Journal on Information Technologies & Communications

MEDICAL REPORT

A Survey on Medical Image and Segmentation

Le Nhi Lam Thuy, Vu Ngoc Thanh Sang, Huynh Trung Hieu, Pham The Bao

Correspondence should be addressed to Pham The Bao: ptbao@sgu.edu.vn

Guiding tutor: _____
Chữ ký của GVHD

Subject:

Falculty: Information Science Faculty,
(Sai Gon University, Vietnam)
Information Technology Faculty,
(Industrial University of Ho Chi Minh City, Vietnam)

HCMC, 6/2022

Table of contents

Contents

Thanks

TBA

Abstract

Medical image processing and analysis are critical for assisting physicians with non-invasive clinical diagnoses. Understanding medical imaging techniques enables you to tackle corresponding medical image processing methodologically. Each imaging technique has distinct properties and produces images of various body regions. Medical image segmentation is crucial to improving treatment accuracy and assisting doctors' diagnosis conclusion. This article discusses medical imaging techniques and approaches to medical image segmentation.

Keywords: Medical images, medical image processing, medical image segmentation.

List of figures

List of tables

I. INTRODUCTION

Medical images are created using specialized techniques and processes to depict information about the human body (or specific parts of the body) for various clinical purposes, including medical procedures, diagnostics, and medical science. It is a subset of bioimaging and addresses radiology, endoscopy, thermography, medical imaging, and microscopy in a larger sense. For example, recording techniques such as electroencephalography (EEG) and magnetoencephalography (MEG) are designed to measure functional and effective brain connectivity. Applied studies and interpretation of medical images frequently focus on assessing pathology in the medical sciences (neuroscience, cardiology, psychiatry, and psychology) using comparable medical images. Numerous medical imaging systems have been created for scientific purposes and recently expanded to industry. Although mathematics was widely utilized for image processing, it had little impact on the biomedical area until the introduction of computed tomography (CT) for radiology, resulting in computer-aided tomography (CAT) procedures. Subsequently, isotope-emission tomography was created, allowing the development of positron emission tomography (PET) scan and single-photon emission tomography (SPECT) scan. Magnetic Resonance Imaging (MRI) has grown in popularity in recent decades, outpacing other imaging techniques [1].

Numerous studies have been researched for ultrasound, EEG applications, and novel optical imaging techniques, apart from these well-established approaches. While medical images are similar in appearance, the technology utilized to create them, and their parameters are relatively different. This distinction is described in their features and usefulness in medicine. Numerous approaches have been developed to enable the physician to generate 3D images from CT, MRI, and ultrasound scanning software. Typically, a single CT or MRI scan results in a two-dimensional image on film. Repetition of the preceding steps results in creating a three-dimensional image [2].

Medical imaging is a subset of biological imaging, which has been developed since the 19th century. The following is a summary of medical imaging:

- In 1895, Roentgen accidentally discovered X-rays [3]. Conventional radiography is by far the most common medical imaging technique in science. Radionuclides have been employed for therapeutic and metabolic tracer research rather than imaging since 1896, leading to the invention of the linear γ-ray image scanner.
- Sonar technology was employed during World War II, and in 1970, ultrasonography became widely available in medicine. A brief history of ultrasound is mentioned explicitly in another document [4].
- In the twentieth century, mathematical expertise for tomographic reconstruction was established, as well as positron emission tomography (PET) and computed tomography (CT) techniques [5]. In magnetic resonance imaging, nuclear magnetic resonance has been employed for imaging (MRI).
- New types of X-rays, MRI, and ultrasound images are being produced in the twenty-first century, particularly those that comprise microscopy and

macroscopic biological structures (thermal imaging, electrical impedance tomography, and probe scanning technique).

Medical images assist doctors in gaining additional information about a patient to make a more accurate diagnosis. Physicians appraise their patients based on their skills and experience. However, as the number of patients increases, the examination puts the doctor under increased pressure, and errors might arise. As a result, medical image segmentation enables physicians to save time on manual image analysis while minimizing potential errors. Physicians can save time when they employ scientific and technical solutions to patients for other medical duties. Medical images range from the simplest, such as a chest X-ray, to the most complicated, such as functional magnetic resonance imaging (fMRI). We have witnessed the phenomenal growth of new, powerful technology for detecting, storing, transmitting, analyzing, and displaying digital medical images over the last several decades. These techniques enable quantitative measurements to be made by biochemists, biologists, medical scientists, and physicists, thereby facilitating the validation of scientific hypotheses and medical diagnoses. The gathering of functional and metabolic data and structural data from medical images will be enhanced in the future. This approach can be accomplished using photon emission tomography (PET) and magnetic resonance spectroscopy [6]. Although computer based medical image processing has been developed remarkably, it is challenging for computers to replace doctors completely. Therefore, the current solution is to develop technology to assist doctors in diagnosing and assessing patients' health. More precisely, medical image segmentation enables physicians to concentrate on areas of interest when making decisions. Additionally, image segmentation ensures the path for future technologies and increases the valuable content of diagnostic images.

II. MEDICAL IMAGE PROCESSING MODEL

Medical image analysis and interpretation are simplified using a three-level processing architecture, as seen in Figure 1. The system's first stage is image generation. Numerous imaging modes were discussed previously, including X-rays and SPECT. The image processing process is divided into low-level and high-level processing stages. Filtering, image augmentation, segmentation, and feature extraction are essential for further analysis. At this level of raw pixels processing, the most critical tasks are tumor identification and disease diagnosis.

The following five steps describe the fundamental procedures of image processing:

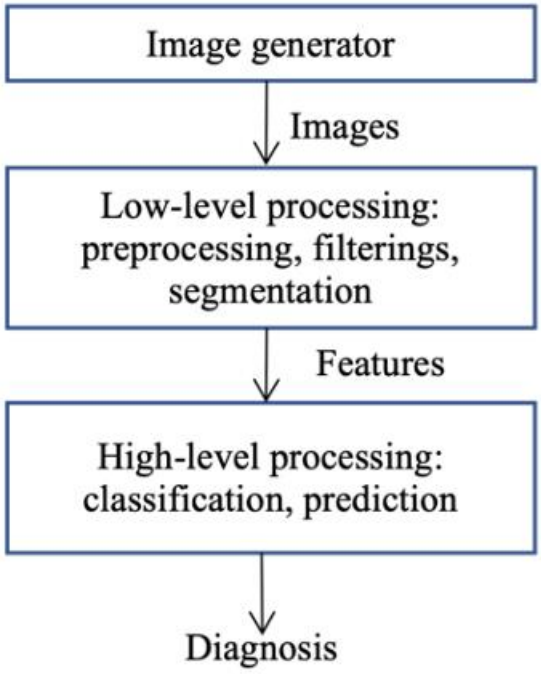

Figure 1. A model of medical imaging diagnostic system *[7]*.

(i) Preprocessing aids in the visualization of low-resolution object boundaries. The most often used techniques are image registration, image transformation, filtering, and Laplace operators.

(ii) Filtering: Filtering includes enhancement, sharpening, and edge detection. Enhancement techniques include linear or nonlinear, local, or global, or wavelet-based filters. Typical sharpening techniques are inversely or Weiner filters. Edge detection techniques include Haar transform, local operators, prediction, and/or classification methods.

(iii) Segmentation: Segmentation can be both region-based and boundary-based. There are different types of segmentation algorithms, including region-based classical algorithms, clustering algorithms, and line and arc detectors. An important issue in medical imaging is that it is possible to perform segmentation for various domains using bottom-up generalization methods without requiring special domain knowledge.

(iv) Shape modeling: Shape modeling is carried out using characteristics that can be employed independently or in conjunction with dimensional data. It is occasionally necessary to characterize the shape of an object in greater detail than is offered by a single but condensed characteristic reflected in the object image. In these cases, the shape descriptor expresses an object's shape in a more concise manner.

(v) Classification: Classification is based on feature selection, texture features, and decisions regarding feature classes. Each abnormality or disease is recognized as belonging to a specific group, and recognition is performed as a classification process.

III. MEDICAL IMAGING TECHNIQUES

Most medical imaging techniques are non-invasive. There are four types of medical images made from: (1) X-rays, (2) γ-rays, (3) ultrasonic echoes, (4) induced magnetic resonance (MR), and (5) microscopic images. Normally images of types (3) and (4) are combined as described in Table I

TABLE I - MEDICAL IMAGE APPLICATION

Techniques	Inspected body parts/organs
X-ray	Chest, lungs, bones
γ-ray	Brain, internal organs, cardiac function
MR	Soft tissue, brain, fetus, pathological changes, internal organs
Microscopic images	Blood cells, bacteria

The conventional way of creating medical images is depicted in Figure 2. Ionizing radiation is seen in Figure 2.a and 2.b. Projection radiography and tomography are based on the passage of X-rays through the body and the selective attenuation of these rays by the body's tissues to generate images. The magnetic resonance effect, depicted in Figure 2.c, is caused by the nuclear magnetic resonance features. This indicates that protons gravitate toward the magnetic field. Selective stimulation of areas within the body can cause these protons to rotate in the opposite direction of the magnetic field, as illustrated in Figure 2.d, an ultrasound image created using high-frequency sound waves transmitted into the body and sound.

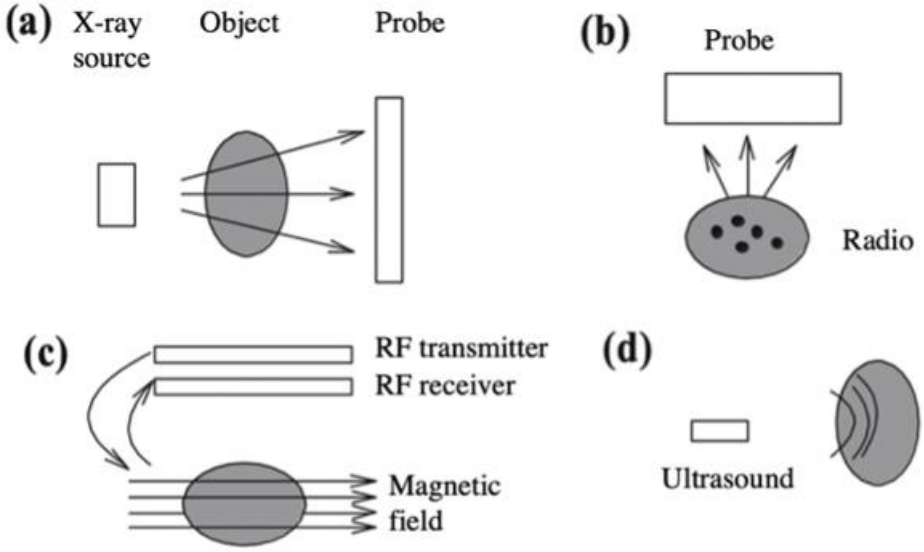

Figure 2. Medical imaging modalities: (a) X-ray images, (b) radionuclide images, (c) MR, and (d) ultrasound.

