6-2021

# The State of Artificial Intelligence in Medical Imaging

Catalin Cristian Veghes

# The State of Artificial Intelligence in Medical Imaging

Catalin Cristian Veghes

Master of Science in Information Technology, Clark University

Capstone Thesis

Professor Richard Aroian

June 1, 2021

## Acknowledgments

The completion of this research thesis could not have been possible without the encouragement and assistance of my advisor, Mr. Richard Aroian, who provided all the necessary advice, tools, and resources for a successful academic endeavor.

I want to thank all my professors at Clark University who contributed to my education throughout my journey as an undergraduate and graduate student. Without their support and dedication, none of my achievements could have been possible.

I am incredibly grateful to my family, girlfriend, and friends for giving me the motivation I needed during these challenging times caused by the COVID-19 pandemic.

## Table of Contents

# Abstract

*This study explores the current state of Artificial Intelligence in medical imaging and provides an accessible assessment of how radiologists perceive the emerging technologies. Throughout the research, we analyze different aspects such as the adoption rate of Artificial Intelligence or the performance of state-of-the-art models, and we identify some of the significant barriers that prevent a wider adoption, such as the lack of collaboration between radiologists and computer scientists. Additionally, we provide a brief theoretical background that explains how deep learning works and how it can be helpful in medical imaging. We describe the architecture of a binary classifier in detail and exemplify several measurements that can be used to evaluate an AI model. The paper concludes with our personal opinion on the subject.*

# Chapter 1: Introduction

## a. General Introduction of the Research Project

Over the last decades, technology has been evolving at an ever-increasing pace. In modern society, technology has become not only an instrument that improves the quality of our lives (Wardlaw, 2004) but also a tool that ensures our survival as humans (Gelsinger, 2017). Starting with mobile devices and personal computers that offer us an impressive processing power right at our fingertips and continuing with complex networks of asteroid-hunting telescopes that are capable of spotting large space bodies heading for Earth, scientists have sought to incorporate technological progress in all areas of life (Davis, 2019). On top of solving intricate problems that generations after generations have tried to resolve, technology has opened a new world full of possibilities and challenges. By leveraging powerful computers and diverse, intelligent algorithms, we managed to find solutions to conventional problems, and we now have a solid foundation to aim even higher.

Technology plays a significant role in almost all industries, but the explicit dependency on computers varies significantly from one sector to another. In the particular field of medicine and healthcare, technology has increasingly become an indispensable asset (Marsch, 2013). Most of its applications aim to automate the tasks and processes that previously had to be done by people manually. Not only it saves valuable time, which can be used for other purposes (such as research, specialized training, or personalized medicine), but it also improves our understanding of complex diseases and has the potential to reduce the overall cost of medical care. Especially during the current pandemic, technology proved to be a critical tool for the health sector. Through telehealth, medical professionals provided valuable services like medication management or online consultations to their patients at a low additional cost (Weiner, 2021).

After writing millions of lines of code and building fairly complex programs, scientists started asking whether computers can think and act like humans. This intense curiosity correlated with continuous technological advancement led to the emergence of a revolutionary field in Computer Science commonly known as Artificial Intelligence (AI). AI can be described as the ability of a digital computer to perform tasks that are generally associated with intelligent beings. In other words, AI systems strive to be capable of mimicking some of the cognitive functions of the human brain, such as learning, planning, reasoning, or problem-solving (Heath, 2020).

As AI moved forward, it branched off into different techniques. One of the most prominent techniques is Machine Learning (ML), which includes all the algorithms that allow computers to learn from data. Unlike traditional programs that represent sets of hard-coded instructions executed in a predefined order based on logical conditions, ML-enabled tools rely on dynamic algorithms that can adapt in response to learned data (Six, n.d.). Among other approaches that fall under the ML umbrella, deep learning (DL) is a promising technique that can achieve state-of-the-art results for classification or detection tasks, sometimes outperforming human subjects. It is based on artificial neural networks and teaches computers to learn by example in a very similar way to how humans learn. *Figure 1* illustrates the relationship between Artificial Intelligence, Machine Learning, and Deep Learning and offers a brief overview of the algorithms they encapsulate.

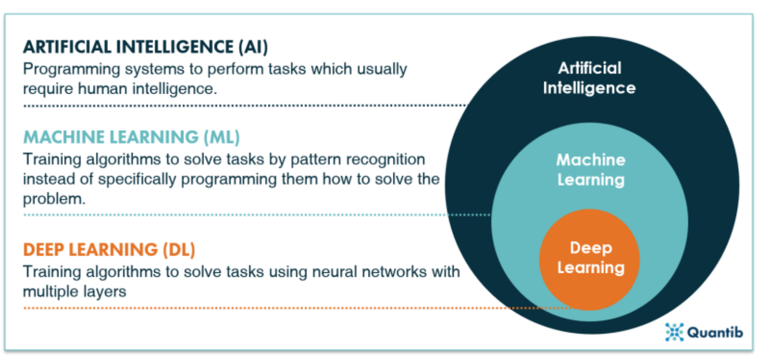


*Figure 1 - Relationship of AI, ML, and DL - source: quantib.com/the-ultimate-guide-to-ai-in-radiology*

Based on the paramount importance of technology in the medical sector and the recent developments of intelligent systems that have been enabled by the advancement of computers and the explosion of big data, more and more AI tools have been adopted and integrated into clinical workflows across the world (Shuaib, 2020). In particular, one of the most notable subfields where these systems have been proliferating is medical imaging, one of the most vital activities used to diagnose and treat a wide range of diseases, including cancer, chronic respiratory illnesses, brain-related injuries, cardiovascular problems, and plenty more (Bresnik, 2018). However, since both AI and medical imaging are still-emerging fields that are continuously evolving in front of our eyes, many unknowns and topics of interest have yet to be explored.

### b. Research Problem

Given the rapid growth of technology and the multitude of human-like algorithms that appear at lightning speed, it becomes a challenge to determine the actual state of AI in specific highly specialized sectors such as medical imaging. Even for a knowledgeable audience capable of staying informed and up-to-date with breakthrough discoveries, it is not easy to assess right away the impact AI brings on biological imaging. This paper strives to investigate this problem and respond to an essential question: *what is the current state of Artificial Intelligence in medical imaging?* Because this is a fairly complex question and a comprehensive response is mandatory, this research analyzes the topic from multiple perspectives. It evaluates the objective performance of the models that have been proposed in the literature, their adoption level, and the informed opinions of radiologists regarding their potential versus the actual usage. Our work aims to explore the topic and offer a holistic analysis of how much radiologists use AI, its limitations, and the future of an AI-powered imaging industry. The results and conclusions are intended to be accessible to a wide range of people, including those who are not necessarily familiar with advanced computer science concepts.

### c. Rationale for Research Project

In general, people interested in technology can keep up to date with the newest discoveries by reading news and searching topics of their interest on the Internet. When it comes to AI, there is a vast pool of resources that can be used to gain knowledge and understand how things are evolving in this field. It is pretty common to see articles or documentaries that discuss various aspects of intelligent technologies in modern society. However, most of these resources refer to AI-based technologies accessible to the large public, like speech recognition devices, mobile applications, self-driving cars, and chatbots. Unfortunately, in the case of medical imaging, the latest news is not readily available, and most of the revolutionary innovations are mainly discussed in research papers or academic journals that require a superior level of understanding. For this reason, we deemed it necessary to conduct this research and provide an accessible assessment of what is going on in the industry at the moment.

The starting point of our work was an interview with Regina Barzilay (Heaven, 2020), a professor at MIT and the first recipient of the Squirrel AI Award, a prize that competes at a financial level with the Nobel Prize and Turing Award. After winning one million dollars for her outstanding research on machine-learning algorithms for detecting cancer and designing new drugs, Barzilay raised the hypothesis that AI does not have the acceptance of society yet, especially in those areas where the cost of failure is very high, such as medical imaging. In the article published by MIT Technology Review, she shared a personal story about the barriers she encountered as a patient suffering from breast cancer. While going through the standard procedures, she realized how much of the AI's potential was not exploited and volunteered to collaborate with doctors to find solutions that would enhance the therapeutic process. Based on her previous experience in natural-language processing, she attempted to research the problem and develop intelligent algorithms that could be useful in answering life-and-death questions. Unfortunately, her efforts barely crystallized into actual discoveries due to a lack of data. Surprisingly, she stated that the main reason behind AI not being more successful in healthcare is not the lack of technology but rather the absence of a collaborative effort between institutions that provide the raw data and computer scientists who develop algorithms. Therefore, we decided to explore this topic and thoroughly understand the wake-up call made by the MIT professor.

