LUISS

Corso di laurea in DATA SCIENCE & MANAGEMENT

Cattedra Big Data and Smart Data Analytics

Healthcare Innovative Technologies & Tumor Detection through CNN

Prof.

IRENE FINOCCHI

Prof.

BLERINA SINAIMERI

Matr.

ANGEL MARFIUC

Anno Accademico 2022/2024

Index

Chapter IV

Brain Tumor Detection with CNN in practice

Introduction

Cancer is still considered one of the toughest challenges in modern medicine, and when it comes to brain tumors, they are described as particularly elusive enemies. This disease is found in the cranial vault, and it creates distinct challenges because it usually develops within this region, where detection is difficult. To overcome these unique problems, advanced medical imaging modalities such as Magnetic Resonance Imaging (MRIs) and Computed Tomography (CTs) had provided clinicians some mighty tools and valuable insights to combat this threat. Artificial intelligence, in particular through Convolutional Neural Networks (CNNs), stands at the forefront of innovation with its capabilities in deciphering the complexities of medical imaging and cancer detection.

This thesis starts with the overview of the progression of digital technologies, of imaging modalities and integration of machine learning algorithms in the field of healthcare (Chapter I). This chapter also discusses the recent developments in the field of digital health, with a focused exploration on the use of smart devices (Internet of Things), Big Data, Artificial Intelligence (AI) and Computer Vision (CV) techniques for solving practical and day-to-day problems in the field of healthcare. This thesis highlights the potential of data-driven solutions in enhancing clinical decisions and detection accuracy, specifically for brain tumor detection.

Convolutional neural networks (CNNs) are one of those techniques that the reader should be aware of. Chapter II of this thesis explains CNNs, their architecture, training procedures and performance metrics.

Moreover, the thesis not only demonstrates the potential of such technologies to extract important insights from complex medical images, but also lays the foundation for their use in brain tumor diagnosis. In order to understand how tumor diagnosis and detection works, the reader needs some knowledge about tumor (Chapter III). In fact, this chapter will provide an overview of cancer and brain tumors.

Finally, in Chapter IV the attention returns to the main topic, which is using CNNs for brain cancer detection and diagnosis, by developing and presenting a CNN model. Furthermore, this section also highlights how these systems can provide more accurate brain tumor detection by describing their architecture, training methods and performances. The practical part of the thesis focuses on the use of CNNs for brain tumor detection, through a code.

In conclusion, this paper aims to demonstrate the transformative and innovative potential of CNNs and image recognition technologies in advancing and developing healthcare, particularly in the context of brain tumor diagnosis and management. It also underline the significance of data-driven approaches in enhancing diagnostic accuracy and clinical decision-making.

Chapter I

Digital and Innovative Technologies in Healthcare

1.1 Medical Technologies

The integration of digital technologies in healthcare is transforming a lot the accessibility and adaptability of medical services to the public. This includes providing details regarding health, diseases, complications, and medical research easily accessible online. Moreover, even in low-income nations the access to diagnostic and medical consultation services is growing. In fact, remote and rural residents can receive excellent consultations through telemedicine, and online pharmacy platforms facilitate it, being easier to get prescription drugs without having to travel.

The use of artificial intelligence (AI) in biomedicine, healthcare, and medical education is another rapidly developing sector. AI has the potential to improve therapy optimization and diagnostic accuracy, which will improve therapeutic outcomes, patient satisfaction, and cost-effectiveness. AI also facilitates clinical trials and biological research, and it simplifies jobs that usually required considerable amounts of manual labor. But even with significant advancements, AI is now unable to fully replace human participation in biomedical research and treatment.

Furthermore, a new era in healthcare is emerging, one in which the abundance of biomedical data are playing increasingly significant roles. Precision medicine, for example, considers various aspects of the patient's data, such as variability in molecular traits, environment, electronic health records (EHRs), and lifestyle, in order to "ensure that the right treatment is delivered to the right patient at the right time. The abundance of these biomedical data presents both enormous potential and difficulties for health care research. In fact, developing accurate medical tools based on data-driven methods and machine learning requires addressing the challenge of examining the relationships between all the multiple pieces of information in these data sets. Prior research has attempted to achieve this goal by tying together various data sources to create collaborative knowledge bases that can be utilized for discovery and prediction analysis.

Predictive tools based on machine learning techniques have not been widely implemented in medicine, despite the considerable potential of existing models. This is due to the high dimensionality, heterogeneity, temporal dependence, sparsity, and irregularity of biological data. Furthermore, there are still many obstacles to overcome before they can be fully utilized [16-17]. For example, the numerous medical ontologies that are utilized to generalize the data which often include contradictions and inconsistencies, intensify these difficulties [4]. In fact, distinct expressions of the same clinical trait could be found throughout the data.

Moving to the Deep Learning, it enabled medical computer vision. Decades of research have been invested into developing computer vision (CV), which aims to give computers the ability to meaningfully comprehend visual stimuli. Machine perception is an extensive area that includes low-level tasks, such edge recognition, and complex tasks, such comprehending entire scenes.

In the past ten years, three elements have contributed significantly to the advancements of this technology:

(1) The development of deep learning (DL), an area of machine learning that allows extremely complicated functions to be learned from scratch using only raw data.
(2) Advancements in localized computing power through GPUs
(3) The open-sourcing of large labeled datasets with which to train these algorithms.

These three factors combined have made possible for individual researchers to access the resources they require to continue researching in this field, and as the research community grew exponentially, so did the progress [18].

Concluding with The Internet of Things, or IoT, is becoming so widely used in various applications that it is becoming more and more important in our daily lives. IoT technology is also being developed in the healthcare monitoring system to help patients receive efficient emergency assistance. Additionally, it is utilized as an E-health application for a variety of purposes, including computer-assisted rehabilitation, emergency alerting, and early medical issue identification.

People's smartphones have become an essential component of daily life, and they are linked to sensors to track their health. The diagnostic equipment provides a variety of data that this sensing-based surveillance systems gathers and analyzes for effective, automated healthcare management. The IoT healthcare system offers effective tracking and monitoring that enhances the management of human resources. Furthermore, healthcare data is handled via cloud computing, which offers resource sharing features such flexibility, early detection of security issues, parallel processing, data service integration with scalable data storage, and adaptability. The cloud data center's high energy usage further raises the price of cloud computing however, low latency and energy consumption cloud services are absolutely required for a health monitoring system[18].

There are also some challenges: for example, the wearable or sensor that is implanted in patients within the Internet of Things (IoT)-based healthcare system has a very short battery life. The frequent charging of these devices and mobile devices may fatigue the patients and require engagement of the nurse, which affects the user experience.

Another issue in healthcare monitoring is the security that the data can be easily corrupted by the attacker or hackers. In fact, for an effective data sharing, an IoT-based healthcare system that protects patient privacy must be developed [19].

In the next sections the thesis will discuss more in detail these technologies, starting with IoT in section 1.2, moving to Big Data and AI respectively in section 1.3 and 1.4, and finally concluding with Computer Vision in the 1.5 paragraph.

1.2 Internet of Things (IoT)

In this section the thesis will discuss more in detail the IoT technology and its structure. IoT is an evolving technology within the internet landscape, characterized by real-time connectivity of various objects. Its capacity to turn common objects into smart devices makes it applicable to a wide range of sectors. This shift has important long-term ramifications for clinical management of patients' physiological data as well for health administration and monitoring. Patients are equipped with sensors that collect data, which is transferred to control units and ultimately to systems that monitor health. This data is frequently kept on cloud servers, which makes it easier to manage huge datasets with improved security.

In fact, security is an important component of IoT because integrity and confidentiality of data may be compromised during the transfer from sensors to cloud storage. Furthermore, it is difficult to encrypt data from devices with limited resources. Despite these challenges, due to its distributed nature, the cloud is a perfect place to store medical data since it allows clinicians to view patient records from any location and offers flexibility for remote patient care [10].

There are different applications of these technologies in the healthcare field, such as early detection of diseases, patient monitoring and provide emergency services especially for hearth patients.

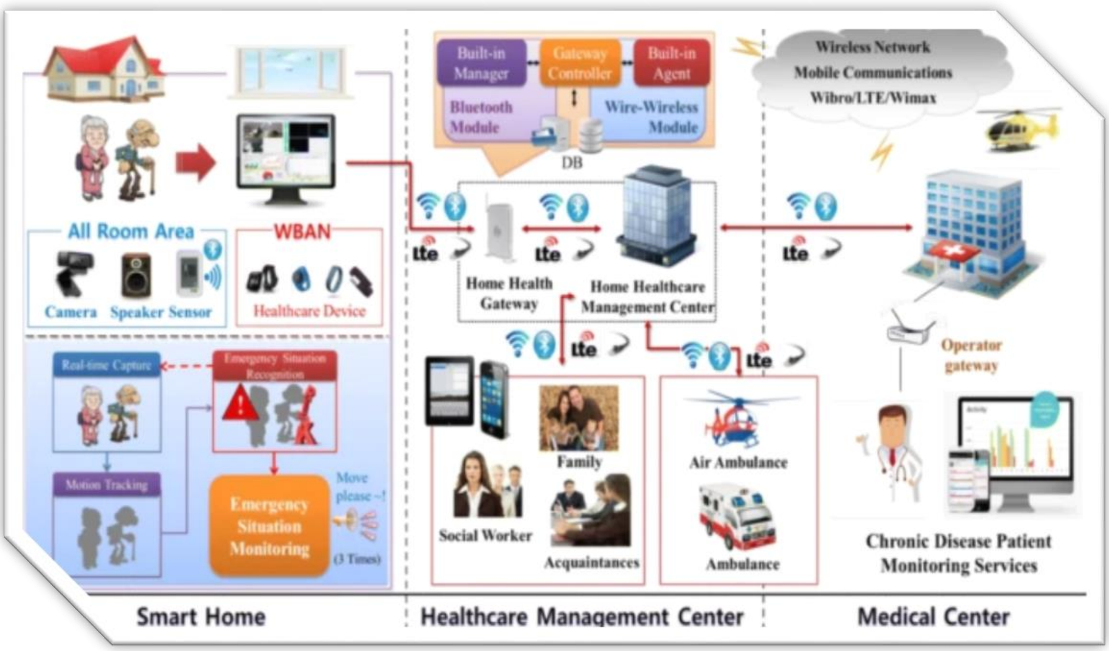


Figure 1 : IoT Architecture. "Challenges and opportunities in IoT healthcare systems" [38]

In Figure 1, we can see how the architecture of IoT is composed.

Architecture Overview:

- Data Collection: Smart home devices and wearable healthcare devices collect health and environmental data.
- Data Transmission: Collected data is sent to the home health gateway via Bluetooth or other wireless modules.
- Data Processing: The home health gateway processes the data and sends it to the home healthcare management center for further analysis and storage.
- Emergency Detection: If an emergency is detected (for example a fall), alerts are sent to family members, social workers, or acquaintances, and emergency services are dispatched if necessary.
- Medical Center Communication: Data is transmitted to the medical center using wireless networks for chronic disease monitoring and further medical analysis.

This interconnected system ensures continuous monitoring and quick response in case of emergencies, enhancing the healthcare and safety of patients, especially those with chronic conditions or the elderly living alone.

1.3 Big Data in Healthcare

This section of the thesis focuses its attention on Big Data technologies.

According to the Gartner Glossary, "Big Data is characterized by high-volume, high-velocity, and/or high-variety information assets that necessitate cost-effective, innovative forms of information processing enabling enhanced insight, decision-making, and process automation." [20] Massive, complex datasets with exponential growth rates in volume, velocity, and diversity of data are known as big data. Big data has been used more and more in the healthcare sector in recent years to improve management, analysis, and prediction capacities. For example, the use of Electronic Health Records (EHR) makes it easier to store, organize, and process patient data more quickly.

Big data technologies have a variety of benefits, including the ability to reduce medical errors in medical practice management and enhance service quality, efficiency, and cost-effectiveness. The usage of big data in healthcare has increased dramatically as a result of the broad use of EHRs.

Beyond clinical uses, big data in healthcare includes genetics, healthcare provider data (pharmaceuticals and logistics), and other pertinent information. Globally, there is an ongoing movement toward adopting big data as a platform to facilitate, improve, accelerate, and streamline biomedical research. It provides numerous opportunities to find innovative solutions in the domains of pharmacology, genetics, epidemiology, medicine, and biology. Big data also plays a crucial role in sorting and analyzing the rapidly expanding volume of medical information. For example, clinical data from a single person can produce about 0.4 terabytes of data over their lifetime, genomics data about 6 terabytes, and additional data up to 1,100 terabytes. Big data platforms manage different challenges such as streaming data analysis, deciphering numerous data formats, dealing with data uncertainty, among other issues relating to the velocity, variety, and authenticity of public health data. In order to maximize diagnoses and treatment through individualized approaches, effective healthcare requires extensive integration of biomedical data from multiple sources, including genetics, pharmaceutical information, and insurance records. In the near future, big data analysis from medical systems may be used to create innovative healthcare plans and establish the foundation for groundbreaking advancements in personalized medicine.

The idea of e-health has received official approval and significant traction in many nations.

E-health platforms enable the automatic and timely collection and analysis of medical information, thus offering safe, equitable, high-quality, and sustainable healthcare services personalized to patients' needs. Leading health organizations worldwide are actively implementing Electronic Health Passports to innovate the primary healthcare, streamlining medical care quality while reducing administrative limitations for the healthcare professionals [20].

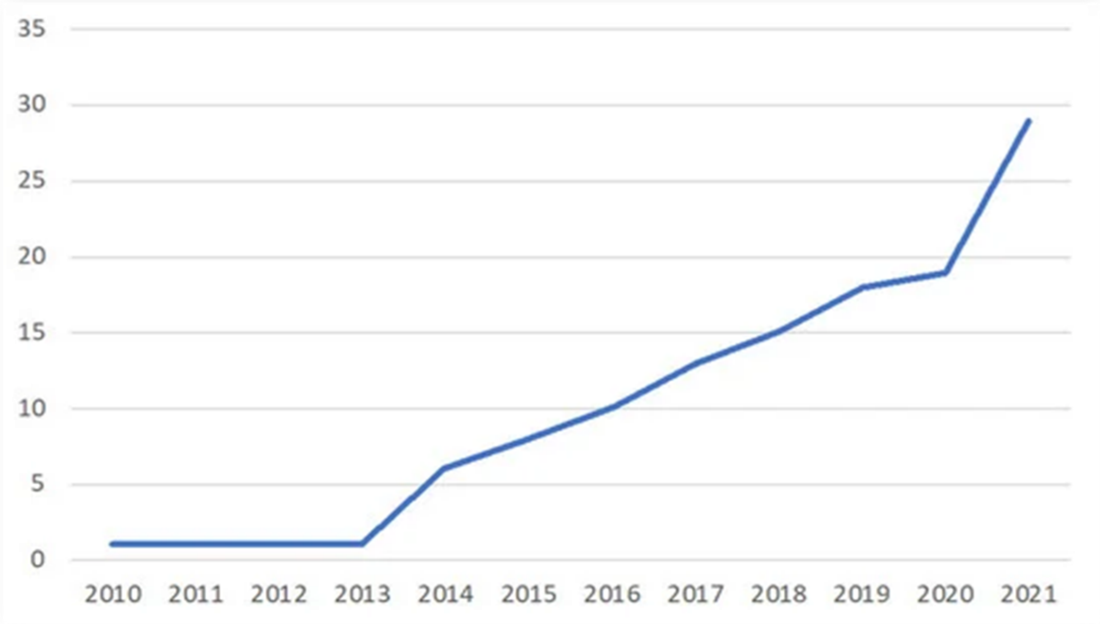


Figure 2 : Annual distribution of publications regarding Big Data [21]

<u>Big Data more detailed</u>

The term Big Data is frequently described not as a specific technology but as a phenomenon characterized by several key attributes, often referred to as the "V's":

- **V**olume: refers to the significant amount of data, posing one of the greatest challenges in Big Data Analytics.
- **V**elocity: the rate at which new data is produced, necessitating effective and real-time data management.
- **V**ariety: reflects the variety of data, which includes various types of data.
- **V**ariability: it implies the inconsistencies of the data.
- **V**eracity: it’s related to how reliable and high-quality the data is.
- **V**isualization: it refers to the ability to interpret data, extract and gather insights.
- **V**alue: the final objective of big data analytics is to exploit the enormous amount of data to extract valuable insights.

Big Data is an information asset with great volume, velocity, and variety, which requires particular technologies and techniques in order to be processed into value.

The concept of Big Data is continually evolving, moving its focus from simply just managing vast quantities of data to actually extracting value from that data.

Furthermore, Big Data is sourced and extracted from a variety of channels, and the data have different characteristics or are processed by different organizational units, resulting in the creation of Big Data Chain. For example, Big Data in the healthcare industry includes a wide range of types of data such as:

- Clinical Data: the data generated by laboratories, pharmacies, imaging centers, electronic medical records, hospital information systems and other healthcare service providers. This data includes also genomic data, genetic information, clinical notes, patient records and data from physiological monitoring.
- Biometric Data: measurements made by equipment that are monitoring variables such as blood pressure, weight, and glucose levels.
- Financial data: the records of economic transactions reflecting operational activities.
- Scientific research data: the results from investigations into new therapeutic approaches, drug development, and the development of innovative medical devices.
- Patient Data: includes self-monitoring data on information such food, exercise and sleep as well as the preferences and the satisfaction ratings.
- Social media data: information taken from various social media platforms.

These data are sourced not only from patients but also from organizations, institutions, and a variety of monitoring devices, sensors, and instruments [22].

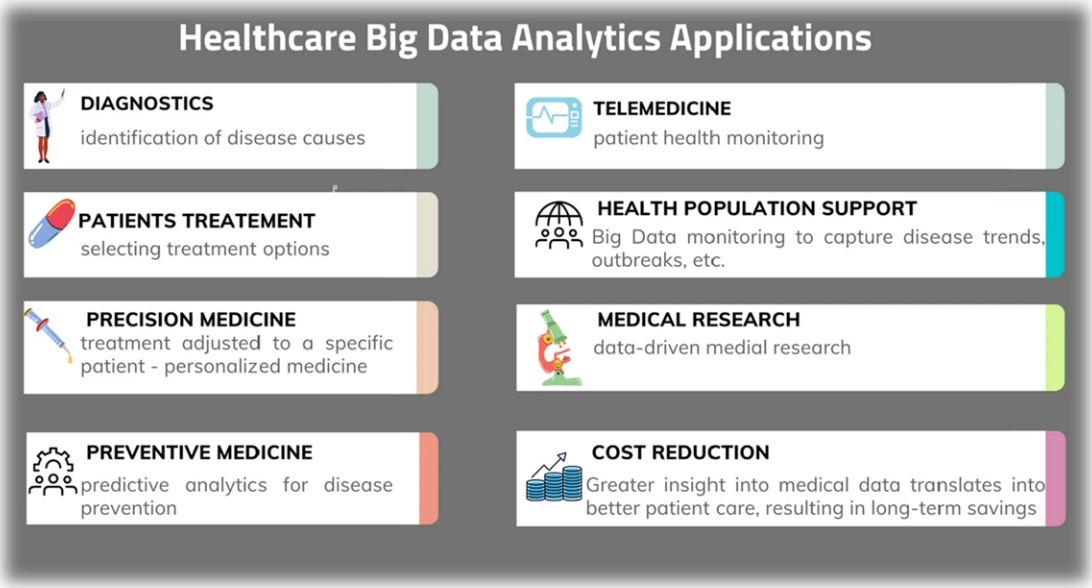


Figure 3 : Big Data Applications in Healthcare [22]

As mentioned before, managing and leveraging massive volumes of data in order to enable data-driven decisions in different industries, is one of the main challenges posed by big data.

Healthcare data analytics systems are engineered specifically to define, integrate, and visualize intricate data in a manner that facilitates a better comprehension (Fig. 4). This enhancement aims to streamline the process of obtaining, storing, analyzing, and visualizing big data within healthcare.

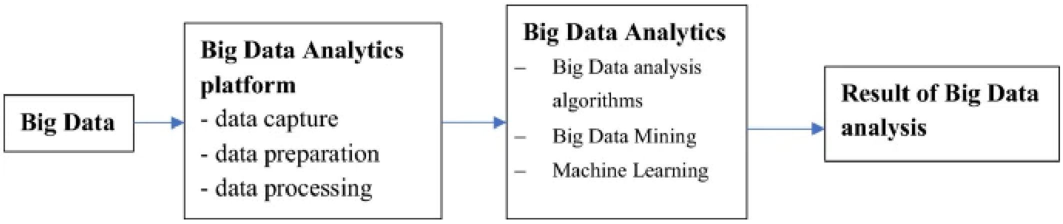


Figure 4 : Processing of Big Data Analytics systems [22]

When examining the impact of Big Data in the healthcare sector, it's crucial to consider it from three perspectives: epidemiological, clinical, and business.

- From a clinical view, Big Data analysis's objective is to enhance patient health outcomes by utilizing long-term health predictions and implementing personalized therapeutic interventions. The ultimate goal of data analysis in medicine is to facilitate personalized therapy, guaranteeing that each patient's needs are specifically met by their treatment.
- From an epidemiological standpoint, the purpose of adopting Big Data is to accurately predict the occurrence and spread of diseases within populations. By analyzing large datasets containing information about factors such as demographics, environmental conditions, lifestyle habits, and health outcomes, epidemiologists aim to identify patterns and trends that can help to anticipate and to mitigate health risks.

  Big Data analysis, for instance, could identify relationships between specific environmental conditions and the frequency of a given disease in a given area. With these information, public health experts can create specific strategies and preventative measures that reduce the disease's overall impact on the general population. In simple terms, the objective of using epidemiological Big Data is to preventatively detect and manage health risks at the population level instead of to only reacting to epidemics after they occur.

- In the business sector, leveraging Big Data analysis involves using large datasets to gain insights of individual disease and infection risks, as well as to organize commercial services accordingly. In fact, by analyzing diverse datasets that include information such as medical histories, genetic profiles, lifestyle choices, and environmental factors, businesses can develop algorithms or models to assess an individual's risk of developing certain diseases or infections. For example, analyzing a person's medical history, family medical history, and genetic predispositions can help predict their likelihood of developing conditions like diabetes, heart disease, or certain types of cancer.

  Furthermore, thanks to the insights of Big Data analysis, businesses can offer personalized or targeted commercial services aiming to promoting health and wellness or managing specific health conditions. For instance, health insurance companies could use data analytics to customize insurance plans based on an individual's health risks and needs. Similarly, companies in the healthcare industry might develop personalized fitness plans, dietary recommendations, or wellness programs customized to each individual's health profile. This personalized approach not only enhances customer satisfaction but also contributes to better overall health outcomes [22].

<u>Big Data Analytics</u>

It includes various technologies primarily rooted in data mining, including text mining, web mining, process mining, audio and video analytics, statistical analysis, network analytics, social media analytics, and web analytics. These techniques can be used to aggregate and visualize Big Data, but also to perform anomaly detection, clustering, classification and other operations on heterogeneous healthcare data. Modern data analytics techniques are also capable of exploring and leveraging unique data characteristics even from high-speed data streams and sensor data.