1. Ionizing radiation images

W.C. Roentgen discovered X-ray in 1895, and it is the most extensively used type of medical imaging. For medical imaging, X-rays are a type of ionizing radiation with an energy range of 25 keV to 500 keV. An X-ray tube in a traditional radiography system

produces a short pulse of X-rays that go through the human body. Non-absorbed or scattered X-ray photons reach the large area detector, which forms an image on the film, as illustrated in Figure 3. Energy loss as a function of the linear attenuation coefficient distribution in the body using a spatial model, Equation (1) captures this matter and energy dependent impact,

$$I_d = \int_0^{E_{max}} S_0(E) E e^{[-\int_0^d \mu(s;E)ds)]} dE \qquad (1)$$

where $S_0(E)$ is the X-ray spectrum and $\mu(s; E)$ is a function that describes the linear attenuation coefficient along the line between the source and the detector. S is the distance from the origin and d is the distance from the source to the detector. Image quality is affected by both noises resulting from the random nature of the X-rays or the transmission process as in Figure 3. Computed tomography (CT), established by Hounsfield in 1972, is another common imaging technique that eliminates the effects of tissue overlap that obstruct an accurate diagnosis. CT collects X-ray projections around

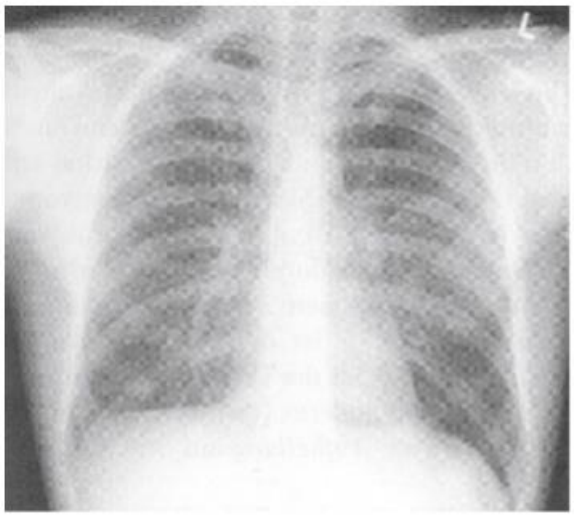

Figure 3. Chest X-ray image.

the patient, which can be seen as a series of X-rays obtained as the patient rotates gently around an axis. The films depict 2-D projections of the 3-D body from various viewpoints. A transverse line in the video represents a 1-D perspective of a 2-D axial cross-section of the body, represented by a collection of horizontal lines at the same height. The Radon transform [8] was used to reconstruct two-dimensional cross-sectional slices of the object using projection data. The Radon transform [8] is an integral transformation established by J. Radon in 1917. This transformation gathers 1-D projections of a 2-D object from various angles and reconstructs them using back propagation. The reconstruction is based on the Fourier slice theorem, which claims that a projection's 1- D Fourier transform is a slice of the object's 2-D Fourier transform as presented in Figure 4.

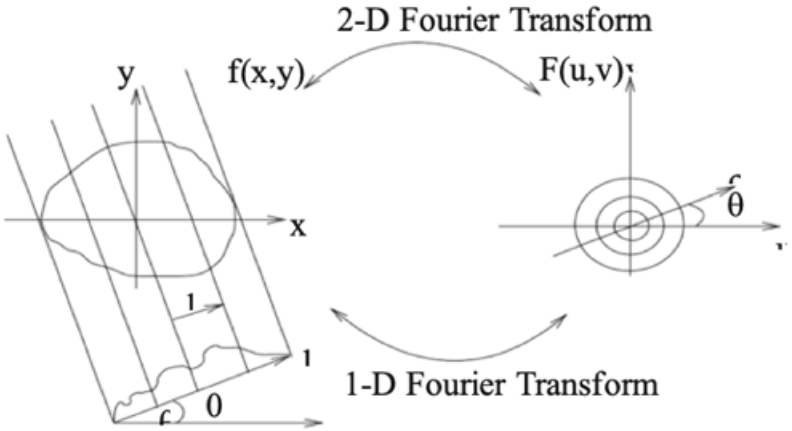

Figure 4. Visualize for projection-slice theorem.

The basic imaging equation is similar to traditional radiography, except for a set of projections employed in cross-sectional image reconstruction as Equation (2).

$$I_d = I_0 e^{[-\int_0^d \mu(s;\bar{E})ds)]} dE \qquad (2)$$

where, I_0 is the standard intensity, and $\bar{}(E)$ is the effective energy is E. CT has several advantages over projection radiography: (1) it eliminates the superimposition of images of structures outside the area of interest, (2) it provides high contrast resolution, allowing differences to be observed, and (3) as a tomographic and potentially three-dimensional approach for analyzing isolated cross-sectional visual slices of the body, to distinguish between tissues with less than 1% physical density. CT scans are a valuable complement to medical imaging since they provide more information than X-rays or ultrasound.

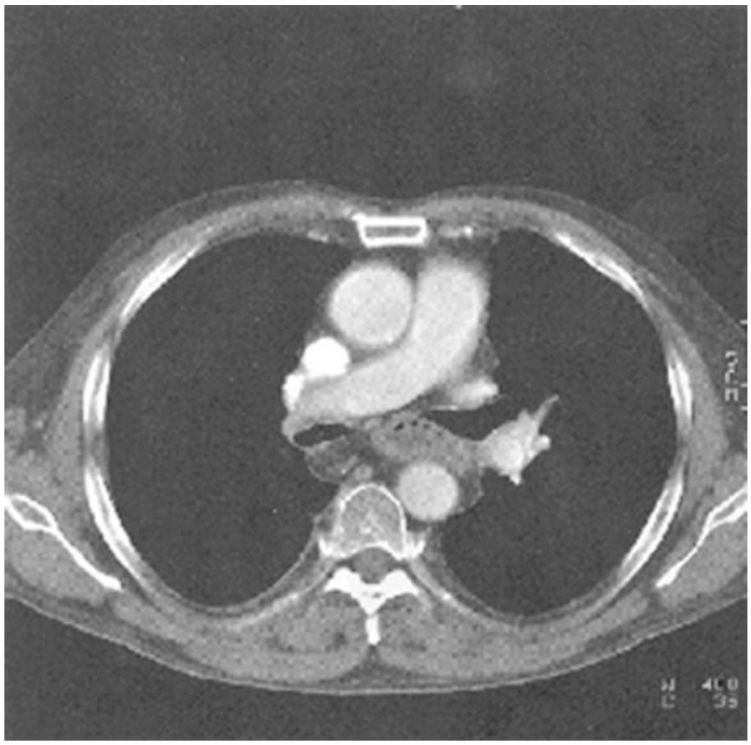

Figure 5. Visualization of the projection slice theorem.

Figure 5 shows how CT is used to diagnose cerebrovascular illnesses, acute and chronic alterations in lung parenchyma, and the comprehensive diagnosis of the abdominal and pelvic organs. Nuclear medicine was first practiced in the late 1930s, and several operations involved radioactive materials. The use of radioactive iodine for the treatment of thyroid illness is known as initiation. Nuclear medicine imaging, like X-ray imaging, has progressed from projection to tomographic imaging. Nuclear medicine uses ionizing radiation and imaging techniques similar to radiography, but it focuses on physiological function rather than anatomy. Radiotherapy equipment, on the other hand, is used in nuclear medicine to receive rays from a radioactive source that is injected into the body. Distinct radiations perform different tasks and provide different information in nuclear medicine. In other words, different radiography technologies can visualize a wide spectrum of physiological and metabolic activities. A camera (external imaging device) records the patient's radiation and converts it into flat, 2-D, or cross-sectional images. Nuclear medicine is a branch of medicine that deals with clinical diagnosis and treatment. It has many applications, including tumor diagnosis and treatment, acute care, cardiovascular, neurological, gastrointestinal, and renal problems. The three primary imaging modalities in nuclear medicine are divided into two main areas based on radiopharmaceutical decay: (1) planar imaging and single-photon emission tomography (SPECT), which uses gamma-emitting as a radiotherapy machine, and (2) positron emission tomography (PET), which uses positrons as an X-ray machine. An anger imaging camera creates a projection image, also known as a planar scan. Equation (3) – the basic image equation – has two crucial components: functioning as the desired parameter and attenuation as an undesired but essential addition,

$$\phi(x,y) = \int_{\infty}^{0} \frac{A(x,y,z)}{4\pi z^2} e^{\left[-\int_0^d \mu(x,y,z',E)dz'\right]} dz \qquad (3)$$

where A(x, y, z) represents the activity in the body and E is the energy of the photon. Image quality is largely determined by the resolution of the camera, and noise comes from system sensitivity, injected activity, and acquisition time.

On the other hand, SPECT collects detection data from several angles using an Anger camera. Single-photon emission is monitored with a gamma camera system and uses nuclei that decay by emitting a single photon. The acquired data is rebuilt using a image to form a 3-D dataset or thin (two-dimensional) slices in SPECT. This imaging technique can be thought of as a series of projections, each of which is a homogenous planar view. We must remember that if x and y are linear coordinates in the plane, then the plane line is expressed as Equation (4),

$$L(l,\theta) = \{(x,y)|x\cos\theta + y\sin\theta = l\} \qquad (4)$$

where l is the lateral position of the line segment and θ is the angle of the one unit normal to the line segment, Figure 6. With the parameterized coordinates according to Equation (5), the integral f(x, y) is calculated according to the Equation (6).

$$\begin{cases} x(s) = l\cos\theta - s\sin\theta \\ y(s) = l\sin\theta + s\cos\theta \end{cases} \tag{5}$$

$$g(l,\theta) = \int_{-\infty}^{\infty} f(x(s), y(s))ds \tag{6}$$

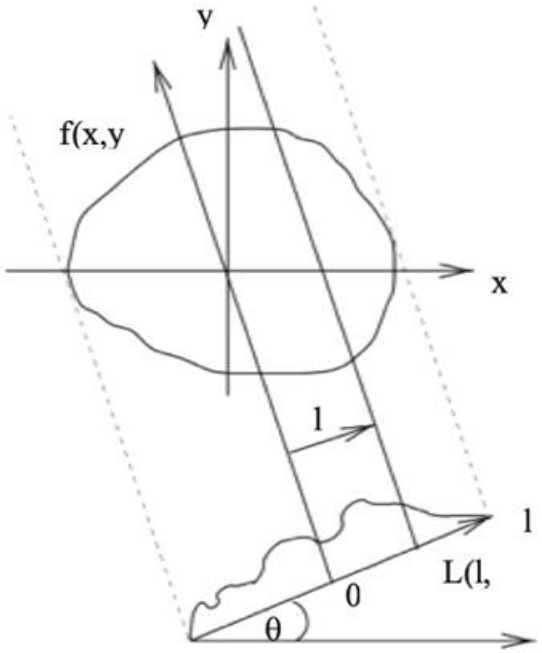

Figure 6. Geometric representation of lines and projections.

