

**UNIVERSITÉ DE STRASBOURG**  
École Doctorale Des Sciences De La Vie Et De La Santé  
ICube Laboratory UMR 7357

**DISSERTATION** presented by :  
**Tony FELEFLY**

Defended on : December 13<sup>th</sup>, 2024

To obtain the grade of : **Strasbourg University Doctor**  
Discipline/Specialty : Signal and Image Processing

<p><b>Quantum-Classical Machine Learning for Brain Tumor Imaging Analysis</b></p>
---

**DISSERTATION supervisors :**

**NOËL Georges**  
**FRANCIS Ziad**

PUPH, Université de Strasbourg  
PR, Université Saint-Joseph

**RAPPORTEURS :**

**PADOVANI Laëticia**  
**CONSTANTIN, Joseph**

PUPH, Assistance Publique Hôpitaux de Marseille  
PR, Université Libanaise

**OTHER MEMBERS OF THE JURY :**

**RICARD Damien**  
**BUND Caroline**  
**ABBOUD MEHANNA Marie**  
**COUBEZ, Xavier**

PR, Université Paris Cité-Centre Borelli  
MCU-PH, Université de Strasbourg  
PR, Université Saint-Joseph  
Chercheur/Biostatisticien - Physicien, Université de Strasbourg

## Avertissement au lecteur / Warning to the reader

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition des membres de la communauté universitaire. Il est soumis à la propriété intellectuelle de l'auteur. Cela implique une obligation de citation et de référencement lors de l'utilisation de ce document. D'autre part, toute contrefaçon, plagiat, reproduction ou représentation illicite encourt une poursuite pénale.

This document is the result of a long process approved by the jury and made available to members of the university community. It is subject to the intellectual property rights of its author. This implies an obligation to quote and reference when using this document. Furthermore, any infringement, plagiarism, unlawful reproduction or representation will be prosecuted.

### [Code de la Propriété Intellectuelle](#)

#### [Article L122-4](#) :

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite. Il en est de même pour la traduction, l'adaptation ou la transformation, l'arrangement ou la reproduction par un art ou un procédé quelconque.

Any representation or reproduction in whole or in part without the consent of the author or his successors in title or assigns is unlawful. The same applies to translation, adaptation or transformation, arrangement or reproduction by any art or process whatsoever.

[Articles L335-1 à L335-9](#) : Dispositions pénales / Penal provisions.

## Licence attributed by the author



<https://creativecommons.org/licenses/?lang=fr-FR>

# Tony FELEFLY

## Quantum-Classical Machine Learning for Brain Tumor Imaging Analysis

### Résumé

La caractérisation des tumeurs cérébrales par des techniques non invasives est nécessaire. L'objectif est d'utiliser l'apprentissage automatique et la technologie quantique sur des imageries pour caractériser les tumeurs cérébrales.

Nous développons un Réseau Neuronal Quantique en utilisant la radiomique des IRM cérébrales pour différencier métastases et gliomes de haut grade. Nous sélectionnons les variables en se basant sur l'information mutuelle et nous utilisons D-Wave pour la solution. Nous entraînons le modèle sur un Simulateur Quantique. Nous utilisons les valeurs de Shapley pour expliquer les prédictions. Nous comparons les résultats à deux modèles classiques performants, DNN et XGB. Le modèle montre une performance comparable.

Ensuite, nous développons un Réseau Neuronal Convolutif 3D en utilisant des TDM cérébrales non injectées pour identifier les patients ayant des métastases cérébrales. Nous avons construit deux cohortes de patients, une avec des métastases cérébrales, et une sans anomalies cérébrales. Le cerveau a été segmenté automatiquement. Nous entraînons plusieurs modèles, et le meilleur a montré une bonne performance.

**Mots clés :** Tumeurs cérébrales ; Métastases ; Gliomes ; Apprentissage automatique ; Apprentissage quantique

### Abstract

Brain tumor characterization using non-invasive techniques is eagerly needed. The objective of this thesis is to use advanced machine learning techniques and quantum technology on brain medical images to characterize brain tumors.

First, we built a Quantum Neural Network using radiomic features from on brain MRI to differentiate between metastases and gliomas. We used a Mutual Information feature selection technique, and solved the resulting heuristic on D-Wave's Quantum Annealer. We trained the model on a Quantum Simulator. We employed instance-wise Shapley values to explain the model predictions. We benchmarked the results against two state-of-the-art classical models, Dense Neural Network and Extreme Gradient Boosting. The model showed comparable performance.

Second, we developed a 3D Convolutional Neural Network using non-enhanced brain CT scans to identify patients with brain metastases. For this purpose, we curated two cohorts of patients, one with brain metastases, and one without brain abnormalities. The brain was automatically segmented. We trained several versions of the model, and the best model showed an impressive performance.

**Keywords :** Brain tumors ; Metastases ; Glioma ; Machine learning ; Quantum machine learning

To my beloved family

# ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisors, Professor Georges Noël and Dr. Ziad Francis, as well as my USJ co-supervisor Dr. Georges Fares, whose expertise, understanding, and patience, added considerably to my research experience. Your guidance and commitment have been a constant source of encouragement and have profoundly shaped my journey.

I am extremely grateful to the members of my thesis committee for their time, and for accepting to be part of this jury. I am also thankful to the administration teams of both Strasbourg and Saint-Joseph Universities for their continuous support, and to the IPHC and ICube laboratories for endorsing my research and allowing me to use their resources.

My sincere thanks also go to Professor Elie Nasr for sharing his pearls of wisdom with me during the course of this research, and for making this work happen. I am indebted to the whole radiation oncology and radiology teams at the Hôtel-Dieu de France Hospital for their assistance, particularly in data collection, which was a great help for my projects. Special thanks to Dr. Camille Roukoz for her invaluable help.

I would like to thank my fellowship supervisor Professor Scott C. Morgan, and my colleagues at the Hôtel-Dieu de Lévis radiation oncology department for their unwavering support throughout this journey.

Completing this PhD has been one of the most challenging and rewarding experiences of my life, and it would not have been possible without the incredible support of my family.

To my parents Hanna and Marina, thank you for instilling in me the values of hard work and perseverance. Your endless encouragement and belief in my abilities have been my anchor throughout this journey. To my sisters Gisèle, Layale, and Carole, your unswerving support and occasional necessary distractions have kept me balanced and focused.

A special word of gratitude goes to my wife Manal. Your love, understanding, and sacrifice have not only been a source of comfort but a cornerstone of my success. You have been my partner in every sense, sharing in the lows and celebrating the highs, always making sure our home was a place where I could focus and recharge.

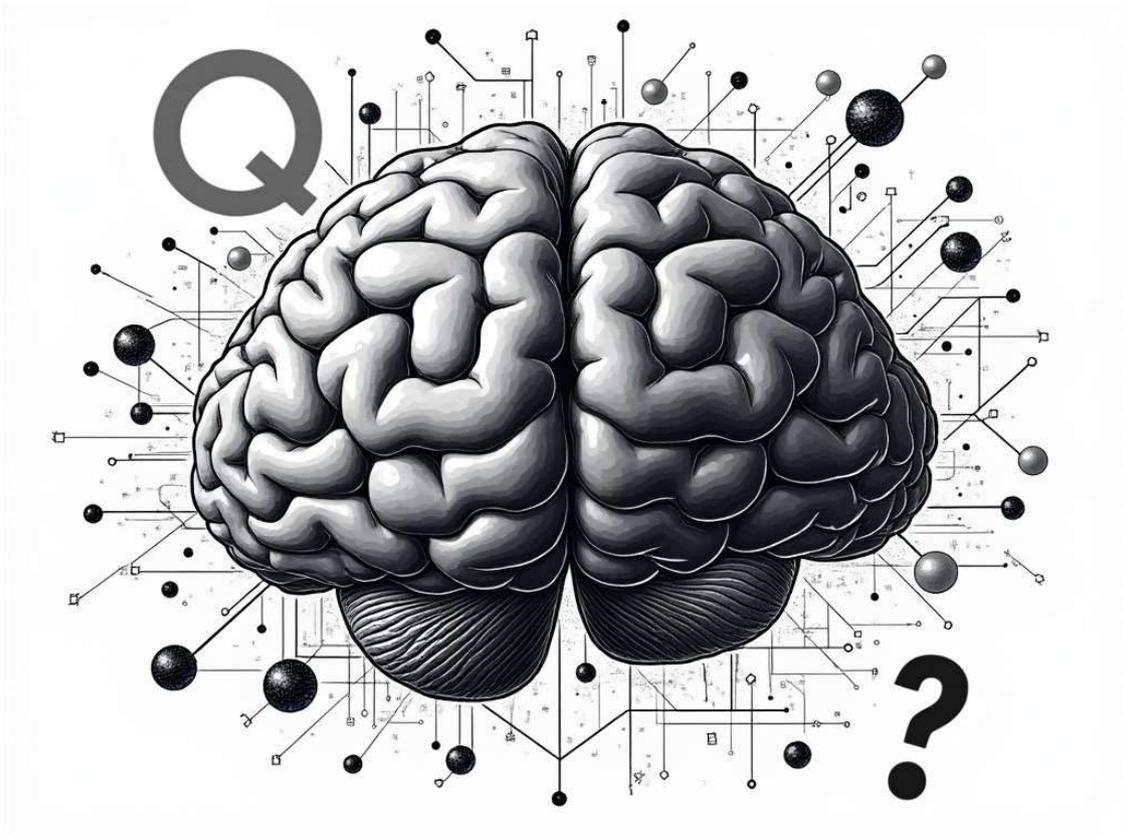
And to my son Ryan, who has been my greatest joy and motivation. Watching you grow as I navigated this journey has been a profound reminder of why I embarked on this

path. You inspire me every day to be the best version of myself, not just as a scholar, but as a father.

Thank you all, from the bottom of my heart, for walking this path with me and for your endless love and patience. This achievement is as much yours as it is mine.

Tony





*Image generated by Grok-xAI*

# CONTENTS

## 1 Table of Contents

1	Introduction.....	22
1.1	Clinical Perspective.....	23
1.1.1	Differentiation between metastases and glioma on T1W MRI.....	24
1.1.2	Detection of brain metastases on non-enhanced CT.....	26
1.2	Technical Perspective.....	27
1.2.1	Differentiation between metastases and glioma on T1W MRI.....	27
1.2.2	Detection of brain metastases on non-enhanced CT.....	34
1.3	Objectives.....	36
2	Technical Overview .....	38
2.1	Radiomics .....	39
2.1.1	Handcrafted Radiomics:.....	40
2.1.2	Deep Radiomics .....	41
2.2	Deep Learning .....	42
2.3	Quantum Computing.....	47
2.3.1	Fundamentals of Quantum Computing.....	47
2.4	Quantum Annealing and Adiabatic Quantum Computing .....	52
2.4.1	The Ising model and QUBO.....	52
2.4.2	Adiabatic Quantum Computing.....	52
2.4.3	Quantum Annealing.....	53
2.5	Quantum Machine Learning.....	57
3	An Explainable MRI-Radiomic Quantum Neural Network to Differentiate Between Large Brain Metastases and High-Grade Glioma Using Quantum Annealing for Feature Selection ...	64
3.1	ABSTRACT .....	65
3.2	INTRODUCTION.....	67
3.3	MATERIALS AND METHODS.....	70
3.3.1	Patients and data collection.....	70

3.3.2	Segmentation.....	70
3.3.3	Image pre-processing.....	70
3.3.4	Feature extraction and data pre-processing.....	71
3.3.5	Feature selection.....	71
3.3.6	Quantum neural network.....	73
3.3.7	Benchmarking and model evaluation.....	75
3.3.8	QNN model interpretation.....	76
3.4	RESULTS.....	77
3.4.1	Patient characteristics.....	77
3.4.2	Feature selection.....	77
3.4.3	QNN model performance and benchmarking.....	78
3.4.4	QNN model interpretability.....	80
3.5	DISCUSSION.....	83
3.6	CONCLUSION.....	86
3.7	APPENDIX.....	87
4	A 3D Convolutional Neural Network Based on Non-Enhanced Brain CT to Identify Patients with Brain Metastases.....	92
4.1	ABSTRACT.....	93
4.2	INTRODUCTION.....	94
4.3	MATERIALS AND METHODS.....	96
4.3.1	Data collection.....	96
4.3.2	Segmentation and image pre-processing.....	96
4.3.3	Deep Learning model development.....	97
4.4	RESULTS.....	98
4.4.1	Patient data.....	98
4.4.2	Model evaluation.....	99
4.5	DISCUSSION.....	102
4.6	CONCLUSION.....	105
5	Conclusion and Future Works.....	107
5.1	Conclusion.....	108
5.2	Perspectives and Future Works.....	110
6	Résumé en Français.....	113
7	BIBLIOGRAPHY.....	120



## LIST OF FIGURES

<b>Figure 1.1</b> Brain MRI images from our dataset for two different patients with Glioblastoma (left) and metastases (right). .....	25
<b>Figure 2.1</b> Differences in workflow between handcrafted radiomics and deep radiomics. Reprinted with permission from Springer Nature from Hosny <i>et al</i> , 2018 (53) .....	39
<b>Figure 2.2</b> Architecture of a Convolutional Neural Network. ReLU: Rectified Linear Unit; FC: Fully Connected. Reprinted from Yamashita <i>et al.</i> (58), no permission is required, open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, Springer Nature .....	42
<b>Figure 2.3</b> A convolutional operation with zero padding. Input size is 5x5, Kernel size is 3x3. Reprinted from Yamashita <i>et al.</i> (58), no permission is required, open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, Springer Nature .....	44
<b>Figure 2.4</b> Max pooling operation. Reprinted from Yamashita <i>et al.</i> (58), no permission is required, open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, Springer Nature .....	44
<b>Figure 2.5</b> A quantum circuit representing two qubits q0 and q1, and one classical bit c2. It shows a sequence of quantum gates and unitary operations, and ends with a measurement of the two qubits. Image created using IBM Qiskit. ....	48
<b>Figure 2.6</b> Most used quantum gates, their respective notations, and matrix representations. Reprinted from Yan <i>et al.</i> (61). No permission required, published under an open access Creative Common CC BY license .....	49
<b>Figure 2.7</b> Graph representation of D-Wave's Chimera (a) and Pegasus (b) topologies. The green highlighted area depict a unit cell in each graph. Red and blue dots respectively denote horizontal and vertical connections to qubits in neighboring unit	

cells. Reprinted from Yarkoni <i>et al.</i> (75); No permission required, published under a Creative Common CC BY license.....	55
<b>Figure 2.8</b> Different steps in implementing a QUBO optimization on D-Wave systems hardware. Reprinted from Yarkoni <i>et al.</i> (75); No permission required, published under a Creative Common CC BY license.....	56
<b>Figure 2.9</b> The four combinations of machine learning relative to the type of data and computing algorithm. Reprinted from (76) by Maria Schuld; No permission required, shared under a Creative Commons CC BY-SA 4.0 license.....	57
<b>Figure 2.10</b> Variational Quantum Algorithm framework, showing a set of training data $\{\rho_k\}$ transformed using an ansatz with a function $f_k(\theta, \rho_k)$ with set of parameters $\theta$ . Parameter optimization is classically performed with a cost function $C(\theta)$ , and then parameters are updated in the quantum circuit. In the figure, the ansatz is shown as a quantum circuit and as a neural network to highlight the analogy. A variety of output data types could be obtained, with the most prominent examples shown in the red box. Reprinted with permission from Springer Nature from Cerezo <i>et al.</i> (82) .....	59
<b>Figure 2.11</b> Key applications for Variational Quantum Algorithms. Reprinted with permission from Springer Nature from Cerezo <i>et al.</i> (82) .....	60
<b>Figure 2.12</b> Anatomy of a two-qubits Variational Quantum Classifier .....	61
<b>Figure 2.13</b> Different encoding techniques, showing basis and amplitude encoding on a 2-qubit system, and Hamiltonian and time-evolution embeddings on an arbitrary unitary. Reprinted with permission from Springer Nature from Schuld and Petruccione (40) .....	61
<b>Figure 3.1</b> Circuit representation for amplitude encoding in the case of a 2-qubit system. $R_y\beta$ denotes the Y-axis rotation angle; the white circle indicates a control on qubit basis state 1; the black circle indicates a control on qubit basis state 0; $q_1$ and $q_2$ are qubit 1 and 2 states respectively.....	74

<b>Figure 3.2</b> The 2-qubit circuit acting as a machine learning classifier. RY: Y-axis rotation; X: Pauli X gate; Rot: arbitrary qubit rotation. Qubits number 1 and 2 are denoted “0” and “1” respectively.....	75
<b>Figure 3.3</b> The results of the MIFS heuristic for all $k$ values between 1 and 17, showing the best combination of $k$ features. Blue squares correspond to selected features.....	78
<b>Figure 3.4</b> Learning curves for training and test accuracy and training loss for the Quantum Neural Network model.....	79
<b>Figure 3.5</b> Mean absolute Shapley Value for each feature, ranked by order of importance.....	81
<b>Figure 3.6</b> Mean Shapley Values and feature values for the different features according to the predicted class (metastasis or glioma) <b>[a]</b> , and the prediction accuracy (true or false) <b>[b]</b> . Colors represent the mean feature values according to the color scale bar. ....	82
<b>Figure 3.7 Supplementary</b> Pipeline for developing the Quantum Neural Network..	87
<b>Figure 3.8 Supplementary</b> Correlation matrix for the whole feature set.....	88
<b>Figure 3.9 Supplementary</b> Correlation matrix for remaining features after eliminating highly correlated features.....	89
<b>Figure 3.10 Supplementary</b> Mutual Information scores in decreasing order for the remaining 54 features following LinearSVC feature selection. The best 17 features were visually selected. The black arrow points past the 17 <sup>th</sup> selected feature.....	90
<b>Figure 4.1</b> Model A network architecture. Numbers indicate the tensor shape of the corresponding layer. Dimensions in the Z-direction (horizontal) were rescaled to 1:100 to allow clear visualization.....	99
<b>Figure 4.2</b> Training and validation learning curves for accuracy and loss for models A <b>(a)</b> , B <b>(b)</b> , and C <b>(c)</b> . train: training; val: validation; acc: accuracy.....	100

**Figure 4.3** Receiver operating characteristic curves for model A for the training and validation sets **(a)**; Confusion matrices for model A on the training **(b)** and validation **(c)** sets. Label “0” corresponds to normal CT, and label “1” denotes a CT with brain metastases. AUC\_Tr: area under receiver operating characteristic curve for the training set; AUC\_Val: area under receiver operating characteristic curve for the validation set.  
..... 101

## LIST OF TABLES

<b>Table 1.1</b> Overview of studies on differentiating between brain metastases and glioblastoma. Reprinted from Jekel <i>et al.</i> (37); <i>No permission is required; published under an open access Creative Common CC BY license. GBM= Glioblastoma; MET = Brainmetastasis; PCNSL = Primary central nervous systemlymphoma; MEN =Meningioma; MED =Medulloblastoma; CV = Cross-validation; LOOCV = Leave-One-Out cross-validation; ML =Machine learning; DL = Deep learning; T1CE = contrast-enhanced T1-weighted sequence; DWI = Diffusion weighted imaging; DTI = Diffusion tensor imaging; PWI = Perfusion weighted imaging; rCBV = relative cerebral blood volume; FLAIR = Fluid-attenuated inversion recovery; TE = Time to echo; AUC = Area under the receiver operating characteristic curve; ADC = Apparent diffusion coefficient; LASSO = Least absolute shrinkage and selection operator; SVM= Support vectormachine; MLP =Multilayer perceptron; NNW= Neural networks; LogReg = Logistic Regression; DNN = Deep neural network; LDA = Linear discriminant analysis; NB = Naïve Bayes; VFI = Voting feature intervals; KNN = k-nearest neighbors; PNN = Probabilistic neural networks; RF = Random Forest; RBF = Radial basis function kernel; n/a = not available. ....</i>	28
<b>Table 1.2</b> TRIPOD Checklist Items for model development. Reprinted from <a href="https://www.tripod-statement.org/resources/">https://www.tripod-statement.org/resources/</a> .....	31
<b>Table 2.1</b> Examples of machine learning subroutines that benefit from quantum algorithms speedups. Complexities are shown with respect to classical algorithms. qRAM: quantum Random Access Memory. Reprinted with permission from Springer Nature, from Biamonte <i>et al.</i> (80).....	58
<b>Table 3.1</b> Models' performance metrics.....	80
<b>Table 4.1</b> Patients and tumors characteristics .....	98
<b>Table 4.2</b> Evaluation metrics for model A.....	100



# **ABBREVIATIONS**

3D CNN – 3D Convolutional Neural Network

ACC - Accuracy

ADC - Apparent Diffusion Coefficient

AI – Artificial Intelligence

ANTs - Advanced Normalization Tools

AUC - Area Under the Receiver Operating Characteristic Curve

bACC – Balanced Accuracy

CAD – Computer Aided Diagnosis

CNN - Convolutional Neural Network

CT - Computed Tomography

DL - Deep Learning

DNN - Dense Neural Network

DTI - Diffusion Tensor Imaging

DWI - Diffusion Weighted Imaging

FC - Fully Connected

FLAIR - Fluid-Attenuated Inversion Recovery

GLCM - Gray-Level Cooccurrence Matrix

GLDZM - Gray-Level Distance Zone Matrix

GLRLM - Gray-Level Run-length Matrix

GLSZM - Gray-Level Size Zone Matrix

HGG – High Grade Glioma

IBSI - Image Biomarker Standardization Initiative

IQR – Interquartile Range

KNN - k-Nearest Neighbors

LASSO - Least Absolute Shrinkage and Selection Operator

LBM – Large Brain Metastases  
LDA - Linear Discriminant Analysis  
LogReg - Logistic Regression  
LOOCV - Leave-One-Out Cross-Validation  
MI – Mutual Information  
MIFS – Mutual Information Feature Selection  
ML - Machine Learning  
MLP - Multilayer Perceptron  
MRI - Magnetic Resonance Imaging  
NB - Naïve Bayes  
NE-CT – Non enhanced Computed Tomography  
NGLDM - Neighborhood Gray-Level Dependence Matrix  
NGTDM - Neighborhood Gray-Tone Difference Matrix  
NISQ - Noisy Intermediate Scale Quantum  
NNW - Neural Networks  
PET CT – Positron Emission Tomography – Computed Tomography  
PNN - Probabilistic Neural Networks  
PWI - Perfusion Weighted Imaging  
QKD – Quantum Key Distribution  
QML - Quantum Machine Learning  
QNN - Quantum Neural Network  
QUBO - Quadratic Unconstrained Binary Optimization  
RBF - Radial Basis Function Kernel  
rCBV - Relative Cerebral Blood Volume  
ReLU - Rectified Linear Unit  
RF - Random Forest  
ROC-AUC - Area-Under-the-Receiver-Operating-Characteristic-Curve

SVM - Support Vector Machine

T1CE - Contrast-Enhanced T1-Weighted

TE - Time to Echo

TRIPOD - Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

VFI - Voting Feature Intervals

VQA - Variational Quantum Algorithm

VQC - Variational Quantum Classifier

VQE - Variational Quantum Eigensolver

XGB - Extreme Gradient Boosting



# 1 Introduction

In this chapter, we will provide context for this thesis, detailing the rationale of the proposed research and its position with regard to published studies.

It will be divided into three parts:

- 1- Clinical perspective: this part will provide the clinical rationale for this work, and explain the importance of addressing the research questions in terms of the inferred benefits for cancer patients
- 2- Technical perspective: this part will include details on the techniques and methods used in previously published works on the same or similar research questions, highlighting caveats and limitations
- 3- Objectives: this part will briefly summarize the objectives of this thesis

## 1.1 Clinical Perspective

Brain metastases are the most frequent intracranial tumors, with an estimated annual incidence of 7 to 14 cases per 100,000 population, followed by high-grade gliomas (1.95 per 100,000 per year) (1,2). Notably, 10-40% of patients with solid tumors are estimated to develop brain metastases (2). It is believed that there is an increasing trend in incidence that can be attributed to advancements in cancer treatment, improved imaging techniques like MRI, and an aging population (3). The most frequent primary cancer sites leading to brain metastases are lung and breast cancers (2).

Magnetic Resonance Imaging (MRI) is the gold standard imaging modality for the diagnosis of brain tumors. It is highly sensitive, significantly surpassing Computed Tomography CT scans in detecting small and deep-seated tumors (4). It can provide detailed information about the number, location, and size of tumors. Advanced MRI techniques like functional MRI, perfusion MRI, and magnetic resonance spectroscopy further aid in tumor evaluation and treatment planning. Undeniably, histopathological diagnosis remains the ultimate ground truth for confirming and characterizing brain malignancies. Obtaining tissue material for histopathology consists of an invasive procedure with either biopsy or surgical resection.

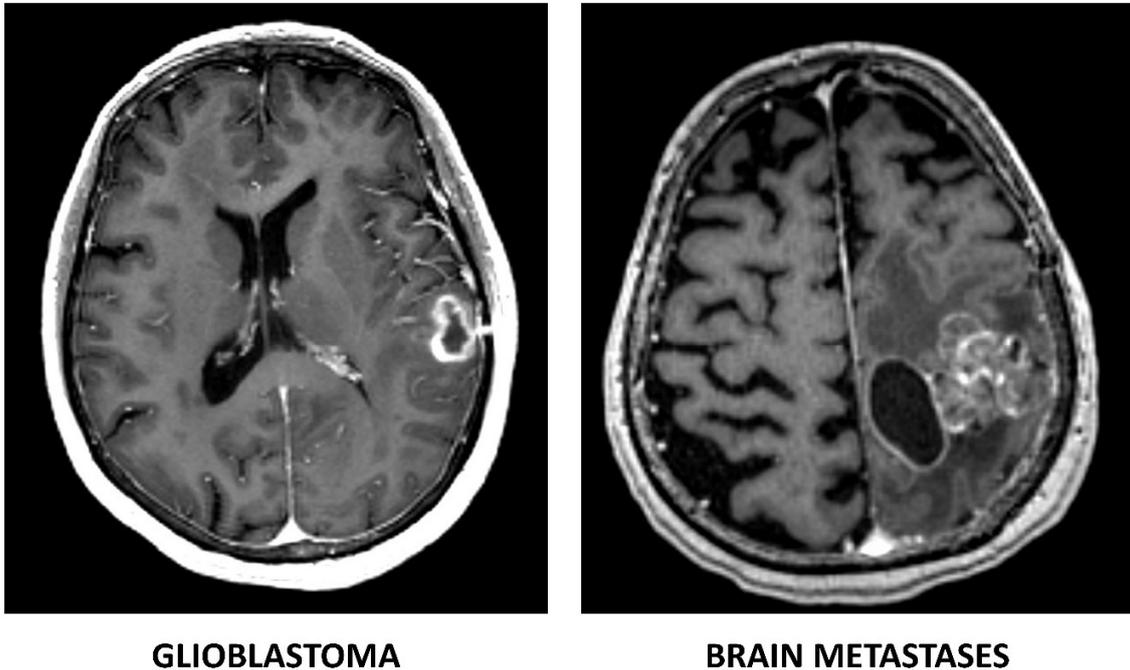
The widespread availability of medical imaging has revolutionized the diagnosis and management of brain tumors. Advanced imaging modalities such as MRI and CT scans are now routinely used to detect, characterize, and monitor brain tumors. These technologies provide high-resolution images that reveal detailed information about the tumor's size, location, and potential effects on surrounding brain tissue. However, a significant amount of data contained within these images remains largely unexplored, mainly due to the limited potential of the human eye to glean such large amount of sometimes subtle information. This untapped potential includes intricate details about tumor composition, vascularization, and growth patterns that could offer deeper insights into the nature of the tumor, its behavior, and related prognosis. Furthermore, slight changes with respect to normal anatomy that could be easily un-noticed by humans, might be key to identify early growing tumors. The complexity and sheer volume of this imaging data present both a challenge and an opportunity for healthcare professionals and researchers. By harnessing advanced computational techniques, such as machine learning and artificial intelligence, there is potential to extract and analyze these nuanced aspects of brain tumor imaging, leading to more personalized and effective treatment strategies. This approach could unveil new biomarkers for tumor characterization, improve diagnostic accuracy, and pave the way for tailored therapeutic interventions, ultimately enhancing patient outcomes in the field of neuro-oncology.