Big Data analytics in the healthcare industry can be divided into four important categories:

- Descriptive Analytics: this is used to comprehend past and current healthcare decisions by converting data into actionable information for analyzing healthcare decisions, outcomes, and quality, as well as making informed decisions. It involves creating reports (e.g., patient hospitalizations, physician performance, utilization management), visualization tools, customized reports, drill-down tables, or running queries based on historical data.
- Predictive Analytics: by operating on past performance, it aims to forecast the future by analyzing historical or summarized health data, identifying patterns of relationships, and extrapolating these relationships to predict outcomes. It can be also used to predict various scenarios such as patient responses to different drugs or dosages, anticipate risks, identify relationships in health data, and detect hidden patterns. Predictive analytics helps in planning healthcare resources, predicting the development of epidemics, anticipating demands for services, enabling accurate diagnosis, and determining appropriate treatments for specific diseases.
- Prescriptive Analytics: it utilizes health and medical knowledge along with data or information to recommend actions. Prescriptive analytics is employed in various healthcare areas, including drug prescriptions, treatment alternatives, personalized medicine, and evidence-based medicine.
- Discovery Analytics: aims to identify new treatments, previously unidentified diseases and medical disorders, medications (drug discovery), and more. It is essential to the advancement of healthcare innovation and medical research [22].

Descriptive, predictive, prescriptive, and discovery analytics all use distinct approaches and techniques, but many applications combine the four.

To conclude, according to research by Wang, Kung, and Byrd, [1] the benefits of Big Data Analytics in healthcare can be categorized into five main areas:

- IT Infrastructure Benefits: includes reducing system costs, avoiding unnecessary IT costs, facilitating quick data transfer among healthcare IT systems, standardizing data processing among different healthcare IT systems, and lowering IT maintenance costs related to data storage.
- Operational Benefits: include enhancing the quality and accuracy of clinical decisions, managing a large number of medical records rapidly, cutting down on patient travel time, providing immediate access to clinical data for analysis, shortening diagnostic

test times, decreasing hospitalizations related to surgeries, and exploring new research avenues.

- Organizational Benefits: this involves detecting interoperability issues quickly compared to manual methods, enhancing cross-functional communication and collaboration among administrative staff, researchers, clinicians, and IT staff, also enabling data sharing with other institutions.
- Managerial Benefits: includes gaining quick insights into changing healthcare market trends, offering decision support information to board members and department heads in daily clinical settings, and optimizing decisions related to business growth.
- Strategic Benefits: providing a comprehensive view of treatment delivery to meet future needs and creating competitive healthcare services [1].

Given the wide range of possible uses for big data analytics in the healthcare industry, this list is not exhaustive. With the use of sophisticated analytical tools, data from multiple sources may be analyzed and cross-analyses can be performed to produce more insightful data. For example, cross-analysis may involve combining patient characteristics with costs and care outcomes to identify the most medically effective and cost-efficient treatments, thereby allowing for better alignment of service provider offerings.

Big Data Challenges

Similarly to every new technology, Big Data also presents different challenges, such as problems with data collection, storage, analyzing, and visualization. The main difficulties are related to data structure: while the ideal concept of Big Data should be user-friendly and transparent, it is often fragmented, dispersed, and lacks standardization, making it challenging to aggregate and analyze effectively.

Moving to security, there are several obstacles to overcome in regard to data security, privacy, and the sensitivity of healthcare data, with confidentiality being one of the main issues.

Another challenge is the Data Standardization: integration and analysis tasks are complicated when data is stored in formats that are not compatible with various applications and technologies.

Furthermore, from a storage and transfers view, the costs associated with securing, storing, and transferring unstructured data, especially in large volumes, present financial challenges for healthcare organizations.
Moving to managerial abilities, to establish policies, ensure compliance, and supervise data-related operations, effective data governance necessitates excellent managerial abilities.
Analytical Skills also present some obstacles. In fact, in the healthcare sector, a lack of suitable analytical skills makes it difficult to analyze and use big data effectively for decision-making.
Concluding with real-time analytics, implementing real-time analytics capabilities is challenging for healthcare organizations.

Despite these challenges, healthcare big data represents a huge potential for the transformation of healthcare: improvement of patients' results, prediction of outbreaks of epidemics, valuable insights, avoidance of preventable diseases, reduction of the cost of healthcare delivery and improvement of the quality of life in general [3,4].

1.4 Artificial Intelligence in healthcare

The phrase "artificial intelligence" (AI) refers to computer systems that simulate human intelligence or cognitive processes. The potential uses of AI in the many sectors of the economy are almost limitless. The application of AI's competencies to healthcare, including the detection and treatment of various diseases, is one of the fascinating rapidly developing modern trends. The main areas in which artificial intelligence is being employed in medicine are imaging, remote consultation, medical statistics, surgery, medicinal products composition, health monitoring, and individualized treatment. Artificial intelligence (AI) and associated technologies are also starting to be used in healthcare in fact nowadays, these technologies are becoming more and more widespread in society and industry. Multiple aspects of patient care as well as administrative procedures and pharmaceutical organizations could be revolutionized by these technologies. Algorithms currently surpass radiologists in recognizing malignant tumors and in assisting researchers in cohort construction for expensive clinical studies.

However, it will be many years before AI replaces human for broad medical process domains.

In this paragraph it will be discussed some of the obstacles of the quick adoption of AI in healthcare as well as the opportunity that this technology offers to automate particular aspects of the patient care.

Types of AI in Healthcare

Artificial intelligence is a group of technologies rather than a single one. Most of these technologies have immediate relevance to the healthcare field, but the specific processes and tasks they support vary widely. The following section describes some of the AI technologies that are very significant to the healthcare industry [23].

Machine Learning and Neural Networks

Machine learning is a technique based on statistics that fits models to data and allows models to "learn" through data-driven model training. One of the most popular types of AI is machine learning; according to a 2018 Deloitte survey of 1,100 US managers whose companies were already exploring AI, 63% of the companies surveyed were using machine learning [25].

In healthcare, the most common application of traditional machine learning is precision medicine, predicting what treatment protocols are likely to succeed on a patient, based on various patient attributes and the treatment context.

Supervised learning refers to the vast majority of machine learning and precision medicine applications that need a training dataset for which an outcome variable (such as the first signs of disease) is known.

Neural networks are a more sophisticated type of machine learning. Developed in the 1960s, these networks have been widely used in healthcare research for several decades, particularly in applications involving categorization, such as predicting the probabilities that a patient will contract a specific illness.

It views problems in terms of inputs, outputs and weights of variables or 'features' that associate inputs with outputs and it has been compared to how neurons interpret signals, even if the analogy to the brain's function is relatively weak.

Natural Language Processing (NLP)

As a result of decades-long efforts by AI researchers to understand human language, Natural Language Processing (NLP) was developed. This field includes a wide range of applications, including text analysis, speech recognition, translation, and other language-related objectives. Semantic and statistical approaches are the two main methods used in NLP. Statistical natural language processing (NLP) is based on machine learning methods, particularly deep learning neural networks, which recently have made significant advancements in recognition accuracy[24].

However, it necessitates a substantial corpus or collection of language data for and effective learning performance.

NLP is essential to the healthcare industry for tasks which includes generating, deciphering, and categorizing research and clinical documentation. NLP systems are effective in transcribing patient interactions, generating reports (for example for radiological exams), evaluating unstructured clinical notes on patients, and facilitating conversational AI. Furthermore, advances in NLP have aided in the creation of complex algorithms that can extract insights from enormous volumes of medical literature, helping doctors and researchers stay updated on the most recent advancements in their respective fields [24].

Rule-based Expert Systems

In the 1980s, the most common artificial intelligence (AI) technology was rule-based expert systems, which were based on a collection of "if-then" rules.

These systems found widespread commercial use, including in healthcare, where they were employed for clinical decision support purposes. Rule-based solutions are still widely involved into the platforms of many electronic health record (EHR) providers.

Human experts and knowledge engineers must create a set of rules relevant to a particular knowledge area in order to build an expert system. These systems have difficulty to handle large numbers of rules (usually more than a few thousand) even if they function properly and are initially easy to understand. As the number of rules increases, conflicts may arise that threaten the functionality of the system. It can also take a lot of effort and time to update the rules in order to reflect the changes or developments made in the knowledge domain.

For instance, machine learning algorithms and the data-driven techniques are progressively replacing rule-based systems in the healthcare industry. Because of their increased scalability, flexibility, and adaptability, these contemporary methods have a way better response to the intricate and always changing nature of healthcare contexts. To conclude, data-driven techniques provide more precise and customized decision support by utilizing vast datasets and sophisticated algorithms, which ultimately improves patient care results [24].

Physical Robots & Robotic process automation

Nowadays, physical robots are used in various industries; globally, approximately 200,000 industrial robots are installed annually [24]. These robots are primarily employed in manufacturing and storage facilities, where they perform preprogrammed tasks including welding, product assembly, lifting, and even supply delivery to medical facilities. Recent advancements have led to the development of collaborative robots that can work alongside humans and are easily trained by physically guiding them through tasks.

Furthermore, AI features are being incorporated into the operating systems of contemporary robots, giving them a higher level of intelligence. Robots can now adapt to changing settings and carry out increasingly complex activities on their own thanks to this intelligence improved performance.

As AI continues to advance, it is expected that physical robots will also benefit from improved intelligence and capabilities over time. Surgical robots have completely changed the way that surgery is performed in the medical field since they were introduced in the USA in 2000. These robots offer less invasive procedures and better patient outcomes by providing surgeons with greater control and precision. Even though surgical robots improve doctors' abilities by providing sophisticated imagery and instrument dexterity, human surgeons still make critical decisions.

Gynecologic, prostate, and head and neck operations are common uses for robotic surgery. Patients benefit significantly from the accuracy and little invasiveness of robotic-assisted procedures in these cases.

Moving to the Robotic Process Automation (RPA), it’s a technology that automates structured digital tasks typically performed for administrative purposes within information systems.

The technique known as robotic process automation (RPA) automates structured digital processes that are normally carried out inside information systems for administrative purposes. RPA operates as computer programs on servers, imitating human users by following predefined scripts or procedures, in contrast to traditional robots. Compared to other types of AI, it has a number of benefits, including being transparent in its operations, affordable, and simple to program.

RPA is utilized in the healthcare sector to automate repetitive procedures such updating patient records and billing. When combined with other technologies, such image recognition, RPA can also be used to extract data from faxed images and input it into transactional systems. By automating administrative processes, this procedure reduces human error [24].

Furthermore, it's becoming increasingly common to combine RPA with AI, which enables RPA bots to think more critically and adapt to changing conditions.

With RPA collaborating with AI-based systems and image recognition tools to create composite solutions, this technological convergence has the potential to produce more productive and efficient automation solutions for the healthcare industry and beyond[24].

Diagnosis and Treatment applications using AI

Since the 1970s, when MYCIN (an early backward chaining expert system that used artificial intelligence to identify bacteria causing severe infections, such as bacteremia and meningitis) was launched, artificial intelligence (AI) has been active in enhancing medical diagnosis and treatment. Despite rule-based systems showed promise in accuracy, such as MYCIN, they were not used in clinical practice regardless of the early successes. Although they have been proven to be competitive to human diagnosticians, they could not be easily integrated into the current medical record systems and processes.

More recently, IBM's Watson (a computer system capable of answering questions posed in natural language, developed as a part of IBM's DeepQA) grabbed headlines for its venture into precision medicine, particularly in cancer diagnosis and treatment. Watson leverages a combination of natural language processing (NLP) and machine learning methods. But when difficulties emerged, the early enthusiasm diminished [24]. Customers realized how difficult it was to train Watson in order to deal with various forms of cancer and how challenging it was to integrate it into the systems and procedures already operating in the healthcare industry. Watson is a collection of cognitive services that include speech, language, vision, and data analysis. It is not a single product that can be accessed by APIs.

Furthermore, it faced competition from free open-source competitors such as Google's TensorFlow which made it more difficult for it to become widely adopted.

Rule-based clinical decision support systems, while effective for certain tasks, often lack the precision of algorithmic systems based on machine learning.

Since these systems are built on previously established criteria, they may find it difficult to adapt as medical knowledge advances and new data sources—like genetic, proteomic, and metabolic information—come into existence. Traditional rule-based systems may not be able to effectively incorporate and understand such various datasets due to the complexity of 'omic-based' approaches to care.

However, there is a growing shift towards integrating machine learning and AI techniques into clinical decision support systems, particularly in research labs and technology firms. Large datasets and complex algorithms are used by these innovative systems to analyze complex medical data and provide more precise and personalized recommendations for patient care. Although these advances have the potential to improve clinical practice, there are still numerous obstacles to overcome before they are widely adopted in healthcare settings. These include concerns about data privacy, regulatory compliance, and workflow integration.

Hardly a week passes without a research laboratory claiming that it has developed an AI or big data approach capable of diagnosing and treating a disease with equal or superior accuracy to human clinicians. While some of these advances involve other image types, such as retinal scans or genomically based precision medicine, many of these advances originate from the interpretation of radiological images. These developments, which are based on statistically-driven machine learning models, contribute to the development of probability-based medicine, which generally is viewed positively. Nonetheless, there are a lot of issues with medical ethics and the dynamics of patient-clinician relationships that come with this change [23,24].

Startups and technology companies are also actively addressing these issues. For example, Google has partnered with health delivery networks to use big data to develop prediction models that would notify clinicians about high-risk conditions like heart failure and sepsis. Furthermore, Google, Enlitic, and several other startups are creating AI-driven algorithms to interpret medical images. In addition, Jvion provides a 'clinical success machine' that identifies patients most likely to benefit from particular treatment procedures as well as those who are most at risk.

Each of these initiatives has the potential to provide valuable decision support to clinicians in their quest to deliver optimal diagnosis and treatment to patients.

Enhancing Patient Engagement and Adherence

Patient involvement and adherence represents a major problem in the healthcare industry and solving this problem is crucial for achieving effective outcomes. The level to which patients actively participate in their own health and care is a key factor in improving patient satisfaction, better financial results, and healthcare utilization. The application of big data and AI technology is becoming more and more common in handling these challenges.

Hospitals and healthcare professionals use their clinical knowledge to create complete therapies that improve patients' acute or chronic medical conditions. On the other hand, if patients don't follow through on follow-up appointments, medication schedules, or lifestyle modifications, the efficacy of these strategies may be compromised. A major barrier to attaining favorable health outcomes is noncompliance, which occurs when patients ignore prescribed plans for treatment.

In fact, More than 300 clinical professionals and healthcare executives participated in a survey, which revealed alarming tendencies in patient involvement [26].

More than 70% of respondents reported that less than half of their patients were highly engaged in their healthcare, with 42% indicating that less than a quarter of their patients demonstrated high levels of engagement. Innovative digital solutions that improve patient engagement and treatment plan adherence are needed to address these challenges.

If greater patient involvement leads to improved health outcomes, can AI-based capabilities effectively personalize and contextualize care? Using machine learning to give personalized assistance over the full care is becoming more and more important, including utilizing messaging alerts and targeted content to prompt actions at critical moments, a field showing promise in ongoing research.

Furthermore, healthcare is increasingly focusing on designing effective 'choice architecture' to guide patient behavior proactively based on real-world evidence. Through the utilization of data from many sources such as cellphones, wearable devices, biosensors, and provider electronic health record (EHR) systems, software can customize treatment recommendations by comparing patient data to well-performing therapies for similar groups of patients. More individualized and efficient care delivery can be achieved by sharing these personalized treatments and recommendations with nurses, doctors, patients, call center representatives, and care coordinators [24].

Streamlining Administrative Processes in Healthcare

Applications for administration are essential to the management of healthcare. Even though AI may not have the same revolutionary effect in this field as it does in patient care, it still has a lot of potential to improve operational efficiency. This is especially important in the healthcare industry, since staff time can be considerably occupied by administrative responsibilities. For example, 25% of a US nurse's working hours are spent on administrative and regulatory duties[27].
Robotic Process Automation (RPA) is one technology that stands out for enhancing administrative effectiveness. Applications of RPA can be found in many areas of healthcare, such as revenue cycle management, clinical documentation, claims processing, and medical record management. Furthermore, a few healthcare institutions have started exploring the application of chatbots for remote medicine services, mental health assistance, and patient connections.

Applications that use natural language processing (NLP) are especially helpful for managing repetitive processes such as scheduling appointments and refilling prescriptions.

However, a survey of 500 US users of the top five healthcare chatbots revealed concerns about data confidentiality, addressing complex health issues, and usability challenges [28].

Machine learning is another AI technique that is applicable to the administration of claims and payments. It could be implemented to match data across several databases in a probabilistic manner. In fact, millions of claims must be verified by insurers, and this is a complex and prone to error procedure. All parties involved, including governments, providers, and health insurers, can save a significant amount of time, money, and resources by accurately detecting, evaluating, and fixing coding errors and false claims. Addressing incorrect claims that slip through the cracks represents a significant opportunity for financial optimization through data matching and claims audits[24].

Impact of AI on Healthcare Workforce

Concerns about how AI can affect workforce displacement and job automation are becoming more and more frequent. Within the next ten to twenty years, artificial intelligence (AI) may automate up to 35% of UK occupations, according to joint study from Deloitte and the Oxford Martin Institute [29].

Other studies, though, indicate that these jobs losses are likely to be constrained by various other external factors, such as automation technologies, labor market dynamics, and regulatory considerations, down to 5% or less.

AI hasn't yet resulted in the loss of any healthcare industry jobs. This lack of effect on employment has been attributed to the industry's slow adoption of AI but also to the difficulties to incorporate AI into clinical processes and electronic health record (EHR) systems. Furthermore, it is expected that healthcare roles most susceptible to automation would involve tasks related to digital information management, such as radiology and pathology, rather than direct patient care.

However, it's estimated that AI adoption will happen gradually, even in areas such pathology and radiology. While deep learning and other technological breakthroughs improve the capacity to analyze and classify medical images, a number of variables will probably postpone the eradication of jobs in these domains.

First, it's necessary to recognize that radiologists have responsibilities beyond simply interpreting images. AI systems in radiology are limited in scope, they are focused on specific tasks, such as detecting hemorrhages on brain MRIs or nodules on chest CT scans. Nowadays available deep learning models are trained on specific image identification tasks and developed by labs and startups.

However, a comprehensive analysis of medical images requires thousands of such tasks, and AI can only perform a fraction of them at present. Beyond just interpreting images, radiologists additionally collaborate with other medical professionals to diagnose and treat patients, administer therapies such as local ablative therapies, perform image-guided medical interventions like vascular stents and cancer biopsies (interventional radiology), tailoring imaging examinations to patient conditions, correlate test results and medical records with image findings, and communicate procedures and outcomes to patients.

Secondly, the integration of AI-based image analysis into clinical workflows is still in its infancy. Various deep learning algorithms and developers of imaging technologies focus on various factors such lesion probability, cancer probability, or nodule characteristics and location. It is difficult to integrate deep learning algorithms into current therapeutic practices because of these different focuses.

Thirdly, millions of images of individuals with verified diagnoses of cancer, fractures, or other diseases are needed to build deep learning algorithms for image recognition. However, the lack of a centralized repository for radiological images (labeled or not) adds difficulties to the development of AI for medical imaging[24,29].

Ethical Considerations in the Age of AI in Healthcare

The use of AI in healthcare presents a number of ethical questions and considerations in addition to technological and practical ones. In fact, in the past most healthcare choices have been made by humans. Nowadays, using AI systems to help with decision-making could create issues with privacy, transparency, accountability, and consent.

With today's AI technology, transparency is a particularly challenging problem, especially when it comes to deep learning algorithms used for image analysis and recognition. Because these algorithms are frequently cryptic and sophisticated, it can be challenging to understand or interpret their conclusions. Patients, understandably, may seek explanations for diagnoses derived from AI systems, yet deep learning algorithms, and even healthcare professionals familiar with their workings, may struggle to provide satisfactory explanations.

Furthermore, AI systems may produce errors in patient diagnosis and treatment, raising questions about its accountability. Establishing responsibility for such errors can be challenging, complicating the issues of liability and patient recourse. Additionally, patients may have the risk of obtaining medical information from AI systems in a way that is lacking the compassion and understanding that human clinicians provide. Managing emotions in these situations is really important. Furthermore, machine learning systems could unintentionally perpetuate algorithmic biases, including estimating the likelihood of a disease based on racial or gender characteristics even in situations where these characteristics are not causal [39].

Evolution of AI in Healthcare

AI will have a significant influence on how healthcare is going to be shaped in the future. Artificial Intelligence (AI) forms the foundation for precision medicine advancement through machine learning, filling in significant deficiencies in healthcare delivery. Although early efforts in diagnosis and therapy recommendation have been challenging, AI will eventually become effective in these areas. Given how quickly AI-powered imaging analysis and recognition is developing, most radiology and pathology images will eventually be analyzed by machines. Speech and text recognition technologies are already being used for patient communication and clinical note recording, and their use is only going to increase in the future.

The primary challenge for AI in healthcare lies not in its technical capabilities, but in ensuring widespread adoption in clinical practice. Achieving this requires regulatory approval, seamless integration with Electronic Health Record (EHR) systems, standardization across products, clinician education, funding support from public or private payers, and continuous updates in real-world settings. It will take time to overcome these obstacles—probably longer than it will for the technologies to mature. Therefore, within the next five years there will be limited use of AI in healthcare settings, and within ten years, there should be broad adoption.

Furthermore, it's becoming more and more clear that AI systems will support human clinicians rather than replace them. AI will improve the talents of healthcare workers rather than replace them, freeing them up to concentrate on jobs that call for human qualities like empathy, persuasion, and comprehensive patient care. In the changing healthcare scene, individuals that welcome AI collaboration are expected to prosper [24].

Google Med-Gemini-3D

One of the most recent technologies developed by Google is Med-Gemini-3D. Trained on a variety of conventional 2D medical images (including chest X-rays, CT slices, pathology slides, etc.) using de-identified medical data with free text labels, Med-Gemini-2D can execute multiple tasks such as classification, visual question answering, and text generation (Fig.5).