Moreover, as AI becomes an omnipresent concept in our daily lives, people start experiencing mixed feelings about its applications in medicine and other related fields. Some people argue that by introducing AI-based systems, most radiologists will end up losing their jobs. Other opposers are extraordinarily skeptical and believe it is far too risky to make medical decisions based on the feedback from a computer, mainly because they consider themselves complex beings that machines cannot understand. Consequently, we want to align multiple variables and determine whether these are simple speculations or factual issues.

### d. Definition and Explanation of Key Terminology

**AI -** intelligent systems capable of mimicking cognitive functions of the human brain.

**DL -** subfield of ML concerned algorithms inspired by biological neurons and their functions.

**Radiology -** field of medicine that leverages imaging technologies to diagnose and cure diseases.

**Model and Algorithm** - used interchangeably in this paper and refer to the entire AI-based system. They include the architecture, learning algorithms, procedures, and training and testing data.

**ANN** – a collection of connected artificial neurons that is able to learn complex relationships.

# Chapter 2: Hypothesis

## a. Brief Overview of Theoretical Foundations Utilized in the Research Study

Despite being a game-changer in multiple industries, Artificial Intelligence is still not a popular reality in medical imaging. For most hospitals and imaging centers, the adoption of intelligent systems has been relatively slow and problematic. Although some performant algorithms have been proposed in the literature, radiologists tend to be reluctant to adopt experimental technologies and prefer the old-fashioned approach to investigate, diagnose, treat, and monitor their patients. In this medical sector, AI faces some recurrent problems caused by the lack of data, laws, standardization, insufficient training programs, and the absence of collaborative efforts between physicians, researchers, and computer scientists. While these are remarkable barriers that prevent the industry from advancing, they are not necessarily indicators of an unsuitable technological solution but rather signs that highlight an emerging field. Given the intrinsic value that medical imaging brings to healthcare, one would typically expect an abundance of AI-based tools designed to streamline the radiology workflow. However, there is still a long way to go before affirming that AI is a substantial component that is actively used in medical imaging.

### The importance of medical imaging in medicine

In the clinical context, medical imaging is generally regarded as identical to radiology ("Medical Imaging", 2021). Under the broad umbrella of medicine, radiology is one of the most popular topics nowadays because it represents the primary diagnostic tool used by specialists to fight against a broad spectrum of medical conditions (Goraczkowski, 2019). Radiology involves using medical imaging techniques to scan specific body areas to identify if any internal problems need to be addressed. The scans offer valuable information about the inside of the human body and increase the ability to diagnose and treat multiple diseases accurately. Because of the sophisticated nature of the human body, doctors cannot manage patients without having detailed insights about structural or disease-related modifications. Therefore, unaccompanied by radiology, professionals would have a hard time analyzing complex illnesses, especially those that do not involve any distinctive external symptoms.

Family doctors and other physicians rely on imaging exams to determine the correct diagnosis and appropriate course of action on multiple occasions. In the particular case of healthy patients who do not suffer from any known preconditions, medical imaging plays an essential role in determining which part of the body is responsible for sudden abnormal behavior. Even though this kind of medical investigation does not always indicate the precise origin of the problem, it is usually able to provide at least a starting point that specialty doctors further explore. After identifying the root cause of the issues and establishing a treatment plan, radiology can be extremely useful in monitoring the disease's progression. Because each patient is unique and standard therapeutic procedures might not be as effective for everyone, radiology can be leveraged to observe whether the chosen treatment plan is efficient and to what extent. By performing subsequent scans of the body part that is not functioning normally, physicians can assess the evolution of the disease and take corrective actions promptly – which is of paramount importance in time-sensitive medical conditions, where fast treatment can make a big difference.

Naturally, doctors who specialize in interpreting the results of these scans are called radiologists. To become a radiologist, students need to earn a bachelor's degree, attend medical school, and complete a medical residency before applying the skills and techniques they have learned. Some radiologists might opt to pursue an optional fellowship in a specialization of radiology, and the entire journey to becoming a licensed professional generally takes thirteen to fifteen years after high school graduation (Global Pre-meds, 2014).

As of 2019, approximately 250,000 radiologists were performing diagnostic imaging examinations on patients in the United States alone (U.S. Bureau of Labor Statistics, 2021). As the share of the population aged 65 or over continues to increase (United Nations, 2019), there will be an increase in medical personnel specialized in radiology. The U.S. Bureau of Labor Statistics indicates that from 2019 to 2029, the number of radiologists is projected to grow 7 percent, which is faster than the average for all professions (U.S. Bureau of Labor Statistics, 2021). According to a study conducted in 2019, 90% of the surveyed radiologists reported their workload had increased significantly over the last three years, and 28% of them indicated that it had increased by more than 20% (Alexander, 2020). Given the intensifying work pressure, radiologists started adopting AI-based tools to automate some of their tasks and compensate for the surging demand.

However, this approach involves a certain amount of training. Radiologists cannot make use of these tools immediately after graduation because AI is not systematically integrated into their curriculum (Collado-Mesa, 2018). Therefore, before working with or developing AI applications, they need to enroll in training programs that provide different types of knowledge, from awareness to basic concepts and hands-on experience. Unfortunately, studies have shown that despite having many AI training programs available, most of them (80%) are stand-alone sessions and are not part of a longer-term learning path (Collado-Mesa, 2018), so the industry does not benefit from having a solid educational foundation.

## Deep Learning in Radiology

In radiology, AI encompasses a wide range of tools that can be virtually used in all steps of treatment, from initial detection to follow-up procedures (Francolini, 2020). Some of its applications, such as AI-powered CRM systems and automated data capture technologies, can even be used to manage the businesses themselves. Therefore, due to its broad applicability and variety of algorithms, concepts, and data it employs, it is challenging to develop structured knowledge in this field. Consequently, we will concentrate exclusively on a subfield of AI and analyze Deep Learning and its ability to solve classification, detection, prediction, and segmentation problems.

Before diving into explaining how Deep Learning actually works in radiology, we need to clarify several theoretical aspects about digital images. More specifically, we need to understand how computers represent images. A digital image is a collection of picture elements, commonly known as pixels, assembled in a grid system. When we talk about the width and height of an image, we are referring to the dimensions of its grid representation. Each pixel represents a color and can be uniquely identified in a picture using a (row, column) pair of coordinates. Depending on the type of image they compose (color, grayscale, binary), pixels store their associated colors using different encoding schemes. However, in medical imaging, we are predominately dealing with grayscale pictures, so we will only discuss this particular model. For grayscale images, each pixel holds an integer value from 0 to 255 (inclusive at both ends), representing an amount of light. The lower bound symbolizes the total absence of light and corresponds to “black,” while the maximum value represents the total presence of light and is associated with “white.” Any fractional value that falls in between these values represents a shade of gray.

In the picture below, the area delimited by the green square has been zoomed so we can observe how a picture looks at pixel level. We can see how the picture elements are organized in a matrix (or a two-dimensional array), and we can understand how each pixel is uniquely located using a pair of coordinates. We can notice several grayscale values and the corresponding amounts of light they represent on the right-hand side. Therefore, computers store digital images as matrices of numerical values. When they are rendered for user display, these values are converted to amounts of light, and users get to see the actual digital images on their screens.