There exist infinite potential oncological benefits from harnessing existing data from brain medical imaging, such as identifying biomarkers for prognostication and prediction of response to therapy, differentiating between different tumor types and sub-categories, between tumor progression and pseudoprogression, correlating imaging patterns to existing biomarkers, and so on and so forth. Definitely, the bottleneck for such applications is the access to high-quality clinical information. In this work, we focus on two separate tasks:

- 1- Differentiating between brain metastases and high-grade gliomas using T1-weighted brain MRI images
- 2- Identifying patients with brain metastases from non-enhanced brain CT images

### **1.1.1 Differentiation between metastases and glioma on T1W MRI**

Differentiating between brain metastases and high-grade gliomas presents a significant challenge, as they often exhibit similar imaging patterns on traditional MRI sequences. Although there might be differences in tumor morphology, with gliomas being more irregular in shape, larger, and exhibiting more heterogeneous enhancement than metastases, there remains much uncertainty and lack of specificity for using such criteria. Figure 1.1 shows two images from our dataset highlighting the exact opposite case scenario.



**Figure 1.1** Brain MRI images from our dataset for two different patients with Glioblastoma (left) and metastases (right).

Accurate distinction between these entities is crucial, as their management strategies differ significantly. Pathology confirmation remains the standard practice for definite diagnosis and treatment planning. However, there are many instances where non-invasive methods to differentiate between these two tumor types are beneficial, and sometimes highly important.

Firstly, upfront characterization may lead to better surgical planning and preparedness for the extent of resection. Moreover, tumors might occur in areas where surgery or biopsy could not be safely performed; in such cases, non-invasive diagnosis becomes the only way to guide treatment. While brain metastases would mostly be treated with stereotactic radiotherapy in this case, treatment for high-grade glioma would consist of conventionally fractionated radiotherapy with temozolomide chemotherapy and tumor-treating fields when available. We acknowledge that this remains a rare occurrence, albeit not impossible. Also, for patients with a history of proven malignancy, an accurate non-invasive method to characterize a brain lesion might be particularly helpful in avoiding unnecessary biopsy-related complications. Taken altogether, optimizing non-invasive easily accessible methods for differentiating between brain metastases and high-grade gliomas represents an important arena to explore as it provides undeniable advantages.

### **1.1.2 Detection of brain metastases on non-enhanced CT**

The early detection of brain metastases is vital, as metastatic complications contribute significantly to cancer-related morbidity and mortality. Furthermore, it is becoming increasingly important in light of recent and ongoing developments in pharmacological anti-cancer agents showing significant central nervous system activity. Identifying patients with asymptomatic brain metastases could lead to changes in therapeutic strategy, both on the level of local and systemic treatment.

MRI is the preferred and most sensitive imaging modality for brain metastases detection. However, for many cancer sites such as breast cancer, its use is limited to patients with symptoms suggestive of brain involvement. Moreover, in the context of advanced and metastatic lung cancer, treatment response evaluation does not systematically include brain imaging for patients without a prior diagnosis of brain metastases (5).

Non-enhanced brain CT scans are increasingly available for cancer patients as a component of whole-body PET-CT whose use has been rising due to its adoption for cancer staging and treatment response monitoring (5). Additionally, non-enhanced CT scans could be performed for a variety of reasons such as investigating patients with head trauma or new onset headache to exclude intracranial bleeding in the Emergency Room setting.

Interpreting non-enhanced brain CT scans is intricate owing to the lack of contrast and the subtle differences in CT values between normal brain parenchyma and metastatic lesions. Computational methods that help addressing these challenges are eagerly needed in order to get the most information out of available imaging data that is otherwise unexploited. This might lead to identifying patients with a high likelihood of having abnormalities suggestive of metastases that might benefit from dedicated brain MRI imaging.

## 1.2 Technical Perspective

The use of artificial intelligence (AI) for medical applications has been growing significantly in the recent years, and we have already witnessed many of these methods and tools being used in the clinical setting. We can cite in this space the Decipher® genomic classifier used for prostate cancer prognostication (6), and more recently a multimodal AI model (ArteraAI®) based on clinical and histopathological data that has been validated as a predictor to benefit from hormonal treatment in prostate cancer (7).

One of the most interesting use of AI in medicine is to analyze medical imaging. Medical images contain too many information, most of which is not discernable to the human eye. Since the introduction of Radiomics in 2012 (8), there has been a continuous rise in the use of computational methods, including machine learning, to extract data from medical images that are then used for endpoint modelling/prediction, prognostication, and computer-aided diagnosis.

We aimed herein to explore advanced machine learning techniques to address the tasks described in the previous section, i.e. to differentiate between brain metastases and high-grade glioma on T1-weighted MRI images, and to identify patients with brain metastases from non-enhanced brain CT scans.

In this section, we will present an overview of the published research for each of these two projects included in this thesis.

### 1.2.1 Differentiation between metastases and glioma on T1W MRI

Many previous papers reported on machine learning models to differentiate between brain metastases and high-grade glioma on MRI images (9–36). These studies are nicely summarized in a systematic review by Jekel *et al* (37) from Yale school of Medicine. Table 1.1 was reprinted and highlighted the most relevant technical details and best classifier performance in these studies.

As shown in Table 1.1, the classification accuracy in these papers ranged between 64 and 98%, with a pooled average of 88.1% from 19 studies that reported on accuracy. These differences highlight the huge heterogeneity between these studies, stemming from a variety of reasons including but not limited to, different inclusion criteria, imaging protocols, quality of data, MRI sequences used, image preprocessing techniques, reported metrics, and machine learning techniques.

**Table 1.1** Overview of studies on differentiating between brain metastases and glioblastoma. Reprinted from *Jekel et al. [37]*; No permission is required; published under an open access Creative Common CC BY license.

Paper	Total Patient Number	Number of Glioma Patients	Number of BM Patients	Ratio of Glioma/BM Patients	Solitary BM Only	GBM Only	Tumor Types Studied	Number of Additional Tumors	Number of Patients (Training)	Number of Patients (Validation)	Testing	External Validation	Source of Data	ML Method	Algorithms Used for Classification	Gold Standard for Accuracy	Extracted Feature Types	Best Performing Classifier
Swinburne et al., 2019 [13]	26	9	9	1.000	no	yes	GBM vs. MET vs. PCNSL	8 (PCSNL)	LOOCV		no	single-center	ML	SVM, MLP	Pathology	Perfusion	MLP (Ktrans on T1CE mask) Accuracy: 83.3% AUC: 0.83	
Park et al., 2020 [14]	276	137	99	2.322	no	yes	GBM vs. MET vs. PCNSL	80 (PCSNL)	216 (109 GBM, 58 PCNSL, 49 MET)	60 (28 GBM, 22 CNSL, 10 MET)	no	multi-center	DL	CNN	Pathology	Perfusion (temporal Patterns of Time-Signal Intensity Curves from DSC)	CNN (DSC, FLAIR, T1CE)—Internally validated AUC: 0.95 Sensitivity: 0.9 Specificity: 0.887	
Shrot et al., 2019 [15]	141	41	38	1.079	no	yes	GBM vs. MET vs. PCNSL vs. MEN	12 (PCSNL), 50 (Meningioma)	LOOCV		no	single-center	ML	Decision tree (SVM)	Pathology	Morphology, Diffusion, Perfusion	Binary hierarchical tree with SVM classifier (T1, T2, DSC) Sensitivity: 0.974 Specificity: 0.969	
Yamashita et al., 2008 [16]	126	95	19	5.000	multiple	no	Glioma vs. MET vs. PCNSL	12 (PCSNL)	LOOCV		no	not specified	ML	3-layered NNW	Pathology	Clinical, Qualitative/ Semantic imaging features	ANN AUC: 0.946 board-certified radiologists without ANN: Accuracy: 87.9% AUC: 0.923 Sensitivity: 0.808 Specificity: 0.903 board-certified radiologists with ANN: Accuracy: 91.5% AUC: 0.946 Sensitivity: 0.868 Specificity: 0.931	
Blanchet et al., 2011 [17]	33	18	15	1.200	solitary	yes	GBM vs. MET		LOOCV		no	single-center	ML	k-means clustering	Pathology	Shape	k-means clustering (T1, T2) Accuracy: 93.9%	
Tsolaki et al., 2013 [18]	49	35	14	2.500	solitary	yes	GBM vs. MET		10-fold CV		no	single-center	ML	SVM, Naive Bayes, KNN	Pathology	Spectroscopy	SVM (MRS: NAA, rCBV)—intratumoral Accuracy: 98% Sensitivity: 0.98 Specificity: 0.99 SVM (MRS: NAA, Cr, rCBV)—intratumoral Accuracy: 95% Sensitivity: 0.94 Specificity: 0.95	
Yang et al., 2014 [19]	48	30	18	1.667	solitary	yes	GBM vs. MET		LOOCV		no	single-center	ML	QDA, NB, SVM, KNN, NNW (MLP architecture)	Pathology	Shape, Diffusion	Neural Network (DT) Accuracy: 97.92% AUC: 0.975 Sensitivity: 100% Specificity: 96.55%	
Tateishi et al., 2020 [20]	127	73	53	1.377	multiple, largest selected for classification	yes	GBM vs. MET		5-fold CV		no	single-center	ML	SVM	Pathology, clinical history of path-proven primary cancer	Texture	SVM (T1CE, T2, ADC) AUC: 0.92	
Abidin et al., 2019 [21]	52	35	17	2.059	solitary	yes	GBM vs. MET		stratified 10-fold CV		no	single-center	ML	AdaBoost	Pathology	First-order statistics, Texture, Higher-order features: Topology (Minkowski functionals), Wavelet-transformed, Local Binary Patterns (LBP)	AdaBoost (Local Binary Pattern, T1CE) AUC: 0.846	
Bae et al., 2020 [22]	248	159	89	1.787	solitary	yes	GBM vs. MET		166 (109 GBM, 57 MET)	82 (50 GBM, 32 MET)	yes	single-center	ML and DL	DNN, AdaBoost, (L-SVM, LDA)	Pathology	DL extracted (DL) Shape, First-order statistics, Texture (traditional ML)	Deep Neural Network (T1CE)—internal AUC: 0.986 Deep Neural Network (T1CE)—external AUC: 0.956 Accuracy: 89% Sensitivity: 0.906 Specificity: 0.88	

Paper	Total Patient Number	Number of Glioma Patients	Number of BM Patients	Ratio of Gliomatous Patients	Solitary BM Only	GBM Only	Tumor Types Studied	Number of Additional Tumors	Number of Patients (Training)	Number of Patients (Validation)	Testing	External Validation	Source of Data	ML Method	Algorithms Used for Classification	Gold Standard for Accuracy	Extracted Feature Types	Best Performing Classifier
Artzi et al., 2019 [23]	439	212	227	0.934	solitary	yes	GBM vs. MET vs. MET-subtypes		5-fold CV		no	single-center	ML	SVM, KNN, decision trees, ensemble classifiers	Pathology	Clinical features, Quantitative/semantic imaging features, Morphology, First-order statistics, Texture, Higher-order features, Wavelet features, Bag-of-features	SVM (TICE) Accuracy: 89% AUC: 0.96 Sensitivity: 0.86 Specificity: 0.85	
Yang et al., 2016 [24]	48	30	18	1.667	solitary	yes	GBM vs. MET		LOOCV		no	single-center	ML	SVM	Pathology	Shape	SVM (DTL Cluster 1 & 4) Accuracy: 95.83% AUC: 0.983 Sensitivity: 0.9444 Specificity: 0.9667	
Dong et al., 2020 [25]	120	60	60	1.000	solitary	n/a	Glioma vs. MET	84 (42 GBM, 42 MET)	36 (18 GBM, 18 MET)		no	single-center	ML	NNW, DT, NB, KNN, SVM	Radiological	Shape, First-order statistics, Texture	Native Bayes (T1, T1CE, T2) Accuracy: 60% Sensitivity: 0.45 Specificity: 0.75 Combined (LOG) (Decision Tree, SVM, NNW, KNN) Accuracy: 64% Sensitivity: 0.5 Specificity: 0.73 Agreement of all 5 classifier: Accuracy: 94% Sensitivity: 1 Specificity: 0.89	
Meier et al., 2020 [26]	109	25	84	0.298	231 lesions in 109 patients	yes	GBM vs. MET		stratified 3-fold CV		no	single-center	ML	SVM	Pathology	Qualitative/Semantic imaging features	SVM (Qualitative image features) F1-Score: 0.865	
Georgiadis et al., 2008 [27]	67	21	19	1.105	no	no	Glioma vs. MET vs. MEN	27 (Meningioma)	external cross-validation (ECV) with 3-fold split		no	single-center	ML	PNN, LSFF, PNN, SVM, RBF, ANN, Cubic, LSFF, PNN, Quadratic, LSFF, PNN	Radiological	Texture	ANN (T1)—Primary tumors vs. Secondary tumors (MET + Meningioma) Accuracy: 100%	
Tsolaki et al., 2015 [28]	126	80	22	3.636	solitary	no	Glioma vs. MET vs. MEN	24 (Meningioma)	10-fold cross validation		no	single-center	ML	SVM, Naive Bayes, k-NN, LDA	Pathology	Spectroscopy, Diffusion, Perfusion	SVM (DWI/DTI/PWI/short TE)—peritumoral Accuracy: 98% (DWI/DTI/PWI/short TE)—intratumoral Accuracy: 96%	
Zacharaki et al., 2009 [27]	98	74	24	3.083	no	no	Glioma vs. MET vs. MEN	4 (Meningioma)	LOOCV		no	single-center	ML	SVM, k-NN, LDA	Pathology	Shape, First-order statistics, texture	SVM (FLAIR, T2, T1ce, rCBV, T1) Accuracy: 84.7% AUC: 0.882 Sensitivity: 0.882 Specificity: 0.865	
Zacharaki et al., 2011 [29]	97	73	23	3.174	no	no	Glioma vs. MET vs. MEN		LOOCV		no	single-center	ML	VFL, KNN, Naive Bayes	Pathology	Clinical, Shape, First-order	kNN with wrapper evaluator Accuracy: 96.91%	
Svolos et al., 2013 [30]	115	73	18	4.056	solitary	no	Glioma vs. MET vs. MEN	24 (atypical Meningioma)	10-fold cross validation		no	single-center	ML	SVM	Pathology	Diffusion, Perfusion	SVM (HGG Grade 4 vs. MET) (ADC, FA, rCBV) — peritumoral Accuracy: 96% Sensitivity: 0.98 Specificity: 0.94	
Sachdeva et al., 2016 [31]	428	177	66	2.682	no	no	Glioma vs. MET vs. MEN vs. MED (Medulloblastoma)	97 (Meningioma), 88 (Medulloblastoma)	40% training, 10% testing, 50% validation	40% training, 10% testing, 50% validation	40% training, 10% testing, 50% validation	no	public dataset (FCIMER and SPL datasets)	ML	GA, GA-SVM, GA-ANN	Radiological	First-order statistics, Texture	GA-ANN—no binary classification Accuracy: 94% (imputed)
Payabvash et al., 2020 [32]	248	99	65	1.523	no	no	Glioma vs. MET vs. Hemangioblastoma vs. Ependymoma	Hemangioblastoma (n = 44), Ependymoma (n = 27), Medulloblastoma (n = 26)	10-fold cross validation		no	single center	ML	NB, RE, NN, SVM	Pathology	Clinical (Age), Qualitative/Semantic imaging features, Diffusion	Random Forest—MET vs. All primary tumors Accuracy: 83% AUC: 0.82 Sensitivity: 55.6 Specificity: 92.6 PPV: 73.9	
Qin et al., 2019 [33]	42	24	18	1.333	solitary	yes	GBM vs. MET		5-fold cross validation		no	single center	ML	Decision trees, LDA, LogReg, linear, SVM, KNN	Pathology	First-order, Second-order (Energy)	kNN Accuracy: 92.9%	

Paper	Total Patient Number	Number of Glioma Patients	Number of BM Patients	Ratio of Glioma/BM Patients	Solitary BM Only	GBM Only	Tumor Types Studied	Number of Additional Tumors	Number of Patients (Training)	Number of Patients (Validation)	Testing	External Validation	Source of Data	ML Method	Algorithms Used for Classification	Gold Standard for Accuracy	Extracted Feature Types	Best Performing Classifier
Chen et al., 2019 [34]	134	n/a	n/a		no	yes	GBM vs. MET		80%	20%		no	single center	ML	LDA, SVM, RF, KNN, Gaussian NB	Pathology	Texture	LogReg + Distance correlation Accuracy: 79% AUC: 0.8 Sensitivity: 0.8 Specificity: 0.71
Ortiz-Ramón et al., 2020 [35]	100	50	50	1.000	no	yes	GBM vs. MET		nested cross-validation		no	single center	ML	random forest (RF), support vector machine (SVM) with linear kernel, k-nearest neighbors (KNN), naïve Bayes (NB) and multi-layer perceptron (MLP)	Radiological	Texture	MLP Accuracy: 81% AUC: 0.91 Sensitivity: 0.91 Specificity: 0.8	
Shin et al., 2021 [36]	741	482	259	1.861	solitary	yes	GBM vs. MET		450	48	100	143	multi-center	DL	CNN (2D)	Pathology	DL extracted	CNN (2D, T1CE, T2)—internal Accuracy: 89% AUC: 0.889 Sensitivity: 0.939 Precision: 0.852 CNN—external Accuracy: 85.9% AUC: 0.835 Sensitivity: 0.889 Precision: 0.807
Priya et al., 2021 [37]	120	60	60	1.000	no	yes	GBM vs. MET		nested cross-validation		no	single center	ML	Linear (LASSO, Elastic Net) and logistic regression, NNW, SVM, MLP, RF, Adaboost	Clinico-Radiological	Shape, First-order statistics, Texture	LASSO (T1, T1CE, T2, FLAIR, ADC) Accuracy: 89.2% AUC: 0.953 Sensitivity: 0.887 Specificity: 0.897	
de Causans et al., 2021 [38]	180	92	88	1.045	multiple, largest selected for classification	yes	GBM vs. MET		143 (71 GBM, 72 BM)	nested cross-validation (10 repeated 5-fold CV)	37 (21, 16)	no	multi-center	ML	LogReg (Voe-Johnson scaling features)	Pathology	Shape, First-order statistics, Texture	LogReg (T1CE) Accuracy: 80% Sensitivity: 0.75 Specificity: 0.86
Liu et al., 2021 [39]	268	140	128	1.094	solitary	yes	GBM vs. MET		208 (110 GBM, 98 BM)	10-fold cross-validation	60 (30, 30)	no	single center	ML	RF, DT, LogReg, Adaboost, Gaussian processing, SVM	Pathology	Shape, First-order statistics, Texture, Higher-order: Wavelet-transformed, Laplace of Gaussian	Random Forest (Boruta selection) (T1CE) Accuracy: 85% AUC: 0.97 Sensitivity: 0.84 Specificity: 0.93
Samani et al., 2021 [40]	136	86	50	1.720	no, 3 patients with multifocal metastasis	yes	GBM vs. MET		108 (66 GBM, 40 BM)	5-fold cross-validation	30 (20, 10)	no	single center	DL	2D CNN	Pathology	Diffusion	CNN (2D, DTI, FW-VP map)—patch wise Accuracy: 85% AUC: 0.9 Sensitivity: 0.87 Specificity: 0.81 CNN—majority vote, subject-wise Accuracy: 93%

GBM= Glioblastoma; MET = Brainmetastasis; PCNSL = Primary central nervous system lymphoma; MEN = Meningioma; MED = Medulloblastoma; CV = Cross-validation; LOOCV = Leave-One-Out cross-validation; ML = Machine learning; DL = Deep learning; T1CE = contrast-enhanced T1-weighted sequence; DWI = Diffusion weighted imaging; DTI = Diffusion tensor imaging; PWI = Perfusion weighted imaging; rCBV = relative cerebral blood volume; FLAIR = Fluid-attenuated inversion recovery; TE = Time to echo; AUC = Area under the receiver operating characteristic curve; ADC = Apparent diffusion coefficient; LASSO = Least absolute shrinkage and selection operator; SVM = Support vector machine; MLP = Multilayer perceptron; NNW = Neural networks; LogReg = Logistic Regression; DNN = Deep neural network; LDA = Linear discriminant analysis; NB = Naïve Bayes; VFI = Voting feature intervals; KNN = k-nearest neighbors; PNN = Probabilistic neural networks; RF = Random Forest; RBF = Radial basis function kernel; n/a = not available

Considering major differences between studies, direct comparison of their results yields little to no information. Nonetheless, in this systematic review, there was no

perceived difference between models utilizing conventional MRI sequences versus those using advanced techniques. Support Vector Machines yielded the highest mean area under the receiver operating characteristic curve, yet there was no statistically significant differences between classifiers used including deep learning models. It is noteworthy however, that while most of these studies included highly imbalanced datasets, balanced accuracy instead of accuracy was rarely reported. Accuracy could be significantly biased in this context. These comparisons should be cautiously interpreted since they stem from a systematic review rather than from a meta-analysis, which was impossible to perform due to missing required criteria. External validation was performed in two of the studies conducted by researchers from the same institution (10,27). While very interesting, these two studies employed a 2D CNN, and no consideration was taken with regard to the size of the lesions.

The TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) (38) initiative published in 2015 reporting guidelines for development and validation of prediction models in the medical field in an effort for standardization. In the published studies above, median TRIPOD adherence score was 48%, ranging from 17 to 79%. Full model specification and model performance were among the items with lowest scores, despite being crucial for model evaluation, interpretability, and reproducibility. Table 1.2 shows the TRIPOD checklist items for model development.

**Table 1.2** TRIPOD Checklist Items for model development. Reprinted from <https://www.tripod-statement.org/resources/>

Section/Topic	Item	Checklist Item	Page
<b>Title and abstract</b>			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	
<b>Introduction</b>			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	
<b>Methods</b>			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	
	5b	Describe eligibility criteria for participants.	
	5c	Give details of treatments received, if relevant.	

Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	
	6b	Report any actions to blind assessment of the outcome to be predicted.	
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	
Sample size	8	Explain how the study size was arrived at.	
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	
Risk groups	11	Provide details on how risk groups were created, if done.	
<b>Results</b>			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	
Model development	14a	Specify the number of participants and outcome events in each analysis.	
	14b	If done, report the unadjusted association between each candidate predictor and outcome.	
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	
	15b	Explain how to use the prediction model.	
Model performance	16	Report performance measures (with CIs) for the prediction model.	
<b>Discussion</b>			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence.	
Implications	20	Discuss the potential clinical use of the model and implications for future research.	
<b>Other information</b>			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	
Funding	22	Give the source of funding and the role of the funders for the present study.	

Many of the previously published studies have included multiple brain metastases in their analysis as shown in Table 1.1. We believe that this practice leads to highly significant bias since gliomas are more likely to occur as a single lesion, and thus the model might rely on multifocality to guide the prediction results, rather than intrinsic imaging characteristics of the tumors.

Brain metastases are frequently diagnosed at a smaller size than high-grade gliomas, mainly due to the frequent surveillance imaging in cancer patients by contrast to rather previously healthy individuals with gliomas. This fact leads naturally to size imbalance between these two tumor categories in any real-life dataset. The vast majority of the

published studies did not report on tumor size and did not include size as an eligibility or exclusion criterion. We believe that this could significantly bias the classification performance. In the de Causans *et al.* (15), authors restricted their included population to patients with tumors larger than 2cm to mitigate this concern. They used a total number of 100 radiomic features, and their best performing model was a Logistic Regression model with an 80% balanced accuracy, numerically lower than most published data not considering size as a factor. Of note, their dataset consisted of MRI images generated on the same 3-Tesla machine using a unified imaging protocol. While data homogeneity boosts the model performance, the ability of the model to generalize to external datasets using different imaging techniques might be negatively impacted.

Although no difference in performance between machine learning techniques was identified in the systematic review by Jekel *et al* (37), one cannot conclude that no difference exists particularly due to the fact that no direct comparison was possible between studies. However, individual studies have shown that deep learning methods might yield a better performance than basic machine learning techniques (27).

More recently, we have been witnessing a surge in quantum technology developments, with more scalable quantum devices, better error-correction methods, and rapid increase in quantum computational algorithms. This has led to the conception of quantum machine learning, a bridging field that uses quantum algorithms and technology for machine learning applications (39–42). For specific tasks and under specific circumstances, there might be an advantage for quantum machine learning over its classical counterpart (43–47). For instance, research in the field of high-energy physics hinted to a better performance of quantum algorithms for small datasets (46). Furthermore, quantum machine learning models seem to generalize better for small training datasets (47). Particularly for medical applications where the number of patients is limited, exploring quantum algorithms could be interesting due to these reasons.

In addition to quantum machine learning, quantum technology offers an advantage as well for feature selection. By using quantum annealing, it is possible to implement a highly efficient feature selection based on mutual information. This method is able to capture non-linear relationships between variables (48), and is intractable to solve on a classical computer. For such combinatorial problems, many classical heuristics have been developed and are commercially available. However, it has been shown that quantum annealing provides an efficient solution that sometimes outperforms classical ones (49). In this context, exploring this novel implementation of mutual information feature selection could further enhance the performance of classification models.

Moreover, one of the major bottlenecks for clinical implementation of machine learning models is interpretability. It is hard for physicians to reconcile with

predictions that would impact the clinical decisions for patients without being able to understand the chain-of-thought that led to them. Deep radiomics models are naturally thought of as “black-box” models despite many efforts to decipher their mechanism. Handcrafted radiomics offer some advantages in this space, i.e. they could be, at least partially, explainable mathematically. Published models have limited explainability, and despite some efforts to rate the features by order of importance, the use of a very large number of features introduces much variability and uncertainty. Furthermore, an “average” working mechanism could not be extrapolated to individual predictions, and thus remains hard to provide sufficient confidence for clinical implementation. Therefore, feature selection methods for dimensionality reduction, as well as instance-wise model interpretability are key for “bench-to-clinic” transitioning.

In summary, although the use of machine learning to differentiate between brain metastases and high-grade gliomas has been extensively studied, there are significant limitations in individual studies and model development strategies, and this leaves room for further refinement and technique exploration before adoption in the clinical setting.

### **1.2.2 Detection of brain metastases on non-enhanced CT**

Despite the extensive work on computer-vision and machine learning to detect brain metastases on MRI images (50), detecting brain metastases on contrast-enhanced CT scans has been the subject of only one study by Kato *et al* (51). The same group published as well the only study on non-enhanced CT scans (52). They included 116 patients with brain metastases, and they developed a Single-Shot Detector model to detect individual metastases on non-enhanced CT scans. They used contrast-enhanced CT scans as ground-truth and performed a 2D slice-wise technique for metastasis detection.

They reported an overall detection sensitivity of only 23.8%, with an extremely low sensitivity of 5.4% for lesions measuring 3 to 6mm, and no detection at all for lesions less than 3mm. For lesions larger than 6mm, the sensitivity remained very low at 35.4%. These results highlight the tremendous challenges, and probably the impossibility to detect individual metastases on non-enhanced CT scans. Furthermore, using contrast-enhanced CT scans as ground-truth might have missed some metastases, probably leading to an over-estimation of the true sensitivity.

While detecting individual metastases on a non-enhanced CT scan is scientifically sound and interesting, we believe that its clinical relevance is very limited, i.e. no meaningful information to guide management, prognostication, or treatment planning

could be safely obtained without further characterization using more sensitive and better-quality imaging. We believe that identifying patients that have brain metastases hard to detect on non-enhanced CT scans is a more interesting task, since these patients would benefit from further imaging to receive proper personalized management in a timely manner.

In the paper by Kato *et al* (52), all patients had brain metastases, and they reported a false positive rate of 15.4 lesions per scan. Consequently, their model was clearly optimized to detect lesions in a context where lesions exist, and thus its ability to identify patients without brain metastases is highly questionable.

In this scope, there is an unmet need to develop more accurate tools to identify patients with a high likelihood of having brain metastases out of non-enhanced CT scans that are widely available and suboptimally exploited.

## 1.3 Objectives

In the context of previous research in this field, this work has two main objectives:

- 1- **Develop an explainable machine learning tool based on conventional MRI images, specifically contrast-enhanced T1-weighted sequences, to differentiate between large brain metastases and high-grade gliomas.**

For this purpose, we will address limitations in existing literature. We aim to restrict our work on tumors larger than 2cm. We will use handcrafted radiomics with state-of-the-art feature selection techniques, as well as instance-wise Shapley values to leverage our model explainability. We will explore quantum technology, namely quantum annealing for feature selection and quantum machine learning for model building and benchmark our model against state-of-the-art classical machine learning and deep learning models.