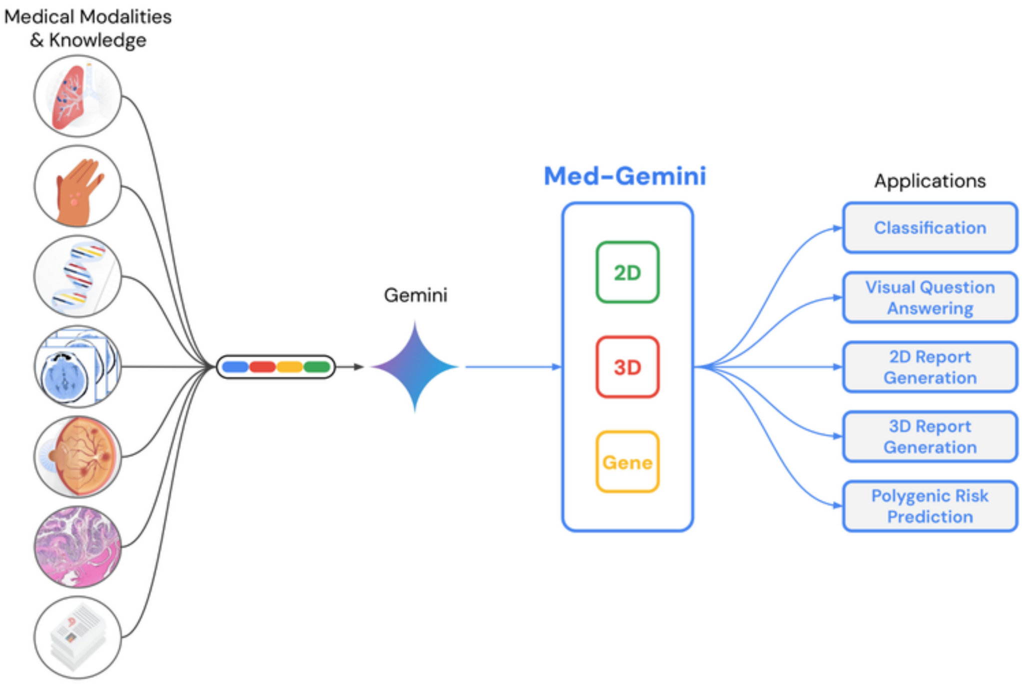


Figure 5 : Med-Gemini applications [30]

For instance, its report generation for chest X-rays has demonstrated improvements over previous state-of-the-art results by up to 12% across normal and abnormal scans from two different datasets. Med-Gemini-3D can interpret as well as generate radiology reports for 3D examinations, such head computed tomography (CT) scans. Modern medicine relies heavily on volumetric, 3D scans because they provide additional context for many diagnostic and therapeutic decisions that necessitate a knowledge of 3D anatomy. In terms of technology, 3D is far more sophisticated than 2D when it comes to multimodal clinical tasks. Furthermore, it was discovered that over half of the Med-Gemini-3D reports that were examined, produced care recommendations that were identical to those made by radiologists. However, further real-world research and validation are necessary to ensure consistent expert-level reporting quality. In addition to producing reports, Med-Gemini-2D has achieved remarkable results in visual question answering for chest X-rays and has demonstrated strong performance in a variety of classification tasks in radiology, dermatology, ophthalmology, and histopathology.

In the figure below (Fig. 6) there is an example of how this technology works, comparing its response to the radiologist one [30].

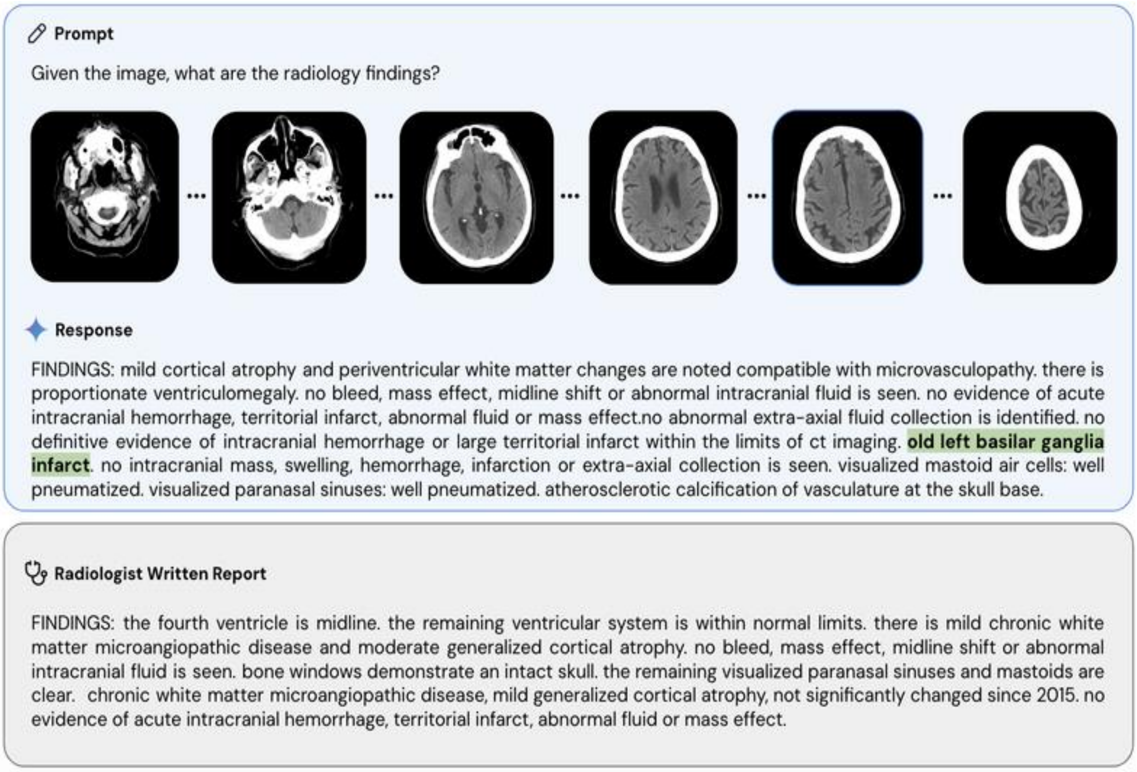


Figure 6 : AI vs Radiologist report [30]

(Fig.6) Overall, the AI's findings align closely with the radiologist's, though there are some additional observations and specific details in each that are not present in the other.

1.5 Computer Vision through Deep Learning

The objective of computer vision (CV), which has a long history spanning decades, is to allow computers to interpret visual stimuli meaningfully. Machine perception covers a spectrum of tasks, from low-level functions such as edge detection to high-level tasks like comprehending entire scenes [2].

Three main factors have contributed to the notable advances made these last ten years:

- The development of deep learning (DL), a machine learning technique that makes it easier to learn complex functions from raw data.
- Enhancements and innovations in localized computing power via GPUs;
- The availability of large labeled datasets for training these algorithms. Together, these components have given individual researchers the tools they need to continue developing this science.

Both the scientific community and advancement grew quickly.

The evolution of modern CV has coincided with the generation of vast amounts of digital data across various scientific fields. Recent medical advancements have been substantial, largely due to DL's exceptional capability to learn numerous tasks from diverse data sources.

With the use of huge datasets, CV models are able to acquire a wide range of pattern-recognition capabilities, supporting anything from medical scene interpretation to physician-level diagnosis [3].

Below, the thesis offer 2 examples of medical computer vision tasks:

- Multimodal Discriminative Model: Deep learning architectures can be designed to simultaneously learn from both image data, typically using convolutional networks, and non-image data, usually employing general deep networks. The resulting annotations can encompass disease diagnostics, prognostics, clinical predictions, and their combinations. (figure 7).

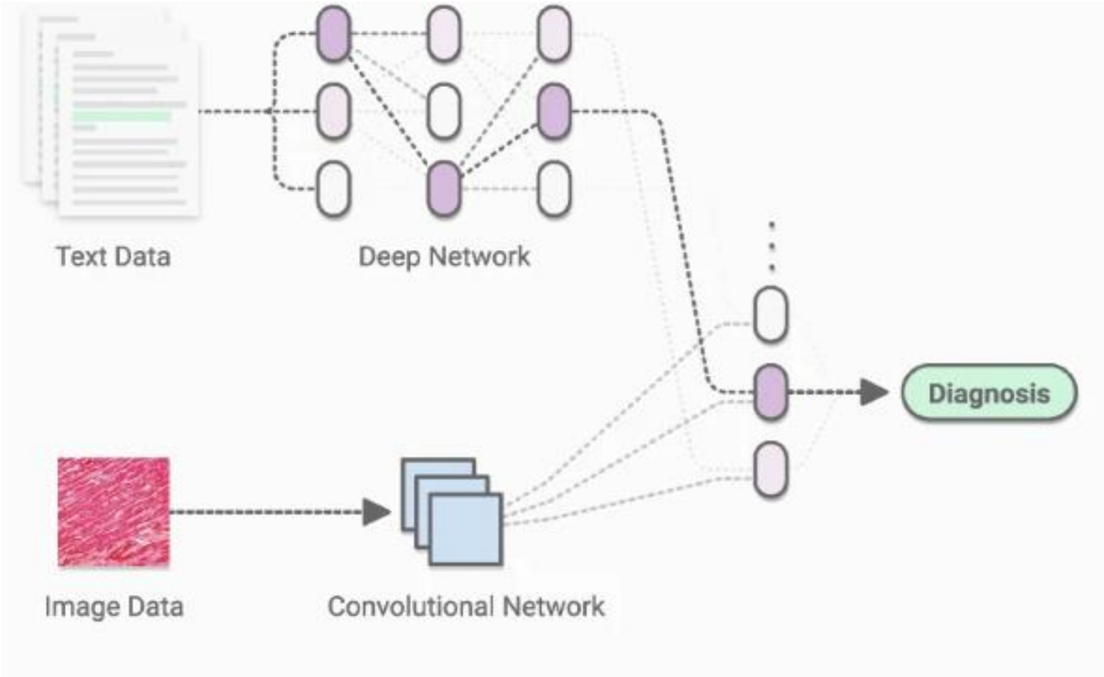


Figure 7 : Example of Multimodal discriminative model [31]

- Generative Model: Convolutional neural networks (CNNs)(fig.8) can be trained to produce images. Applications include image-to-image regression (as depicted in the image below), super-resolution image enhancement, novel image creation, and more.

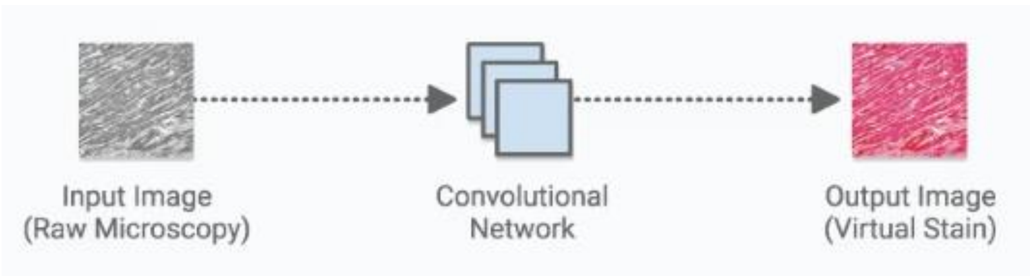


Figure 8 : Generative model: CNN can be trained to generate images [31]

Object classification, localization, and detection involve identifying an object type in an image, pinpointing the locations of objects within the image, and determining both the type and location simultaneously. Over the past ten years, the ImageNet Large-

Scale Visual Recognition Challenge (ILSVRC) has been essential in the advancement of these tasks [7].

It encouraged a large community of deep learning (DL) researchers who competed and collaborated to enhance techniques across various computer vision (CV) tasks.

Notably, during this time, classification accuracy performed at a level comparable to the performance of humans. In the medical field, fine-grained adaptations of these methods have been successfully applied to classify and detect numerous diseases (Fig. below).

In the images below we can see how CNNs were trained to classify diseases states, have been extensively tested across diseases and compared against physicians' performances.

As we can see their performance is usually similar to the one of the physicians.

In figure 9 we can see the AI performances in radiology:

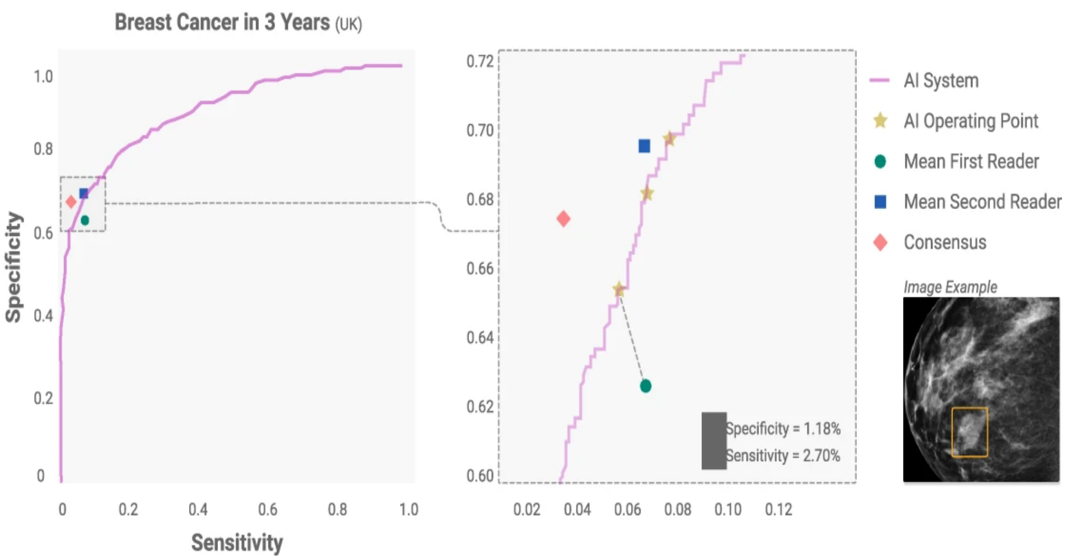


Figure 9 : Experts vs AI performance in Breast Cancer Recognition [31]

This graph shows the performance of an AI system in predicting breast cancer risk over three years compared to human radiologists.

<u>Main Graph (Left Side)</u>

Axes: The x-axis represents sensitivity, while the y-axis represents specificity.

Sensitivity measures the true positive rate (correctly identifying those with the disease), and specificity measures the true negative rate (correctly identifying those without the disease).

Curves and Points:

Pink Line: Represents the performance of the AI system across different thresholds.

Gold Stars: Show the operating points of the AI system at specific thresholds.

Green Circle: Represents the mean performance of the first reader (radiologist).

Blue Square: Represents the mean performance of the second reader (radiologist).

Red Diamond: Indicates the consensus performance when both readers agree.

Zoomed-in Graph (Right Side)

This zoomed-in view focuses on a specific region of the main graph, highlighting the differences in performance more clearly.

Axes: Same as the main graph but on a smaller scale for better detail.

Points:

Gold Stars: Still represent the operating points of the AI system at specific thresholds.

Green Circle, Blue Square, Red Diamond: Represent the same as in the main graph.

These techniques predominantly rely on supervised learning, which utilizes datasets containing both data points (such as images) and corresponding labels (like object classes). Due to the scarcity and access difficulties of medical data, transfer learning (where an algorithm is initially trained on a large, unrelated dataset, such as ImageNet, and then fine-tuned on a specific dataset, like medical data) has been crucial in the developing progress.

Also, techniques for creating synthetic data, such as data augmentation and generative adversarial networks (GANs), are being developed to reduce the expenses related to data collection and categorization. Additionally, researchers have shown that efficient medical algorithms may be produced using crowdsourcing image annotations [5].

These technologies are applied in different areas such as cardiology, pathology, dermatology, ophthalmology and medical video surgical applications.

Cardiology: the use of cardiac imaging is growing in many clinical procedures and diagnosis. Deep learning (DL) in cardiac imaging has several important clinical applications including screening and diagnosis. The most common imaging modality in cardiovascular medicine is cardiac ultrasound, also known as an echocardiogram. Because of its simple data collecting and analysis, this approach is especially well-suited for DL. It is also radiation-free and cost-effective. Echocardiography is routinely used in most acute inpatient facilities, outpatient centers, and emergency rooms.

In addition to echocardiograms, 3D imaging techniques such as CT (computed tomography) and MRI (magnetic resonance imaging) are used to understand cardiac anatomy and better characterize supply-demand mismatches in the heart. DL has a wide range of uses in cardiac imaging. Large databases of echocardiographic studies can be used to train DL models, which frequently outperform board-certified echocardiographers in tasks such view classification. Additionally, diseases such as cardiac amyloid, pulmonary arterial hypertension, and hypertrophic cardiomyopathy can be evaluated by computational DL processes. One notable advancement in the field is EchoNet, a DL model capable of recognizing cardiac structures, estimating function, and predicting systemic phenotypes that are not readily identifiable by human interpretation [5].

To address challenges related to data access, data-efficient echocardiogram algorithms have been developed, including semi-supervised generative adversarial networks (GANs) that perform well in downstream tasks like predicting left ventricular hypertrophy. A noteworthy advance is the recent release of 10,000 annotated echocardiography videos, which were previously privately held medical imaging datasets, for public use.

Alongside this data release, a video-based model named EchoNet-Dynamic was developed, capable of estimating ejection fraction (a measurement, expressed as a percentage, of how much blood the left ventricle pumps out with each contraction) and assessing cardiomyopathy. This model includes a comprehensive evaluation criterion based on results from external datasets and human experts. [6] In order to guarantee that the models can be widely applied in a variety of clinical contexts, it requires enhancing the algorithms to better handle variations in image quality across different patient groups [30,31].

Pathology: in order to detect and treat cancer, pathologists are essential. The visual examination of tissue samples under a microscope, or pathological analysis, is naturally subjective. Evaluations of prognosis and diagnosis may differ due to differences in visual perception and clinical training. Deep learning (DL) can help with critical medical tasks such as disease monitoring, segmentation of pathology, prognostication of outcomes and responses to treatment, and diagnostics.

Recent advancements have seen the introduction of sub-micron-level resolution tissue scanners.

These developments, combined with advances in computer vision (CV), have sparked significant research and commercialization in AI-driven digital histopathology.

y improving the efficiency and accuracy of routine tasks, this field has the potential to:

- Address the limitations of human visual perception and cognition.
- Develop new disease and therapy signatures from morphological structures not visible to the human eye.
- Integrate pathology with radiological, genomic, and proteomic data to improve diagnosis and prognosis.

The last research had focused on automating the routine, time-consuming tasks of localizing and quantifying morphological features.

Examples include detecting and classifying cells, nuclei, and mitoses, as well as localizing and segmenting histological primitives (the basic structural elements or components within tissue samples that pathologists examine) such as nuclei, glands, ducts, and tumors [31].

Dermatology: the key clinical tasks for DL in dermatology include lesion-specific differential diagnostics, finding concerning lesions amongst many benign lesions, and helping track lesion growth over time. A number of studies have demonstrated that CNNs are just as effective at differentiating between benign and malignant skin lesions as board-certified dermatologists. These studies have sequentially tested on big numbers of dermatologists, consistently demonstrating a sensitivity and specificity in classification that matches or even exceeds physician levels. These studies were restricted to the binary classification task of discerning benign vs malignant cutaneous lesions, classifying either melanomas from nevi or carcinomas from seborrheic keratoses.

Recently, these research has expanded to incorporate differential diagnosis for a broad spectrum of skin conditions, including rashes and genetic disorders as well as non-neoplastic lesions, and incorporating non-visual metadata (such patient demographics) as inputs for the classifier.

These advancements have been driven by open-access image repositories and AI competitions that encourage teams to compete on established benchmarks.

Integrating these algorithms into clinical workflows could enhance their utility in other essential tasks, such as large-scale detection of malignancies in patients with multiple lesions and tracking lesions across images to capture temporal changes like growth and color variations. With new research concentrating on teaching CNNs to simultaneously detect and track lesions, this area is still relatively unexplored [8].

In the Figure 10 we can see AI performances in Dermatology:

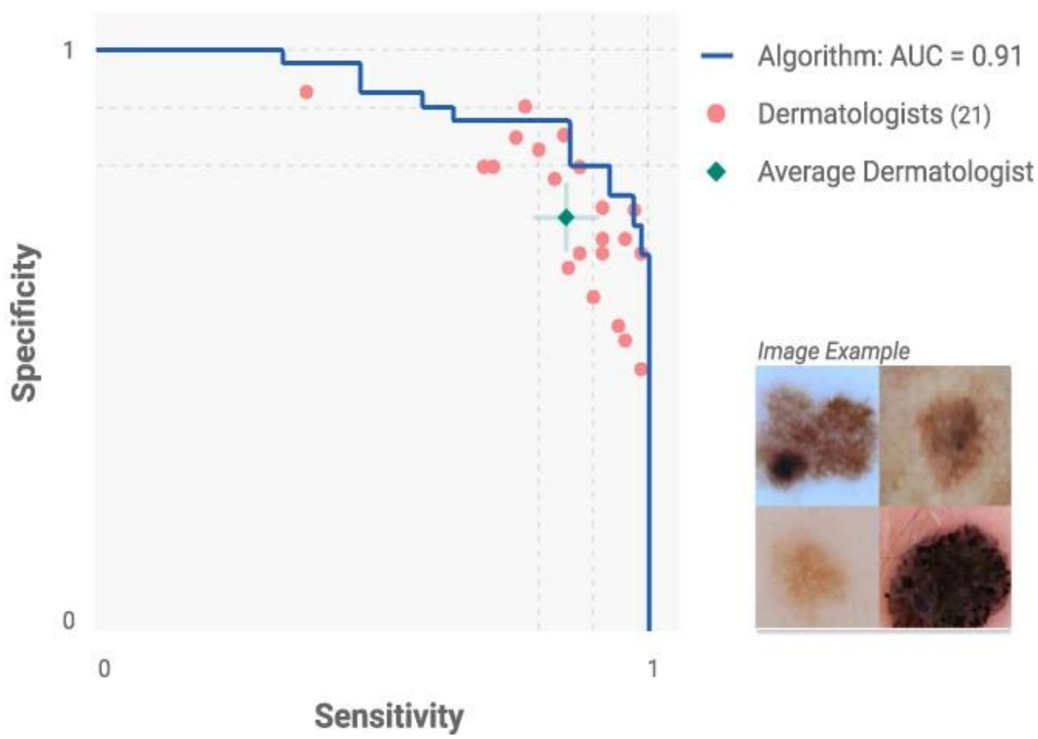


Figure 10 : Experts vs AI performance in Dermatology [31]

Ophthalmology: In recent years, ophthalmology has seen a substantial increase in AI research, with numerous studies showcasing diagnostic and analytical capabilities that surpass current human performance. For the diagnosis and treatment of patients, this field primarily uses optical coherence tomography (OCT) and fundus imaging (images of the retina). Convolutional neural networks, or CNNs, have proven to be highly accurate in diagnosing a wide range of illnesses. There has been a great deal of research done on diabetic retinopathy, a disorder where blood vessels in the eyes leak and can cause blindness. Analogously, these networks are also capable of diagnosing or forecasting the progression of age-related macular degeneration (an eye disease that can blur your central vision), glaucoma, visual field loss, juvenile blindness, glaucoma, , and more.

Many non-human-interpretable characteristics found in the eyes contain important medical information that CNNs can identify. It has been impressively demonstrated that CNNs are capable of classifying a wide range of cardiovascular and diabetes risk variables, such as age, gender, smoking status, hemoglobin-A1c levels, body mass index, systolic and diastolic blood pressure, from fundus photographs. Furthermore, using fundus photos, CNNs can identify indicators of anemia and chronic kidney disease. By enabling eye exams to screen for both ocular and non-ocular disorders, this creates interesting opportunities for future AI research to predict non-ocular information from eye scans, potentially changing care[9,30].

Medical Video in Surgery applications: computer vision (CV) holds significant potential in procedural fields like surgery and endoscopy. Enhancing surgeon performance through training, skills examinations, and real-time contextual awareness are some of the key clinical uses for deep learning (DL). These objectives have been the subject of early research, mainly in video-based robotic and laparoscopic surgery (a surgical technique in which short, narrow tubes are inserted into the abdomen through small incisions). When evaluating a surgeon's performance, some study looks at tool movement and other indicators.

These evaluations and their ratings are compared to established standards like the Global Operative Assessment of Laparoscopic Skills (GOALS) criteria for laparoscopic surgery.

Another research avenue employs CV to recognize distinct phases of surgery during operations, aiming to develop context-aware computer assistance systems. [30,31].

