 40 thousand rubles, its cost is taken into account in full.

If you use available equipment, then you need to calculate depreciation:

$$A = \frac{C_{\text{перв}} * \text{Н}_{\text{а}}}{100}$$

A - annual amount of depreciation;

Сперв - initial cost of the equipment;

$\text{Н}_{\text{а}} = \frac{100}{\text{Т}_{\text{сл}}}$ - rate of depreciation;

Тсл - life expectancy.

Table 17 b Depreciation of special equipment (+software)

№	equipment identification	Quantity of equipment	Total cost of equipment, rub.	Life expectancy, year	Depreciation for the duration of the project, rub.
1.	Personal Computer	1	75000	6	12500

6.6 Basic salary

This point includes the basic salary of participants directly involved in the implementation of the work on this research. The value of salary costs is determined based on the labor intensity of the work performed and the current salary system.

The basic salary (Sb) is calculated according to the following formula:

$$S_b = S_a \cdot T_w ,$$

Where S_b – basic salary per participant;

T_w – the duration of the work performed by the scientific and technical worker, working days;

S_a - the average daily salary of an participant, rub.

The average daily salary is calculated by the formula:

где S_m – monthly salary of an participant, rub .;

M – the number of months of work without leave during the year:

at holiday in 48 days, M = 10.4 months, 6 day per week;

at holiday in 24 days, M = 11.2 months, 5 day per week;

F_v – valid annual fund of working time of scientific and technical staff.

Table 18 The valid annual fund of working time

Working time indicators	
Calendar number of days	365
The number of non-working days - weekend - holidays	 52 14
Loss of working time - vacation - sick absence	48
The valid annual fund of working time	251

Monthly salary is calculated by formula:

$$S_{month} = S_{base} \cdot (k_{premium} + k_{bonus}) \cdot k_{reg},$$

where S_{base} – base salary, rubles;

$k_{premium}$ – premium rate;

k_{bonus} – bonus rate;

k_{reg} – regional rate.

Table 19 Calculation of the base salaries

Performers	S_{base}, rubles	$k_{premium}$	k_{bonus}	k_{reg}	S_{month}, rub.	W_d, rub.	T_p,	W_{base}, rub.
Supervisor	35120				1404,80	1,699	34	81149,68
Student	16139				645,56	1,699	113	123939,12
Total:								205088,79

6.7 Additional salary

This point includes the amount of payments stipulated by the legislation on labor, for example, payment of regular and additional holidays; payment of time associated with state and public duties; payment for work experience, etc.

Additional salaries are calculated on the basis of 10-15% of the base salary of workers:

$$W_{add} = k_{extra} \cdot W_{base},$$

where W_{add} – additional salary, rubles;

k_{extra} – additional salary coefficient;

W_{base} – base salary, rubles.

6.8 Labor tax

Tax to extra-budgetary funds are compulsory according to the norms established by the legislation of the Russian Federation to the state social insurance (SIF), pension fund (PF) and medical insurance (FCMIF) from the costs of workers.

Payment to extra-budgetary funds is determined of the formula:

$$P_{social} = k_b \cdot (W_{base} + W_{add})$$

where kb – coefficient of deductions for labor tax.

In accordance with the Federal law of July 24, 2009 No. 212-FL, the amount of insurance contributions is set at 30%. Institutions conducting educational and scientific activities have rate - 27.1%.

Table 20 Labor tax

	Project leader	**Engineer**
Coefficient of deductions	27.1%	
Salary (basic and additional), rubles	81149,68	123939,12
Labor tax, rubles	21991.56	3358.8

6.9 Overhead costs

Overhead costs include other management and maintenance costs that can be allocated directly to the project. In addition, this includes expenses for the maintenance, operation and repair of equipment, production tools and equipment, buildings, structures, etc.

Overhead costs account from 30% to 90% of the amount of base and additional salary of employees.

Overhead is calculated according to the formula:

$$C_{ov} = k_{ov} \cdot (W_{base} + W_{add})$$

where k_{ov} – overhead rate.

Table21 Overhead

	Project leader	**Engineer**
Overhead rate	30%	
Salary, rubles	81149,68	123939,12
Overhead, rubles	24344,90	37181,74

6.10 Other direct costs

Energy costs for equipment are calculated by the formula:

$$C = P_{el} \cdot P \cdot F_{eq},$$

where P_{el} – power rates (5.8 rubles per 1 kWh);

P – power of equipment, kW;

F_{eq} – equipment usage time, hours.