For a fixed angle θ, $g(l, \theta)$ is a projection, while for all l and θ, it is called a 2-D Radon transformation of $f(x, y)$. The image equation for SPECT is given as Equation (7), where $A(x(s), y(s))$ represents the inverse 2-D Radon transformation of (l, θ) and describes the radioactivity inside 3-D objects. As a result, there is no closed-form solution for attenuation correction in SPECT. SPECT is a significant imaging technology for obtaining functional images of organs because it provides accurate placement in 3-D space. Figure 7 shows a critical application in cardiac and brain function imaging.

$$\phi(l,\theta) = \int_{-\infty}^{\infty} A(x(s), y(s))ds \tag{7}$$

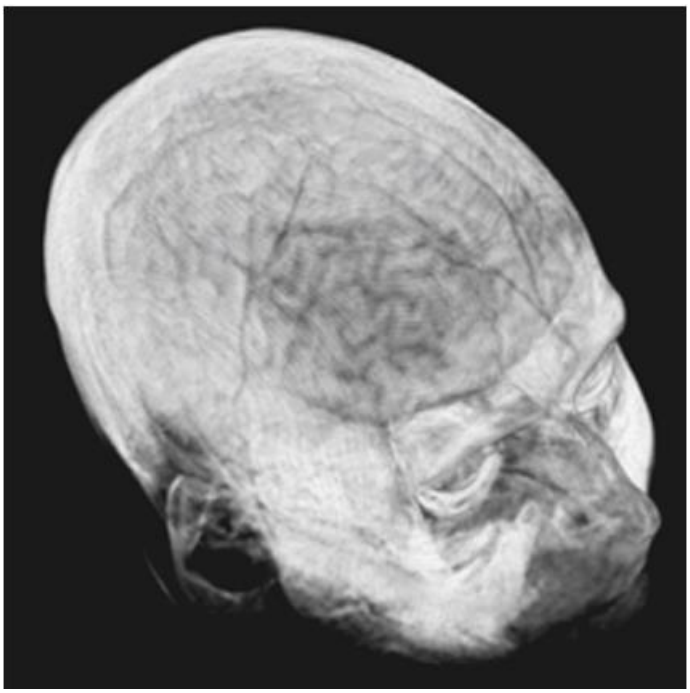

Figure 7. Brain imaging with SPECT.

PET is a special technique that has no similarities with other imaging modalities. PET radionuclides release positrons rather than x-rays. The positrons are the anti-electrons, and their locations are determined and measured. The reconstruction is built using the back-projection approach. PET imaging equations are identical to SPECT imaging equations with one exception: the integration limit for the attenuation component is extended over the entire body due to the simultaneous detection of paired γ-rays, referred to as the canceling photon as Equation (8),

$$\phi(l,\theta) = K\int_{-\infty}^{\infty} A(x(s),y(s))ds \qquad (8)$$

where K is a constant that accounts for the influence of constant factors such as detector area and efficiency ϕ. Both SPECT and PET have image quality limitations due to resolution, scattering, and noise. PET is widely used in oncology, neurology, and psychiatry. A significant part focuses on neurological illnesses such as Alzheimer's disease, dementia, and epilepsy.

2. Magnetic resonance images

Magnetic resonance imaging (MRI) is a non-invasive imaging technique that produces images of the interior of the body. Over the last three decades, this technique has developed into a critical bio imaging modality in medicine. Nuclear medicine and MR imaging both provide information on pathological and physiological changes in bodily tissues, as well as structural details about organs. The magnetic resonance signal is generated by the nuclear magnetism of hydrogen atoms found in the body's fat and water and is based on the basic basis of nuclear magnetic resonance (NMR). NMR is associated with the electric charge and angular moment that particular nuclei possess. The nucleus possesses a positive charge and, when the atomic or mass number is odd,

an angular momentum ϕ. Rotating these nuclei activates NMR. A tiny magnetic field should also be present around each revolving nucleus. When a magnetic field is applied externally, the direction of rotation tends to match the magnetic field. This is referred to as nuclear magnetism. Consider a particular rotating system (hydrogen atom) in a sample in the MR images. The term "sample" refers to a small piece of tissue – a voxel. The rotation system becomes magnetized when a static magnetic field B_0 is applied and can be described using a bulk magnetization vector M. M will attain equilibrium M_0 in the undisturbed state, but parallel to the direction of B_0, Figure 9a. The critical point is that $M(r, t)$ is a function of time and that external radio frequency excitations and magnetic fields can be used to modify the three-dimensional coordinate r in space. The value of an MR image at a specific voxel is determined by two critical factors: tissue characteristics and the imaging methodology used by the scanner. The T_1 and T_2 relaxation parameters and the proton density are the most important tissue features. The term "proton density" refers to the number of target nuclei per unit volume. The scanner's software and hardware manage the temporal and spatial M magnetization vectors based on the pulse sequence. The equation of motion of $M(t)$ with respect to time t is based on the Bloch equation and the resonant frequency ω_0 or the Larmor frequency. $M(t)$ is composed of two components:

- The vertical magnetization is given by $M_z(t)$, with the z component of $M(t)$.
- $M_{xy}(t)$ is a complex quantity composed of two orthogonal components, Equation (9). The complex number M_{xy}, referred to as the phase angle ϕ, which is determined using the Equation (10).

$$M_{xy}(t) = M_x(t) + jM_y(t) \tag{9}$$

$$\phi = tan^{-1}\left(\frac{M_x}{M_y}\right) \tag{10}$$

A moment is generated when an external time-varying magnetic field $B(t)$ is applied because $M(t)$ is a magnetic moment. $B(t) = B_0$ if the field is static and parallel to the z-axis. It rotates if the magnetization vector M is initially pointed away from B_0. RF signals can also be used to stimulate rotating systems. This RF stimulation is accomplished by applying B_1 at the Larmor frequency rather than maintaining it constantly, so tracking the position of $M(t)$. This value, however, is not permanent, and we shall demonstrate two distinct ways for reducing motion and obliterating the received signal: longitudinal and transverse stretch. $M(t)$ is pushed downward by an angle with respect to the plane due to RF stimulation with $B1$ is perpendicular to the y-axis. For α = 0, we have M_z + 0 and the magnetization vector rotates in the xy-plane with a frequency equal to the Larmor frequency. The pulse B_1 required for an angle $\alpha = \pi/2$ is called pulse 90. The magnetization vector returns to its equilibrium and the

dilation process is described by Equation (11) and will depend on the relaxation time according to vertically or in vertical rotation (T1), Figure 8.

$$M_z(t) = M_0 \left[1 - e^{-\frac{t}{T_1}}\right] \qquad (11)$$

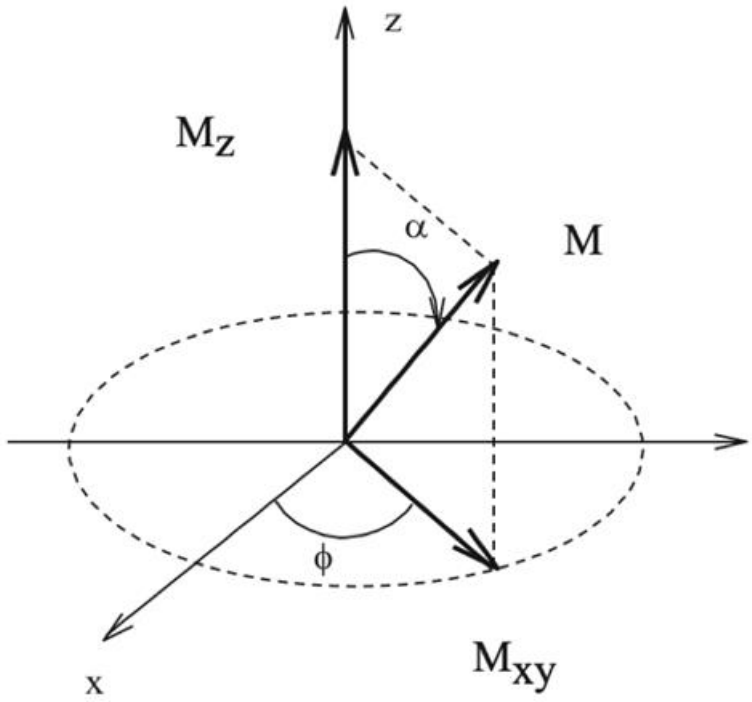

Figure 8. Magnetization vector M processed along the z-axis.

Horizontal stretching, also known as rotation-rotation, is the result of perturbations generated by other rotations when their phase is altered in relation to other rings. As a result of this attenuation, the reception antenna's signal is lost. The term "free inductive decay" refers to the received signal (FID). Equation describes the process of horizontal magnetization M_{xy} returning to equilibrium Equation (12),

$$M_{xy}(t) = M_{x_0 y_0} e^{-\frac{t}{T_1}} \qquad (12)$$

where T_2 is the rotational dilation time. T_2 is tissue-dependent and will produce contrast in the MR image. However, the received signal decays faster than T_2. Local noise in the static field B_0 gives rise to a faster time constant than T_2^* where $T_2^* < T2$, Figure 9b. The decay process involving external field effects is modeled by the time constant T_2'. The relationship between the three horizontal expansion constants is modeled by Equation (13).