Chapter II

CONVOLUTIONAL NEURAL NETWORK (CNN)

2.1 CNNs

Chapter II focuses its attention on one particular technology: Convolutional Neural Networks. This Chapter starts with a general introduction of CNNs in the section 2.1, moving to a more in depth examination in section 2.2 and 2.3 with the CNN's architecture and training, and finally concluding with a comparison across the different CNNs architectures.

As a subclass of artificial neural networks, Convolutional Neural Networks (CNNs) are becoming more and more important in a wide range of computer vision applications. CNNs use a variety of fundamental components, including convolutional layers, pooling layers, and fully connected layers, to automatically and adaptively learn spatial hierarchies of data. This chapter offers an overview of the fundamental concepts behind CNNs and discusses how they are applied to various tasks.

One type of deep learning model that is designed for processing data having a grid-like architecture, such images, is the Convolutional Neural Network (CNN). It is inspired by the structure of the animal visual cortex and is engineered to automatically and adaptively learn spatial hierarchies of features, ranging from low-level to high-level patterns. The following primary kinds of layers constitute CNNs: fully connected, pooling, and convolutional layers.

The convolutional and pooling layers are the layers responsible for feature extraction, while the fully connected layers map the extracted features to the final output, such as a classification. Convolution, which is a specialized linear operation, is one of the mathematical operations that compose the convolutional layer, which is at the core of a CNN. In this operation, an array of numbers, or pixel values recorded in a two-dimensional (2D) grid, is used to represent digital images (Fig.11)[32]. Every position in the image is subjected to a small grid of parameters called a kernel or filter, which serves as an optimizable feature extractor. Because features can be found anywhere in the image, CNNs are especially effective for processing images.

Next, the retrieved features get increasingly complex as one layer's output becomes the next layer's input. During the training phase, parameters like kernels are optimized to reduce the difference between the outputs of the CNN and the labels that correspond to the ground truth. This optimization is achieved through algorithms like backpropagation and gradient descent, among others.

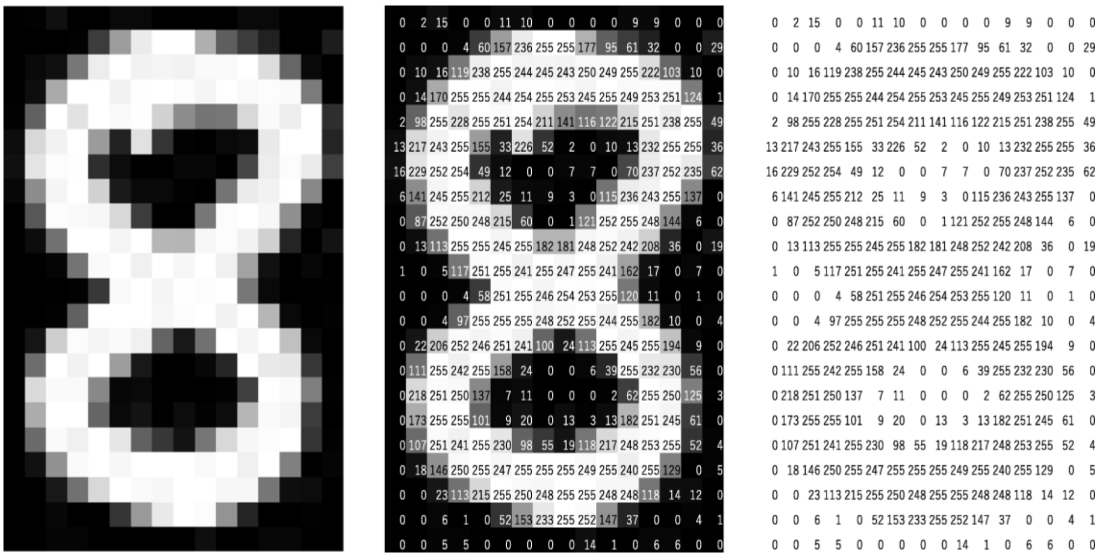


Figure (11) : A computer sees an image as an array of numbers [32]

A computer interprets an image as a matrix of numbers. The matrix on the right displays values ranging from 0 to 255, each representing the pixel brightness in the left image. The middle image illustrates both the numerical matrix and the corresponding visual representation.

2.2 Convolutional Neural Network Layers and Architecture

The CNN architecture is composed of different building blocks, such as convolution layers, pooling layers, and fully connected layers. The standard architecture involves repeating a stack of several convolution layers and a pooling layer, followed by one or more fully connected layers.

Forward propagation is the method by which input data is converted into output across these levels. Even if this chapter focuses on convolution and pooling processes for two-dimensional (2D) CNN, similar techniques can also be used for three-dimensional (3D) CNN.

Convolution Layer

Convolution is a specialized type of linear operation used for feature extraction, where a small array of numbers, called a kernel, is applied across the input, which is an array of numbers, referred to as a tensor. A feature map, which is the output value at the corresponding position in the output tensor, is produced by computing an element-wise product between each element of the kernel and the input tensor at each location in the tensor. This final product is then summed.

In order to create an arbitrary number of feature maps, each representing a distinct aspect of the input tensors, this process is repeated with numerous kernels. Here, various kernels function as distinct feature extractors (Fig. below). The size and number of kernels are two essential hyperparameters that determine the convolution operation. Although 3x3 is the standard size, it can alternatively be 5x5 or 7x7. The depth of the output feature maps is determined by the arbitrary number of kernels. Figure 12 shows how CNN works [32].

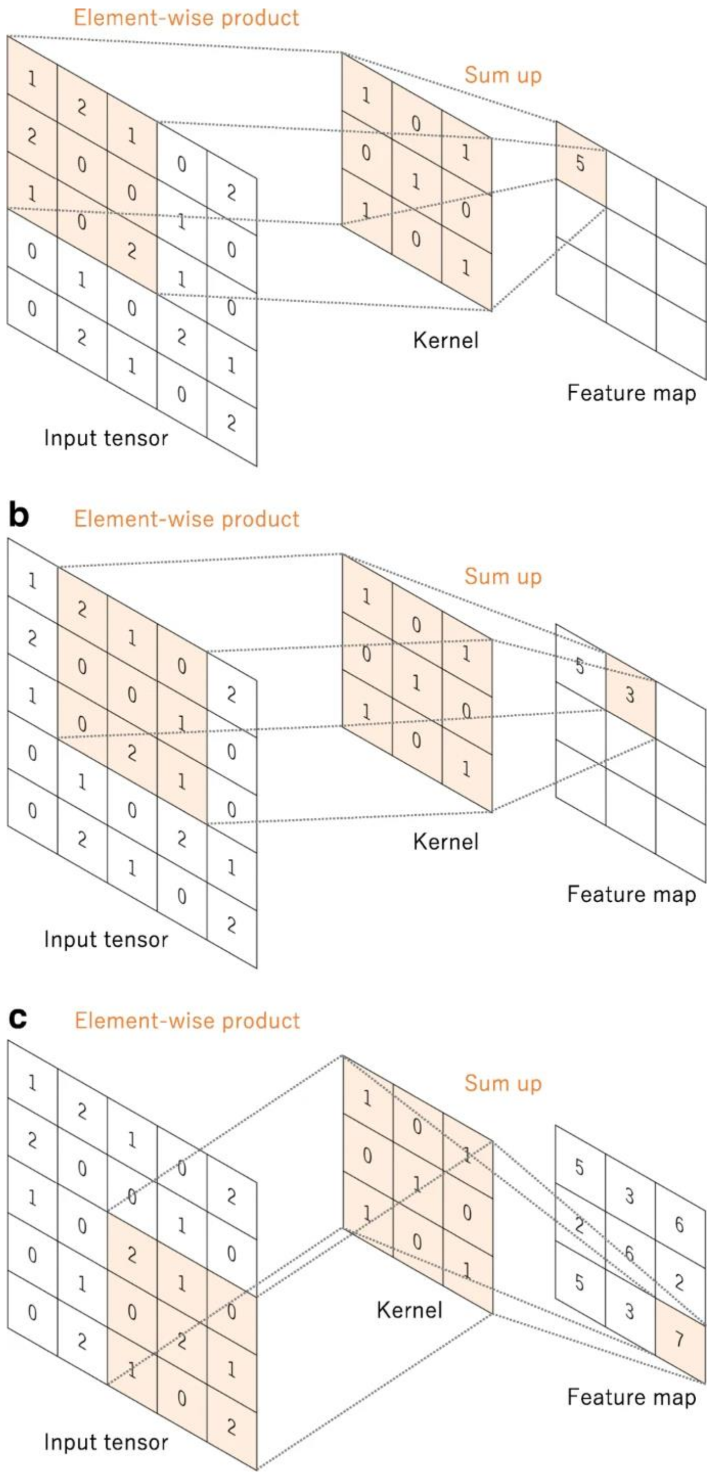


Figure 12 : How CNN works [32]

Image explanation: this is an example of convolution operation with a kernel size of 3 × 3, no padding, and a stride of 1. The image consists of three panels (a, b, c), each depicting different steps in the convolution process.

- **Input Tensor**: The grid on the left in each panel represents the input tensor, which could be an image. Each cell in the grid corresponds to a pixel value in the image.
- **Kernel:** The smaller grid overlaying the input tensor is the kernel (or filter). It is a small matrix of weights that slides over the input tensor to perform the convolution operation.
- **Feature Map:** The grid on the right in each panel is the feature map, which contains the results of the convolution operation.

Steps in the Convolution:

- The kernel is placed over a specific location on our input tensor.
- Each element of the kernel is then multiplied by the corresponding element of the input tensor (element-wise product). The term "element-wise product" refers to the multiplication of corresponding elements from two matrices (or arrays). In the context of a convolutional neural network (CNN), this operation is performed between a kernel (the filter) and a specific region of the input tensor (such as an image).
- The results of the element-wise products are summed up to produce a single value.
- Finally, this single value is placed in the corresponding location of the feature map.

The convolution operation described above does not allow the center of each kernel to overlap the outermost elements of the input tensor, resulting in a reduction of the height and width of the output feature map compared to the input tensor. One way of dealing with this is to utilize padding, usually in the form of zero padding. The center of the kernel can be placed over the outermost elements by adding rows and columns of zeros on each side of the input tensor, keeping the same in-plane dimensions throughout the convolution process (Fig. 13) [32].

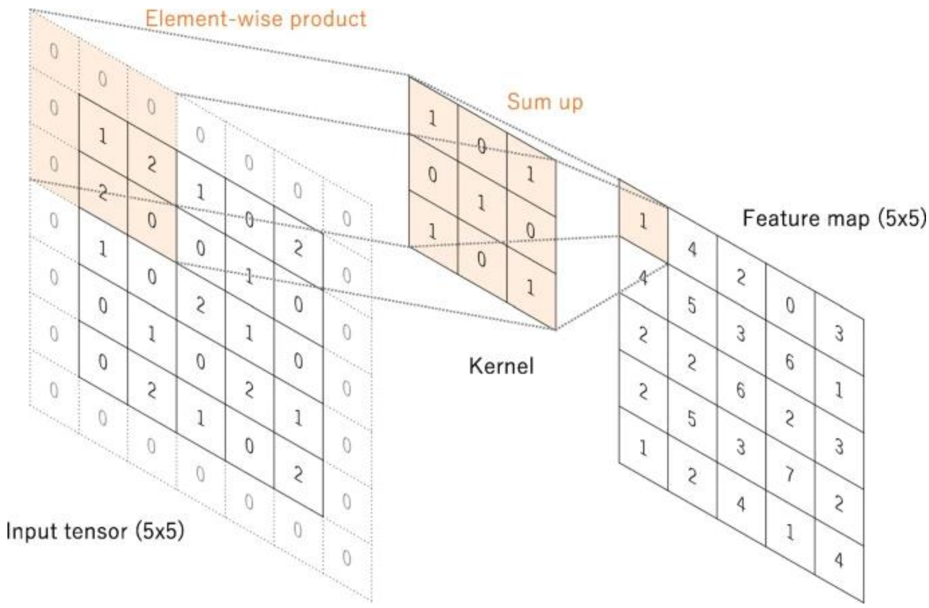


Figure 13 : Padding example [32]

In a convolutional neural network (CNN), the "stride" refers to how far the CNN moves the filter (also called a kernel) across the input image. This motion takes place during the convolution process, which functions similarly to moving a window across an image in order to identify patterns. The filter typically moves one pixel at a time (stride = 1), but occasionally we move it in many directions to reduce the output size and streamline the network. Another way to shrink the output size is by using pooling, where we group nearby pixels and keep only the most important one. Convolution has also an interesting feature called "weight sharing." This means that the same filter is used all over the image. It helps the network with different aspects.

Starting with staying invariant to translation, it can recognize patterns no matter where they are in the image.

Next it acquires knowledge of patterns at various levels of detail in fact convolution and downsampling (similar to pooling) are combined to allow the network to progressively examine more of the image.

Finally, efficiency, in fact the network needs to learn less parameters by sharing filters. In order to identify the most effective filters for our task and data, we adjust them while training a CNN. These filters are the only things the network learns automatically. Before training starts, we select various options including padding, stride, filter size, and number of filters.

Non Linear Activation Functions

After performing linear operations like convolution, the results are fed through a nonlinear activation function. Since they replicated the behavior of biological neurons, smooth nonlinear functions like the sigmoid or hyperbolic tangent (tanh) were once popular. However, the rectified linear unit (ReLU) is currently the most commonly utilized nonlinear activation function. This is a basic function that computes f(x) = max (0, x).

Figure 14 displays three different activation functions commonly used in neural networks: ReLu, Sigmoid and Tanh.

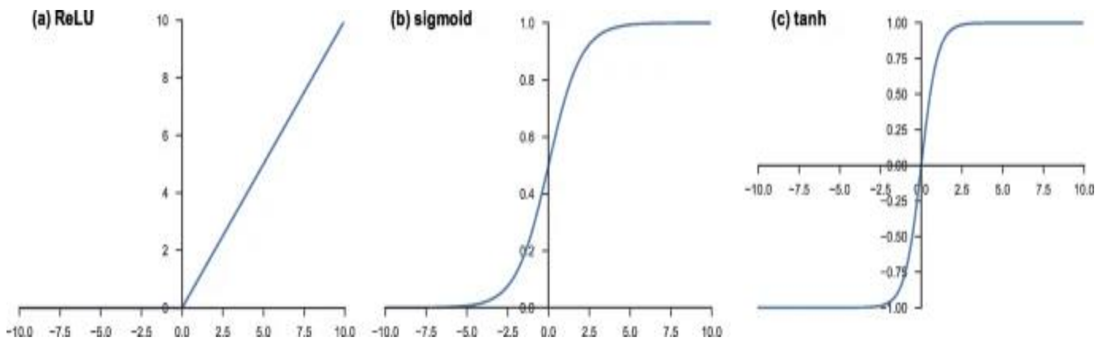


Figure 14 : Common Activation Functions applied to Neural Networks [32]

a) The ReLU function outputs zero for any input less than zero and outputs the input value for any input greater than or equal to zero. This function introduces non-linearity to the model and it activates only a few neurons at a time, which can lead to sparse representations. Furthermore, it helps mitigate the vanishing gradient problem which occurs when the gradients of the loss function with respect to the network parameters (weights) become very small as they are propagated back through the network during training. One disadvantage of this function is that it could possibly lead to dead neurons, which are neurons that never activate during the training.

   Equation: $f(x) = \max(0,x)$

b) Moving to the sigmoid function, it outputs a value between 0 and 1 and it has an S-shaped curve. The fact that it outputs values between 0 and 1 makes it useful for binary classification problems.

This function is good for models where a probabilistic interpretation is needed but can cause the vanishing gradient problem which slows down the training process for deep networks.

c) Concluding with Than function, outputs values between -1 and 1 and also has a S-shaped curve, symmetric around the origin. It outputs values centered around zero which can lead to faster convergence in some cases. Similarly to the sigmoid function, also the Than function could suffer from the vanishing gradient problem [32,33].

Pooling Layer

A pooling layer performs downsampling, which reduces the size of the feature maps. This helps the network handle small shifts and distortions in the input and lowers the number of parameters that need to be learned in the following layers. Notably, there are no parameters for pooling layers to learn. However, in pooling operations, the filter size, stride, and padding are settings (hyperparameters) that we define before training the network, similar to those in convolution layers.

Max pooling: Max pooling (Fig. 15) [32] is the most used kind of pooling procedure. This method takes small regions (patches) from the input feature maps, keeping the highest value from each patch, and discarding the rest. Usually, max pooling reduces the width and height of the feature maps by half by using a 2x2 filter with a stride of 2. It does not, however, alter the feature maps' depth [32,33].

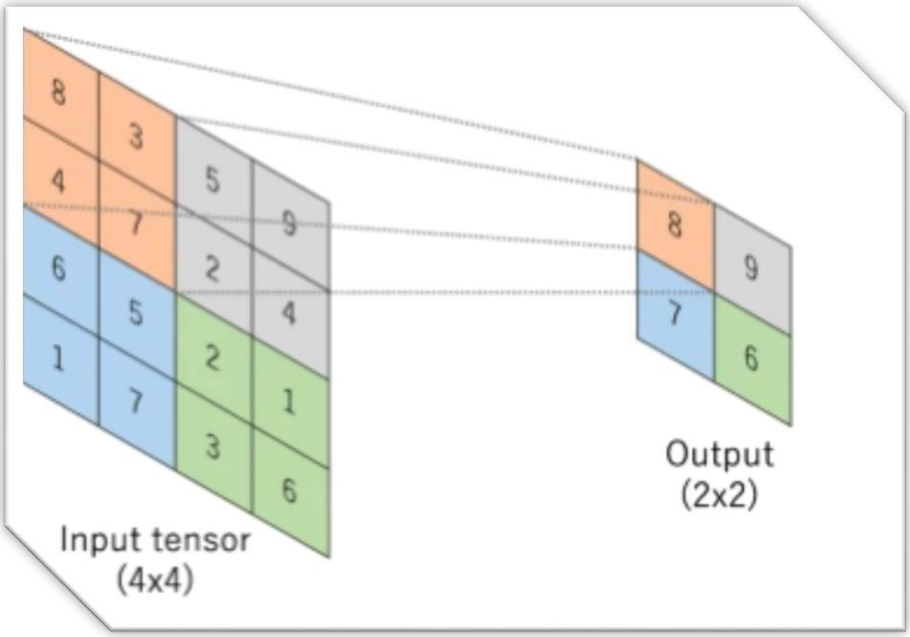


Figure 15 : Representation of Max Pooling [32]

Global Average Pooling: Global average pooling is an additional noteworthy pooling operation. By averaging every element in each feature map, this method reduces a feature map with dimensions of height × width to a 1 × 1 array, therefore performing an extreme type of downsampling. The feature maps' depth doesn't change. This technique is usually applied once, just before the fully connected layers. Global average pooling has two benefits: it lowers the quantity of learnable parameters and enables the CNN to process inputs of various sizes [32].

Fully Connected Layer

The output feature maps from the last convolutional or pooling layer are usually flattened into a one-dimensional (1D) array of numbers, also known as a vector.

After that, this vector is passed into one or more completely connected layers, often referred to as dense layers, in which a learnable weight connects each input to every output. The final outputs of the network, such as class probabilities in classification tasks, are generated by running these features through the fully connected layers after the convolutional layers extract them and the pooling layers downsample them. Usually, the number of output nodes in the last fully connected layer equals the number of classes. Finally, a nonlinear function such as ReLU follows each fully connected layer. [33]

2.3 Training a Network

Training a neural network involves finding the best kernels (our filters) in the convolutional layers and the optimal weights in the fully connected layers to minimize the difference between the network's predictions and the actual labels in the training dataset. For such types of training, the backpropagation algorithm is frequently used. It relies heavily on the loss function and the gradient descent optimization algorithm. It works in 2 phases:

- The first phase is called Forward Propagation: in this phase, the model makes its predictions on the training data using the current kernels and weights. The performance of this model is measured by a loss function, which calculates how far off the predictions are from the actual labels. A loss function, also referred to as a cost function, measures the compatibility between output predictions of the network through forward propagation and given ground truth labels.

- The second phase is called Backpropagation and Gradient Descent: in this phase, based on the loss value, the backpropagation algorithm adjusts the kernels and weights to reduce the loss. This adjustment process uses the gradient descent optimization algorithm to update the parameters step-by-step, aiming to minimize the loss over time[32].

Importance of Data and Ground Truth Labels

In deep learning and other machine learning techniques, data and ground truth labels are essential components. The expression “Garbage in, garbage out” highlights how crucial having high-quality data is to model testing and training. A deep learning project’s success depends on obtaining correct data and reliable ground truth labels, even if doing so can be costly and time-consuming. Although there are a number of publicly accessible medical picture datasets, it’s crucial to confirm the accuracy of their ground truth labels.

Data is typically divided into three sets: training, validation, and test sets (Fig. 16), though variations like cross-validation also exist. When dividing our dataset for a machine learning task, a common approach is to use 70/80% of the dataset for the training set, 10/15% for the validation set and the remaining 10/15% for the Test set.

The network is trained using the training set, where learnable parameters are updated through backpropagation and loss values are calculated during forward propagation. The validation set is used to evaluate the model during training, fine-tune hyperparameters, and assist in model selection. Ideally, the test set is used only once at the end of the project to assess the performance of the final model, which has been fine-tuned and selected using the training and validation sets.

Because training a model involves adjusting hyperparameters and selecting the optimal model based on performance, it is essential to separate validation and test sets. During this process, information about the validation set can unintentionally influence the model, leading to overfitting, even if the model is not directly trained on it. For instance, fine-tuning hyperparameters on the validation set ensures that the model performs well on that specific data. However, a totally different test set is necessary to assess the model’s generalizability to unobserved data. In fact, a separate test set is required since we need to evaluate the model’s performance on data that it has never seen before [32,33].

It's also important to remember that machine learning and medicine have distinct definitions for the word "validation". In medicine, it frequently refers to confirming the accuracy of a prediction model, which is similar to what is known as a “test” in machine learning. In machine learning, it usually relates to optimizing and choosing models during training. In order to prevent misunderstandings, “development set” is occasionally used in place of “validation set.”