Table 22 Electricity costs for technical purposes

Equipment	P_{el}	P	F_{eq}	**Per/kWh**	**Cost, rubles**
Supervisor of personal computers	0,6	0.234	163,2	5.8	133,66
Engineer's personal computer	0,8	0.308	723,2		779,61
Workstations/personal computers with installed components purchased for the project	0.9	0.486	720,0		1224,72
Inkjet printer	1.0	0.550	4,0		7,70
Total:					2145,69

6.11 Formation of budget costs

After calculating all articles of the development cost estimate, the total cost of the project "Mathematics and software for localization of rectal tumor areas from computed tomography data" could be determined. The calculation of the project implementation cost is shown in Table 23. The cost of the project was 366, 315.47 rubles.

Table 23 Items expenses grouping

Name	**Cost, rubles**
1. Material costs	48704.00
2. Equipment costs	23500
3. Basic salary	205088,79
4. Additional salary	0
5. Labor tax	25350,36
6. Overhead	61526,63

7. Other direct costs	2145,69
Total planned costs	**366315,47**

6.12 Evaluation of the comparative effectiveness of the project

Determination of efficiency is based on the calculation of the integral indicator of the effectiveness of scientific research. Its finding is associated with the definition of two weighted average values: financial efficiency and resource efficiency.

The integral indicator of the financial efficiency of a scientific study is obtained in the course of estimating the budget for the costs of three (or more) variants of the execution of a scientific study. For this, the largest integral indicator of the implementation of the technical problem is taken as the calculation base (as the denominator), with which the financial values for all the options are correlated.

The integral financial measure of development is defined as:

$$I_f^d = \frac{C_i}{C_{\max}}$$

where I_f^d – integral financial measure of development;

C_i – the cost of the i-th version;

C_{max} – the maximum cost of execution of a research project (including analogues).

The obtained value of the integral financial measure of development reflects the corresponding numerical increase in the budget of development costs in times (the value is greater than one), or the corresponding numerical reduction in the cost of development in times (the value is less than one, but greater than zero).

Since the development has one performance, then $I_f^d = 1$.

The integral indicator of the resource efficiency of the variants of the research object can be determined as follows:

$$I_m^a = \sum_{i=1}^{n} a_i b_i^a \qquad I_m^p = \sum_{i=1}^{n} a_i b_i^p$$

where I_m – integral indicator of resource efficiency for the i-th version of the development;

a_i – the weighting factor of the i-th version of the development;

b_i^a, b_i^p – score rating of the i-th version of the development, is established by an expert on the selected rating scale;

n – number of comparison parameters.

The calculation of the integral indicator of resource efficiency is presented in the form of Table 24.

Table 24 Evaluation of the performance of the project

Criteria	**Weight criterion**	**Points**
1. Energy efficiency	0.12	13
2. Reliability	0.20	12
3. Safety	0.10	13
4. Functional capacity	0.18	13
Economic criteria for performance evaluation		
1. Development cost	0.14	12
2. Market penetration rate	0.18	13
3. Expected lifecycle	0.08	9
Total	1	85

The integral indicator of the development efficiency (I_e^p) is determined on the basis of the integral indicator of resource efficiency and the integral financial indicator using the formula:

$$I_{\text{e}}^{p} = \frac{I_{m}^{p}}{I_{f}^{d}}, \quad I_{\text{e}}^{a} = \frac{I_{m}^{a}}{I_{f}^{a}}$$

$$I_{\text{исп.2}} = \frac{I_{p\text{–исп2}}}{I_{\text{финр}}^{\text{исп.2}}} \text{ и т.д.}$$

Comparison of the integral indicator of the current project efficiency and analogues will determine the comparative efficiency. Comparative effectiveness of the project:

$$E_{c} = \frac{I_{a}^{p}}{I_{e}^{a}}$$

Thus, the effectiveness of the development is presented in Table 25.

Table 25 Efficiency of development

№	Indicators	Points
1	Integral financial measure of development	12
2	Integral indicator of resource efficiency of development	15
3	Integral indicator of the development efficiency	13

Comparison of the values of integral performance indicators allows us to understand and choose a more effective solution to the technical problem from the standpoint of financial and resource efficiency.

6.13 Conclusion

No profit is expected from the direct implementation of the current project. It is assumed that the implementation of this project could make a significant contribution to the start-up of large projects involving the use of artificial intelligence tools to perform radiation diagnostics.

According to BusinesStat estimates, the natural number of studies in the tomography diagnostics market reaches 10.9 million in 2021. The most common type of tomography in 2021 is computed tomography (68.5% of the natural market volume). In addition, within the framework of this study, it was found that among the disease types, respiratory diseases are the most common group: in 2021, the prevalence of this group is 24.8% of the total national indicator. According to the above-mentioned study, more than 1.5 million CT scans are performed annually to diagnose the rectal cancer area, so more than 11,000 were performed in the Tomsk region.

