People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Echikh El Arbi Tebessi University – Tebessa
Faculty of Exact Sciences and Sciences of nature and life

Dèpartment
Mathematics and Informatics

Final thesis For obtaining the Master's degree

Domain : Informatics
Speciality : Systems and Multimedia

Deformable Medical Image Registration using advanced AI Techniques

Realized by :
Hachichi Aimene Nedjemeddine

Publicly defended, in front of the jury composed of:

Dr.Gettal Abdeldjalil	**MCA**	**President**
Dr.Abbas Faycel	**MCA**	**Examiner**
Dr.Bennour Akram	**MCA**	**Supervisor**

University year: 2022/2023

بسم الله الرحمن الرحيم

Thanks

First and foremost, I would like to thank Allah for giving me the strength and courage I needed to complete this modest work.

I would like to express my deep gratitude to my supervisor, Dr. Bennour Akram, for supporting me, advising me, guiding me, and assisting me throughout this project.

I would also like to thank the entire faculty of the Department of Mathematics and Computer Science at El Arbi Tebessi University for their contribution to my education.

Lastly, a big thank you to my family and friends who have encouraged and supported me throughout this project.

Dedication

"I DEDICATE THIS HUMBLE WORK TO MY DEAR MOTHER, WHO SUPPORTED ME AND ENCOURAGED ME DURING THE TOUGHEST TIMES… AND TO MY DEAR FATHER, MAY HE REST IN PEACE. AND TO MY BROTHERS AND ALL MY FRIENDS…"

AIMENE

Abstract :

Image medical registration is one of the most active areas in image processing, and has attracted particular interest in many areas such as medical image analysis, remote sensing, and mapping. The main idea of these techniques is to identify the spatial shift between two images which makes it possible to match the equivalent properties. Registration can be local or international. Several deep-learning-based registration techniques have been developed that have shown impressive performance in terms of matching, which can be grouped into two main classes: Flexible and Diffeomorphic feature registration based on extraction between two images.

In the graduation project, we established an accurate learning model based on convolutional neural networks that were trained to record images. Where the model proved acceptable accuracy and speed

:

خلاصة

يعد مطابقةالصور أحد أكثر المجالات نشاطًا في معالجة الصور ، وقد اجتذب اهتمامًا خاصًا في العديد من المجالات مثل تحليل الصور الطبية والاستشعار عن بعد ورسم الخرائط. الفكرة الرئيسية لهذه التقنيات هي تحديد التحول المكاني بين صورتين مما يجعل من الممكن مطابقة الخصائص المكافئة. يمكن أن يكون التسجيل محليًا أو دوليًا. تم تطوير العديد من تقنيات التسجيل القائمة على التعلم العميق والتي أظهرت أداءً مثيرًا للإعجاب من حيث المطابقة ، والتي يمكن تجميعها في فئتين رئيسيتين: تسجيل ميزة مرنة ومختلفة بناءً على الاستخراج بين صورتين

في مشروع التخرج ، أنشأنا نموذجًا تعليميًا دقيقًا يعتمد على الشبكات العصبية التلافيفية التي تم تدريبها على تسجيل الصور. حيث أثبت النموذج الدقة والسرعة المقبول

Résumé :

Le recalage d'images medicales est l'un des domaines les plus actifs du traitement d'images et a suscité un intérêt particulier dans de nombreux domaines tels que l'analyse d'images médicales, la télédétection et la cartographie. L'idée principale de ces techniques est d'identifier le décalage spatial entre deux images qui permet de faire correspondre des propriétés équivalentes. L'enregistrement peut être local ou international. Plusieurs techniques de recalage basées sur l'apprentissage en profondeur ont été développées qui ont montré des performances impressionnantes en termes de correspondance, qui peuvent être regroupées en deux classes principales : le recalage de caractéristiques flexible et difféomorphique basé sur l'extraction entre deux images.

Dans le projet de fin d'études, nous avons établi un modèle d'apprentissage précis basé sur des réseaux de neurones convolutifs formés pour enregistrer des images. Lorsque le modèle s'est avéré une précision et une vitesse acceptables

Contents :

List of Figures :

Chapter 1 : Imaging Medical

Chapter 2 : Image Registration

Chapter 3 : Artificial intelligence techniques

Chapter 4 : Concept and programming

List of tables :

Chapter 1

Chapter 3

General Introduction

Deformable image medical registration plays a vital role in the field of medical imaging by aligning and mapping images obtained from different modalities or at different time points. It allows medical images to be compared, combined, and analyzed, providing valuable insights for diagnosis, treatment planning, and monitoring of various diseases. With advances in artificial intelligence (AI) technologies, anamorphic image registration has been revolutionized, allowing for more accurate and robust image alignment.

In recent years, advanced AI technologies have emerged as powerful tools in medical image recording. These technologies make use of deep learning algorithms and other AI methodologies to overcome the limitations of traditional logging methods. It enables automatic and efficient deformable registration of medical images, reducing reliance on manual intervention and autonomy.

One of the main advantages of using advanced AI techniques to record deformable images is their ability to learn complex patterns and deformations from large datasets. Deep learning models, such as convolutional neural networks (CNNs) and generative adversarial networks (GANs), have shown remarkable capabilities in capturing complex image features and distortions, allowing highly accurate and adaptive recording.

Advanced AI techniques in registration distorted images have facilitated the development of personalized medicine. By learning from individual patient data, these techniques can optimize registration parameters and create deformable, patient-specific models. This customized approach improves the accuracy of image registration, resulting in more accurate diagnosis, treatment planning, and monitoring tailored to each patient's unique characteristics.

The combination of AI technologies with anamorphic registration enabled real-time interactive recording workflows. With the ability to quickly process large amounts of data, AI algorithms can provide rapid and continuous registry updates, enhancing intraoperative guidance, image-guided interventions, and adaptive radiotherapy.

Despite the great progress, there are still challenges in applying advanced AI techniques to record distorted images. Issues such as robustness to different imaging conditions, limited availability of annotated training data, and interpretability of deep learning models must be addressed to ensure reliable and widespread adoption of these techniques in clinical practice.

The combination of anamorphic image registration and advanced AI technologies has revolutionized medical imaging by enhancing the accuracy, efficiency and personalized nature of image alignment. These technologies have the potential to dramatically improve clinical outcomes, aid treatment decisions, and advance our understanding of complex diseases. With continuous research and technological advancements, registration deformed images using advanced AI techniques holds great promise for the future of medical imaging and healthcare.

This memory is structured as follows :

Chapter 1 : This chapter present Medical imaging and their types in detail.

Chapter 2:This chapter present image registration and methode for transformation

Chapter 3 : This chapter constains several AI technologies , in partiqular the technique on which the study based , such as deep learning and convolutional neural networks .

Chapter 4 : This chapter present the desing of our contribution that corresponds to a deep neural networks architecture for deformable image medical registration.

Chapter 1

Imaging Medical

Introduction

The medical field relies heavily on image comparisons to track the progression of a patient's pathology. This involves comparing images acquired at different times, as well as superimposing images from multiple patients or using a reference such as an anatomical or functional atlas. With the increasing use of different acquisition methods, each with their own advantages and drawbacks, interpreting image data has become a complex task.

One common approach to address this complexity is to merge images acquired with different modalities, but this process can be challenging due to geometric differences between the images, and the presence of distortions or deformations within the images themselves. Several factors contribute to these challenges, including differences in the position of the subject and device parameters at each acquisition, as well as errors inherent to each acquisition system that can introduce distortions to the image. Additionally, movements of certain elements of the subject during imaging can further complicate the comparison process.

To overcome these challenges, medical professionals use a process of registration, which aims to align images as accurately as possible. This technique involves applying mathematical algorithms to manipulate the images to adjust for any differences in position or orientation, and to compensate for any distortions or deformations within the images. By employing this process, medical professionals can more accurately compare images and identify any changes that may have occurred over time, ultimately improving patient outcomes.

1.The image :

Images can be defined as the representation of a real-life scene, either as an exact or analog reproduction. They are also a structured collection of information that holds meaning when displayed to the human eye. In terms of their mathematical description, images can be characterized as a function I(x, y), where x and y are the spatial coordinates of a point within the image, and I is a continuous analog brightness function defined within a bounded domain that is dependent on light intensity and color. However, this representation is unusable for machines, which require the image to be digitized before processing. [1]

1.1 Image formats:

In the field of image processing, the choice of image format is crucial as it affects the quality and suitability of the image for specific applications. Different image formats are designed to accommodate specific types of images and data, with varying levels of color depth and compression capabilities. For instance, GIF is a basic image format with limited color depth, suitable for storing simple images like logos and icons, while JPEG is a widely-used format for consumer imaging, capable of storing up to 24-bit RGB color images. TIFF is a highly-adaptable format that can store a wide range of image data forms, making it useful for scientific and medical imaging applications. Other formats like BMP and PNG are also common and have their own advantages and limitations. The choice of image format ultimately depends on the specific requirements of the application at hand.

type	Name	Properties
GIF	**Graphics interchange format**	Limited to only 256 colors (8 bit); losslesscompression
JPEG	**Joint Photographic Experts Group**	-In most common use today; -with loss of compression; -lossless variants exist
BMP	**Bit mappicture**	_Base picture format; - limited (usually) lossless compression; - lossy variants exist
PNG	**Portable network graphics**	-New lossless compression format; -designed to replace GIF
TIF/TIFF	**Tagged image (file) format**	Very flexible, detailed and adaptable format; -compressed/uncompressed variants exist

Table 1: Image formats and their properties

1.2.Medical imaging:

Medical imaging has revolutionized the field of medicine, allowing doctors to peer inside the human body without invasive procedures. With advanced imaging techniques, medical professionals can generate high-resolution images that provide a visual representation of what is happening inside the body. This enables doctors to make accurate diagnoses and plan appropriate treatments.

Anatomical imaging is a type of medical imaging that focuses on creating detailed images of the physical structure of organs and tissues. This type of imaging is commonly used to detect and diagnose conditions such as tumors, fractures, and other abnormalities. Some common anatomical imaging techniques include X-rays, CT scans, and MRI.

Functional imaging, on the other hand, is a type of medical imaging that focuses on the activity and function of organs and tissues. This type of imaging provides information on how different parts of the body are functioning, such as blood flow or metabolic activity. Functional imaging is often used to diagnose and monitor conditions such as cancer, neurological disorders, and heart disease. Some common functional imaging techniques include PET scans and fMRI.

In the field of plant biology, medical imaging is also a valuable tool for studying the internal structures and functions of plants. Researchers use a variety of imaging techniques to visualize plant anatomy, monitor plant growth and development, and investigate how plants respond to environmental stressors. This helps to advance our understanding of plant physiology and can inform agricultural practices. [2]

1.3. Definition of a digital image:

A digital image is a representation of an analog image that has been digitized through a process of sampling and quantization. The surface of the image is divided into fixed-size cells or pixels, each with a specific color or gray level, and a corresponding position in the real image. The digitization process involves converting the analog image into a digital form, represented by a two-dimensional matrix of numerical values f(x, y).

The matrix contains n rows and p columns, where each element (x, y) represents a pixel in the image, and its value is the gray level represented by m bits (2^m gray levels, ranging from 0 for black to 2^m-1 for white). The value of each pixel represents the intensity of light measured by the sensor at that location.

To display or manipulate digital images using a computer, the image must be scanned and digitized. This process involves converting the physical image into a digital image state. The resulting digital image is represented as a code or numerical value, which can be saved, processed, and displayed using a variety of computer tools.

In general, a digital image can refer to any image that has been digitized, regardless of the specific method used. Digital images are widely used in fields such as medicine, biology, engineering, and entertainment, and have revolutionized the way we capture, store, and process visual information. [3]

1.4. Characteristics of a digital image:

1.4.1. The pixel:

Pixels are small squares that make up the screen and display one color at a time. Therefore, the screen consists of millions of pixels in height and width. All these pixels form the image of the screen. The pixels are so small that they are almost invisible to the naked eye. The more screens, the clearer the image! [4] .

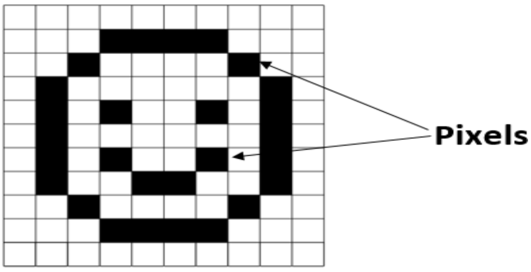


Figure 1: Image and pixels

1.4.2. Intensity level:

Intensity level refers to the brightness or luminosity of the pixels within an image. It is a measure of the amount of light that is present in each pixel, and is often expressed in terms of luminous intensity per unit area. The intensity level of an object can be affected by various factors, including the amount of light that is reflected or absorbed by the object, as well as the distance between the object and the observer. In general, a higher intensity level corresponds to a brighter or more luminous image, while a lower intensity level indicates a darker or less luminous image.

1.4.3.Types of images:

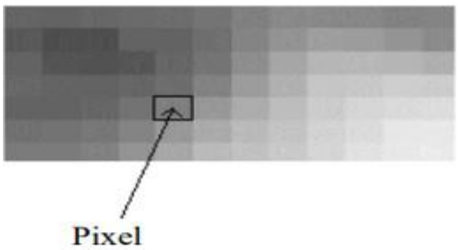
there are three types of images:

1.4.3.1. Binary images (in black and white):

Binary images are digital images that consist of only two values: black and white, or 0 and 1. These images are the simplest form of digital images, and their encoding and decoding can be directly done using binary code. In a binary image, each pixel is either black (0) or white (1), depending on its intensity level. Binary images are commonly used in various applications, such as document scanning and recognition, fingerprint recognition, and image segmentation. They are also used in medical imaging for tasks such as identifying bone structure and detecting tumors. Binary images are easy to process and analyze, and can be efficiently stored and transmitted, making them useful for a variety of applications. [5]

1.4.3.2. Grayscale Images (Monochrome):

be capable of displaying different shades of gray. Grayscale images are also referred to as monochrome images because they use only one color (gray) to represent the image. The value assigned to each pixel in a grayscale image represents the intensity of light or darkness at that point in the image. The higher the value, the brighter the pixel, and the lower the value, the darker the pixel. Grayscale images are commonly used in medical imaging, document scanning, and black and white photography. [6]



107	99	99	107	107	115	132	132	140	140	132
90	82	82	99	99	115	132	148	165	156	148
99	82	90	90	99	115	140	165	181	181	181
99	107	99	99	115	132	156	173	198	198	206
99	99	99	115	$\boxed{123}$	156	173	189	206	214	222
115	115	123	132	148	173	181	189	206	214	231
123	123	132	148	165	189	198	198	206	222	231

Niveau de gris

Figure 2: Example of grayscale image

1.4.3.3. Color images (Polychromes):

Color images, also known as polychromes, are composed of a combination of red, green, and blue (RGB). In order to represent color images digitally, each pixel is assigned a value for the amount of red, green, and blue that it contains. This value is usually represented using 24 bits, with 8 bits assigned to each of the red, green, and blue components. This allows for a total of 16,777,216 different color possibilities, resulting in a highly detailed and realistic representation of the original image. Different color models can also be used, such as CMYK (Cyan, Magenta, Yellow, and Key), which is commonly used in printing. [7]

1.5. History of medical imaging:

Medical imaging has a rich history that dates back to the discovery of X-rays by Wilhelm Conrad Roentgen in 1895. Roentgen's discovery revolutionized the field of medicine and provided a non-invasive way of looking inside the human body. Shortly after Roentgen's discovery, the first X-ray image of the human body was taken, which was a radiograph of Roentgen's wife's hand.

Over the years, various medical imaging techniques have been developed, each with their own strengths and limitations. In the early days, X-rays were the primary imaging modality, and their use quickly became widespread in medical practice. However, the risks associated with ionizing radiation were not fully understood at the time, and many patients suffered radiation burns and other adverse effects as a result.

In the 1950s, the development of the ultrasound machine brought a new era of non-invasive imaging. Ultrasound uses sound waves to create images of the body, and it is particularly useful for imaging soft tissues such as the liver, kidneys, and heart.

In the 1970s, the development of computed tomography (CT) revolutionized medical imaging once again. CT scans use X-rays to create cross-sectional images of the body, providing a more detailed and accurate view of the internal organs and tissues.

In the 1980s, magnetic resonance imaging (MRI) was developed, which uses a strong magnetic field and radio waves to create images of the body. MRI is particularly useful for imaging the brain and spinal cord, as well as soft tissues such as the heart and muscles.

In recent years, medical imaging has continued to evolve with the development of new technologies such as positron emission tomography (PET) and single-photon emission computed tomography (SPECT), which use radioactive tracers to create images of the body at the molecular level. [8].

Figure 3: Wilhelm Röntgen's first hand X-ray 1895

1.6.Medical image acquisition method:

Medical image acquisition is the process of capturing images of the human body for diagnostic or therapeutic purposes. There are various methods for medical image acquisition, including:
X-ray imaging: This method uses ionizing radiation to produce images of internal body structures. X-rays are absorbed differently by different tissues, producing varying levels of brightness in the image. X-ray imaging is commonly used for detecting bone fractures, lung abnormalities, and certain tumors.
Computed tomography (CT): CT scans use a series of X-rays taken from different angles to create cross-sectional images of the body. This method provides more detailed images than traditional X-rays and is commonly used to detect internal injuries, such as brain bleeding or abdominal trauma.
Magnetic resonance imaging (MRI): This method uses a powerful magnetic field and radio waves to produce detailed images of the body's soft tissues. MRI is commonly used for brain and spinal cord imaging, as well as for detecting tumors, inflammation, and other abnormalities.
Ultrasound imaging: This method uses high-frequency sound waves to produce images of internal organs and tissues. Ultrasound is commonly used for fetal imaging, detecting abnormalities in the liver and gallbladder, and guiding certain medical procedures.
Nuclear medicine imaging: This method involves the injection of radioactive substances into the body, which are then detected by a special camera to produce images of internal organs and tissues. Nuclear medicine imaging is commonly used for detecting cancer, heart disease, and certain neurological disorders.
Optical imaging: This method uses light to produce images of internal tissues. Optical imaging is commonly used for imaging the eye and certain types of cancer.
Positron Emission Tomography (PET) imaging: PET scans use a small amount of radioactive material to produce images of the body's metabolic activity.
Single Photon Emission Computed Tomography (SPECT) imaging: SPECT scans use a small amount of radioactive material to produce images of the body's metabolic activity.
Each of these methods has its advantages and limitations, and the choice of imaging method depends on the clinical question, the anatomical region of interest, and the patient's medical history. [9].

1.6.1. X-ray imaging:

- Definition:

X-rays, a form of electromagnetic radiation with a high intensity and brief wavelength that may penetrate materials, including human tissues, are defined as such. A shadow-like image is created on a detector or film when X-rays are focused at a patient's body because different tissues absorb the radiation differently [10].

- Process:

A beam of X-rays is emitted by the X-ray machine and is focused on the bodily portion being scanned. An image receptor detects the X-rays as they pass through the body and records them after they have done so. The resulting image demonstrates the varying tissue densities, with bones looking white and soft tissues showing in a range of gray tones. [10]

- Uses:

X-rays are frequently used in medical imaging to identify and track a variety of illnesses, including dental issues, lung infections, and fractured bones. High-energy X-rays are also employed in the treatment of cancer because they can kill malignant cells. [10].

- Advantages:

X-rays are a rapid and non-invasive diagnostic and monitoring tool for a variety of illnesses. Additionally, they are readily accessible and reasonably priced. [10].

- Limitations:

Ionizing radiation, which is used in X-rays, has the potential to harm cells and raise cancer risk. Children and pregnant women are especially vulnerable to the effects of ionizing radiation, thus care must be made to minimize their exposure. Additionally, X-rays are not very good at differentiating between different kinds of soft tissues, which makes it challenging to find problems like tiny injuries or early-stage malignancies. [10].

Figure 4: X-ray image of the human chest

1.6.2. Computed tomography (CT):

- Definition:

Computed tomography (CT) is a medical imaging technique that utilizes X-ray technology and computer processing to produce detailed cross-sectional images of the internal structures of the body.

- Process:

In a CT scan, the patient is positioned on a table that moves through a circular machine known as a CT scanner. The scanner emits X-ray beams, which are detected by sensors located on the opposite side of the machine. These sensors gather information that is then analyzed by a computer to produce comprehensive cross-sectional images of the body. [11].

- Uses:

CT scans are frequently employed to diagnose and track a variety of medical ailments, including cancer, heart disease, bone conditions, and injuries. Additionally, they can assist in directing medical procedures like surgeries and biopsies. [11]

- Advantages:

CT scans provide highly detailed images of the body, allowing doctors to see internal structures and organs more clearly than with traditional X-rays. They are also a relatively fast and non-invasive imaging technique that can be used on different parts of the body. [11].

- Limitations:

One of the main limitations of CT scans is that they involve exposure to ionizing radiation, which can potentially increase the risk of cancer over time, especially with repeated scans. Additionally, CT scans can be expensive, and some patients may experience discomfort from lying still on the table during the scan. Finally, CT scans may not be appropriate for certain patient groups, such as pregnant women, due to the potential risks of radiation exposure to thedevelopingfetus.

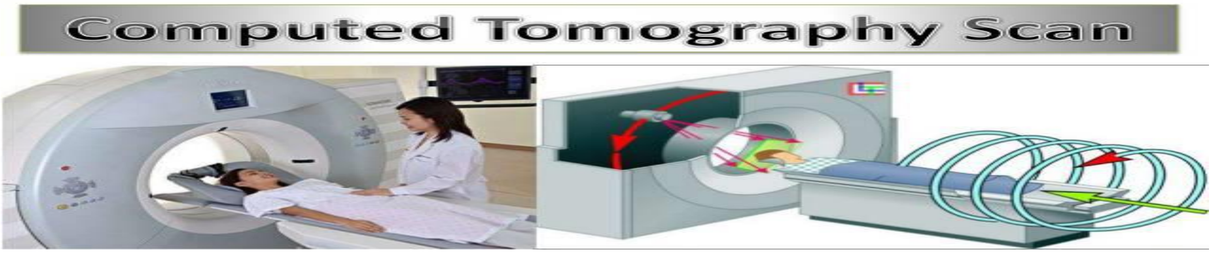


Figure 5: CT Scan

1.6.3. Magnetic resonance imaging (MRI):

- Definition:

Magnetic resonance imaging (MRI) is a medical imaging technique that employs a potent magnetic field, radio waves, and computer analysis to generate comprehensive images of the internal structures of the body. [12].

- Process:

In an MRI scan, the patient is positioned on a table that is transferred into a sizable tube-shaped apparatus. The apparatus produces a powerful magnetic field that causes the hydrogen atoms in the body's tissues to align. Radio waves are then utilized to temporarily disturb this alignment, causing the hydrogen atoms to release signals that are picked up by sensors within the apparatus. The signals are analyzed by a computer to produce extremely intricate images of the body.. [12].