- 2- **Develop a deep learning model to identify patients with brain metastases based on non-enhanced brain CT images.**

For this purpose, we will train our model on a dataset including patients with and without brain metastases. We will include patients with brain metastases larger than 5mm considering previous research. We will design our model for a patient-wise prediction.



## 2 Technical Overview

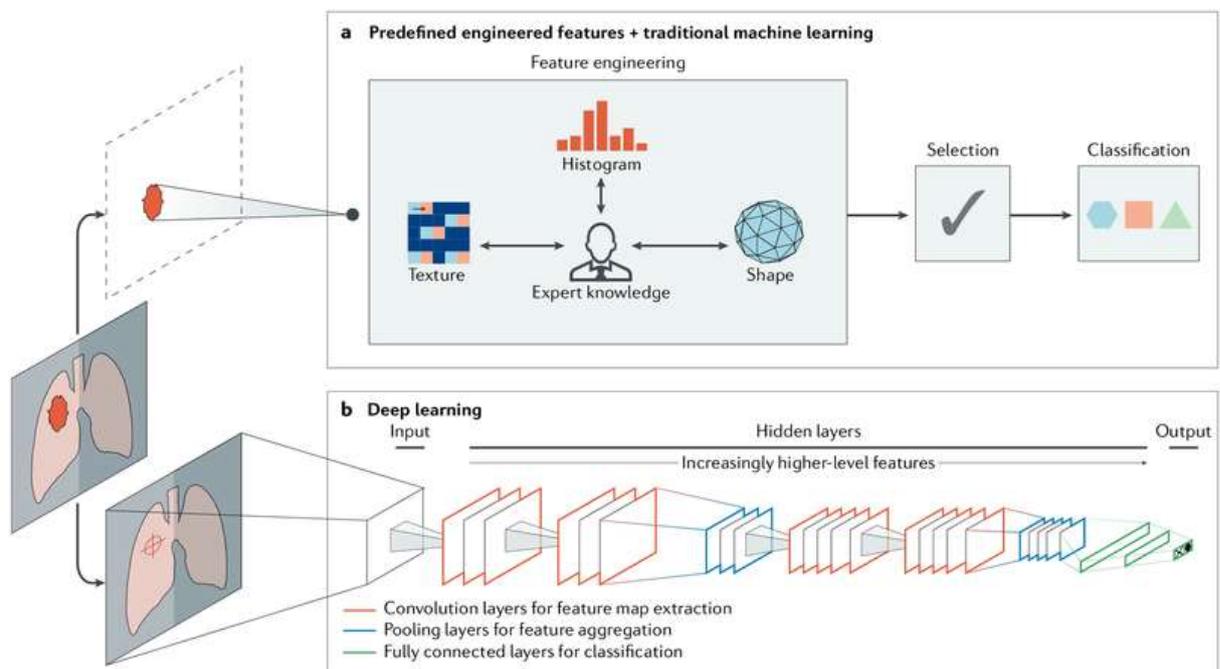
In this chapter, we provide an overview on the technical methods used in this work. It will consist of five parts:

- 1- Radiomics
- 2- Deep Learning
- 3- Quantum Computing
- 4- Quantum Annealing
- 5- Quantum Machine Learning

## 2.1 Radiomics

Radiomics was first introduced by Lambin *et al* (8) in 2012, and consists of extracting quantitative metrics, or "radiomic features," from medical images. These features capture various tissue and lesion characteristics like heterogeneity and shape, and texture. Radiomics can be integrated with demographic, histologic, genomic, or proteomic data for clinical problem-solving. This field is particularly significant in the era of precision medicine, where it is used to identify potential biomarkers from various imaging modalities like X-ray, CT, MRI, PET, and ultrasound. The integration of radiomic features with other data types facilitates a comprehensive assessment for diagnosis and personalized treatment planning.

Two distinct types of Radiomics exist: Handcrafted Radiomics and Deep Radiomics. Figure 2.1 shows the differences in workflow between these two subtypes.



**Figure 2.1** Differences in workflow between handcrafted radiomics and deep radiomics. Reprinted with permission from Springer Nature from Hosny *et al*, 2018 (53)

### 2.1.1 Handcrafted Radiomics:

These features are typically extracted using radiomics software. The process involves registration, segmentation, pre-processing images, followed by extracting radiomic features using specialized software that relies on pre-defined mathematical functions. These features are then typically used as inputs for machine learning model development or multivariable statistical analysis.

For a detailed description of the Radiomics features used in this work, please refer to (54): <https://pyradiomics.readthedocs.io/en/latest/>

In addition to first order statistical features and shape features, there are texture features matrices as described in the following:

- 1- The Gray-Level Cooccurrence Matrix (GLCM), initially introduced by Haralick *et al.* (55), is a form of second-order gray-level histogram, analyzes spatial correlations between pairs of pixels/voxels. These correlations are defined based on their gray-level intensities, orientations, and the distance between them. Key features of GLCM include entropy, which quantifies the inhomogeneity of gray levels; energy, indicating gray-level homogeneity; and contrast, highlighting the differences in gray levels between pairs of pixels/voxels.
- 2- The Gray-Level Run-length Matrix (GLRLM) offers insights into the spatial arrangement of sequences of adjacent pixels that share an identical gray level. This matrix can analyze these pixel runs in one or more directions and is applicable in both two-dimensional and three-dimensional spaces.
- 3- The Gray-Level Size Zone Matrix (GLSZM) quantifies interconnected pixel or voxel groups, referred to as zones, that exhibit identical gray levels. A texture characterized by more homogeneity will lead to a broader and more level matrix.
- 4- The Gray-Level Distance Zone Matrix (GLDZM), a variant of GLSZM, extends its analysis to include not only zones of connected pixels or voxels with uniform gray levels but also their equal distance from the edge of the region-of-interest.
- 5- The Neighborhood Gray-Tone Difference Matrix (NGTDM) evaluates the cumulative differences between the gray level of a specific pixel or voxel and the average gray level of its adjacent pixels or voxels within a set distance.
- 6- The Neighborhood Gray-Level Dependence Matrix (NGLDM) assesses the gray-level relationships between a central pixel or voxel and its neighboring pixels or voxels. In this matrix, a neighboring pixel or voxel is considered connected to the central one if it falls within a predefined distance and meets the dependence criterion, which is based on a specific range of gray-level differences.

Handcrafted Radiomic features suffer from much variability and instability due to differences in image acquisition, image processing, and extraction techniques. This has led to the Image Biomarker Standardization Initiative (IBSI) in a perspective to provide guidelines for standardization the feature extraction process for the use in medical applications (56). Abiding by these guidelines is highly recommended to avoid significant loss of generalizability of developed models.

### **2.1.2 Deep Radiomics**

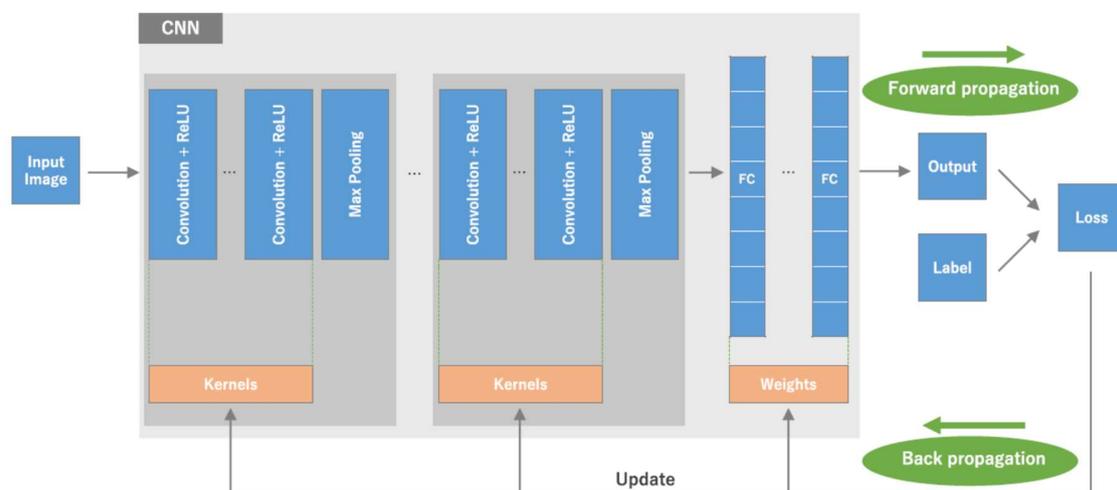
Deep radiomics utilize deep learning algorithms for feature extraction. Deep learning methods have shown tremendous potential in automating tasks like image segmentation, reconstruction, recognition, and classification. Deep radiomics goes beyond conventional feature extraction by employing deep learning techniques to analyze images, addressing limitations such as manual lesion annotation and inadequate feature criteria. These features are typically part of the deep layers of a convolutional neural network. This approach has shown sometimes enhanced performance compared to handcrafted methods (53), however there remains much controversy on the superiority of either of these approaches.

## 2.2 Deep Learning

Deep learning is a subset of machine learning based on representation learning methods, where a network of distributed layers of communication nodes fed initially with raw data, transforms one representation into another through its deep layers, amplifying task-oriented discriminatory aspects of the data through a general-purpose learning process, in order to learn complex functions required for detection or classification (57).

Many types of deep learning architectures exist such as deep belief networks, recurrent neural networks, convolutional neural networks, and transformers, that involve different working mechanisms, each type being adapted to a certain type of tasks. For medical image analysis, convolutional neural networks are by far the most used type of deep learning architecture; hence, we will limit our overview on this specific type.

Convolutional Neural Networks (CNNs) are a specialized type of neural networks used primarily for processing data that has a grid-like topology, such as images. They are particularly well-suited for analyzing visual imagery and have become a cornerstone technique in the field of computer vision. Figure 2.2 shows the typical architecture of a CNN.



**Figure 2.2** Architecture of a Convolutional Neural Network. ReLU: Rectified Linear Unit; FC: Fully Connected. *Reprinted from Yamashita et al. (58), no permission is required, open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, Springer Nature*

Key Components of CNNs:

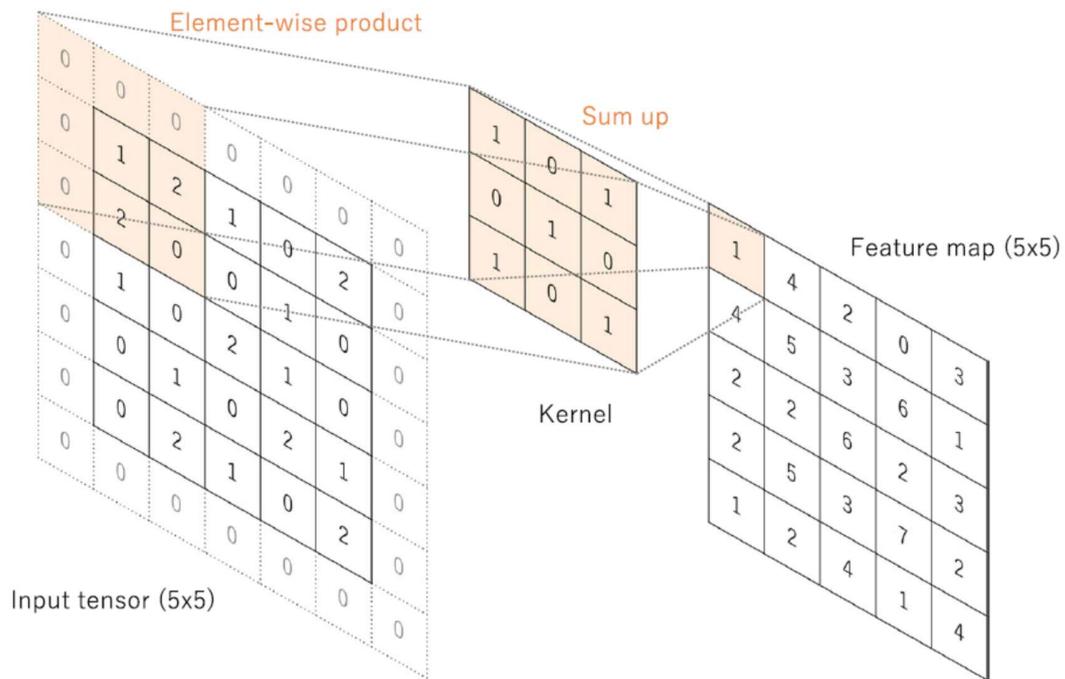
- 1- Convolutional Layers (Figure 2.3):

These layers perform a convolution operation that filters the input data to extract features. They apply a set of learnable filters (kernels) to the input. Each filter activates certain features from the input, like edges or shapes. Convolution involves sliding these filters over the input data and computing the dot product between the filter and the input at each position. Depending on the kernel dimension, there exist 2D and 3D CNNs. 2D CNNs analyze medical images by slice, whereas 3D networks perform tridimensional convolution on the 3D image series directly.
- 2- Activation Function:

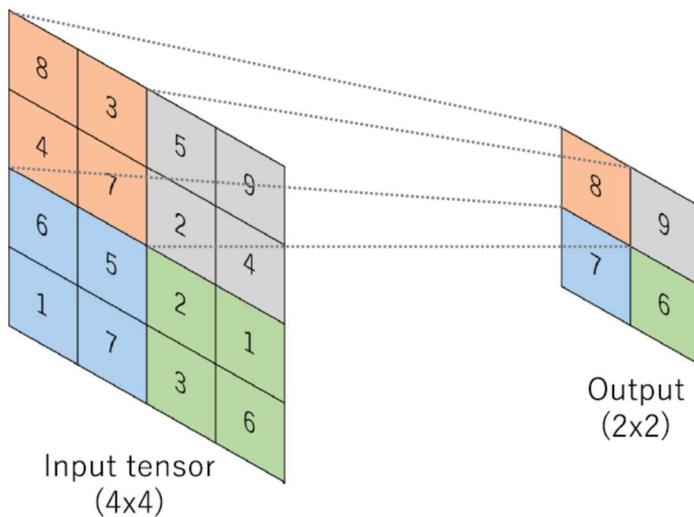
After the convolution operation, the result is passed through an activation function, like the Rectified Linear Unit (ReLU). This function introduces non-linear properties to the system, allowing the network to learn more complex patterns.
- 3- Pooling Layers (Figure 2.4):

Pooling (also known as subsampling or downsampling) reduces the dimensionality of each feature map but retains the most important information. Common pooling methods include max pooling, which returns the maximum value from the portion of the image covered by the filter, and average pooling.
- 4- Fully Connected Layers:

After several convolutional and pooling layers, the high-level reasoning in the neural network is done via fully connected layers. In a fully connected layer, neurons have connections to all activations in the previous layer. These layers are typically placed towards the end of CNN architectures and can be used to optimize the output for the task at hand (like classification).



**Figure 2.3** A convolutional operation with zero padding. Input size is 5x5, Kernel size is 3x3. Reprinted from Yamashita et al. (58), no permission is required, open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, Springer Nature



**Figure 2.4** Max pooling operation. Reprinted from Yamashita et al. (58), no permission is required, open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, Springer Nature

When an image is input into a CNN, it passes through a series of convolutional, non-linear, pooling (downsampling), and fully connected layers. The initial layers detect simple features like edges, and as the data progresses through the layers, more complex features are identified. By the time the information reaches the fully connected layers, the CNN has understood the main features and patterns in the image and can interpret what the image represents.

CNNs present many advantages:

- 1- **Parameter Sharing:** The same filter (weights) is used for each pixel in the layer; this significantly reduces the number of parameters, making the network efficient.
- 2- **Local Connectivity:** Each neuron is connected only to a small region of the input, making the network focus on low-level features in the initial layers. As we move deeper, the layers have a broader view of the input, capturing more abstract features.
- 3- **Translation Invariance:** Once trained, CNNs can recognize objects in an image, regardless of their location in the frame.

Overfitting in CNNs occurs when a model becomes excessively complex and begins to memorize rather than generalize from the training data, resulting in poor performance on new, unseen data. This issue often arises in scenarios with limited training data or when the network has an overly complex architecture with too many parameters relative to the amount of training data available. To mitigate overfitting, several strategies can be employed:

- 1- **Data Augmentation:** By artificially expanding the training dataset using transformations like rotation, scaling, and flipping, data augmentation can effectively increase the diversity of the training set, helping the model generalize better.
- 2- **Dropout:** This technique involves randomly 'dropping out' a subset of neurons during training. This prevents neurons from co-adapting too much and encourages the network to learn more robust features that are useful in conjunction with many different random subsets of the other neurons.
- 3- **Regularization:** Techniques like L1 and L2 regularization add a penalty to the loss function based on the complexity of the model (the magnitude of the weights), discouraging the model from fitting too closely to the training data.
- 4- **Early Stopping:** This involves monitoring the model's performance on a validation set during training and stopping the training process once the model's performance on the validation set stops improving. This prevents the model from learning noise and irrelevant details in the training dataset.

- 5- Simplifying the Model: Reducing the number of layers or the number of neurons in each layer can help prevent overfitting, especially when the amount of training data is limited.
- 6- Using Transfer Learning: Leveraging pre-trained models on large datasets and fine-tuning them on the specific task can help in avoiding overfitting, as these models have already learned a robust set of features.

By implementing these techniques, the robustness and generalization ability of CNNs can be significantly improved, enhancing their performance on unseen data and real-world applications.

## 2.3 Quantum Computing

Quantum computing represents a fundamental shift in our approach to computation, harnessing the principles of quantum mechanics to process information in ways that traditional computers cannot. Unlike classical computers, which use *bits* as the basic unit of information, quantum computers use quantum bits or *qubits*, which can exist in multiple states simultaneously. This unique feature allows quantum computers to perform certain types of calculations much more efficiently than their classical counterparts (59).

### 2.3.1 Fundamentals of Quantum Computing

#### 2.3.1.1 Qubits and Superposition:

Qubits are the heart of quantum computers. Unlike a classical bit, which is either 0 or 1, a qubit can be in a state of 0, 1, or any quantum superposition of these states. Using Dirac notation, a qubit ( $\psi$ ) state is given by:  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , where  $\alpha$  and  $\beta$  are complex numbers representing the probability amplitudes of states 0 and 1 respectively.

The probability of measuring the qubit ( $\psi$ ) in state “0” is given by the Born rule, and is equal to  $|\alpha|^2$ , with  $|\alpha|^2 + |\beta|^2 = 1$ .

The vector representation of each quantum state is:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Also, the basis state of 2 qubits could be represented in a four-dimensional linear vector, the tensor product of the two states. For instance, the basis state of  $|01\rangle$  is  $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ .

So  $n$  qubits map to a  $2^n$  dimensional Hilbert space. This property allows quantum computers to process a vast amount of information simultaneously.

#### 2.3.1.2 Entanglement:

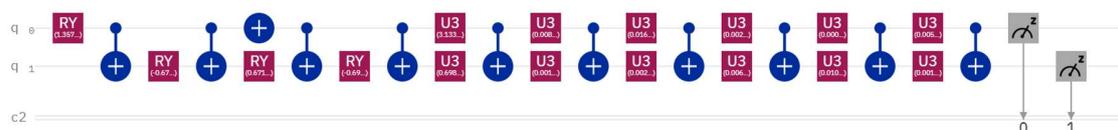
Quantum entanglement is a pivotal phenomenon in quantum computing, representing one of the most profound and unique aspects of quantum mechanics. When qubits are entangled, their quantum states become interdependent, meaning the state of one

cannot be fully described without the state of the other, regardless of the physical distance between them. This entanglement allows for the creation of highly correlated quantum states, which are fundamental for quantum computing. In practical terms, entanglement enables quantum computers to perform complex calculations at unprecedented speeds and efficiencies, using parallelism. It is the basis for quantum teleportation, superdense coding, and quantum key distribution, which are essential for quantum communication and cryptography. The power of quantum computers to solve certain problems much more efficiently than classical computers largely hinges on this phenomenon. However, maintaining entanglement over time and against environmental interference, known as decoherence, is one of the major challenges in the development of quantum computing technology (60).

### 2.3.1.3 Quantum Gates and Circuits:

In the realm of quantum information theory, a quantum circuit functions as a framework for quantum computation. It mirrors the concept of classical circuits, wherein the computation process consists of a series of quantum gates, qubit initializations, measurements, and other possible operations.

In the representation of quantum circuits, the progression of time is depicted along the horizontal axis, commencing from the left and concluding on the right. Qubits are represented by horizontal lines, while classical bits are denoted by double lines (Figure 2.5). These lines are interconnected with various elements that represent operations on the qubits, such as gates or measurements. The sequence of these operations is outlined by these lines, which conceptually represent the event sequence rather than physical connections.



**Figure 2.5** A quantum circuit representing two qubits q0 and q1, and one classical bit c2. It shows a sequence of quantum gates and unitary operations, and ends with a measurement of the two qubits. Image created using IBM Qiskit.

A quantum logic gate, often simply referred to as a quantum gate, is a fundamental element in quantum circuits, analogous to classical logic gates in traditional digital circuits, and it operates on a defined number of qubits. These gates are the primary building blocks for constructing quantum circuits.

One of the key distinctions of quantum logic gates, as opposed to many classical logic gates, is their reversible nature. In terms of their operational mechanics, quantum gates function as unitary operators. They are represented by unitary matrices, which are defined relative to a specific orthonormal basis. This unitary property ensures the reversible nature of quantum gates and is central to the operation of quantum computing.

Figure 2.6 shows the most frequently used quantum gates, along with their matrix representation.

Gate	Notation	Matrix
NOT ( Pauli- $X$ )		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Pauli- $Z$		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Hadamard		$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
CNOT ( Controlled NOT )		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

**Figure 2.6** Most used quantum gates, their respective notations, and matrix representations. Reprinted from Yan et al. (61). No permission required, published under an open access Creative Common CC BY license

### 2.3.1.4 Types of quantum computers

Quantum computing has diversified into several types, each leveraging different physical systems and mechanisms:

- 1- **Gate-based ion trap processors** utilize ions trapped in electromagnetic fields, where quantum gates are implemented using laser pulses. This approach is known for its high accuracy in qubit manipulation and long coherence times, essential for maintaining quantum states.
- 2- **Gate-based superconducting processors** use superconducting circuits cooled to extremely low temperatures. They operate by generating quantum states in microwave cavities, offering fast gate speeds and scalability potential.

- 3- **Photonic processors** use photons as qubits, manipulating them through optical components like mirrors and beam splitters. This method stands out for its potential in creating quantum networks due to photons' suitability for long-distance quantum communication.
- 4- **Neutral atoms processors and Rydberg atom processors** exploit the properties of atoms in highly excited states (Rydberg states). They utilize laser-induced interactions between atoms to perform quantum operations, offering a promising path towards scalability and high-fidelity operations.
- 5- **Quantum annealers**, different from universal quantum computers, are designed for solving optimization problems. They use quantum tunneling and superposition to find the minimum of a complex landscape, making them suitable for specific tasks like optimization and material science simulations.

### 2.3.1.5 Applications of Quantum Computing

Quantum computing offers groundbreaking potential in various fields. Its unique capabilities, such as handling complex calculations at incredible speeds, promise to transform industries ranging from cryptography to pharmaceuticals. We describe in the following some promising applications of quantum computing:

- 1- **Cryptography and Cybersecurity:**  
One of the most immediate applications of quantum computing lies in the field of cryptography. Quantum computers pose a significant threat to current cryptographic algorithms. Shor's algorithm, for instance, can factorize large numbers exponentially faster than classical computers, rendering current encryption methods like RSA vulnerable (62). On the flip side, quantum computing also provides solutions to these vulnerabilities through quantum key distribution (QKD), which offers theoretically unbreakable encryption. This dual role highlights the urgency in developing quantum-resistant cryptographic algorithms and the potential of quantum cryptography in enhancing cybersecurity (63).
- 2- **Drug Discovery and Materials Science:**  
In pharmaceuticals and materials science, quantum computing presents a transformative approach to molecular modeling and drug discovery. Traditional computers struggle to simulate complex molecular interactions due to the exponential scaling of quantum states in molecules. Quantum computers can handle these simulations more efficiently, potentially speeding up the discovery of new drugs and materials. Algorithms like the Variational Quantum Eigensolver (VQE) enable researchers to study molecular structures at an

unprecedented level of detail, fostering innovations in drug development and material science (64).

3- Optimization Problems:

Quantum computing also excels in solving complex optimization problems, which are prevalent in logistics, finance, and artificial intelligence. Quantum annealers, like those developed by D-Wave Systems, have been applied to optimization problems in various industries. These systems use quantum properties to explore numerous potential solutions simultaneously, identifying optimal solutions more efficiently than classical algorithms in specific scenarios (65).

4- Artificial Intelligence and Machine Learning

In artificial intelligence (AI) and machine learning, quantum computing offers the potential for significant advancements. Quantum algorithms can potentially process and analyze large datasets much faster than classical algorithms, enhancing the capabilities of machine learning models. Quantum machine learning algorithms, such as quantum neural networks, promise to improve pattern recognition and decision-making processes in AI systems, with applications ranging from natural language processing to autonomous vehicles (40,66).

5- Climate Modeling and Environmental Research

Quantum computing's ability to handle complex systems can significantly impact climate modeling and environmental research. Traditional models often struggle with the vast and intricate data involved in climate simulations. Quantum computers, with their superior processing power, can potentially model climate change scenarios with greater accuracy and detail, aiding in better prediction and understanding of environmental changes (67).

6- Financial Modeling

In finance, quantum computing can revolutionize risk assessment, portfolio optimization, and algorithmic trading. Quantum algorithms can analyze market data and simulate financial scenarios with a level of complexity and speed unattainable by classical computers. This capability could lead to more efficient markets, better investment strategies, and more robust risk management practices (68).

## 2.4 Quantum Annealing and Adiabatic Quantum Computing

### 2.4.1 The Ising model and QUBO

The Ising model, named after the physicist Ernest Ising, is a theoretical framework in statistical mechanics for explaining ferromagnetism. It is characterized by discrete variables known as spins possessing binary states  $\sigma_x \in \{-1, +1\}$ . These spins are organized on a lattice, facilitating interactions among adjacent spins. The model's primary objective is to elucidate the large-scale magnetic behaviors of materials by examining the atomic-scale interactions, marking it as a pivotal tool in studying phase transitions. The energy of a lattice configuration is given by the Hamiltonian:

$$H(\sigma) = - \sum_{\langle ij \rangle} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i$$

Where  $i$  and  $j$  represent adjacent sites with an interaction factor  $J_{ij}$ , and  $h_i$  denotes a transverse field strength acting on site  $i$ .

Quadratic Unconstrained Binary Optimization (QUBO) is a mathematical formulation that represents a large variety of combinatorial optimization problems that are considered NP-Hard problems. QUBO has been used in many applications including problems in finance, economics, and machine learning. The cost function of a QUBO for a set of  $n$  binary values  $x_k \in \{0, 1\}$  is very similar to an Ising model Hamiltonian and given by:

$$f(x) = \sum_{i < j}^n Q_{ij} x_i x_j + \sum_i^n Q_{ii} x_i$$

Where  $Q$  denotes the weight assigned to a pair of indices.

The similarity between QUBO formulation and the Ising model Hamiltonian has led to the increased utilization of Ising machines such as Quantum Annealers in solving combinatorial problems.

### 2.4.2 Adiabatic Quantum Computing

The evolution of a system of qubits is governed by the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = H |\Psi(t)\rangle$$

where  $t$  is time,  $|\Psi(t)\rangle$  is the state vector of the quantum system at time  $t$ , and  $H$  is the Hamiltonian operator acting on the corresponding state.

This implies that the temporal evolution of the quantum system state depends on the evolution of the Hamiltonian.

The Adiabatic Theorem in quantum mechanics, first formulated by Max Born and Vladimir Fock in 1928, posits that “a quantum system remains in its instantaneous eigenstate if the perturbations acting upon it are slow enough and there is a clear energy gap between the eigenvalue and the rest of the Hamiltonian's spectrum” (69). Essentially, under slow gradual changes (adiabatic changes), a system that starts in one of its eigenstates will remain in a corresponding eigenstate of the gradually evolving Hamiltonian.