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'} \qquad (13)$$

Notably, T_1 and T_2 are tissue-dependent, and $T_2 \leq T_1$ for all materials. The time path of T_1/T_2 relaxation following the deployment of the RF pulse sequence provides critical information. The Fourier transform translates this measurable time process from the

time domain to the frequency domain. The spectrum's amplitude corresponds to the resonant frequency of hydrogen nucleons in water, Figure 10. Contrast across tissues can be detected if the signals measured in those tissues are different. There are two ways to accomplish this: using intrinsic NMR properties such as P_D, T_1 and T_2 or using attributes of externally delivered stimuli. Angle control is possible, as are sophisticated pulse sequences such as rotation-feedback. The 900-second pulse has a TR second repetition period followed by an 1800- second pulse after TE seconds. This second pulse provides an echo signal by replaying a partial rotation. The brain scan image is shown in Figure 11. The weights indicate that the observed difference in intensity between distinct tissues is primarily related to the tissues' P_D, T_1 and T_2 values. Contrast is created using the settings listed in Table II. MR images with pixel intensity $I(x, y)$ generated using the rotation-response series are defined by Equation (14).

$$I(x,y) \propto P_D(x,y)(1 - e^{-\frac{T_R}{T_1}})e^{-\frac{T_E}{T_2}} \qquad (14)$$

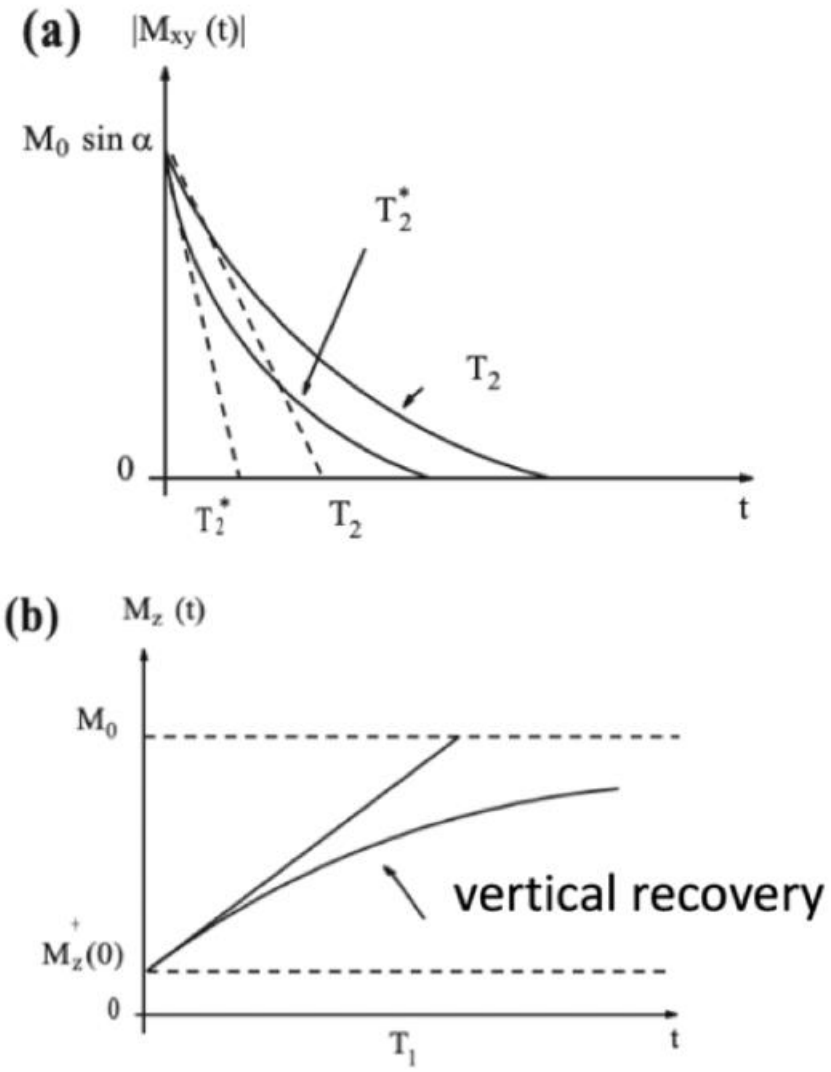

Figure 9. (a) Horizontal stretch and (b) Vertical stretch.

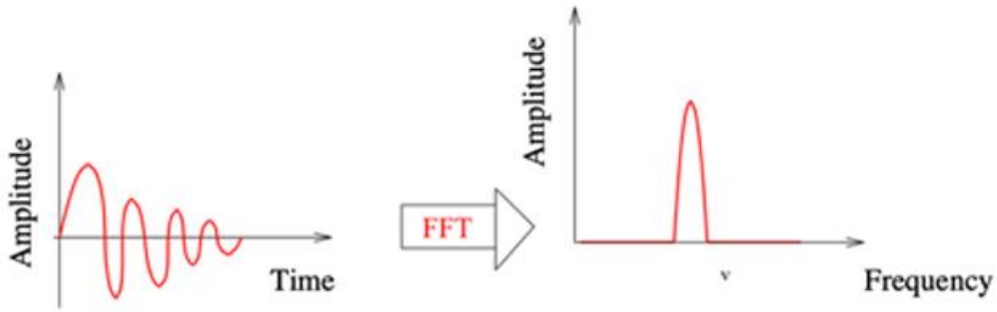

Figure 10. Frequency domain transformation of measured time process. Amplitude in the spectrum is represented at the Larmor frequency.

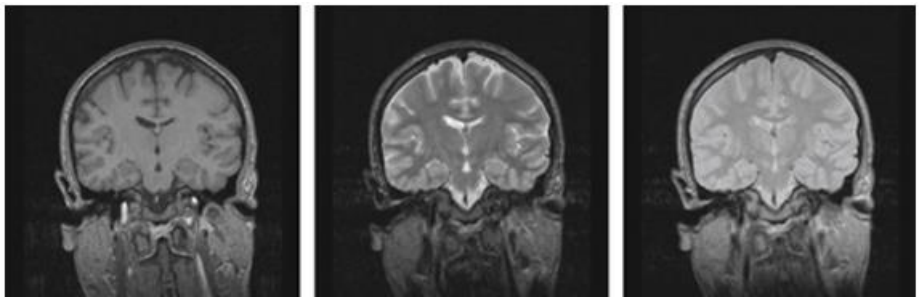

Figure 11. Brain MRI presetting (a) T_1, (b) T_2, and (c) hydrogen density weighted images.

TABLE II – THE BASIC WAY TO CREATE CONTRAST DEPENDS ON T_1,T_2 AND P_D

Contrast	Scan parameters
P_D	Long T_R, FID, or short T_E
T_2	Long T_R, $T_E \approx T_2$
T_1	FID or short T_E, $T_R \approx T_1$

Adjusting the T_R and T_E values control the sensitivity of the signal to T_1/T_2 dilation and produce differently weighted contrast images. For instance, if T_R is significantly greater than T_1 for all tissues within the ROI, the T_1 weights converge to zero, indicating that signal sensitivity to T_1 dilation is absent. The same holds true for T_E, which is significantly less than T_2 in all tissues. When both the T_1 and T_2 sensitivities are decreased, the pixel density is solely determined by the proton density $P_D(x, y)$. The quality of an MR image is determined not only by contrast but also by sampling and noise. Magnetic resonance imaging with multiple dimensions is a critical method in MR. A sequence of three-dimensional magnetic resonance images with the same ROI was captured, provided that the images were registered correctly. This type of image enables the differentiation of various types of tissue. Contrast agents (CA) were utilized to correct the dilation time to further increase the contrast between tissue types. CA is injected intravenously, and throughout this period, signal increase for vascular-enhanced tissues occurs. Functional Magnetic Resonance Imaging (fMRI) is a novel non-invasive technique for examining the brain's cognitive activities [9]. The MR signal is vulnerable to changes in hemodynamic parameters such as blood flow, blood volume, and oxygen output during psychomotor activity, which is the basis for this approach. Blood Oxygenation Level-Dependent (BOLD) contrast is the most often used fMR

signal. The BOLD time response changes when the local deoxyhemoglobin concentration falls in a neuronal active region, as shown in the T_2^* and T_2 weighted MR images.

The hemodynamic effect is defined by two fundamental characteristics: space and time. While the vascular system is primarily responsible for spatial effects, temporal effects account for the delay in detecting changes in the magnetic resonance signal in response to cerebral activity and time. Hemodynamic changes disperse more slowly. Due to timing constraints, two distinct forms of fMR testing are required: "block" and "event-related" designs. The block designs are defined by an alternating sequence of 20–60 second blocks, including a test task. Multiple stimuli are provided at random, and the hemodynamic response to each stimulus is assessed in event-related designs. fMR with the excellent spatial and temporal resolution is an extremely powerful approach for visualizing the human brain's quick and fine activation processes as function localization is predicated on a solid connection between neuronal activity and MR signaling alterations. As is known from theoretical estimations and experimental results [10], the variability of the activated signal on clinical scans appears to be relatively low. This variation encourages analytical techniques to ascertain the response waveform and its related active regions. The primary advantages of this technique are that it allows for non-invasive recording of brain signals without the risk of radiation exposure associated with CT, (2) excellent spatial and temporal resolution, and (3) integration of fMR with other techniques such as MEG and EEG for studying the human brain. In pharmacology, fMR is a valuable technique for determining the brain's response to a specific medication. Apart from therapeutic uses, the value of fMR in comprehending neurological and mental illnesses and in optimizing diagnostics continues to grow.

3. Ultrasound images

Since the early 1950s, ultrasound has been a leading and extensively investigated imaging modality. It is a noninvasive way of imaging using 1–10 MHz vibrations passing through soft tissues and fluids. The method's cost-effectiveness and portability have made it immensely popular. It is diagnostically significant because it enables the "visualization" of pathological alterations in internal organs and blood vessels, which aids in the identification of cancer. Ultrasound imaging works on a fairly simple principle: when sound waves released by the transducer contact with tissue and blood, some of the unabsorbed energy returns to the transducer and is measured. The resulting "ultrasound signature" is formed when ultrasound radiation interacts with a variety of tissue types and is then used for diagnosis. Table 3 shows the speed of sound in tissue as a function of tissue type, temperature, and pressure (some examples of acoustic properties of some biological materials and tissues). Sound waves are "seen" as a result of scattering, absorption, or reflection. $A(x) = A_0 e^{-\alpha x}$, where A is a function of amplitude, A_0 is a constant, and α is the attenuation coefficient, and x is the distance. The amplitude and phase of the feedback signal contain critical information about the interaction and the type of interference medium. The fundamental image equation is the inverse pulse-echo equation, which describes the relationship between the excitation pulse, the

transducer face, the object's reflectivity, and the received signal. In ultrasound, we have the following image modes:

- A-mode or amplitude-mode: the simplest model in which the envelope of the reflected pulse is shown against time. Primarily used in ophthalmology to establish the relative distances between various eye parts and to locate the midbrain or myocardial infarction, figure 12.
- B-mode or light mode: created by scanning the transducer beam in a plane, figure 13, and is employed for both stationary and moving structures, such as the heart valve.
- M-mode or motion mode: displays the A-mode signal corresponding to repeated pulses in a distinct column of the 2D image; this mode is typically used in conjunction with an ECG to determine valve movement.

TABLE III - ACOUSTIC PROPERTIES OF SOME BIOLOGICAL MATERIALS AND TISSUES

Type	Speed of sound (m/s)	Impedance ($106kg/m^2s$)	Level of decline (dB/cm at 1 MHz)
Air	344	0.0004	12
Water	1480	1.48	0.0025
Fat	1410	1.38	0.63
Tissue	1566	1.70	1.2 – 1.3
Liver	1540	1.65	0.94
Bone	4080	7.80	20.0

The two basic techniques used to achieve better sensitivity of echoes along the dominant direction are:

- Beam: increase the sensitivity of the probe for a particular direction.
- Dynamic focus: increases the probe's sensitivity to a particular point in space at a specific time.