- Uses:

MRI scans are frequently utilized to diagnose and monitor a diverse array of medical ailments, including neurological conditions, joint traumas, and cancer. Additionally, they can assist in directing medical procedures like surgeries and biopsies.

- Advantages:

A significant benefit of MRI scans is that they offer extremely intricate images of the body without exposing the patient to ionizing radiation, which could potentially heighten the risk of cancer. MRI scans can also supply information about the body's soft tissues, which is not feasible with customary X-rays or CT scans. Furthermore, MRI scans can be utilized on various parts of the body and can be customized to meet the specific requirements of each patient. [12].

- Limitations:

One of the main limitations of MRI scans is that they can be expensive and time-consuming. Patients must lie still for an extended period of time, which can be uncomfortable for some individuals. MRI scans may also not be appropriate for certain patient groups, such as those with metallic implants or pacemakers, due to the strong magnetic field generated by the machine. Finally, MRI scans may not be as effective at detecting some conditions as other imaging techniques, such as CT scans. [12].

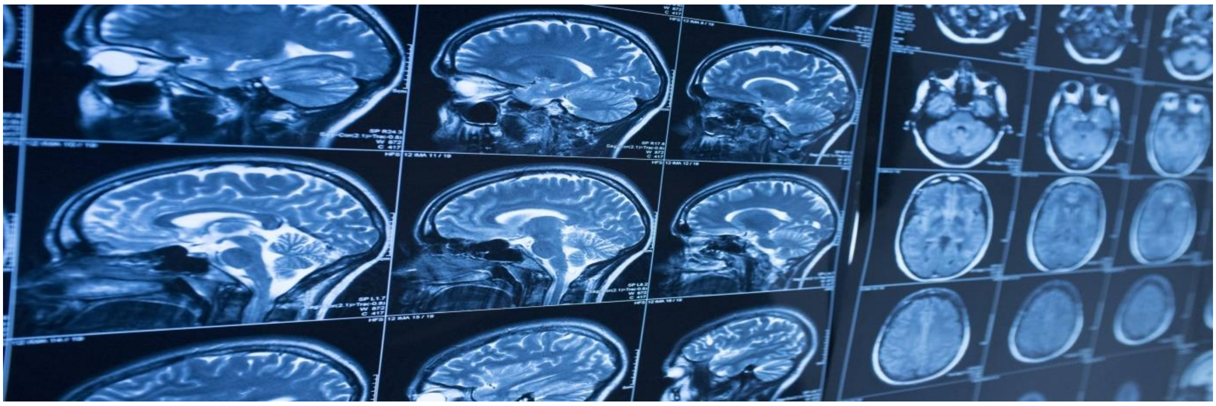


Figure 6: Results of magnetic resonance imaging of the human head (MRI)

1.6.4. Ultrasound imaging:

- Definition:

Ultrasound imaging, which is also referred to as sonography, is a medical procedure that employs high-frequency sound waves to generate pictures of the internal organs and tissues of the body.

- Process:

When conducting an ultrasound scan, a technician will first apply a unique gel to the skin over the region of interest. Next, they will utilize a handheld device known as a transducer to emit high-frequency sound waves over the gel. These sound waves will bounce off the body's tissues and organs, and the transducer will detect the echoes produced. Finally, a computer will process the echoes to create images of the body's internal structures.

- Uses:

Ultrasound scans have a broad range of medical applications, which include monitoring pregnancies, diagnosing and tracking heart, liver, kidney, and other organ conditions, as well as assisting in the guidance of biopsies and other medical procedures. [13].

- Advantages:

An essential benefit of ultrasound scans is that they are non-intrusive and do not require exposure to ionizing radiation. They are also relatively low-cost and easily accessible. Moreover, ultrasound scans can be conducted in real-time, enabling physicians to observe the body's organs and tissues in motion. [13].

- Limitations:

A significant drawback of ultrasound scans is that they might not generate images of the body that are as comprehensive as those produced by other imaging methods like MRI or CT scans. Besides, ultrasound scans may not be useful in diagnosing conditions that are situated deep within the body or that affect the bone. Lastly, ultrasound scans may be reliant on the operator's skills and experience, which can result in variations in the quality of the images produced. [13].

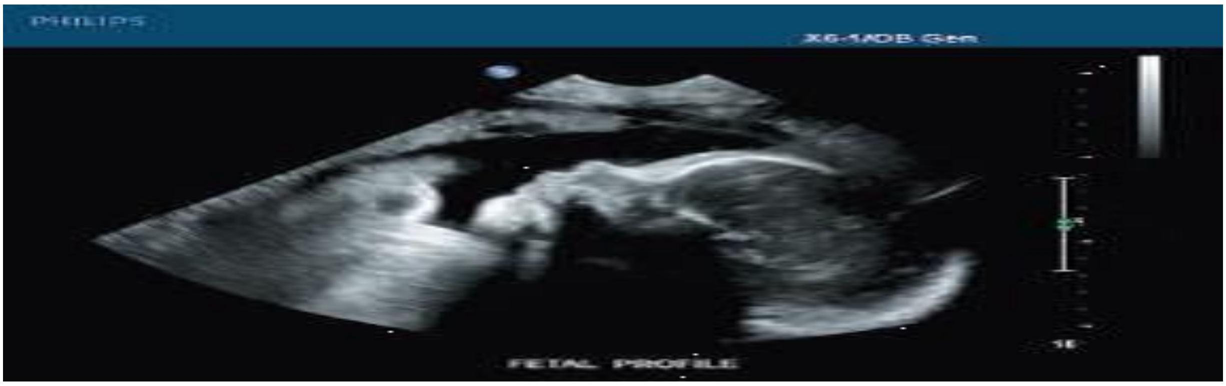


Figure 7: Ultrasound image of a fetus

1.6.5. Nuclear medicine imaging:

- Definition:

Nuclear medicine imaging is a medical technique that employs minute quantities of radioactive substances to develop pictures of the body's tissues and organs.. [14]

- Process:

In nuclear medicine imaging, a minute dosage of a radioactive material, referred to as a radiopharmaceutical, is administered to the patient through injection, inhalation, or ingestion. The radiopharmaceutical travels to the targeted organ or tissue and releases gamma rays in the form of radiation. A gamma camera, a specialized device, captures pictures of the radiation emitted by the radiopharmaceutical. [14]

- Uses:

Nuclear medicine imaging has diverse applications in diagnosing and tracking various medical disorders, like cancer, heart disease, and neurological ailments. It can also be utilized to evaluate organ function and blood circulation, as well as to direct medical procedures such as surgeries and biopsies. [14]

- Advantages:

A significant benefit of nuclear medicine imaging is its ability to offer insights into the function and metabolism of the body's organs and tissues, which cannot be achieved through other imaging methods such as MRI or CT scans. Furthermore, nuclear medicine imaging is a reasonably safe and non-intrusive imaging technique that can be used repeatedly without posing any threat to the patient. [14]

- Limitations:

A primary drawback of nuclear medicine imaging is that it entails exposure to minimal amounts of radiation, which can potentially raise the risk of cancer over time, particularly with repeated scans. Furthermore, nuclear medicine imaging can be a costly and time-consuming process. Lastly, the images generated by nuclear medicine imaging may not be as intricate as those produced by other imaging methods, like MRI or CT scans. [14]

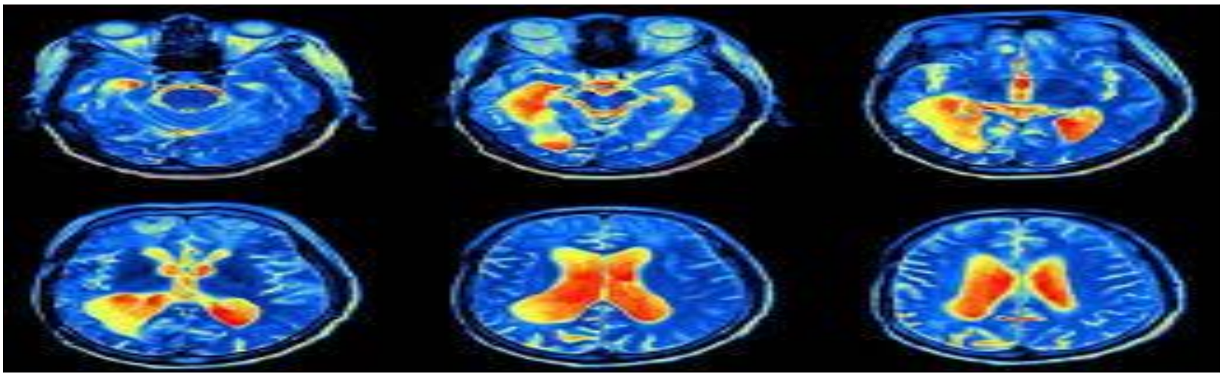


Figure 8: Human head section with nuclear medicine imaging

1.6.6. Optical imaging:

- Definition:

Optical imaging is a medical imaging technique that uses light to produce images of the inside of the body. It involves shining light into the body and detecting the light that is reflected or transmitted through the tissue. [15]

- Process:

Optical imaging can be performed using a variety of methods, including optical coherence tomography (OCT), diffuse optical tomography (DOT), and fluorescence imaging. In OCT, a light source is used to create detailed images of tissue layers, similar to ultrasound. In DOT, near-infrared light is used to create 3D images of tissue that can be used to detect tumors or blood flow changes. In fluorescence imaging, fluorescent dyes are used to make cancer cells or other targets visible under special light. [15]

- Uses:

Optical imaging has a variety of applications in medicine, including cancer detection and treatment, ophthalmology, and dermatology. It can also be used to monitor brain activity and detect changes in blood flow. [15]

- Advantages:

A significant benefit of optical imaging is that it is non-intrusive and does not require exposure to ionizing radiation, making it a safer substitute for other imaging methods such as X-ray, CT, or PET. Besides, optical imaging can produce intricate images of tissue structure and function, making it a valuable resource for diagnosis and treatment planning. [15]

- Limitations:

One of the main limitations of optical imaging is that it is limited to imaging tissues close to the body's surface, and is not well-suited for imaging deeper tissues or organs. Additionally, optical imaging can be affected by factors such as tissue absorption and scattering, making it less accurate in certain situations. Finally, optical imaging devices can be expensive and require highly trained operators, making it less accessible in some settings. [15]

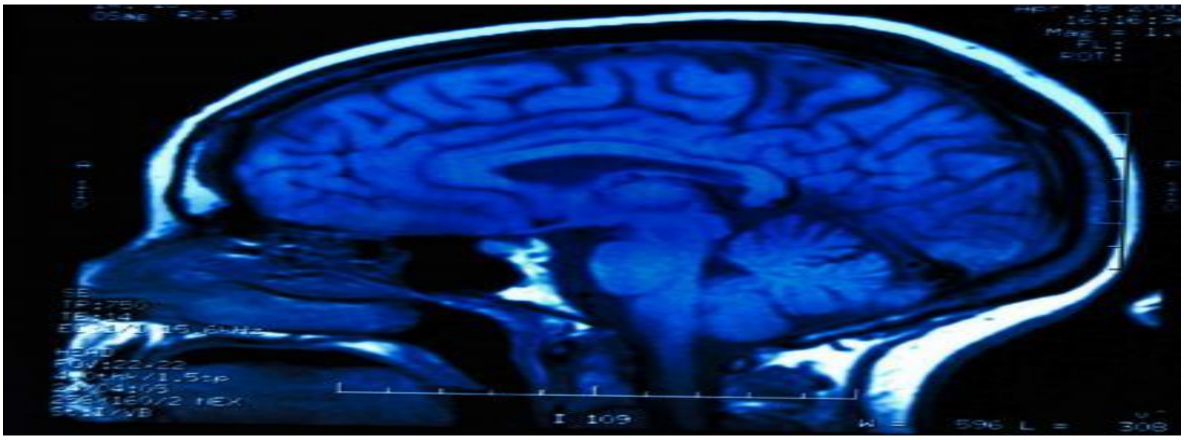


Figure 9: Human head visual imaging

1.6.7. Positron Emission Tomography (PET) imaging:

- Definition:

Positron Emission Tomography (PET) imaging is a medical imaging method that employs a tracer, a radioactive substance, to visualize metabolic processes in the body. [16]

- Process:

In a PET scan, a patient is administered a small quantity of a tracer substance that emits positively charged particles known as positrons. As the tracer substance circulates throughout the body, it releases positrons that bind with electrons in the body, generating gamma rays. These gamma rays are captured by a specialized camera known as a PET scanner, which generates images of the body's metabolic processes. [16]

- Uses:

PET imaging is a versatile medical technique that can identify and track various medical disorders, such as cancer, heart disease, and neurological ailments. It can also evaluate organ function, blood circulation, and the efficacy of medical treatments.

- Advantages:

One of the main advantages of PET imaging is that it can provide detailed information about the metabolic activity of tissues and organs, which can be used to detect disease at an early stage and monitor disease progression. Additionally, PET imaging can be used to guide treatment decisions and evaluate the effectiveness of therapy. [16]

- Limitations:

One of the main limitations of PET imaging is that it involves exposure to ionizing radiation, which can potentially increase the risk of cancer over time, especially with repeated scans. Additionally, PET imaging is relatively expensive and may not be widely available in all healthcare settings. Finally, the images produced by PET imaging may not be as detailed as those produced by other imaging techniques, such as MRI or CT scans. [16]

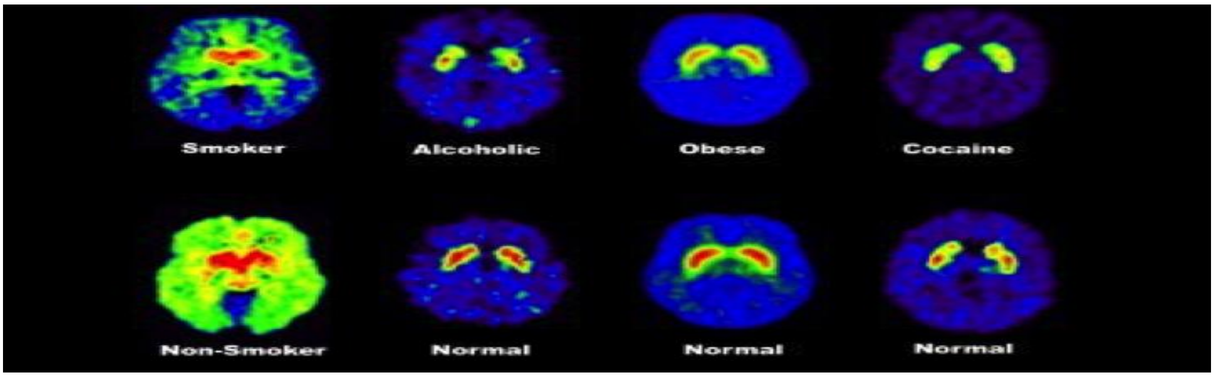


Figure 10: PET scan result

1.6.7.Single Photon Emission Computed Tomography (SPECT) imaging:

- Definition:

Single Photon Emission Computed Tomography (SPECT) imaging is a medical imaging method that employs a radioactive substance to produce visuals of the internal organs and tissues of the body. [17]

- Process:

In a SPECT scan, a patient is given a small dose of a radioactive substance, known as a tracer. The tracer emits gamma rays, which are identified by a specialized camera called a gamma camera. The camera revolves around the patient, capturing numerous images from varying perspectives. These images are subsequently reconstructed by a computer to generate a three-dimensional image of the body's internal organs and tissues. [17]

- Uses:

Heart disease, neurological diseases, and cancer are just a few of the many medical illnesses that may be detected and monitored with SPECT imaging. Additionally, it is utilized to assess how well a treatment is working and how blood flow and organ function.[17]

- Advantages:

The ability to offer extensive information on the function of organs and tissues is one of the key benefits of SPECT imaging. This information may be utilized to diagnose disease early on and track the course of disease. SPECT imaging is furthermore more affordable than other imaging modalities like MRI or PET, making it more available in various healthcare settings.

- Limitations:

Ionizing radiation exposure is one of the primary drawbacks of SPECT imaging, as it may eventually raise the chance of developing cancer, especially with repeated scans. Additionally, SPECT imaging pictures might not be as detailed as those generated by other imaging methods like MRI or CT scans. Finally, radioactive tracers are needed for SPECT imaging, which might be costly and not always be readily available in medical facilities. [17]

Figure 11: Single Photon Emission Tomography (SPECT):

2. Medical image processing:

Medical image processing refers to the application of various techniques to improve the quality of medical images and extract relevant information from them. The process typically involves several stages, including image acquisition, pre-processing, analysis, segmentation, and interpretation.

Image acquisition involves capturing the image using a camera or other imaging device. Pre-processing includes several operations such as binarization, localization, segmentation, noise elimination, and normalization, which are aimed at improving the quality of the image and making it easier to segment.

Analysis involves the extraction of relevant information from the image, such as the fundamental frequency, harmonics, energy, occlusions, concavities, and contours. This information can be used to identify and diagnose medical conditions.

Segmentation is the process of partitioning the image into subsets or regions, which can then be analyzed further. Segmentation can be used to identify tumors, organs, blood vessels, and other structures in the image.

Interpretation involves classifying the image based on the information extracted during analysis and segmentation. This can involve learning and decision-making algorithms, as well as combining and fusing multiple images to provide a more accurate diagnosis.

Overall, medical image processing plays a critical role in the diagnosis and treatment of a wide range of medical conditions, from cancer and heart disease to neurological disorders and musculoskeletal injuries. By improving the quality of medical images and extracting relevant information from them, this field has the potential to revolutionize the way that healthcare is delivered and improve patient outcomes.

Conclusion

In this chapter we have discussed the general notions of the image and its characteristics, and in particular the notion of the digital image. We have defined medical imaging as all the techniques making it possible to maintain a picture of an organ or a portion of the human body with the intention of performing a diagnostic or directing a therapeutic action. such as a puncture, or to evaluate a treatment's progress over time. Then we quoted some methods used in the field of imaging such as: radiology, ultrasound but we were particularly interested in MRI and CT on the one hand because of their digitization and their considerable progress in quality of images and in terms of precision and on the other hand for the availability of their images segmented manually by the experts.

Chaptre 2

Image Registration

Introduction

Image registration is the process of aligning two or more images to a common coordinate system. This is typically done by identifying corresponding points or features in the images and computing a transformation that maps one image onto the other. The transformation can be rigid, affine, or non-rigid, depending on the degree of deformation between the images.

Image registration is an important technique in many fields, including medical imaging, remote sensing, computer vision, and astronomy. In medical imaging, it is used to combine images from multiple modalities (such as CT and MRI) for more accurate diagnosis and treatment planning. In remote sensing, it is used to create maps and monitor environmental changes over time. In computer vision, it is used for object tracking, recognition, and 3D reconstruction. And in astronomy, it is used to combine images from different telescopes to create high-resolution images of the universe.

Overall, image registration is a powerful tool that allows us to combine and analyze images from multiple sources, providing us with a deeper understanding of the world around us.

1.Objective:

Image registration, also known as image alignment or image enrollment, is the process of finding a transformation that can spatially align two or more images, typically acquired at different times, from different viewpoints, or under different imaging modalities. The goal of image registration is to bring the images into a common coordinate system, so that corresponding features in the images can be compared and analyzed.

The process of image registration involves identifying corresponding features in the images, such as points, edges, or regions, and then finding a mathematical transformation that can map the features from one image to the other. This transformation can be rigid, meaning that it only involves translation, rotation, and scaling, or non-rigid, meaning that it involves more complex transformations such as warping, bending, or stretching.

In medical imaging, image registration plays a crucial role in diagnosis, treatment planning, and monitoring of diseases. For example, in radiation therapy, image registration is used to align pre-treatment and post-treatment images to ensure that the radiation dose is delivered to the target area while minimizing damage to surrounding healthy tissues. In neuroimaging, image registration is used to map brain structures across individuals, time points, or imaging modalities to better understand brain function and disease.

Image registration can be a challenging task, especially when dealing with complex and deformable structures, such as organs or tissues that change shape or size over time. To overcome these challenges, a wide range of registration techniques and algorithms have been developed, including intensity-based methods, feature-based methods, and deformable models. These methods rely on different assumptions, constraints, and optimization strategies, and can be tailored to specific applications and imaging modalities.

Overall, image registration is a fundamental problem in image processing and computer vision, with applications in diverse fields such as medical imaging, remote sensing, robotics, and computer graphics. [18]

Schematically, we can interpret the registration as shown in the following Figure 1 :

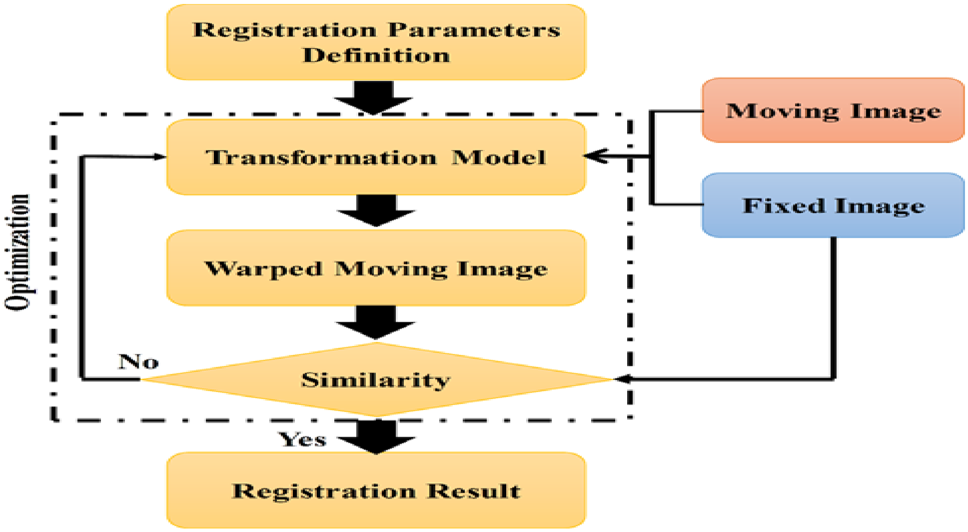


Figure 1: image registration framework flowchart for the medical images

2. Registration applications :

2.1. The stereovision :

Stereo vision is an important application of image registration in computer vision, which involves using multiple photos of the same scene taken from various angles to create a 3D representation of the scene. Image registration plays a crucial role in stereo vision by enabling accurate alignment of the images and subsequent reconstruction of the scene.

In stereo vision, image registration involves finding matching points between two or more images of the same scene, which allows for the computation of the 3D structure of the scene. This is typically done by identifying and matching

features, such as corners or edges, in the different images, and then applying mathematical transformations to align the images.

One common technique used in stereo vision for image registration is called the normalized cross-correlation method, which involves comparing the intensity values of each pixel in one image with the corresponding pixels in another image. By computing the correlation between the two images, it is possible to identify and match corresponding points, even when The photographs were taken from different viewpoints or in various orientations.

Other techniques used in stereo vision for image registration include the use of epipolar geometry, which involves constraining the search for corresponding points to a specific line in each image, and the use of feature-based methods, which involve identifying and matching distinctive features in the images.