The tool for analyzing CT images using artificial intelligence was used only for the task of analyzing the lung region in the Tomsk region and was priced at 250 rubles. For an automated study it can provide a profit of more than 1.3 million rubles per year.

7 Social responsibility

7.1 Introduction

The developed project aims to use the neural network model to accurately segment rectal cancer tumor images and provide a reference for doctors to diagnose and analyze the disease. The development of the program is only carried out with the help of computer.

In this section, harmful and dangerous factors affecting the work of personnel will be considered, the impact of the developed program on the environment, legal and organizational issues, measures in emergency situations will be considered.

The work was carried out in the hall of residence of TPU (8th floor). Room 85B was a research execution place.

The layout of the room is shown in Figure 15.

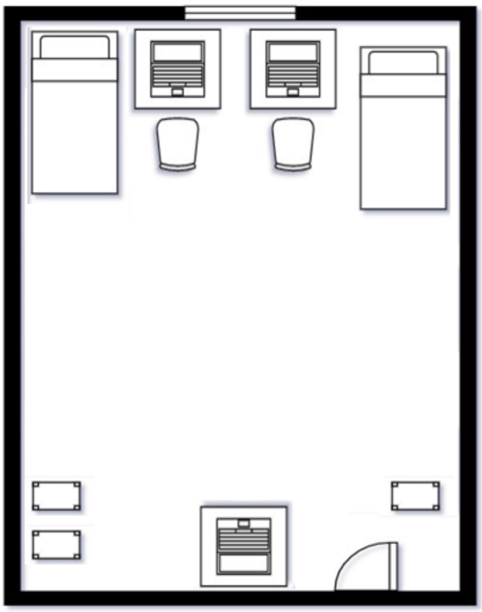

Figure 15 Room layout 85B

7.2 Legal and organizational issues of occupational safety

Today, one of the main approaches to radically improve all prevention efforts in order to reduce the overall accident rate and occupational morbidity is the widespread implementation of an integrated occupational safety and health management system. This means integrating isolated activities into a single system with targeted actions at all levels and stages of the production process.

Occupational safety is a system of legislative, socio-economic, organizational, technological, hygienic and therapeutic and prophylactic measures and tools that ensure the safety, preservation of health and human performance in the work process.

The Labor Code of the Russian Federation stipulates that normal working hours may not exceed 40 hours per week, and employers must record the hours worked by each employee[1].

In order to prevent accidents and ensure safe working conditions for workers, rules on labor protection and safety measures were introduced, which are mandatory for workers, managers, engineers and technicians.

7.3 Basic ergonomic requirements for the correct location and arrangement of researcher's workplace

The workplace when working with a PC should be at least 6 square meters. The legroom should correspond to the following parameters: the legroom height is at least 600 mm, the seat distance to the lower edge of the working surface is at least 150 mm, and the seat height is 420 mm. It is worth noting that the height of the table should depend on the growth of the operator[2].

The following requirements are also provided for the organization of the workplace of the PC user: The design of the working chair should ensure the maintenance of a rational working posture while working on the PC and allow the posture to be changed in order to reduce the static tension of the neck and shoulder muscles and back to prevent the development of fatigue.

In accordance with the GOST 12.2.032-78 SSBT[3], general ergonomic requirements for seated workstations in the new and modern design of existing equipment and production processes are specified. Every employee has the right:

A workplace sitting in work is organized for light work that does not require workers' free movements, as well as moderate work due to specialty of technical processes;

The relative position of the workplace design and all the elements (the way the seat, control device, the display information, etc.) must in line with human measurement, physiological and psychological requirements and working nature;

The workstation must be organized according to the requirements of the Motivation Security, Specification and / or Guide;

Women should be considered when designing equipment and organizing workplace, if only women work) and male (if only men's work); if the equipment is repaired by women and male - women and men's overall average;

The design of production equipment and workplace should ensure the best position of the workers, which is achieved by adjustment;

In accordance with the studio shape of each device according to the nature of the work being performed. It can be that, there is a slit or desktop groove groove for working. If necessary, the armrest should be installed on the studio;

The height of the footrest must be terrible. The width must be at least 300 meters, and the length must be at least 400 meters.

The type of working chair should be selected taking into account the growth of the user, the nature and duration of work with the PC. The working chair should be lifting and swivel, adjustable in height and angle of inclination of the seat and back, as well as the distance of the back from the front edge of the seat, while the adjustment of each parameter should be independent, easy to carry out and have a secure fit[3].

7.4 Occupational safety

Workplace safety is the responsibility of everyone in the organization.

Occupational hygiene is a system of ensuring the health of workers in the process of labor activity, including legal, socio-economic, organizational and technical, sanitary and hygienic, treatment and prophylactic, rehabilitation and othermeasures.