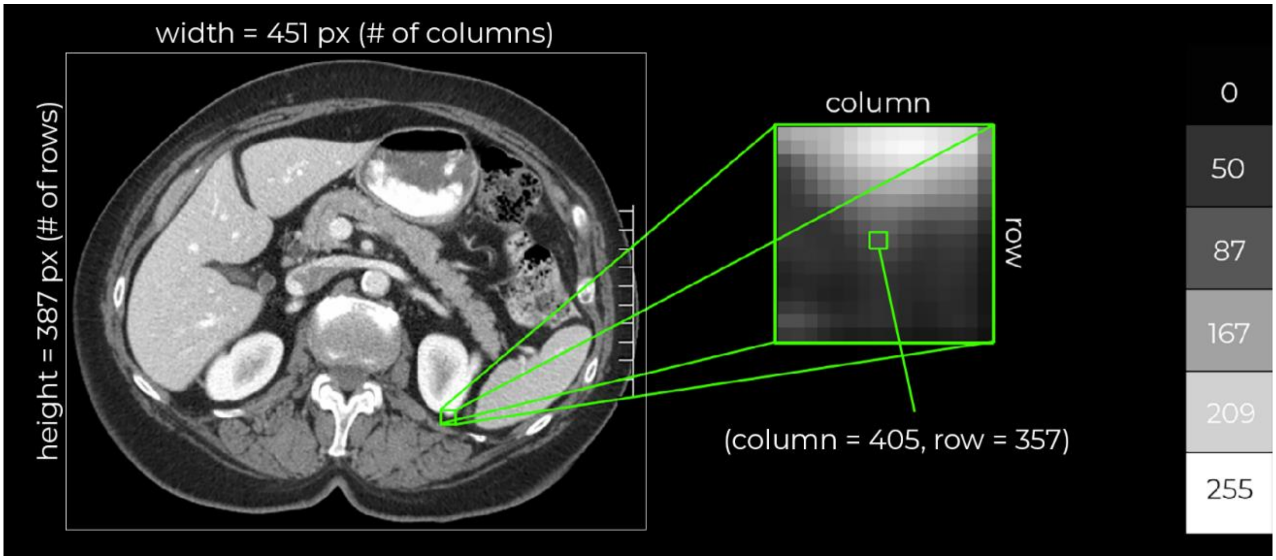


*Figure 2 – The representation of a digital image*

Artificial Neuronal Networks represent the very core of Deep Learning. They consist of layers of interconnected functional units called nodes. These nodes are inspired by the biological neurons we find in our brains and have the responsibility of performing basic calculations – they take an input, apply some mathematical formulas to it, and produce an output that is forwarded to the nodes located in the next layer. *Figure 3* shows a brief overview of what happens in a neuronal node. Each input is correlated with a certain weight, and then the weighted sum of the inputs is plugged in a predefined mathematical formula to compute the final output.

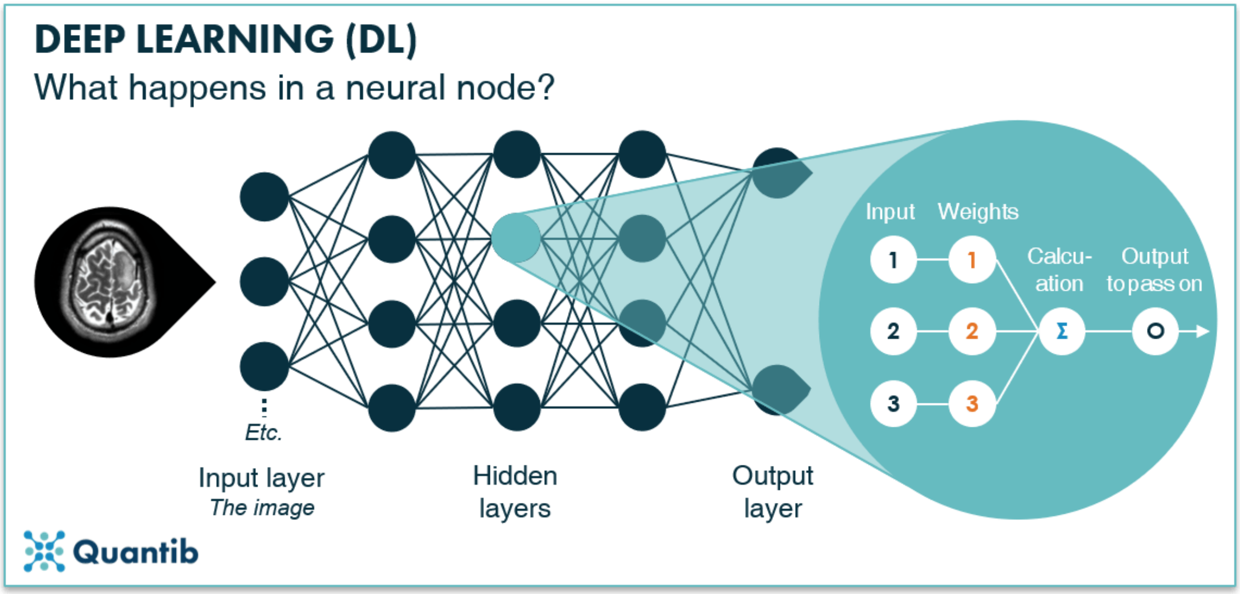


*Figure 3 - How do neuronal nodes work? - source: quantib.com/the-ultimate-guide-to-ai-in-radiology*

The architecture of an ANN includes an input layer, followed by a hidden layer and an output layer. The input layer consists of external data belonging to an image used for training or testing. For example, if we wanted to train our model using the image in *Figure 2*, we would need an input layer consisting of 174,537 (451*387) neurons. In this case, each of the neurons would be initialized to the grayscale value of a pixel (like a bijective function). The next layer is the hidden layer which is responsible for extracting features from the image. Lastly, there is the output layer, whose purpose is to answer the question of interest such as: *is the tumor in the image an oligodendroglioma or an astrocytoma?* In general, images contain multiple features, and a single hidden layer is not able to extract all of them. Thus, we normally utilize many hidden layers – hence ANNs become Deep Neuronal Networks. *Figure 4* illustrates the standard architecture for a classification problem. We feed a radiography containing a tumor into a trained network and we ask the system to tell us which kind of tumor is present in the scan. We can design the model to answer with a categorical result such as 0 or 1 (since in our case there are only two possibilities) or we can employ probabilities as part of the prediction (30% oligodendroglioma, 70% astrocytoma).

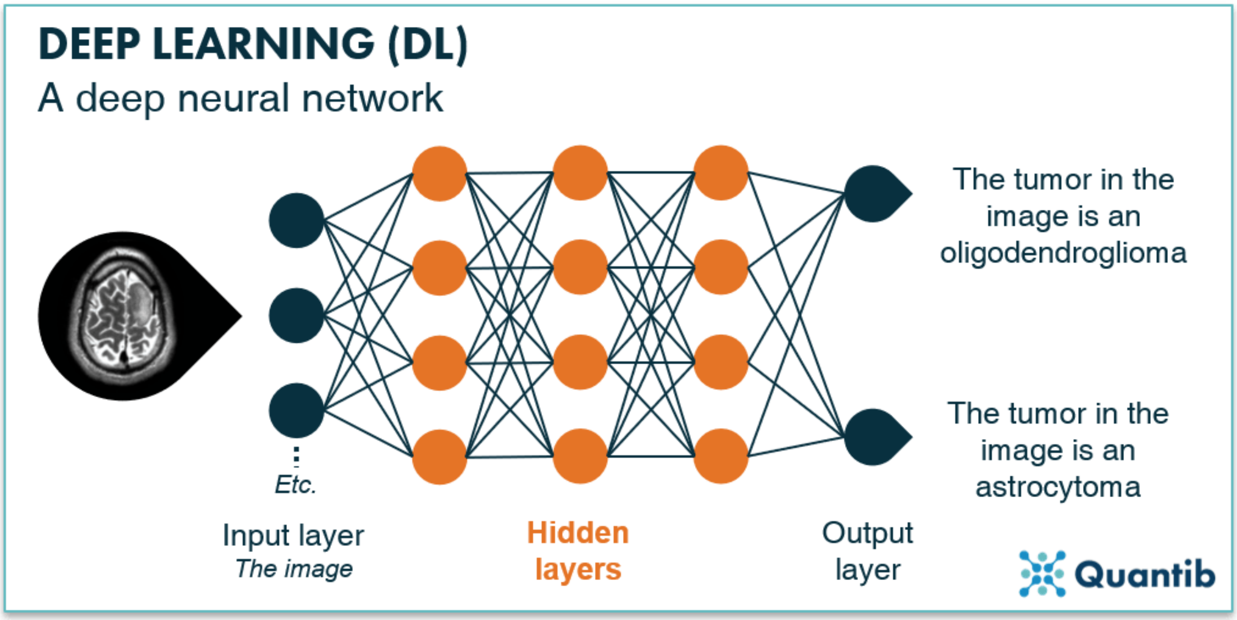


*Figure 4 - Architecture of a DNN - source: quantib.com/the-ultimate-guide-to-ai-in-radiology*

After selecting and implementing an architecture, developers choose a data set containing images which are meaningful to the problem they are trying to solve. In our case, we would select a collection of, let us say, 10000 brain scans and label them accordingly. All images that correspond to oligodendroglioma would be labeled with 0, and all images containing an astrocytoma would be labeled with 1. If we had other categories, they would normally receive subsequent values 2, 3, 4, and so on. After labeling the images, they are split in three different sets training, validation, and test, based on arbitrary split ratios such as 70%, 15% and 15%. For the majority of algorithms, the number of training images is considerably larger compared to the other two groups.