Adiabatic Quantum Computing was introduced by Farhi *et al.* in 2000 (70). The adiabatic quantum algorithm consists of slowly evolving a quantum system over a time  $T$ , initialized in its ground state with a Hamiltonian  $H_0$ , to a final Hamiltonian  $H_f$  encoding the desired QUBO problem. This ensures the system stays in its ground state at the end of the evolution, thus getting the global minimum of the desired optimization. Hamiltonian evolution would thus be given by:

$$H(t) = (1 - t/T)H_0 + t/T H_f$$

The time  $T$  required for evolving the system in the ground state is inversely proportional on the square of the energy gap between the ground state and the lowest excited state. For hard optimization problems, the energy gap typically grows exponentially small, thus requiring a practically infinite time  $T$  for adiabatic transformation.

### 2.4.3 Quantum Annealing

Abiding by the requirements of adiabatic quantum computing implies respecting the minimum transformation time, which is very hard to calculate, as well as eliminating any external and internal interference with the quantum system. These conditions are extremely hard to meet practically.

Quantum annealing (71) provides a more realistic way to solve the same problem, in rather less perfect conditions, and thus providing a practical way to reach a solution very close to the sought minimum, in a non-universal non-deterministic technique.

It consists of applying the adiabatic quantum algorithm more rapidly, and repeating the process for a large number of iterations to increase the probability of reaching the

global minimum of the desired Hamiltonian. Decreasing the overall evolution time has been shown to increase the probability of reaching the global minimum in a paper by Crosson *et al.* (72).

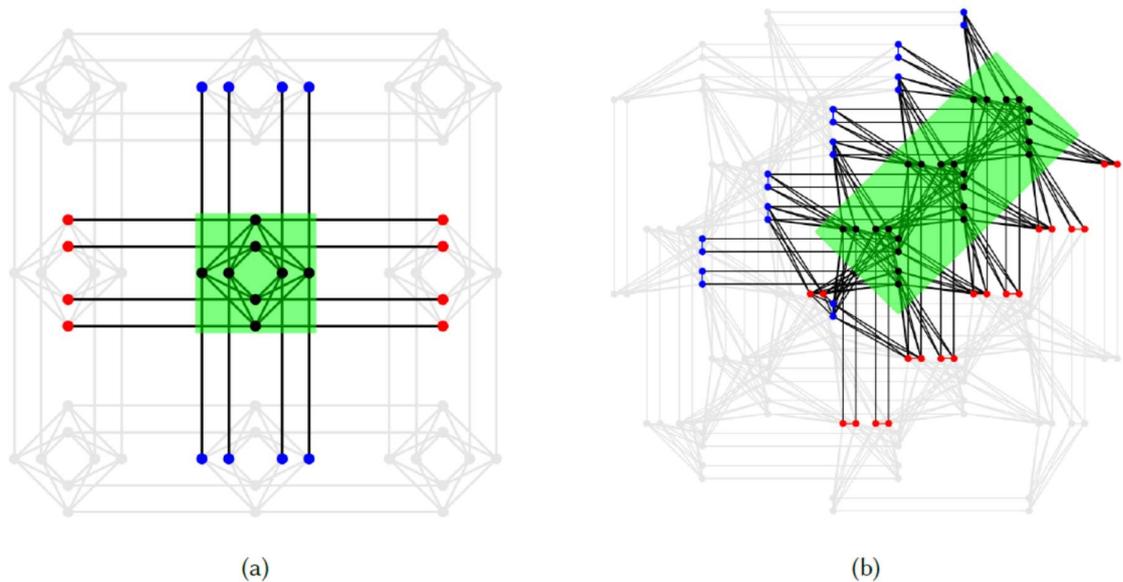
The quantum annealing heuristic closely mirrors the classical simulated annealing technique in optimization, using quantum fluctuations instead of thermal ones (73). In quantum annealing, the strength of the transverse field applied initially to the qubit system and then gradually decreased, also named tunneling field, governs the quantum tunneling effect that hypothetically provides advantage over classical simulated annealing in terms of transitioning the qubit states from local minima towards the global ground state of the system, even in the presence of high energy barriers between these states (74).

#### **2.4.3.1 Hardware implementation on D-Wave systems**

Implementing a QUBO optimization on D-Wave hardware entails several steps that we will describe briefly hereafter:

- 1- QUBO formulation: the first step is to write the optimization problem in a QUBO format that maps to an Ising model in order to represent it in a graph format to facilitate embedding
- 2- Minor embedding: this is the technique of aligning the logical structure of a problem with the physical architecture of a quantum processor. It necessitates grouping several physical qubits to depict a single logical variable and methodically linking them to emulate the problem's interaction dynamics. This process is carried out keeping in mind the physical constraints and unique attributes of the hardware's qubits, aiming to authentically capture the problem's framework and limitations within the quantum system.

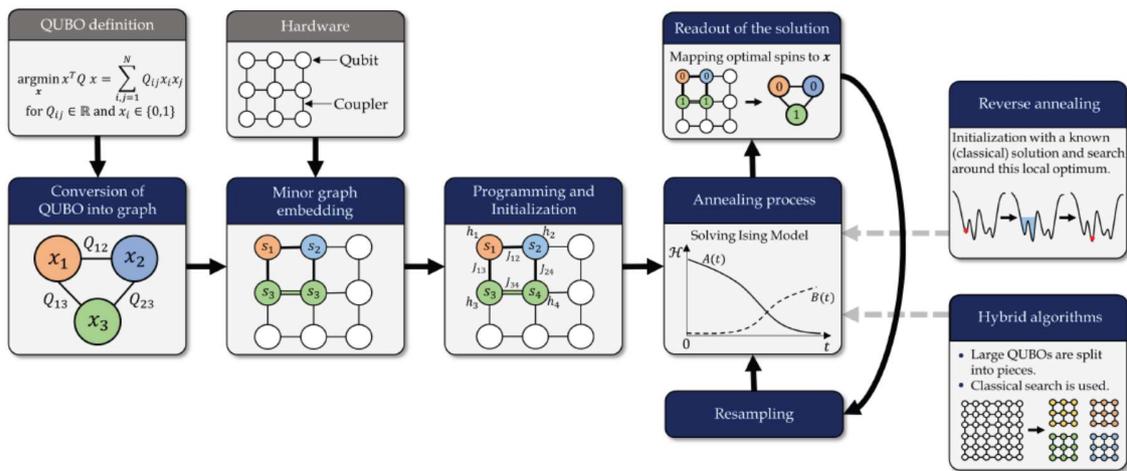
On D-Wave hardware, the qubits are arranged and connected in either the Chimera or the Pegasus topologies, with Pegasus featuring more connectivity and coupling potential between the qubits (fully connecting 6x6 qubits versus 4x4 qubits in Chimera). Representative graphs are shown in Figure 2.7. D-Wave is currently developing a more complex Zephyr topology which allows an 8x8 qubit coupling.



**Figure 2.7** Graph representation of D-Wave's Chimera (a) and Pegasus (b) topologies. The green highlighted area depict a unit cell in each graph. Red and blue dots respectively denote horizontal and vertical connections to qubits in neighboring unit cells. *Reprinted from Yarkoni et al. (75); No permission required, published under a Creative Common CC BY license*

- 3- Problem-solving process: this involves a sequence of processes, starting by initializing the quantum system in a superposition of quantum states which represents a ground state configuration, and applying a transverse tunneling field on the qubits. The next step is evolving the system using predefined parameters to solve the Ising Hamiltonian. This step could be enhanced using additional techniques such as reverse annealing and hybrid search. After finishing the annealing process, the qubit states are measured, and these values represent the solution of the optimization problem.
- 4- Resampling: the problem-solving process is repeated a large number of times, and the best solution, i.e. with the lowest energy, is kept as the best solution to the heuristic.

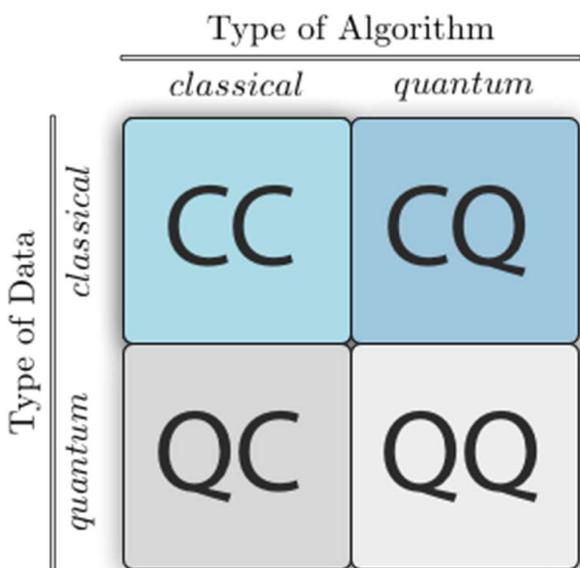
These steps are summarized in Figure 2.8.



**Figure 2.8** Different steps in implementing a QUBO optimization on D-Wave systems hardware. Reprinted from Yarkoni et al. (75); No permission required, published under a Creative Common CC BY license

## 2.5 Quantum Machine Learning

Quantum machine learning (QML) is a bridging discipline that combines principles from quantum computing and learning information theory. There are four combinations of machine learning and quantum computing: CC (classical data with classical processing), QC (quantum data with classical processing), CQ (classical data with quantum processing), and QQ (quantum data with quantum processing), as shown in Figure 2.9. Each scenario utilizes different aspects of quantum theory and machine learning. CC adopts quantum-inspired approaches for traditional machine learning, QC leverages machine learning to analyze quantum systems, CQ employs quantum computing for classical data analysis, and QQ involves quantum devices processing quantum-generated data. These integrations vary from utilizing quantum insights in classical algorithms to fully quantum methods for data handling and processing.



**Figure 2.9** The four combinations of machine learning relative to the type of data and computing algorithm. *Reprinted from (76) by Maria Schuld; No permission required, shared under a Creative Commons CC BY-SA 4.0 license.*

For this thesis, we will focus on the CQ case, denoting the use of quantum algorithms for classical data manipulation.

We can basically distinguish between two extreme types of QML algorithms that bound a spectrum of techniques, those that require fault-tolerant quantum devices, and those

that could be implemented on a near-term quantum device in the Noisy Intermediate Scale Quantum (NISQ) era.

Many quantum algorithms have been developed in a perspective to achieve a quantum advantage that could be used in the field of machine learning. To name a few, Grover's algorithm (77) that relies on quantum amplitude amplification, quantum phase estimation (78) that lies at the heart of Peter Shor's prime factorization algorithm, and Harrow–Hassidim–Lloyd (HHL) algorithm named after its developers that provides theoretically an exponential speedup in solving systems of linear equations (79). Table highlights some machine learning applications that would theoretically derive a quantum advantage of select quantum subroutines, with the corresponding speedups with respect to their classical counterparts. As we can see, these algorithms offer quadratic to exponential speedup advantage.

**Table 2.1** Examples of machine learning subroutines that benefit from quantum algorithms speedups. Complexities are shown with respect to classical algorithms. qRAM: quantum Random Access Memory. *Reprinted with permission from Springer Nature, from Biamonte et al. (80)*

Method	Speedup	Amplitude amplification	HHL	Adiabatic	qRAM
Bayesian inference <sup>106,107</sup>	$O(\sqrt{N})$	Yes	Yes	No	No
Online perceptron <sup>108</sup>	$O(\sqrt{N})$	Yes	No	No	Optional
Least-squares fitting <sup>9</sup>	$O(\log M)^*$	Yes	Yes	No	Yes
Classical Boltzmann machine <sup>20</sup>	$O(\sqrt{N})$	Yes/No	Optional/No	No/Yes	Optional
Quantum Boltzmann machine <sup>22,61</sup>	$O(\log M)^*$	Optional/No	No	No/Yes	No
Quantum PCA <sup>11</sup>	$O(\log M)^*$	No	Yes	No	Optional
Quantum support vector machine <sup>13</sup>	$O(\log M)^*$	No	Yes	No	Yes
Quantum reinforcement learning <sup>30</sup>	$O(\sqrt{N})$	Yes	No	No	No

\*There exist important caveats that can limit the applicability of the method<sup>51</sup>.

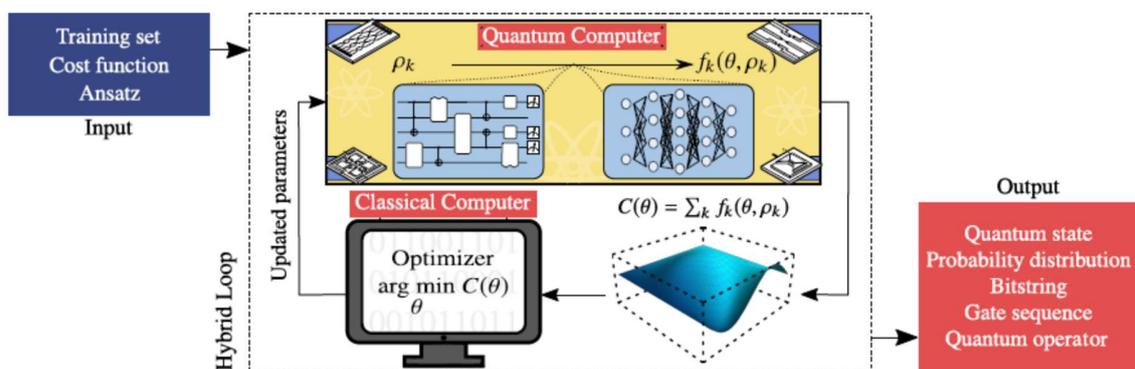
While these algorithms might revolutionize the field of machine learning, their implementation is not possible on near-term quantum devices. As an example, a practical implementation of the HHL algorithm to provide an advantage over classical

algorithms would require a quantum circuit width between 340 and  $10^8$ , and a depth between  $10^{25}$  and  $10^{29}$  quantum gates (81). Although we are witnessing a rapid surge in quantum device development and quantum error-correction algorithms, we are far from being able to perform such implementations in the near future.

Considering these limitations and accompanying skepticism, there has been a strategic shift in recent years in quantum software development, with a trend towards focusing on algorithms that could be implemented on near-term devices. In this scope, foregoing the idea of quantum speedup and rather focusing on the ability of quantum computers to process data differently has led to the emergence of a different subset of algorithms that treats the framework of quantum computation as a machine learning model, or as a part of it.

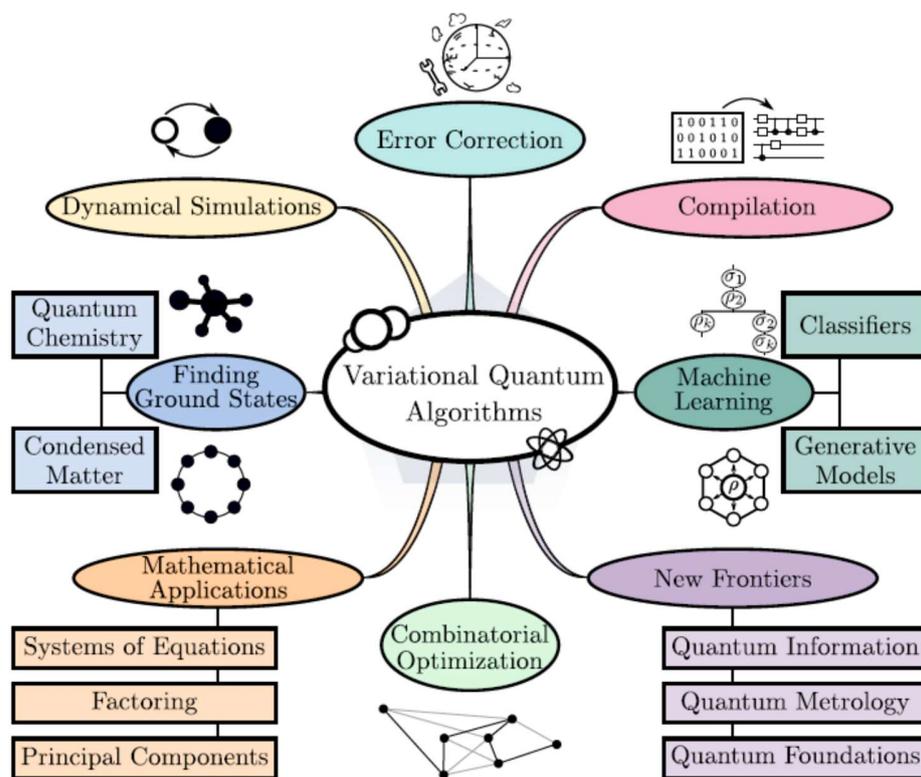
Variational Quantum Algorithms (VQAs) have emerged as a prominent method to harness the potential of NISQ devices. These algorithms, akin to machine-learning techniques, utilize a hybrid quantum-classical approach. They employ parameterized quantum circuits, optimized by classical algorithms, to navigate the limitations of NISQ technology effectively. This strategy allows for shallow circuit depth, reducing noise and making VQAs a promising practical solution for early quantum advantage.

VQAs' general framework is shown in Figure 2.10. It consists of encoding classical data into the circuit of a quantum device, and then applying a template of parametrized quantum gates known as *ansatz*, that could be repeated several times. The parameters are then classically optimized using a predefined cost function, and updated again into the quantum circuit until an optimal solution is found.



**Figure 2.10** Variational Quantum Algorithm framework, showing a set of training data  $\{\rho_k\}$  transformed using an ansatz with a function  $f_k(\theta, \rho_k)$  with set of parameters  $\theta$ . Parameter optimization is classically performed with a cost function  $C(\theta)$ , and then parameters are updated in the quantum circuit. In the figure, the ansatz is shown as a quantum circuit and as a neural network to highlight the analogy. A variety of output data types could be obtained, with the most prominent examples shown in the red box. *Reprinted with permission from Springer Nature from Cerezo et al. (82)*

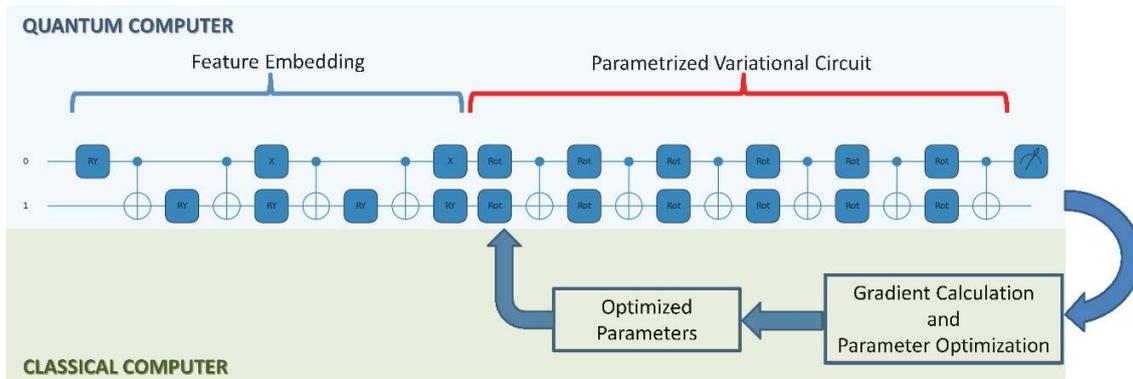
A variety of applications for VQAs has been explored, from quantum chemistry to combinatorial optimization to machine learning. The main arenas are showcased in Figure 2.11.



**Figure 2.11** Key applications for Variational Quantum Algorithms. *Reprinted with permission from Springer Nature from Cerezo et al. (82)*

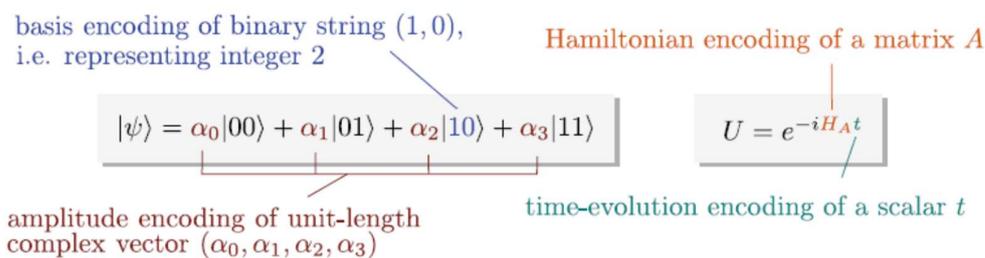
The widely known Variational Quantum Eigensolvers (VQEs) are mainly used in quantum chemistry and used for ground state energy estimation of a molecule (64). VQE in this setting consists of minimizing the expectation value of a Hamiltonian  $H$  encoding the interactions of the studied molecular system instead of finding the ground state. This operation approximates the ground state energy by preparing a state  $|\psi(\theta)\rangle$  using an ansatz with a set of parameters  $\theta$ , and then using the VQE to minimize the expectation of the Hamiltonian  $\langle\psi(\theta)|H|\psi(\theta)\rangle$ .

For classification tasks, Variational Quantum Classifiers (VQCs) have been developed using the same VQA architecture. An example of a VQC is shown in Figure 2.12.



**Figure 2.12** Anatomy of a two-qubits Variational Quantum Classifier

The first step in VQC is feature encoding or embedding, which entails different techniques to incorporate the classical input data into the quantum circuit. Many techniques for feature embedding may be used, namely basis embedding, amplitude embedding, Hamiltonian embedding, and time-evolution embedding. Figure 2.13 shows the different strategies for encoding on a 2-qubits system and a general unitary.



**Figure 2.13** Different encoding techniques, showing basis and amplitude encoding on a 2-qubit system, and Hamiltonian and time-evolution embeddings on an arbitrary unitary. *Reprinted with permission from Springer Nature from Schuld and Petruccione (40)*

Following feature embedding, a parametrized unitary is applied - the ansatz - which corresponds to a linear transformation of the prepared state. The readout of the first qubit typically takes place then, and its measurement statistics correspond to a non-linear continuous output of the system. Then the result is binarized using a threshold function with the possibility of introducing a bias parameter.

The parameters are optimized on classical computer to minimize the cost function, and these are updated in the ansatz, and the process is repeated to perform the training until an optimal solution is reached. Strongly entangling ansatz are believed to yield

theoretically better results. This VQC architecture mimics in some ways the classical neural networks framework.



### **3 An Explainable MRI-Radiomic Quantum Neural Network to Differentiate Between Large Brain Metastases and High-Grade Glioma Using Quantum Annealing for Feature Selection**

This chapter presents a quantum machine learning approach for brain tumor classification. The core of this research was the development of a Quantum Neural Network model to differentiate between brain metastases and high-grade gliomas on contrast-enhanced T1-weighted brain MRI images. The model was benchmarked against two classical machine learning models: extreme gradient boosting (XGBoost) and dense neural network (DNN). We evaluated the model interpretability using Shapley values. This research paves the way for further exploration and development of quantum machine learning algorithms in oncology.

This work has been published in the *Journal of Digital Imaging* in July 2023 (1) and has been modified herein to a minimal extent. A supplementary appendix has been added at the end of this chapter showing Figures not included in the published manuscript.

## 3.1 ABSTRACT

### Purpose

High-grade gliomas (HGG) and solitary large brain metastases (LBM) are sometimes hard to differentiate on MRI. The management differs significantly between these two entities, and non-invasive methods that help differentiate between them are eagerly needed to avoid potentially morbid biopsies and surgical procedures. We explore herein the performance and interpretability of an MRI-radiomics Variational Quantum Neural Network (QNN) using a quantum-annealing Mutual-Information (MI) feature selection approach.

### Materials and methods

We retrospectively included 423 patients with HGG and LBM (>2cm) who had a contrast-enhanced T1-weighted (CE-T1) MRI between 2012 and 2019. After exclusion, 72 HGG and 129 LBM were kept. Tumors were manually segmented, and a 5mm peritumoral ring was created. MRI images were pre-processed, and 1813 radiomic features were extracted.

A set of best features based on MI was selected. MI and conditional-MI were embedded into a quadratic unconstrained binary optimization (QUBO) formulation that was mapped to an Ising-model and submitted to D'Wave's quantum annealer to solve for the best combination of 10 features.

The 10 selected features were embedded into a 2-qubits QNN using PennyLane library. The model was evaluated for balanced-accuracy (bACC) and area-under-the-receiver-operating-characteristic-curve (ROC-AUC) on the test set. The model performance was benchmarked against two classical models: Dense-Neural-Networks (DNN) and Extreme-Gradient-Boosting (XGB). Shapley values were calculated to interpret sample-wise predictions on the test set.

### Results

The best 10-features combination included 6 tumor and 4 ring features. For QNN, DNN, and XGB respectively: training ROC-AUC was: 0.86, 0.95, and 0.94; test ROC-AUC was: 0.76, 0.75, and 0.79; and test bACC was: 0.74, 0.73, and 0.72. The two most influential features were tumor Laplacian-of-Gaussian-GLRLM-Entropy and Sphericity.

### Conclusion

We developed an accurate interpretable QNN model with quantum-informed feature selection to differentiate between LBM and HGG on CE-T1 brain MRI. The model performance is comparable to state-of-the-art classical models.

## 3.2 INTRODUCTION

The most frequent malignant brain tumors in adults are metastases (BM) and high-grade gliomas (HGG) with incidence rates of 7 to 14 and 1.95 per 100000 population per year respectively (1,83). MRI is the gold standard imaging modality for diagnosing and characterizing BM and HGG. However, differentiating these two entities on MRI is often challenging, owing to shared imaging characteristics such as central necrosis, ring enhancement, and peritumoral edema. Furthermore, both tumor types might present as a solitary mass or multifocal disease, albeit with different frequencies.

The management strategy differs significantly between these two disease categories (84–87). Although biopsy and/or surgical resection with histopathological and molecular analysis (88) is generally recommended to establish a definite diagnosis and guide personalized treatment, there remains an undeniable need to improve the diagnostic ability of imaging to differentiate between these two disease entities. In rare cases where an invasive procedure could not be undertaken due to medical contraindications or eloquent tumor location, imaging accuracy becomes paramount. Moreover, for patients with a known history of malignancy, avoiding a confirmatory biopsy for brain metastases could prevent unnecessary toxicity, and thus leveraging confidence in the diagnostic capabilities of MRI imaging is eagerly needed.

MRI techniques that help differentiate between BM and HGG represent an active area of research, and significant improvements over conventional MRI techniques have been made in the last few years using spectroscopy, perfusion, and diffusion weighted imaging techniques (89–95). A major drawback of such techniques with respect to conventional MRI is the need for additional special imaging acquisitions that would have a negative impact on the procedural cost, patient convenience, and the time to get a diagnosis. Further, many of these techniques are not widely available.

Radiomics extraction is a promising quantitative method that could glean useful information from imaging data that are not discernable to the human eye (96,97). Two types of radiomics exist: “handcrafted radiomics” that are distinct mathematically-defined features generated through a stand-alone process, and “deep radiomics” that represent task-oriented features extracted by an auto-encoder or as a part of the deep layers’ data of a convolutional neural network (98). In many tasks, deep radiomics outperformed engineered features (99–101). In one BM-Glioblastoma classification study by Bae *et al.* comparing both techniques, deep radiomics yielded slightly better results than their counterparts (102).

Quantum computing is gaining much momentum as advances in quantum device engineering and computational algorithms are surging, and healthcare is a promising arena for these technologies (103). Quantum machine learning (QML) is an emerging

bridging discipline that combines principles from both quantum computing and classical machine learning (39,40,42,104). Few studies have pointed to a potential superiority of QML over classical algorithms under specific circumstances (43–45), and classification tasks with small training datasets might derive some benefit in terms of performance and generalization (46,47). However, a clear quantum advantage is yet to be demonstrated. Variational quantum classifiers (VQC), also quoted as quantum neural networks (QNN) due to their analogy with classical neural networks, are hybrid quantum-classical models consisting of three sequential blocks: data encoding or state preparation, trainable parametrized unitary evolution, and measurement of the quantum state. Such algorithms are considered kernel classifiers in that they map classical data to a high-dimensional feature space where a linear separation could be defined by measurement (41). Although the hope of QML is to achieve an advantage over classical methods through computational speed-ups, such results might not be possible on near-term quantum devices.