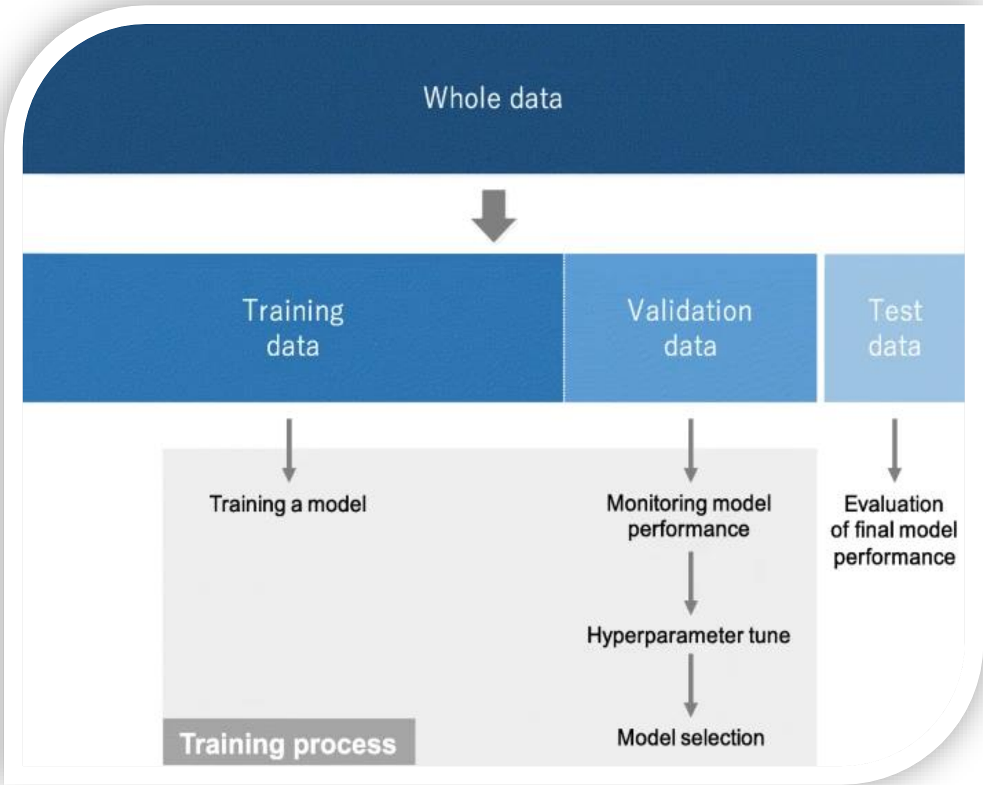


Figure 16: Data is typically divided into 3 sets [32]

<u>Overfitting:</u> previously we mentioned overfitting, but what is it? (Fig.17)

When a model is overfitted to the training set, it becomes extremely sensitive to the noise or small variations in the data and loses sight of the underlying patterns. Consequently, when presented with new, unseen data, the overfitted model struggles to make accurate predictions because it has essentially "memorized" the training data rather than learning the true underlying relationships. In machine learning tasks, where the ultimate goal is to construct models that can efficiently generalize to unseen data beyond the training set, this lack of generalizability presents a significant challenge. There are various methods to address overfitting, some examples could be:

- Acquiring more training data, however this isn't always possible, particularly in medical imaging.
- Regularization techniques, like dropout, which randomly deactivate some neurons during training to prevent reliance on specific weights.

- In order to support simpler models, weight decay, also known as L2 regularization, penalizes excessive weights.
- Moving to batch normalization, it normalizes input values to stabilize training and improve gradient flow.
- Finally, data augmentation modifies training data with random transformations to expose the model to diverse inputs.

Even with these techniques, overfitting to the validation set can still happen when selecting a model and fine-tuning hyperparameters. To guarantee the final model's generalizability, it is crucial to evaluate its performance on a different test set and, if possible, on external validation datasets.

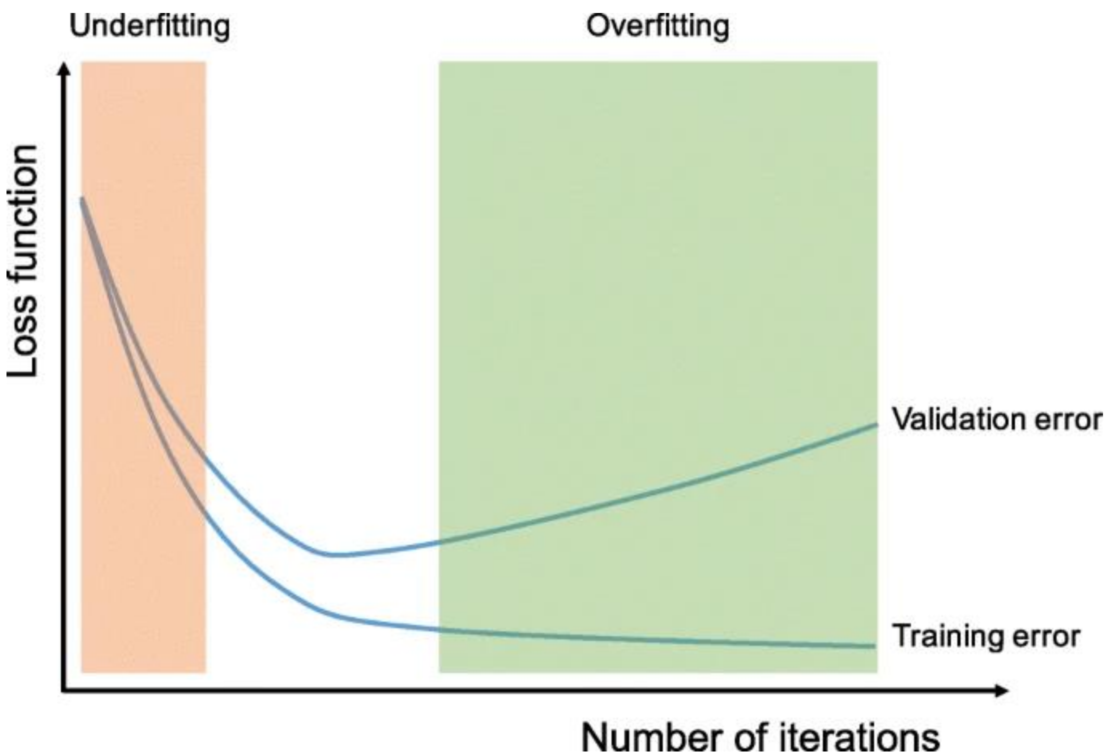


Figure 17 : Loss function, Underfitting & Overfitting [32]

The image illustrates the concepts of underfitting and overfitting.

The X axis represents the number of iterations (or epochs) during the training process.

The Y axis represents the loss function, which indicates us how well the model is performing. Lower values of the loss function indicate better performance.

Training error curve: indicates the loss on the training data

Validation error curve: indicates the loss on the validation data

Underfitting: it occurs when the model is too simple to capture the underlying patterns in the data. (Fig 17)

In the underfitting region (marked in orange on the left side of the graph), both training and validation errors are high. This indicates that the model is not performing well on both the training and validation datasets. To conclude, the model has not learned enough from the data, leading to poor generalization.

Overfitting: This happens when the model becomes too complex and starts to learn the noise in the training data as if it were a true pattern.

In the overfitting region (marked in green on the right side of the graph), the training error continues to decrease, but the validation error starts to increase, indicating that the model is performing very well on the training data but poorly on the validation data.

In fact, the model has learned the training data too well, including its noise, which lowers its performance on new, unseen data.

Optimal Point: The optimal point lies somewhere in between underfitting and overfitting, where the validation error is at its minimum; at this point, the model has learned enough to generalize well to new data without being overly complex [32,33].

2.4 Comparison of Different CNN Architectures

Based on convolutional neural network principles, the most popular CNN designs are LeNet, AlexNet, VGGNet, GoogLeNet, ResNet, and ZFNet[37]. The most popular options for picture segmentation tasks are U-Net, SegNet, and ResNet18. LeNet-5, a popular variant with five CNN layers, achieved an amazing 99.2% accuracy in single character recognition.

Next, AlexNet, developed by Alex Krizhevsky, gained significant popularity as the first widely adopted convolutional network. It consists of five convolutional layers followed by three fully connected layers, totaling eight primary layers. To enhance speed and accuracy, AlexNet incorporates ReLU activation functions.

The concept of residual neural networks, where layers are reconfigured to learn residual functions, was first presented by Microsoft Research's ResNet. This innovation allows for deeper networks with improved accuracy and easier optimization.

Subsequently, GoogLeNet, created by Szegedy et al., is notably deeper than AlexNet, with 22 layers. Despite its depth, GoogLeNet achieves efficiency with only 4 million parameters compared to AlexNet's 60 million.

In medicine, CNN architectures have been extensively explored, with various post-processing methods proposed to refine prediction outcomes. [33] In Figure 18 there's a comparation across all these architectures.

Architectures	Layers	Advantages	Disadvantages
LeNet-5	7 layers	Ability to process higher resolution images need larger firmer layers.	Overfitting in some cases and no built-in mechanism to avoid this
AlexNet	8 layers 60 M parameters	A very rapid downsampling of the intermediate representations through convolutions and max-pooling layers.	The use of large convolution filters (5 × 5) is not encouraged shortly after that, Is not deep enough rather than another techniques.
ZFNet	8 layers	Improved image classification rate error in compared with Alexnet, winner of ILSVRC2012	Feature maps are not divided across two different GPU, Thus connections between layers are dense.
GoogleNet	22 layers 4–5 M parameters	Winner of ILSVRC2014, Decreased the number of parameters from 60 million (AlexNet) to 4 million so network can have a large width and depth.	Consists of a hierarchy of complex inception modules/blocks that consist of operations over different scales in each of the modules.
VGGNet	Between 11 to 19 layers the best one is 16 layers 138 M parameters	At present it is the most prefer election for extracting features from images.	Consists of 138 million parameters, which can be a bit challenging to handle.
ResNet	152 layers	Network learns difference to an identity mapping (residual), Faster convergence if identity is closer to the optimum.	Lower complexity than VGGNet, Overfitting would increase test but decrease training error.

Figure 18 : Comparison of the different architecture of CNN [33]

Chapter III

General Overview on Tumors

3.1 Understanding Cancer

In order to fully understand the next chapter, which covers the topic of CNN applied to brain tumor recognition, the reader needs to have some knowledge about cancer. For this purpose, this chapter introduces a slight overview on cancer starting with its understanding in section 3.1 and concluding with its detection in section 3.2.

Cancer includes a diverse array of over 100 diseases characterized by the unregulated proliferation of cells within the body. The fundamental processes that lead to cancer are quite similar in all forms of the disease, despite the fact that cancer can grow in almost any tissue in the body and that every variety of cancer has its own distinct characteristics.

Cancer initiates when a cell liberates itself from the usual checks on cell division and starts proliferating independently (Figure 18)[34]. There are four mutations that lead to the emergence of this tumor. However, the precise number of mutations implicated in other tumor types can vary. Although the precise number of changes required for a normal cell to become completely malignant is unknown, it is most likely less than ten.

This original cell, along with its descendants resulting from division, also exhibits abnormal proliferation.

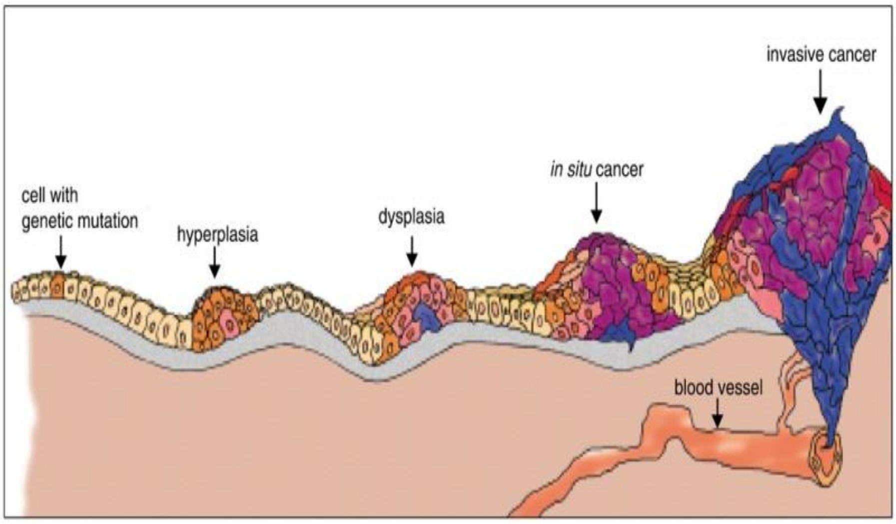


Figure 18 : Stages of tumor Development [34]

These abnormal cells aggregate to form a tumor or mass, which can either remain confined within its tissue of origin, known as in “situ cancer”, or extend into nearby tissues, termed “invasive cancer”.

A malignant tumor denotes an invasive growth, with cells capable of spreading through the bloodstream or lymphatic system to establish new tumors, a process known as metastasis. The growth of tumors becomes life-threatening when it disrupts vital tissues and organs necessary for sustaining life.

Scientists have been intrigued by the mechanism that causes cells to change into malignant ones for a long time. They were unable to come up with a cogent response to this subject thirty years ago. Although the cause of cancer was known to be uncontrolled cell multiplication within the body, the precise mechanisms were still unknown. Scientists recognized that various factors like chemicals, radiation, and viruses could trigger this transformation, yet the precise mechanisms remained elusive. However, in the last thirty years, our understanding of cancer has experienced an important evolution. Advances in molecular biology tools have played a major role in this evolution, enabling researchers to examine and clarify the complex characteristics of individual cells in ways that were previously unimaginable. Today, we understand cancer as a disease rooted in the molecular and genetic domains, with many of the involved molecules and genes identified.

Indeed, our expanding knowledge of these genetic components is paving the way for the development of innovative strategies to prevent, delay, and even reverse the cellular changes that precipitate cancer [34].

3.2 Tumor Detection

Examining brain tumor images provides crucial information about the patient and aids in clinical diagnosis. Image processing techniques (IPTs) that are most commonly utilized are segmentation and classification. These methods are crucial for feature extraction, analysis, and interpretation in medical image (MI) analysis. Both methods are widely used to identify the location of tumors, classify blood cells, and categorize tumor cells. A brain tumor is characterized by the abnormal growth of cells within the brain. A brain tumor is a cluster of these abnormal cells in the brain that can damage brain cells and induce inflammation. Brain tumors have been observed in a variety of forms. Tumors are generally classified as either malignant (cancerous) or benign (not cancerous). Malignant tumors are further divided into two categories: primary tumors, which start in the brain, and secondary tumors, which start in the body and spread to the brain through other organs like the breast or lungs. The diagnostic process for brain tumors involves three main steps: tumor detection, segmentation, and classification as the Figure 19 below displays [11].

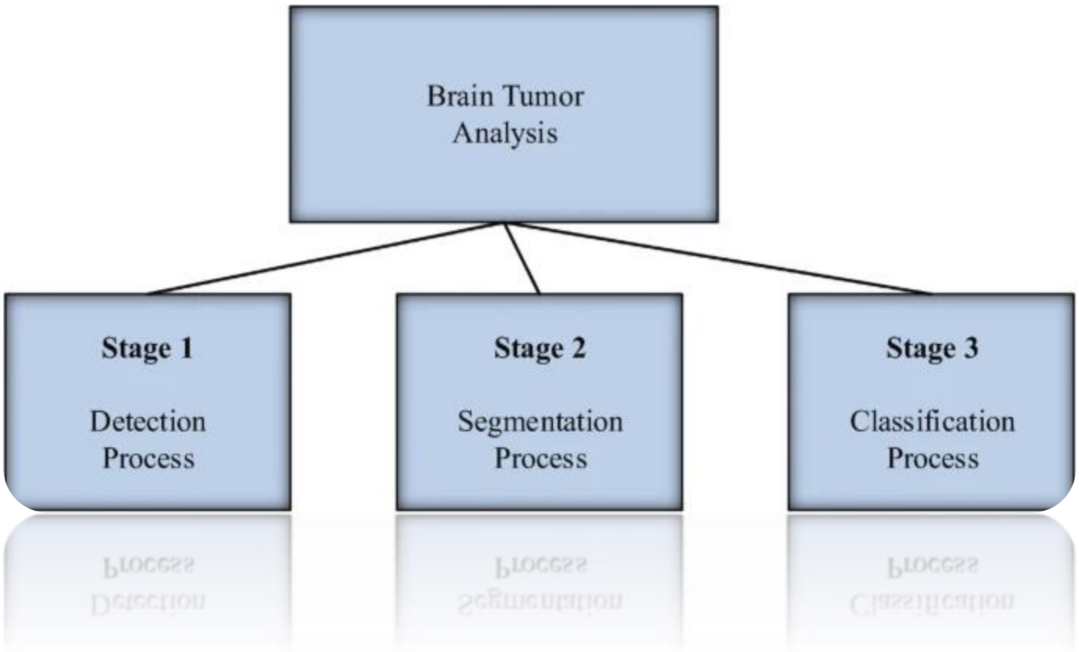


Figure 19: Three stages for Brain Tumor Diagnosis [35]

MRI (Magnetic Resonance Images) tumor images from the database are classified using detection procedures, which are known to be labor-intensive. To recognize and separate the different tumor tissues shown in the MRI images, segmentation techniques are applied. Ultimately, the tumor's malignant or benign nature is determined by the classification process. The World Health Organization (Kleihues et al., 2013) has divided malignant brain tumors into four classes, which are depicted in Figure 20. The first two grades are considered semi-malignant tumors with minimal aggression, while the remaining grades are malignant tumors that can lead to human death [12].

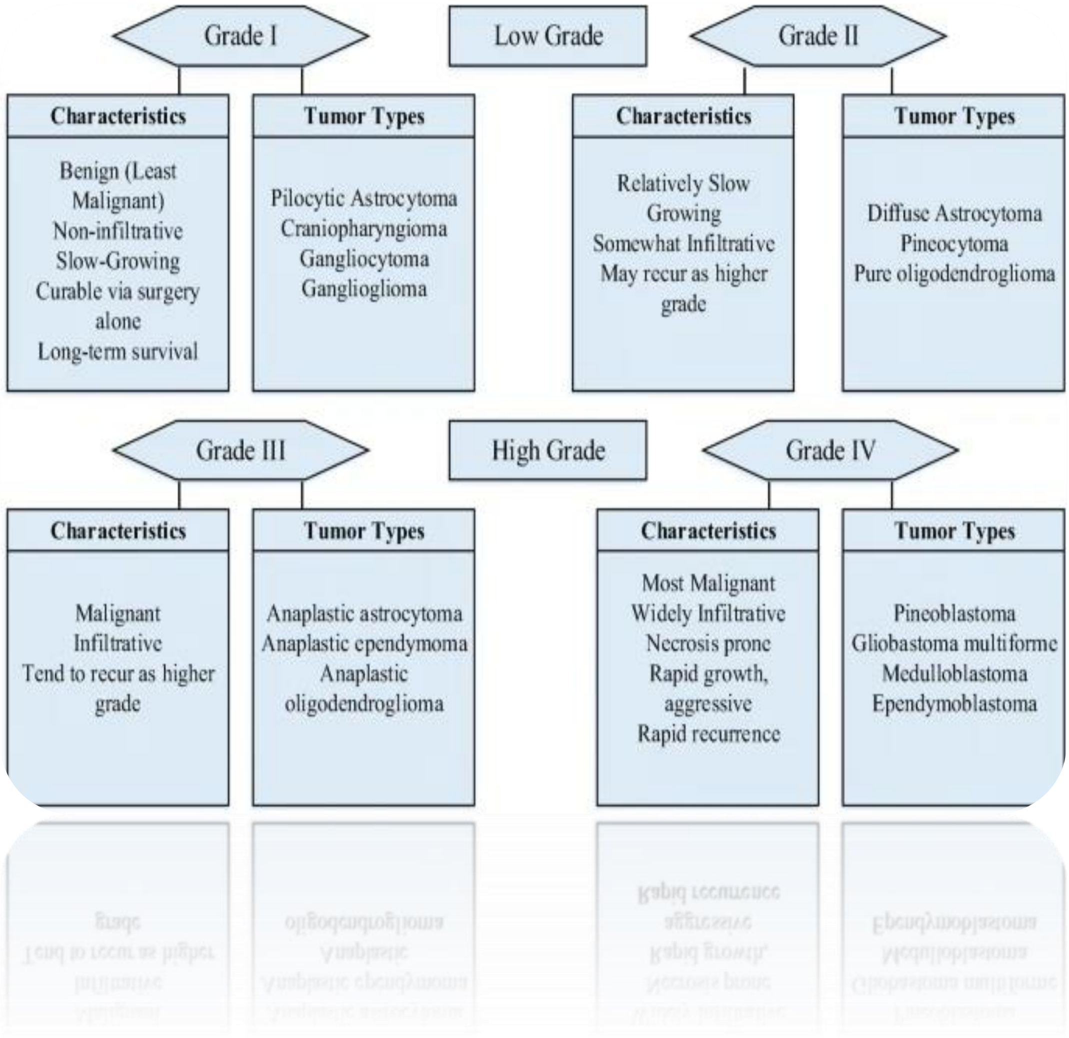


Figure 20 : Brain Tumor Grades [35]

Early tumor detection can save lives and enhance treatment options (Darling 2017). Different brain imaging techniques assist specialists detect brain tumors by providing essential details about their size, location, type, and shape. These techniques include CT (Computed Tomography), SPECT (Single-Photon Emission CT), MRI (Magnetic Resonance Imaging), MRS (Magnetic Resonance Spectroscopy), and PET (Positron Emission Tomography). Clinicians prefer MRI above the other methods. MRI offers a detailed view of the brain, vascular anatomy, and spinal cord, and can visualize the brain structure in all planes (axial, sagittal, and coronal). Since MRI is so widely available and has excellent soft tissue contrast, it is considered as the standard method. It is mostly utilized for medical problems related to neurology and neurosurgery. MRI employs radio frequency signals within a powerful magnetic field. MRI has the following benefits over others techniques:

- High capacity for contrast

- Minimal exposure to radiation
- The capacity to identify vascular anomalies and blood flow problems
- Accurate identification of disorders affecting the nervous system
- Beam-hardening artifacts are absent. To be more clear, beam-hardening artifacts are distortions that occur in imaging techniques due to the way X-ray beams interact with dense tissues. As X-rays pass through these dense structures, they can become "hardened," meaning that lower-energy X-rays are absorbed more than higher-energy ones, resulting in streaks, shadows, or false areas of high density on the images.

In the Figure 21, we can see some tumor representation and the basic types of MRI.