The first set of 7000 images would be fed into the network for training purposes. Technically speaking, the training process is equivalent to determining the appropriate weights of the nodes so that their output is optimal, and the overall system produces the intended results. During the training phase, the validation set is frequently used to fine-tune the model hyperparameters. Even though the model occasionally "sees" the validation dataset, it never actually "learns" from it.

After training, the test dataset is used to provide an unbiased evaluation of the model. Since this set contains data that is unknown to the model, it provides an accurate measurement of how well the algorithm performs on new images that have not been included in the training set. In other words, we use the test set to evaluate the algorithm's capability to generalize what it has learned. This measurement is further used to improve the model by either providing more training data, adjusting the architecture, tweaking parameters, and so on. By using the performance on the test set, developers get a sense of how reliable their model is when it comes to actual patient scans. Because those real scans are also unknown to the model, they are very similar to test images.

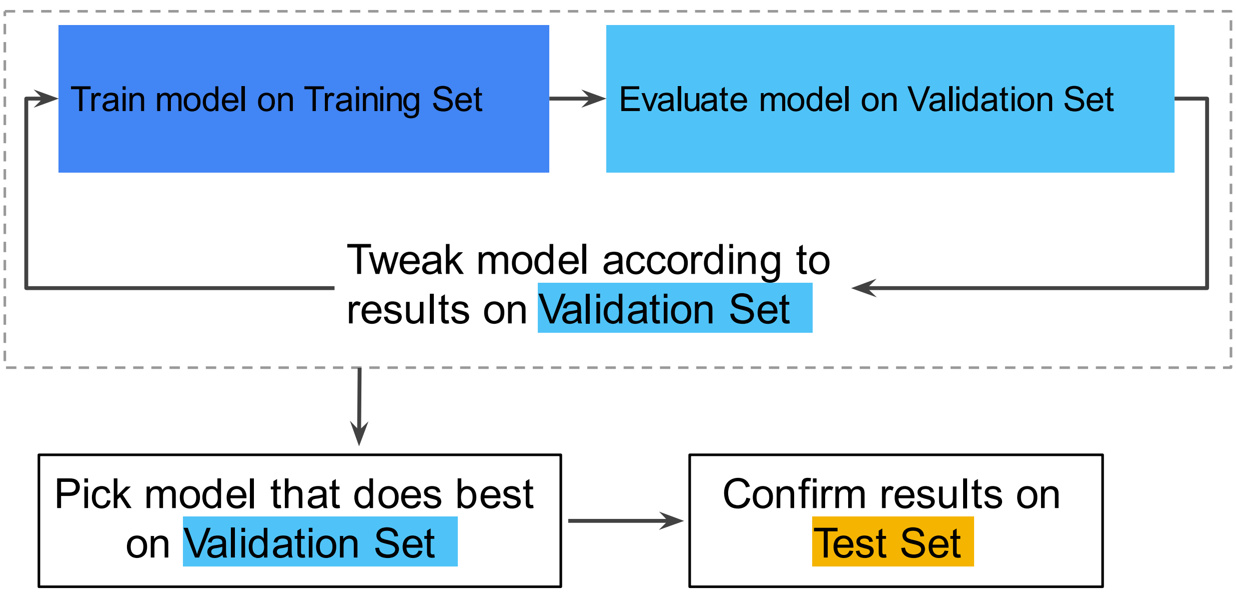


*Figure 5 - Workflow for a DNN model - source: developers.google.com/machine-learning/crash-course/validation/another-partition*

### b. Brief Overview of Literature Reviewed, Discussed and Applied

Most of the consulted literature includes research papers published by reputable specialists who propose different deep learning architectures and evaluate their performances on predefined data sets that the authors thoroughly explain. By their very nature, these papers are incredibly technical and include various advanced concepts from multiple disciplines such as mathematics, computer science, and medicine. Because these papers represent the ongoing efforts of pioneers in our area of interest, they are usually hard to assimilate fully. Therefore, we developed a systematic way of analyzing their work, results, and implications. First, we identify the problem they are trying to solve, establish the architecture of their model, and then determine the relative performance compared to human subjects.

## Chapter 3: Methods

### a. Study Method and Study Design

This thesis is secondary research intended to synthesize the available resources regarding the state of AI in medical imaging. To provide a holistic view of the subject matter, it employs a combination of quantitative data (reports, surveys with close-ended questions, and numerical observations) and qualitative data (podcasts, interviews, blog posts, articles, and literature review).

We use the quantitative data to assess the adoption level of AI in medical imaging and the objective performance of two state-of-the-art AI-based models. We analyze statistical information provided by research institutions, technology companies, and government agencies to determine the adoption level. Most of this data has been collected through reports or surveys and offers an impartial perspective on how many organizations (hospitals or imaging centers) utilize intelligent systems in patient care imaging. The performance of state-of-the-art models is determined by examining published research papers that describe the theoretical results of some specific models. Because many models have been developed to solve isolated problems, we arbitrarily consider only two of them. There is a natural correlation between a specific adoption rate and the performance of the available algorithms, so we want to explore this relationship.

It is essential to understand the technical characteristics and limitations of the systems specialists can take advantage of. Without reliable solutions, radiologists are susceptible to refuse the emerging technologies. However, regardless of their capabilities, AI systems can be adopted or rejected by professionals based on multiple other factors that are more subjective, such as personal beliefs or unpleasant past experiences with computers. Therefore, we also employ a qualitative approach to investigate the problem even further and analyze the opinion of different radiologists who are at the forefront of innovation in medicine. Their hands-on experience combined with their theoretical knowledge can offer valuable insights into the usage of AI in medical imaging.

### b. Sample Used in the Study

The sample population we analyze consists of computer scientists, radiologists, and professionals with vast experience in both areas of interest (medical imaging and computer science). This selection was dictated by the technical nature of the topic we are exploring. It is essential to analyze the numbers and opinions provided by those at the forefront of research in this field to draw reliable conclusions. While we sporadically consider the impressions coming from untrained authors for a broader perspective, we exclusively base our reasoning and arguments on data coming from experts.

### c. Explanation of Measurements

As previously mentioned at the beginning of the chapter, this research employs two distinct types of data: quantitative and qualitative. While the first category is statistical, rigid, well-defined, and can be measured using numbers or values, the second category is non-statistical, unstructured, and represents opinions, beliefs, or feeling that cannot be represented using numerical values. This fundamental difference requires two sets of measurements, definitions, and rules.

**For quantitative data:**

This type of data is used to evaluate the adoption level and theoretical performance of AI-based systems. The first metric refers to the rate of adoption of intelligent tools and describes how many institutions avail themselves of some sort of AI. It is computed using a simple formula:

$$rate\ of\ adoption = \frac{number\ of\ organizations\ that\ use\ AI}{number\ of\ surveyed\ organizations} * 100$$

The second metric is more complex and aims to provide an aggregate measure of performance for deep learning models. It is called Area Under the ROC Curve (AUC) and represents the area underneath the entire ROC curve. A ROC curve (receiver operating characteristic curve) is a visual representation that shows the performance of a classification model by plotting two rates: True Positive Rate (TPR) $TPR = \frac{TP}{TP+FN}$ and False Positive Rate (FPR) $FPR = \frac{FP}{FP+TN}$. TPR, also known as sensitivity, recall, or hit rate, is used to measure the percentage of actual positives that have been correctly identified. FPR, also known as fall-out, is used to measure proportion of all negatives that yield positive test outcomes. In this context, we define TP – the number of total positives, TN – the number of total negatives, FN – the number of false negatives, and FP – the number of false positives.