In order to enhance model interpretability, dimensionality reduction through feature selection is a pivotal step before training. Mutual information feature selection (MIFS) is a model-agnostic filtering method to select the features that share the maximal amount of information with the output variable, and is able to capture non-linear relationships (48). Conditional mutual information scores have been used to account for feature interaction (105), and such computation is classically intractable. As inspired by the work of Nguyen and colleagues (106), the MIFS problem could be mapped to a quadratic unconstrained binary optimization (QUBO) and solved heuristically using an Ising machine. Quantum annealing is a metaheuristic that consists of evolving an initial multi-qubit system prepared as the ground state of a simple Ising-spin Hamiltonian to a final Hamiltonian whose ground state represents the solution of a desired optimization problem. Quantum annealing has been suggested to match and sometimes outperform classical solvers in combinatorial optimization (49), hypothetically due to quantum tunneling. While quantum annealing is pushing its way toward commercial utilization, its superiority over classical heuristics remains questionable.

Radiomics and deep learning techniques have been widely used for brain tumor classification (102,107–111). Deep learning algorithms suffer from the lack of interpretability and are regarded as “black-box” methods, drawing much skepticism around their use in clinical practice. QNNs represent a new family of machine learning algorithms that are interesting to explore on real-world datasets. In this work, we evaluate the performance of a QNN to differentiate between BM and HGG on a real-world conventional MRI dataset, with a particular focus on model interpretability. We use a state-of-the-art method for feature selection, MIFS, taking into consideration feature interaction, and we implement it on a quantum annealer. Moreover, since BM

are often smaller than HGG at diagnosis, there exists a significant confounding effect of tumor size that is seldom considered in many published studies (107). We herein limit our analysis to large BM. Finally, we compare the QNN performance to proven classical machine learning algorithms.

## **3.3 MATERIALS AND METHODS**

### **3.3.1 Patients and data collection**

This study was conducted at the Hôtel-Dieu de France University Medical Center after obtaining approval from the institutional Ethics board. Medical records from the radiation oncology and radiology departments were retrospectively screened for patients diagnosed with brain metastases and high-grade gliomas (WHO grade 3-4) between 2012 and 2019. A total of 423 patients aged more than 18 years who had a confirmatory pathology report and a pre-treatment gadolinium-enhanced T1-weighted (CE-T1W) brain MRI were included. Brain MRI images were thoroughly reviewed, and patients who had significant motion artifacts, lesions without ring-enhancement, or tumors with largest diameter of less than 2cm were excluded. The final dataset consisted of 129 BM and 72 HGG patients. Imaging performed at our institution as well as imported images from other radiology centers were allowed. Different imaging protocols and 2D or 3D acquisition were permitted as this is more representative of the real-world setting. Since this was a tumor-wise analysis, patients who had multiple brain metastases or multifocal gliomas were accepted for inclusion, yet only one tumor per patient was selected. A 2D largest diameter cutoff of 2cm was chosen to exclude very small lesions that are more likely to represent metastases rather than gliomas, and thus might lead to classification bias.

### **3.3.2 Segmentation**

Tumors on CE-T1W brain MRI images were manually segmented by a senior radiation oncology resident and a radiation oncologist using Eclipse™ treatment planning system (Varian Medical Systems, Palo Alto, CA, USA). A 5mm isotropic peritumoral ring extension was created for each tumor volume, and was manually edited to carve out parts extending beyond anatomical barriers such as bone, cerebral falx, etc. All segmentations were verified by both physicians for consistency.

### **3.3.3 Image pre-processing**

First, the Advanced Normalization Tools (ANTs) library (112) was used to apply the bias field correction on the MRI images using the N4ITK algorithm (113). Spatial resampling to a 1mm x 1mm x 1mm voxel size was then performed using ANTs. Automatic brain segmentation was done using two publicly available brain extraction

tools, BET from the Functional MRI of the Brain Library (FSL) (114), and the HD-BET (115). The brain masks were then checked visually, and the best fitting mask was retained for each MRI. The masks were used for Z-Score intensity normalization (116) of the resampled MRI images. We used the Z-Score method as it resulted in the best performing radiomics-based models when used with a fixed bin number intensity discretization method (117).

### 3.3.4 Feature extraction and data pre-processing

Radiomic extraction was performed using PyRadiomics library version 3.0.1 (118) on Python 3.7.9. Most of the features used in PyRadiomics conform to the Imaging Biomarker Normalization Initiative (IBSI) (119,120). A fixed bin number of 32 was used for gray level intensity discretization as recommended by Carré *et al.* (117). A total number of 913 tumor features and 900 ring features were extracted for each patient. Seven feature classes were considered: shape features (13 features), first order statistics (17 features), grey-level co-occurrence matrix – GLCM (22 features), grey-level run length matrix – GLRLM (16 features), grey-level size zone matrix – GLSZM (16 features), grey-level dependence matrix – GLDM (14 features), and neighbourhood grey tone difference matrix – NGTDM (5 features). Aside from original images, 9 filters were used for first and second order features prior to extraction: Laplacian of Gaussian – LoG (1 filter, sigma value: 1), and wavelet filters (8 filters, with the different combinations of high – H and low – L pass filters in the three dimensions: LLH, LHL, LHH, HLL, HLH, HHL, HHH, and LLL). Shape features were omitted for the ring volume due to irrelevance.

The extracted features were rescaled using Scikit-learn StandardScaler function by subtracting the mean value and scaling to unit variance. The dataset was split into a training set (70%) and a validation set (30%) using a stratified approach. The number of patients in the BM cohort is significantly higher than the HGG dataset. To minimize the effect of class imbalance, a combined over-sampling (Synthetic Minority Oversampling Technique – SMOTE) and under-sampling (RandomUnderSampler) strategy was used. Highly correlated features with a Spearman’s rank correlation coefficient absolute value  $|\rho|$  of more than 0.8 were eliminated.

### 3.3.5 Feature selection

Primary feature selection was performed using linear support vector classification (LinearSVC), and then the best combination of 10 features was determined using MIFS.

Mutual information (MI) between two variables  $X$  and  $Y$  quantifies the amount of information about one variable derived from observing the other variable. It can be expressed as a function of Shannon entropy (SE) and conditional Shannon entropy (CSE) of these variables as follows:

$$I(X;Y) = H(Y) - H(Y|X)$$

where  $H(Y) = - \int p(Y) \log p(Y) dY$  is the SE of  $Y$ , a function of its probability distribution  $p(Y)$ , and  $H(Y|X) = H(X,Y) - H(X)$  is the CSE of  $Y$  conditional of  $X$ , a function of SE of  $X$  and joint SE of both variables.

CMI between two variables  $X$  and  $Y$  given the previous selection of a variable  $Z$  takes into account the interaction between these variables, and is given by:

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$$

where  $H(X|Z)$  is the CSE of  $X$  conditional of  $Z$ , and  $H(X|Y,Z)$  is the CSE of  $X$  conditional of  $Y$  and  $Z$ .

Given a set of  $n$  features  $X = \{X_1, X_2, X_3, \dots, X_n\}$ , selecting the subset of variables  $S \subset X$  that shares the maximum information with a variable  $Y$  requires maximizing  $I(S;Y)$ , which is an NP-hard problem. The method used herein is inspired by the work of Nguyen *et al.* (106), and a relevant example can be found at (121).

Under the assumption of variable conditional independence, this feature selection problem could be approximated (106) by:

$$\arg \max_{S \subset X} \left\{ \sum_{X_i \in S} I(X_i; Y) + \sum_{X_i, X_j \in S} I(X_i; Y|X_j) \right\}$$

and could be written in a matrix representation as follows:

$$\arg \max_{x \in \{0,1\}^n} \{x^T Qx\}$$

where  $Q$  is a  $n \times n$  matrix with  $Q_{ii} = I(X_i; Y)$  and  $Q_{ij} = I(X_i; Y|X_j)$  for  $i \neq j$ , and  $x$  is a  $n \times 1$  vector with  $x_i = \begin{cases} 1, & i^{th} \text{ feature is selected} \\ 0, & \text{otherwise} \end{cases}$

Evidently, this problem fits to the following QUBO formulation:

$$\arg \min \left\{ - \sum_i^n Q_{ii} x_i - \sum_{i < j}^n Q_{ij} x_i x_j \right\}$$

This optimization could be naturally mapped to an Ising model Hamiltonian, and thus could be solved by approximating its ground state on an Ising machine. We use for this purpose a quantum annealing approach on D-Wave Systems Advantage Quantum Processing Unit (QPU).

In order to solve for a specific number of features  $k$ , a penalty factor is added to the QUBO to penalize solutions with  $|S| \neq k$ :

$$P = -\alpha \left( \sum_{i=1}^n x_i - k \right)^2$$

where  $\alpha$  is a tunable penalty amplitude.

We selected a set of most important features based on MI scores, and then clique embedding was performed using D-Wave's minorminer heuristic tool. The MIFS heuristic was solved for all  $k$  values, each with 5000 reads. The set of 10 features was selected to build the classification model.

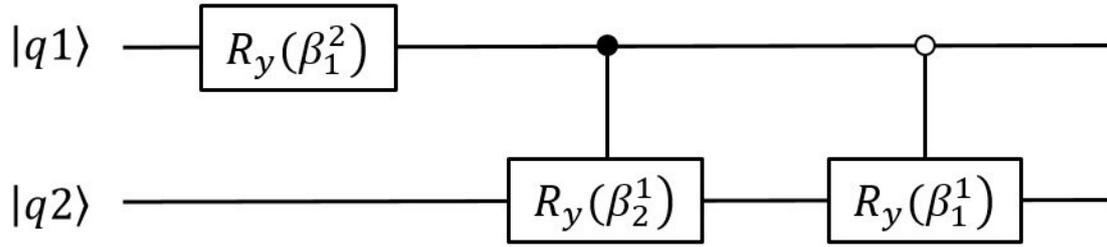
### 3.3.6 Quantum neural network

We used herein the PennyLane quantum machine learning library (122) to build a QNN, which consists of a parametrized variational circuit that acts as a binary classifier. For this purpose, we used PennyLane's default\_qubit quantum simulator, with a circuit of two qubits using an architecture similar to that described in Farhi *et al.* (123) and Schuld *et al.* (124). A Python code developed by the PennyLane team to implement a variational classifier could be accessed at (125).

An amplitude encoding method was employed to map the feature vector to the Hilbert space of the 2-qubit system, as shown in Möttönen *et al.* (126) and the book by Schuld and Petruccione (40). For an  $n$ -qubit circuit, a feature vector of  $2^n$  dimensions is encoded, i.e. four features in the current work. Amplitude encoding entails a sample-wise feature vector normalization to unit norm to satisfy the Born rule, which leads to different normalization factors and significant distortion of the feature dataset. To tackle this issue, three informative features were used, and were subsequently padded with a non-zero constant term that would ultimately bear information on the normalization factor.

Principle component analysis (PCA) was used to reduce the dimensionality of the feature dataset from 10 to 3 while retaining the maximal amount of information. Rescaling was performed using MinMaxScaler to ensure positivity of the features, since this would lead to positive amplitudes and thus avoidance of a cascade of Z-axis rotations for the state preparation subroutine (40).

To achieve the amplitude encoding of the normalized feature vectors, a set of controlled Y-axis rotations is performed according to a reversed scheme of the algorithm used by Möttönen *et al.* (126). The corresponding circuit representation in our case is shown in Figure 3.1, where  $\beta$  denotes the Y-axis rotation angle, and the white and black circles indicate a control on qubit basis state 1 and 0 respectively.



**Figure 3.1** Circuit representation for amplitude encoding in the case of a 2-qubit system.  $R_y(\beta)$  denotes the Y-axis rotation angle; the white circle indicates a control on qubit basis state 1; the black circle indicates a control on qubit basis state 0;  $|q1\rangle$  and  $|q2\rangle$  are qubit 1 and 2 states respectively.

The angles  $\beta$  are given by (40):

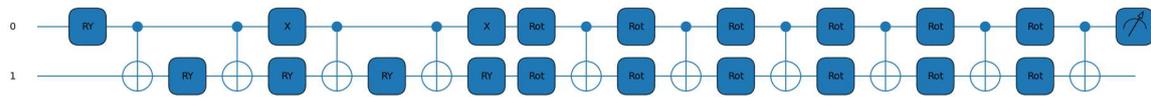
$$\beta_j^s = 2 \sin^{-1} \left( \frac{\sqrt{\sum_{l=1}^{2^{s-1}} |\alpha_{(2j-1)2^{s-1}+l}|^2}}{\sqrt{\sum_{l=1}^{2^s} |\alpha_{(j-1)2^s+l}|^2}} \right)$$

where  $\alpha_i$  is the  $i^{\text{th}}$  element of the amplitude vector (i.e. the normalized feature vector).

These controlled Y-axis rotations are further decomposed into basic circuit components including Controlled-NOT (CNOT) and rotation around Y-axis (Ry) gates.

We define a parametrized entangling circuit model composed of 6 layers of a unitary consisting of an arbitrary single qubit rotation (Rot gate) on each qubit followed by a CNOT gate. A Rot gate is parametrized by 3 angle values  $\varphi$ ,  $\theta$ , and  $\omega$ , and could be written as a function of Y and Z axis rotations as  $RZ(\omega)RY(\theta)RZ(\varphi)$ . This circuit acts on the encoded 'ket' vector, and the expectation value of the Pauli Z operator acting on the first qubit is calculated. The quantum circuit (or ansatz) is shown in Figure 3.2. If the value is negative, the prediction  $p = -1$  is a brain metastasis; otherwise,  $p = +1$  for glioma.

The parameters are classically optimized using AdamOptimizer, a gradient-descent optimizer with an adaptive learning rate.



**Figure 3.2** The 2-qubit circuit acting as a machine learning classifier. RY: Y-axis rotation; X: Pauli X gate; Rot: arbitrary qubit rotation. Qubits number 1 and 2 are denoted “0” and “1” respectively.

### 3.3.7 Benchmarking and model evaluation

To evaluate the performance of the QNN relative to classical machine learning models on our 10-features dataset, we benchmarked our results against two well-established classical models: extreme gradient boosting (XGBoost) and dense neural network (DNN).

For XGBoost, we used Bayesian optimization to tune 9 hyper-parameters to get the best area under receiver operating characteristic curve (ROC-AUC). The search range for the hyper-parameters were  $1 \times 10^{-5}$  to  $1 \times 10^{-1}$  for the learning rate, 300 to 1000 for the number of trees, 1 to 4 for the minimum sum of instance weight needed in a child, 3 to 5 for maximum tree depth, 0.2 to 0.5 for the subsample ratio of the training instance, 0.2 to 0.5 for the subsample ratio of columns when constructing each tree, 0 to 0.1 for the minimum loss reduction, and 0 to 75 for L1 and L2 regularization terms on weights. Five-fold cross validation was done on the training set to reduce overfitting.

The DNN used Keras’ Adam optimizer, and included three hidden layers, with the output layer using a sigmoid activation function. We used a dropout layer with a rate of 30% after the input and each hidden layer to minimize overfitting. Early stopping with a patience of 10 was utilized during training.

Balanced accuracy (bACC) and ROC-AUC were used for model comparison. Precision, recall, and F1-scores were also reported.

### **3.3.8 QNN model interpretation**

Shapley Value (127,128), a coalitional game-theory solution concept to determine a player's contribution to an overall cooperative gain was used herein to compute sample-wise feature attribution to an overall prediction, as implemented by Lundberg and colleagues (129). We used the model-agnostic KernelExplainer to estimate Shapley Values for the purpose of the current study. The mean absolute Shapley Values for all features were calculated for impact ranking. Class-wise mean Shapley Values and mean feature values were also calculated, and the same was done for true and false predictions.

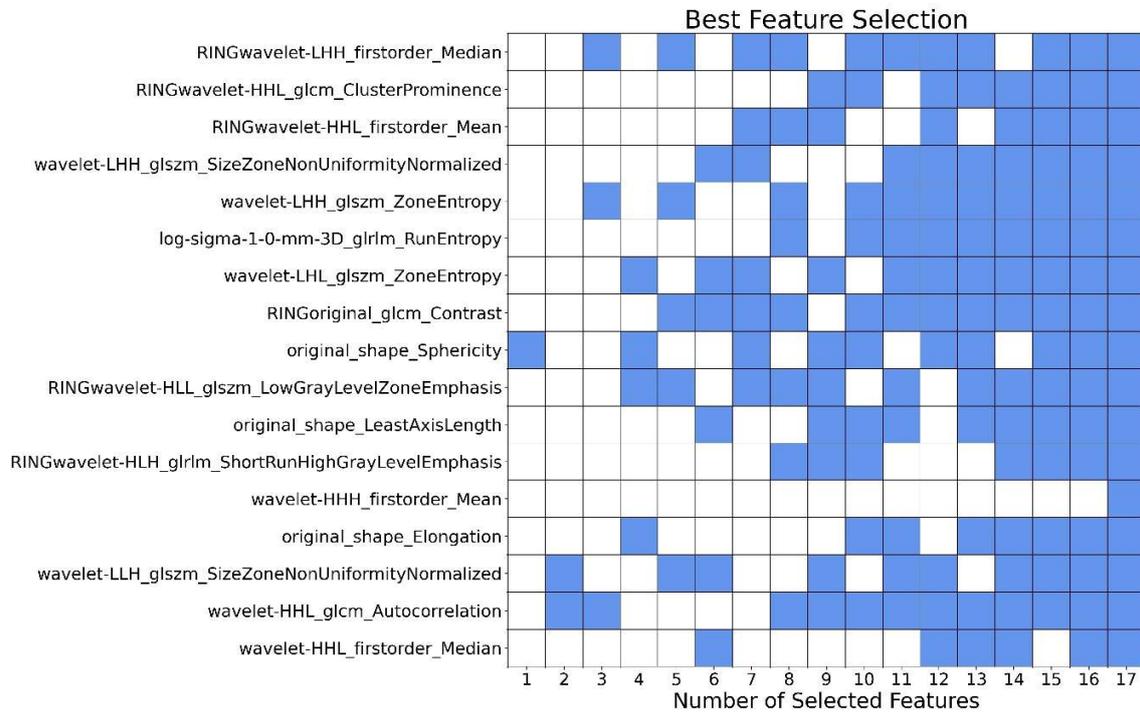
## **3.4 RESULTS**

### **3.4.1 Patient characteristics**

The BM cohort included 56 patients (43.4%) with primary lung cancer, 40 patients (31%) with breast cancer, and 33 patients (25.6%) with other pathologies. Median tumor 3D-largest diameter was 34mm (IQR: 26-42.6mm) for BM and 47.9mm (IQR: 38.4-61.1mm) for HGG. 3D MRI acquisition was performed in 91 patients with BM (70.5%) and 43 patients with HGG (59.7%).

### **3.4.2 Feature selection**

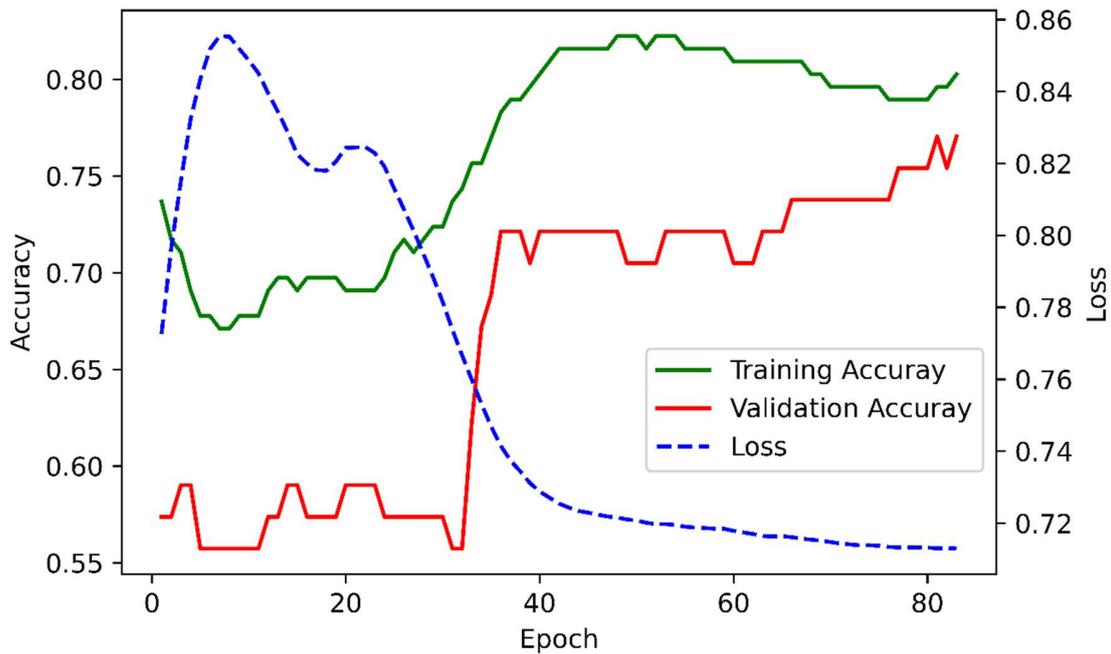
After eliminating highly correlated features, a set of 209 features was kept. Using LinearSVC, 54 important features were classically selected. The MI scores for these remaining features were calculated, and 17 most important features were retained. The MIFS heuristic was solved on D-Wave's quantum annealer for  $k$  between 1 and 17 (Figure 3.3). The best combination of 10 features was determined and was used to build the machine learning classifiers.



**Figure 3.3** The results of the MIFS heuristic for all  $k$  values between 1 and 17, showing the best combination of  $k$  features. Blue squares correspond to selected features.

### 3.4.3 QNN model performance and benchmarking

The variational quantum classifier training and validation accuracy learning curves, as well as the training loss curve are shown in Figure 3.4. Model training accuracy was 0.8, and validation bACC was 0.74. ROC-AUC was 0.86 on the training set, and 0.76 on the validation set.



**Figure 3.4** Learning curves for training and test accuracy and training loss for the Quantum Neural Network model.

Respective bACC on the test set for XGBoost and DNN were 0.72 and 0.73. Training and test ROC-AUC were 0.94 and 0.79, and 0.95 and 0.75 for XGBoost and DNN respectively. These metrics along with the precision, recall, and F1-scores for all models are summarized in Table 3.1. The difference in accuracy between training and validation sets are 6%, 17%, and 15% for QNN, XGBoost, and DNN respectively, and the respective differences in ROC-AUC are 10%, 15%, and 20%.

**Table 3.1** Models' performance metrics

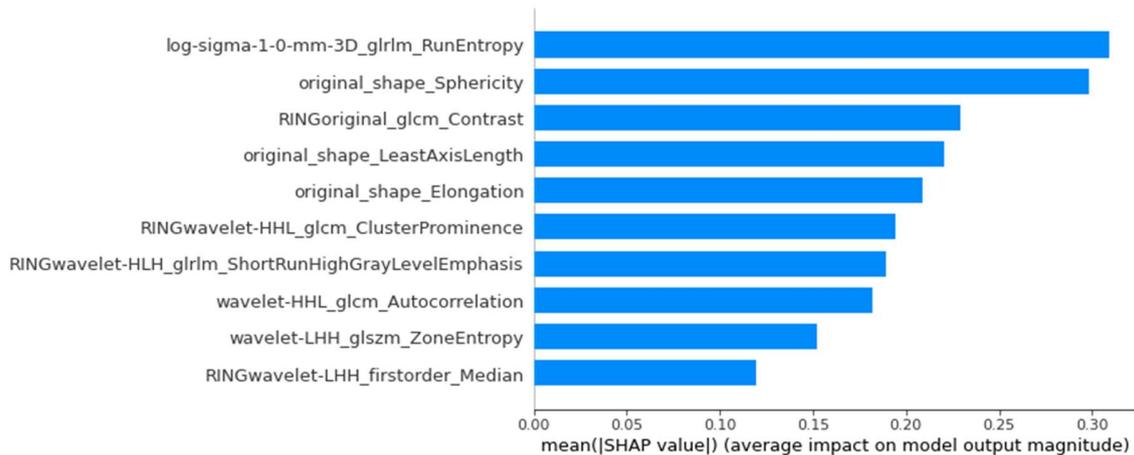
	QNN	XGBoost	DNN
Training ACC	0.8	0.89	0.88
Test bACC	0.74	0.72	0.73
Training ROC-AUC	0.86	0.94	0.95
Test ROC-AUC	0.76	0.79	0.75
Test Precision	0.77	0.75	0.77
BM	0.80	0.79	0.79
HGG	0.7	0.68	0.72
Test Recall	0.77	0.75	0.77
BM	0.85	0.85	0.87
HGG	0.64	0.59	0.59
Test F1-score	0.77	0.75	0.76
BM	0.83	0.81	0.83
HGG	0.67	0.63	0.65

*QNN: quantum neural network; XGBoost: extreme gradient boosting; DNN: dense neural network; ACC: accuracy; bACC: balanced accuracy; ROC-AUC: area under receiver operating characteristic curve; BM: brain metastases; HGG: high-grade glioma*

#### 3.4.4 QNN model interpretability

Instance-wise Shapley Values were calculated, and each feature's contribution to the overall model prediction was determined. Figure 3.5 shows the mean absolute values for each feature, ranked by order of importance. Tumor LoG-3D-GLRLM-Run-Entropy

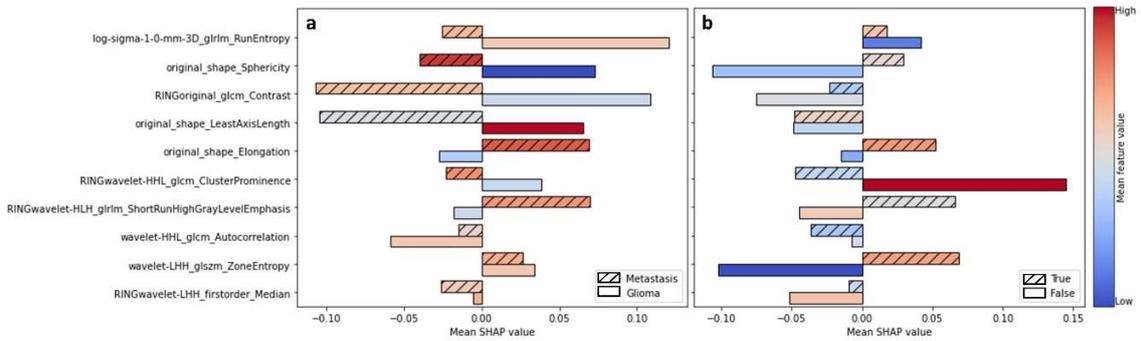
and Sphericity were the most important features. Ring GLCM-Contrast ranked third and was the most important ring feature.



**Figure 3.5** Mean absolute Shapley Value for each feature, ranked by order of importance.

Figure 3.6A shows the mean Shapley Value for each feature according to the predicted class, and colors represent the mean feature value according to the color scale bar. A positive Shapley Value shifts the prediction output towards HGG from the baseline value i.e. the mean Shapley Value for all predictions. For instance, low tumor Sphericity is associated with a positive Shapley Value for HGG predictions, whereas for BM, a high Sphericity value leads to an average negative Shapley Value. Further, a high tumor LoG-3D-GLRLM-Run-Entropy that indicates increased texture heterogeneity contributes to negative Shapley Values and shifts the prediction towards BM. A lower average GLRLM-Run-Entropy value is associated with a high positive mean Shapley Value shifting the outcome significantly towards HGG. On another hand, high values of the two most important ring features, GLCM-Contrast (correlating with disparity in intensity values among neighboring voxels) and Wavelet-HHL-GLCM-Cluster-Prominence (implying more asymmetry about the mean voxel value) are associated

with negative Shapley Values pushing the prediction towards BM, and the opposite is true for HGG.



**Figure 3.6** Mean Shapley Values and feature values for the different features according to the predicted class (metastasis or glioma) **[a]**, and the prediction accuracy (true or false) **[b]**. Colors represent the mean feature values according to the color scale bar.

Figure 3.6B highlights differences in Shapley Value and feature values between true and false predictions on the test set. The largest mean Shapley Value magnitudes for the false predictions were observed for low tumor Sphericity combined with a negative Shapley Value, high ring Wavelet-HHL-GLCM-Cluster-Prominence combined with a positive Shapley Value, and low tumor Wavelet-LHH-GLSZM-Zone-Entropy combined with a negative Shapley Value.