4. Microscopic images

A microscope enables the physician to view images of cells invisible to the human eye. A microscope image was captured during observation with the aid of a microscope, as illustrated in Figure 14. Objects with a diameter of fewer than 75 m cannot be identified with the naked eye. Additionally, cells (10 m in diameter), bacteria (1 m), viruses (100 nm), molecules (2 nm), and atoms are frequently observed under a microscope (0.3 nm) [11].

Observing an image under a microscope entails several processes, including adjusting the light source, finding the specimen, and fine-tuning the focus. Locating significant cells is a critical step in the diagnostic use of microscopy. Chemical staining of red blood cells, white blood cells, platelets, and parasites facilitates microscopic examination.

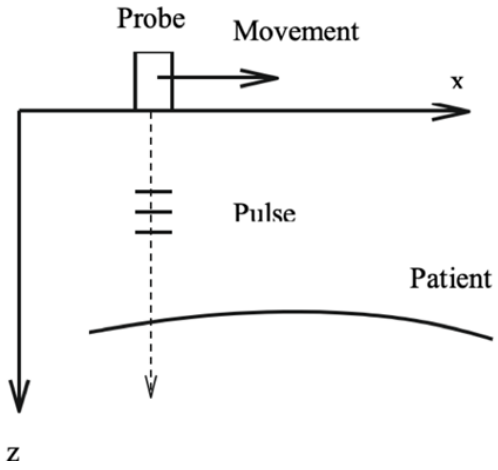

Figure 12. A mode.

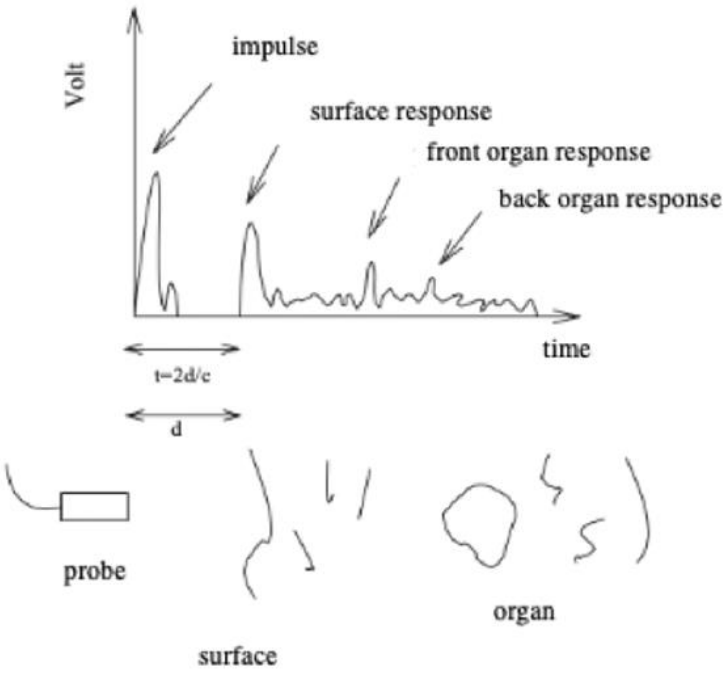

Figure 13. B mode.

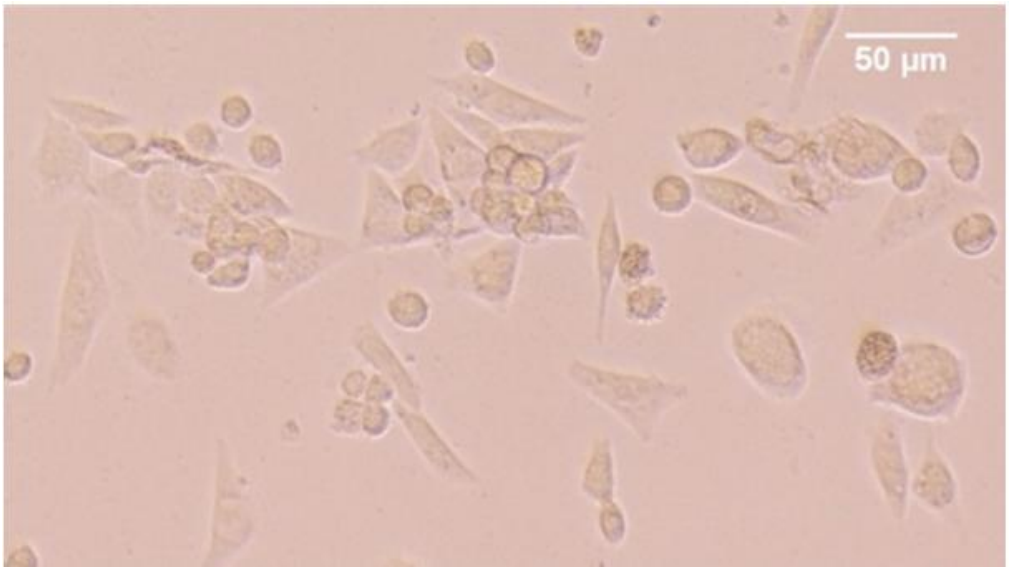

Figure 14. Cell culture medium.

IV. MEDICAL IMAGING SEGMENTATION

Automating image processing and analysis procedures is critical for clinical diagnosis and treatment planning assistance. To identify regions of interest, consistent algorithms are necessary. (Regions of Interest – ROI) and anatomical structure. Computer-aided diagnosis is based on precise medical imaging analysis and utilizes information technology to produce rapid results. Segmentation is the process of breaking down a image into regions with similar attributes, such as color, gray level, contrast, brightness, and texture. The primary function of image segmentation in the medical industry is to define the region of interest (ROI), such as lesions, tumors, and any other anomalies, investigate the anatomical structure, quantify tissue volume, and aid in treatment planning. Automated medical image segmentation is challenging because of inhomogeneity, poor contrast, partial mass effect, artifacts, and relative gray level values of various soft tissues impacting segmented images. There is no uniform algorithm effective with all sorts of medical images, as each imaging technique has distinct advantages and disadvantages. Therefore, many segmentation approaches have been developed and utilized in medical applications, such as (i) shape-based segmentation (ii) interactive segmentation, and (iii) map-based segmentation. Segmentation techniques are further classified as follows: (i) histogram-based segmentation, region-based segmentation, and edge-based segmentation, all of which are based on gray-level features; and (ii) texture-based segmentation characteristics [12].

1. Histogram-based segmentation

Histogram-based image segmentation depends mainly on the threshold value of the histogram. This technique is used with uniformly bright areas in the image under consideration, where thresholding is used to segment the ROI and the background. The threshold can have one value or more values based on the number of ROIs in the image [13] or an appropriate threshold according to the characteristics of the image. Several algorithms are used to perform threshold segmentation efficiently. Typically, the entire image is scanned pixel by pixel to label the pixel as an object or background based on the gray level value compared to the threshold function (T), Algorithm 1.

Numerous improvements have been made to the fundamental histogram-based segmentation method, including the following: adaptive histogram [14], local histogram [15], Kapur/Otsu [16], Gaussian mixture model [17].

Algorithm 1: Segmentation Algorithm based on Global threshold

Inputs: Gray Image I and threshold T (from histogram of I)
Outputs: foreground R_1 and a set of and background R_2
begin
  for each pixel p in the image **do**
    if $p > T$ **then**
      p is assigned to set R_1;
    else
      p is assigned to set R_2
    end
    ;
  end
  Calculate $mean_1$ and $mean_2$ of R_1 and R_2;
  Calculate new threshold $T_{new} = \frac{mean_1+mean_2}{2}$;
  Repeat the previous step until the difference in T is less than the predefined threshold T_0 in subsequent iterations;
end

2. Region-based segmentation

Segmentation is the process of splitting an image into distinct sections in order to ascertain the image's discrete regions directly [18], Algorithm 2.

Subsequently, numerous studies have been conducted to produce region-increasing algorithms with a method for selecting initial locations that maximize efficiency, such as evolving in the direction of regional spacing, such as k-means [19], fuzzy c-means [20], Level Set [21], Fast Marching [22].

Algorithm 2: Region-based segmentation algorithm (Region growing algorithm)

Inputs: Gray Image I and value of similarity
Outputs: object's regions
begin
  Iterative merge of a small region initially set based on the similarity;
  Select an arbitrary pixel;
  Compare that pixel with its neighboring pixels;
  Add similar neighbor pixels to increase the area size from the original pixel;
  if *The growth rate of an area stops* **then**
    Select another original pixel that does not belong to any region;
  end
  Repeat the whole process until all the pixels fit several areas;
end

3. Split and merge segmentation

Split and merge segmentation are based on the quartile tree data descriptor, which divides the image into four quadrants and provides heterogeneous root segmentation. Then, the squares' four neighbors are merged based on their segment identities [23]. This procedure of split/merge is repeated until no further splits/merges are possible, Algorithm 3.

Algorithm 3: Segmentation Algorithm based on Split and Merge regions

Inputs: Gray Image I, homogeneity condition and uniformity criteria for split region or merge regions
Outputs: object's regions
begin
    Define uniformity criteria;
    Divide the image into 4 quadrants;
    if *the quadrant is not uniform* **then**
        Divide the quadrant into 4 smaller quadrants
    end
    Merge two or more neighbors that satisfy the homogeneity condition at each level;
    Continue to split/merge until no more regions can be split or merged;
end

4. Edge-based segmentation

Edge-based segmentation is a powerful segmentation technique because edges contain a wealth of information about the image. Edges denote the boundary between two dissimilar regions; they have information about objects' placement, size/shape, and texture [24]. As a result, gradients can determine pixel values that differ between locations where the image intensity shifts from high to low or vice versa [25]. Different methods of edge-based segmentation include the Hough transform, contour detection, and edge stretching [26], Algorithm 4.

Algorithm 4: Edge-based segmentation algorithm

Inputs: Gray Image I and threshold T
Outputs: object's regions
begin
    Taking the derivative of the image to detect edges;
    Measure gradient amplitude calculate edge strength;
    Preserve all edges with magnitude greater than threshold T;
    Determine the dashed edges;
    Repeat the previous two steps with different threshold values to get closed margins;
end

5. Graph-based segmentation

Recently, a method for segmenting graph images using global functions has been proposed that produces better final segmentation results and fits the criteria of a wide variety of visual applications [27], [28], Algorithm 5.