Overall, image registration is a crucial step in stereo vision, enabling accurate alignment of images and subsequent reconstruction of the 3D structure of a scene. It has a wide range of applications in fields such as robotics, autonomous navigation, and virtual reality. [19]

2.2.Indexing:

In image registration, indexing is a crucial step that involves assigning unique identifiers to images based on their content or features. These identifiers are then used to match corresponding features in different images, allowing for accurate alignment and registration of the images.

Indexing in image registration typically involves identifying key features or landmarks in images, such as corners, edges, or other distinctive shapes. These features are then assigned Scale-invariant feature transforms (SIFT) and SURF (Speeded-Up Robust Features) are examples of distinctive descriptors that may be used to locate and match similar features across various pictures.The matching of corresponding features in different images is an important step in image registration, as it allows for accurate alignment of the images even when they are distorted or have different orientations. Once corresponding features have been identified and matched, mathematical transformations can be applied to align the images, allowing for accurate analysis and comparison.

Overall, indexing is an essential component of image registration, enabling accurate and efficient alignment of images. By assigning unique identifiers to images based on their content, it allows for the matching of corresponding features and the accurate alignment of images, with applications in fields such as medical imaging, remote sensing, and computer vision. [19]

2.3. The medicine :

Medical image registration is a crucial technique that finds its application in a wide range of medical image analysis tasks. It is considered to be the most important field of application with regard to publications. The main objective of medical image registration is to match similar anatomical structures in order to find the transformation that separates them. This technique has numerous potential applications in the medical field,[19] including:

1. Treatment planning: Medical image registration enables precise targeting and planning of treatment for different medical conditions by aligning pre-operative images with real-time images during surgery.
2. Disease progression monitoring: Image registration helps track changes in anatomical structures over time, allowing for early detection and monitoring of disease progression.
3. Image-guided interventions: Medical image registration assists in guiding minimally invasive procedures, such as needle biopsy and catheterization, providing accurate targeting of the affected area.
4. Multimodal image fusion: Medical image registration enables the fusion of multiple types of medical images, such as MRI, CT, and PET scans, resulting in a comprehensive view of the affected area.
5. Brain connectivity analysis: Image registration plays a vital role in aligning and comparing images of the brain, leading to the identification of regions of the brain that are involved in specific functions or affected by disease.

3. The image registration problem :

3.1. General principle of a registration system :

This summary is a very accurate description of the steps involved in the image registration process. Here is a more detailed explanation of each step:

Step 1: Attribute extraction

In this step, features or attributes are extracted from the IMS and IMC images that can be used to guide the registration process. The goal is to identify salient points or regions in the images that can be used as landmarks or control points. This is done using preprocessing functions FS and FC, which can include techniques such as edge detection, corner detection, or feature extraction. These functions can help reduce the complexity of the images and eliminate noise, making it easier to identify the relevant features.[20]

Step 2: Transformation estimation

Once the features have been extracted from the images, the next step is to estimate the transformation that maps the points in the source image to their corresponding points in the target image. This transformation is represented by the displacement field u, which can be a rigid, affine, or non-rigid transformation depending on the application. The search space E is defined as the set of all possible transformations that can be used to find the desired transformation.[20]

Step 3: Cost function definition

The goal of this step is to define a cost function that measures the similarity or dissimilarity between the transformed source image and the target image. The cost function C can be based on different criteria such as intensity, gradient, or texture. The optimal value of the cost function is achieved when the transformed source image and the target image are perfectly aligned.[20]

Step 4: Optimization

In the final step, an optimization algorithm is used to find the optimal transformation that minimizes (or maximizes) the cost function C over the search space E. This involves searching through the search space to find the transformation that produces the best match between the images. This can be done using techniques such as gradient descent, genetic algorithms, or simulated annealing. The end result is a transformation that can be applied to the source image to align it with the target image. [20]and the optimization problem can thus be formulated as follows:

$$\boxed{\hat{T} = \arg\min_{T \in E} C(I_C, T(I_S))}$$

3.2.The attributes:

In image registration, there are several attributes that can be used to estimate the transformation between two images. Here are some common attributes:

Intensity-based approach:

This approach uses the pixel intensity values of the images to estimate the transformation. It relies on the assumption that corresponding pixels in the two images have similar intensity values. Examples of intensity-based methods include correlation-based methods, mutual information-based methods, and normalized cross-correlation.

Geometry-based approach:

This approach uses geometric primitives such as lines, circles, or contours to estimate the transformation. It relies on the assumption that corresponding geometric features in the two images have similar properties, such as orientation or shape. Examples of geometry-based methods include point-based registration, feature-based registration, and contour-based registration.

Hybrid approach:

This approach combines both intensity and geometry information to estimate the transformation. It relies on the assumption that corresponding pixels and geometric features in the two images have similar properties. Examples of hybrid methods include mutual information-based registration with feature extraction, and edge-based registration with intensity correlation.

In practice, the choice of the attribute depends on the specific application and the nature of the images being registered. Some attributes may be more suitable for certain types of images, while others may be more robust to noise or artifacts in the images. A hybrid approach may also be useful in cases where no single attribute is sufficient to accurately estimate the transformation.[21]

- **Similarity criteria:**

The criterion of similarity in image registration is critical for accurately aligning images. One limitation of many traditional similarity measures, such as those based on pixel intensity, is their inability to handle large changes in brightness between images. This is especially problematic in the registration of retinal images, which often exhibit significant brightness variations due to factors such as illumination and disease.

To address this issue, new similarity evaluation criteria have been developed that use edge point detection. Instead of relying solely on pixel intensity, these methods detect edge points in the

images and use them to assess the correspondence between image areas. By focusing on edges, which are more robust to changes in brightness, these methods can produce more accurate registration results for retinal images and other challenging datasets.
A new similarity evaluation criterion based on edge point detection is presented. The process involves detecting edges in both images, and then identifying corresponding edge points between the two images. The similarity between corresponding edge points is then used to evaluate the overall similarity between the images. Two prepared registration algorithms are used to compensate for any geometric misalignments between the images, using the similarity measure generated by the edge point detection process.
Overall, the use of edge-based similarity measures can significantly improve the accuracy and robustness of image registration, especially in challenging datasets with significant brightness variations.[21]

- **Transformations considered:**

In image registration, the mathematical relationships that allow the alignment of two images are called transformations. There are two main types of transformations: linear and non-linear. Linear transformations are geometric transformations that can be represented by a polynomial of degree 1. Some examples of linear transformations include Thin-Plate Splines and B-spline combinations. Linear transformations are often referred to in the image processing literature as elastics, deformable, or not rigid.
Non-linear transformations are more complex and allow for more flexible transformations. They can be represented by polynomials of degree higher than 1 or by other non-linear functions. Examples of non-linear transformations include Free-Form Deformations (FFD) and Demons algorithm. Non-linear transformations are better suited for handling complex deformations in images.
Transformations can also be classified based on their application to the entire image or only to specific parts of it. A global transformation applies to the entire image, while a local transformation divides the image into parts, and the transformation is applied to each part separately. Local transformations are useful for handling images with complex structures or local deformations, while global transformations are suitable for images with overall similarity in structure.
Overall, the choice of transformation depends on the nature of the images to be aligned and the level of accuracy required. Linear transformations are faster and simpler to compute but are limited in their ability to handle complex deformations, while non-linear transformations are more powerful but can be computationally expensive.[21]

- **Optimization method:**

The optimization method is an important aspect of non-rigid registration techniques for medical images. One popular approach involves maximizing the mutual information of the two images, which requires a deformation field parameterized by cubic B-splines. The process involves an iterative optimization procedure to find the coordinate mapping between the two images.
In this context, eight different optimization methods were compared in terms of their performance, including Gradient descent using two alternative techniques for choosing the step size, quasi-Newton, nonlinear conjugate gradient, Kiefer-Wolf witz, simultaneous perturbation,

Robbins-Monro, and strategy devolution. The goal of the study was to shorten calculation time by computing the cost function and its derivatives on fewer voxels.
To test the optimization methods, manually distorted CT images of the heart, follow-up chest CT scans, and MRIs of the prostate acquired using different protocols were used. The registration accuracy was evaluated by calculating the overlap of segmented edges, and the strain fields were compared to study accuracy and convergence properties.
As a consequence of employing a tiny portion of the picture to compute the derivative of mutual information, the calculation time per iteration was lowered by almost 500 times without compromising the convergence rate, proving that the Robbins-Monro approach was the best option in most cases. Although they needed more computing time, the quasi-Newton and nonlinear conjugate gradient approaches produced somewhat greater accuracy..[21]

3.2.1. Geometric approaches:

The geometric approach to image registration involves identifying and matching geometric primitives in two images, such as points, curves, surfaces, and volumes. The registration process consists of two main steps: segmentation and matching. In segmentation, the images are divided into homogeneous primitives, which can be done manually or automatically using differential constants such as angles and curves. In matching, the primitives in the two images are compared and matched using various methods such as reducing Euclidean distance or using a distance map. The quality of the registration depends on the number and accuracy of the extracted primitives, which in turn relies on the quality of the segmentation. The geometric approach may not be as effective for low resolution images or multimodal registration,such as with PET and SPECT images.[20]

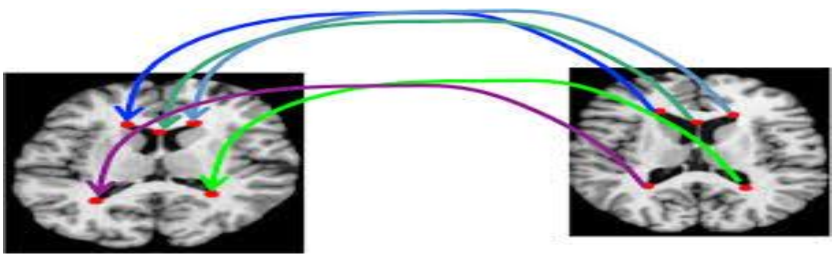


Figure 2: shows the smooth points of the two images

3.2.2. Iconic Approaches:

The iconic style directly depends on the density of the images. Unlike Geometric Approach, the object or hash geometry understanding step is no longer necessary. This approach mainly consists of refining the similarity criterion. Solely on the basis of density comparisons. The low-level method is another Name the Iconic approach. Therefore, this approach is more suitable for Multimedia registration. Thus, it is from the intensity of the images that the objective function (or cost function) is constructed. Then, reset synchronization is implemented by optimizing this function. Thanks to the success of algorithms based on mutual information, this approach has been restored to interest over the past fifteen years. Moreover, one can find many variations of this approach. Restricted with matching and settlement by settlement.[20]

- **Criterion of similarity:**

This is the information that makes it possible to calculate the accuracy of the registration.
It is a question of associating it with a criterion of similarity making it possible to define a certain distance between the two images. These functions can be classified into two categories:

1. **Geometric Category:**

The primitives to be considered are set of points. Several methods can be used to estimate the distance between 2 geometric primitives[23]:

- Sum of Squared Differences (SSD): $ST = \sum (I\ x\ y - J\ (T\ x\ y))\ 2$
- Sum of differences in absolute value (SAD): st= $\sum_{x,y} \mid I\ x\ y - J\ (T\ x\ y) \mid$
- Mutual information: $I\ m\ (x,y) = \sum_{x,y} p(x,y) \log(\frac{p(x,y)}{p(x)p(y)})$
- Sum of Squared Differences (SSD): This method calculates the sum of the squares of the differences between the pixel intensities of two images. It is used to measure the similarity between two images and is often used in keypoint matching algorithms.
- Sum of Absolute Value Differences (SAD): This method is similar to the SSD method, but it calculates the total difference in pixel brightness between two pictures, measured in absolute terms. It is frequently used in keypoint matching algorithms and is also used to calculate the degree of similarity between two photos.
- Mutual Information: This method is used to measure the correlation between two images by comparing the probability distributions of their pixels. It is often used for edge detection and image segmentation.

2. **Iconic category:**

In the context of this category, it is necessary to compare the intensity between the two reference and source images to obtain the degree of similarity which can be calculated by the relations:

Normalized correlation: $N\ C\ (i\ ,\ j\ ) = \frac{\sum_I^N IiJj}{\sqrt{\sum_I^N I^2 J^2}}$

3. **The concept of joint histogram:**

$f\ (i1\ ,i2) = card\{(x,y \mid I1(x,y)= i1 \wedge I2(x,y)=i2\}$

3.3. Nature of the transformations:

To align two images, it is essential to know which transformation is used to link a Two Images file. As such, we distinguish two classes of Transformations:

3.3.1.Linear transformations:

There is only translation and rotation.[19]

- **Rigid Transformation:**

  A Rigid Transformation is a type of geometric transformation that preserves distances and angles. It includes rotations, translations, and reflections.

  The general formula for a Rigid Transformation is:

  (x', y') = R(x, y) + T

  where:

  (x, y) are the original coordinates

  (x', y') are the transformed coordinates

  R is a 2x2 rotation matrix that preserves distances and angles

T is a 2x1 translation vector that moves the coordinates to a new position
A rotation matrix can be represented as:
R = [cos(theta) -sin(theta)]
[sin(theta) cos(theta)]

where:

theta is the angle of rotation
A translation vector can be represented as:
T = [tx][ty]

where:

tx, ty are the translation parameters
A reflection is a special case of Rigid Transformation that flips the coordinate system across a line. It can be represented as:
(x', y') = (x, -y) (reflect across x-axis)
(x', y') = (-x, y) (reflect across y-axis)
(x', y') = (-x, -y) (reflect across origin)

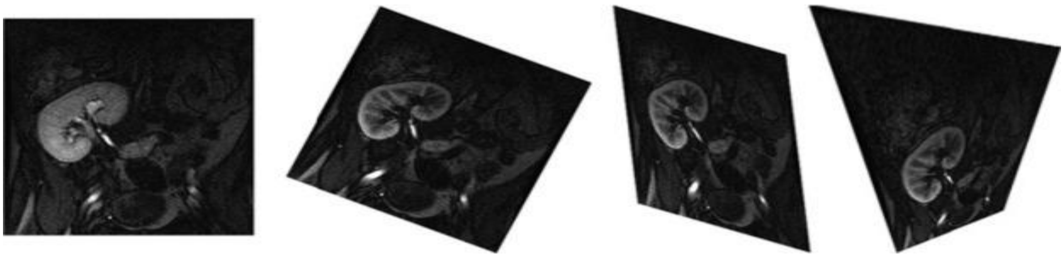


Figure 3: Rigid transformations

We define rigid complex transformation Translation and Rotation only.

Translation:

Translation is a transformation that involves moving objects in the image in one direction by a two-coordinate vector. If a point (,y) undergoes a translation (TTx,TY) its new position (x ,yy) verifies:

x1=x+Tx
y2=y+Ty

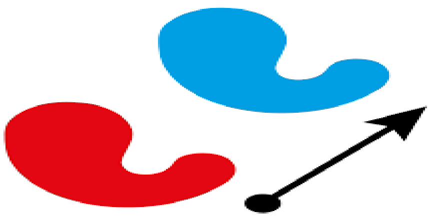


Figure 4: Translation by (TTx,TY)

Rotation:

Rotation is a transformation that rotates objects in the image with an angle $\theta\theta$t around a center (two coordinates) the rotation of angle $\theta\theta$ and centers (0,0), expressed by:
xt=xcoso-ysino

yt=xsino+ycoso

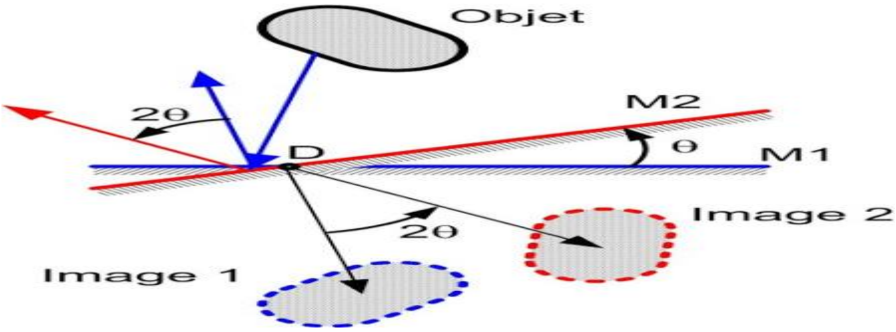


Figure 5: rotation with ongle0

3.3.2. Affine Transformation:

An Affine Transformation is a type of geometric transformation that preserves parallelism and straight lines. It can be represented as a linear transformation followed by a translation.[24]
The general formula for an Affine Transformation is:
(x', y') = (a b c d) (x y) + (tx ty)

where:

(x, y) are the original coordinates
(x', y') are the transformed coordinates
a, b, c, d are scaling, rotation, shearing and flipping parameters tx, ty are the translation parameters .A homothety is a specific type of Affine Transformation, which corresponds to a uniform scaling of the coordinate system. It is represented by the equations:
x' = Sx + tx
y' = Sy + ty

where:

S is the scaling factor
tx, ty are the translation parameters.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

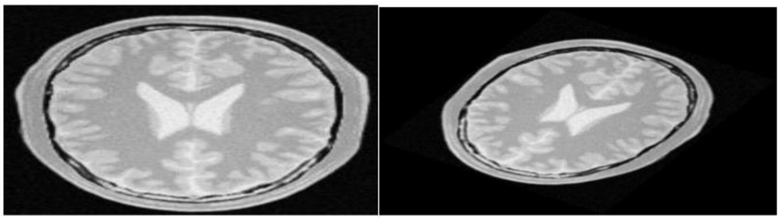


Figure 6: Affine transformations

- **Projective:**

that is another form of the term "projective" that is used in mathematics, specifically in the field of geometry. In this context, the idea of projective refers to the concept of a projective plane or space, which can be used to define a set of points and lines that satisfy certain projective properties. The transformations between these spaces are called projectivities, and they are of fundamental importance in projective geometry.

In this mathematical context, there are certain properties and measures that remain invariant under the action of a projectivity, and these are called projective invariants. Some examples of projective invariants include collinearity, concurrency, tangency, and incidence.
While the study of projective geometry is purely mathematical in nature, there are practical applications of these concepts in fields such as computer vision, where certain types of projectivities, such as perspective projection onto an image plane or planar projection onto another plane, are of particular interest.[25]

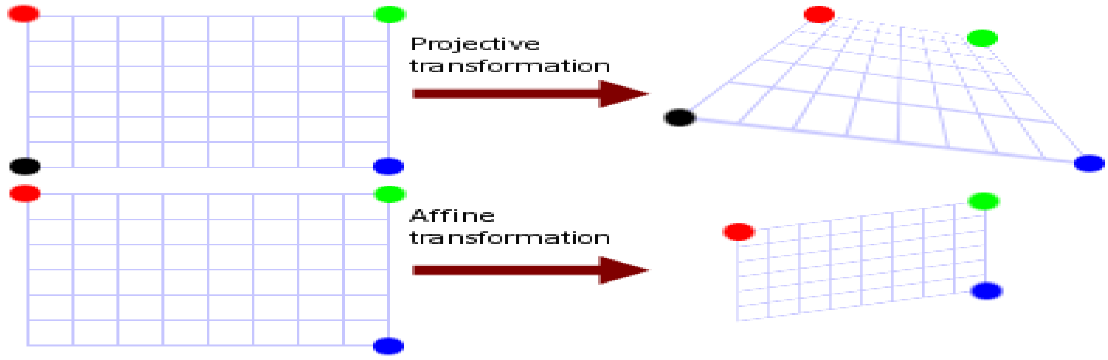


Figure 7: Projecting an image

- **Polynomial transformations (parametric) :**

that is another form of the term "projective" that is used in the context of polynomial transformations. In this context, a polynomial transformation refers to a mathematical function that transforms one set of points in space to another set of points, using a polynomial equation of a fixed order.
The order of the polynomial determines the precision and complexity of the transformation, with higher-order polynomials allowing for more complex transformations but also requiring more control points to define them accurately. In this context, affine transformations are a special case of polynomial transformations where the order of the polynomial is equal to 1, and only six parameters are needed to define the transformation.
Polynomial transformations can be used in a variety of applications, including image and signal processing, computer graphics, and geometric modeling. They are particularly useful in cases where precise and complex transformations are needed, but the underlying geometry is not well understood or defined.[25]

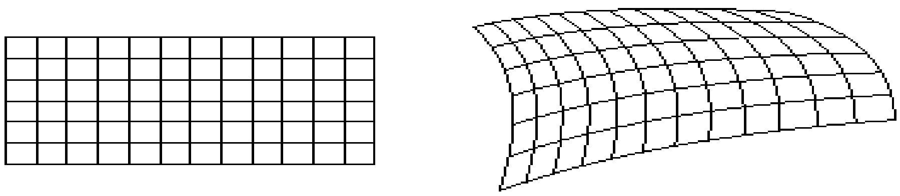


Figure 8: polynomial transformations

3.4. Deformable Transformation:

3.4.1. Elastic model:

by having elastic properties, an object can deform in a well-defined way under the following conditions: The role of external loads. The elastic material returns to its original state when unloaded. The external force applied to the deformable object can be point-like or distributed over the entire object. A straight line in the surface or volume of an object. Therefore, the force can be passed A few points and unknowns. In deformable solid mechanics, four criteria are considered: applied load, stress, deformation and displacement.
There are two main types of elasticity: • linear elasticity, • non-linear elasticity
Linear elasticity mainly involves small strains proportional to stress. Stress is proportional to stress. With the large being Deformation, one can resort to nonlinear elasticity. In addition, one of the characteristics of the material is isotropy. It characterizes The physical properties of the medium vary with the direction. In this study, the image is treated as an isotropic linear elastic material. From Navier-Stokes we have partial differential equations (PDE) [18][26]:

$$\mu\nabla^2 u(x, y, z) + (\lambda + \mu)\nabla(\nabla u(x, y, z)) + g(x, y, z) = 0$$

with f the external force, u the displacement field of a point X; λ and μ the Lamé coefficients

Figure 9: elastic transformations

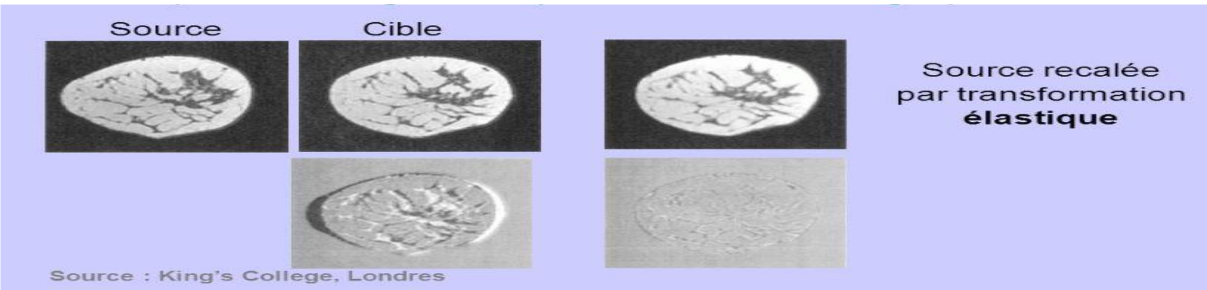


Figure 10: source readjusted by elastic T

4.4.1.1.Déformation :