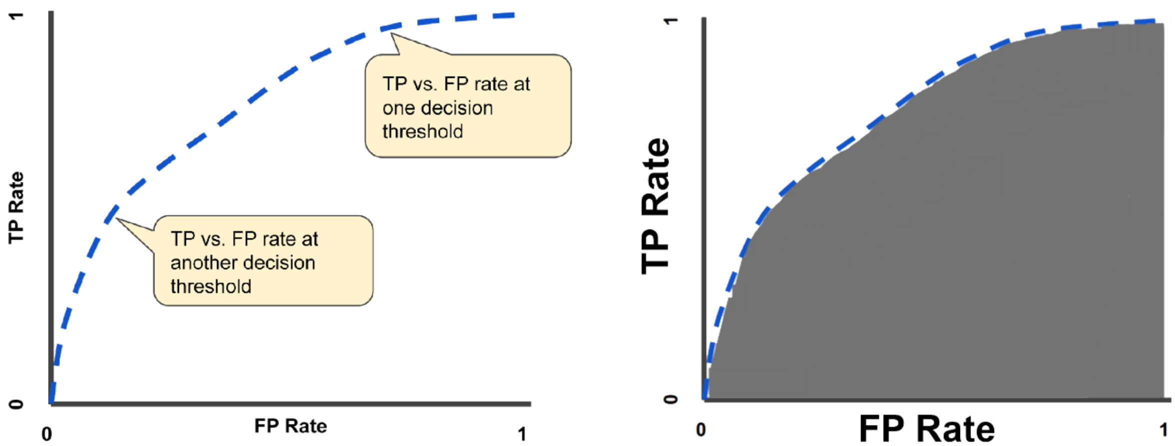


*Figure 6 - ROC Curve – source: developers.google.com/machine-learning/crash-course/classification/roc-and-auc*

Explaining precisely how each of these numbers is calculated or why AUC is a generally accepted measure of performance is beyond this paper's scope, but the main idea we are trying to convey is that we are interested in AUC because it helps us understand how well an algorithm is performing. As we can see in the graph on the right, AUC ranges in value from 0.0 to 1.0. For a model whose predictions are 100% right, the corresponding AUC is 1.0, while for a model whose predictions are 100% wrong, the associated AUC is 0.0. In practice, it is almost impossible to get an AUC of 1, so we usually strive for values that are close to 1. Depending on the problem we are trying to solve, even smaller areas can indicate a useful model – in fact, every model with an AUC above 0.5 can be considered better than a random guess.

**For qualitative data:**

This type of data is used to describe the characteristics of the problem we are investigating, and therefore it cannot be measured with numerical values. It mainly consists of opinions, beliefs, and personal impressions belonging to the people who have been interviewed in the papers or articles we analyze. To some extent, we can categorize this information into two groups: positive and negative attitudes towards AI in medical imaging, but it is conceptually impossible to measure our data precisely because most of it comes from unstructured interviews, surveys, podcasts, or blogs.

We consider our study method and study design both valid and reliable, primarily because they employ data that any other researcher can verify. While our quantitative conclusions can be tested using other analytical tools, the qualitative results rely on personal beliefs and cannot be labeled right or wrong. Their sole purpose is to convey to our audience the thoughts and ideas that come from experts.

### d. Description and Justification of Analytical Techniques Applied

Since we are conducting secondary research and do not have any prior insights about the topic and no information has been collected through primary research, we want to establish a reliable benchmark using statistical techniques. The general idea is to get a sense of what is available and how many individuals or institutions are taking advantage of it. We acknowledge that this approach provides some valuable details that are not necessarily comprehensive. For instance, imagine we find out that 90% of imaging centers and hospitals in the US integrated AI solutions in their practice, but they primarily use them on a small scale to automate one or two tasks. Based on this analytical technique alone, we can affirm that AI is widespread and therefore benefits from a high status. However, it is essential to understand how much of the AI's potential is being exploited by those adopters before drawing the final conclusions. Therefore, we also perform an analysis on specialists' opinions to determine if there is room for expansion or if a peak has been reached.

### e. Assumptions and Implied Limitations of Study Method and Design

As secondary research, this paper uses data collected by other people to explore the topic and formulate conclusions. Thus, our study method and design involve inherited assumptions and limitations. First of all, we assume that all the results published by other authors are reliable and verifiable, and we presume the sample population they selected during their research is a representative subset of the larger population of professionals they analyze. Moreover, we assume there are no hidden biases or financial interests behind the authors' findings and take for granted the honesty and integrity of the surveyed individuals. In terms of limitations, our data was gathered by the researchers in a given context to answer questions that are slightly different from ours. By decontextualizing their data, we might lose some of its value or significance.

Because of time and knowledge constraints, we focus on two of the top-performing models. This approach does not ensure that our findings are generally applicable to all algorithms on the market. However, the main goal is to scratch the surface and observe whether some of the current AI-based algorithms can achieve performance levels that are at least as good as those attained by human subjects. Additionally, a significant limitation of our methods is the tremendous amount of knowledge we needed in order to fully understand the mathematical formulas and proofs described in technical documents. We concentrated on identifying the context in which a particular algorithm would be useful but depended entirely on the author's ability to prove its performance.

# Chapter 4: Findings

## a. Brief Overview

As a result of this academic endeavor, AI in medical imaging can be seen as an emerging field that is slowly but surely becoming an essential part of our lives. We identified several distinct groups (radiologists, computer/data scientists, and researchers) who constantly strive to improve the existing solutions and take AI to the next level. However, it seems like the sector as a whole is highly conservative, and each group takes incremental steps in a familiar direction. Radiologists acknowledge the importance of AI but prefer to use traditional methods, scientists value the theoretical performance of their models more than their clinical applicability, and researchers try to bridge the gap between the first two categories.

## b. Results of the Method of Study

The results of our research project indicate that AI in medical imaging is still underdeveloped. While state-of-the-art models' adoption rate and performance give the impression of a relatively well-established industry, the specialists' opinions prove this is not necessarily the case. Multiple barriers hinder a wider adoption and a coherent development of more advanced algorithms. Based on the two types of research data we analyzed (quantitative and qualitative), we can classify our findings (in arbitrary order) into two categories:

**Quantitative analysis:**

1. Based on the most recent data, half of the hospitals and imaging centers are currently using some sort of AI to enhance their imaging capabilities.
2. Performant AI-based algorithms are currently available on the market, but they have not been thoroughly tested in clinical settings.

3. There are Deep Learning models that outperform human subjects, but they are usually optimized to address only isolated tasks and cannot be integrated with other similar tools.
4. A very small percentage or radiologists describe themselves as “very familiar” with AI.

**Qualitative analysis:**

1. The radiologists themselves are the main barrier that prevent a wider adoption.
2. There are collaboration issues between radiologists and algorithm developers.
3. There is a lack of publicly available image repositories which prevents a systematic collection of data.
4. U.S. Food and Drug Administration (FDA) plays a major role in a wider adoption of AI.

## Chapter 5: Discussion

### a. Discussion of Findings and Implications

#### The adoption rate of AI in Medical Imaging

According to a study conducted in 2019 by Definitive Healthcare, a technology company that provides intelligent solutions to help customers grow their businesses, nearly one-third of the surveyed organizations employed AI, Machine Learning, or Deep Learning to improve their imaging or business activities (Definitive Healthcare, 2021). The study analyzed 207 institutions (including 72 imaging centers and 135 hospitals) and revealed a slight difference in the adoption rate between the two facility types – 34.7% for imaging centers and 31.9% for hospitals. The researchers argue that this natural difference is due to distinctions in the primary objectives the two types of organizations strive to achieve. While hospitals offer a broad range of services aimed at helping patients with various diseases, imaging centers are highly specialized in diagnostic imaging. Therefore, they are more flexible and willing to integrate new technologies.

Almost one-third of the responding institutions that were not using AI declared they would be utilizing intelligent technologies within the next two years. Again, the trend was consistent, and imaging centers showed a stronger inclination for a wider adoption than their hospital counterparts. Both facility types identified cost as the most significant barrier, followed by “lack of IT infrastructure and technical expertise” in the case of medical centers, and “lack of strategic direction” in the case of hospitals (Definitive Healthcare, 2021).