TABLE IV – STUDIES APPLIED CONVENTIONAL SEGMENTATION METHODS

Ref.	Medical images	Image type	Segmentation method	Segmentation algorithm	Result
[13]	(1) Blood cell (2) Bacteria (3) Medium	Microscopic	Histogram	Threshold	(1) 120 (2) 122 (3) 125
[14]	(1) Retina (2) Knee joint (3) Brain (4) Breast (5) Uterus	(1) X-ray (2) X-ray (3) MRI (4) X-ray (5) X-ray	Histogram	Histogram equalization Cumulative histogram equalization Quadrant dynamic histogram equalization Contrast limited adaptive histogram equalization	MSE (1) 223 (2) 280 (3) 670 (4) 902 (5) 765
[16]	Chest	CT	Histogram	Kapur/Otsu	Accuracy: 97.62%
[18]	Brain	MRI	Region-based	Probabilistic active contour	DICE 90%
[19]	Brain	MRI	Region-based	K-mean clustering	Accuracy: 99.80%
[24]	Heart	MRI	Edge-based	Threshold	Proposed method and parameters for segmentation
[25]	(1) Brain (2) Chest (3) Chest	(1) CT (2) MRI (3) CT	Edge-based	Active contour	Proof-of-concept
[29]	Chest	X-Ray	Graph-based	Minimum spanning tree	Sensitivity: 88.89% Specificity: 95.83%
[30]	Retina	Optical tomography	Graph-based	Shorted path	Error: Mean 2.17 Standard deviation 1.25 Median 1.53
[31]	Brain	MRI	Graph-based	N-cuts	Qualitative comparison with related studies

Algorithm 5: Segmenting image by N-cuts

Inputs: Gray Image I
Outputs: object's regions
begin
    Convert medical image to graph $G = (V, E)$, calculate edge weight and matrix D and W;
    Solve $(D - W)y = \lambda Dy$ for the second smallest eigenvalue;
    Use eigenvectors to create eigenvalues and separate the graphs into 2 subgraphs;
    Decide whether to split the current partition or not. Recursively partition the partitioned parts, if necessary;
end

TABLE V – STUDIES APPLIED MODERN METHODS FOR SEGMENTATION

Ref.	Medical images	Image type	Segmentation method	Segmentation algorithm	Result
[32]	Brain	MRI	Metaheuristics	Particle Swarm Optimization	Accuracy: 92.3%
[33]	Brain	MRI	Metaheuristics	Shuffled Frog Leaping	DICE: 0.8652
[34]	Brain	MRI	CNN	2D CNN	DICE: 0.8503 ± 0.0227
[35]	Chest	X-ray	CNN	2D CNN	AUC: 0.93
[36]	(1) Brain (2) Chest (3) Heart	(1) MRI (2) MRI (3) CT	CNN	2.5D CNN	DICE: (1) 0.7–0.9 (2) 0.75 (3) 0.6
[37]	Brain	MRI	CNN	3D CNN	DICE: 0.84
[38]	Brain	MRI	CNN	3D CNN	DICE: 0.66
[39]	Brain	MRI	LSTM	U-net	DICE: 0.73

- The procedure for determining the graph's minimum spanning tree (Minimal Spanning Tree – MST): MST techniques for image segmentation describe the image by constructing a weighted graph against the feature space [29]. The grouping of pixels treated as a search for the graph's minimum spanning tree displays the image corresponding to the input image. By eliminating edges from the graph, subgraphs are created with a minimum sum of weights.
- Random walker: a proposed approach for performing interactive image segmentation with multiple labels. Utilizes a small number of user-defined labeled pixels to assess and quickly estimate the likelihood that a random walk beginning at each unlabeled pixel will end up at one of the labeled pixels. By labeling each pixel and calculating the maximum probability, high-quality image segmentation can be achieved. This technique has been used to segment the spinal cord using magnetic resonance imaging (MRI) images [40].
- Graph-cut technique using value functions [41]: similar to MST, this method also defines graphs using weighted graphs and global graph partitioning. The graph-cut method can be thought of as a template for image segmentation applications that use graph partitioning techniques. It is advantageous for various applications since it enables the specification of distinct 'cut' criteria and the optimization of global value computation functions for graph partitioning. Typical examples include the max-flow/min-cut algorithm and the Markov random field model. Typical graph-cut methods include the minimal cut, N-cuts, mean cut, and ratio cut [42], [43].
- The method is based on the object's or image's boundary to determine the shortest path through the graph [30]. Using standard graph theory techniques

such as Dijkstra, the edge set is the shortest path between two vertices of a graph. Calculating the border of an object or partition in an image is equivalent to computing the shortest path between two graph vertices. Additionally, other applications demand user participation, such as selecting the boundary's starting point.

- N-cut problems [31].
  – Let $d(i) = \sum_i w_{ij}$
  – D is $n \times n$ diagonal matrix with the diagonal d.
  – Let W is the $n \times n$ symmetric matrix and $w_{ij} = w_{ji}$
  – Find the minimal $\min_{(S,\bar{S})} ncut(S,\bar{S}) = \min_{y} \frac{y^T(D-W)y}{y^T D_y}$ where $y_i \epsilon \{1, -b\}$ and $y^t D_y = 0$
  – Minimize $\frac{y^T(D-W)y}{y^T D_y}$ with the constraint y, the problem becomes finding the second smallest eigenvalue of the problem $(D - W)y = \lambda Dy$

6. Machine Learning

Numerous studies have investigated the segmentation of various organs to extract problematic areas from medical images [44]. Active Appearance Models (AAM), Support Vector Machine (SVM), and Artificial Neural Network (ANN) are three supervised learning methods that are used to interpret medical images. The use of nonlinear statistics in ANN and SVM demonstrates complex modeling relationships between inputs and outputs. The classifier's weight is determined by maximizing organ, structural, and cell features by maximizing the differentiation function. These weights are redistributed after each sample in the training set is processed. The extracted information from the training set gives basic structural information, such as shape, location, and intensity, that can be evaluated as corresponding information for image segmentation. On the other hand, AAMs are statistical models of a structure's shape, with training samples used to extract shape parameter ranges, mean shape, and mean shape. To ensure comparability between segmentation results and training samples, conformational parameters should be restricted in areas where the segmentation technique is used to find the higher position based on appearance information of shape points. Thus, the classifier's algorithms can be widely applied to images of internal organs such as the brain and heart in medical imaging. Metaheuristic-based magnetic resonance imaging: Brain tumors originate when the cell shape within the brain becomes aberrant; they are classified into two forms, benign tumors, and malignant tumors. Primary malignant tumors are those that have not metastasized elsewhere. Secondary malignant tumors are those that have metastasized elsewhere. MR imaging is one of the most advanced tools available for diagnosing brain tumors today. Additionally, segmentation is critical for extracting suspicious regions from complicated medical imaging of the brain. Automated brain tumor diagnosis through MR imaging represents an exciting promise for early and more accurate detection of brain malignancies. Gopal and Karnan [32] developed an intelligent method for diagnosing brain malignancies by

combining PSO (Particle Swarm Optimization) and GA in the clustering phase of MR image processing (Genetic Algorithms). Ladgham et al. [45] developed a novel metaheuristic method dubbed SFLA (Shuffled Frog Leaping Algorithm) or modified MSFLA for brain MR image segmentation. MSFLA enables segmentation without the use of attenuating filters or with the use of three-dimensional Otsu.

7. Deep Learning

- CNN: CNN [45] is a subtype of neural network that consists of a stack of layers, each of which performs a distinct action, such as convolution, aggregation, or loss computation. Each intermediate layer receives the output of the preceding layer as input, as seen in Figure 15. The first layer is an input layer connected directly to the input image. The following collection of layers is composite classes that produce the outcome of converting the input data through a specified number of filters and operate as an automatic feature extraction tool. Filters are frequently referred to as multipliers and have a size that is based on the data. Each convolutional layer's output is handled as an activation map, emphasizing the effect of a particular filter applied to the input data. The structure described above is referred to as the conventional CNN.

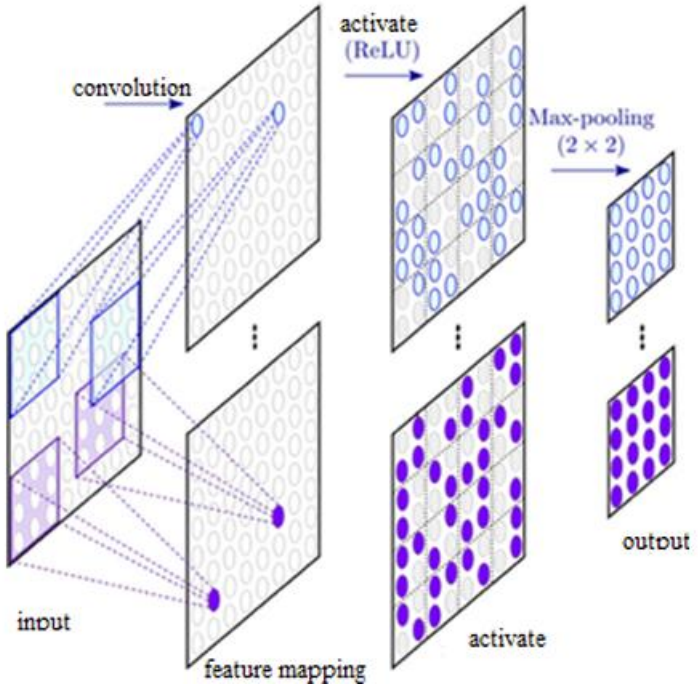

Figure 15. CNN structure.

– 2D CNN: With CNN's promising image classification and pattern recognition skills. The general concept pertaining to medical images is segmentation and applying 2D filters to the 2D input image. Zhang [34] improves segmentation results by combining numerous 2D images with distinct color channels (e.g., R, G, B). The results indicated that a particular type was more productive. Bar et al. [35] merged low-level characteristics from ImageNet with transfer learning. Originate from PiCoDes [46].

– 2.5D CNN: 2.5D Approaches [36], [47], [48], The authors in [47] applied this idea to knee cartilage segmentation. Moeskops et al. evaluated fractional multitasking using the 2.5D architecture [36]. Whether a single network design is capable of segmenting many organs, the author extended the concept even further by incorporating several modalities (for brain MR, breast MR, and cardiac CTA). 2.5D techniques make use of 2D labeled data, which is more accessible than 3D data and produces findings that are more consistent with existing technology. Additionally, breaking the cube into a random collection of 2D images alleviates the size issue [49].