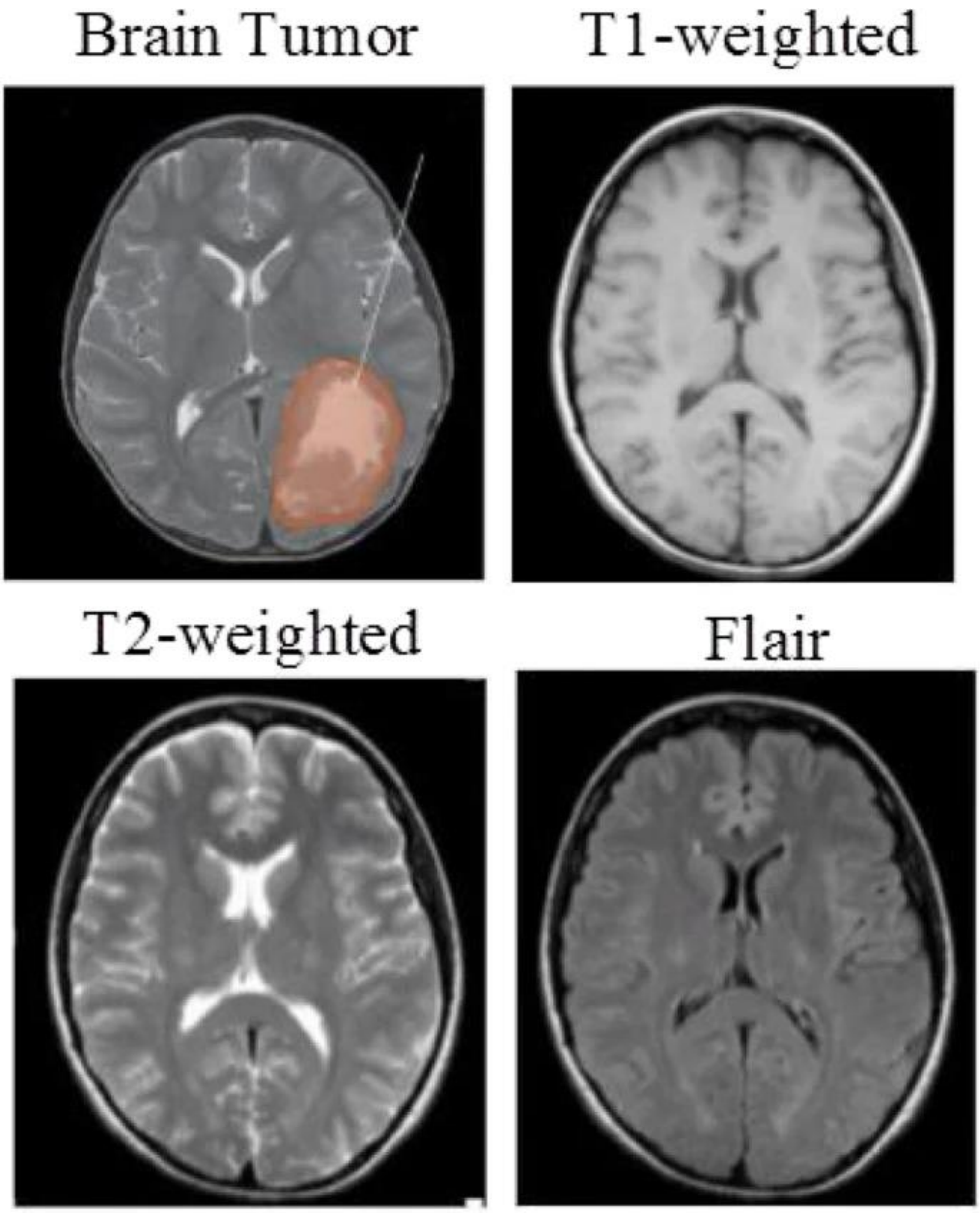


Figure 21 : Representation and the basic types of MRI [35]

The visibility of the tumor varies due to the different imaging techniques used. More in detail:

- Brain Tumor (upper left): this image specifically highlights the tumor, possibly with a contrast agent or post-processing to make it more visible. The tumor's distinct red color helps in its differentiation from the surrounding brain tissue.
- T1-Weighted (upper right): T1-weighted images are better at displaying anatomical detail than they are at emphasizing anomalies or lesions.
- T2-Weighted (lower left): This scan highlights fluid content in the brain
- FLAIR (lower right): Fluid-attenuated inversion recovery (FLAIR) imaging is used to suppress the effects of fluid, making the tumor more conspicuous by appearing brighter against a dark background, which helps in identifying lesions close to fluid-filled spaces like ventricles.

Deep architectures of CNN are normally utilized nowadays for MRI analysis of brain as pre-processing data, segmentation and identification of lesions and segmenting tumors [13,35].

Image Processing

One of the rapidly advancing technologies today is image processing (IP). It involves performing several operations to an image in order to improve its quality or extract important information. Tumor region detection and removal is a very difficult task in biomedicine. However, MRI imaging assists medical practitioners by providing clearer and more detailed images.

Numerous procedures are involved in the analysis of medical images, such as modeling, quantization, segmentation , registration, enhancement, and visualization. The steps involved in identifying tumor regions in IP include image acquisition, pre-processing, segmentation, and classification. Since no processing can take place without an image, image acquisition is the process of getting an image from a specified source. This is the first stage in the detection of brain tumors.

To reduce noise and enhance image quality, pre-processing is done once the image is acquired.

In order to remove ringing, noise, and blurring from the image, the pre-processing phase is necessary. The result of this phase is a clearer image which is more appropriate for further processing.

In the MRI acquisition phase, several pre-processing steps are necessary for an effective MRI segmentation and classification. The steps in pre-processing shown in Figure 22, guarantee that the image is refined and ready for various uses.

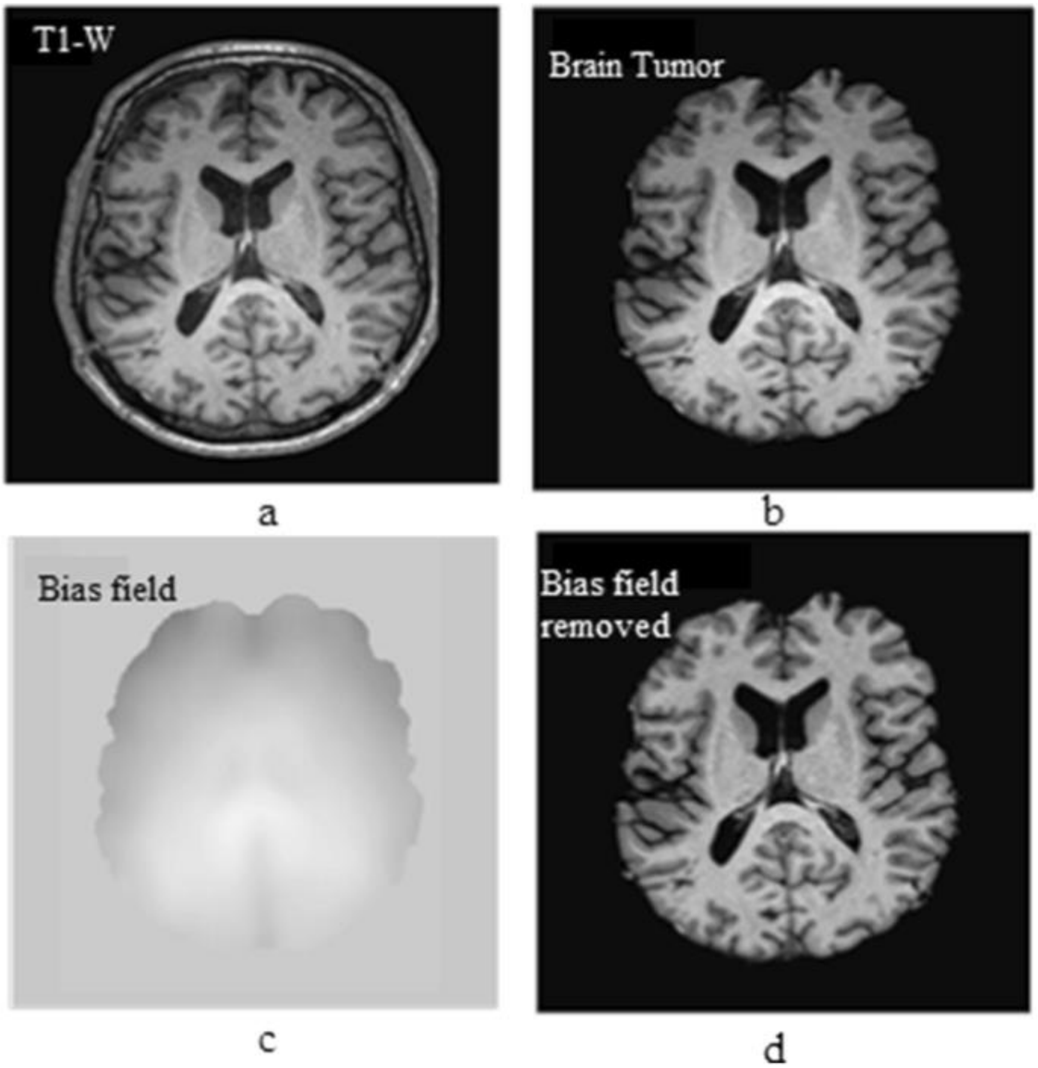


Figure 22 : Pre-processing example [35]

A) Adult brain T1-weighted MRI
B) Image of the brain after removing the non-brain structures
C) As in Data Science also in medicine there are biases. Bias Field indicates the relevant area for our task. In this case the Bias are the non brain structures.
D) The image of the brain after the pre-processing

Segmentation

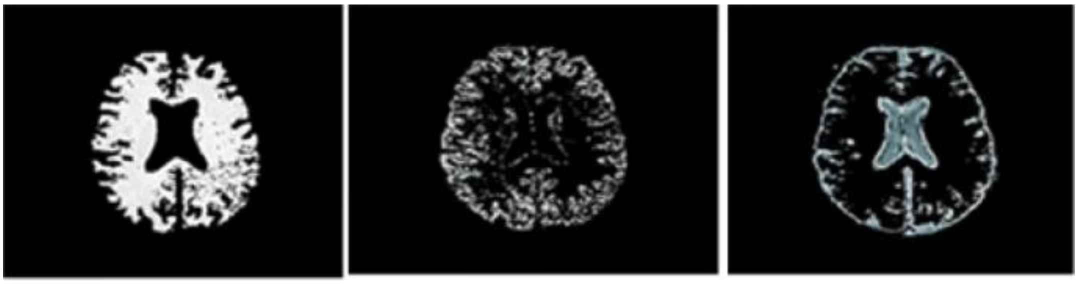
Segmentation splits one image into multiple segments. This partitioning is done among various tumor tissues (necrosis, edema, and solid tumor) and brain tissues, namely cerebrospinal fluid, gray matter, and white matter. These are classified using various principles. MRI segmentation techniques are divided into four categories for therapeutic applications: manual, semi-automatic, fully-automatic, and hybrid.

In the manual segmentation method, labeling of the image and segmentation of boundaries are performed by expert human operators using their hands.

Figure 22 displays the various brain segmentation categories. In reality, the manual segmentation approach takes a lot of time.

Both completely and semi-automated solutions are used to reduce the complexity involved in the manual process. Boundary segmentation is carried out automatically by a computer in automatic segmentation. Semi-automatic segmentation combines automatic segmentation with manual inspection.

Finally, fully automatic segmentation is considered superior to semi-automatic methods due to its optimal and highly effective outcomes [14,35].



i ii iii

Figure 22 ; The image displays the segmentation of brain categories: [35]

I) Gray Matter

II) White Matter

III) Cerebrospinal fluid

Classification

Classification is the final step of the tumor detection process, where the segmented regions of the MRI images are analyzed in order to determine the type and characteristics of the tumor.

For instance, one of the main objectives of classification is to categorize the segmented tumor regions as either being benign or malignant (non-cancerous and cancerous). Furthermore, if the analyzed tumor is malignant, the classification can further determine the specific type and grade of the tumor.

Classification is crucial for diagnosis, treatment planning and prognosis.

In fact, an accurate classification helps in diagnosing the specific type and grade of the tumor which is critical for determining the appropriate treatment plan, such as surgery, radiation therapy or chemotherapy. Furthermore, understanding whether a tumor is benign or malignant assists whit the prediction of the tumor's course and the with the final nature of the disease.

There are different techniques used for the classification such as post-processing techniques like 3D-CRF (3D Conditional Random Fields) and SIOX (Simple Interactive Objective Extraction). These techniques are used to refine the classification results in order to reduce errors and improve accuracy.

In conclusion, machine learning and deep learning approaches are the most commonly employed methods for classification. In fact, because of their exceptional capacity to extract intricate patterns and features from data, complex algorithms, especially neural network architectures (CNNs), are frequently employed for classification tasks [14, 35].

Enhancing Brain Tumor Classification with Deep Learning and CNNs

Nowadays, deep learning has shown remarkable results in medical image analysis, especially in the classification of brain tumors. In terms of accuracy, deep learning networks have outperformed traditional machine learning techniques. In this context, Convolutional Neural Networks (CNNs) have raised significant attention for its ability to automatically extract relevant insights, which are crucial for the final model output, and for their adaptability to minor variations in the images [15].

Figure 23 displays the architecture of a typical CNN-based framework for brain tumor classification.

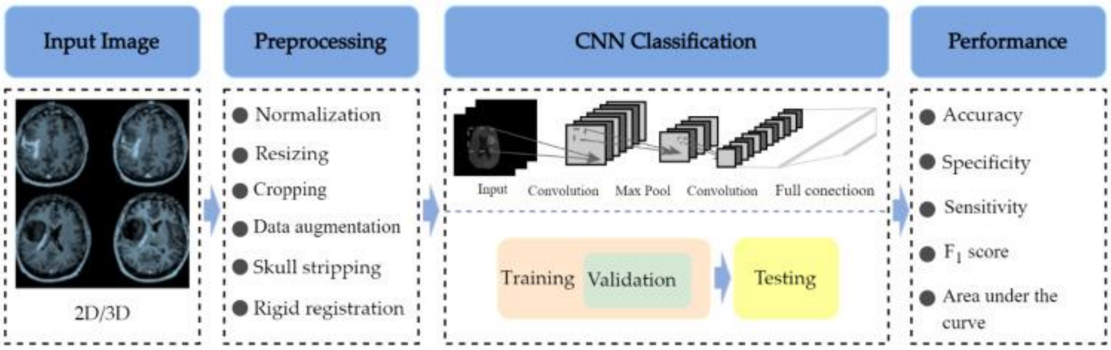


Figure 23 : Architecture of a typical CNN-based framework for brain tumor classification [36]

A CNN-based brain tumor classification study usually involves the following four steps in its workflow:

- Input image: The classification model receives 2D or 3D brain MRI scans as input
- Preprocessing: Images are resized, the skull is removed, and further training examples are introduced using a variety of preprocessing techniques.
- CNN Classification: The preprocessed dataset is fed into the CNN model, going through training, validation and testing processes.
- Performance Evaluation: Metrics including accuracy, specificity, F1 score, area under the curve (AUC), and sensitivity are used to evaluate the CNN algorithm's classification performance. To be more precise, the F1 score is a measure of a model's accuracy in binary classification and is particularly useful when the class distribution is imbalanced.

  Concluding with the AUC is a performance measurement for classification models at various threshold settings and basically tells how much the model is capable of distinguishing between classes. The closer the AUC value is to 1, the better the model's performance.

Chapter IV

Brain Tumor Detection with CNN in Practice

4.1 First steps

The last Chapter of the thesis presents a real application of CNNs models for the brain tumor recognition. This chapter through section 4.1 will introduce the first steps for the preparation of the model, moving to a breakdown of the model in section 4.2 and concluding with its performance evaluation in section 4.3.

As mentioned in the previous chapters, an abnormal mass or aggregation of brain cells is called a brain tumor. Since the skull provides a limited amount of space around the brain, any growth inside this area can cause issues. Brain tumors can be classified as benign (noncancerous) or malignant (cancerous). These benign or malignant tumors present a life-threatening risk because they may cause brain injury as they grow by increasing the pressure inside the skull.

Medical imaging research is crucial in the fields of brain tumor classification and early diagnosis. Precise classification assists in determining the best course of action for medical intervention, perhaps saving patients' lives.

Medical diagnosis has advanced significantly as a consequence of the adoption of deep learning techniques in the medical sector. The World Health Organization (WHO)[12] states that identifying the tumor, determining its precise position inside the brain, and classifying it according to its grade, type, and malignancy are all necessary for an accurate diagnosis of brain tumors.

The goal of this chapter is to utilize magnetic resonance imaging (MRI) to diagnose brain cancers, through CNN, highlighting its potential and accuracy.

The process includes tumor detection, classification by grade and type, and identification of the tumor's location. Instead of using separate models for each classification task, this approach utilizes a single model for classifying brain MRI images across different tasks.

To conclude, the CNN-based model is employed to segment the brain and identify the location of the tumor.

By leveraging machine learning techniques, such as CNN, we can significantly aid the advancement and development of neurology and ultimately improve healthcare outcomes for people affected by tumors.

Platform and Technical description

This code has been built using Google Colab which is a cloud-based Jupyter notebook environment, provided by Google. This platform is typically suited for these tasks such as deep learning, thanks to its support for GPU and TPU acceleration. The code has been constructed starting from an originary code about Medical Prediction, which was later adapted using different snippets of code [40,41,42] in order to reach the objective of Tumor Recognition.

During the code preparation, the following specifications were utilized:

Platform : Google Colab

CPU: Intel(R) Xeon(R) CPU @ 2.20GHz

GPU: NVIDIA Tesla

RAM: 13,290,464Kb → 12978.96 MB

Storage: size 108Gb, used 81Gb

Total Time Preprocessing, Training, Evaluation: 3 h and 3 min and 20 s

Training Time: 2 h and 53 min and 20 s

About the Dataset

The Dataset utilized to build the CNN comprehend a total of 7032 human MRI images. These images are categorized into four different classes (Fig. 24):

Glioma: cancerous brain tumor in glial cells.

Meningioma: Non-cancerous tumors originating from meninges.

No Tumor: Normal brain scans, healty brains with no detectable tumors.

Pituitary: Tumors affecting the pituitary gland, which can be cancerous or non-cancerous.

This Dataset focus its attention of brain tumors and their classification.

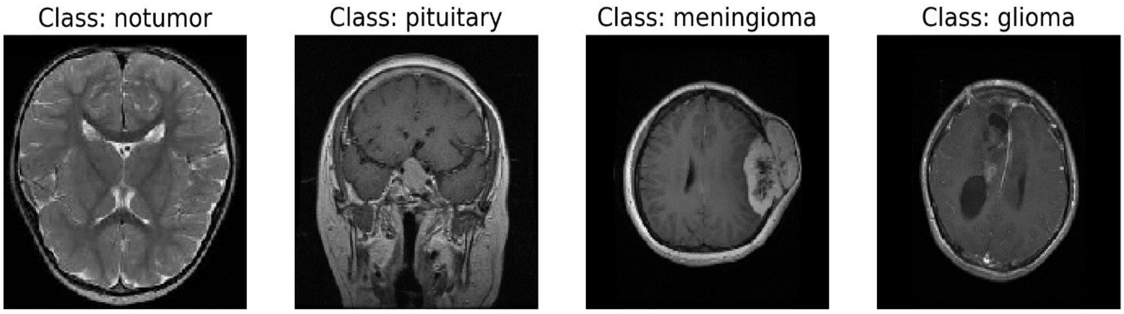


Figure 24 : Example of the class images

<u>Data Distribution</u>

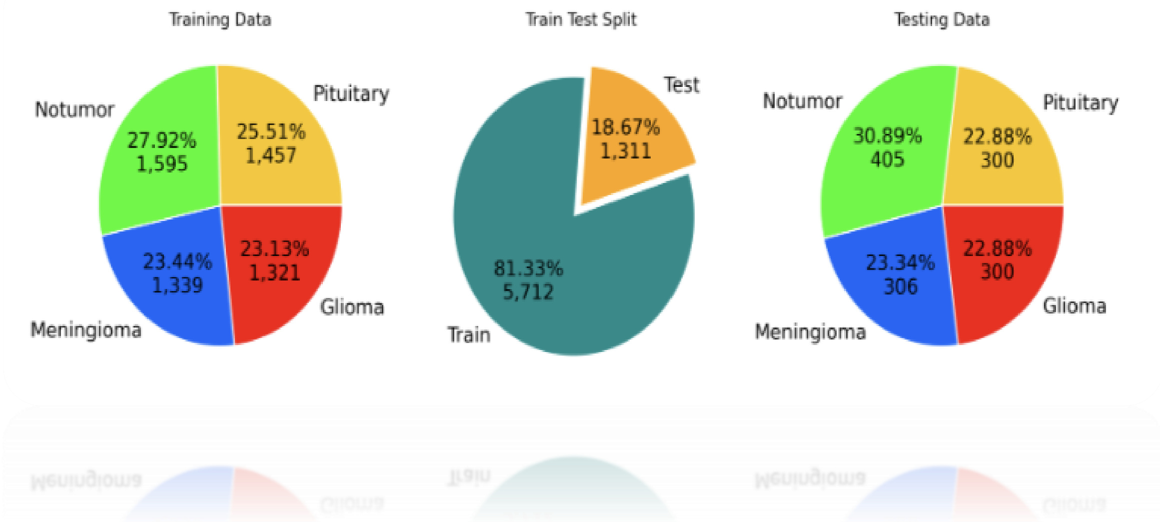


Figure 25 : Data distribution

As mentioned in the previous chapters, data distribution (Fig. 25) is one of the most important phases to train a CNN. For instance, for the best model learning, the training data must include a well-balanced distribution of categories. It makes it easier to fully comprehend the characteristics of each category, prevents against biases, improves generalization capabilities, and allows iterative improvement. Consequently, this results in enhanced capability for precisely classifying new data.

Following the data distribution, the next phase is the <u>data preprocessing and augmentation</u> for the training and testing of our CNN. In summary this part of the code prepares the training and testing data by augmenting the training images to improve our model robustness and normalizing all images to be sure that they are on the same scale. In fact, scale is crucial for training deep learning models effectively.

The next phase is the Training the Setup Values phase, built with the following code.

```
image_shape = (image_size[0], image_size[1], 3)

epochs = 40

steps_per_epoch = train_generator.samples // batch_size

validation_steps = test_generator.samples // batch_size

print(f'Image shape: {image_shape}')
print(f'Epochs: {epochs}')
print(f'Batch size: {batch_size}')
print(f'Steps Per Epoch: {steps_per_epoch}')
print(f'Validation steps: {validation_steps}')
```

```
Image shape: (150, 150, 3)
Epochs: 40
Batch size: 32
Steps Per Epoch: 178
Validation steps: 40
```

Image Shape: *the images are 150x150 pixels with 3 color channels.*

Epochs: *the model will train for 40 epochs*

Batch Size: *each batch contains 32 images.*

Steps per Epoch: *there will be 178 steps per epoch.*

Validation steps: *there will be 40 steps during validation.*

[40,41,42]

After training the setup values, we have to define the functions and the metrics for the CNN. The definition of function and metrics provide insights into how well the model is performing in predicting the correct class labels. In fact, especially when dealing with multi-class classification where the model can predict one of several possible classes, the evaluation metrics have to be adapted in a proper way.

Since we have multiple classes a confusion matrix is required: it is a table that allows the visualization of the performance of a classification model by providing a breakdown of predictions compared to the actual class labels. How does it work?

Each row of the matrix represents the instances in a class, while each column represents the instances in a predicted class.

There are some particular terms in the confusion matrix, such as:

- True Positives (TP): the number of instances correctly classified as a specific class.
- False Positives (FP): the number of instances incorrectly classified as a specific class that does not actually belong to it.
- False Negatives (FN): the number of instances belonging to a specific class but incorrectly classified as other classes.

There are also some different relevant metrics that must be known, such as Precision, Recall and F1-Score. The Precision measures the model's accuracy, identifying positive instances for each class among all instances predicted as positive. The Recall evaluates the model's ability to correctly identify positive instances for each class among all actual positive instances. Moving to the F1-Score, it represents the harmonic mean of precision and recall, providing a balanced measure that incorporates both metrics for each class. Finally, Accuracy it measures the overall correctness of the model's predictions across all of our classes.

Initial CNN Model Tests: in the initial CNN model tests, different combinations of filter sized and pooling sizes have been tried in our convolutional layers. Through these tests, we can notice that certain configurations performed better, resulting in higher test accuracies.

filter size of `filter_size = (4, 4)` and `pool_size = (2, 2)` the accuaracy was `Test Accuracy: 0.97`

filter size of `filter_size = (3, 3)` and `pool_size = (3, 3)` the accuaracy was `Test Accuracy: 0.97`

filter size of `filter_size = (4, 4)` and `pool_size = (3, 3)` the accuaracy was `Test Accuracy: 0.98`

filter size of `filter_size = (3, 3)` and `pool_size = (2, 2)` the accuaracy was `Test Accuracy: 0.98`

[40,41,42]

Among these combinations, the configurations that achieved the highest test accuracy have been selected and utilized for further tuning and improvements.