Under the action of force, non-rigid objects deform in a certain way [27].
The strain tensor of a point is a function of the displacement of the point and is defined by the following formula:

$$\varepsilon = \begin{pmatrix} \varepsilon 11 & \varepsilon 12 & \varepsilon 13 \\ \varepsilon 21 & \varepsilon 22 & \varepsilon 23 \\ \varepsilon 31 & \varepsilon 32 & \varepsilon 33 \end{pmatrix}$$

U (u1, u2, u3) is the displacement of the point X (x1, x2, x3) on the point X ' By construction, we see that the tensor ε is symmetric and the deformation is small, We can write:

$$\varepsilon = \begin{bmatrix} \varepsilon 11 & \varepsilon 12 & \varepsilon 13 \\ \varepsilon 21 & \varepsilon 22 & \varepsilon 23 \\ \varepsilon 31 & \varepsilon 32 & \varepsilon 33 \end{bmatrix} \approx \begin{pmatrix} \varepsilon\, 11 \\ \varepsilon\, 22 \\ \varepsilon\, 33 \\ 2\varepsilon\, 11 \\ 2\varepsilon\, 12 \\ 2\varepsilon\, 13 \end{pmatrix}$$

3.4.1.2. Constraints:

We imagine a virtual cut defined by its normal unit n at point X, we can consider the surface force flux of C(X,n) through the material through this cut point. This process Represents the local stress of the material at the position of this incision

Therefore, the stress at point X can be considered as a characteristic of the force flow. Solicit documents at this stage [27]. The stress tensor can be written:

$$\sigma = \begin{bmatrix} \sigma 11 & \sigma 12 & \sigma 13 \\ \sigma 21 & \sigma 22 & \sigma 23 \\ \sigma 31 & \sigma 32 & \sigma 33 \end{bmatrix}$$

With a voluminal force fi , we have the relation:

$$\text{fi} = \sum_j \frac{\sigma\, i\, j}{x j}$$

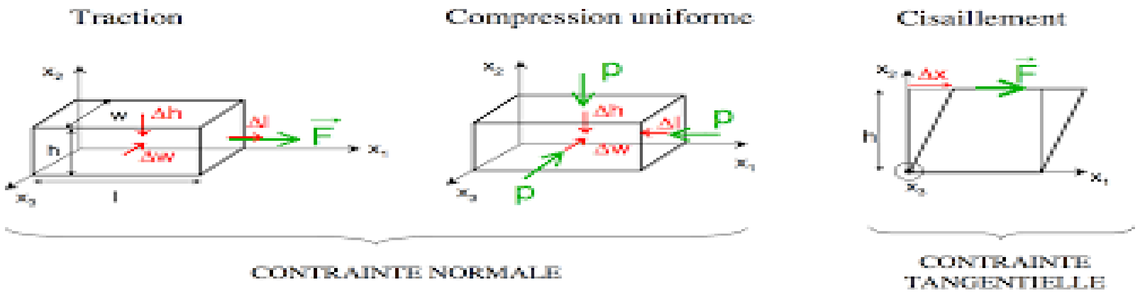


Figure 11: Representation of constraints

3.4.1.3. Relationship between Déformation and Constraints:

According to Hooke's law, in linear and isotropic elasticity, the following relations can be considered: σ = Lε

Moreover, we have the relations:

N $\quad\quad\quad$ $\sigma ij = \lambda tr(\varepsilon)\delta ij + 2\mu\varepsilon ij$

Avec : $\quad\quad\quad$ $tr(\varepsilon) = \varepsilon xx + \varepsilon yy + \varepsilon zz$

Et $\delta ij = \{1(i=j), 0(i\neq j)\}$

3.4.2. Finite element method :

A numerical approach for solving partial differential equations is the finite element method. by approximating the solution of the problem. The continuous domain Ω is divided into several subdomains, called finite elements, and the problem is solved by sections. However, the result varies depending on the type of grid used. The process involves discretizing a complex field, such as an image, using simple geometric elements like triangles, rectangles, tetrahedrons, and hexahedrons. There are two main types of meshes: the uniform mesh and the adaptive mesh. The uniform mesh consists of equally sized elements, whereas the adaptive mesh uses smaller elements in regions where the solution changes rapidly and larger elements in regions where the solution is smooth.[28]

3.4.2.1. The uniform and adaptive mesh:

The image is partitioned into polygons or polyhedra of the same shape, treating all elements as regular collages. This approach ensures that the grid used is not dependent on the information in the image. On the other hand, when an adaptive grid is used, the tiling can be uneven as the construction of the grid depends on the content of the image. In this case, each element in the image provides important and meaningful information, and the quality of the grid determines the quality of the outcome. Although using an adaptive grid can be fun, it takes longer to construct compared to uniform grids. Additionally, when considering material uniformity, adaptive discretization of the mesh is almost useless. Therefore, the study uses a uniform mesh since it is simple to construct and takes a relatively short time to complete.

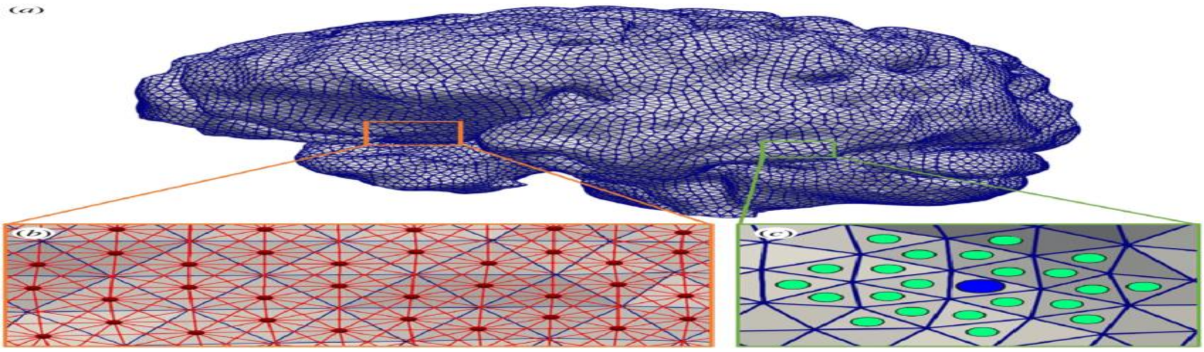


Figure 12: The uniform mesh

By creating a linear system, partial differential equations are solved using the finite element method: K is the global stiffness matrix, and KU = F, U is the displacement field to be determined, and F is the equivalent force. Each finite element is characterized by a specific number of nodes or vertices, which form a discrete set of points used to compute the physical properties of the system. Specifically, the displacement of any point in the image is dependent on the displacement of the node that belongs to the element containing that point. In our study, we consider tetrahedrons as the elements used in the finite element method.[28]

3.4.2.2.Moving a point inside an element:

Note that each point in the image belongs to at least one element. Since we have a tetrahedron, there are four types of point B:

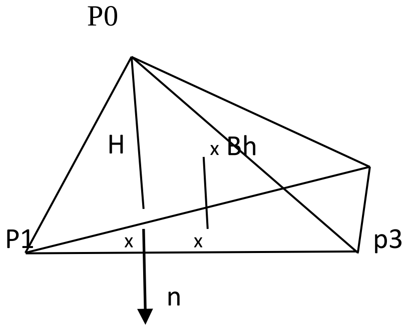
-B is a node. It belongs to all the elements to which the node belongs.Vertex

- Located on the edge. Then, M belongs to all elements with this type of edge, for example border,
- Point B is part of a surface (triangle). Therefore, it belongs to two elements Juxtaposed to this face,
- B is neither a node, nor an edge or a face. It belongs to an element, belongs to Solo

Whatever the type of point B, its displacement can be expressed from a single element. The displacement is given by the following relation:

$$U(B)=\sum_{ii=0..3} NiU\ (pi)$$

with u(Pi) the displacement of the node Pi and Ni is a barycentric coordinate.



n is a unit vector and normal to the face (P1 , P2 , P3) . It is oriented outward. According to the Figure, N0 can be expressed by:

$$N0 = \frac{h}{H}$$

with H is the norm of the orthogonal projection of P1 P0 on n and we have:

$$\overrightarrow{H} = \|\overrightarrow{projn}\text{-}(P0P1)\| = \frac{p0p1.n}{\|n\|^{2}} = p1p0.n$$

$$\overrightarrow{H} = \|\overrightarrow{projn}\text{-}(P1B)\| = \frac{p1b.n}{\|n\|^{2}} = p1B.n$$

Regarding the calculation of n , we know that n is parallel to P1 P2ΛP1 P3 then:

$$n = \frac{P1P2 \wedge P1P3}{\|P1P2 \wedge P1P3\|}$$

$$Ve = \frac{1}{6}\overrightarrow{P1P0}.(\overrightarrow{P1P2} \wedge \overrightarrow{P1P3})$$

then, we can write:

$$N0 = \frac{P1\ B.(\ P1\ P2 \wedge P1\ P3)}{6Ve}$$

3.4.3.Relation between deformations and displacements in an element:

The deformation of each point of the element can be obtained from the Displacement function. From the equations we can use a matrix to express: [ε] = [Be][Ue]

Use the vertex displacement vector [Ue], [Be] gradient matrix [Ne], [Ne] The barycentric coordinate matrix Ni. As a result, we can write:

$$[Ue] = (\ U0\ x\ U1\ x\ U2\ x\ U3\ x\ U0\ y\ U1\ y\ U2\ y\ U3\ y\ U0\ z\ U1\ z\ U2\ z\ U\ 3\ z)$$

- **Stiffness matrix of an element [Ke]:**

The stiffness matrix [Ke] depends on the elastic behavior of the element. It is expressed by:

$$[ke]=\int_{ve}^{t}[Be]\,[De][Be]dV$$

With Ve is the volume of the element, [De] the elasticity matrix. It is the matrix which relates the stress tensor and the strain tensor in the element:

$$[\sigma\,] = [De][\varepsilon\,]$$

3.4.4. Elastic registration:

- **Fixed non-rigid registration issue:**

To solve the non-rigid registration problem, a transformation that aligns two pictures must be found. that are not simply translated or rotated relative to each other. Modeling is one way to solve this issue. the deformation between the images as if they were isotropic linear elastic materials. This means that the behavior of elastomers must be taken into account. In a previous article, we covered the elastic model and the finite element method, which can be used to simulate this behavior.

However, in order to apply these methods to non-rigid registration, we need to specify an external load that acts on the image. In this case, the load is represented by a similar force F. By using the finite element method, we cancalculate the deformation caused by this external load and use it to align the two images.

It should be noted that there are other approaches to non-rigid registration that do not rely on the elastic model and the finite element method. For example, some methods use feature-based approaches, where corresponding points or descriptors are used to find the transformation that minimizes the error between the two images. The choice of method will depend on the specific application and the type of data being analyzed.[29]

3.4.5 Similarity strength F:

As for rigid registration, according to Mattes 2003. Therefore, it constitutes a similar energy. This energy is It is applied to the elastic body or image in the form of external energy. Similar energy produces similar power. This power is enormous, it is established On the gradient of mutual information. At node i, the component of the nodal force fi is x, En y and z. fi is proportional to the gradient component of the similarity, relative to the displacement of node i:

$$\text{Fi}=\begin{bmatrix} fix \\ fix \\ fix \end{bmatrix}= -\text{p}\begin{bmatrix} \frac{\partial \text{s}}{\partial \text{uix}} \\ \frac{\partial \text{s}}{\partial \text{uiy}} \\ \frac{\partial \text{s}}{\partial \text{uiz}} \end{bmatrix}$$

3.4.6.deformable registration:

Deformable image registration establishes dense voxel correlation between two images in order to align a moving image with a reference or fixed image. Deformable registration is far more flexible than rigid or affine transformations and is capable of handling significant deformations. The stationary picture is designated as F, the moving image is designated as M, and the displacement field is designated as in the deformable registration process. The objective is to identify the ideal displacement field that maintains a smooth deformation field while minimizing the difference between the still picture F and the deformed image M. This is an issue that can be solved by optimization:

$$\phi * = \text{argmin } \phi \text{ Lsim}(F, M(\phi)) + \text{Lreg}(\phi)$$

where Lsim(·,·) represents the dissimilarity function and Lreg(·) represents the regularization function that ensures the deformation field is smooth. Typically, affine and scaling transformations are factored out so that the only source of misalignment between images is non-linear.

In practice, deformable registration is commonly used in medical imaging to align scans acquired at different times or using different modalities. The goal is to obtain accurate and consistent information across different scans, which is important for diagnosis and treatment planning. In the experiments, the brain scans were finely registered in the MNI152 space during the preprocessing phase.[30]

3.4.6.1.diffeomorphic registration:

A displacement field u is frequently used in deformable image registration techniques to parameterize the deformable model. For massive and messy deformations, the real inverse transformation of the displacement field could not exist, and the deformable model need not enforce a one-to-one mapping in the transformation. Diffeomorphic deformations are favored in order to get around these restrictions. A time 1 strain, which is a member of a Lie group, is created using this method by exponentiating the strain field, which is represented as a member of the Lie algebra. A stationary velocity field can create the path of diffeomorphic strain fields The velocity field is integrated across time using the scaling and squaring procedure to get the time 1 strain field. It is guaranteed that the resultant time 1 strain field ϕ (1) is diffeomorphic and invertible. [31]

3.4.7.Fluid transformations:

Fluid transformations, or the use of a fluid model of deformation, is suggested as an alternative to the elastic model for estimating large deformations. Instead of treating the image as an elastic material in a Lagrangian framework, it is considered as a fluid in an Eulerian framework. This means that the partial differential equations used are the same as those used in the elastic model, but they act on the speed rather than the displacement. However, one major disadvantage of using fluid transformations is that they tend to be more computationally expensive compared to elastic methods.

3.4.8.Diffusion transformations:

The deformable grid model is utilized to deform the image, with the image being treated as a deformable grid. In this model, the outlines of the objects in the reference image are oriented in a manner that extends from the inside of the object to the outside.

3.5. Optimization methods:

Three steps are typically taken throughout the image registration procedure [Leila M. Fonesca, 1997]. Selecting desirable attributes is the first step. Each of these traits is compared to potential counterpart features in the other image in the following step, and any matched pair is approved as a control point (CP). Finally, a transformation between the pictures is created by modeling the deformation using these CPs. Several algorithms have been suggested and categorized into the following categories [BS Reddy, 1996] in order to automate this process:

(1) Algorithms that directly utilize the pixel values of the image
(2) Algorithms that operate in the frequency domain
(3) Algorithms that employ low-level features such as edges and corners
(4) Algorithms that utilize high-level functionality such as identified objects or features.
After an evaluation of various algorithms, four distinct methods were identified to cover one from each category.

3.5.1-Wavelet-Modulus Maxima method:

Similar to other approaches mentioned before, the Wavelet-Modulus Maxima method proposed in [37] employs picture pixel values to decide feature selection. After completing wavelet decomposition up to two levels, the probable control points are found by identifying the largest local moduli of the wavelet transform applied to both the input and reference pictures. Only the pairwise best fit among all pairs of feature points is taken into account to be an actual control point, and the correlation coefficient is used to quantify similarity. The deformation between the images is then modeled using a polynomial transformer that can accommodate translation and rotation errors, and their parameters are determined in a rough-to-fine manner. Using the deformed image and the collection of distinguishing points found in the image reference, refinement matching is accomplished. The final parameters are chosen and utilized to alter the original input image after all levels have been processed.

3.5.2-Fast Fourier Transform Technique (FFT):

With no need for control points, the Fast Fourier Transform Technique (FFT) is a frequency domain method. Instead, it computes the FFT ratio between the two images that are provided. Calculating F1.conj(F2) / |F1.F2| yields a pulse-like function that is nearly zero everywhere except at the displacement, which indicates the translation error between the frames. The inverse of this ratio yields the displacement between the images. Rotation and scaling errors can alternatively be expressed as offsets by transforming the images from rectangular coordinates to log-polar coordinates and computing a comparable ratio. The image is then geometrically corrected in relation to the reference image using these three parameters to create the mathematical model.[38]

3.5.3. Morphological pyramid image registration algorithm:

While taking into account radiometric changes, the Morphological Pyramid Image Registration Algorithm uses low-level shape features to establish the global affine transformation model between pictures. A morphological pyramid (MP), which has the capacity to omit details and keep form features, is used to depict the photos. The matching parameters of translation, rotation, and scaling errors are then estimated with sub-pixel accuracy using the Levenberg Marquardt nonlinear optimization algorithm. This method combines the geometric mapping function with the intensity mapping function to obtain precise registration.[39]

3.5.4.Image registration using genetic algorithms (GA):

In order to identify the best transformation between two images, the image registration utilizing genetic algorithms (GA) uses a population-based search technique. The process starts by creating a population of potential solutions at random, which are represented as strings of chromosomes. Based on their fitness scores, which are calculated using a similarity measure between the altered image and the reference image, these candidates are assessed. The fittest candidates are then chosen for reproduction, and to produce a new generation of candidate solutions, their genetic material is mixed and modified. Until convergence or the maximum number of generations is reached, this process continues. The benefit of employing GA is its capacity to quickly search across a sizable solution space needed for image registration with sub-pixel accuracy.. The risk that the algorithm will converge to a local rather than a global optimum is the disadvantage, though..[40]

Conclusion

In this chapter, we have provided an overview of the concept of registration and its various applications. We have also discussed the two main registration approaches: geometric and iconic. Additionally, we have examined the different parameters that must be taken into account during registration, including similarity criteria and transformation optimization.

When dealing with images that have an irregular shape, such as the brain, which is composed of multiple materials, an adaptive grid can be used to simulate the real appearance of the human body. However, if the image is modeled as having uniform, linear, and isotropic materials, then a uniform mesh is more suitable for this shape. By understanding the isotropic linear elastic model and the finite element method, we can use them for unregistered rigid registration.

Chapter 3

Artificial intelligence techniques

Chapter 3

Introduction

Artificial intelligence has become a transformative force in today's world, driving advancements across a wide range of fields such as healthcare, sports, food, commerce, and industry. Despite its widespread impact, however, many people still lack a clear understanding of what artificial intelligence is and how it works. In this chapter, we will provide a comprehensive overview of artificial intelligence, including its definition, history, and key technologies.

In particular, we will focus on the contributions of artificial intelligence and its technologies in the field of image processing. Machine learning and deep learning, two of the most important artificial intelligence technologies, have revolutionized the way we analyze and interpret images. We will explore the ways in which these technologies have been applied to a variety of image processing tasks, including image recognition, object detection, and image generation.

By the end of this chapter, you will have a solid understanding of the fundamentals of artificial intelligence and its applications in image processing. Whether you are a researcher, a student, or simply someone interested in the future of technology, this chapter will provide you with the knowledge you need to stay informed and engaged in this rapidly evolving field.

1. Artificial Intelligence :

1.1. Definition :

Artificial intelligence, as defined within the field of computer science, is concerned with the development of computer programs that exhibit behavior and characteristics which mimic those of human cognitive abilities. This can include the ability to learn, reason, and react to novel situations, even in the absence of explicit programming or direction. The overarching goal of artificial intelligence is to create machines that can function autonomously and independently, using their own intelligence to make decisions and take actions. By leveraging advanced algorithms and data-driven approaches, researchers and engineers are working to build AI systems that can perform a wide range of complex tasks with increasing accuracy and efficiency.[41]

1.2. History of artificial intelligence :

In the first half of the twentieth century, some scientists explored a new approach to building intelligent machines, the first of them being the mathematician, computer scientist, logician and cryptanalyst Alan Mathison Turing, who made a device that helped to decipher the signals of the German Enigma machine during the Second World War [42]. Later it was called Turing's machine. This machine proved the reality of the machine's ability to work with mathematical logic without human intervention, according to prior inputs and commands [43]. This stage has been considered the first stage of artificial intelligence.
Alan Turing was not the only one in this field (the development of artificial intelligence), but two scientists: Warren McCulloch, a neurophysiologist, and Walter Pitts, a young mathematician, who published in 1943 a scientific article which speaks of the how the neural network works, they proposed a model for the first artificial neural network, using electrical circuits, and they attribute the basic idea to the artificial neural network we use today [44].
Computer scientists began to apply this idea in their work. In the 1950s, scientist Arthur Samuel created a checkers game program that played better than its developer, and he was the first to coin the term machine learning[45].
The official beginning of the emergence of the term "artificial intelligence" was launched in 1956 by the scientist John McCarthy during a conference held at Dartmouth College in the state of Hanover in the United States of America. Attending this conference were Marvin Minsky, Allen Noel, and Herbert Simon who, along with their students, have been leaders in AI research for decades [46].
The field of artificial intelligence entered a period of great stagnation that extended from 1974 to 1980 due to the low funding of this field. Among the problems he encountered at that time were:

- The processing capacity of the computers was low
- Storage memory was limited
- The amount of data was insufficient due to lack of storage memory.
- There were many difficulties in finding the appropriate mathematical operations to find the optimal solution for certain algorithms.

These major issues and many others have led to the restructuring of multiple areas related to the computer field, and led to work on the development of many important aspects such as the development of computers and artificial intelligence simultaneously.

The success of "expert systems" in the 1980s made it possible to relaunch research projects in the field of artificial intelligence. Like Xcon systems, which have become very popular as specialized systems that simulate the decision-making process like specialized (human) experts, and are able to solve problem ssuch as the diagnosis of infectious diseases [47], thanks to which the profits of research in artificial intelligence on the market reached more than a billion dollars, which motivated various governments to resume financial support for the projects.

In 1987, due to the accumulation of problems with artificial intelligence devices, starting with their cost and the difficulty of updating them, and because consumers no longer needed to buy expensive machines specialized only in the operation of artificial intelligence devices. This led to the collapse of the market for these machines and the decrease in research by scientists. For these reasons, artificial intelligence has experienced another crisis [48].

In 1997, IBM built a chess machine (Deep Blue) that beat world champion Gary Kasparov. This event was the reason for the return of interest in the field of artificial intelligence [49]. Since this event, many works in the field of artificial intelligence began to appear, and the studies in this field increased, and many giant companies such as Facebook and Google joined, so that the intelligence artificial intelligence has become a nascent and necessary field.