Working conditions - a set of factors of the working environment and the laborprocess that affect human health and performance.

Harmful production factor is a factor of the environment and the work process that can cause occupational pathology, temporary or permanent decrease in working capacity, increase the frequency of somatic and infectious diseases, and lead to impaired health of the offspring.

Hazardous production factor is a factor of the environment and the labor process that can cause injury, acute illness or sudden sharp deterioration in health, death.

In this subsection it is necessary to analyze harmful and hazardous factors thatcan occur during research in the laboratory, when development or operation of the designed solution (on a workplace).

GOST 12.0.003-2015 " Hazardous and harmful production factors. Classification" must be used to identify potential factors, that can effect on a worker(employee).

Table 26 Potential hazardous and harmful production factors

Factors (**GOST 12.0.003-2015)**	Stages of work			Legislation documents
	developing	manufacturing	operation	

Increased levels of noise	+	+		**GOST 12.1.003-2014** Occupational safety standards system. Noise. General safety requirements
Lack or lack of natural light, insufficient illumination	+			**SanPiN 2.2.1/2.1.1.1278-03** Hygienic requirements for natural, artificial and mixed lighting of residential and public buildings
Electromagnetic fields	+	+	+	**SanPiN 2.2.4.1329-03** Requirements for protection of personnel from the impact of impulse. electromagnetic fields.
Abnormally high voltage value in the circuit, the closure which may occur through the human body		+	+	**Sanitary rules GOST 12.1.038-82 SSBT** Electrical safety. Maximum permissible levels of touch voltages and currents.

7.4.1 Increased levels of noise

Noise worsens working conditions; have a harmful effect on the human body, namely, the organs of hearing and the whole body through the central nervous system. It results in weakened attention, deteriorated memory, decreased response, and increased number of errors in work.

Noise can be generated by operating equipment, air conditioning units, daylight illuminating devices, as well as spread from the outside.

When working on a PC, the noise level in the workplace should not exceed 50 dB[4]. In order to study in a quiet environment, irrelevant applications of the computer should be closed to reduce computer power consumption, thereby reducing computer noise, and windows should also be closed to reduce environmental noise.

7.4.2 Lack or lack of natural light, insufficient illumination

Light sources can be both natural and artificial. The natural source of the light in the room is the sun, artificial light are lamps. With long work in low illumination conditions and in violation of other parameters of the illumination, visual perception decreases, myopia, eye disease develops, and headaches appear[5].

According to the SanPiN 2.2.1/2.1.1.1278-03 standard[4], the illumination on the table surface in the area of the working document should be 300-500 lux. Lighting should not create glare on the surface of the monitor. Illumination of the monitor surface should not be more than 300 lux.

The brightness of the lamps of common light in the area with radiation angles from 50 to 90° should be no more than 200 cd/m, the protective angle of the lamps should be at least 40°. The ripple coefficient should not exceed 5%.

7.4.3 Electromagnetic fields

In this case, the sources of increased intensity of the electromagnetic field computer. 8 kA / m is considered acceptable [6]. An hour's working day for an employee at his workplace, with the maximum permissible level of tension, should be no more than 8 kA / m, and the level of magnetic induction should be 10 mT. Compliance with these standards makes it possible to avoid the negative effects of electromagnetic radiation.

To reduce the level of the electromagnetic field from personal it is recommended to connect no more than two computers to one outlet, make a protective grounding, connect the computer to the outlet through an electric field neutralizer.

Sources of electromagnetic radiation in the workplace are system units and monitors of switched-on computers. To bring down exposure to such types of radiation, it is recommended to use such monitors, the radiation level is reduced, as well as to install protective screens and observe work and rest regimes.

According to the intensity of the electromagnetic field at a distance of 50 cm around the screen along the electrical component should be no more than[6]:

- in the frequency range 5 Hz - 2 kHz - 25 V / m;
- in the frequency range 2 kHz - 400 kHz - 2.5 V/m.

The magnetic flux density should be no more than:

- in the frequency range 5 Hz - 2 kHz - 250 nT;
- in the frequency range 2 kHz - 400 kHz - 25 nT.

There are the following ways to protect against EMF:

- increase the distance from the source (the screen should be at least 50 cm from the user);
- the use of pre-screen filters, special screens and other personal protective equipment.

When using a computer, the source of EM is a monitor. Under the influence of EM in the body, there may be a violation of normal blood coagulability, an increase in the fragility of blood vessels, a decrease in immunity, etc. The dose of irradiation at a distance of 20 cm to the display is 50 μrem/hr. According to the norms[7], the design of the computer should provide the power of the exposure dose of x–rays at any point at a distance of 0,05 m from the screen no more than 100 μR/h.