### 3.5 DISCUSSION

The integration of machine learning models in conjunction with other clinical instruments holds the potential to significantly enhance the accuracy of diagnosing and planning treatment for brain tumors, specifically LBM and HGG, using MRI scans. By leveraging the power of quantum computing, a QNN model can uncover subtle patterns and relationships within radiomic data, leading to more precise and reliable classification of brain tumors. This enhanced accuracy can aid healthcare professionals in making more informed decisions regarding treatment strategies, such as surgical intervention, radiation therapy, or targeted drug therapies. Developing an accurate QNN model could serve as a valuable tool for non-invasive and automated analysis of medical imaging data, reducing subjectivity and interobserver variability in tumor classification.

In this work, we developed an interpretable radiomic classifier to differentiate between large BM and HGG on CE-T1W brain MRI. We used to this end a QNN model, and we showed that its performance on our dataset is comparable to two state-of-the-art classical algorithms: XGBoost and DNN. We employed a MIFS method and solved the resulting combinatorial optimization on D-Wave's quantum annealer. To the extent of our knowledge, this is the first study to use quantum algorithms for brain tumor classification on a real-world dataset, and our results shed light on the potential value of these algorithms, and the need for further exploration and development.

Since the role of feature selection is cardinal for reducing model complexity and leveraging interpretability, we applied a pipeline of two feature selection methods as this strategy was shown to improve the selection results (130,131). We used a MIFS method to get the best 10 features, since this method captures non-linear relationships and accounts for interaction effects between features. The resulting QUBO, a computationally expensive optimization, was solved using a quantum annealing metaheuristic as it was shown to match or outperform other classical methods (49). Nonetheless, the limited number of qubits and connectivity in quantum annealers can restrict the scale and complexity of implementable circuits, and this remains a considerable limitation in the current era.

While the QNN model developed herein showed similar performance on the test set as for XGBoost and DNN, the difference between training and validation metrics was smaller for QNN than for the classical algorithms despite using overfitting techniques for the latter, hinting to a better generalization and lower overfitting for the QNN. This is in line with the work of Caro *et al.* that showed a favorable generalization error for quantum machine learning algorithms (47), and it would be worthwhile testing this further in future studies.

Differentiating between BM and HGG has been the subject of many previously published studies (107), including handcrafted radiomics-based machine learning techniques as well as deep learning models, with reported accuracy results ranging from 64% to 98%. Comparing results indirectly from separate studies is trivial, and such heterogeneity in classification performance could stem from many factors including the nature and quality of the data, the methods used, and the quality of reporting. In a systematic review by Jekel *et al* (107), less than half of the TRIPOD items reflecting critical points in model development were reported on average. Moreover, most of the published papers do not report the tumor size, which could introduce significant classification bias because BM are more likely to be diagnosed at an earlier growth stage than HGG owing to the close surveillance pattern of cancer patients.

De Causans *et al* (132) developed a tumor-radiomics machine learning algorithm based on CE-T1W MRI to classify glioblastoma and solitary BM larger than 2cm, yielding a balanced accuracy of 0.8 and a ROC-AUC of 0.85 on the test set. Their results were slightly better than ours, however it should be noted that imaging was done on the same 3-Tesla MRI machine using the same protocol, by contrast to our heterogeneous dataset including images from different institutions using different machines and acquisition parameters. Although heterogeneity of training data might affect the performance of a model, its ability to generalize to unseen data with different acquisition protocols might be better. Nonetheless, the ultimate test of generalization is through external validation. Heterogeneity has been shown to compromise the robustness of texture features, yet most of the selected texture features in this work were shown to have good repeatability for T1-weighted MRI with various MRI scanners and scanning parameters, particularly after pre-processing (133). In this study, we adhered to the image pre-processing pipeline suggested by Carré *et al.* (117) to maximize radiomic feature stability. Moreover, 3 out of 10 features were shape-based, and thus are very robust with respect to imaging techniques.

In this work, we focused on the model interpretability. Sphericity ranked second in terms of feature attribution to an overall prediction. This is consistent with the results of De Causans *et al.* (132) and Priya *et al.* (134) where tumor sphericity was the most important discriminating feature. From a clinical perspective, being able to see how each feature is influencing the decision process for a certain prediction would give more confidence in using the algorithm in practice. For instance, knowing that BM are more spherical on average than HGG, if an algorithm uses low sphericity of a tumor to significantly bias a prediction towards a BM prediction, the probability that this prediction is false increases. Our results support this rationality, and as we can see in Figure 6B for sphericity, false predictions had high mean absolute Shapley Values attributing low sphericity to BM prediction. We acknowledge that some features are

more complex to understand, and significant uncertainty remains as to extrapolating average values to individual instances.

Our analysis had some limitations. Despite being one of the largest studies to report on this subject, the sample size remains relatively small. Additionally, in this study, external validation was not conducted to provide a better evaluation of generalizability, although the heterogeneous nature of our dataset could potentially be advantageous in this aspect. This highlights the difficulty in gathering sufficient data for imaging-based machine learning studies in neuro-oncology, and stresses on the need for data-sharing initiatives to improve model development and validation and push this field further towards clinical integration. Furthermore, we limited our model to CE-T1W MRI sequence. Adding more sequences might lead to improved prediction performance, although it might increase model complexity.

We used PennyLane's quantum simulator for the purpose of this study since our scope is to test the algorithm rather than the quantum device. Nevertheless, this algorithm could be easily implemented on a near-term quantum computer with reasonable error. The model architecture is scalable and could include a larger number of features, however as the number of required qubits increases, the circuit becomes exponentially expensive to simulate classically. The rapid development in quantum device engineering and quantum error-correction algorithms is likely to allow practical implementation of such algorithms in the near future. However, we acknowledge that quantum technology is still in its early stages, and further developments are needed to fully exploit the advantages of quantum computing over its classical counterpart. Efforts are eagerly needed to explore and establish the utility of quantum machine learning algorithms in the field of medicine, more particularly in oncology.

## 3.6 CONCLUSION

In conclusion, this work led to the creation of a cutting-edge CE-T1W MRI-based radiomic Quantum Neural Network classifier, employing quantum-informed Mutual Information Feature Selection. We were able to demonstrate that our developed model effectively discerns between large brain metastases and high-grade gliomas with remarkable accuracy. Notably, our model's performance was found to be on par with two prevailing state-of-the-art algorithms, namely XGBoost and Dense Neural Network, while seemingly exhibiting less susceptibility to overfitting issues. This algorithm warrants further external validation. We have also provided a game-theory approach for model interpretability using Shapley Value, and our results were in-line with published data. Further development of such algorithms using quantum technology has significant implications for the field of medical imaging, paving the way for enhanced diagnostic capabilities and improved patient care in the realm of brain tumor classification.

### 3.7 APPENDIX

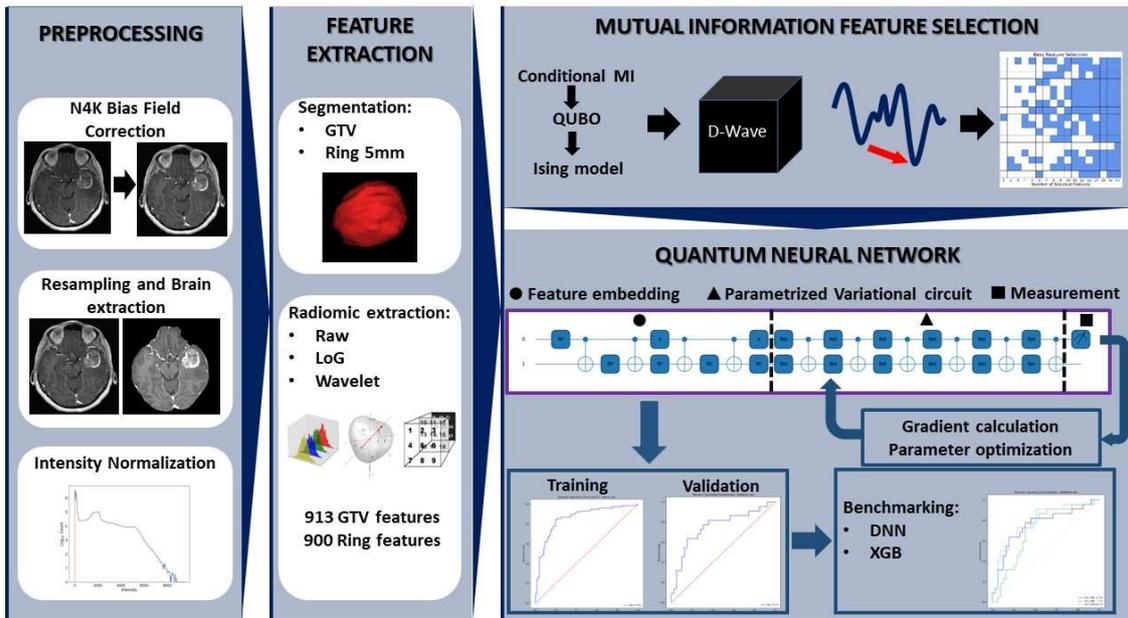
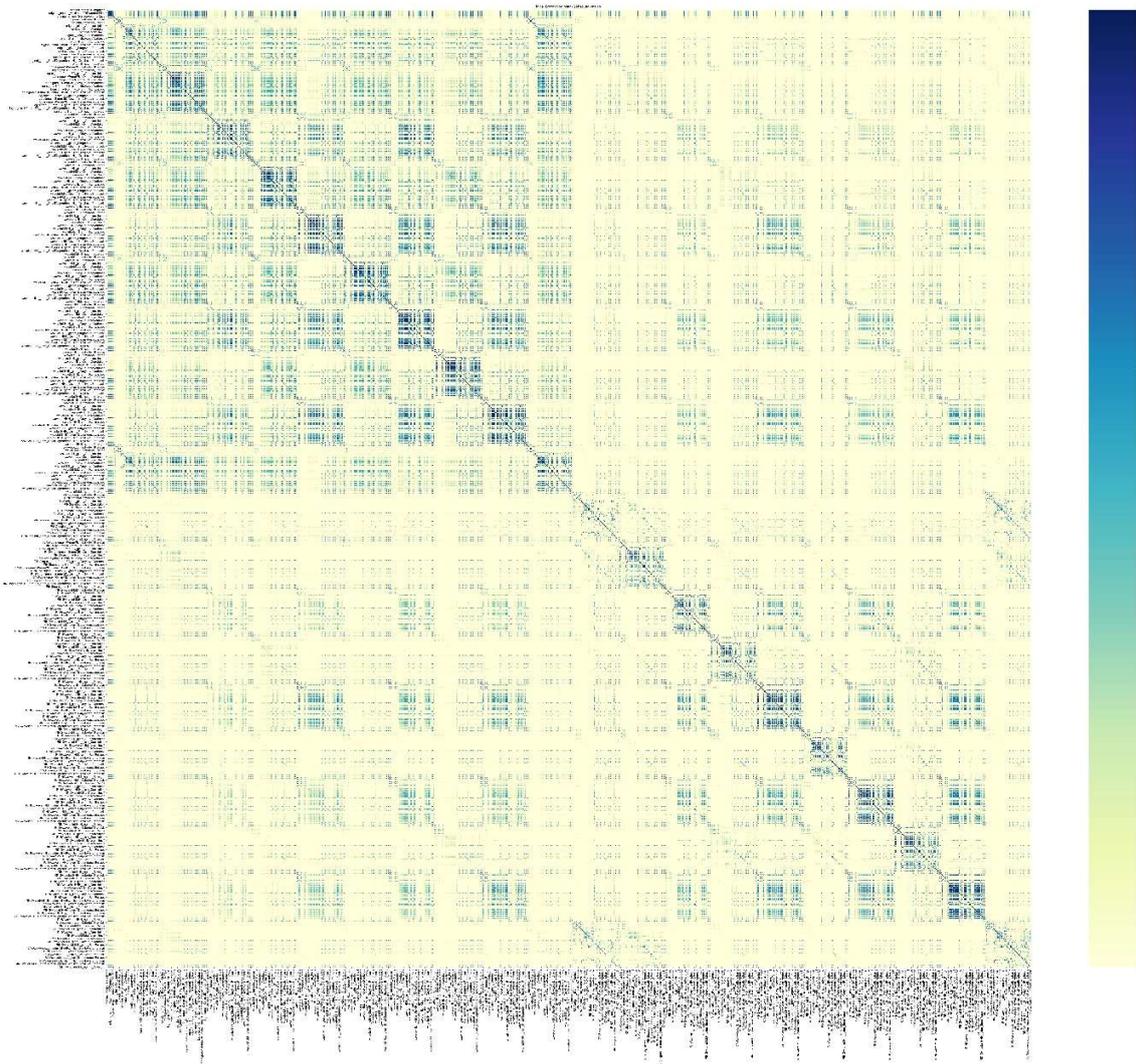
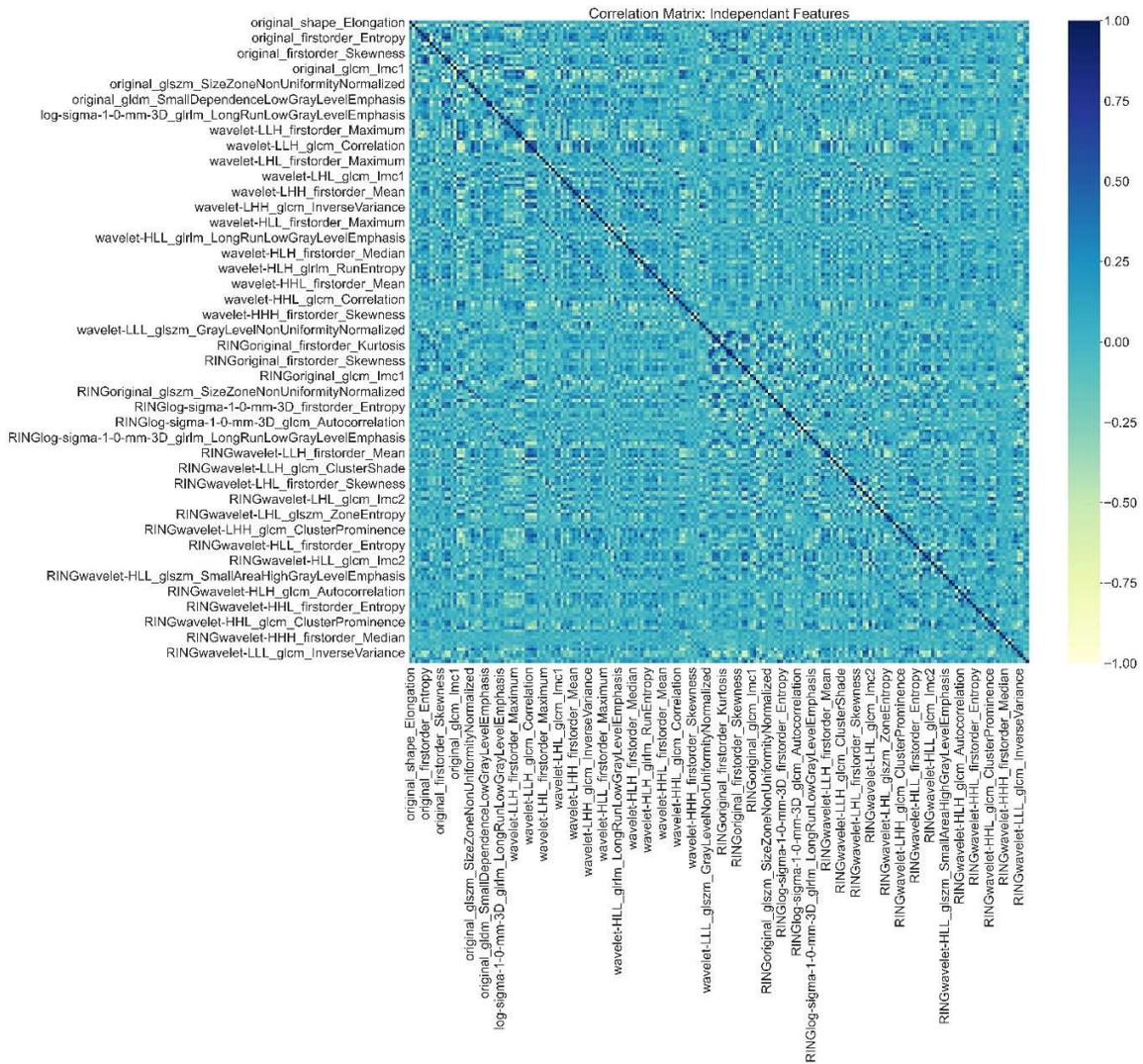


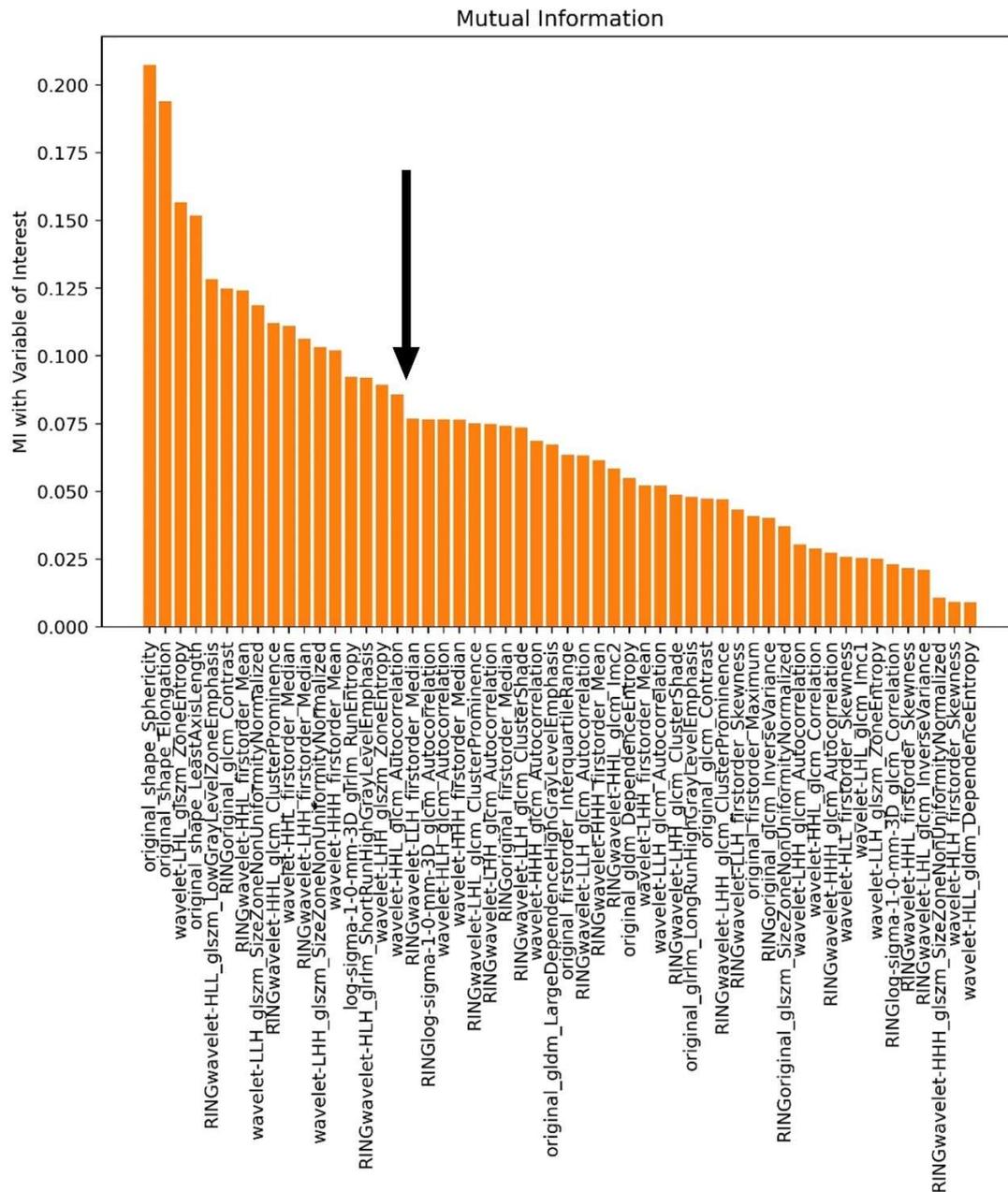
Figure 3.7 Supplementary Pipeline for developing the Quantum Neural Network



**Figure 3.8** Supplementary Correlation matrix for the whole feature set



**Figure 3.9 Supplementary** Correlation matrix for remaining features after eliminating highly correlated features



**Figure 3.10 Supplementary** Mutual Information scores in decreasing order for the remaining 54 features following LinearSVC feature selection. The best 17 features were visually selected. The black arrow points past the 17<sup>th</sup> selected feature.



## **4 A 3D Convolutional Neural Network Based on Non-Enhanced Brain CT to Identify Patients with Brain Metastases**

This chapter highlights a deep learning approach to diagnosing brain metastases in patients using non-enhanced brain CT scans. Non-enhanced CT scans are widely available due to the rise of using PET-CT in cancer staging and treatment response evaluation but are often limited in detecting brain metastases. This research aims to develop a 3D Convolutional Neural Network that can distinguish between patients with and without brain metastases using non-enhanced CT scans.

This work has been published in the *Journal of Imaging Informatics in Medicine* and has been modified herein to a minimal extent.

## **4.1 ABSTRACT**

### **Introduction**

Dedicated brain imaging for cancer patients is seldom recommended in the absence of symptoms. There is increasing availability of non-enhanced CT (NE-CT) of the brain, mainly owing to a wider utilization of Positron Emission Tomography-CT (PET-CT) in cancer staging. Brain metastases (BM) are often hard to diagnose on NE-CT. This work aims to develop a 3D Convolutional Neural Network (3D-CNN) based on brain NE-CT to distinguish patients with and without BM.

### **Materials and methods**

We retrospectively included NE-CT scans for 100 patients with single or multiple BM and 100 patients without brain imaging abnormalities. Patients whose largest lesion was <5mm were excluded. The largest tumor was manually segmented on a matched contrast-enhanced T1 weighted Magnetic Resonance Imaging (MRI), and shape radiomics were extracted to determine the size and volume of the lesion. The brain was automatically segmented, and masked images were normalized and resampled. The dataset was split into training (70%) and validation (30%) sets. Multiple versions of a 3D-CNN were developed, and the best model was selected based on accuracy (ACC) on the validation set.

### **Results**

The median largest tumor Maximum-3D-Diameter was 2.29cm, and its median volume was 2.81cc. Solitary BM were found in 27% of the patients, while 49% had >5 BMs. The best model consisted of 4 convolutional layers with 3D average pooling layers, dropout layers of 50%, and a sigmoid activation function. Mean validation ACC was 0.983 (SD: 0.020) and mean area under receiver operating characteristic curve was 0.983 (SD: 0.023). Sensitivity was 0.983 (SD: 0.020).

### **Conclusion**

We developed an accurate 3D-CNN based on brain NE-CT to differentiate between patients with and without BM. The model merits further external validation.

## 4.2 INTRODUCTION

Brain metastases represent the most common type of malignant brain tumors, affecting up to 20% of cancer patients throughout their disease course (2). There is growing evidence to suggest that the incidence of brain metastases is rising (135,136), likely due to advances in cancer therapy and increased patient survival, as well as improved diagnostic capabilities stemming from wider utilization of advanced imaging modalities.

Lung, breast, colorectal, melanoma, and kidney cancers encompass the majority of primary malignancies among patients with brain metastases (2). Current guidelines do not explicitly endorse dedicated brain imaging in the absence of symptoms for many types of cancer, even in the metastatic setting (5), primarily due to the lack of data demonstrating a survival advantage with early diagnosis of asymptomatic brain metastases in those cases.

Early detection of brain metastases is becoming increasingly vital in the management of cancer, especially with the advent of Central Nervous System (CNS) active drug development. Current guidelines provide a nuanced stance on the management of asymptomatic brain metastases in patients with breast, lung, and melanoma cancers (137,138). However, these guidelines endorse the deferral of local therapy for patients with asymptomatic brain metastases eligible for CNS-directed systemic therapy, albeit conditional on patient-centered decision-making and multidisciplinary discussion including size, location, and number of brain metastases, expected response rates and durability with systemic therapy, and access to close neuro-oncologic surveillance among other factors. This endorsement reflects a broader shift in the oncology community, further stressing on the need for early detection of brain metastases to enable access to innovative treatments, contribute to the evolving clinical trial landscape, and ultimately, enhance cancer outcomes and quality of life.

Magnetic Resonance Imaging (MRI) is widely recognized as the gold standard for the diagnosis of brain metastases owing to its superior soft tissue contrast. Nevertheless, its use is primarily recommended in the scope of neurological symptoms rather than as a routine part of the initial staging workup (5). On the other hand, there has been an increasing adoption and endorsement of Positron Emission Tomography-Computed Tomography (PET-CT) for whole-body staging (5), which has been demonstrated to have very low sensitivity in diagnosing brain metastases (139). Furthermore, many cancer patients undergo brain non-enhanced CT scans (NE-CT) for head trauma or headache investigation. In this context, efforts to glean as much information as possible from NE-CT images are highly needed to detect brain metastases, whether performed as a standalone test or as part of a PET-CT.

Detecting brain metastases on NE-CT presents substantial challenges for radiologists due to the inherent limitations in contrast and resolution compared to enhanced imaging modalities. Drawing from previous experiences in other imaging modalities such as MRI and contrast-enhanced CT scans (CE-CT) where deep learning algorithms and computer-aided diagnosis (CAD) have significantly contributed to better detection capabilities for brain metastases (51,140), it is conceivable that similar methods could significantly mitigate the challenges faced for NE-CT brain scans. Kato *et al.* published two studies on automatic detection of brain metastases on NE-CT using Single-Shot Detectors (SSD) and its use as a CAD tool to help non-experienced radiologists in this task (52,141), thus providing a proof of concept in this space. Of note, their model yielded very low sensitivity for lesions smaller than 6mm. However, their work was limited to patients with metastases; hence, the efficacy of this algorithm in distinguishing between patients with and without metastases remains unclear.

In this work, we aimed to develop a 3D Convolutional Neural Network (3D-CNN) based on NE-CT of the brain to differentiate between patients with and without brain metastases.

## **4.3 MATERIALS AND METHODS**

### **4.3.1 Data collection**

This study was retrospectively conducted at the Hôtel-Dieu de France University Medical Center. Approval from the institutional Research Ethics Board was obtained prior to study start. The medical records of patients treated for brain metastases at the Radiation Oncology department were reviewed to construct a cohort of 100 patients (C1). Such diagnosis was made using clinical judgement based on histopathological data, history of cancer, and imaging results. Patients 18 years of age and older who had both a brain NE-CT and a gadolinium-enhanced T1-weighted (CE-T1W) brain MRI at the time of their brain metastases diagnosis were included. A delay up to two weeks between the NE-CT and the CE-T1W MRI was allowed for inclusion. Patients who had prior craniotomy, those who had significant artefacts on the NE-CT or the CE-T1W MRI, and those whose largest lesion measured less than 5mm in greatest diameter were excluded. Patients with single or multiple metastases were eligible. Imaging studies performed at our institution or elsewhere were allowed for inclusion. For each patient, the primary tumor anatomical site was collected. Medical records at the Radiology department were also reviewed to assemble another cohort of 100 patients (C2) without a cancer diagnosis, who had a NE-CT coupled with a CE-CT, and whose results were strictly normal. The search was conducted in a decreasing order with respect to time, starting at 2020, and reaching 2013 for patients with brain metastases, and 2018 for the cohort of patients with normal brain NE-CT.

### **4.3.2 Segmentation and image pre-processing**

For each patient in cohort C1, the CE-T1W MRI was automatically matched to the NE-CT, and the largest tumor on the CE-T1W MRI was manually segmented by either a senior radiation oncology resident or a radiation oncologist using Eclipse™ treatment planning system (Varian Medical Systems, Palo Alto, CA, USA). Each physician segmented 50% of the cases, and all segmentations were reviewed by both physicians to check and resolve discrepancies. Tumor boundaries were defined as the peripheral enhancement on the CE-T1W MRI. Two shape features - Maximum-3D-Diameter and Mesh-Volume - were extracted using PyRadiomics library version 3.0.1 (118) on Python 3.7.9 to report tumor size characteristics.