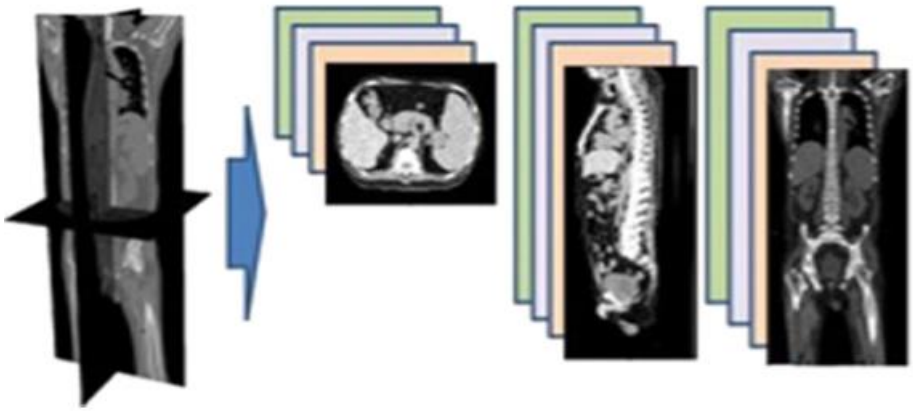

Figure 16. 2.5D example.

– 3D CNN: Using a 2.5D framework enables the enrichment of spatial information. However, the 2.5D approaches remain restricted to the 2D kernel, preventing 3D filters. 3D CNNs are used to extract data along all three axes (X, Y, and Z). The network's construction is similar to that of a 2D CNN, except that 3D modules are used in each required part, for example, 3D composite layers and 3D sub-sampling layers. In one of the first pure 3D models, brain tumors of any size were segmented [37]. The proposal of Kamnitsas [38] is a bidirectional 3D CNN design; this enables the analysis of greater areas surrounding the voxel and enhances the entire system when dealing with multi-scale contexts. With the reduced kernel size usage, this alteration 3 × 3 improved accuracies. Dou [50] offers to employ a collection of spatially shared 3D kernels to address the size and processing time issues. To segment an organ from complicated three-dimensional images. However, training such deep networks presents a substantial difficulty in 3D modeling. Three-dimensional compositing is used at the subsampling layer as feedback filtering in a small neighborhood to stabilize the learned features against local displacement in three-dimensional space. This enables a substantially faster convergence rate than pure 3D CNNs when a convolution mask with the same size as the input volume is used. Kleesiek [37] tackled the difficult task of brain boundary detection using 3D CNNs. The author applied binary segmentation via the cutoff function and mapped the outputs to the desired label, achieving a nearly 6% improvement over other conventional methods.

– Additionally, there are several enhancements, including: Fully Convolutional Network (FCN), Cascaded FCN (CFCN), Multi-Stream FCN, U-Net, 2D U-Net, 3D U-Net, V-Net, Convolutional Residual Networks (CRNs).

- Recurrent Neural Networks (RNNs): RNN Allows the network to retain knowledge from previous layers and use the preceding layer's output data as input data. Since the ROI in medical images is frequently scattered over multiple contiguous slices (e.g., in CT or MR), subsequent slices exhibit correlation. As a result, the RNN can extract adjacent slice contexts as sequential data from the input slices. RNN Structures are composed of two major components for the extraction of internal information that can be accomplished by any CNN model or RNN.
- LSTM: LSTM [51] is often regarded as the most well-known RNN model. A conventional LSTM network requires vectorization of the input, which is inconvenient for medical image segmentation because spatial information is lost. However, the excellent idea is to use convolutional LSTM (CLSTM) [52], [53], where vector multiplication is substituted with convolution. F. Xu and colleagues presented a brain tumor segmentation architecture using MRI images [39].
- Contextual LSTM (CLSTM): CLSTM is used to the output layer of the CNN in [54] to achieve crisper segmentation by utilizing context information from nearby slices. The approach significantly outperforms the well-known U-Net structure [55]. Chen [56] modified the CNN by adding a two-way CLSTM (BD-CLSTM) to the U-Net structure. Because BD-CLSTM can process data sequentially in two directions z+ and z- rather than only one, it outperforms tower LSTM [57] in terms of information acquired in six directions. (x+, x-, y+, y-, z+, and z-). While the LSTM tower moves in six directions, combining the six outputs generated in each direction results in a loss of spatial information. As a result, BDC-LSTM should perform slightly better in the z direction.
- Gated Recurrent Unit (GRU): GRU is an LSTM version in which memory cells are omitted, and the structure becomes simpler without sacrificing speed [58]. Poudel [59] used GRU to analyze the FCN and regression FCN (RFCN) systems. The advantages of RFCN are that it can perform both detection and segmentation in a single build and can be trained for both FCN and GRU.
- Clockwork RNN (CW-RNN): The CW-RNN proposed in [60], [61] has promised long-term information modeling with fewer parameters than a pure RNN. This structure has been used to the portion of connective tissue that surrounds the perimuscular muscle [62]. Since only one component of the CW-RNN is active at any given time, it is more efficient than other approaches (requiring 100 times the runtime of the modified CNN) [63], and a comparison of CW-RNN and U-Net demonstrates that, on average, a 5% increase in accuracy. It should be mentioned that parallelizing RNNs on a restricted number of GPUs is a complex operation, even more so when the data is three-dimensional [57]. Additionally, decoupling training for separate RNN modules increased the

complexity and duration of the training process. In that case, RNN performs better with greater internal organs data and more interlayer information than with a small lesion section with the entire ROI captured in a single slice.

V. CHALLENGES AND OPPORTUNITIES

1. Challenges

At the time this article was written, no one ideal segmentation method exists for all types of medical images, such as coronary imaging [64], Figure 16. There are numerous causes for this, including the following:

- Sophisticated organ structures: The human body's blood arteries are complex. The coronary artery system is comprised of the right and left coronary arteries as well as several small blood vessels that surround the heart and are responsible for the circulatory system's pumping function. The number and position of these blood arteries within the coronary artery network frequently differ between individuals. As a result, it is impossible to apply a model of one patient's vascular structure to another.
- Noise: is inevitability and the most concerning source of error in image processing. Depending on the type of imaging, whether X-ray or MR. Different strategies are required to deal with noise in different kinds of images.
- The contrast between topic and background: in some image areas, such as figure 18, the background will have the same gray level as the subjects. Depending on the time of the image, the volume of blood pushed up varies. Increased or decreased heart rate affects both the quality of blood vessels and the image.
- Non-uniform distribution of light: When capturing X-ray images, the light impact from the lamp and ambient light significantly affects the image quality, causing areas of the image to be bright and dark. However, in the absence of light, the items contained within the image will be unrecognizable. As a result, non-uniform light distribution is one of the constraints of X-ray image databases that must be recognized. Then, we should have strategies for resolving this circumstance.

2. Opportunities

Each medical category will provide information about a specific area of the body. The relevance study is based on a characteristic of each image kind and employs analytical relevancy. Medical problem analysis is one of the most complex problems; it is also critical for classification and subsequent analysis. The results of the problem segmentation, the medical image, will be used to determine the problem's level of accuracy later. With the issues described above and numerous more obstacles encountered during implementation, these are the formulas and an opportunity for many researchers to tackle the medical image analysis problem.

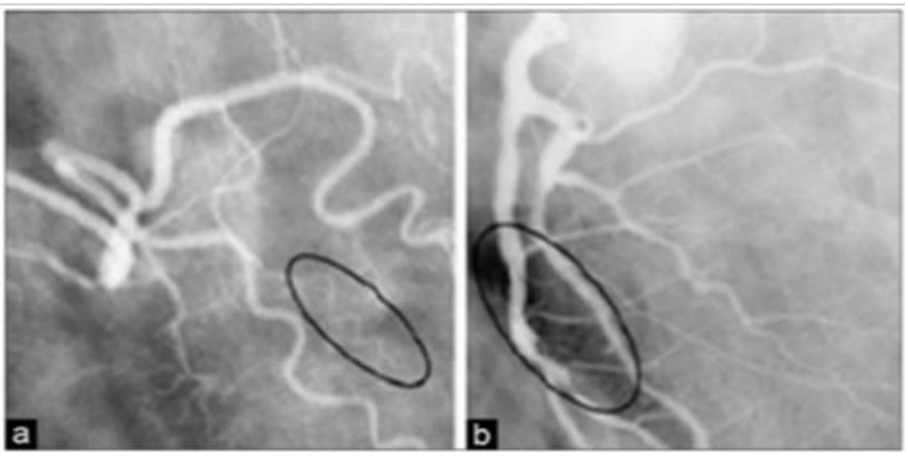

Figure 17. Limitations of coronary angiography: (a) low contrast, and (b) non-uniform light distribution.

REFERENCES

[1] S. Webb, "The physics of medical imaging-medical science," 1996.

[2] ScienceDirect Topics, "Medical imaging, medical imaging - an overview," [Online]. Available: https://www.sciencedirect.com/topics/computerscience/medical-imaging. [Accessed 01 January 2022].

[3] History, "German scientist discovers x-ray," [Online]. Available: https://www.history.com/this-day-in-history/german-scientist-discovers-x-rays. [Accessed 01 January 2022].

[4] Ultrasound schools info, "History of ultrasound," [Online]. Available: https://www.ultrasoundschoolsinfo.com/history/. [Accessed 01 January 2022].

[5] The evolution of imaging technology, "History of the pet/ct scanner," [Online]. Available: https://www.theevolutionofimagingtechnology.net/history-of-the-pet-ct-scanner/. [Accessed 01 January 2022].

[6] H. Zaidi, "Medical imaging: Current status and future perspectives," 1997.

[7] M. Akay, "Time frequency and wavelets in biomedical signal processing," 1998.

[8] Z. H. Cho, J. P. Jones and M. Singh, Foundations of medical imaging, vol. 228, Wiley New York, 1993.

[9] S. Ogawa, D. W. Tank, R. Menon, J. M. Ellermann, S. G. Kim, H. Merkle and K. Ugurbil, "Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging," *Proceedings of the National Academy of Sciences*, vol. 89, no. 13, pp. 5951-5955, 1992.

[10] S. Ogawa, T. Lee and B. Barrere, "The sensitivity of magnetic resonance image signals of a rat brain to changes in the cerebral venous blood oxygenation," *Magnetic resonance in medicine,* vol. 29, no. 2, pp. 205-210, 1993.

[11] A. Maier, S. Steidl, V. Christlein and J. Hornegger, "Medical imaging systems: An introductory guide," 2018.

[12] T. Heimann, B. Van Ginneken, M. A. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel and G. Bekes, "Comparison and evaluation of methods for liver segmentation from ct datasets," *IEEE transactions on medical imaging,* vol. 28, no. 8, pp. 1251-1265, 2009.

[13] O. J. Tobias and R. Seara, "Image segmentation by histogram thresholding using fuzzy sets," *IEEE transactions on Image Processing,* vol. 11, no. 12, pp. 1457-1465, 2002.