7.4.4 Abnormally high voltage value in the circuit

The mechanical action of current on the body is the cause of electrical injuries. Typical types of electric injuries are burns, electric signs, skin metallization, tissue tears, dislocations of joints and bone fractures.

The following protective equipment can be used as measures to ensure the safety of working with electrical equipment:

- disconnection of voltage from live parts, on which or near to which work will be carried out, and taking measures to ensure the impossibility of applying voltage to the workplace;
- posting of posters indicating the place of work;
- electrical grounding of the housings of all installations through a neutral wire;
- coating of metal surfaces of tools with reliable insulation;

inaccessibility of current-carrying parts of equipment (the conclusion in the case of electroporation elements, the conclusion in the body of current carrying parts)[8].

7.5 Ecological safety

Presently section discusses the environmental impacts of the project development activities, as well as the product itself as a result of its implementation in production. The software product itself, developed during the implementation of the master's thesis, does not harm the environment either at the stages of its development or at the stages of operation. However, the funds required to develop and operate it can harm the environment.

There is no production in the laboratory. The waste produced in the premises, first of all, can be attributed to paper waste - waste paper, plastic waste, defective parts of personal computers and other types of computers. Waste paper is recommended accumulate and transfer them to waste paper collection points for further processing. Place plastic bottles in specially designed containers.

Modern PCs are produced practically without the use of harmful substances hazardous to humans and the environment. Exceptions are batteries for computers and mobile devices. Batteries contain heavy metals, acids and alkalis that can harm the environment by entering the hydrosphere and lithosphere if not properly disposed of. For battery disposal it is necessary to contact special organizations specialized in the reception, disposal and recycling of batteries[9].

Fluorescent lamps used for artificial illumination of workplaces also require special disposal, because they contain from 10 to 70 mg of mercury, which is an extremely dangerous chemical substance and can cause poisoning of living beings, and pollution of the atmosphere, hydrosphere and lithosphere. The service life of such lamps is about 5 years, after which they must be handed over for recycling at special reception points. Legal entities are required to hand over lamps for recycling and maintain a passport for this type of waste. An additional method to reduce waste is to increase the share of electronic document management[9].

7.6 Safety in emergency

An emergency situation (ES) is a situation in a certain territory that has developed as a result of an accident, hazardous natural phenomenon, catastrophe or other disaster, which may entail human casualties, damage to human health or the environment, significant material losses and violation of the living conditions of people. Emergency for the presented work space is a fire. This emergency can occur in the event of non-compliance with fire safety measures, violation of the technique of using electrical devices and PCs, violations of the wiring of electrical networks and a number of other reasons.

A fire in the working environment of the PC operator can be caused by non-electrical and electrical causes. Causes of a non-electrical nature include negligent and careless handling of fire (smoking, leaving heating devices unattended)[10]. Electrical causes include:

- short circuit.

- dangerous overload of networks, which leads to strong heating of live parts and ignition of insulation.
- start-up of equipment after incorrect and unqualified repairs.

To prevent emergencies, it is necessary to comply with fire safety rules in order to ensure the state of protection of employees and property from fire

To protect against short circuits and overloads, it is necessary to correctly select, install and use electrical networks and automation equipment.

To prevent the occurrence of fires, it is necessary to exclude the formation of a combustible environment, to monitor the use of non-combustible or hardly combustible materials in the construction and decoration of buildings.

It is necessary to carry out the following fire prevention measures:

- organizational measures related to the technical process, taking into account the fire safety of the facility (personnel briefing, training in safety rules, publication of instructions, posters, evacuation plans).
- operational measures that consider the operation of the equipment used (compliance with equipment operating standards, ensuring a free approach to equipment, maintaining conductor insulation in good condition).
- technical and constructive measures related to the correct placement and installation of electrical equipment and heating devices (compliance with fire safety measures when installing electrical wiring, equipment, heating, ventilation and lighting systems).

To increase the resistance of the working room to emergencies, it is necessary to install fire alarm systems that react to smoke and other combustion products, install fire extinguishers. Also, two times a year to conduct drills to practice actions in case of fire.

An evacuation plan is presented in the presented working room at the entrance, a fire alarm system is installed. The room is equipped with OU-2 type carbon dioxide fire extinguishers in the amount of 2 pieces per one working area. There is an electrical panel within the reach of workers, with the help of which it is possible to completely de-energize the working room.

In the event of a fire, you must call the fire department by phone 101 and inform the place of the emergency, take measures to evacuate workers in accordance with the evacuation plan. In the absence of direct threats to health and life, make an attempt to extinguish the resulting fire with existing carbon dioxide fire extinguishers. In case of loss of control over the fire, it is necessary to evacuate after the employees according to the evacuation plan and wait for the arrival of the fire service specialists.