The NE-CT images from both cohorts C1 and C2 were converted from DICOM to NIFTI format using the NiBabel library on Python. The brain was automatically segmented for each CT using the SwissSkullStripper tool (142), and the resultant masked brain

image was used thereafter. Sample-wise Z-score normalization was performed for all CT images on both cohorts, then spatial resampling to a 2mm x 2mm x 2mm voxel size was applied using the Advanced Normalization Tools (ANTs) library (112).

### **4.3.3 Deep Learning model development**

Processed NE-CT images from cohorts C1 and C2 were combined into a single dataset, and this was further split into a training set (70%) and a validation set (30%) using a stratified technique.

A 3D-CNN was developed in Python using Keras Application Programming Interface version 2.5.0. The network architecture was inspired by the work of Zunair *et al.* (143). It consists of four 3D convolutional blocks, a fully connected layer, and an output layer. Many variations of this 3D-CNN were created by modifying the convolutional layers, the pooling layers, the dropout rate, and the activation function.

On the fly data augmentation was performed for training using random angle rotations between -20 and +20 degrees. Parameter optimization was done using Adam optimizer with a binary cross-entropy loss function. Training was accelerated with a 16GB NVIDIA Tesla P100 GPU. Early stopping with patience of 15 epochs was employed to avoid overtraining.

Each model was repeatedly trained five times using different dataset shuffling to evaluate the variance and minimize reporting bias. The performance comparison for the best three models was assessed using a one-way ANOVA test on IBM® SPSS version 20. The model architecture yielding the best mean accuracy on the validation set was selected.

The final model performance was evaluated using metrics that included accuracy (ACC), area under receiver operating characteristic curve (ROC-AUC), precision, recall, and F1-score.

## 4.4 RESULTS

### 4.4.1 Patient data

Patient and tumor characteristics are summarized in Table 4.1. The most frequently occurring primary anatomical site was the lung (51%), and 81% of the patients had either a lung or breast malignancy. The median Maximum-3D-Diameter for the largest tumor per-patient was 2.29 cm (IQR: 1.72 - 2.64 cm), while the median Mesh-Volume was 2.81 cc (IQR: 0.93 - 4.20 cc). Twenty-seven percent of the patients had solitary metastases, whereas 49% had more than five metastases.

**Table 4.1** Patients and tumors characteristics

	Number of patients N=100
Primary cancer site	
Lung	51
Breast	30
Gastro-intestinal	6
Genito-urinary	5
Melanoma	5
Gynecologic	3
Number of metastases	
1	27
2	8
3	8
4	5
5	3
>5	49
Maximum 3D diameter - largest lesion	
≤ 1 cm	8
1 - 2 cm	25
2 - 3 cm	53
> 3cm	14
Volume - largest lesion	
≤ 2 cc	39
2 - 4 cc	33
> 4 cc	28

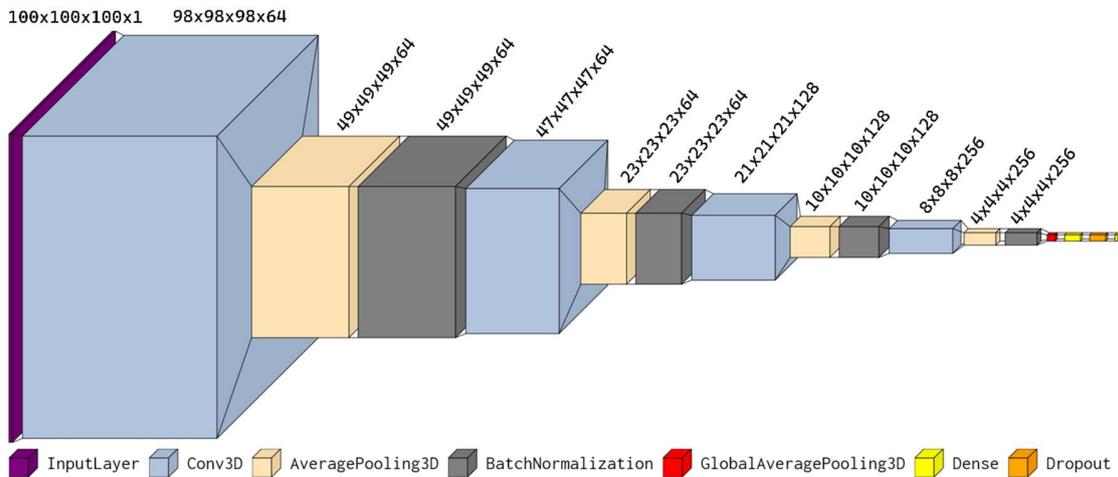
*cc: cubic centimeters*

## 4.4.2 Model evaluation

We tested several variants of the 3D-CNN, and we have chosen to limit our results to the three models that showed the highest ACCs on the validation set: models A, B, and C with respective mean ACCs of 0.983 (SD: 0.020 – 95%CI: 0.957-1.000), 0.970 (SD: 0.021 – 95%CI: 0.943-0.996), and 0.963 (SD: 0.021 – 95%CI: 0.936-0.990).

Although model A yielded numerically the best results, the difference in mean ACC did not reach statistical significance with a p-value of 0.351.

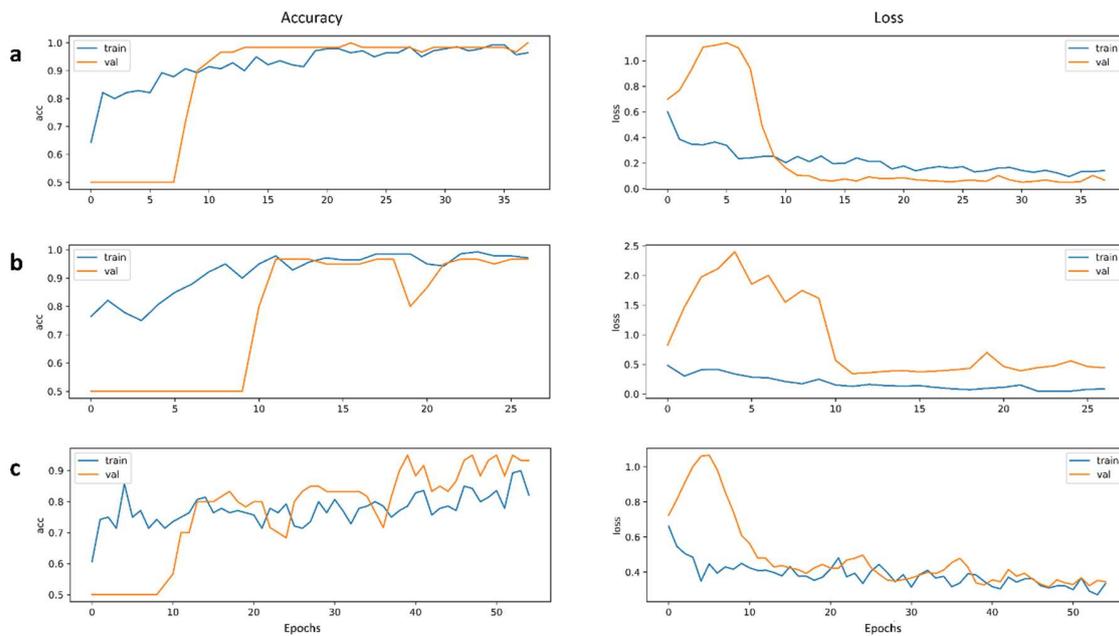
Model A architecture is illustrated in Figure 4.1. Each convolutional block included a convolutional layer with a 3x3x3 kernel size and ReLU activation, a 3D average pooling layer with a stride of 2, and a batch normalization layer. The first two convolutional layers included 64 filters, then 128 filters in the third layer, and 256 in the fourth. The output of the fourth block was passed through a 3D global average pooling layer and fed into the dense layer of 512 neurons. A dropout layer with a rate of 50% and a Sigmoid activation function were used.



**Figure 4.1** Model A network architecture. Numbers indicate the tensor shape of the corresponding layer. Dimensions in the Z-direction (horizontal) were rescaled to 1:100 to allow clear visualization.

Model B differs from model A in that it included 3D maximum pooling layers, rather than 3D average pooling layers, into its convolutional blocks. Model C has the same architecture as model B, with a dropout rate of 30% instead of 50%. Figure 4.2 shows

the learning curves for the three models for both ACC and Loss using the same dataset split. The ROC curves for the corresponding version of model A as well as the training and validation confusion matrices are shown in Figure 4.3.



**Figure 4.2** Training and validation learning curves for accuracy and loss for models A (a), B (b), and C (c). *train*: training; *val*: validation; *acc*: accuracy

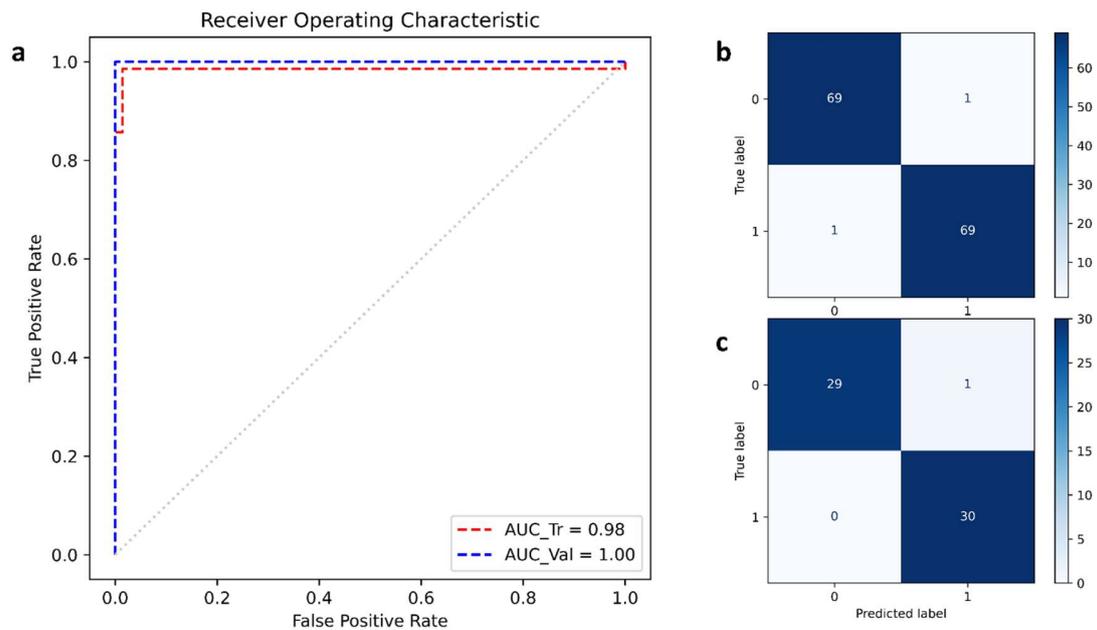
All evaluation metrics for model A are reported in Table 4.2. Mean ROC-AUC was 0.979 (SD: 0.022) on the training set and 0.983 (SD: 0.023) on the validation set. Mean ACC on the training set was 0.985 (SD: 0.011). Mean precision and recall were 0.987 (SD: 0.012) and 0.987 (SD: 0.013) respectively for training, and 0.983 (SD: 0.020) for both on the validation set. Mean F1-scores were 0.987 (SD: 0.013) and 0.983 (SD: 0.020) on the training and validation sets, respectively.

**Table 4.2** Evaluation metrics for model A

	Training	Validation
Accuracy	0.985 ± 0.011	0.983 ± 0.020
ROC-AUC	0.979 ± 0.022	0.983 ± 0.023
Precision	0.987 ± 0.012	0.983 ± 0.020

Recall	$0.987 \pm 0.013$	$0.983 \pm 0.020$
F1-score	$0.987 \pm 0.013$	$0.983 \pm 0.020$

All metrics are reported in the format: Mean  $\pm$  SD; ROC-AUC: area under receiver operating characteristic curve



**Figure 4.3** Receiver operating characteristic curves for model A for the training and validation sets **(a)**; Confusion matrices for model A on the training **(b)** and validation **(c)** sets. Label “0” corresponds to normal CT, and label “1” denotes a CT with brain metastases. *AUC\_Tr*: area under receiver operating characteristic curve for the training set; *AUC\_Val*: area under receiver operating characteristic curve for the validation set.

## 4.5 DISCUSSION

In this work, we developed an accurate 3D-CNN based on NE-CT brain images to distinguish between patients with brain metastases and those with normal brain imaging. To the best of our knowledge, this is the first study to address this specific task.

Many publications have reported on brain metastases detection using MRI (144–151) with a pooled sensitivity of 88.6% (95%CI: 84.4%-92.2%) across studies (50), whereas only one study used CE-CT with a reported sensitivity of 88.1% (51), and one used NE-CT with a sensitivity of 35.4% for lesions larger than 6mm (52). Most of these studies reported on lesion-wise detectability.

In the study by Kato *et al.* (52), 116 patients with 590 brain metastases were included. They trained a 2D SSD model on 3697 axial slices all including abnormalities from 426 lesions, while slices with (835 slices) and without (2709 slices) metastases were included in the test set for a total of 164 lesions. They reported an overall sensitivity of 23.8%. The model could not detect lesions less than 3mm, and had a very low sensitivity of 5.4% for lesions 3-6mm. A total of 176 false-negative and 415 false-positive lesions were reported on the test set. The false positive rate was 15.4 lesions per scan. Of note, the presence of surrounding edema increased the model performance. Considering the high rate of false positive lesions as well as the fact that all included patients had metastases, it is difficult to infer on the capability of this model to identify patients without metastases. However, the results seem to suggest suboptimal performance in this regard, acknowledging that the model was not optimized to tackle this task.

Model A sensitivity in the current study was 98.3% (SD: 2%) on the validation set. We could not compare our findings to those from Kato *et al.* (52) due to different methodology, and to the fact that we performed a patient-wise rather than lesion-wise classification. Although a lesion-wise detection is interesting in terms of locating the metastases, we feel that patients who have a high-probability of having intracranial metastases would be better evaluated with an MRI, and thus, we believe that further lesion characterization on a NE-CT would likely be futile. Therefore, our main objective was to flag NE-CT scans suspicious for metastases that would be further characterized with dedicated imaging.

By contrast to the study by Kato *et al.* (52), we used a 3D network architecture. A 2D network processes each slice independently and might therefore overlook important volumetric information that could help the classification task. However, it is worth noting that 3D-CNNs come with a higher computational cost and require a robust

hardware infrastructure to handle the increased dimensionality and the consequent growth in the size of the parameter space. Additionally, we have used the SwissSkullStripper tool (142) to segment the brain, thus making the whole process fully automated and more practical to use.

In model B, more epochs were needed to reach a solution as evidenced in Figure 2, and the model performed worse than model A with more fluctuations. The main difference between the two models are the pooling layers, where model A included average pooling by contrast to maximum pooling in model B. In the study by Kato *et al* (52), the surrounding edema was important to the model efficacy. We believe that the inferior results might be partially explained by the fact that maximum pooling, at least for some convolutional kernels, would likely miss information from small areas of edema that exhibit low raw values, and therefore negatively affect the model performance. This might be different with contrast-enhanced images where maximum pooling would likely increase the signal-to-background ratio. Model C included less dropout rate and showed more fluctuations in accuracy than models A and B, hinting to some degree of overtraining.

Our study has several limitations. First, the number of patients included is relatively small despite using data augmentation, and we did not perform an external validation to test our model. Also, the model was trained on a balanced dataset that does not reflect the true rate of brain metastases in an unselected dataset, and this could lead to a classification bias. Furthermore, NE-CT scans for patients with metastases were mainly collected from the Radiation Oncology department, whereas those for patients with normal brain were curated from the Radiology medical records. This might introduce significant differences between the two datasets, particularly in terms of CT values, image resolution, surrounding structures, and image orientation. To mitigate this risk, we developed the model on the segmented brain, rotated randomly the images to reduce bias from patient positioning, normalized the images, and performed spatial resampling prior to model training. Nonetheless, we cannot totally exclude residual differences that might have biased the model classification. Moreover, we did not perform feature map visualization to understand the prediction process; this might be done at a later stage to better explain the model work mechanism and the differences between the different models.

In this study, we excluded patients with only very small lesions less than 5mm based on the findings by Kato *et al.* (52). However, patients with multiple lesions were allowed, and these likely included metastases smaller than 5mm. Nonetheless, in the absence of edema that is unusual with very small lesions, we do not believe that there would be a meaningful result in detecting these metastases, more particularly in the context of a 2x2x2mm spatial resampling. Consequently, our results might probably be inapplicable in this setting. Further, asymptomatic lesions smaller than 5mm are

sometimes managed by observation, more so in the setting of CNS-active systemic therapy.

The control cohort herein consisted of strictly normal brain NE-CTs. Subsequently, in the setting of other types of abnormalities, our model might lead to false positive results, or miss serious conditions unrelated to cancer. On these grounds, we must stress on the fact that this CNN could be used as a CAD tool to help detect brain metastases, but never as a standalone tool.

While dedicated brain imaging is not routinely performed for asymptomatic cancer patients, brain NE-CT scans are becoming increasingly available as part of staging PET-CT scans. Therefore, detecting the presence of brain metastases on NE-CT could be of paramount importance to guide further high-yield brain imaging such as MRI, and inform local and systemic therapy. Although we acknowledge that our results might not be directly extrapolated to CT scans done in the scope of a PET-CT due to a lower resolution in the latter, we provide at least a proof of principle that merits further validation to determine its applicability in clinical practice. Furthermore, 67% of the patients in our cohort had lesions larger than 2cm, and 14% had lesions larger than 3cm. Many of these patients were likely symptomatic and many lesions would have been easily detected by the radiologist. Hence, the magnitude of the benefit from implementing this model is yet to be substantiated.

## **4.6 CONCLUSION**

In conclusion, we developed a 3D Convolutional Neural Network based on non-enhanced brain CT scans to distinguish between patients with and without brain metastases. We demonstrated that the model is very accurate for detecting lesions larger than 5mm. This algorithm merits further validation in external cohorts and might be used as a computer-aided diagnosis tool to guide subsequent dedicated brain imaging.



# 5 Conclusion and Future Works

This chapter includes two parts:

- 1- A summary on the main results of the studies included in this thesis, and the contribution to the existing body of literature
- 2- Perspectives and future works

## 5.1 Conclusion

The integration of machine learning and quantum technology into the medical field represents a significant paradigm shift, particularly in the fields of radiology and oncology. These technologies are not just enhancing existing diagnostic and therapeutic processes but are fundamentally reshaping how clinicians approach patient care, research, and personalized medicine. Quantum machine learning represents a groundbreaking convergence of quantum computing and machine learning technologies, offering transformative potential in the field of cancer imaging. Furthermore, quantum machine learning algorithms represent a new class of models that are worth exploring, particularly in the field of oncologic radiology.

In Chapter 3, we discussed the development of a Quantum Neural Network model using contrast-enhanced T1-weighted MRI radiomic features to differentiate between large brain metastases and high-grade gliomas. We have tackled most of the limitations that exist in the current body of literature and have abided by most of the TRIPOD criteria for model development.

We have limited our analysis to large tumors measuring more than 2cm to minimize size-related classification bias. We also performed a lesion-wise analysis to avoid multifocality bias. Our dataset is heterogeneous in terms of imaging techniques and machines used, and despite this being sometimes regarded as a limitation, we believe that heterogeneity might be key to real-life model generalizability.

We have utilized handcrafted radiomic features for the model development to maintain reasonable model interpretability. Moreover, we have used instance-wise Shapley values to explain the model predictions. We believe that model interpretability is a bottleneck for the use of AI models in clinical practice, and efforts in this direction are eagerly needed to facilitate this “bench-to-clinic” transition.

Feature selection is a pivotal step in model building, and failure to capture the best correlations with the outcome variable is one of the main reasons of model under-performance and lack of generalizability. We have employed a model-agnostic Mutual-Information Feature Selection scheme to get the most important features linearly and non-linearly correlated with the prediction class, while considering feature interaction. Since this method is too computationally expensive, we have used a Quantum Annealing method to solve the problem heuristic.

To the best of our knowledge, our study is the first to use a quantum machine learning algorithm for brain tumor classification on a real-world dataset. Our model was

benchmarked against two great classical machine learning models and showed comparable performance. This study paves the way for more utilization of these promising algorithms in the medical field.

Chapter 4 highlighted the development of a 3D Convolutional Neural Network that differentiates between patients with brain metastases and those without abnormalities based on their non-enhanced brain CT scan. This is the first study to address this problem to the extent of our knowledge. Despite many limitations fully discussed above, the model achieved a remarkable accuracy of 98.3% on the validation set. Our results represent at least a proof-of-concept and a starting point to guide efforts in gleaning more information from these under-explored imaging sets.

## 5.2 Perspectives and Future Works

Quantum Machine Learning models constitute a promising novel class of algorithms. We have shown herein that these models at least rival the best-in-class classical models. However, with the rapidly advancing quantum hardware developments, the best is yet to come in terms of the potential benefits that could result from these algorithms, particularly when used on actual fault-tolerant quantum devices in the future, with the ability of including a much larger number of features.

In the current NISQ era, the model we developed could be tested on available quantum computers, such as the IBM platform. We have elected to use a quantum simulator rather than an actual device to avoid a decrease in performance due to decoherence. This has led us to use a limited number of qubits, and thus a limited number of features, due to the inability of a simulator to classically simulate large-qubit quantum systems. Nonetheless, with the increase in number of qubits available on recent devices, it would be interesting to explore the performance of QML algorithms with a much larger number of features, even on a noisy quantum computer.

The 3D CNN model that identifies patients with brain metastases from non-enhanced brain CT scans showed an impressive performance. In subsequent work, we plan to explore feature activation maps to understand the model working mechanism. This would also confer more confidence in using the model in a clinical setting.

For both models developed within this thesis, no external validation was performed. Our hope is to be able to test our models on external datasets to better evaluate their generalizability. So far, this has not been possible mainly due to data-sharing complexity. Our aim is to overcome this hurdle soon. Furthermore, it would be very interesting to compare the results of the models in both studies to the reading results of experienced neuro-radiologists to get a better understanding of the added value of these models in current clinical practice. We hope we will be able to address this in the near future.

To expand our research space, we have started curating a brain MRI dataset to build a machine learning model that predicts the primary cancer site for brain metastases. We aim to use for this purpose advanced techniques including the use of quantum technology. We have also believe that this methodology could be applied to help differentiate radiation necrosis from true progression of brain metastases following stereotactic radiotherapy. This would have a significant impact on the follow-up and further management for these patients.

Undeniably, we are witnessing the « big-bang » of a transformative era in terms of computational technology that will definitely reshape the future of many fields, particularly the field of data science with all its applications.



# 6 Résumé en Français

## CONTEXTE

### Perspective Clinique

Les métastases cérébrales représentent les tumeurs intracrâniennes les plus courantes, avec une incidence annuelle estimée entre 7 et 14 cas pour 100 000 personnes, suivies par les gliomes de haut grade. On estime que 10 à 40 % des patients atteints de tumeurs solides développent des métastases cérébrales, une tendance à la hausse attribuable aux progrès dans le traitement du cancer, l'optimisation des techniques d'imagerie comme l'utilisation de l'IRM et à une population démographiquement vieillissante. Les cancers du poumon et du sein sont les sites primitifs les plus fréquents essaimant sous forme de métastases cérébrales.

L'IRM cérébrale est l'imagerie de choix pour le diagnostic des tumeurs cérébrales, surpassant largement la tomodensitométrie (TDM) cérébrale dans la détection des tumeurs petites et profondes, et fournissant des informations détaillées sur le nombre, l'emplacement et la taille des tumeurs. L'histopathologie demeure toutefois la référence absolue pour la confirmation et la caractérisation des tumeurs cérébrales, malgré la nature invasive des procédures pour obtenir du matériel tissulaire.

L'augmentation de disponibilité de l'imagerie médicale a révolutionné le diagnostic et la gestion des tumeurs cérébrales, avec des modalités avancées telles que l'IRM et les scanners CT utilisés en routine. Cependant, une quantité considérable de données reste inexplorée, principalement en raison des limitations à interpréter de grandes quantités d'informations. L'exploration de techniques computationnelles avancées, comme l'apprentissage automatique et l'intelligence artificielle, présente un large potentiel pour extraire et analyser ces aspects nuancés de l'imagerie des tumeurs cérébrales, menant à des stratégies de traitement plus personnalisées et efficaces.

Cette thèse se concentre sur deux études distinctes :

- 1- Différencier les métastases cérébrales des gliomes de haut grade à l'aide d'images IRM pondérées en T1:

La distinction entre ces deux types de tumeurs présente un défi significatif en raison de leurs similitudes dans les séquences d'imagerie traditionnelles, bien que des différences morphologiques puissent exister. Une caractérisation précise est cruciale car les stratégies thérapeutiques diffèrent largement entre ces entités.

- 2- Identifier les patients atteints de métastases cérébrales à partir d'images TDM cérébrales non-injectées:

La détection précoce des métastases cérébrales est essentielle en raison des complications qui en résultent. Bien que l'IRM soit la modalité d'imagerie la plus sensible pour la détection des métastases cérébrales, les scanners CT cérébraux non rehaussés sont de plus en plus disponibles et pourraient, avec l'aide de méthodes computationnelles, permettre d'identifier les patients présentant des anomalies suggérant des métastases et pouvant bénéficier d'une IRM cérébrale dédiée.

### **Perspective Technique:**

L'utilisation de l'intelligence artificielle (IA) dans le domaine médical présente une croissance significative, avec des applications cliniques déjà en place (Decipher®, ArteraAI®, etc.). L'analyse des images médicales représente un domaine prometteur pour l'IA, permettant d'extraire des données pour la modélisation, le pronostic et le diagnostic des tumeurs.

Les études publiées ont montré une grande variabilité dans la précision de la classification entre les métastases cérébrales et les gliomes de haut grade, avec une précision allant de 64 à 98%. Cette hétérogénéité résulte de différents critères d'inclusion, protocoles d'imagerie, techniques de prétraitement des images, et méthodologies d'apprentissage automatique. Malgré les différences, aucune supériorité n'a été observée entre l'utilisation des séquences d'IRM conventionnelles et avancées ou entre différents classificateurs, y compris les modèles d'apprentissage profond.

L'initiative TRIPOD vise à standardiser le développement et la validation de modèles de prédiction, mais l'adhésion à ces directives reste sous-optimale. De plus, plusieurs biais sont identifiés dans les études, tels l'inclusion de plusieurs métastases par patient, et l'inclusion de métastases de petites tailles, ce qui augmente fictivement la précision des modèles. Bien qu'aucune différence de performance n'ait été identifiée entre les techniques d'apprentissage automatique, l'apprentissage profond semble être supérieur aux techniques de base dans certaines études. La technologie quantique et l'apprentissage automatique quantique offrent des perspectives prometteuses pour l'amélioration de la sélection des variables et de la performance des modèles d'apprentissage, en particulier dans le contexte où le nombre de patients est limité.

Détecter des métastases cérébrales sur des scans CT non-injectés présente d'énormes défis, comme le montre une sensibilité de détection très faible dans les études existantes. Cette tâche a peu de pertinence clinique sans caractérisation supplémentaire par des techniques d'imagerie plus sensibles telle l'IRM, ce qui signifie que l'identification des patients atteints de métastases plutôt que la caractérisation de celles-ci est demandée.

## **OBJECTIFS**

- 1- *Développer un modèle d'apprentissage automatique interprétable basé sur des images IRM-T1 injectées pour différencier entre les larges métastases cérébrales et les gliomes de haut grade.*

À cette fin, nous aborderons les limitations des études publiées. Nous visons à exclure les tumeurs plus petites que 2 cm. Nous utiliserons la radiomique ainsi que les « valeurs Shapley » pour interpréter le modèle. Nous explorerons la technologie quantique, notamment le « Quantum Annealing » pour la sélection des variables et l'apprentissage quantique pour la construction du modèle, et nous comparerons notre modèle aux modèles d'apprentissage classique et d'apprentissage profond les plus avancés.

- 2- *Développer un modèle d'apprentissage profond pour identifier les patients présentant des métastases cérébrales en se basant sur des images de scanner CT non-injectées.*

À cette fin, nous entraînerons notre modèle sur un ensemble de données incluant des patients avec et sans métastases cérébrales. Nous excluons les patients présentant des métastases cérébrales de taille inférieure à 5 mm, compte tenu des résultats des études précédentes. Nous concevrons notre modèle pour une prédiction par-patient et non par-lésion.