Using inductive reasoning and considering that this study was carried out about two years ago, we can combine the findings (the adoption percentage at the time of the survey + the forecasted adoption percentage in the next two years) and conclude that the current adoption rate is around 53%. In other words, almost half of the hospitals and imaging centers nowadays are using AI to enhance their imaging or other business activities. However, in the absence of more recent data, this adoption rate is more of an informed guess rather than an actual statistic. Especially if we consider the impact of the COVID-19 pandemic and the fact that hospitals had to rearrange their priorities to take care of an unprecedented number of patients, we should not be surprised to see future research demonstrating a lower adoption rate.

**State-of-the-art Models**

The literature we examined proposed algorithms optimized for either classification, detection, prediction, or segmentation of medical images. Generally, these algorithms are trained to solve isolated tasks and are not one-size-fits-all solutions. Some of the most notable discoveries are presented in the following paragraphs.

A group of Stanford researchers developed a deep-learning algorithm (Armitage, 2018) called CheXNet, which is able to interpret chest X-ray images and detect 14 different pathologies.

This algorithm can take as input any chest scan and tell, with a certain confidence level, which diseases are present in that particular X-ray. For each disease, the system can also highlight those parts of the image that serve as indicators for determining one pathology or the other, which is extremely helpful for doctors. They have the ability to take a closer look at the areas indicated by the software and explore why the algorithm made certain particular predictions rather than observe the final list of identified pathologies.

The group of researchers has thoroughly tested their algorithm, and the results are promising. They used several radiologists from around the U.S. and compared the performance of their algorithm with the performance of the human subjects. Some of the most important metrics they considered include accuracy, sensitivity, specificity, and speed. For 10 diseases, the model was able to perform as well as radiologist; for three it underperformed compared with humans; and for one, the solution outdid the professionals. However, it is essential to realize that one of the significant differences in performance between the computer and experts was speed. On average, radiologists needed around four hours to complete their analysis, while CheXNet could interpret the same image in less than two minutes (Armitage, 2018).

On top of this impressive performance, the team was able to deploy the solution on a cloud infrastructure and created a mobile application that enables average phone users to analyze scans by simply taking a picture of an actual X-ray and uploading it to the cloud. While this might not sound like a big deal, it is, in fact, an extremely useful feature for many under-developed countries in this world that suffer from an acute shortage of radiologists.

Another notable model (Wu JT, 2020) we have investigated was created by IBM Research in collaboration with the University of Southern California. Using a novel Deep Learning architecture, the group built a model to analyze anteroposterior (AP) frontal chest radiographs and trained it to identify a total of 72 findings. The model was trained using 342126 frontal chest images collected in urgent care settings, and its performance was judged using the conventional AUC metric – the mean AUC across labels was 0.8. For testing, the team used 1998 images whose ground truth was determined through a triple consensus by a team of experienced radiologists. The study results indicate that the model was able to detect the findings with the same level of sensitivity as the human subjects (5 third-year radiology residents). However, the algorithm performed better than radiologists when it comes to the specificity and positive predictive value (the fraction of misses and overcalls). Above all, this research proved that AI-based systems can provide reliable interpretations of chest radiographs in clinical workflows and highlighted the AI's potential to improve accuracy, reduce costs, and save time (Wu JT, 2020).

**What do radiologist think about AI?**

When it comes to understanding what radiologists think about AI in medical imaging, we identified a paradoxical situation. Based on the surveys we analyzed, a large majority of the radiologist believes that AI is either essential or extremely important and just a tiny portion of them say they are yet to be convinced of its potential (Alexander, 2020). However, those who are currently using AI have reported very little adoption. Surprisingly, the most significant barrier seems to be the radiologists themselves. Contrary to popular belief, they do not fear AI's potential to replace them, but they do question the current diagnostic capability of intelligent tools. Especially in the case of complex diseases, they consider that AI is not ready yet to tackle intricate problems, so they prefer traditional methods that are, in their opinion, safer and more consistent.

56% of the responders surveyed in a study (Alexander, 2020) confirmed that they use AI, but only half of them reported exposure to the most popular five use cases. These numbers and the author's comments indicate that AI has not been adopted at scale in this subfield of medicine despite the undeniable interest in intelligent solutions. AI's capabilities are still limited by specific conditions, imaging techniques, or disease states, and professionals express skepticism for innovation.

### A Collaboration Issue

A frequent problem reported by many specialists is the lack of collaborative efforts between radiologists and computer programmers. It seems like most of the AI-based solutions on the market were developed in a vacuum, without the active involvement of medical professionals (Kent, 2020). It goes without saying that programmers need to build algorithms while keeping their clinical purpose in mind to create systems that can seemingly integrate into radiology workflows. One group needs to be able to speak the other side's language so the clinical needs match technological capabilities. Some efforts have been made in this direction. Recently, Brigham and Women's Hospital created a data science path for their fourth-year radiology residents. The program teaches students advanced analytical tools and gives them the opportunity to get involved in every stage of algorithm development (Kent, 2020).

A strong collaboration between the two groups can also prepare radiologists for problems that can appear in real-word scenarios. Having a better understanding of how these tools work can enhance clinicians' ability to manage unexpected errors. In the long run, knowledgeable clinicians can become the principal source of feedback and play an active role in monitoring and updating deployed systems (Kent, 2020).

A concrete example that illustrates the implications of a poor communication between radiologists and computer scientists is a personal story that we found during our research. It was shared by a reputable specialist in a podcast (Dania, 2020) hosted by the Radiological Society of

North America. The expert described that he and his team were working on developing an algorithm to distinguish between several pathologies. To train that model, the radiologists in his group collected a set of medical images from a nearby hospital and sent them to the computer scientists in the same team, who managed to build an extremely accurate model. After taking a closer look at the system, the team realized that the incredible performance was in fact an error. Instead of learning useful features about the images, the model learned the name of the hospital from which the images originated because it was written in the bottom part of each picture. Since all the pictures that came from that hospital had the same label, the system was able to classify them with a perfect accuracy. If there was a better communication between radiologists and computer scientists, this issue could have been avoided right away as it is a rookie mistake. Any computer scientist who works with AI knows that training data needs to be flawless because any imperfection can be involuntarily learned by the algorithm.

**A Lack of Data**

Deep Learning models are known for the tremendous amounts of data they require during the training phase. Given that the volume of data is practically exploding these days, most of AI systems have enough information to train and operate within optimal parameters. Unfortunately, in the particular case of medical imaging, the lack of data (images), is still a fundamental problem (Sewell, 2019). The entire process to collect and preprocess radiographs and other types of medical images is extremely costly. On top of being created by expensive equipment, medical images also need to be manually analyzed by radiologists for labeling purposes. There are no publicly available repositories where institutions can share their data and individual institutions are generally unable to produce in-house the necessary amount of data to train a model efficiently. Even if they had those capabilities, there would be another problem regarding the variability of that data. If only one organization creates large datasets and trains different models, the resulting systems would be extremely biased and dependent on characteristics that describe the patients who visit that particular organization. As long as we do not have a coordination between hospitals, research units, government, and technology companies, we cannot create large repositories of data and much of AI's potential remains unreachable (Sewell, 2019).

**A Lack of Regulations**

Radiologists consider the lack of regulatory approval a major barrier to the broader adoption of AI (Alexander, 2020). As an emerging field, AI represents a new reality that must be understood before it can be regulated. It seems like the U.S. Food and Drug Administration (FDA), the government body responsible for managing AI technologies in healthcare, has failed to keep up with the rapid technological advancements. Before any medical software or hardware enters the U.S. market, the parent company needs to submit elaborated documentation to FDA for evaluation and approval, a process that is currently both time-consuming and ambiguous. The main problem of AI/ML algorithms is that they fall under the category of adaptive algorithms, which is not currently regulated (Benjamens, 2020). Because AI employs algorithms that are constantly changing based on the new data they learn, FDA realized that it is impossible to monitor intelligent systems. They proposed "a total product lifecycle-based regulatory framework" to overcome this problem, but no concrete solutions have been adopted in that direction. Some reports suggest that the agency has repeatedly failed to provide clear guidelines, so the current approval process is unclear. Without the active support of the regulatory agency, developers have a hard time monetizing their products since they cannot enter the market, while radiologists become more reluctant to embrace new tools because they do not have any assurance that the models they use are up to standard (Benjamens, 2020).