7.7 Conclusion

Each employee must carry out professional activities with taking into account social, legal, environmental and cultural aspects, issues health and safety, be socially responsible for the solutions, be aware of the need for sustainable development.

In presently section covered the main issues of observance of rights employee to work, compliance with the rules for labor safety, industrial safety, ecology and resource conservation.

It was found that the researcher's workplace satisfies safety and health requirements during project implementation, and the harmful impact of the research object on the environment is not exceeds the norm.

Reference for social responsibility

1. Federal Law "On the Fundamentals of Labor Protection in the Russian Federation" of 17.07.99 № 181 – FZ.
2. SanPiN 2.2.2/2.4.1340-03 Hygienic requirements for personal electronic computers and organization of work.
3. GOST 12.2.032-78 SSBT. Workplace when performing work while sitting. General ergonomic requirements.
4. GOST 12.1.003-2014 SSBT. Noise. General safety requirements.
5. SanPiN 2.2.1 / 2.1.1.1278-03. Hygienic requirements for natural, artificial and combined lighting of residential and public buildings.
6. SanPiN 2.2.4.1329-03 Requirements for protection of personnel from the impact of impulse electromagnetic fields.
7. GOST R ISO 1410-2010. Environmental management. Assessment of life Cycle. Principles and structure.
8. GOST 12.1.038-82 Occupational safety standards system. Electrical safety.
9. GOST R 52105-2003 Resources saving. Waste treatment.
10. GOST 12.2.003-91 Occupational safety standards system. Industrial equipment. General safety requirements.

Reference

1. Zhang S H, Yang D , Liu Y W. An improved edge detection algorithm based on Canny operator[J]. Computer Engineering and Applications, 2004, 39(29):113-115.
2. Qin Xu-jia, Jiang Jiong-hui, Wang Wei-hong, et al. Canny operator based level set segmentation algorithm for medical images[C]: ICBBE 2007: The 1st International Conference on Bioinformatics and Biomedical Engineering, Wuhan, July 6-8, 2007. New Jersey: IEEE Xplore digital library, 2007.
3. Mera D, Cotos J M, Varela-Pet J, et al. Adaptive thresholding algorithm based on SAR images and wind data to segment oil spills along the northwest coast of the Iberian Peninsula[J]. Marine Pollution Bulletin, 2012, 64(10):2090-2096.
4. Zhang Y J, A survey on Evaluation methods for image segmentation, Pattern Recognition, 1996, 29(8):1335-1346.
5. Luo H G ,Zhu L M ,Ding H. Active contour model and level set method for image segmentation technique[J]. China Journal of Graphic Imaging, 2006,11(3): 301-309.
6. Mharib A M, Ramli A R, Mashohor S, et al. Survey on liver CT image segmentation methods[J]. Artificial Intelligence Review, 2012, 37(2):83-95.
7. Li W, Jia F, Hu Q. Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks[J]. Journal of Computer & Communications, 2015, 03(11):146-151.
8. Udupa J K, Samarasekera S. Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation[J]. Graphical models and image processing, 1996, 58(3): 246-261.
9. Unser M. Multigrid adaptive image processing[C]. International Conference on Image Processing. IEEE Computer Society, 1995: 49.
10. Cheewatanon J, Leauhatong T, Airpaibon S, et al. A New White Blood Cell Segmentation Using Mean Shift Filter and Region Growing Algorithm[J]．2011,4:30-35.
11. Borst H, Abmayr W, Gais P.A Thresholding Method for Automatic Cell Image

Segmentation.[J].Journal of Histochemistry & Cytochemistry Official Journal of the Histochemistry Society, 1979,27(1):180-7.

12. Zafari S.Segmentation of Overlapping Convex Objects[J]. Lutpub, 2014.
13. Ahmed M N, Farag A A. Two-stage neural network for volume segmentation of medical images[C]. International Conference on Neural Networks. IEEE, 1997:1373-1378 vol.3.
14. Wang Z, Vemuri B C. DTI segmentation using an information theoretic tensor dissimilarity measure[J]. IEEE Transactions on Medical Imaging, 2005, 24(10):1267.
15. Deng W, Iyengar S S. A new probabilistic relaxation scheme and its application to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(4): 432-437.
16. Tian J, Bao S L, Zhou M Q. Medical image processing and analysis [M]. Beijing: Electronic Industry Press, 2003: 34-51.
17. Worth A J, Makris N. Semeautomatic segmentation of brain exterior in magnetic resonance images driven by empirical procedures and anatomical knowledge[J]. Medical Image Analysis, 1998, 2(4):315-324.
18. Sluimer I, Schilham A, Pro Kop M, et al. Computer analysis of computed tomography scans of the lung[J]. IEEE transactions on medical imaging, 2006, 25(4): 385-405.
19. Bieniecki W, Grabowski S. Nearest neighbor classifiers for color image segmentation[C]. Modern Problems of Radio Engineering, Telecommunications and Computer Science, 2004. Proceedings of the International Conference. IEEE, 2004:209-212.
20. Boiman O, Shechtman E, Irani M. In defense of Nearest-Neighbor based image classification[C]. Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008:1-8.
21. Coupé P, Manjón J V, Fonov V, et al. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation[J]. Neuroimage, 2011, 54(2):940.