2. Machine Learning :

2.1. Definition :

Machine learning is a subfield of artificial intelligence that centers around the creation of algorithms and methodologies which empower computers to learn from data. Its primary objective is to develop models that can detect patterns in data and make predictions or decisions based on that information. Fundamentally, machine learning algorithms aim to enhance their performance over time by assimilating information from prior experiences. This enables computers to execute sophisticated tasks like natural language processing, speech recognition, and image recognition with greater precision and efficiency. The domain of machine learning is swiftly expanding, and it has a wide range of applications in various industries such as healthcare, finance, and transportation.[50]

2.2. Types of machine learning systems :

1. Supervised learning :

In supervised learning, a corresponding label or output value is assigned to each training sample in the dataset. As a result, the algorithm has the ability to anticipate the output labels or values.51]

2. Unsupervised learning :

For train data, there are no labels in unsupervised learning. The machine learning algorithm makes an effort to discover the fundamental properties or distributions in the data..[51]

3. Reinforcement learning :

In reinforcement learning, the algorithm chooses the best course of action to maximize a reward (represented as a number) in order to accomplish a particular goal..[51]

3.The artificial neural networks :

3.1. Neural networks :

A model of thinking based on the human brain is what is meant by the term "neural network." The brain is made up of a network of neurons, or nerve cells, which are the simplest type of information processing. Nearly 10 billion neurons and 60 trillion synapses, or connections between neurons, make up the human brain. The brain can accomplish its tasks considerably more quickly than the current quickest computers because it uses numerous neurons at once.

Even though each neuron has a fairly basic structure, an army of these components may process information incredibly quickly. Figure 1 depicts the components of a neuron, including the cell body (soma), nucleus, dendrites (groups of short fibers), axon (a single long fiber), and axonal terminal.

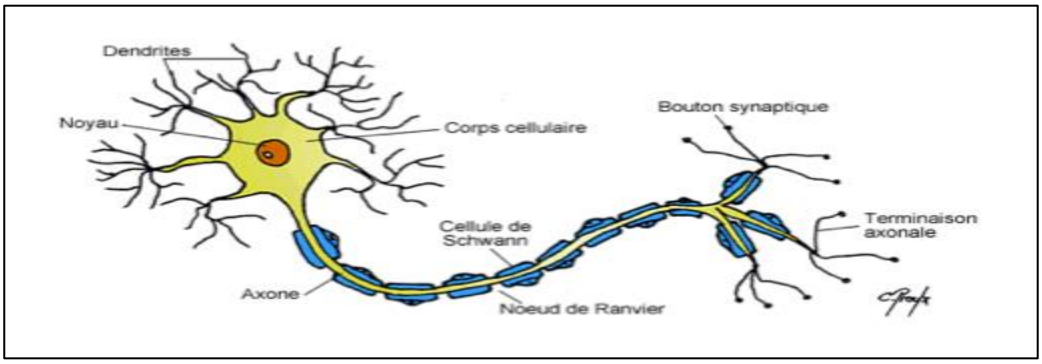


Figure 1: A schematic drawing of a neuron

When the dendrites branch in a network around the soma, the axon extends to the dendrites and the soma of other neurons, the neurons are interconnected at the synapses and the neuronal networks are formed in this way (figure 2) . [52]

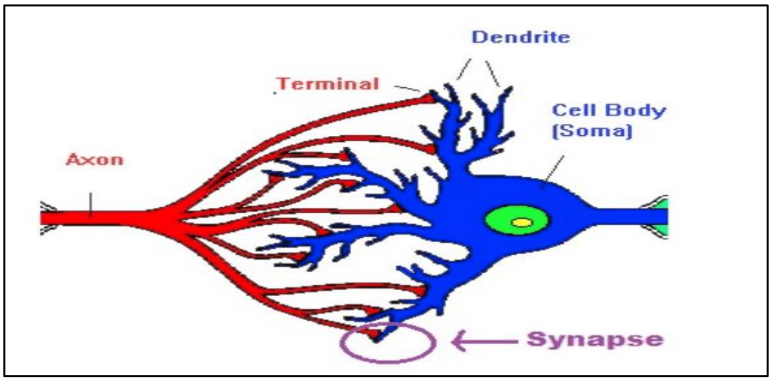


Figure 2: Schematic of the connection between neurons

Neurons communicate with one another by means of intricate electrochemical processes. The electrical potential of the cell body changes as a result of the substances released by the synapses. The action potential, an electrical impulse, is delivered down the axon when the voltage hits its threshold. The synapses' potential changes as a result of the impulse's propagation and eventual arrival there. However, the fact that a brain network is changeable is the most intriguing discovery. Neurons show long-term alterations in the strength of their connections in response to the stimulation pattern. Additionally, neurons and other neurons can

interact in new ways. Sometimes even complete neuronal clusters might move from one location to another. These processes serve as the brain's building blocks for learning.
Our brains are thought to be extremely sophisticated, non-linear, parallel information processing systems. Instead than being processed at certain points, information in a neural network is concurrently stored and processed all around the network. In other words, data and its processing in neural networks are global rather than local.
The connections between neurons that lead to the "right answer" are strengthened by plasticity, whereas the connections that lead to the "wrong answer" are diminished. Neural networks can therefore gain knowledge through experience.
A key and crucial characteristic of biological brain networks is learning. There have been attempts to simulate a biological neural network in a computer due to how simple and naturally they can learn.

3.2. From biology to simulation :

An artificial neural network is comprised of a group of interconnected processors, referred to as neurons, which are simple and highly interlinked. These neurons are analogous to biological neurons present in the brain. The neurons are linked through weighted connections, which pass signals from one neuron to another. Each neuron obtains multiple input signals through these connections, but it generates only one output signal. The outgoing connection of the neuron (equivalent to the biological axon) transmits the output signal. This connection further divides into numerous branches that carry the same signal (there is no division of the signal between these branches). The outgoing branches ultimately lead to incoming connections from other neurons present in the network. [54], [53].

3.3. The artificial neuron :

A simple computing component of an artificial neural network (ANN), an artificial neuron receives numerous signals from its input links, computes a new value, and transmits it as an output signal through the output links. The input signal may be made up of raw data or neuronal outputs. The output signal may serve as both a conclusive resolution to the issue and an input for further neurons[54]. Figure 3 depicts a synthetic neuron.

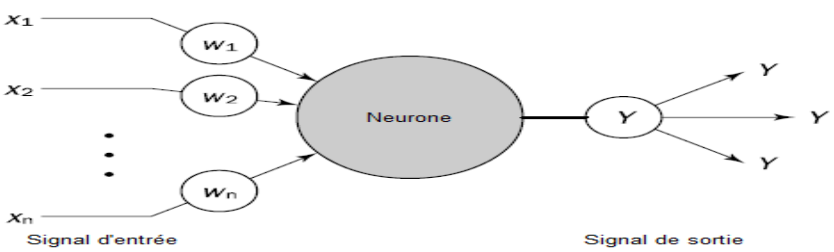


Figure 3: The artificial neuron

3.4. The Perceptron :

The threshold logic unit (TLU) or linear threshold unit (LTU), a slightly modified version of an artificial neuron, is the foundation of the perceptron, a simple type of artificial neural network design. The TLU employs numerical values as inputs and outputs (rather than binary on/off values), and each input connection is associated with a particular weight, in contrast to the artificial neuron. In order to produce an output, the TLU first calculates a weighted sum of its inputs (z = w_1 x_1 + w_2 x_2 + + w_n x_n= xTw), and then applies a step function to the total (h_w (x) = step(z), where z = xTw. Figure 4 gives a graphic illustration of this procedure.

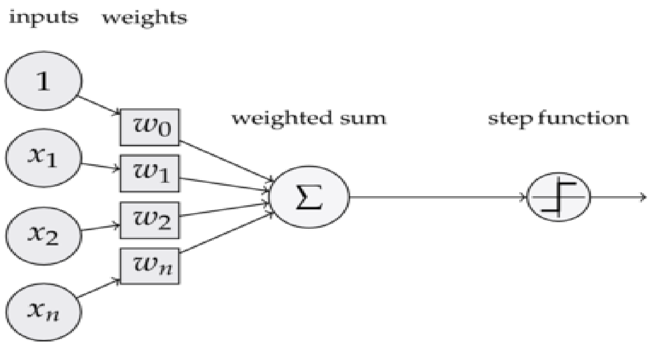


Figure 4: The structure of a perceptron

3.5. Activation functions :

In neurons, the inputs are added together by weights before a function—the activation function—is applied. The figure depicts where the activation function is located within the neural network. The output of nodes is propagated from one layer to the next (up to and including the output layer) using activation functions. The neuron is activated by the activation functions, which are a scalar to scalar function. We add nonlinearity to a neural network's modeling capabilities by using activation functions for hidden neurons. A class of logistic transforms that (when graphed) resembles a S includes several activation functions. A sigmoid is a class of functions in this category. The sigmoid function is one of several variations that make up the sigmoid function family. Now let's examine a few helpful activation features in neural networks.[55].

3.5.1. Linear :

In a linear transformation, the dependent variable and independent variable have a direct and proportional relationship, and f(x) = Wx is the identity function. This essentially indicates that the signal is passed through the function unchanged.

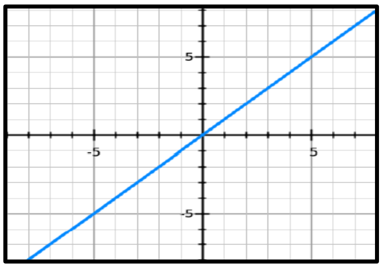


Figure 5: linear function

3.5.2. Sigmoid :

A sigmoid function ($f(x) = \frac{1}{1+e^{-x}}$) is a function that converts independent variables of nearly infinite range into simple probabilities between 0 and 1, and most of its results will be very close to 0 or 1.

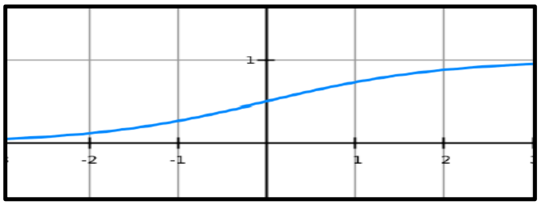


Figure 6: Sigmoid function

3.5.3. Tanh :

Tanh's normalized range, in contrast to the Sigmoid function, is from -1 to 1. Tanh is 1 to -1. Tanh has the benefit of being more adept at dealing with negative numbers.

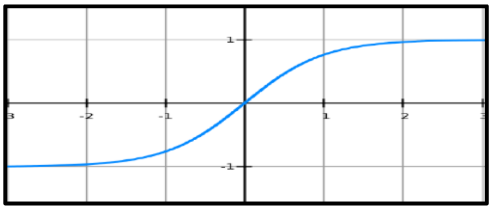


Figure 7: Tanh function

3.5.4. ReLU :

A more interesting transformation is Rectified linear Units (ReLU), which only activates a node if the input exceeds a predetermined threshold. The output is zero when the input is less than zero, but there is a linear relationship between the input and the dependant variable when the input reaches a particular threshold.

$$\text{f(x)} = \begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases}$$

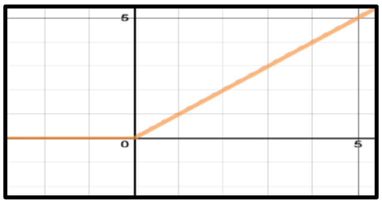


Figure 8: ReLU function

4. Deep learning :

4.1. Definition :

A branch of machine learning called "deep learning" uses neural networks with three or more layers. These neural networks are designed to behave similarly to how the human brain functions so they can learn from enormous volumes of data. A neural network with a single layer can still make rough predictions, but the accuracy can be improved and strengthened by adding more hidden layers. [56].

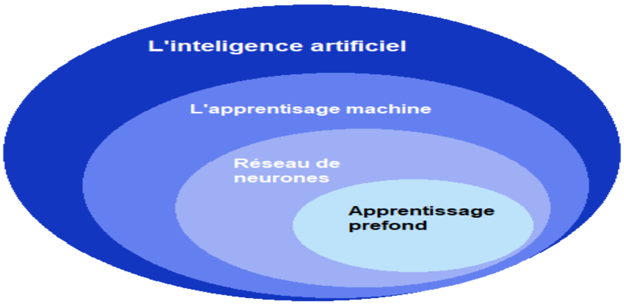


Figure 9: Diagram showing the positioning of AI concepts

The depth of the layers in a neural network is expressly referred to by the term "deep" in deep learning, it should be emphasized. Deep neural networks are defined as having more than three layers. Although this truth has already been indicated in the context of neural networks, it is crucial to expressly declare it to remove any potential for confusion.This is usually represented using the following diagram (Figure 10):

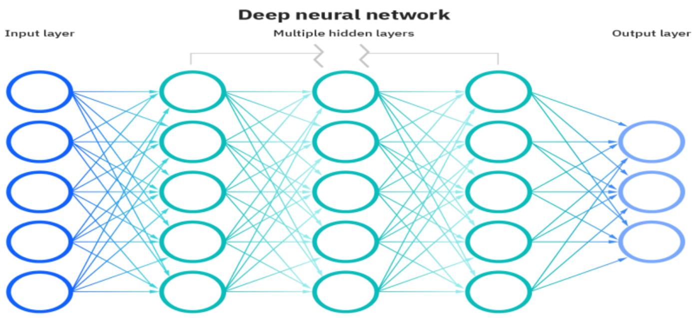


Figure 10: Deep neural network

4.2. History of deep learning :

Année	Contributeur	Contribution
300 AC	Aristotle	introduction de l'associationnisme, début de l'histoire des humains qui essayent de comprendre le cerveau
1873	Alexander Bain	introduction du Neural Groupings comme les premiers modèles de réseaux de neurone
1943	McCulloch and Pitts	introduction du McCulloch–Pitts (MCP) modèle considérer comme L'ancêtre des réseaux de neurones artificielles
1949	Donald Hebb	considérer comme le père des réseaux de neurones, il introduit la règle d'apprentissage de Hebb qui servira de fondation pour les

		réseaux de neurones modernes
1958	Frank Rosenblatt	introduction du premier perceptron
1974	Paul Werbos	introduction de la retro propagation
1980		introduction des cartes auto organisatrices
1980	Kunihiko Fukushima	introduction du Neocognitron, qui a inspiré les réseaux de neurones convoluti
1982	John Hopfield	introduction des réseaux de Hopfield
1985	Hilton and Sejnowski	introduction des machines de Boltzmann
1986	Paul Smolensky	introduction de Harmonium, qui sera connu plus tard comme machines de Boltzmann restreintes
1986	Michael I. Jordan définition et	introduction des réseaux de neurones récurrent
1990	Yann LeCun	introduction de LeNet et montra la capacités des réseaux de neurones profond
1997	Schuster and Paliwal	introduction des réseaux de neurones récurrent bidirectionnelles
1997	Hochreiter and Schmidhuber	introduction de LSTM, qui ont résolu le problème du vanishing gradient dans les réseaux de neurones récurrent
2006	Geoffrey Hinton	introduction des Deep belief Network
2009	Salakhutdinov and Hinton	introduction des Deep Boltzmann Machines
2012	Alex Krizhevsky	introduction de AlexNet qui remporta le challenge ImageNet

Table 1: Summary of the history of deep learning

4.2.1. Deep learning vs. Machine learning :

Deep learning is a subset of machine learning that differs in the type of data it utilizes and the learning methods employed. Traditional machine learning algorithms work with structured and labeled data to make predictions, where specific features are defined from input data and organized into tables. Although machine learning can use unstructured data, it typically requires preprocessing to organize the data into a structured format.

In contrast, deep learning models can learn from unstructured and unlabeled data, such as images, audio, and text, without the need for preprocessing. Deep learning models also employ different learning techniques, including supervised learning, unsupervised learning, and reinforcement learning [57]. This is in contrast to traditional machine learning, where human experts manually establish these characteristics.

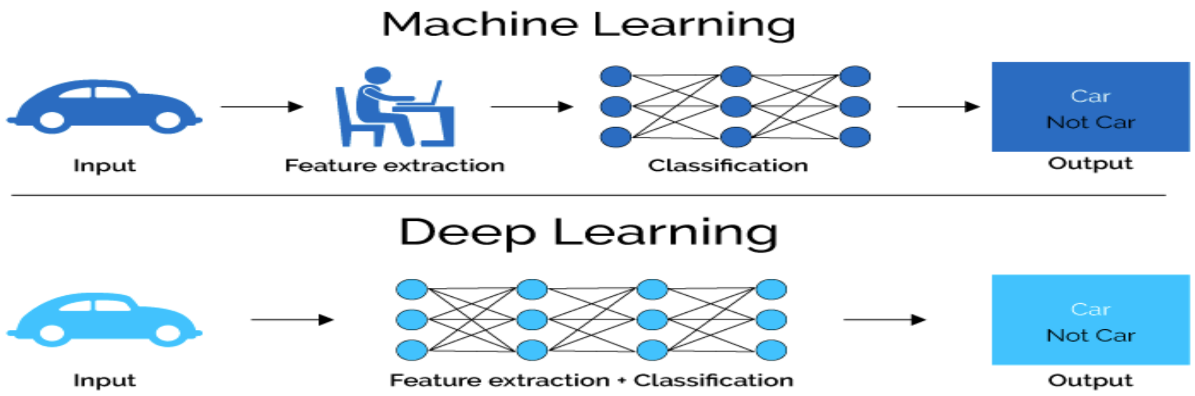


Figure 11: machine learning and deep learning

5. Convolutional Neural Networks :

Since our study is based on medical images, we will talk about the convolutional neural network in more detail, since it is the most specialized structure in image recognition:

5.1. The architecture of a convolutional neural network :

There are many variations of the CNN architecture but they are based on the layered model as shown in Figure 12 :

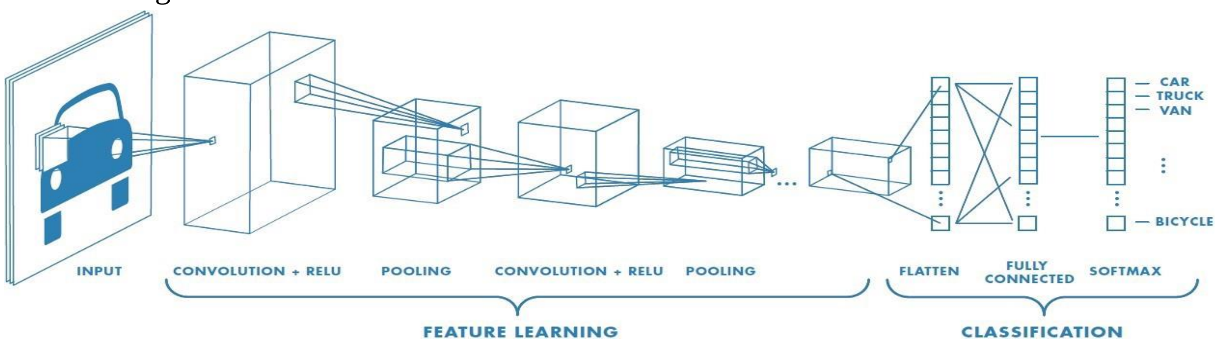


Figure 12: The architecture of a convolutional neural network

- Input layer
- Feature-extraction layers
- Classification layers

The input layer typically accepts three-dimensional data in the spatial form of the image's dimensions (width height) and a depth indicating the color channels (typically three for RGB color channels).

Multiple characteristics are found in images via feature extraction layers, which then gradually develop higher-order features. This fits perfectly with the current deep learning trend. by means of which, as opposed to conventional manual engineering, features are taught automatically.

The classification layers, which comprise one or more fully connected layers, are the last to come into play. These layers use higher order characteristics to generate probabilities or class scores.

As their name implies, these layers are completely connected to every neuron in the preceding layer.

5.2. Convolutional Neural Network Layers :

As we saw in the previous figure 12, the convolative neural network is composed of 3 main types of layers: convolutional layer, pooling layer and fully connected layer.

5.2.1. Convolutional layer :

A convolutional layer, which is the primary component of a convolutional neural network (CNN), is where the majority of computations are carried out. Three elements are needed for the convolutional layer: input data, a filter, and a feature map. The input will have three dimensions (RGB in an image) for height, width, and depth if it is a color image. Convolution is the process of detecting the feature by moving a feature detector or kernel over the receptive fields of the image.

A 2D array of weights serving as the feature detector represents a section of the image. The size of the receptive field is typically determined by the filter size, which is typically a 3x3 matrix. A section of the image is subjected to the filter, and the dot product between the input pixels and the filter is computed. The output matrix is then fed with the scalar product, and the filter shifts once more in order to scan the entire image. A feature map, activation map, or convolved feature is the output of the series of dot products between the input and the filter.

As you can see in Figure 13, It is not necessary to link each pixel value of the input image to each output value of the feature map. Only the receptive field, where the filter is applied, needs to be connected..[58]

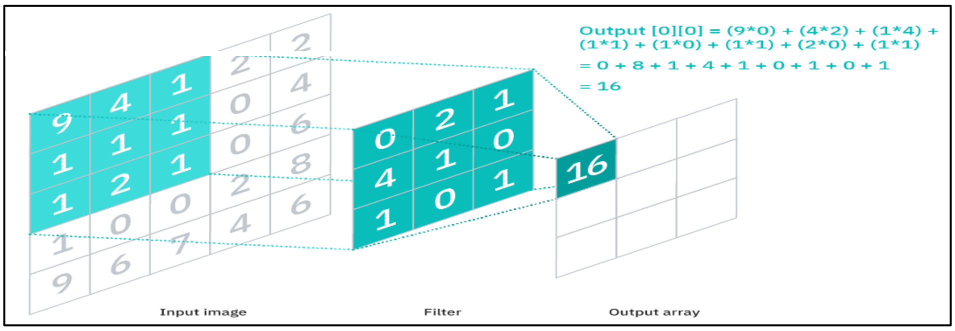


Figure 13: How the convolutional layer works

The majority of the computations are carried out by convolutional layers, which are regarded as the essential parts of a convolutional neural network (CNN). They need input data, a filter, and a feature map, which total three components. For instance, if the input is a color image, the result will be a 3D pixel matrix where the RGB values for height, width, and depth are matched. A 2-D array of weights known as the feature detector, or kernel, scans the image's receptive fields in search of the feature.

This method is referred to as convolution.Due to the fact that not all input values in the output matrix must be directly connected, convolutional and pooling layers are referred to as "partially connected" or "locally connected layers." The feature detector uses parameter sharing, which keeps its weights constant as it traverses through the image. The size of the output volume is affected by three hyperparameters that must be established prior to starting the neural network training. Gradient descent and backpropagation are used during learning to modify some parameters, such as weight values.

5.2.1.1. The number of filters :

It determines the depth of the exit. For example, three separate filters would produce three maps of different characteristics, creating a depth of three. Related works

5.2.1.2. The strides :

The distance—or quantity of pixels—that the core travels on the input matrix is known as the stride. Despite the rarity of stride values of two or higher, a bigger stride has a negative impact on output.

5.2.1.3. padding :

When the filters do not adjust to the input image, padding is typically utilized. By zeroing out all elements outside the input array, this results in an output that is larger or equal in size. Three different fillings are available.:

- Valid padding: This option is also known as "no padding". In this case, the last convolution is discarded if the dimensions are not aligned.
- Same padding: This padding ensures that the output layer has the same size as the input layer.
- Full padding: This type of padding increases the size of the output by adding zeros to the border of the input.