After selecting the best configurations, optimizers had to be implemented in order to determine their effect on the model's accuracy. Different optimizers provided different accuracies:

- Adam: test Accuracy: 0.982
- RMSprop: test Accuracy: 0.972
- Nadam: Test Accuracy: 0.964

The optimizer "Adam" was the one with the highest test accuracy, so that's the one used for this model.

4.2 Model Breakdown

In order to fully understand our CNN, let's breakdown it layer by layer. Figure 26 displays these layers. In each layer there are some transformations applied to the data.

- 1st Layer Convolutional:

  Input shape (150,150,3): indicates that the input images are 150x150 pixels with 3 color channels.

  Filters (32): the layer applies 32 different kernels (filters) to the input image

  Filter size (4,4): each filter is 4x4 pixels in size

  Activation ReLU: this activation function introduces non-linearity to the model, allowing it to learn more complex patterns.

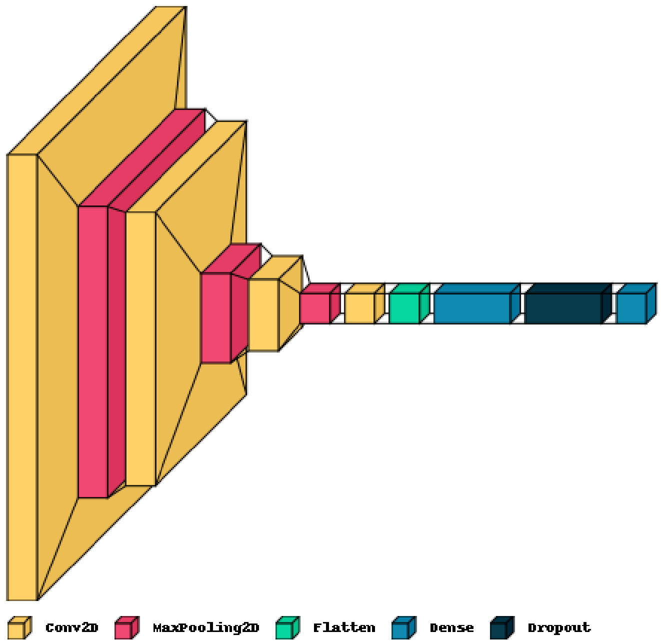


Figure 26: Model's Layers

  Output Shape (147,147,32): the output dimensions are calculated based on the input size and filter size: (150 - 4) + 1 = 147. The number of channels in the output corresponds to the number of filters.

- 2nd Layer Max Pooling:

  Pool size (3,3): this layer downsamples the input by taking the maximum value in each 3x3 region.

  Output shape (49,49,32): the output dimensions are reduced by a factor of the pool size: 147 / 3 = 49.

- 3rd Layer Convolutional:

  Input shape (49,49,32): The input to this layer is the output from the previous max pooling layer.

  Filters (64): this layer applies 64 filters.

  Output shape (46,46,64): The output dimensions are calculated as: (49 - 4) + 1 = 46.

- 4th Layer Max Pooling:

Pool size (3,3)

Output shape (15,15,64): The dimensions are reduced by a factor of the pool size: 46 / 3 = 15

- 5^{th} Layer Convolution:

  Input shape: (15, 15, 64) Filters: 128

  Filter size: (4, 4) Activation: ReLU

  Output shape: (12, 12, 128)

  The output dimensions are: (15 - 4) + 1 = 12.

- 6^{th} Layer Max Pooling:

  Pool size: (3, 3) Output shape: (4, 4, 128)

  The dimensions are reduced by a factor of the pool size: 12 / 3 = 4.

- 7^{th} Layer Convolutional:

  Input shape: (4, 4, 128) Filters: 128

  Filter size: (4, 4) Activation: ReLU

  Output shape: (1, 1, 128)

  The output dimensions are: (4 - 4) + 1 = 1.

- 8^{th} Flatten Layer:

  This layer with its function converts the 3D output from the previous layer into a 1D array.

  Output shape (128): this layer takes the (1, 1, 128) output and reshapes it into a 1D array of size 128.

- 9^{th} Fully Connected Layer (Dense Layer):

  Neurons: 512 Activation: ReLU

  Output shape (512): This layer has 512 neurons, each fully connected to the previous layer, applying the ReLU activation function.

- 10^{th} Dropout Layer:

  Dropout rate: 0.5. The Dropout layer is a regularization technique used in neural networks to prevent overfitting. In this case, this layer randomly sets 50% of the input units to 0 at each update during the training, in order to prevent overfitting.

  Output shape (512)

- 11^{th} Output Layer:

  Neurons: N_TYPES (number of output classes, which is 4 in this case)

  Activation: the Softmax activation function is a function utilized for multi-class classification.

Output Shape: N_TYPES (4)

- Summary of Model Parameters:

  Total trainable parameters: 495,972 (1,89MB). This is the total number of parameters (weights and biases) that will be adjusted during the training process.

  To summarize, this CNN model is built to classify images into four different categories. The architecture includes several convolutional layers with ReLU activation, which helps in learning complex features from the input images. Each convolutional layer is followed by a max pooling layer that reduces the spatial dimensions, in order to decrease the computational load and help in generalizing the model better.

  The model also includes a flatten layer to convert the 3D outputs into a 1D array, which is then fed into fully connected (dense) layers. A dropout layer is used to prevent overfitting by randomly deactivating half of the neurons during each training step. Finally, a Softmax layer is used to output the probability distribution across the four classes, enabling the model to make multi-class classifications.

Optimization:

This phase consisted in optimizing our model. We applied the Adam optimizer which is used with customized values for the learning rate, beta_1 and beta_2 parameters. In fact, by trying different values for beta_1 and beta_2, the training process is optimized for better convergence and performance. Finally, the model is compiled with the chosen optimizer and loss function, and accuracy is used as our evaluation metric.

Final Model Considerations:

This section represents the final version of our CNN, along with the optimizations and adjustments made to improve its performance.

Firstly, BatchNormalization layers were tested but they weren't so helpful for this specific model. Furthermore, we noticed that certain filter sizes, such as the last filter (3,3) did not significantly affect the model accuracy.

Regarding the Fine-Tuning, adjustments were made to beta_1 and beta_2 parameters of the Adam optimizer to optimize our training process. We explored a range of values for beta ∈ [0.7, 0.995] and beta_2 ∈ [0.9, 0.9995] in order to achieve the highest validation accuracy.

Adam Parameters explained:

Learning_rate: it determines the step size for adjusting the model weights during training.

Beta_1: exponential decay rate (for the first moment estimates), controls how quickly the moving average of past gradients decays.

Beta_2: works like beta_1 but for the second moment estimates.

Epsilon: a small value added to the denominator just for numerical stability.

Decay: gradually decreases the learning rate over time

Amsgrad: a Boolean value indicating whether to use the AMSGrad variant of the Adam optimizer.

Clipnorm and Clipvalue: Parameters for capping the norm or value of the gradients. It prevents gradients from becoming too large by capping them at a specified maximum value.

Callback for Training Enhancement: there are 2 callbacks in this model. The first one is EarlyStopping, which stops the training if the loss stops decreasing in the validation set. The second one is ReduceLROnPlateau, which reduces the learning rate if the validation loss plateaus[42].

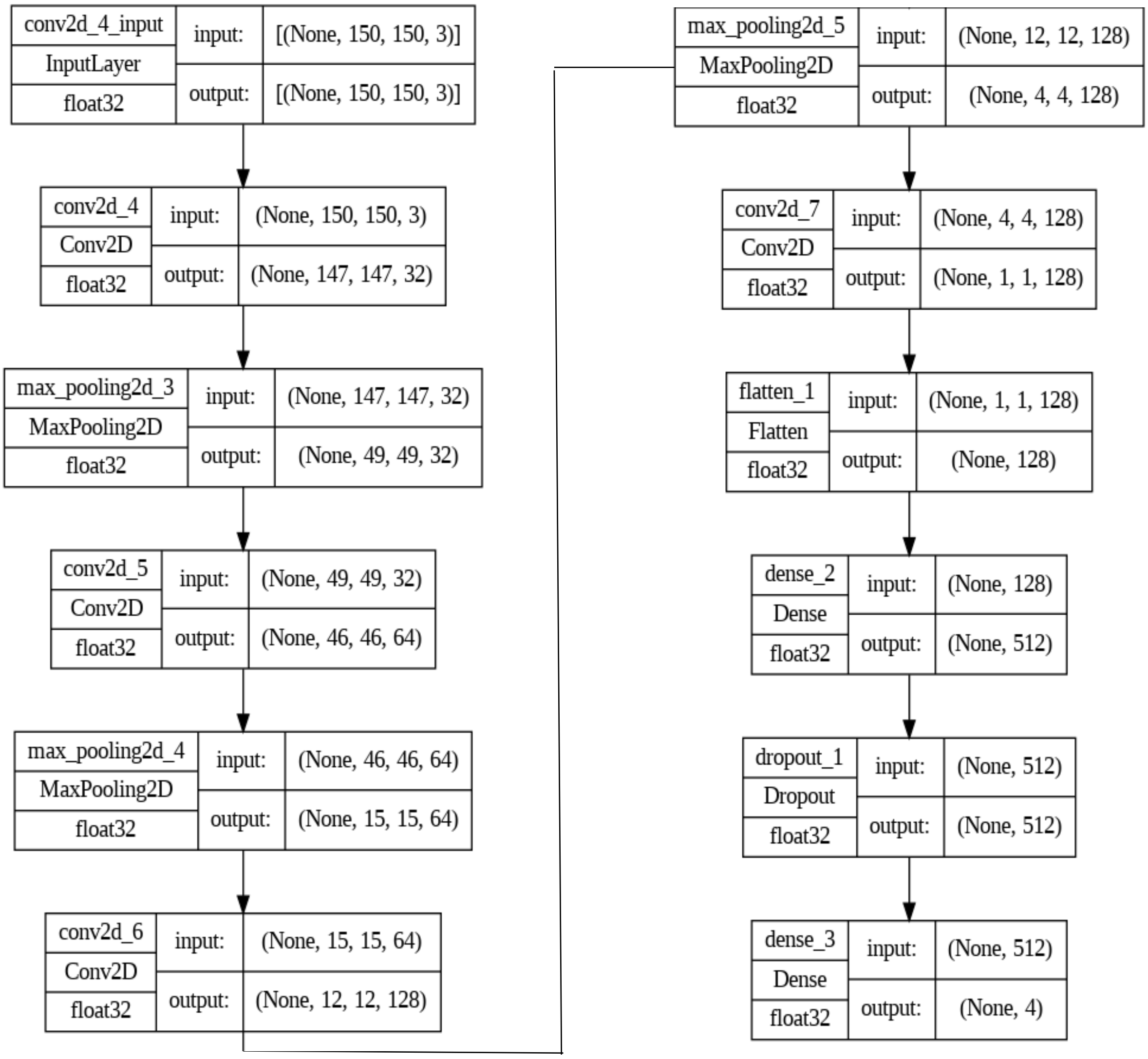


Figure 27 : Representation of the model's layers

4.3 Model Evaluation

As mentioned before, the model is evaluated on the test data. These are the evaluations:

Test Loss: 0.07839 Test Accuracy: 0.98906

These evaluations are displayed in the plots below (Fig.28).

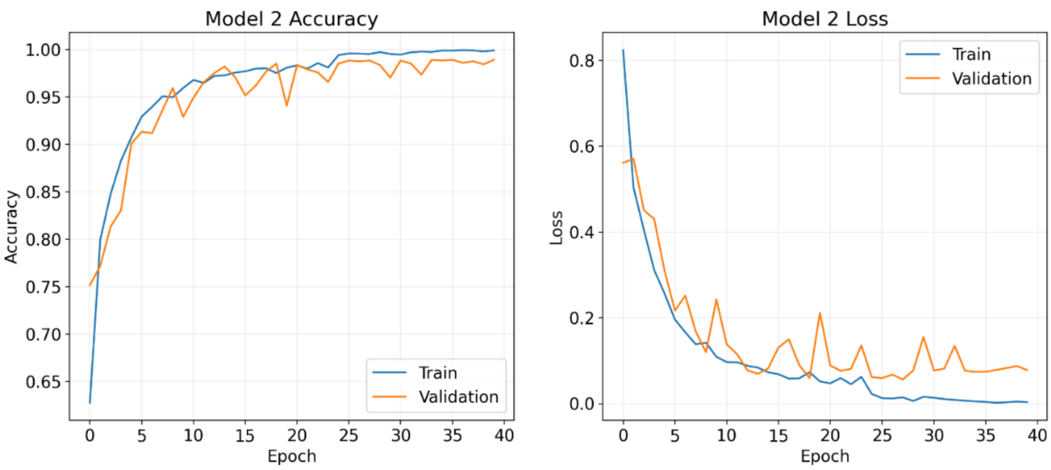


Figure 28 : Accuracy & Loss Visualizations

Following the plots of accuracy and loss, there is the confusion matrix (Fig. 29) which is plotted to visualize the performance of the model on each class. This plot helps to see the number of correct and incorrect predictions for each class.

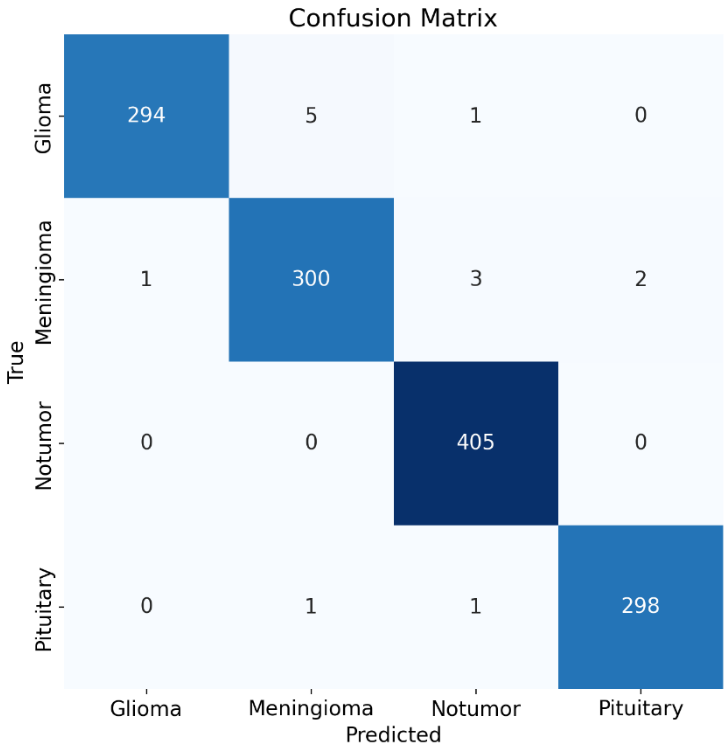


Figure 29

Confusion Matrix (Fig. 29) explanation: it shows the performance of the CNN model on the test dataset.

- Glioma:

  True Positives (TP): 294

  False Positives (FP): 1 (1 instance of notumor classified as Glioma)

  False Negatives (FN): 6 (5 instances of Glioma classified as Meningioma and 1 instance of Glioma classified as Notumor)

  True Negatives (TN): Sum of all other correctly classified instances

- Meningioma:

  TP: 300

  FP: 5 (5 instances of Glioma classified as Meningioma)

  FN: 6 (3 instances of Meningioma classified as Notumor, 2 instances of Meningioma classified as Pituitary)

- Notumor

  TP: 405

  FP: 2 (2 instances of other classes classified as Notumor)

  FN: 0 (No instances of Notumor misclassified)

- Pituitary:

  TP: 298

  FP: 2 (2 instances of other classes classified as Pituitary)

  FN: 1 (1 instance of Pituitary classified as Meningioma)

Below (Fig. 30) the final evaluation of our model is displayed.

	precision	recall	f1-score	support
Glioma	1.00	0.98	0.99	300
Meningioma	0.98	0.98	0.98	306
Notumor	0.99	1.00	0.99	405
Pituitary	0.99	0.99	0.99	300
accuracy			0.99	1311
macro avg	0.99	0.99	0.99	1311
weighted avg	0.99	0.99	0.99	1311

Accuracy: 0.989

Figure 30 : Model Evaluation

The model performs well, with precision, recall and F1-scores being high across all our classes. Furthermore, the overall accuracy of 0.989% (0.99%) indicates a robust model and our confusion matrix demonstrates that the model is particularly strong in identifying tumors.

True vs Predicted output

In the Figure 31, there are represented the true labels vs the predicted labels for a set test of images. The green labels indicate correct predictions, showing our model's ability to accurately classify the images. This visualization helps in understanding the general performance of the model between the different classes.

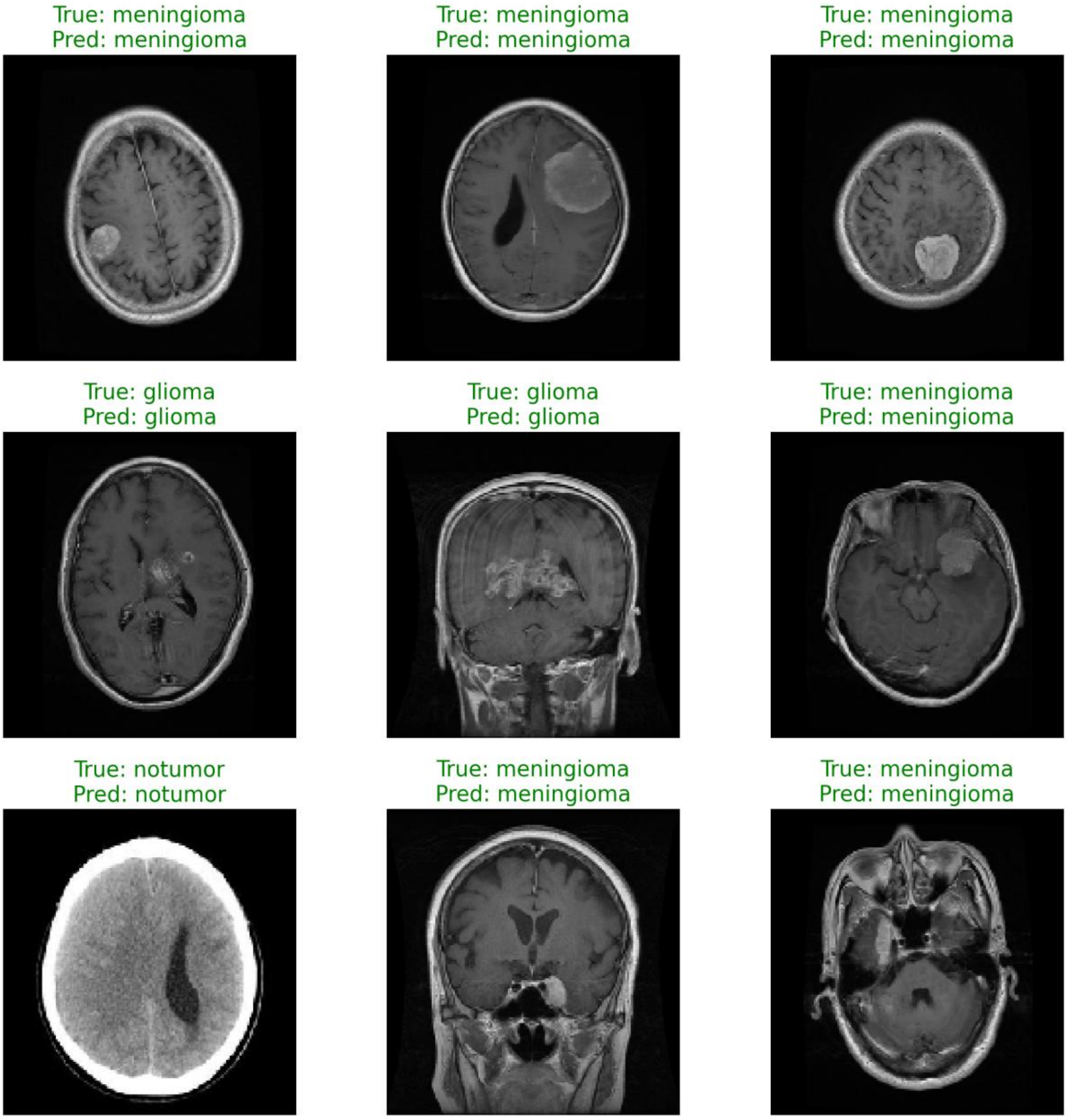
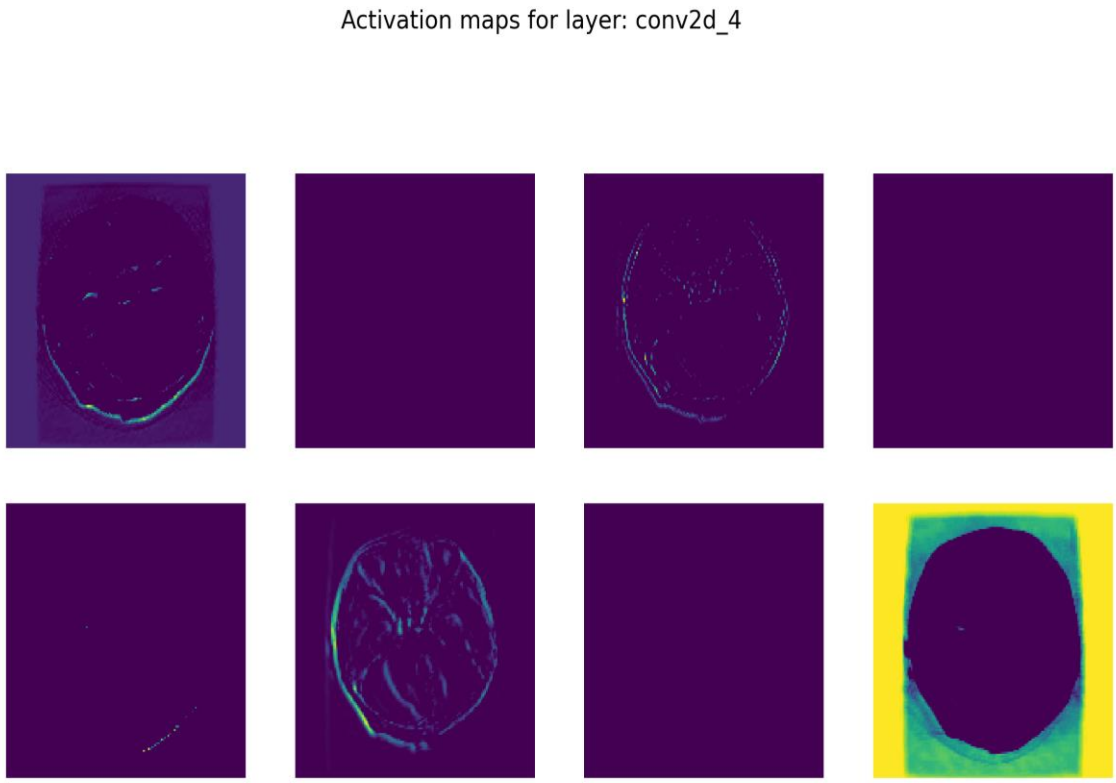