22. Rousseau F, Habas P A, Studholme C. A supervised patch-based approach for human brain labeling.[J]. IEEE Transactions on Medical Imaging, 2011, 30(10):1852-62.
23. Konukoglu E, Glocker B, Zikic D, et al. Neighbourhood approximation using randomized forests.[J]. Medical Image Analysis, 2013, 17(7):790-804.
24. Wazarkar S, Keshavamurthy B N, Hussain A. Region-based Segmentation of Social Images Using Soft KNN Algorithm[J]. Procedia Computer Science, 2018, 125:93-98.
25. Yadav S, Patra S, Arora C, et al. Deep CNN with color lines model for unmarked road segmentation[C]. IEEE International Conference on Image Processing. IEEE, 2017:585-589.
26. Liu Z, Li X, Luo P, et al. Deep Learning Markov Random Field for Semantic Segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2016, PP(99):1-1.
27. Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 40(4):834-848.
28. Mishra R, Daescu O. Deep learning for skin lesion segmentation[C]. IEEE International Conference on Bioinformatics and Biomedicine. IEEE, 2017:1189-1194.
29. Rumelhart, Geoffrey Hinton, Ronald Williams.Learning representations by back-propagating errors,1986.
30. M. Arbib, "Review of 'Perceptrons: An Introduction to Computational Geometry' (Minsky, M., and Papert, S.; 1969)," IEEE Transactions on Inf ormation Theory, vol. 15, pp. 738-739, 1969.
31. G. Hinton, S. Osindero, M. Welling, and Y. W. Teh, "Unsupervised discovery of nonlinear structure using contrastive backpropagation," Cognitive Science, vol. 30, p. 725, 2006.
32. D. Hubel, T. Wiesel. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat’s Visual Cortex. The Journal of Physiology, 160, 106-

154, 1962.

33. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
34. Krizhevsky A, Sutskever I, Hinton G E. Image Net classification with deep convolutional neural networks[C].International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
35. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[J]. 2015:770-778.
36. A. Giusti, D. C. Ciresan, J. Masci, L. M. Gambardella, J. Schmidhuber. Fast Image Scanning with Deep Max-Pooling Convolutional Neural Networks. ICIP, 2013.
37. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, ICCV, 2015.
38. Huang Y J, Dou Q, Wang Z X, et al., 3d roi-aware u-net for accurate and efficient colorectal tumor segmentation. 2018.
39. Arif, S M M R A, Knapp K, and Slabaugh G, Shape-Aware Deep Convolutional Neural Network for Vertebrae Segmentation. 2017.
40. Lin T Y, Goyal P, Girshick R, et al., Focal Loss for Dense Object Detection. 2017. PP(99): p. 2999-3007.
41. Milletari F, Navab N and Ahmadi S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. in 2016 Fourth International Conference on 3D Vision (3DV). 2016. IEEE.
42. Cai Z, Fan Q, Feris R S, et al., A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection.
43. Kervadec H, Bouchtiba J, Desrosiers C, et al., Boundary loss for highly unbalanced segmentation.

Appendix A Program code based on U-Net network model

Initial settings

```
im_height = 512
im_width = 512
batch_size = 32
epochs = 10
if not os.path.exists("save_weights"):
os.makedirs("save_weights")
image_path = "../input/ Rectal_cancer_picture/"
train_dir = image_path + "train"
validation_dir = image_path + "test"
test_dir = image_path + "valid"
```

Data pre-processing

```
train_image_generator =
ImageDataGenerator( rescale=1./255,shear_range=0.2,z
oom_range=0.2,
horizontal_flip=True)
validation_image_generator =
ImageDataGenerator(rescale=1./255)
test_image_generator =
ImageDataGenerator(rescale=1./255)
```

Generate data

```
train_data_gen =
train_image_generator.flow_from_directory(directory=t
rain_dir,
```

```
batch_size=batch_size,
shuffle=True,
target_size=(im_height, im_width),
class_mode='categorical')
total_train = train_data_gen.n
val_data_gen =
validation_image_generator.flow_from_directory(direct
ory=validation_dir,
batch_size=batch_size,
shuffle=False,
target_size=(im_height, im_width),
class_mode='categorical')
total_val = val_data_gen.n
test_data_gen =
test_image_generator.flow_from_directory( directory=t
est_dir,
batch_size=batch_size,
shuffle=False,
target_size=(im_height, im_width)
class_mode='categorical')
total_test = test_data_gen.n
```

Result

Found 2423 images belonging to 2 classes.