## Chapter 6: Conclusions

We believe the most important takeaway from our paper is the lack of collaboration between radiologists and algorithm designers. We consider it a central problem that prevents AI from advancing to the next level. Currently, the industry is evolving and heading in the right direction, but the steps are incremental, and as a result, many patients do not benefit from what we currently have. Some people might even argue that patients are losing their lives because things are not progressing faster. We find that attitude too extreme but agree on the idea that things should be accelerated.

The decent adoption rate of AI combined with the reasonable performance of state-of-the-art models prove that humans understand the tremendous potential AI can bring to the table. However, radiologists, a target audience for AI, are still skeptical about the current capabilities of intelligent solutions. While we understand their legitimate concerns, we reckon this attitude is faulty. There is no need for AI systems to be extremely capable or reliable as long as we do not empower them to make the ultimate decision. The critical point is that there should be a clear separation of diagnosis and prediction from treatment and curative procedures. AI does not aim to replace radiologists because their job involves more than image analysis, and computers are not empathic nor astute. Therefore, intelligent systems should have a consultative role and provide a second opinion just because it might be helpful in some situations. For this reason, we believe AI-based tools should be far more popular despite their limitations. They are not perfect, but it does not mean they are not helpful.

Additionally, all the barriers we have identified can serve as excellent starting points for further research. It would be interesting to explore in great detail how different solutions can bridge the gap between radiologists and algorithm designers. In this context, big tech companies can play an essential role as they have the financial resources to create homogenous teams of professionals. Another aspect that can be considered is the necessity of adjusting the current academic curriculum for either computer scientists or radiologists so that each side is better at understanding the other side’s language.

By and large, we believe that AI in medical imaging is on the right track, but several issues need to be addressed as soon as possible to ensure optimal technological development. Radiologists need to embrace AI solutions and aid their development, while computer scientists need to pay more attention to clinical needs and develop algorithms that can be seemingly integrated into the radiology workflows.

*“The advance of technology is based on making it fit in so that you don't really even notice it, so it's part of everyday life.”* Bill Gates, Co-founder of Microsoft

## Chapter 7: Bibliography

Alexander A, Jiang A, Ferreira C, Zurkiya D. An Intelligent Future for Medical Imaging: A Market Outlook on Artificial Intelligence for Medical Imaging. J Am Coll Radiol. 2020 Jan;17(1 Pt B):165-170. doi: 10.1016/j.jacr.2019.07.019. PMID: 31918875.

Armitage, H. (2018, November 20). *Artificial intelligence rivals radiologists in screening X-rays for certain diseases*. News Center. https://med.stanford.edu/news/all-news/2018/11/ai-outperformed-radiologists-in-screening-x-rays-for-certain-diseases.html.

Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. Npj Digital Medicine, 3(1). https://doi.org/10.1038/s41746-020-00324-0

Bresnick, J. (2018, November 5). *Top 5 Use Cases for Artificial Intelligence in Medical Imaging*. HealthITAnalytics. https://healthitanalytics.com/news/top-5-use-cases-for-artificial-intelligence-in-medical-imaging.

Collado-Mesa, F., Alvarez, E., & Arheart, K. (2018). The Role of Artificial Intelligence in Diagnostic Radiology: A Survey at a Single Radiology Residency Training Program. Journal of the American College of Radiology, 15(12), 1753–1757. https://doi.org/10.1016/j.jacr.2017.12.021

Dania, D. (Host). (2020–present). Radiology AI Podcasts | RSNA [Video podcast]. Spotify. https://open.spotify.com/show/2iVoRSFnUPXAEBTjsOC3pr

Davis, J. (2019, September 24). *NASA to Build New Asteroid-Hunting Space Telescope*. The Planetary Society. https://www.planetary.org/articles/nasa-to-build-asteroid-telescope#:~:text=The%20yet%2Dto%2Dbe%2D,the%20Earth%20within%2010%20years.

Definitive Healthcare, D. H. (2021, April 8). *The Future of the AI Market: 2019 Study Results*. Definitive Healthcare Blog. https://blog.definitivehc.com/2019-artificial-intelligence-study.

Francolini, G., Desideri, I., Stocchi, G. et al. Artificial Intelligence in radiotherapy: state of the art and future directions. Med Oncol 37, 50 (2020). https://doi.org/10.1007/s12032-020-01374-w

Gelsinger, P. (2017, November 8). *This is how tech can help us all survive a modern-day apocalypse*. Vox. https://www.vox.com/2017/11/8/16621066/tech-benefit-global-planet-innovation-websummit-health-climate-poverty-extremism.

Global Pre-Meds. (2014, May 20). *Exploring Careers in Radiology: Global Pre*. Meds. https://www.globalpremeds.com/2014/05/20/exploring-careers-in-radiology/.

Goraczkowski, S. (2019, October 17). *The Importance of Radiology*. Healthcare Staffing Agency Blog. https://blog.fusionmedstaff.com/the-importance-of-radiology#:~:text=Radiology%20is%20the%20main%20diagnostic%20tool%20for%20many,have%20made%20diagnosing%20more%20effective%20more%20than%20ever.

Heath, N. (2020, December 11). *What is AI? Everything you need to know about Artificial Intelligence*. ZDNet. https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/.

Heaven, W. D. (2020, December 10). *We're not ready for AI, says the winner of a new $1m AI prize*. MIT Technology Review. https://www.technologyreview.com/2020/09/23/1008757/interview-winner-million-dollar-ai-prize-cancer-healthcare-regulation/.

Kent, J. (2020, November 23). *Top Challenges of Applying Artificial Intelligence to Medical Imaging*. HealthITAnalytics. https://healthitanalytics.com/features/top-challenges-of-applying-artificial-intelligence-to-medical-imaging.

Marsch, L. A., & Gustafson, D. H. (2013). The Role of Technology in Health Care Innovation: A Commentary. Journal of Dual Diagnosis, 9(1), 101–103. https://doi.org/10.1080/15504263.2012.750105

Medical Imaging (2021, June 1). In Wikipedia. https://en.wikipedia.org/wiki/Medical_imaging

Sewell, J. (2019, May 10). *Four challenges in developing AI algorithms for medical imaging*. Global Engage. https://www.global-engage.com/life-science/four-challenges-in-developing-ai-algorithms-for-medical-imaging/.

Shuaib, A., Arian, H., & Shuaib, A. (2020). <p>The Increasing Role of Artificial Intelligence in Health Care: Will Robots Replace Doctors in the Future?</p>. International Journal of General Medicine, Volume 13, 891–896. https://doi.org/10.2147/ijgm.s268093

Six, O. (n.d.). *The ultimate guide to AI in radiology*. Artificial Intelligence in Healthcare & Radiology. https://www.quantib.com/the-ultimate-guide-to-ai-in-radiology.

U.S. Bureau of Labor Statistics. (2021, April 9). *Radiologic and MRI Technologists : Occupational Outlook Handbook*. U.S. Bureau of Labor Statistics. https://www.bls.gov/ooh/healthcare/radiologic-technologists.htm.

United Nations. (2019). *World population ageing, 2019 highlights*. United Nations.

Wardlaw, J. M., Seymour, J., Cairns, J., Keir, S., Lewis, S., & Sandercock, P. (2004). Immediate Computed Tomography Scanning of Acute Stroke Is Cost-Effective and Improves Quality of Life. Stroke, 35(11), 2477–2483. https://doi.org/10.1161/01.str.0000143453.78005.44

Weiner, J. P., Bandeian, S., Hatef, E., Lans, D., Liu, A., & Lemke, K. W. (2021). In-Person and Telehealth Ambulatory Contacts and Costs in a Large US Insured Cohort Before and During the COVID-19 Pandemic. JAMA Network Open, 4(3), e212618. https://doi.org/10.1001/jamanetworkopen.2021.2618

Wu JT, Wong KCL, Gur Y, et al. Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. JAMA Netw Open. 2020;3(10):e2022779. doi:10.1001/jamanetworkopen.2020.22779

# Appendix

## a. Project Charter

**The state of AI in Medical Imaging**

(Submitted on 03/24/2021)

### 1. Project Overview

**1. 1 Introduction** (The introduction provides a brief summary of what the project is designed to achieve, along with some background information on why the project is being done – the business drivers, the opportunity to be exploited, costs to be reduced etc.)