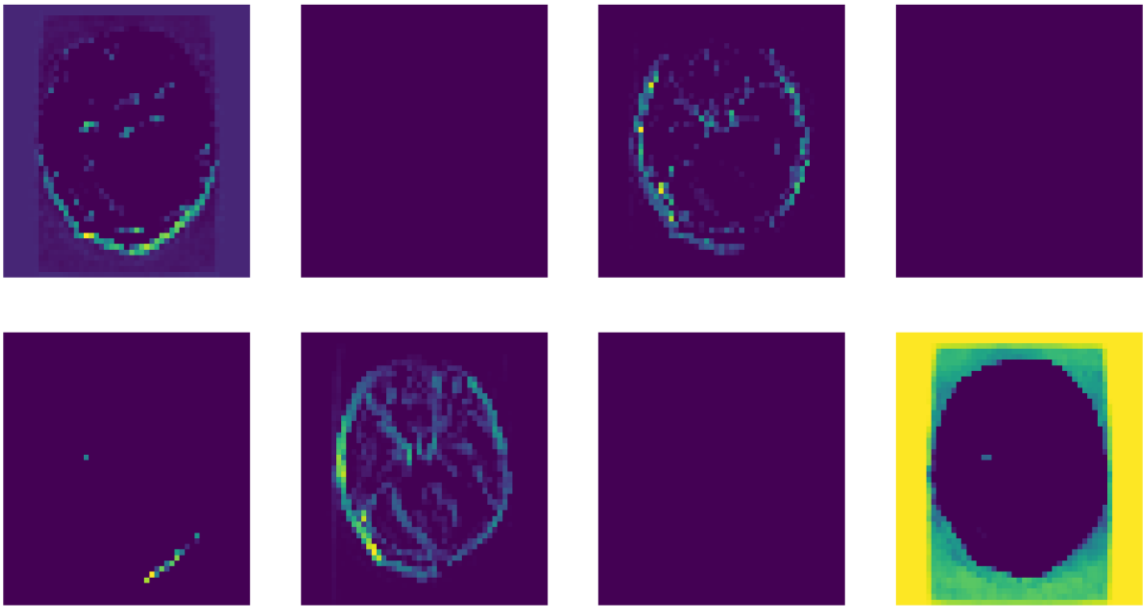
Figure 31 : True Labels vs Predicted Labels

Activation Map

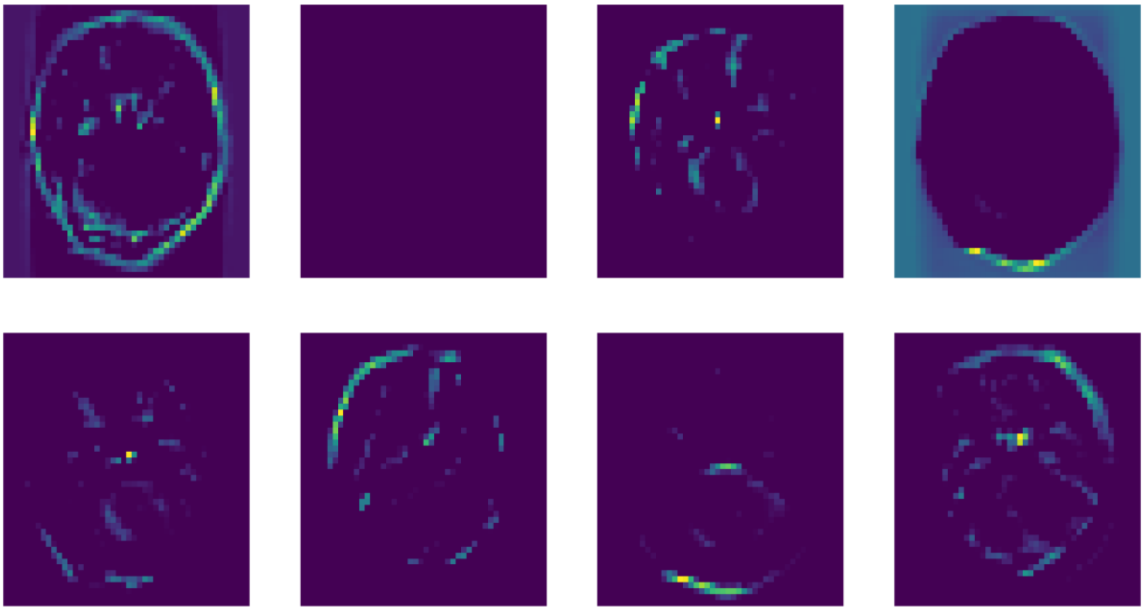
Figure 32 below, displays the activation maps from a specific convolutional layer (conv2d_6) in our model. These images highlight the regions of the input image that activate specific neurons in the convolutional layer. These maps can provide insight into what parts of the image are being used by the model to make its predictions. In this case, the activation maps show the parts of the MRI scans that the model focuses on, indicating which features are deemed important for classification.

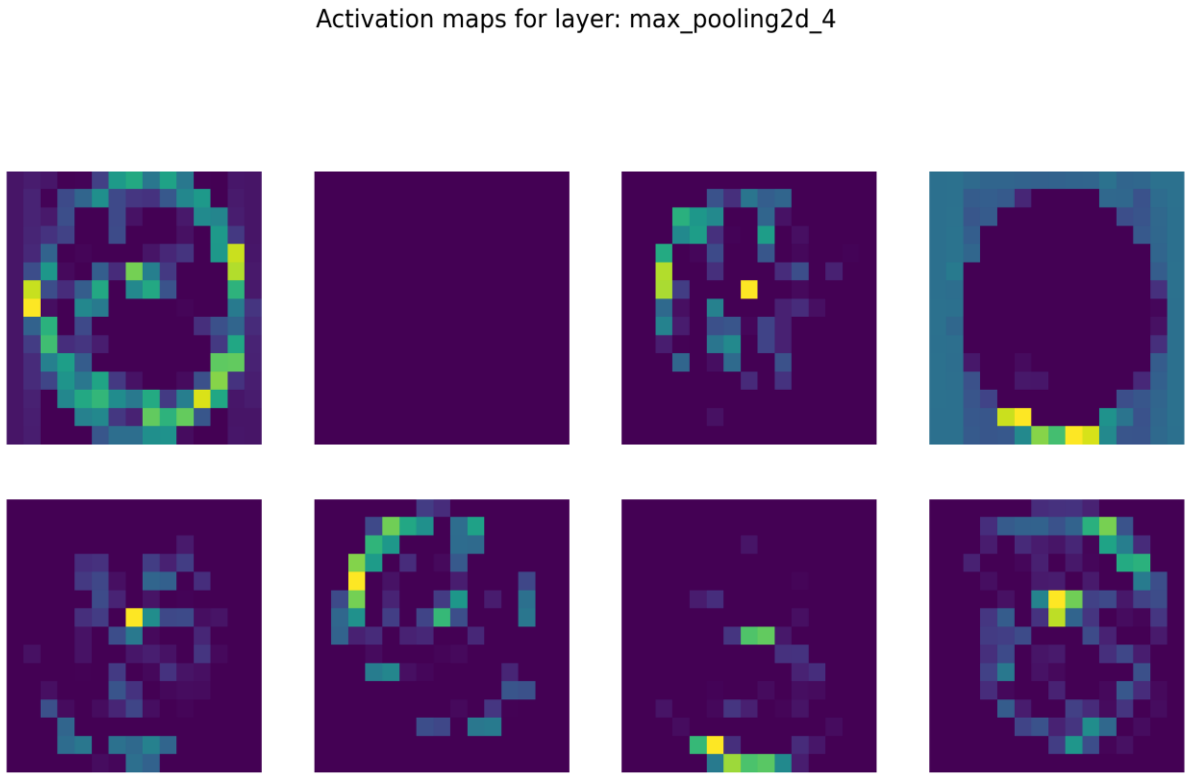


Activation maps for layer: max_pooling2d_3



Activation maps for layer: conv2d_5





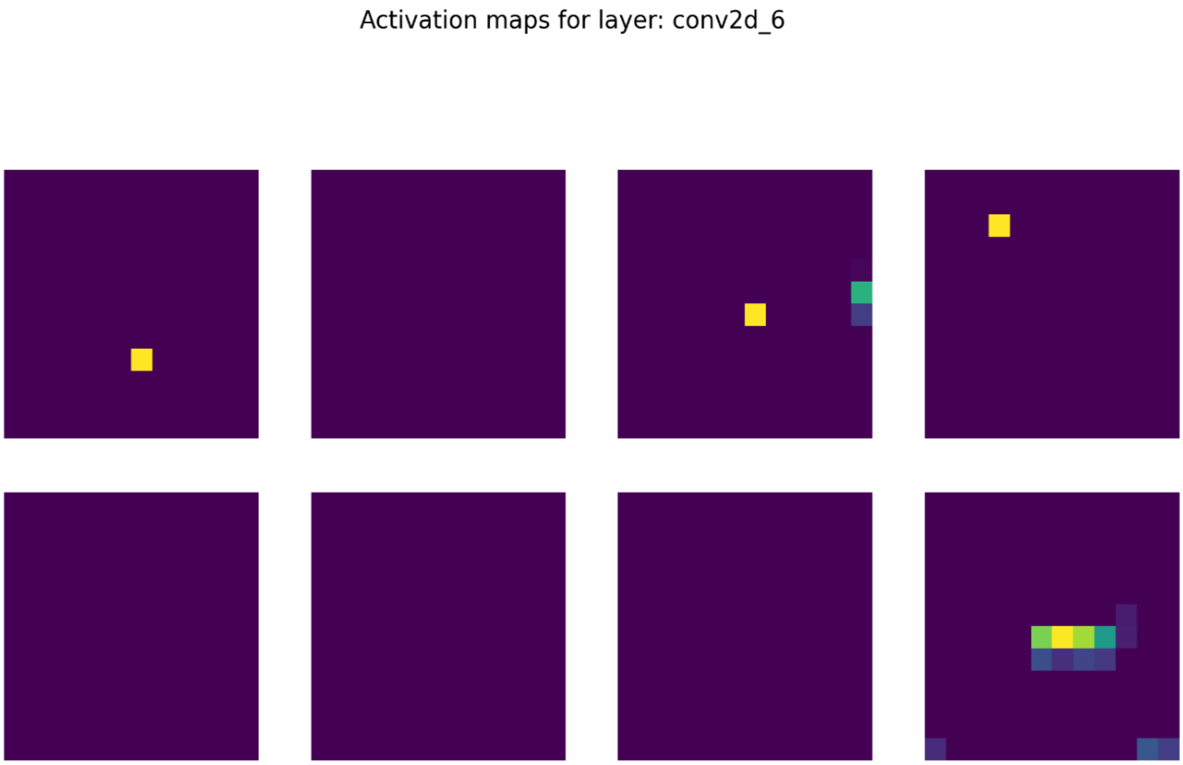


Figure 32 : Activation Maps

Misclassified Tumors

Concluding with the misclassification, Figure 33 focuses on misclassified images, showing true labels and the model's incorrect predictions. This image is crucial for error analysis. By examining the misclassified instances, we can identify potential patterns or common features that the model struggles with. For example, certain misclassifications might be due to similarities in tumor appearance between different classes. This information can guide further model refinement or the need for additional training data.

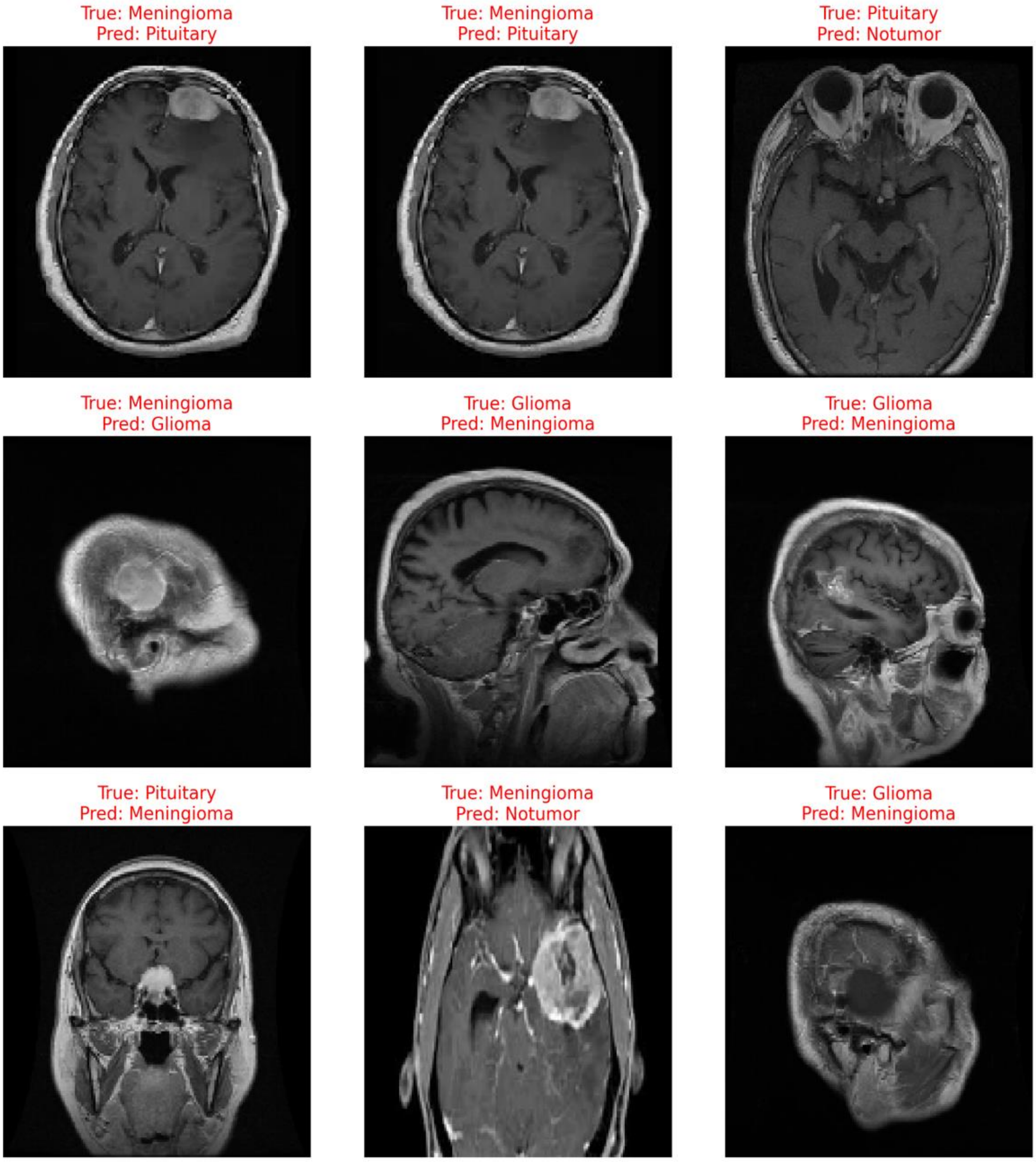


Figure 33 : Misclassification

Chapter V

Conclusions

Deep Learning techniques, in particular Convolutional Neural Networks, have revolutionized the process of identification and classification of brain tumors.

The thesis has explored different key areas, including the integration of innovative medical technologies, a general overview of the fundamentals of CNNs, the characteristics and challenges of brain tumors and to conclude, a practical application of CNNs for tumor detection. Through this journey, different findings emerged.

Starting with the Medical Technologies, advancements in medical technology (such as MRI) have improved by a lot the detection and monitoring of brain tumors. In fact, these imaging modalities provide essential information about the size, location and type of tumors. These information are crucial for planning a treatment and for patient management.

Moving to CNNs, these models have demonstrated a better and superior performance in feature extraction and image classification, compared to traditional machine learning techniques. Thanks to their architecture and key components (such as convolutional and pooling layers) these models have the capability to automatically learn and identify intricate patterns in medical images. These characteristics make them highly effective for tasks such brain tumor detection.

CNNs represent an innovative and transformative advancement in medical image analysis, and it offers an unprecedented accuracy in detecting and classifying brain tumors.

Despite their success, there are also many challenges to overcome such as overfitting, data scarcity and the need for standardized datasets. These challenges could be solved through continue research to develop more robust models that maybe could generalize better to diverse and unseen data and through the creation of large and standardized datasets.

Furthermore, in order to have a wide clinical adoption of these technologies, it’s required the collaboration between technologists and healthcare professionals with the goal of creating accurate models usable in real-world settings.

Below there is a table that highlights the application areas and research fields along with the AI techniques used, explored in the thesis:

Application/Area/Research	**AI Technique Used**
Medical Image Analysis	Convolutional Neural Networks (CNNs)
Brain Tumor Detection	Deep Learning, CNNs
Image Segmentation	U-Net, SegNet, ResNet18
Image Classification	LeNet, AlexNet, VGGNet, ResNet, ZFNet
Preprocessing Techniques	Normalization, Skull Removal, Image Augmentation
Feature Extraction	CNNs, Deep Features
Overfitting Mitigation	Data Augm., Dropout, Batch Normalization, Weight Decay
Performance Evaluation	Accuracy, F1 Score, AUC, Sensibility

To conclude, the use of CNNs in brain tumor detection is a big advancement in medical image analysis. The potential of these technologies can be completely exploited by addressing the existing obstacles and conducting further model refinement, which will ultimately improve medical tools and patient outcomes.

Bibliography:

[1] Wang Y, Byrd TA. Business analytics-enabled decision-making effectiveness through knowledge absorptive capacity in health care. J Knowl Manag. 2017;21(3):517–39.

[2] Szeliski, R. *Computer Vision: Algorithms and Applications* (Springer Science & Business Media, 2010).

[3] Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).

[4] Papolos, A., Narula, J., Bavishi, C., Chaudhry, F. A. & Sengupta, P. P. U. S. Hospital
use of echocardiography: insights from the nationwide inpatient sample. J. Am.
Coll. Cardiol. 67, 502–511 (2016).

[5] HeartFlowNXT—HeartFlow Analysis of Coronary Blood Flow Using Coronary CT Angiography—Study Results—ClinicalTrials.gov.

[6] Ouyang, D. et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature 580, 252–256 (2020).

[7] Russakovsky, O. et al. ImageNet Large Scale Visual Recognition Challenge. Int. J. Comput. Vis. 115, 211–252 (2015).

[8] Esteva, A. & Topol, E. Can skin cancer diagnosis be transformed by AI? Lancet 394, 1795 (2019)

[9] Ting, D. S. W. et al. Artificial intelligence and deep learning in ophthalmology. Br. J. Ophthalmol. 103, 167–175 (2019).

[10] Dehury CK, Sahoo PK (2017) Design and implementation of a novel service management framework for IoT devices in cloud. J Syst Softw 119:149–161

[11] Wadhwa A, Bhardwaj A, Verma VS (2019) A review on brain tumor segmentation of MRI images[J]. Magn Reson Imaging

[12] Lewis EB, Fox NC (2004) Correction of differential intensity inhomogeneity in longitudinal MR images. NeuroImage 23(1):75–83

[13] Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bendszus M, Biller A (2016) Deep MRI brain extraction: a 3D convolutional neural network for skull stripping. Neuroimage 129:460–469

[14] Murgasova M (2008) Segmentation of brain MRI during early childhood. Dissertaion, Imperial College London

[15] Deepak S., Ameer P.M. Brain tumor classification using deep CNN features via transfer learning. *Comput. Biol. Med.* 2019;111:103345.

[16] Fakoor R , Ladhak F , Nazi A , Huber M . Using deep learning to enhance cancer diagnosis and classification . A conference presentation . The 30th International Conference on Machine Learning , 2013 .

[17] Vial A , Stirling D , Field M et al . The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review . Transl Cancer Res 2018 ; 7 : 803 – 16 .

[18] Deep learning for healthcare: review, opportunities and challenges Riccardo Miotto, Fei Wang, Shuang Wang, Xiaoqian Jiang, Joel T Dudley

[19] Challenges and opportunities in IoT healthcare systems: a systematic review Sureshkumar Selvaraj & Suresh Sundaravaradhan

[20] The Recent Progress and Applications of Digital Technologies in Healthcare: A Review, Maksut Senbekov,[1]Timur Saliev,[1]Zhanar Bukeyeva et al.

[21] Towards the Use of Big Data in Healthcare: A Literature Review, Grazia Dicuonzo et al.

[22] The use of Big Data Analytics in healthcare, Kornelia Batko & Andrzej Slezak

[23] Bush J. How AI is taking the scut work out of health care . Harvard Business Review 2018

[24] The potential for artificial intelligence in healthcare, Thomas Davenport and Ravi Kalakota

[25] Deloitte Insights . State of AI in the enterprise . Deloitte , 2018 . www2.deloitte.com/content/dam/insights/us/articles/4780_Stateof-AI-in-the-enterprise/AICognitiveSurvey2018_Infographic.pdf .

[26] Davenport TH , Hongsermeier T , Mc Cord KA . Using AI to improve electronic health records . Harvard Business Review 2018 . https://hbr.org/2018/12/using-ai-to-improve- electronichealth-records

[27] Berg S. Nudge theory explored to boost medication adherence . Chicago : American Medical Association , 2018 . www.ama-assn.org/ delivering-care/patient-support-advocacy/nudge-theory-exploredboost-medication-adherence .

[28] Utermohlen K. Four robotic process automation (RPA) applications in the healthcare industry . Medium , 2018 . https://medium.com/@ karl.utermohlen/4-robotic-process-automation-rpa-applications-inthe-healthcare-industry-4d449b24b613 .

[29] Deloitte . From brawn to brains: The impact of technology on jobs in the UK . Deloitte , 2015 . www2.deloitte.com/content/dam/Deloitte/uk/Documents/Growth/deloitte-uk-insights-from-brawns-to-brain.pdf .

[30] https://research.google/blog/advancing-medical-ai-with-med-gemini/

[31] Deep learning-enabled medical computer vision, Andre Esteva et al.

[32] Convolutional neural networks: an overview and application in radiology, Rikiya Yamashita et al.

[33] Deep Learning for Smart Healthcare—A Survey on Brain Tumor Detection from Medical Imaging, Mahsa Arabahmadi et al.

[34] Understanding Cancer, National Library of Medicine, National Center for Biotechnology Information

[35] A comprehensive review on brain tumor segmentation and classification of MRI images. Champakamala Sundar Rao et al.

[36] Convolutional Neural Network Techniques for Brain Tumor Classification (from 2015 to 2022): Review, Challenges, and Future Perspective , Leonardo Rundo et al.

[37] Daimary, D.; Bora, M.B.; Amitab, K.; Kandar, D. Brain Tumor Segmentation from MRI Images using Hybrid Convolutional Neural Networks. Procedia Comput. Sci. 2020, 167, 2419–2428.

[38] Kim SH, Chung K (2015) Emergency situation monitoring service using context motion tracking of chronic disease patients. Clust Comput 18(2):747–759

[39] Davenport TH , Dreyer K. AI will change radiology, but it won't replace radiologists . Harvard Business Review 2018 .

[40] Kaggle, Qaid_Jamali https://www.kaggle.com/code/qaidjamali/brain-tumor-classification-98-acc

[41] Kaggle, Halemo GPA, https://www.kaggle.com/code/halemogpa/brain-tumor-classification-pytorch-99-3-test

[42] Kaggle, Supriyo Ain, https://www.kaggle.com/code/supriyoain/brain-tumor-detection-pytorch-99-accuracy