[14] N. Salem, H. Malik and A. Shams, "Medical image enhancement based on histogram algorithms," *Procedia Computer Science,* vol. 163, pp. 300-311, 2019.

[15] S. Patel, K. Bharath, S. Balaji and R. K. Muthu, "Comparative study on histogram equalization techniques for medical image enhancement," in *Soft Computing for Problem Solving,* Springer, 2020, pp. 657-669.

[16] S. C. Satapathy, D. J. Hemanth, S. Kadry, G. Manogaran, N. M. Hannon and V. Rajinikanth, "Segmentation and evaluation of COVID-19 lesion from CT scan slices-a study with kapur/otsu function and cuckoo search algorithm," 2020.

[17] F. Riaz, S. Rehman, M. Ajmal, R. Hafiz, A. Hassan, N. R. Aljohani, R. Nawaz, R. Young and M. Coimbra, "Gaussian mixture model based probabilistic modeling of images for medical image segmentation," *IEEE Access,* vol. 8, pp. 16846-16856, 2020.

[18] E. R. Arce-Santana, A. R. Mejia-Rodriguez, E. Martinez-Pena, A. Alba, M. Mendez, E. Scalco, A. Mastropietro and G. Rizzo, "A new probabilistic active contour region-based method for multiclass medical image segmentation," *Medical & biological engineering & computing,* vol. 57, no. 3, pp. 565-576, 2019.

[19] D. M. Kumar, D. Satyanarayana and M. Prasad, "An improved gabor wavelet transform and rough k-means clustering algorithm for MRI brain tumor image segmentation," *Multimedia Tools and Applications,* vol. 80, no. 5, pp. 6939-6957, 2021.

[20] K. M. Aljebory and T. S. Mohammed, "Modified fuzzy c-means clustering algorithm application in medical image segmentation," *JEA Journal of Electrical Engineering,* vol. 2, no. 1, pp. 1-9, 2018.

[21] Z. Zhang and J. Song, "An adaptive fuzzy level set model with local spatial information for medical image segmentation and bias correction," *IEEE Access,* vol. 7, pp. 27322-27338, 2019.

[22] D. Jia and X. Zhuang, "Directional fast-marching and multi-model strategy to extract coronary artery centerlines," *Computers in Biology and Medicine*, vol. 108, pp. 67-77, 2019.

[23] J. Freixenet, X. Munoz, D. Raba, J. Martí and X. Cufí, "Yet another survey on image segmentation: Region and boundary information integration," in *European conference on computer vision*, Springer, 2002, p. 408–422.

[24] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE transactions on medical imaging,* vol. 22, no. 2, pp. 137-154, 2003.

[25] R. Hemalatha, T. Thamizhvani, A. J. A. Dhivya, J. E. Joseph, B. Babu and R. Chandrasekaran, "Active contour based segmentation techniques for medical image analysis," *Medical and Biological Image Analysis,* vol. 4, no. 17, p. 2, 2018.

[26] R. Hapsari, M. Utoyo, R. Rulaningtyas and H. Suprajitno, "Iris segmentation using hough transform method and fuzzy c-means method," in *Journal of Physics: Conference Series*, Vols. 1477, no. 2, IOP Publishing, 2020, p. 022037.

[27] B. Peng, L. Zhang and D. Zhang, "A survey of graph theoretical approaches to image segmentation," *Pattern recognition,* vol. 46, no. 3, p. 1020–1038, 2013.

[28] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International journal of computer vision,* vol. 59, no. 2, p. 167–181, 2004.

[29] S. Nithya, S. Bhuvaneswari and S. Senthil, "Robust minimal spanning tree using intuitionistic fuzzy c-means clustering algorithm for breast cancer detection," *Am. J. Neural Netw. Appl.,* vol. 5, pp. 12-22, 2019.

[30] X. Liu, D. Liu, T. Fu, Z. Pan, W. Hu and K. Zhang, "Shortest path with backtracking based automatic layer segmentation in pathological retinal optical coherence tomography images," *Multimedia Tools and Applications,* vol. 78, no. 12, p. 15817–15838, 2019.

[31] N. Ghorpade and H. Bhapkar, "Evaluation and performance analysis of brain MRI segmentation methods," *Int. J. Computer Sci. Eng.,* vol. 6, no. 8, p. 687–696, 2018.

[32] N. N. Gopal and M. Karnan, "Diagnose brain tumor through MRI using image processing clustering algorithms such as fuzzy c means along with intelligent optimization techniques," in *2010 IEEE international conference on computational intelligence and computing research*, IEEE, 2010, pp. 1-4.

[33] A. Ladgham, F. Hamdaoui, A. Sakly and A. Mtibaa, "Fast MR brain image segmentation based on modified shuffled frog leaping algorithm," *Signal, Image and Video Processing,* vol. 9, no. 5, p. 1113–1120, 2015.

[34] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji and D. Shen, "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *Neuro Image,* vol. 108, p. 214–224, 2015.

[35] Y. Bar, I. Diamant, L. Wolf and H. Greenspan, "Deep learning with non-medical training used for chest pathology identification," in *Medical Imaging 2015: Computer-Aided Diagnosis*, vol. 9414, International Society for Optics and Photonics, 2015, p. 94140V.

[36] P. Moeskops, J. M. Wolterink, B. H. v. d. Velden, K. G. Gilhuijs, T. Leiner, M. A. Viergever and I. Isgum, "Deep learning for multi-task medical image segmentation in multiple modalities," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2016, p. 478–486.

[37] G. Urban, M. Bendszus, F. Hamprecht and J. Kleesiek, "Multi-modal brain tumor segmentation using deep convolutional neural networks," *MICCAI BraTS (brain tumor segmentation) challenge. Proceedings, winning contribution,* p. 31–35, 2014.

[38] K. Kamnitsas, L. Chen, C. Ledig, D. Rueckert and B. Glocker, "Multi-scale 3d convolutional neural networks for lesion segmentation in brain MRI," *Ischemic stroke lesion segmentation,* vol. 13, p. 46, 2015.

[39] F. Xu, H. Ma, J. Sun, R. Wu, X. Liu and Y. Kong, "LSTM multi-modal unet for brain tumor segmentation," in *2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC)*, IEEE, 2019, p. 236–240.

[40] D. Brindha and N. Nagarajan, "An efficient automatic segmentation of spinal cord in MRI images using interactive random walker (RW) with artificial bee colony (ABC) algorithm," *Multimedia Tools and Applications,* vol. 79, no. 5, pp. 3623-3644, 2020.

[41] J. Dogra, S. Jain and M. Sood, "Analysis of graph cut technique for medical image segmentation," in *International Conference on Advanced Informatics for Computing Research*, Springer, 2019, p. 451–463.

[42] S. Wang and J. M. Siskind, "Image segmentation with minimum mean cut," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 1, IEEE, 2001, p. 517–524.

[43] S. Wang and J. M. Siskind, "Image segmentation with ratio cut," vol. 25, no. 6, p. 675–690, 2003.

[44] D. L. Pham, C. Xu and J. L. Prince, "Current methods in medical image segmentation," *Annual review of biomedical engineering,* vol. 2, no. 1, p. 315–337, 2000.

[45] F. Commandeur, M. Goeller, J. Betancur, S. Cadet, M. Doris, X. Chen, D. S. Berman, P. J. Slomka, B. K. Tamarappoo and D. Dey, "Deep learning for

quantification of epicardial and thoracic adipose tissue from non-contrast CT," *IEEE transactions on medical imaging,* vol. 37, no. 8, pp. 1835-1846, 2018.

[46] A. Bergamo, L. Torresani and A. Fitzgibbon, "Picodes: Learning a compact code for novel-category recognition," *Advances in neural information processing systems,* vol. 24, 2011.

[47] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *International conference on medical image computing and computer-assisted intervention*, Springer, 2013, pp. 246-253.

[48] H. R. Roth, L. Lu, A. Seff, K. M. Cherry, J. Hoffman, S. Wang, J. Liu, E. Turkbey and R. M. Summers, "A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations," in *International conference on medical image computing and computer-assisted intervention*, Springer, 2014, pp. 520-527.

[49] R. Fakoor, F. Ladhak, A. Nazi and M. Huber, "Using deep learning to enhance cancer diagnosis and classification," in *Proceedings of the international conference on machine learning*, vol. 28, New York, ACM, 2013, pp. 3937-3949.

[50] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin and P. A. Heng, "3D deeply supervised network for automated segmentation of volumetric medical images," *Medical image analysis,* vol. 41, p. 40–54, 2017.

[51] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation,* vol. 9, no. 8, p. 1735–1780, 1997.

[52] N. Srivastava, E. Mansimov and R. Salakhudinov, "Unsupervised learning of video representations using LSTMs," in *International conference on machine learning*, PMLR, 2015, p. 843–852.

[53] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong and W. C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Advances in neural information processing systems,* vol. 28, 2015.

[54] J. Cai, L. Lu, Y. Xie, F. Xing and L. Yang, "Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function," *arXiv preprint arXiv:1707.04912,* 2017.

[55] O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234-241.

[56] J. Chen, L. Yang, Y. Zhang, M. Alber and D. Z. Chen, "Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation," *Advances in neural information processing systems,* vol. 29, 2016.

[57] M. F. Stollenga, W. Byeon, M. Liwicki and J. Schmidhuber, "Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation," *Advances in neural information processing systems*, vol. 28, 2015.

[58] D. Cheng and M. Liu, "Combining convolutional and recurrent neural networks for Alzheimer's disease diagnosis using pet images," in *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*, IEEE, 2017, p. 1–5.

[59] R. P. Poudel, P. Lamata and G. Montana, "Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation," in *Reconstruction, segmentation, and analysis of medical images*, Springer, 2016, p. 83–94.

[60] D. J. Rezende, S. Mohamed and D. Wierstra, "Proceedings of the 31st international conference on international conference on machine learning," 2014.

[61] J. Koutnik, K. Greff, F. Gomez and J. Schmidhuber, "A clockwork RNN," in *International Conference on Machine Learning*, PMLR, 2014, p. 1863–1871.

[62] Y. Xie, Z. Zhang, M. Sapkota and L. Yang, "Spatial clockwork recurrent neural network for muscle perimysium segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2016, p. 185–193.

[63] D. Ciresan, A. Giusti, L. Gambardella and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," *Advances in neural information processing systems*, vol. 25, 2012.

[64] P. V. Truc, M. A. Khan, Y. K. Lee, S. Lee and T. S. Kim, "Vessel enhancement filter using directional filter bank," *Computer Vision and Image Understanding*, vol. 113, no. 1, p. 101–112, 2009.

AUXILIARY

A1. Details of experimental data

TBA

A2. Details of calculation steps

TBA

A3. Details of simulation diagrams

TBA