Convolutional Neural Networks (CNNs) use After each convolution operation, a Rectified Linear Unit (ReLU) activation function is applied. to introduce nonlinearity into the model. This is because linear transformations are not sufficient to capture the complex relationships and patterns in images. CNNs can have multiple convolutional layers, and when this happens, the structure of the network becomes hierarchical, as later layers can see the pixels located in the receptive fields of the previous layers. This hierarchical structure allows the network to detect increasingly complex features in the image. For example, in an image of a bicycle, the lower-level patterns would be the individual parts of the bike, such as the frame and handlebars, while the higher-level pattern would be the combination of these parts that represents the complete bicycle. Finally, The convolutional layer numerically transforms the image so that the neural network can understand and extract patterns from it.[58]

5.2.2. The Pooling Layer :

By lowering the number of parameters, layer pooling, often referred to as subsampling, aids in reducing the dimensionality of the input. In that it employs a filter to scan the full input, the pooling operation is comparable to the convolutional layer; the distinction is that the pooling filter lacks weights. Instead, it creates the output array by applying an aggregate function to the data in its receptive field. Two popular forms of pooling exist.:

5.2.2.1. Average pooling layer :

The average value in the filter's receptive field is calculated as it passes over the input, and it is then given to the output network.

5.2.2.2. Max pooling layer :

The filter chooses the highest value inside its receptive field as it passes over the input and assigns it to the output array. Max pooling is employed more frequently in reality than average pooling.

5.2.3. Fully connected layer :

Every node in the preceding layer is connected to every node in the current layer in a completely connected layer, sometimes referred to as a dense layer. Based on the features retrieved by the preceding layers and their respective filters, this layer performs the classification operation. Fully connected layers often utilize a softmax activation function to appropriately categorize inputs, producing a probability value between 0 and 1 for each class, as opposed to convolutional and pooling layers, which frequently use ReLu functions..[58]

5.3. Some convolutional neural network architectures :

5.3.1. Lenet-5 :

Lenet-5 was initially introduced in a research paper titled "Gradient-Based Learning Applied to Document Recognition" published in 1998 by Yann LeCun and colleagues. The model was developed for recognizing machine-printed characters and handwritten characters using a convolutional neural network architecture.

The simplicity and ease of implementation of Lenet-5 made it one of the most widely used pre-trained models. The architecture of Lenet-5 consists of multiple convolutional layers, with three sets of convolutional layers combined with average pooling. The output of these convolutional layers is then passed through two fully connected layers before being classified using Softmax activation.[59]

Using Lenet-5, researchers have been able to achieve high accuracy in image classification tasks. However, to avoid plagiarism, it is important to rephrase the text in your own words and provide proper attribution when necessary.

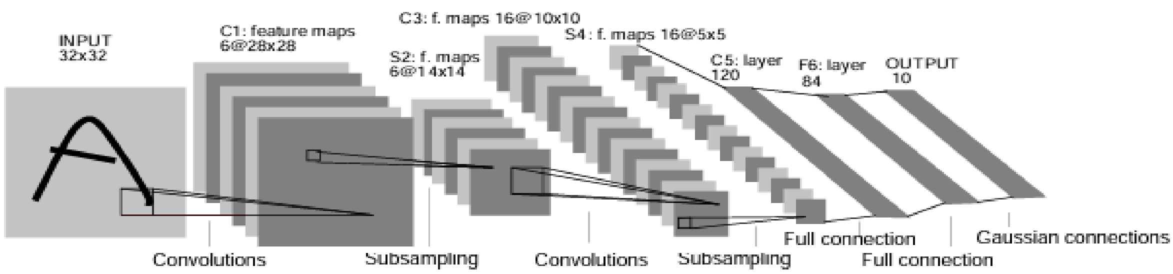


Figure 14: Archetecture of Lenet

5.3.2.VGGNet :

VGG is an abbreviation for Visual Geometry Group, which is a deep convolutional neural network (CNN) structure with either 16 or 19 convolutional layers known as VGG-16 and VGG-19, respectively. VGGNet is a cutting-edge object recognition model architecture that is based on the VGG design. It has been shown to surpass existing benchmarks in a wide range of tasks and datasets beyond just ImageNet. Furthermore, it is still one of the most widely used image recognition architectures in use today, making it highly relevant and important in the field of computer vision.[60]

The VGG network is constructed using incredibly tiny convolutional filters. Thirteen convolutional layers and three fully connected (FC) layers make up the VGG-16:

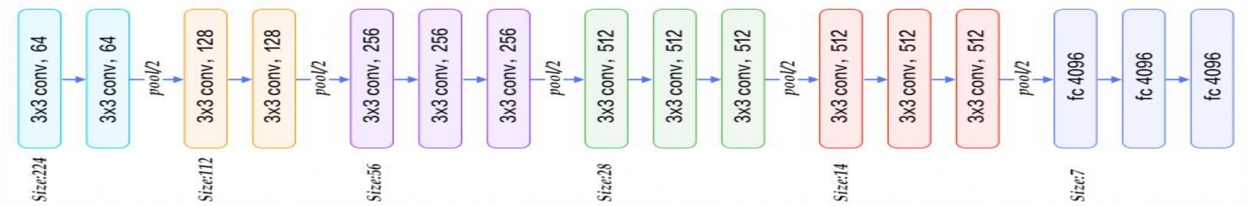


Figure 15: The architecture of VGG-16

Compared to VGG16, VGG19 includes three extra convolutional layers.:

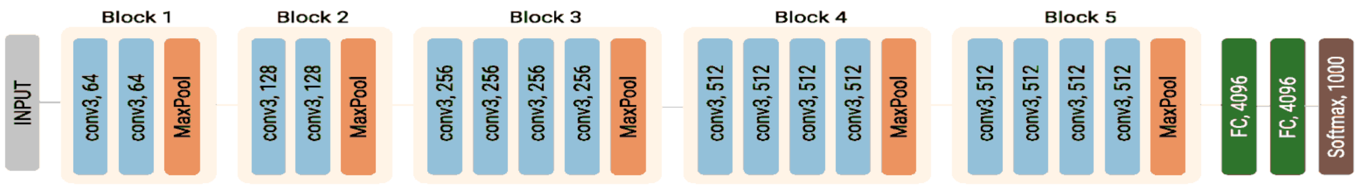


Figure 16: The architecture of VGG-19

5.3.3. GoogleNet (Inception) :

GoogLeNet is a convolutional neural network (CNN) architecture developed by Christian Szegedy et al. from Google Research, which achieved an exceptional performance by winning the 2014 ILSVRC Image Classification Challenge with a top-5 error rate of less than 7%. The network's depth was significantly greater than that of previous CNNs, owing to the use of inception modules, which allowed GoogLeNet to use its parameters more effectively.

The input signal is replicated and passed into four separate levels in the initialization module, each of which has a 33 kernel, stride 1, and SAME padding. The second set of convolutional layers detects patterns at various scales by using kernels of various sizes, including 11, 33, and 55. All layers, including the maxpooling layer, employ ReLU activation, with a stride of 1, and the same padding. The feature mappings from the first four convolutional layers are stacked in the final depth concat layer. The TensorFlow tf.concat() technique can be used to implement this concatenation layer.

The first two layers reduce the height and width of the image by 4, followed by local response normalization to promote feature diversity. Two more convolutional layers and a max pooling layer reduce dimensionality, and nine inception modules follow. Max pooling is also used within the inception modules to further reduce dimensionality and improve speed.[61]

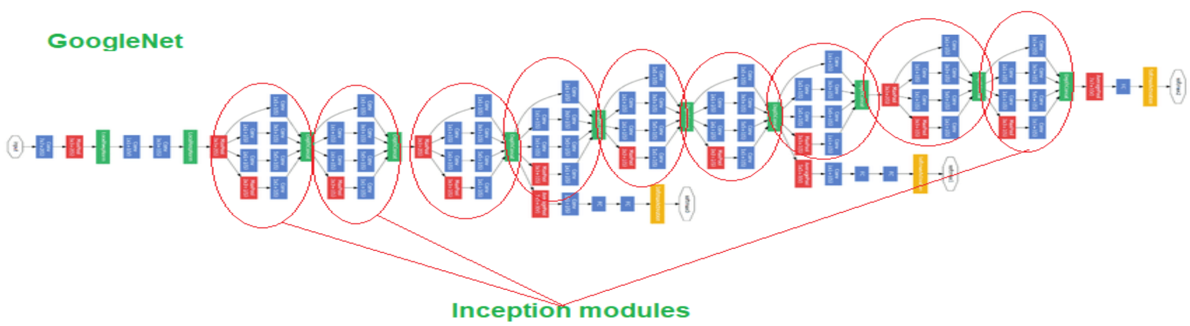


Figure 17: The architecture of GoogleNet

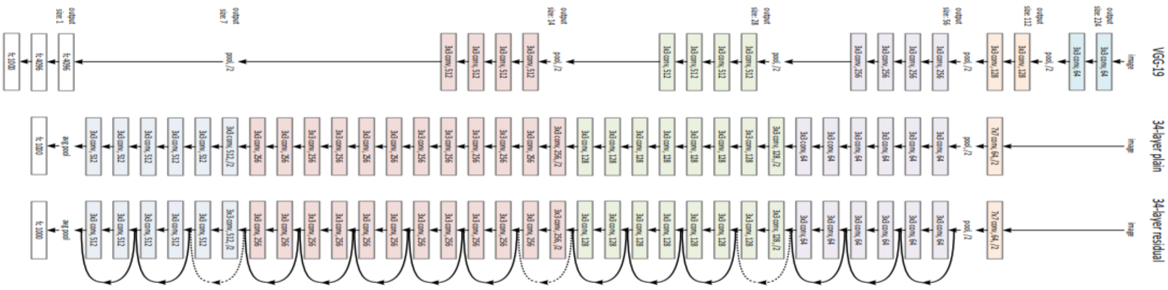
The output is routed into a global pooling average layer after going through all nine inception modules. This layer computes the average of each feature map while omitting any residual spatial data. This is advantageous since the GoogLeNet's input images are typically anticipated to be 224 ×224 pixels, and the feature maps are reduced to 7 ×7 after five max pooling layers, each of which divides the height and width by 2. The object's location is unimportant because this is a categorization challenge rather than a location task.

Since there are 1000 classes, the network's final layers consist of a dropout layer for regularization, a fully connected layer with 1000 units, and a softmax activation function to output the estimated class probabilities.

5.3.4.ResNet :

In the 2015 ILSVRC challenge for classification, the winning model was a Residual Network (ResNet), developed by Kaiming He et al. This model achieved an impressive top-5 error rate of less than 3.6%, using a deep CNN with 152 layers. This continued the trend of models becoming deeper with fewer parameters. The reason ResNet was able to achieve such depth is the use of skip connections, or shortcut connections, which add the input signal to the output of a higher layer in the stack. This approach forces the network to learn the residual function f(x) = h(x) - x instead of directly modeling the target function h(x). This is known as residual learning.[62]

The ResNet network uses a simple 34-layer network architecture inspired by VGG-19 in which



skip connections are then added:

Figure 18: ResNet-34 architecture

6. Image Registration Related Work :

Over the years, medicine and disease detection techniques have evolved significantly. However, the analysis of medical images has posed numerous challenges and obstacles due to distortions and other factors. To overcome these difficulties, various solutions have been devised, primarily centered around the practice of registration medical images.
The process of image registration holds immense significance within the realm of medical imaging. It encompasses a range of methodologies and strategies that aim to align and compare images to extract valuable information. In this section, we will present an overview of several cutting-edge studies conducted in recent years, which leverage deep learning techniques for medical image registration.

6.1. **Wu et al**.(2013) ,Unsupervised Deep Feature Learning for Deformable Registration of MR Brain Images [63]:

They randomly selected a subset of 20 magnetic resonance (MR) images from the ADNI data set, distinct from the training images. Pre-processing steps included skull removal, bias correction, and intensity normalization. To align all subject images with the template image, linear registration was performed using FLIRT. Then, six deformable registration methods were used to further normalize all subjects in the template space. Dice ratios, which measure the overlap between recorded and floor tissue types (ventricular, gray matter, white matter) were calculated for each recording method. Whereas, the H + ISA method yielded the highest dice score of 87.3%, similar to the results obtained in the IXI dataset. Additionally, compare the hippocampal overlap ratio utilizing recording techniques that use various features; for example, I learnt about the ISA and PCA features. In the case of HAMMER, H+ISA is obtained.a general improvement of 2.74% when compared to the similarly performing models H + PCA and HAMMER. The overall improvement for M+ISA, however, was 0.19%. 0.24%
Only H + ISA showed a significant improvement (0.05) over the baseline technique, when compared to M + PCA and the original daemons (Hummer).

6.2. **Liao et al**.(2017) ,An artificial agent for robust image registration [64] :

Two medical 3D image registration datasets were used in the experiments. The first dataset (E1) consisted of abdominal spine CT and CBCT images, and the success rate was evaluated based on a threshold of TRE (target registration error) $\leq$ 10 mm. The second dataset (E2) included cardiac CT and CBCT images, and the success rate was evaluated using a threshold of MME (mean registration error) $\leq$ 20 mm. To ensure reliability, cross-checks were performed by five different blinded data splitters for both E1 and E2.
The training and network architecture were consistent for both applications. The same network architecture and descriptive parameters were used. The training process involved two stages: coarse registration and fine registration. For the coarse-registration CNN, a solid-body

perturbation was randomly generated within specific ranges to cover a large field of view. For E2, the perturbation range was [±30 mm, ±30 mm, ±30 mm, ±30°, ±30°, ±30°], while for E1, it was [±30 mm, ±30 mm, ±150 mm, ±30°, ±30°, ±30°] to accommodate the head-to-foot direction of CT spine images. To refine the registration, a narrower range of solid body disturbance was used, namely [±5 mm, ±5 mm, ±5 mm, ±5°, ±5°, ±5°].

The RMSprop update algorithm without momentum was employed for training the network, with a batch size of 32. The learning rate was set to 0.00006 with a decay of 0.7 every 10,000 small batch-based iterations.

The proposed hierarchical scoring method was evaluated using dataset E1. Applying the first CNN reduced the mean error from 3.4 mm to 2.5 mm, and further applying the second CNN yielded better results. For E2, due to significant non-rigid deformation between CT and CBCT, the MME was noticeable even with ground truth shift. The success rate achieved was 92%.

6.3. **Yang et al**.(2016) , Fast predictive image registration [65] :

Based on image appearance-wise correction, they have offered a method to forecast image distortions. on the OASIS brain picture dataset, both in 2D and 3D.

Concentrate on Model for large deformation mapping of different shapes (LDDMM). When compared to GPU-based optimization of 2D/3D picture registration, it achieves 1500p / 66x faster computation times while maintaining the ideal theoretical qualities of an LDDMM by anticipating the momentum parameters of an LDDMM. This method produces anomaly transitions and has better forecast accuracy than predicting deformation fields or velocity. Additionally, they developed a probabilistic Bayesian version that allows for dropout Monte Carlo sampling to quantify the uncertainty in the deformation field at the time of the test. And you get the lowest error rate deformation of the case in 2D / 3D

6.4. **Fan et al.** (2019) ,Birnet: Brain image registration using dual supervised fully convolutional networks [66] :

They suggested a deep learning method for picture registration using appearance-based deformation prediction. Due to the difficulty in obtaining Ground Truth Distortion Fields for training, they created a fully convolutional network that performs two types of routing: ground-truth routing using the deformation fields obtained with the current registration method, and direct image contrast using the difference between the registered images. The latter advice aids in avoiding an excessive reliance on supervision of training distortions that may be erroneous. They optimize the deep convolutional network using bridging, hierarchical loss, and many sources for efficient training.In contrast to numerous contemporary distortion scoring methods. Using the LPBA40, they trained the BIRNet. where the opening picture appeared in the LPBA40 Images 1 through 30 were chosen as training samples, whereas images 31 through 40 served as validation data. then, without further ado, apply to four distinct test datasets. BIRNet outperformed all other networks with DSC values of 69.2%.

6.5. Fu et al.(2021) , Biomechanically constrained non-rigid mr-trus prostate registration using deep learning based 3d point cloud matching [67] :

a non-rigid MR-TRUS image registration model for prostate procedures has been proposed. The registration architecture consists of convolutional neural networks (CNNs) for MRI prostate segmentation, a CNN for TRUS prostate segmentation, and a cloud-based network for quick 3D point cloud matching. With the aid of segmented prostate masks, volumetric prostate point pulls were produced.Tetrahedron network. The deformation field produced by finite element analysis served as the training data for the point cloud matching network. For the purposes of testing and training the network, 50 patient data sets were utilized. Three measures, including the dice similarity coefficient (DSC), mean surface distance (MSD), and Hausdorff distance (HD), were used to assess the prostate forms' alignment following registration. Using the measurable Recording error (TRE), point-to-point internal scoring accuracy was assessed. Jacobean determining factors and stress tensors of the anticipated deformation field were calculated for the deformation field's physical fidelity study. On average, generally, and standard For DSC, MSD, HD, and TRE, the deviations were 0.94 ± 0.02 mm, 0.90 ± 0.23 mm, 2.96 ± 1.00 mm, and 1.57 ± 0.77 mm, respectively. Our findings demonstrated that the suggested approach may swiftly record MR-TRUS images with good recording accuracy and power.

6.6. He et al.(2019) , 3D Convolutional Neural Networks Image Registration Based on Efficient Supervised Learning from Artificial Deformations [68]

.Proposed supervised, non-static image recording ,The training approach was used to compare three network architectures using artificial displacement vector fields (DVF).Offer a plan of action can create DVFs artificially and record them on chest CT scansThis mimicked breathing action should be increased. Suggest a multi-stage technique is used to integrate the architecture in order to expand the network capture area and more accurately predict significant displacements. In the SPREAD and DIR-Lab-4DCT experiments, the suggested technique RegNet was tested against various chest CT scan databases and achieved a target registration error of 2.32± 5.33 mm and 1.86 ± 2.12 mm, respectively. RegNet with two stages has an average inference time of roughly 2.2 seconds.

6.7. Hu et al.(2018) , Weakly supervised convolutional neural networks for multimodal image registration [69] :

presented a technique to determine voxel-level shift using high-level correspondence data from anatomical labels.During the SmartTarget clinical trials, 108 pairs of T2-weighted MR and TRUS pictures from 76 patients were collected. While only untagged image pairings are employed as network inputs for inference, the proposed end-to-end convolutional neural network approach tries to anticipate displacement fields to align multiple labeled matching structures for individual image pairs during training. In validation trials, 108 pairs of multimodal photos from 76 individuals were examined using high-quality anatomical posters.

The results showed a mean scoring error of 3.6 mm on conspicuous centromeres and a mean dice of 0.87 on prostate glands.

6.8. Hering et al.(2021) , Cnn-based lung ct registration with multiple anatomical constraints [70] :

Identify the main strategies used in conventional lung recording techniques. And I succeeded in creating a deep learning equalizer. Use a series of rules based on the Gaussian pyramid level framework to address the problem of image registration optimization from coarse to fine,Provide an overall assessment of the scalability, accuracy, robustness, and reasonableness of our recording method's estimation associated with deformation fields. The latest findings in the COPDGene dataset achieve significant improvements (TRE < 1.2 mm) over other deep learning approaches, with much shorter execution times in exhalation-to-inhalation lung recordings compared to the traditional recording system. With a smaller standard deviation of 0.026 vs. 0.046, they have significantly superior average Dice scores than the conventional scoring technique (0.95 vs. 80.92). Additionally, 1.72 0.89 mm vs. 1.97 1.24 mm and 26.84 14.27 vs. 27.24 13.70 mm, respectively, were obtained for the mean symmetric surface and the Hausdorff distance.

6.9. Xu et al.(2021) , Multi-scale neural odes for 3d medical image registration [71] :

proposed to use a multiscale neural ODE model to learn the registration optimizer.Because the neural ODE learns of training data for effective gradient conditioning at each iteration, the inference comprises of iterative gradient updates similar to the traditional gradient descent optimizer but in a much faster manner. Additionally, he recommended developing a similarity metric that is independent of media to address image appearance variations across various image contrasts.Use distinct general and specialized data sets. voluntary set of data. The public collection, called Brain Tumor Segmentation (BraTS) 2020 collection, consists of 494 glioblastoma patients' multimodal 3D brain MRI images with four different contrasts (T1, T2, T2-FLAIR, and T1Gd). Dice = 81.1% (9.9), RMSE(x) = 6.28 (2.38), RMSE() = 1.52 (1.00), Time = 0.56 (0.09)s.

6.10. Ti et al.(2022) , GraformerDIR: Graph convolution transformer for deformable image registration [74] :

They proposed a Graformer-based DIR framework called GraformerDIR by placing a graph convolution converter (Graformer) layer at the bottom of the feature extraction grid. The Graformer layer consists of the Graformer module and the Cheby-shev graph convolution module. Among them, the Graformer module is designed to capture high-quality long-running dependencies. The Cheby-shev graph convolution unit is used to further expand the receiving field. The performance and generalizability of GraformerDIR was evaluated on publicly available brain datasets including the OASIS, LPBA40, and MGH10 datasets. Compared to VoxelMorph, GraformerDIR achieved performance improvements of 4.6% in dice similarity coefficient (DSC) and 0.055 mm in mean asymmetric surface distance (ASD) while reducing the non-positive rate for the specific Jacobin index (Npr.Jac) by approximately 60 times on a dataset Publicly available OASIS. In the MGH10 unseen dataset, GraformerDIR obtained performance improvements of 4.1% in DSC and 0.084 mm in ASD compared to VoxelMorph, demonstrating a GraformerDIR with better generalizability.

Conclusion

In this chapter, we talked about artificial intelligence, got acquainted with many of its technologies, and now we know the technology on which we base our study, which is the convolutional neural network.

We reviewed relevant work related to recording medical images using deep learning and machine learning in 3D and 2D, and reviewed the latest algorithms used.

Chapter 4

Concept and programming

Introduction

We will present the experimental portion of our research in this chapter. We employed the FCN pre-training based on medical image because we decided to apply diffeomorphic registration of 2D images. We connected this network to reposition the base's medical imaging (brain MRI from OASIS). using 2D medical images, we validated and tested this model.

1.Work Environment and Tools :

We will see in this section, the hardware and software used to create our model:

1.1.Google Colaboratory :

It can be CPU/GPU expensive to train a deep learning model, thus we made use of Google's Colab cloud computing platform. Colaboratory is a Google research initiative designed to advance machine learning research and education. It is a cloud-based Jupyter notebook environment that may be used immediately without any configuration.[72]

Figure 1:Google Colaboratory Logo

1.2. Python programming language :

Powerful and simple to learn, Python is a programming language. High-level data structures are included, and object-oriented programming is made simple but powerful. Python is a fantastic language for scripting and quick application development in many domains and on most platforms because of its clean syntax, dynamic typing, and interpretability.[73]

Figure 2: Python Logo

If Python is usable in so many areas, it is thanks to the richness of its libraries
So, the most used Python libraries in our approach are:

1. NumPy :

The cornerstone Python module for scientific computing is called NumPy. It is a Python library that offers a multidimensional array object, various derived objects (like hidden arrays and matrices), and a variety of routines for carrying out quick operations on arrays. These operations include math, logic, shape manipulation, sorting, selection, E/S, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation, and much more.