The primary purpose of the project is to analyze the current role of Artificial Intelligence in medical imaging and predict how these new AI-related technologies will be used by healthcare professionals in the next few years. In other worlds, it aims to consolidate our knowledge in the field of medical imaging, which is a topic of interest to any human. Because our computational capabilities have considerably expanded lately, various AI applications have become an integral part of our daily lives. I consider it is important to understand how we can leverage these discoveries not only in sectors of industry such as financial/banking, commerce, social media, but also in health. I strongly believe that when it comes to healthcare and medical treatments, we are going to automate many of the steps involved in the decision-making process. Therefore, it is critical to understand the current role of AI and how it is going to fit in the new picture. The main factor that motivates this project is a Machine Learning research I am currently working on. Even though the purpose of that project is extremely different (it studies types of folds in crumpled sheets of paper), it convinced me of the actual power that AI is bringing to the table. Because of how effective convolutional neuronal networks are when it comes to image recognition and classification, I believe that medical imaging should be a top priority for anyone interested in the impact of AI in the healthcare industry.

**1.2 Major Stakeholders** *(List all the key stakeholders (decision makers and anyone who will be impacted by the project outcomes)*

Primarily, this work will benefit professionals who work in medical imaging. However, every human is a potential stakeholder since sooner or later we all need to undergo certain medical investigations.

### 2 Project Goal and Scope

**2.1 Project Goal** *(Define the high-level goals of the project)*

- Understand the state of the art of AI in medical imaging, identify challenges, and recognize future directions.
- Establish the areas of medical imaging where AI represents a revolutionary discovery.
- Analyze ML models that are currently being used and determine their limitations.

**2.2 Project Scope** *(The project scope details the work to be taken in order to achieve the project goal. It is just as important to explicitly state what is not included in scope as it is to state what the project will deliver).*

**In Scope:**

- read research papers and academic journals to formulate solid arguments.
- study the limitations and observe whether people are working on overcoming them.
- build/recreate certain ML models to test the effectiveness of the model and to explain how the whole mechanism works.
- identify areas of improvement.
- ethical concerns (what if AI is wrong? Who is responsible?)
- analyze surveys and interviews offered by doctors.

**Out of Scope:**

- explain in detail the biology/chemistry terminology.
- analyze how AI is used in other areas (the primary focus is imaging).
- focus on outdated tools that are no longer relevant.

**3. Assumptions** *(An assumption is anything the project team or client considered to be true, real or certain often without any proof or demonstration. List in bullet format).*

- AI is safe and currently used in medical imaging.
- There is room for improvement and new solutions are needed to address new challenges.
- Published papers and articles present valid and verifiable information.
- In certain cases, AI-based solutions might perform better than humans.
- We all care about health and have a vested interest in finding ways to improve the quality of our lives.
- The future is predictable but not predetermined so any prevision about the future might or might not happen.

**4. Constraints** *(Anything that restricts or dictates the actions of the project team. These can include the so-called 'Triple Constraint'- the 'triangle' of time, cost and scope - and every project as project drivers has one or two, if not all three project constraints).*

- Limited amount of time to complete the project.
- Some papers might not be available for free.
- Not all tools/practices have been documented/investigated.
- Barriers imposed by the lack of technical knowledge.

**5. Risks** *(Risk is any unexpected event that might affect the people, processes, technology, and resources negatively or positively by the project)*

- Certain aspects might require biology/chemistry expertise.
- Insufficient information available (online or at the library).
- The scope might be too broad.
- Dead end topics.

**6. Measures of Success** *(Detailed measurements that will indicate that the project is a success)*

| Project Outcomes | Measure of Success |
|---|---|
| Clearly understood the role of AI in medical imaging | Yes/No |
| Identified at least 2 areas where there are limitations | Yes/No |
| Consulted at least 50 published articles and analyzed 2 models | Yes/No |
| Made clear forecasts about the future. | Yes/No |
| Analyzed multiple benefits of AI-related tools | Yes/No |

**7. Stakeholder Sign-off** *(For capstone thesis/case study students only capstone advisor signature is required)*

This project charter has been signed off by the client, capstone advisor and project team members.

____Catalin Veghes____ ________________ __03/24/2021___

Name Title Date

________________ ________________ ________________

Name Title Date

**b. Literature Review**

***Note: Literature has been reviewed only for a small portion of the research, to understand DNN and establish the performance of two state-of-the-art models. Therefore, we did not have enough information to complete all sections.***

## The Status of AI in Medical Imaging

The purpose of this research thesis is to explore the state of Artificial Intelligence in medical imaging, a field that is not extensively covered by the mainstream media. Its lack of publicity is understandable given that it is a topic that requires a certain amount of prior experience. Because we acknowledge the importance of intelligent algorithms in other industries such as finance or business management, we would like to investigate how things are evolving in this particular field that is of interest to every single one of us.

The main objectives for undertaking the research are rather broad as we want to accumulate as much knowledge as possible. We want to be able to determine the adoption rate of AI among different institutions (hospitals and imaging centers), the performance of state-of-the-art algorithms compared to human subjects, and the barriers that prevent AI from advancing further. Additionally, we strive to understand what is the general attitude that radiologists have towards new technologies and the role of regulatory agencies in the whole ecosystem.

To draw reliable conclusions, we need to focus on a knowledgeable population of radiologists, computer scientist and researchers. Because this is an emerging industry and distinct people have contradictory opinions, we concentrate on reputable specialists who are primarily affiliated with academic institutions rather than commercial companies. Our main goal is to analyze unbiased opinions and obtain an accurate image of what is going on. Therefore, we believe that professionals who are not affiliated with for-profit companies are more likely to provide honest insights and perspectives.

***Question:*** What is the status of Artificial Intelligence in medical imaging?

**Introduction to Literature Review**

As part of the research project, we reviewed the existing literature to understand how Deep Learning works in general, and the performance levels of two state-of-the-art models. We also took a look at other architectures such as CNNs that are extensively used by AI applications, but we decided to focus only on Deep Learning due to time and knowledge limitations. The selection process for the two models was relatively straightforward because we could find a blog post that included an organized list of top-performing models. Therefore, we thoroughly examined only three papers.

**Literature Review Components**

We analyzed three papers: one that was describing the general concepts employed by Deep Learning and two papers that proposed state-of-the-art models. The first paper provided a detailed explanation of how deep networks can extract features from images using subsequent hidden layers and enhanced our general understanding about various concepts associated with deep networks. We were able to understand how the training process works and why this category of algorithms is capable of achieving impressive results that sometimes outperform human subjects. The other two paper were mainly used to extract the performance levels, to prove our point that there are some models available that can produce results comparable to human subjects.

**Types of Published Documentation – Academic and Commercial**

All papers we analyzed were published by individuals who are affiliated with either academic or research institutions. In general, our topics of interest are primarily addressed by researchers who have a solid theoretical knowledge. So, our selection was dictated by the technical nature of the topic we are exploring. It is essential to analyze the numbers and opinions provided by those at the forefront of research in this field to draw reliable conclusions.

## Conclusions

The literature we examined proposed algorithms optimized for either classification, detection, prediction, or segmentation of medical images. Generally, these algorithms are trained to solve isolated tasks and are not one-size-fits-all solutions. After reviewing the three articles we had solid arguments to demonstrate that there are performant models on the market.

## Lessons Learned

- *Better understanding of Deep Learning concepts.*
- *AI-based systems can provide reliable interpretations of chest radiographs in clinical workflows.*
- *For 10 diseases, the ChestXNet is able to perform as well as radiologist; for three it underperformes compared with humans; and for one, the solution outdoes the professionals.*
- *ChestXNet is extremely useful for many under-developed countries in this world that suffer from an acute shortage of radiologists.*
- *The model proposed by IBM Research and University of Southern California is able to detect the findings with the same level of sensitivity as the human subjects (5 third-year radiology residents) and performs better than radiologists when it comes to the specificity and positive predictive value.*

## Cited References

Shrestha, A., & Mahmood, A. (2019). Review of Deep Learning Algorithms and Architectures. IEEE Access, 7, 53040–53065. https://doi.org/10.1109/access.2019.2912200

Armitage, H. (2018, November 20). *Artificial intelligence rivals radiologists in screening X-rays for certain diseases*. News Center. https://med.stanford.edu/news/all-news/2018/11/ai-outperformed-radiologists-in-screening-x-rays-for-certain-diseases.html.

Wu JT, Wong KCL, Gur Y, et al. Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. JAMA Netw Open. 2020;3(10):e2022779. doi:10.1001/jamanetworkopen.2020.22779