Found 303 images belonging to 2 classes.

Found 303 images belonging to 2 classes.

Building the model

```
def build_unet(self, n_filters=16, dropout=0,
batchnorm=True, padding='same'):
def conv2d_block(input_tensor, n_filters=16,
kernel_size=3, batchnorm=True, padding='same'):
  # the first layer
  x = Conv2D(n_filters, kernel_size,
padding=padding)(input_tensor)
if batchnorm:
x = BatchNormalization()(x)
x = Activation('relu')(x)
# the second layer
x = Conv2D(n_filters, kernel_size, padding=padding)(x)
if batchnorm:
x = BatchNormalization()(x)
X = Activation('relu')(x)
return X
img = Input(shape=self.shape)
# contracting path
c1 = conv2d_block(img, n_filters=n_filters * 1,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
p1 = MaxPooling2D((2, 2))(c1)
p1 = Dropout(dropout * 0.5)(p1)
c2 = conv2d_block(p1, n_filters=n_filters * 2,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
p2 = MaxPooling2D((2, 2))(c2)
```

```
p2 = Dropout(dropout)(p2)
c3 = conv2d_block(p2, n_filters=n_filters * 4,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
p3 = MaxPooling2D((2, 2))(c3)
p3 = Dropout(dropout)(p3)
c4 = conv2d_block(p3, n_filters=n_filters * 8,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
p4 = MaxPooling2D((2, 2))(c4)
p4 = Dropout(dropout)(p4)
c5 = conv2d_block(p4, n_filters=n_filters * 16,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
# extending path
u6 = Conv2DTranspose(n_filters * 8, (3, 3), strides=(2,
2), padding='same')(c5)
u6 = concatenate([u6, c4])
u6 = Dropout(dropout)(u6)
c6 = conv2d_block(u6, n_filters=n_filters * 8,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
u7 = Conv2DTranspose(n_filters * 4, (3, 3), strides=(2,
2), padding='same')(c6)
u7 = concatenate([u7, c3])
u7 = Dropout(dropout)(u7)
c7 = conv2d_block(u7, n_filters=n_filters * 4,
```

```
kernel_size=3, batchnorm=batchnorm,
padding=padding)
u8 = Conv2DTranspose(n_filters * 2, (3, 3), strides=(2,
2), padding='same')(c7)
u8 = concatenate([u8, c2])
u8 = Dropout(dropout)(u8)
c8 = conv2d_block(u8, n_filters=n_filters * 2,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
u9 = Conv2DTranspose(n_filters * 1, (3, 3), strides=(2,
2), padding='same')(c8)
u9 = concatenate([u9, c1])
u9 = Dropout(dropout)(u9)
c9 = conv2d_block(u9, n_filters=n_filters * 1,
kernel_size=3, batchnorm=batchnorm,
padding=padding)
output = Conv2D(1, (1, 1), activation='sigmoid')(c9)
return Model(img, output)
self.unet = self.build_unet()
self.unet.summary()
```

Compilation Network

```
optimizer = Adam(0.0002, 0.5)
self.unet.compile(loss='mse', optimizer=optimizer,
metrics=['accuracy'])
def metric_fun(self, y_true, y_pred):
fz = tf.reduce_sum(2 * y_true * tf.cast(tf.greater(y_pred,
```

```
0.1), tf.float32)) + 1e-8
fm = tf.reduce_sum(y_true + tf.cast(tf.greater(y_pred,
0.1), tf.float32)) + 1e-8
return fz / fm
self.unet.compile(loss='mse',
optimizer=optimizer,
metrics=[self.metric_fun])
```

Training Network

```
x_train, x_label, y_train, y_label = self.load_data()
callbacks = [EarlyStopping(patience=100, verbose=2),
ReduceLROnPlateau(factor=0.1, patience=20,
min_lr=0.0001, verbose=2),
ModelCheckpoint('./weights/best_model.h5', verbose=2,
save_best_only=True)
# training
results = self.unet.fit(x_train, x_label, batch_size=32,
epochs=200, verbose=2,
callbacks=callbacks, validation_split=0.1, shuffle=True)
```

Precision

```
mask = self.unet.predict(x_train[index:index +
batch_size]) > 0.1
mask_true = x_label[index, :, :, 0]
if (np.sum(mask) > 0) == (np.sum(mask_true) > 0):
  n += 1
```