Figure 3: NumPy Logo

2. Matplotlib :

The Python programming language module Matplotlib is used to plot and visualize data as graphs. It can be used in conjunction with the Python libraries for NumPy and SciPy for scientific computing. A BSD-style license governs the free and open distribution of Matplotlib.[73]

Figure 4: Matplotlib Logo

3. Pandas :

Built on top of the Python programming language, pandas is an open source data analysis and manipulation tool that is quick, strong, adaptable, and simple to use.

Figure 5: Pandas Logo

4. Tensorflow :

An open source library for machine learning and numerical computation is called TensorFlow. It makes it easier to gather data, train machine learning models, make predictions, and improve future outcomes.

TensorFlow combines Deep Learning and Machine Learning models and methods. The front-end API provided by the Python programming language makes it easy and enjoyable to create applications utilizing this framework. This framework can be used to train and run deep neural networks for recurrent neural networks, image recognition, lexical embedding, and the classification of handwritten digits.[73]

Figure 6: Tensorflow Logo

5. Keras :

A high-level neural network API called Keras was created in Python and can interact with either TensorFlow or Theano. Rapid experimentation was a major consideration in its development. Effective research relies on being able to move as quickly as feasible from concept to conclusion. Its primary author and maintainer is François Chollet, a Google engineer, and it was created as a component of the ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System) project's research endeavor. The Google TensorFlow team made the decision to include Keras support in the primary TensorFlow library in 2017. Keras was created as an interface rather than an entire learning platform, according to Chollet. It presents a higher-level, more understandable collection of abstractions, making configuration simpler.independent of the backend computing library, neural networks. Microsoft is also attempting to give Keras a CNTK backend.[73]

Figure 7: Keras Logo

2.General Architecture :

2.1. Method :

In traditional anamorphic registration methods, a pair of input images is typically set as a still image and a motion picture, and only one mapping from the still image to the motion image is considered. The reverse mapping is often ignored in these methods. However, in symmetric registration settings, the static and animated identities of the input images are not assumed.
Let X and Y be two volumes of holograms defined in a reciprocal spatial domain R^3. The distortion registration problem can be represented as a function :

$$f\theta(X, Y) = (\phi(1)\_XY, \phi(1)\_YX$$

where θ represents the learning parameters in a convolutional neural network (CNN).

Here, $\phi(1)\_XY = \phi\_XY(x; 1)$ and $\phi(1)\_YX = \phi\_YX(y; 1)$ represent the time 1 deformation fields that deform the identity position of some anatomical position x in X towards y in Y and warp y in Y towards x in X, respectively.
Motivated by conventional non-learning-based symmetric image normalization methods, the proposal is to learn two 0.5-time-discrete deformation fields that deform both X and Y towards their mean M-shapes along the geodesic trajectory. Once the model converges, the time 1 deformation fields that warp X to Y and Y to X can be obtained by composing the two estimated 0.5-time deformation fields. This configuration is possible because anisotropy is a differentiable map and guarantees the existence of a differentiable inverse.
The transformation from X to Y is represented as :

$$\phi(1)\_XY = \phi(-0.5)\_YX(\phi(0.5)\_XY(x))$$

while the transformation from Y to X is represented as :

$$\phi(1)\_YX = \phi(-0.5)\_XY(\phi(0.5)\_YX(y))$$

Therefore, the function fθ can be rewritten as :

$$f\theta(X, Y) = (\phi(-0.5)\_YX(\phi(0.5)\_XY(x)), \phi(-0.5)\_XY(\phi(0.5)\_YX(y)))$$

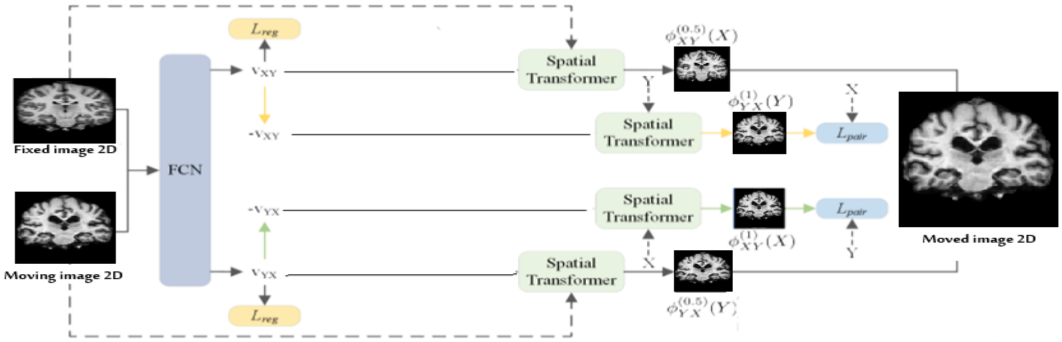


Figure 8: General Overview

2.1.1. symmetric diffeomorphic neural network :

Using a fully convolutional neural network (FCN), several scaling and quadrature layers, and differentiable spatial transformers, as seen in Figure 12, we parameterized the function f.[34] Scaling and squaring are used to calculate (0.5)XY and (0.5)YX using the estimated velocity fields (vXY and vYX, respectively.

Our FCN's architecture resembles that of U-Net [35], which has a jump-connected 5-level hierarchical encoder-decoder as illustrated in Figure 13. The suggested FCN combines X and Y into a single input for two channels and learns to estimate two dense and nonlinear velocity fields, v XY and v YX, from X and Y jointly from the start. We use two successive convolution layers with a 3 ˣ 3 ˣ 3 structure for each encoder level.

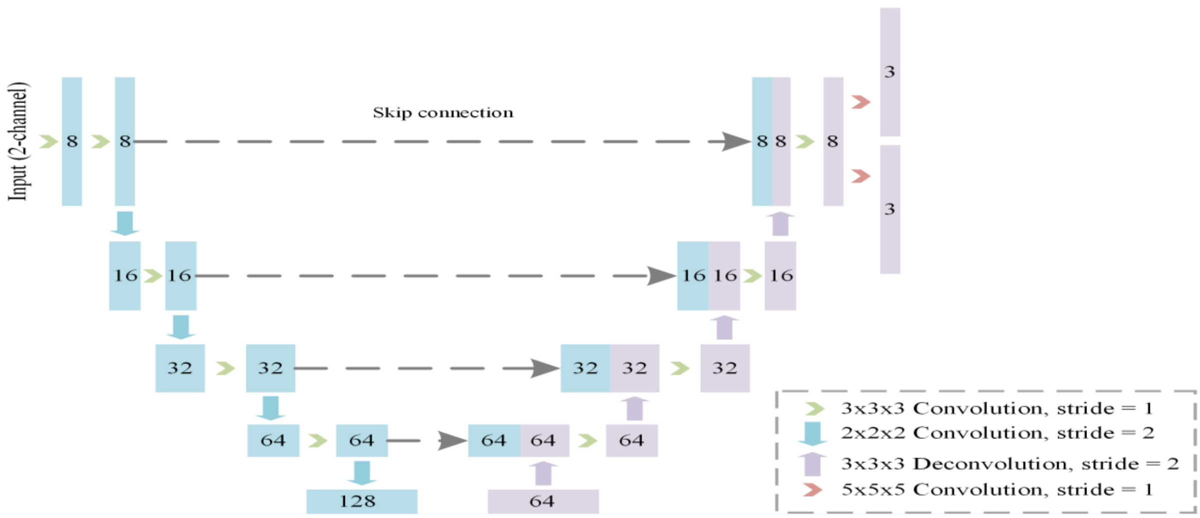


Figure 9: FCN architecture

2.1.2. symmetrical similarity:

The topology preservation, invertibility, and reverse consistency of transformation are three important diffeomorphic qualities that are frequently disregarded by current CNN-based approaches.[36]. Inspired by classical iterative symmetric normalization methods, Our method calculates transformations that deform X into Y and Y into X, such as (0.5)XY and (0.5)YX that change X into the mean form M from X and Y. We propose to minimize the symmetric mean shape similarity loss Lmean and the pairwise similarity loss Lsim by gradient descent, , which force the projected transformations' invertibility and inverse consistency. Our suggested method is compatible with all differentiable similarity measures, including normalized cross-correlation (NCC), root mean square error (MSE), sum of squared distance (SSD), and mutual information (MI), just like previous CNN-based methods. To determine the level of alignment between two images, we utilize the NCC normalized cross-correlation as the similarity measure. Let I and J represent two input image volumes, with I(x) and J(x) representing the local means of I and J, respectively, at location x. With w = 7 in our studies, the local mean is computed over a local window w3 centered at each position x. Following is a definition of the NCC:

NCC (I , J) =

$$\sum_{x\in\,\theta} \frac{\sum_{xi}\ (I(xi)-\bar{I}(x))(J(xi)\,-\bar{J}(x))}{\sqrt{\sum_{xi}(I(xi)-\bar{I}(x))^2\ \sum_{xi}(J(xi)-\bar{J}(x))^2}}$$

where xi designates the position in the local windows w3 centered in x. Specifically, our proposed similarity loss function.

Mean shape similarity loss Lmean and pairwise similarity loss Lpair make up the two symmetric loss terms that make up Lsim. The Lpair measures the pairwise dissimilarity between the deformed X at Y and the deformed Y at X, whereas the Lmean measures the dissimilarity between the distorted X and the deformed Y, which to the average shape M. After that, the suggested loss of similarity function is written as:

$$L_{sim}\ = L_{mean}\ + L_{pair}$$

With

$$L_{mean}\ = -NCC(X\left(\phi_{XY}^{0.5}\right), Y\left(\phi_{YX}^{0.5}\right))$$

And

$$L_{pair} = -NCC\left(X\left(\phi_{XY}^{1}\right), Y\right) - NCC(Y\left(\phi_{YX}^{1}\right), X)$$

And in diffeomorphic space XY and YX can be divided into (0.5) YX (0.5) XY and (0.5) XY (0.5) YX. In other words, increasing the similarity of bidirectionally deformed images tends to minimize the Lsim. Furthermore, as the suggested pairwise similarity loss function takes into account the transformation in both directions, our method automatically guarantees the inverse consistency in addition to inheriting the topology-preserving and invertibility properties of the diffeomorphic deformation model.

2.2. Data preprocessing :

We used the 425 T1 brain MRI dataset obtained from the OASIS dataset. Where we extract the middle slice from the brain scans, we obtain a 2D dataset.

The data set was divided into two subsets: the training set and the validation set. The training set consists of 208 volumes, which were used to train our method. This subset allowed the model to learn the necessary features and patterns from the data. The Verification Collection included 140 volumes. It served as a separate subset used to evaluate the performance of the trained model and to adjust any hyperparameters. This step ensured that the model generalized well to unseen data and helped prevent overfitting.

This group was reserved for final evaluation and provided an unbiased evaluation of the method's performance on unseen data. The test set allowed us to measure the effectiveness and robustness of the approved recording method,Figure .10

```
# separer les donnes de train et de validation

npz = np.load('OASIS_data.npz')
x_train = npz['train']
x_val = npz['validate']

print('train shape:', x_train.shape)
vol_shape = x_train.shape[1:]
print('val shape:', x_val.shape)
vol_shape = x_val.shape[1:]

# affichage (number_of_samples, height, width)
```

```
train shape: (208, 192, 160)
val shape: (141, 192, 160)
```

```
# affiche un exemple de données de train
nb_vis = 8
idx = np.random.randint(0, x_train.shape[0], [8,])
example_digits = [f for f in x_train[idx, ...]]

# affichage
ne.plot.slices(example_digits, cmaps=['gray'], do_colorbars=True);
```

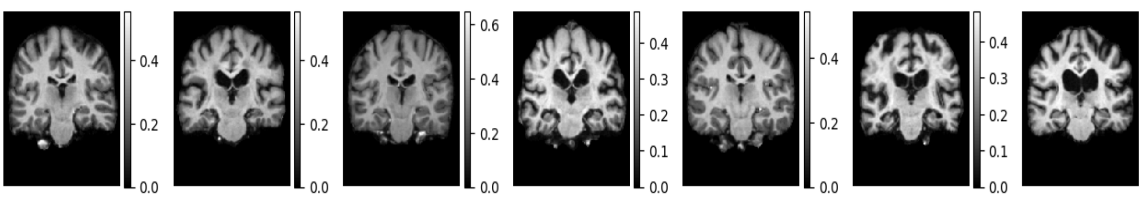


Figure 10 : Divide the dataset and exhibit a sample of the training dataset.

3. Programming :

3.1.Mount drive in colab :

After connecting google Colaboratory with google Drive (GPU , storage and RAM allocation) , there are two ways to connect to google Drive from google Colab

- The first method is to mount the drive manually as shown in figure 11 :

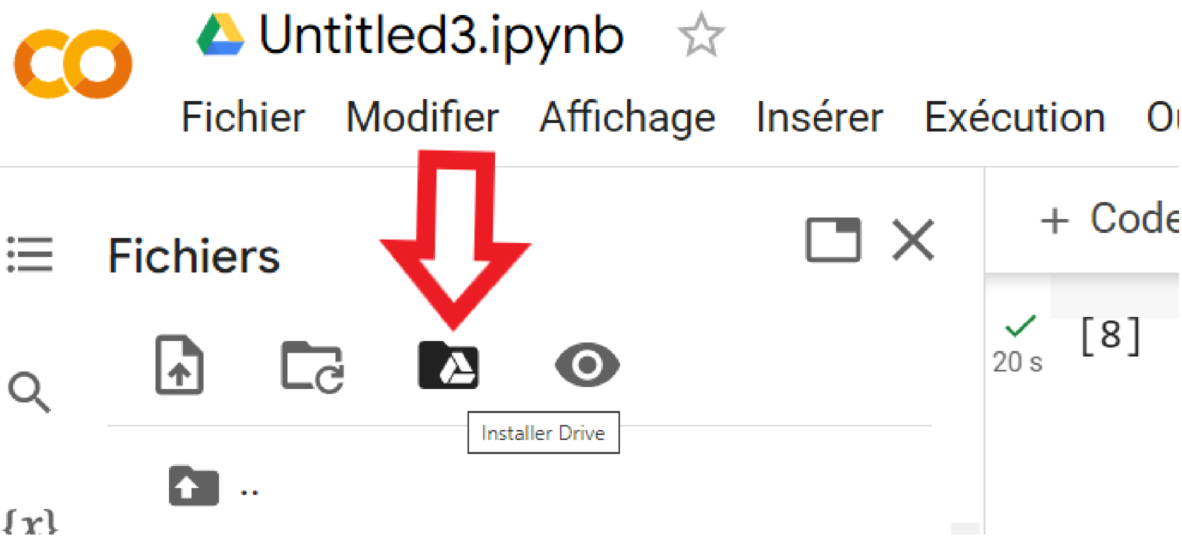


Figure 11: Access google Drive from Colab

- The second method is to execute this command in google Colaboratory to get access as shown in Figure 12 :

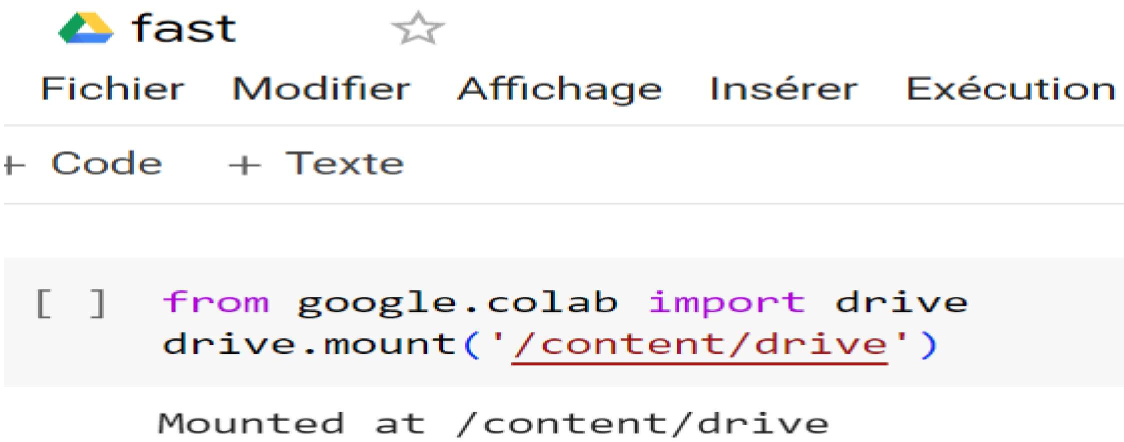


Figure 12: Access google Drive from Colab

In both cases , you must grant access permission by proccing the option “allow” when this window appears :

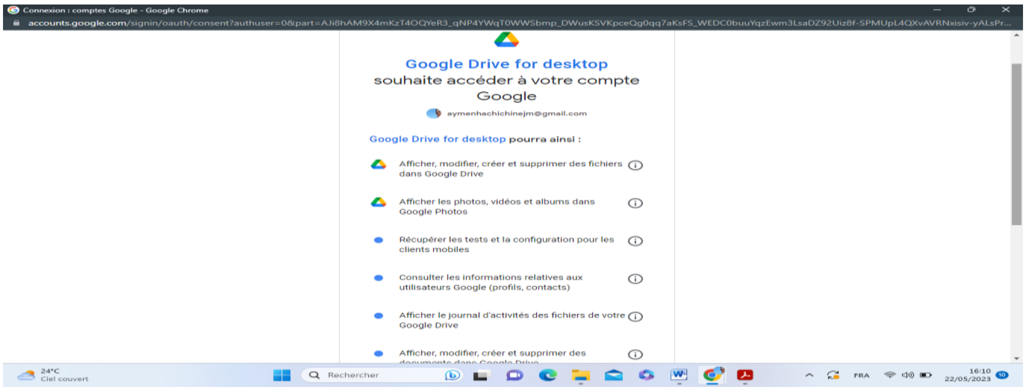


Figure 13: Drive Permission window

3.2. Importing libraries :

To use the previously mentioned libraries and its classes , it is necessary to call them using the following instructions :

```
import numpy as np
import torch.utils.data as Data
import nibabel as nib
import torch
import itertools
from scipy.ndimage import zoom

import matplotlib.pyplot as plt
```

Figure 13: Importing Lib

4. Results :

Results for registration moving image to fixed image , Figure 14 .

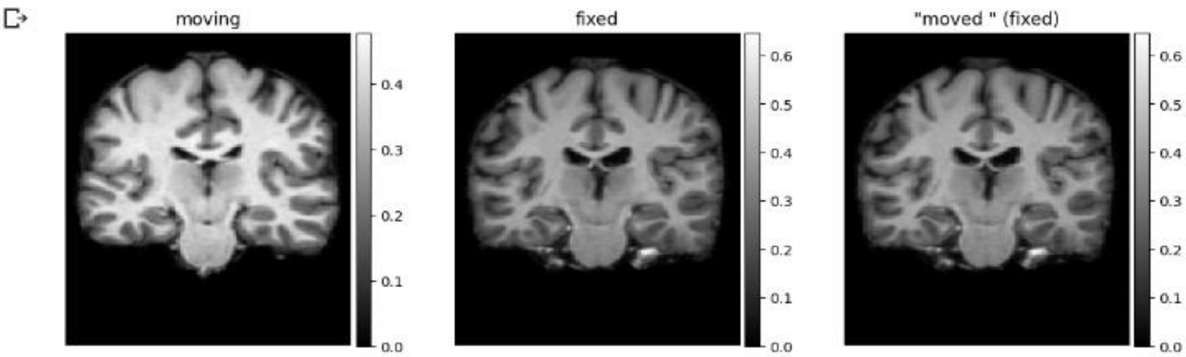


Figure 14: Result 1

- We see that the registration between the fixed image and the moving image was done well .

- To check the registration quality we present in the next figure the losses function :

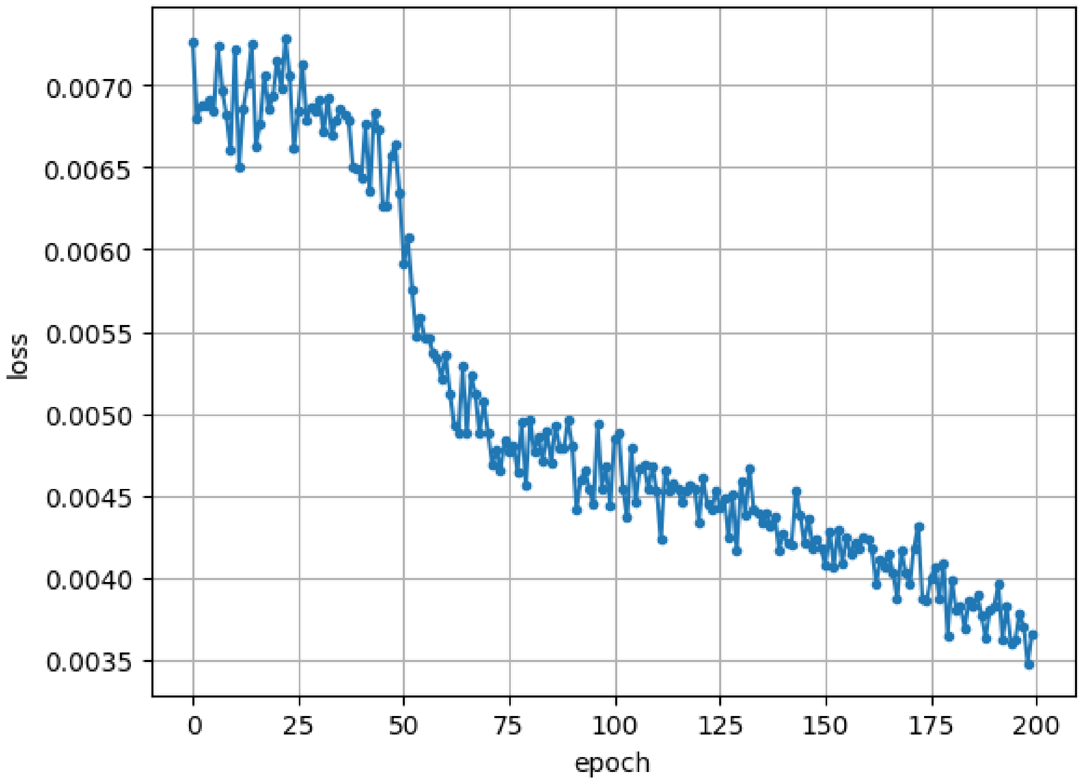


Figure 15: losses Result 1

- We see thet the loose decries over the time and converge to a minimal value after 200 epoch; which mean that the registration accuracy has been increased and maximized over the epochs .

Other results :

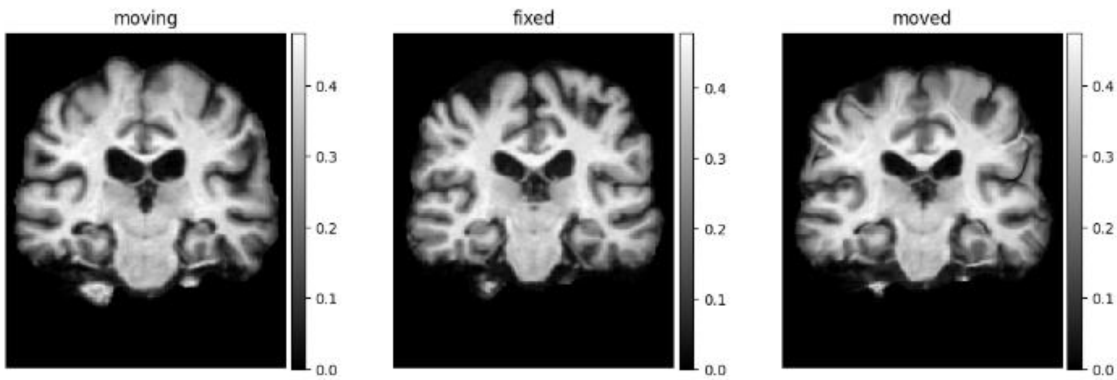


Figure 16: Result 2

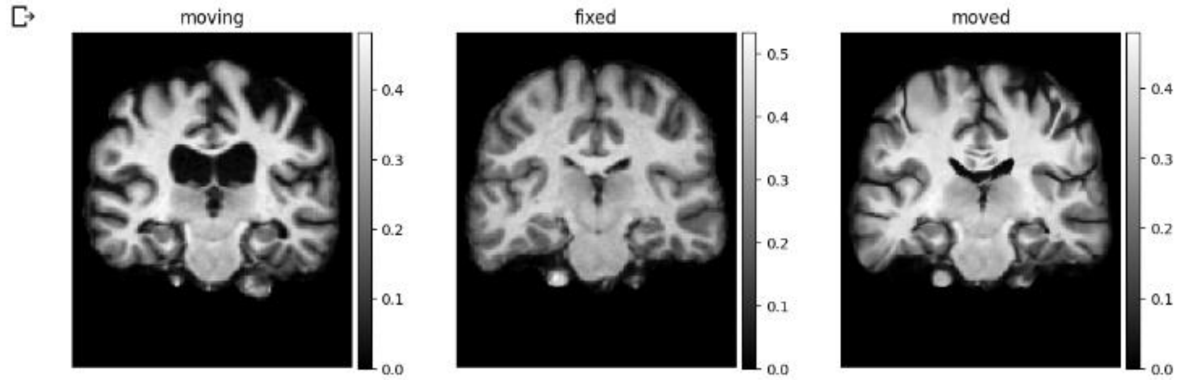


Figure 17: Result 3

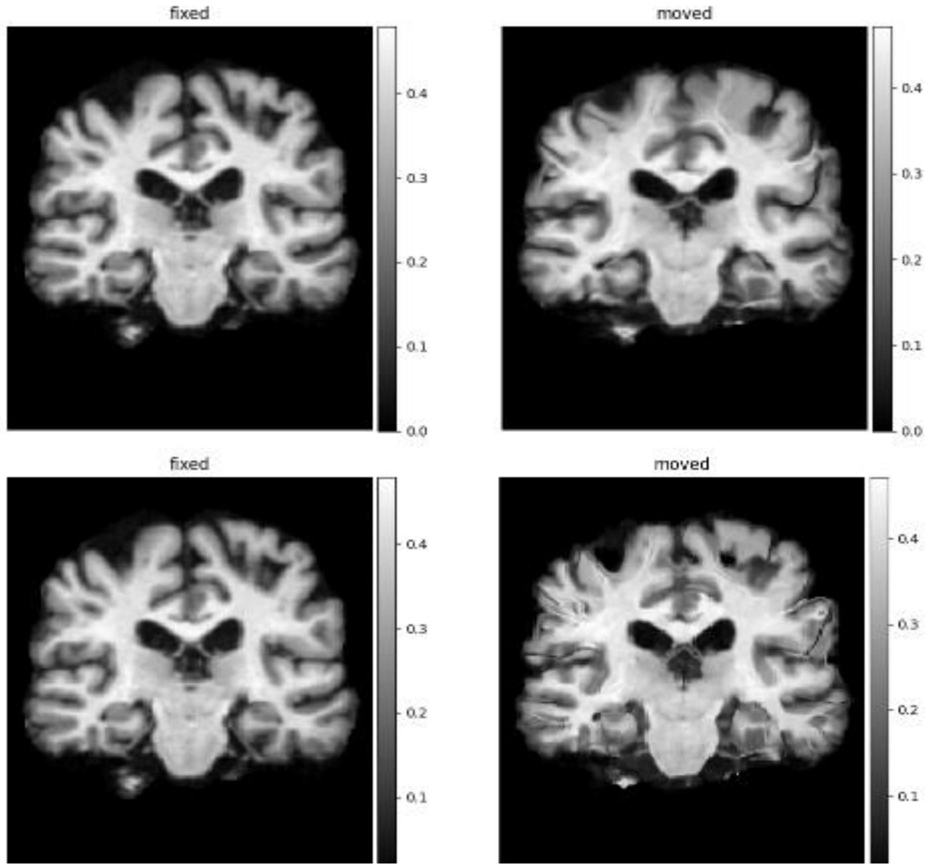


Figure 18: Result

- We re-established that bulk rumors over time converge to a minimum value after 200 epochs in other results which means that the accuracy of registration has been increased and maximized throughout the ages.

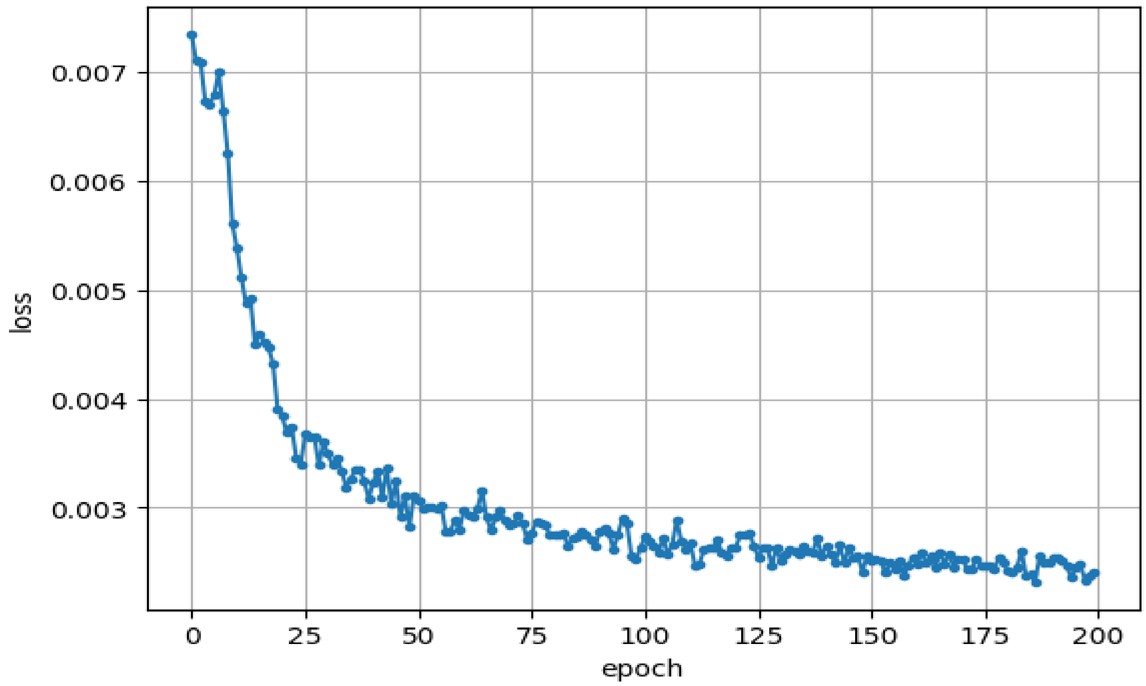


Figure 19: losses

- You also verified the recording using the “checkerboard method” as described :

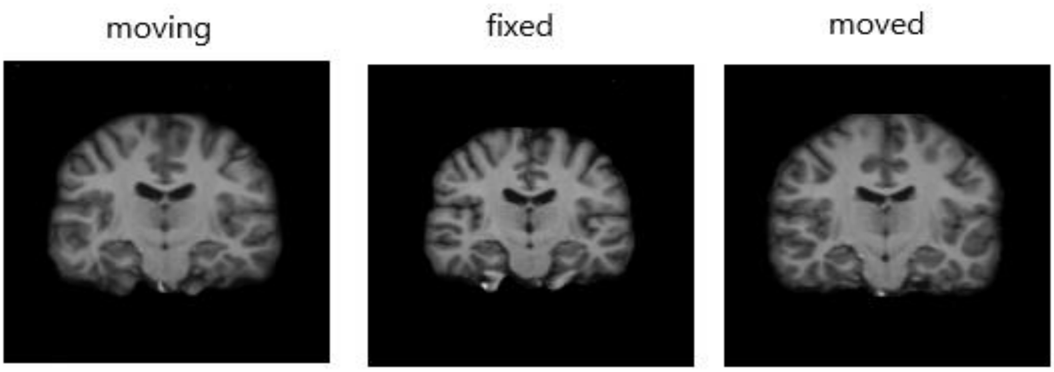


Figure 20: Result 5

- We see that the result obtained using the ‘checkerboard method’ in the image of the brain(figure 21)a very clear image with a large percentage. This indicates the accuracy of the registration .

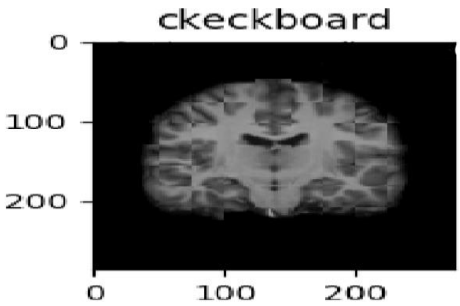


Figure 21: ckekboard fixed and moved

Chapter 4

Conclusion

This chapter contains The first section of this chapter gives a basic overview of our work. In the second section, we provided a study on Google Colab and the features that this platform offers as resources and constraints. and learn how to run any code on this platform.

The explanation of the results, which show that our method can outperform both the conventional method and the learning-based methods in terms of registration accuracy and quality of the strain fields, took up the majority of the last section.

General Conclusion

This project explored a deep learning application in registration distorted medical images. We used a 425 T1 brain MRI dataset obtained from the OASIS dataset. The main objective of this project was to develop an accurate and rapid form for registration medical images. through various stages of assembly Data, exploration, preparation and modeling, we were able to record the images This project highlights the growing importance of artificial intelligence in the field health, in particular the registration of medical images.

In conclusion, this thesis project has contributed to the advancement of medical image registration by exploiting techniques of artificial intelligence and deep learning. It paves the way for future research and applications in the field of Medicine .

In future research, we will develop our CNN models to:

- Registration 3D medical images.
- Increase accuracy to a higher level.
- Building other models based on deep learning and verifying their accuracy in other ways.

Reference

[1] EL-Hachemi Guerrout and Samy Ait-Aoudia. “Performance in the segmentation of medical images”. Doctoral thesis. July 2018. doi: 10.13140/RG.2.2.36804.88964 (see p. 6)
[2] Grist, T. M., & Korosec, F. R. (2018).Medical Imaging. In R. A. Meyers (Ed.), Encyclopedia of Sustainability Science and Technology (pp. 7279-7309). Springer**.**
[3] Gonzalez, R. C., & Woods, R. E. (2018). Digital Image Processing (4th ed.). Pearson.
[4] Morrison, M. (2019). Understanding Digital Color (2nd ed.). Routledge.
[5] "Handbook of Medical Image Processing and Analysis" edited by Isaac Bankman (Chapter 23: Binary Image Processing and Analysis)
[6] Jain, A. K. (1989). Fundamentals of digital image processing. Prentice-Hall, Inc.
[7] Forsyth, D. A., & Ponce, J. (2012). Computer vision: A modern approach (2nd ed.). Prentice Hall.
[8] "History of Medical Imaging." Society for Imaging Informatics in Medicine, 2019, https://siim.org/page/history_of_medical_imaging.
[9]"Medical imaging." National Institute of Biomedical Imaging and Bioengineering. https://www.nibib.nih.gov/science-education/science-topics/medical-imaging (accessed April 23, 2023).
[10] Radiological Society of North America. (2021). X-ray (Radiography) - Body. Retrieved from https://www.radiologyinfo.org/en/info.cfm?pg=xraybody
[11] American College of Radiology. (2021). Computed tomography (CT). Retrieved from https://www.acr.org/Clinical-Resources/Radiology-Safety/Radiation-Protection/Computed-Tomography-CT
[12] National Institute of Biomedical Imaging and Bioengineering. (2021). Magnetic Resonance Imaging (MRI) - Body. Retrieved from https://www.nibib.nih.gov/healthcare-providers/treatments-and-therapies/magnetic-resonance-imaging-mri-body.
[13] "Ultrasound - Mayo Clinic." Mayo Clinic, Mayo Foundation for Medical Education and Research, 19 Nov. 2020, www.mayoclinic.org/tests-procedures/ultrasound/about/pac-20395177.
[14] RadiologyInfo. (2021). Nuclear Medicine Imaging. Retrieved from https://www.radiologyinfo.org/en/info.cfm?pg=nuclearmed.
[15] Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., ... & Puliafito, C. A. (1991). Optical coherence tomography. Science, 254(5035), 1178-1181.
[16] National Institute of Biomedical Imaging and Bioengineering. Positron Emission Tomography (PET). Retrieved from https://www.nibib.nih.gov/science-education/science-topics/positron-emission-tomography-pet
[17] "What is SPECT Imaging?" Johns Hopkins Medicine, 11 June 2021, www.hopkinsmedicine.org/radiology/nuclear_medicine/what-is-spect-imaging.html.
[18] Fenohery Tiana ANDRIAMANAMPISOA ‘ Recalage multimodal 3D utilisant le modèle élastique, la méthode des éléments finis et l'information mutuelle dans un environnement parallèle’ Novembre 2008.

[19] Szeliski, R. (2010). Computer Vision: Algorithms and Applications. Springer Science & Business Media.
[20]Maintz, J. B., & Viergever, M. A. (1998). A survey of medical image registration. Medical image analysis, 2(1), 1-36.
[21] Pluim, J. P., Maintz, J. B., & Viergever, M. A. (2003). Mutual-information-based registration of medical images: a survey. IEEE Transactions on medical imaging, 22(8), 986-1004.
[23] Lamraoui Mohamed ,Recalage des paires bilatérales et temporelles dans le cadre de l’analysedes mammographies 2018/2019
[25] Hartley, R., & Zisserman, A. (2003). Multiple View Geometry in Computer Vision (2nd ed.). Cambridge University Press.
[26] Galdi 1994 G. P. Galdi. An introduction to the mathematical theory of the Navier-Stokes
[27] Gourinat 2001 Yves Gourinat. Introduction à la dynamique des structures. Cépaduès éditions.Mars 2001
[28] "An Introduction to the Finite Element Method" by J. N. Reddy.2013
[29] Medical Image Registration" by Jasjit S. Suri, Paul Suetens, and Yongxia Zhou.2011
[30] Vladimir Fonov, Alan C Evans, Kelly Botteron, C Robert Almli, Robert C McKinstry, D Louis Collins, Brain Development Cooperative Group, et al. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*, 54(1):313–327, 2011.
[31] "Diffeomorphic Demons: Efficient Non-parametric Image Registration",Tom Vercauteren, Xavier Pennec, Aymeric Perchant, Nicholas Ayache, NeuroImage,2009 DOI: 10.1016/j.neuroimage.2008.10.040
[32] "A Review of Classical Diffeomorphic Image Registration Methods" Chunfeng Lian, Tamas Ungi, Gabor Fichtinger,Journal of Medical Imaging and Health Informatics, 2016,DOI: 10.1166/jmihi.2016.1923
[33] "Deep Learning Techniques for Deformable Medical Image Registration: A Survey",Guorong Wu, Miaomiao Zhang, Dijia Wu, Dinggang Shen, Journal of Medical Image Analysis,2020,DOI: 10.1016/j.media.2019.101660
[34] EL-Hachemi Guerrout et Samy Ait-Aoudia. « Performances dans la segmentation d’images médicales ». Thèse de doct. Juil. 2018. doi : 10.13140/RG.2.2.36804.88964 (cf. p. 6)
[35] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. Unet: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.
[36]Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. An unsupervised learning model for deformable medical image registration. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 9252–9260, 2018 [37]Leila M.Fonesca and Max H.M.Costa,”Automatic Regestration of Satellite Images”,Proceedings on IEEE transaction of computer society,1997,pp 21 9-226.
[38]BS Reddy, 1996,Y.Keller, 2002
[39]Zhongxiu Hu and Scott T.Acton,”Morphological Pyramid Image Registration” 4th IEEE south west symposium 2000.P22
[40] "Genetic Algorithms for Image Registration: A Review", N. Karimian, F. A. B. Tahir, R. Sallehuddin, S. Sulaiman, R. Z. R. M. Hashim,2017,DOI: 10.1016/j.jvcir.2017.06.012

[41] Russell, S. J., Norvig, P., & Davis, E. (2019). Artificial Intelligence: A Modern Approach. Pearson.

[42] Jack Copeland and Didier Galmiche, Dominique Larchey-Wending et Joseph Vidal-Rosset (dir.), vol. 157. open edition journal, 2012.

[43] A.M. Turing, “Computing Machinery and Intelligence,” in Parsing the Turing Test,

[44] T. Claude, LES RESEAUX DE NEURONES ARTIFICIELS, INTRODUCTION AU CONNEXIONNISME, vol. 130. 1992.

[45] M. John and A. F. Edward, “Arthur Samuel: Pioneer in Machine Learning,” 1990.

[46] M. Minsky, C. Shannon, and N. Rochester, “The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years,” 2006. Accessed: May 05, 2022. [Online].

[47] J. Mcdermott, “RI: A Rule-Based Configurer of Computer Systems*,” 1982. Accessed: May 05, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/0004370282900212

[48] L. MILTON, “History of AI Winters,” Sep. 05, 2018. https://www.actuaries.digital/2018/09/05/history-of-ai-winters/ (accessed May 05, 2022).

[49] L. Rougetet, “Un ordinateur champion du monde d’Échecs : histoire d’un affrontement homme-machine,” Sciences du jeu, no. 5, Feb. 2016, doi: 10.4000/sdj.598.

[50] H. Gorny, “Les apports du machine learning dans la synthèse de molécules médicamenteuses.” [Online]. Available: https://dumas.ccsd.cnrs.fr/dumas-03188573

[51] F. David, “Types of Machine Learning Algorithms You Should Know | by David Fumo | Towards Data Science,” Jun. 15, 2017. https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861 (accessed May 16, 2022).

[52] M. C. Nwadiugwu, “Neural Networks, Artificial Intelligence and the Computational Brain.” Accessed: May 06, 2022. [Online]. Available: https://www.researchgate.net/publication/348675268_Neural_Networks_Artificial_Intelligence_and_the_Computational_Brain

[53] L. Medsker, “Integrating Expert Systems and Neural Computing for Decision Support,” 1994.

[54] T. Claude, *LES RESEAUX DE NEURONES ARTIFICIELS, INTRODUCTION AU CONNEXIONNISME*, vol. 130. 1992.

[55] Lambert R., “Page 4 – Pensée Artificielle,” 2019. Accessed: Jun. 08, 2022. [Online]. Available: https://penseeartificielle.fr/page/4/?cat=-1

[56] “What is Deep Learning? | IBM.” https://www.ibm.com/cloud/learn/deep-learning (accessed May 08, 2022).

[57] “Apprentissage profond et apprentissage automatique - Azure Machine Learning | Microsoft Docs.” https://docs.microsoft.com/fr-fr/azure/machine-learning/concept-deep-learning-vs-machine-learning (accessed May 20, 2022).

[58] "Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville
URL: http://www.deeplearningbook.org/

[59] L. Yann, B. Léon, B. Yoshuma, and H. Patrik, “Gradient-Based Learning Applied to Document Recognition,” 1998.

[60] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.1556

[61] C. Szegedy et al., “Going Deeper with Convolutions.”

[62] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Dec. 2015, [Online]. Available: http://arxiv.org/abs/1512.03385
[63]. Wu G, Kim M, Wang Q, Gao Y, Liao S, Shen D. Unsupervised deep feature learning for deformable registration of mr brain images. International Conference on Medical Image Computing and Computer-Assisted Intervention (Springer) (2013) 18(2) p. 649–56.
[64]. Liao R, Miao S, de Tournemire P, Grbic S, Kamen A, Mansi T, et al. (2017). An artificial agent for robust image registration, in: Proceedings of the AAAI conference on artificial intelligence Palo Alto, California USA: AAAI Press, Vol. 31.
[65]. Yang X, Kwitt R, Niethammer M. (2016). Fast predictive image registration, in: Deep Learning and Data Labeling for Medical Applications Springer International Publishing AG:48–57.
[66]. Fan J, Cao X, Yap PT, Shen D. Birnet: Brain image registration using dual supervised fully convolutional networks. Med Image Anal (2019) 54:193–206. doi: 10.1016/j.media.2019.03.006
[67]. Fu Y, Lei Y, Wang T, Patel P, Jani AB, Mao H, et al. Biomechanically constrained non-rigid mr-trus prostate registration using deep learning based 3d point cloud matching. Med Image Anal (2021) 67:101845. doi: 10.1016/ j.media.2020.101845
[68]. Hessam Sokooti, Bob de Vos, Floris Berendsen, Mohsen Ghafoorian, Sahar Yousefi, Boudewijn P.F. Lelieveldt, Ivana Iˇsgum, and Marius Staring. 3D Convolutional Neural Networks Image Registration Based on Efficient Supervised Learning from Artificial Deformations(2019) 1908.10235v1.
[69]. Hu Y, Modat M, Gibson E, Li W, Ghavami N, Bonmati E, et al. Weakly supervised convolutional neural networks for multimodal image registration. Med Image Anal (2018) 49:1–13. doi: 10.1016/j.media.2018.07.002
[70]. Hering A, Häger S, Moltz J, Lessmann N, Heldmann S, van Ginneken B. Cnn-based lung ct registration with multiple anatomical constraints. Med Image Anal (2021) 72:102139. doi: 10.1002/mp.14674
[71]. Xu J, Chen EZ, Chen X, Chen T, Sun S. (2021). Multi-scale neural odes for 3d medical image registration, in: International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2021 Springer Nature Switzerland AG:213–23.
[72] "Welcome to Colaboratory" - This is the official introduction to Google Colab provided by Google. It gives an overview of the platform, its features, and how to get started.
[73] "Python Crash Course"(2015) by Eric Matthes.
[74] Tiejun Y, Xinhao B, Xiaojuan C, Yuehong G, Lei L , (2022) GraformerDIR: Graph convolution transformer for deformable image registration. https://doi.org/10.1016/j.compbiomed.2022.105799.