### **ÉTUDE 1 : Un réseau neuronal quantique interprétable basé sur la Radiomique des images IRM-T1 injectée pour différencier entre les métastases cérébrales volumineuses et les gliomes de haut grade en utilisant le « Quantum Annealing » pour la sélection des variables**

#### Objectif :

Les métastases cérébrales solitaires de grande taille et les gliomes de haut grade sont parfois difficiles à différencier sur les images IRM. La prise en charge diffère considérablement entre ces deux entités, et des méthodes non invasives permettant de

les différencier sont absolument nécessaires afin d'éviter des biopsies et des interventions chirurgicales potentiellement morbides. Nous explorons ici la performance et l'interprétabilité d'un réseau neuronal quantique variationnel basé sur la radiomique des images IRM-T1 injectée en utilisant une approche de sélection de variable basée sur l'information mutuelle en utilisant le « Quantum Annealing » pour le calcul.

#### Matériels et méthodes :

Nous avons inclus rétrospectivement 423 patients diagnostiqués de gliomes de haut grade ou de métastases cérébrales volumineuses (> 2 cm) entre 2012 et 2019, et qui ont eu une IRM pondérée T1 injectée. Après exclusion, 72 patients atteints de gliome de haut grade et 129 patients atteints de métastases cérébrales ont été retenus pour l'analyse. Les tumeurs ont été segmentées manuellement, et un anneau péri-tumoral de 5 mm a été créé. Les images IRM ont été pré-traitées et 1 813 variables radiomiques ont été extraites.

Un ensemble des variables les plus importantes basées sur l'information mutuelle a été sélectionné. L'information mutuelle et l'information mutuelle conditionnelle ont été intégrées dans une formulation QUBO (« Quadratic Unconstrained Binary Optimization »), qui a été analoguée à un modèle Ising et soumise au processeur quantique de D'Wave pour déterminer la meilleure combinaison de 10 variables.

Les 10 variables sélectionnées ont été intégrées dans un réseau de neurones quantique (QNN) à 2 Qubits à l'aide de la bibliothèque digitale PennyLane. Le modèle a été évalué pour la précision équilibrée (bACC) et la ROC-AUC sur la sous-cohorte de validation. Les performances du modèle ont été comparées à celles de deux modèles classiques : Dense-Neural-Networks (DNN) et Extreme-Gradient-Boosting (XGB). Les valeurs de Shapley ont été calculées pour interpréter les prédictions par échantillon sur la sous-cohorte de validation.

#### Résultats :

La meilleure combinaison de 10 variables comprenait 6 caractéristiques tumorales et 4 caractéristiques d'anneau. Pour QNN, DNN et XGB respectivement : la ROC-AUC d'entraînement était de : 0,86, 0,95 et 0,94 ; la ROC-AUC de validation était : 0,76, 0,75 et 0,79 ; et la bACC de validation était : 0,74, 0,73 et 0,72. Les deux caractéristiques les plus influentes étaient l'Entropie Laplacienne-de-GLRLM-Gaussienne et la sphéricité de la tumeur.

#### Conclusion :

Nous avons développé un modèle QNN interprétable précis avec une méthode de sélection de variables quantique pour différencier les gliomes de haut grade des métastases cérébrales volumineuses sur l'IRM cérébrale T1 injectée. La performance du modèle est comparable à celle des modèles classiques de pointe.

## **ÉTUDE 2 : Un réseau neuronal convolutif 3D basé sur des tomodensitométrie cérébrale non-injectée pour identifier les patients atteints de métastases cérébrales**

### Introduction :

L'imagerie cérébrale pour les patients atteints de cancer est rarement recommandée en l'absence de symptômes. La tomodensitométrie non-injectée du cerveau est de plus en plus réalisée, principalement en raison d'une utilisation plus large de la TEP-CT dans le bilan d'extension du cancer et de l'évaluation de la réponse aux traitements. Les métastases cérébrales sont souvent difficiles à diagnostiquer sur la tomodensitométrie non-injectée. Cette étude vise à développer un réseau neuronal convolutif 3D (3D-CNN) basé sur les tomodensitométries cérébrales non-injectées pour distinguer les patients avec et sans métastases cérébrales.

### Matériels et méthodes :

Nous avons inclus rétrospectivement des tomodensitométries cérébrales non-injectées pour 100 patients présentant des métastases cérébrales, unique ou multiple, et 100 patients présentant une imagerie cérébrale sans anomalie. Les patients dont la plus grande métastase était <5 mm ont été exclus. La plus grosse tumeur a été segmentée manuellement sur une IRM pondérée en T1 avec contraste, et des variables radiomiques de forme ont été extraites pour déterminer la taille et le volume de la lésion. Le cerveau a été automatiquement segmenté et les images masquées ont été normalisées et rééchantillonnées. La cohorte a été divisée en sous-cohortes d'entraînement (70 %) et de validation (30 %). Plusieurs versions d'un 3D-CNN ont été développées et le meilleur modèle a été sélectionné en fonction de la précision (ACC) sur la sous-cohorte de validation.

### Résultats :

Le diamètre 3D maximal médian de la plus grande tumeur était de 2,29 cm et son volume médian était de 2,81 cc. Des métastases solitaires ont été trouvées chez 27% des patients, tandis que 49% avaient plus que 5 métastases. Le meilleur modèle était constitué de quatre couches convolutives avec des couches de regroupement moyen

3D, des couches de « dropout » de 50 % et une fonction d'activation sigmoïde. L'ACC de validation était de 0.983 (SD: 0.020) et la ROC-AUC était de 0.983 (SD: 0.023). La sensibilité était de 0.983 (SD: 0.020).

#### Conclusion :

Nous avons développé un réseau de neurones convolutionnel 3D précis basé sur les tomodensitométries cérébrales non-injectées pour différencier entre les patients avec et sans métastases cérébrales. Le modèle mérite une validation externe plus extensive.

### **CONCLUSION**

Dans le contexte de cette thèse, nous avons développé un modèle d'apprentissage quantique qui différencie avec précision entre les métastases cérébrales volumineuses et les gliomes de haut grade sur les images IRM-T1 injectées, et un réseau de neurones convolutionnel 3D qui identifie les patients atteints de métastases cérébrales à partir des tomodensitométries cérébrales non-injectées.

Dans des études futures, nous pourrions tester le modèle quantique sur un ordinateur quantique, explorer les cartes d'activation du 3D CNN, et valider l'efficacité de ces modèles sur des cohortes externes. Nous avons commencé à construire une base de données d'IRM cérébrales pour prédire le site du cancer primitif à partir des images des métastases cérébrales.



## 7 BIBLIOGRAPHY

1. Ostrom QT, Price M, Neff C, Cioffi G, Waite KA, Kruchko C, *et al.* CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2015–2019. *Neuro-Oncology*. 2022 Oct;24(Supplement\_5):v1–95.
2. Achrol AS, Rennert RC, Anders C, Soffiotti R, Ahluwalia MS, Nayak L, *et al.* Brain metastases. *Nat Rev Dis Primers*. 2019 Jan 17;5(1):5.
3. Kromer C, Xu J, Ostrom QT, Gittleman H, Kruchko C, Sawaya R, *et al.* Estimating the annual frequency of synchronous brain metastasis in the United States 2010-2013: a population-based study. *J Neurooncol*. 2017 Aug;134(1):55–64.
4. Graif M, Bydder GM, Steiner RE, Niendorf P, Thomas DG, Young IR. Contrast-enhanced MR imaging of malignant brain tumors. *AJNR Am J Neuroradiol*. 1985 Dec;6(6):855–62.
5. National Comprehensive Cancer Network, Inc. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) 2023 [Internet]. [cited 2023 Nov 25]. Available from: [nccn.org](http://nccn.org)
6. Erho N, Crisan A, Vergara IA, Mitra AP, Ghadessi M, Buerki C, *et al.* Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. *PLoS One*. 2013;8(6):e66855.
7. Esteva A, Feng J, van der Wal D, Huang SC, Simko JP, DeVries S, *et al.* Prostate cancer therapy personalization via multi-modal deep learning on randomized phase III clinical trials. *NPJ Digit Med*. 2022 Jun 8;5(1):71.
8. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, *et al.* Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012 Mar;48(4):441–6.
9. Sachdeva J, Kumar V, Gupta I, Khandelwal N, Ahuja CK. A package-SFERCB-“Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors.” *Applied Soft Computing*. 2016;47:151–67.
10. Shin I, Kim H, Ahn SS, Sohn B, Bae S, Park JE, *et al.* Development and Validation of a Deep Learning-Based Model to Distinguish Glioblastoma from Solitary Brain Metastasis Using Conventional MR Images. *AJNR Am J Neuroradiol*. 2021 May;42(5):838–44.

11. Ortiz-Ramón R, Ruiz-España S, Mollá-Olmos E, Moratal D. Glioblastomas and brain metastases differentiation following an MRI texture analysis-based radiomics approach. *Phys Med*. 2020 Aug;76:44–54.
12. Qin J, Li Y, Liang D, Zhang Y, Yao W. Histogram analysis of absolute cerebral blood volume map can distinguish glioblastoma from solitary brain metastasis. *Medicine (Baltimore)*. 2019 Oct;98(42):e17515.
13. Chen C, Ou X, Wang J, Guo W, Ma X. Radiomics-Based Machine Learning in Differentiation Between Glioblastoma and Metastatic Brain Tumors. *Front Oncol*. 2019;9:806.
14. Liu Z, Jiang Z, Meng L, Yang J, Liu Y, Zhang Y, *et al*. Handcrafted and Deep Learning-Based Radiomic Models Can Distinguish GBM from Brain Metastasis. Lou JC, editor. *Journal of Oncology*. 2021 Jun 4;2021:5518717.
15. de Causans A, Carré A, Roux A, Tauziède-Espariat A, Ammari S, Dezamis E, *et al*. Development of a Machine Learning Classifier Based on Radiomic Features Extracted From Post-Contrast 3D T1-Weighted MR Images to Distinguish Glioblastoma From Solitary Brain Metastasis. *Frontiers in Oncology [Internet]*. 2021;11. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2021.638262>
16. Payabvash S, Aboian M, Tihan T, Cha S. Machine Learning Decision Tree Models for Differentiation of Posterior Fossa Tumors Using Diffusion Histogram Analysis and Structural MRI Findings. *Frontiers in Oncology [Internet]*. 2020;10. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2020.00071>
17. Samani ZR, Parker D, Wolf R, Hodges W, Brem S, Verma R. Distinct tumor signatures using deep learning-based characterization of the peritumoral microenvironment in glioblastomas and brain metastases. *Scientific Reports*. 2021 Jul 14;11(1):14469.
18. Priya S, Liu Y, Ward C, Le NH, Soni N, Pillenahalli Maheshwarappa R, *et al*. Machine learning based differentiation of glioblastoma from brain metastasis using MRI derived radiomics. *Scientific Reports*. 2021 May 18;11(1):10478.
19. Georgiadis P, Cavouras D, Kalatzis I, Daskalakis A, Kagadis GC, Sifaki K, *et al*. Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features. *Computer Methods and Programs in Biomedicine*. 2008;89(1):24–32.
20. Svolos P, Tsolaki E, Kapsalaki E, Theodorou K, Fountas K, Fezoulidis I, *et al*. Investigating brain tumor differentiation with diffusion and perfusion metrics at 3T MRI using pattern recognition techniques. *Magnetic Resonance Imaging*. 2013;31(9):1567–77.

21. Artzi M, Bressler I, Ben Bashat D. Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis. *J Magn Reson Imaging*. 2019 Aug;50(2):519–28.
22. Yang G, Jones TL, Howe FA, Barrick TR. Morphometric model for discrimination between glioblastoma multiforme and solitary metastasis using three-dimensional shape analysis. *Magn Reson Med*. 2016 Jun;75(6):2505–16.
23. Tsolaki E, Svolos P, Kousi E, Kapsalaki E, Fezoulidis I, Fountas K, *et al*. Fast spectroscopic multiple analysis (FASMA) for brain tumor classification: a clinical decision support system utilizing multi-parametric 3T MR data. *Int J Comput Assist Radiol Surg*. 2015 Jul;10(7):1149–66.
24. Zacharaki EI, Kanas VG, Davatzikos C. Investigating machine learning techniques for MRI-based classification of brain neoplasms. *Int J Comput Assist Radiol Surg*. 2011 Nov;6(6):821–8.
25. Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER, *et al*. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn Reson Med*. 2009 Dec;62(6):1609–18.
26. Meier R, Pahud de Mortanges A, Wiest R, Knecht U. Exploratory Analysis of Qualitative MR Imaging Features for the Differentiation of Glioblastoma and Brain Metastases. *Frontiers in Oncology* [Internet]. 2020;10. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2020.581037>
27. Bae S, An C, Ahn SS, Kim H, Han K, Kim SW, *et al*. Robust performance of deep learning for distinguishing glioblastoma from single brain metastasis using radiomic features: model development and validation. *Scientific Reports*. 2020 Jul 21;10(1):12110.
28. Dong F, Li Q, Jiang B, Zhu X, Zeng Q, Huang P, *et al*. Differentiation of supratentorial single brain metastasis and glioblastoma by using peri-enhancing oedema region-derived radiomic features and multiple classifiers. *European Radiology*. 2020 May 1;30(5):3015–22.
29. Yang G, Jones TL, Barrick TR, Howe FA. Discrimination between glioblastoma multiforme and solitary metastasis using morphological features derived from the p:q tensor decomposition of diffusion tensor imaging. *NMR in Biomedicine*. 2014;27(9):1103–11.
30. Yamashita K, Yoshiura T, Arimura H, Mihara F, Noguchi T, Hiwatashi A, *et al*. Performance Evaluation of Radiologists with Artificial Neural Network for Differential Diagnosis of Intra-Axial Cerebral Tumors on MR Images. *American Journal of Neuroradiology*. 2008;29(6):1153–8.

31. Abidin AZ, Dar IA, DSouza AM, Lin EP, Wismüller A. Medical Imaging: Biomedical Applications in Molecular, Structural, and Functional Imaging - Investigating a quantitative radiomics approach for brain tumor classification [Internet]. Vol. 10953, Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging. SPIE; 2019. 36 p. Available from: <https://lens.org/057-346-840-096-911>
32. Park JE, Kim HS, Lee J, Cheong EN, Shin I, Ahn SS, *et al.* Deep-learned time-signal intensity pattern analysis using an autoencoder captures magnetic resonance perfusion heterogeneity for brain tumor differentiation. *Sci Rep.* 2020 Dec 8;10(1):21485.
33. Tateishi M, Nakaura T, Kitajima M, Uetani H, Nakagawa M, Inoue T, *et al.* An initial experience of machine learning based on multi-sequence texture parameters in magnetic resonance imaging to differentiate glioblastoma from brain metastases. *J Neurol Sci.* 2020 Mar 15;410:116514.
34. Swinburne NC, Schefflein J, Sakai Y, Oermann EK, Titano JJ, Chen I, *et al.* Machine learning for semi-automated classification of glioblastoma, brain metastasis and central nervous system lymphoma using magnetic resonance advanced imaging. *Ann Transl Med.* 2019 Jun;7(11):232.
35. Tsolaki E, Svolos P, Kousi E, Kapsalaki E, Fountas K, Theodorou K, *et al.* Automated differentiation of glioblastomas from intracranial metastases using 3T MR spectroscopic and perfusion data. *International Journal of Computer Assisted Radiology and Surgery.* 2013 Sep 1;8(5):751–61.
36. Shrot S, Salhov M, Dvorski N, Konen E, Averbuch A, Hoffmann C. Application of MR morphologic, diffusion tensor, and perfusion imaging in the classification of brain tumors using machine learning scheme. *Neuroradiology.* 2019 Jul 1;61(7):757–65.
37. Jekel L, Brim WR, von Reppert M, Staib L, Cassinelli Petersen G, Merkaj S, *et al.* Machine Learning Applications for Differentiation of Glioma from Brain Metastasis—A Systematic Review. *Cancers* [Internet]. 2022;14(6). Available from: <https://www.mdpi.com/2072-6694/14/6/1369>
38. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* 2015 Jan 6;162(1):55–63.
39. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S. Quantum machine learning. *Nature.* 2017 Sep 1;549(7671):195–202.
40. Schuld M, Petruccione F. *Supervised Learning with Quantum Computers.* 1st ed. Springer Cham; XIII, 287. (Quantum Science and Technology).

41. Schuld M. Supervised quantum machine learning models are kernel methods [Internet]. arXiv; 2021. Available from: <https://arxiv.org/abs/2101.11020>
42. Dunjko V, Briegel HJ. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Rep Prog Phys*. 2018 Jul;81(7):074001.
43. Liu Y, Arunachalam S, Temme K. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*. 2021 Sep 1;17(9):1013–7.
44. Abbas A, Sutter D, Zoufal C, Lucchi A, Figalli A, Woerner S. The power of quantum neural networks. *Nature Computational Science*. 2021 Jun 1;1(6):403–9.
45. Huang HY, Broughton M, Mohseni M, Babbush R, Boixo S, Neven H, *et al*. Power of data in quantum machine learning. *Nature Communications*. 2021 May 11;12(1):2631.
46. Mott A, Job J, Vlimant JR, Lidar D, Spiropulu M. Solving a Higgs optimization problem with quantum annealing for machine learning. *Nature*. 2017 Oct 1;550(7676):375–9.
47. Caro MC, Huang HY, Cerezo M, Sharma K, Sornborger A, Cincio L, *et al*. Generalization in quantum machine learning from few training data. *Nature Communications*. 2022 Aug 22;13(1):4919.
48. Vergara JR, Estévez PA. A review of feature selection methods based on mutual information. *Neural Computing and Applications*. 2014 Jan 1;24(1):175–86.
49. McGeoch CC, Wang C. Experimental Evaluation of an Adiabatic Quantum System for Combinatorial Optimization. In: *Proceedings of the ACM International Conference on Computing Frontiers* [Internet]. New York, NY, USA: Association for Computing Machinery; 2013. (CF '13). Available from: <https://doi.org/10.1145/2482767.2482797>
50. Ozkara BB, Chen MM, Federau C, Karabacak M, Briere TM, Li J, *et al*. Deep Learning for Detecting Brain Metastases on MRI: A Systematic Review and Meta-Analysis. *Cancers* [Internet]. 2023;15(2). Available from: <https://www.mdpi.com/2072-6694/15/2/334>
51. Amemiya S, Takao H, Kato S, Yamashita H, Sakamoto N, Abe O. Automatic detection of brain metastases on contrast-enhanced CT with deep-learning feature-fused single-shot detectors. *European Journal of Radiology*. 2021;136:109577.
52. Kato S, Amemiya S, Takao H, Yamashita H, Sakamoto N, Abe O. Automated detection of brain metastases on non-enhanced CT using single-shot detectors. *Neuroradiology*. 2021 Dec;63(12):1995–2004.

53. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018 Aug;18(8):500–10.
54. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, *et al*. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res*. 2017 Nov 1;77(21):e104–7.
55. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*. 1973;SMC-3(6):610–21.
56. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, *et al*. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020 May 1;295(2):328–38.
57. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015 May 1;521(7553):436–44.
58. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*. 2018 Aug 1;9(4):611–29.
59. Nielsen MA, Chuang IL. *Quantum Computation and Quantum Information: 10th Anniversary Edition* [Internet]. Cambridge University Press; 2010. Available from: <https://books.google.ca/books?id=-s4DEy7o-a0C>
60. Horodecki R, Horodecki P, Horodecki M, Horodecki K. Quantum entanglement. *Rev Mod Phys*. 2009 Jun;81(2):865–942.
61. Yan F, Ilyyasu AM, Jiang Z. Quantum Computation-Based Image Representation, Processing Operations and Their Applications. *Entropy*. 2014;16(10):5290–338.
62. Shor PW. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal on Computing*. 1997;26(5):1484–509.
63. Bennett CH, Brassard G. Quantum cryptography: Public key distribution and coin tossing. In: *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*. India; 1984. p. 175.
64. Peruzzo A, McClean J, Shadbolt P, Yung MH, Zhou XQ, Love PJ, *et al*. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*. 2014 Jul 23;5(1):4213.

65. Johnson MW, Amin MHS, Gildert S, Lanting T, Hamze F, Dickson N, *et al.* Quantum annealing with manufactured spins. *Nature*. 2011 May 12;473(7346):194–8.
66. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S. Quantum machine learning. *Nature*. 2017 Sep 1;549(7671):195–202.
67. Lloyd S, Garnerone S, Zanardi P. Quantum algorithms for topological and geometric analysis of data. *Nature Communications*. 2016 Jan 25;7(1):10138.
68. Orús R, Mugel S, Lizaso E. Quantum computing for finance: Overview and prospects. *Reviews in Physics*. 2019 Nov 1;4:100028.
69. Born M, Fock V. Beweis des Adiabatenatzes. *Zeitschrift für Physik*. 1928 Mar 1;51(3):165–80.
70. Farhi E, Goldstone J, Gutmann S, Sipser M. *Quantum Computation by Adiabatic Evolution*. 2000.
71. Apolloni B, Carvalho C, Falco D de. Quantum stochastic optimization. *Stochastic Processes and their Applications*. 1989;33(2):233–44.
72. Crosson E, Farhi E, Lin CYY, Lin HH, Shor P. Different Strategies for Optimization Using the Quantum Adiabatic Algorithm. 2014.
73. Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by Simulated Annealing. *Science*. 1983;220(4598):671–80.
74. Razavy M. *Quantum Theory of Tunneling* [Internet]. World Scientific; 2003. (G - Reference, Information and Interdisciplinary Subjects Series). Available from: <https://books.google.ca/books?id=Ek0R62Ahu8UC>
75. Yarkoni S, Raponi E, Bäck T, Schmitt S. Quantum annealing for industry applications: introduction and review. *Reports on Progress in Physics*. 2022 Sep;85(10):104001.
76. Wikipedia contributors. Quantum machine learning — Wikipedia, The Free Encyclopedia [Internet]. 2024. Available from: [https://en.wikipedia.org/w/index.php?title=Quantum\\_machine\\_learning&oldid=1193939929](https://en.wikipedia.org/w/index.php?title=Quantum_machine_learning&oldid=1193939929)
77. Grover LK. Quantum Mechanics Helps in Searching for a Needle in a Haystack. *Phys Rev Lett*. 1997 Jul;79(2):325–8.
78. Kitaev AY. Quantum measurements and the Abelian Stabilizer Problem. 1995.

79. Harrow AW, Hassidim A, Lloyd S. Quantum Algorithm for Linear Systems of Equations. *Phys Rev Lett*. 2009 Oct;103(15):150502.
80. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S. Quantum machine learning. *Nature*. 2017 Sep 13;549(7671):195–202.
81. Scherer A, Valiron B, Mau SC, Alexander S, van den Berg E, Chapuran TE. Concrete resource analysis of the quantum linear-system algorithm used to compute the electromagnetic scattering cross section of a 2D target. *Quantum Information Processing*. 2017 Jan 25;16(3):60.
82. Cerezo M, Arrasmith A, Babbush R, Benjamin SC, Endo S, Fujii K, *et al*. Variational quantum algorithms. *Nature Reviews Physics*. 2021 Sep 1;3(9):625–44.
83. Achrol AS, Rennert RC, Anders C, Soffiatti R, Ahluwalia MS, Nayak L, *et al*. Brain metastases. *Nature Reviews Disease Primers*. 2019 Jan 17;5(1):5.
84. Le Rhun E, Guckenberger M, Smits M, Dummer R, Bachelot T, Sahm F, *et al*. EANO-ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up of patients with brain metastasis from solid tumours. *Ann Oncol*. 2021 Nov;32(11):1332–47.
85. Mohile NA, Messersmith H, Gatson NT, Hottinger AF, Lassman A, Morton J, *et al*. Therapy for Diffuse Astrocytic and Oligodendroglial Tumors in Adults: ASCO-SNO Guideline. *J Clin Oncol*. 2022 Feb 1;40(4):403–26.
86. Weller M, van den Bent M, Preusser M, Le Rhun E, Tonn JC, Minniti G, *et al*. EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nature Reviews Clinical Oncology*. 2021 Mar 1;18(3):170–86.
87. Gondi V, Bauman G, Bradfield L, Burri SH, Cabrera AR, Cunningham DA, *et al*. Radiation Therapy for Brain Metastases: An ASTRO Clinical Practice Guideline. *Pract Radiat Oncol*. 2022 Aug;12(4):265–82.
88. WHO Classification of Tumours Editorial Board. Central Nervous System Tumours: WHO Classification of Tumours. 2021.
89. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ. Perfusion MRI as a diagnostic biomarker for differentiating glioma from brain metastasis: a systematic review and meta-analysis. *European Radiology*. 2018 Sep 1;28(9):3819–31.
90. Tsolaki E, Svolos P, Kousi E, Kapsalaki E, Fountas K, Theodorou K, *et al*. Automated differentiation of glioblastomas from intracranial metastases using 3T MR spectroscopic and perfusion data. *International Journal of Computer Assisted Radiology and Surgery*. 2013 Sep 1;8(5):751–61.

91. Blasel S, Jurcoane A, Franz K, Morawe G, Pellikan S, Hattingen E. Elevated peritumoural rCBV values as a mean to differentiate metastases from high-grade gliomas. *Acta Neurochirurgica*. 2010 Nov 1;152(11):1893–9.
92. Lin L, Xue Y, Duan Q, Sun B, Lin H, Huang X, *et al*. The role of cerebral blood flow gradient in peritumoral edema for differentiation of glioblastomas from solitary metastatic lesions. *Oncotarget*. 2016 Oct 18;7(42):69051–9.
93. Bauer AH, Erly W, Moser FG, Maya M, Nael K. Differentiation of solitary brain metastasis from glioblastoma multiforme: a predictive multiparametric approach using combined MR diffusion and perfusion. *Neuroradiology*. 2015 Jul;57(7):697–703.
94. Kadota Y, Hirai T, Azuma M, Hattori Y, Khant ZA, Hori M, *et al*. Differentiation between glioblastoma and solitary brain metastasis using neurite orientation dispersion and density imaging. *J Neuroradiol*. 2020 May;47(3):197–202.
95. Suh CH, Kim HS, Jung SC, Kim SJ. Diffusion-Weighted Imaging and Diffusion Tensor Imaging for Differentiating High-Grade Glioma from Solitary Brain Metastasis: A Systematic Review and Meta-Analysis. *AJNR Am J Neuroradiol*. 2018 Jul;39(7):1208–14.
96. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016 Feb;278(2):563–77.
97. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, *et al*. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012 Mar;48(4):441–6.
98. Afshar P, Mohammadi A, Plataniotis KN, Oikonomou A, Benali H. From Handcrafted to Deep-Learning-Based Cancer Radiomics: Challenges and Opportunities. *IEEE Signal Processing Magazine*. 2019 Jul;36(4):132–60.
99. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, *et al*. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLoS Med*. 2018 Nov;15(11):e1002711.
100. Carneiro G, Oakden-Rayner L, Bradley AP, Nascimento J, Palmer L. Automated 5-year mortality prediction using deep learning and radiomics features from chest computed tomography. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). 2017. p. 130–4.
101. Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, *et al*. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal*. 2017 Jan;35:303–12.

102. Bae S, An C, Ahn SS, Kim H, Han K, Kim SW, *et al.* Robust performance of deep learning for distinguishing glioblastoma from single brain metastasis using radiomic features: model development and validation. *Scientific Reports*. 2020 Jul 21;10(1):12110.
103. Valdez F, Melin P. A review on quantum computing and deep learning algorithms and their applications. *Soft comput*. 2022 Apr 7;1–20.
104. Schuld M, Sinayskiy I, Petruccione F. The quest for a Quantum Neural Network. *Quantum Information Processing*. 2014 Nov 1;13(11):2567–86.
105. Brown G, Pocock A, Zhao MJ, Luján M. Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection. *J Mach Learn Res*. 2012 Jan;13(1):27–66.
106. Nguyen XV, Chan J, Romano S, Bailey J. Effective global approaches for mutual information based feature selection. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2014;
107. Jekel L, Brim WR, von Reppert M, Staib L, Cassinelli Petersen G, Merkaj S, *et al.* Machine Learning Applications for Differentiation of Glioma from Brain Metastasis- A Systematic Review. *Cancers (Basel)*. 2022 Mar 8;14(6).
108. Rastogi R, Chaturvedi DK, Sagar S, Tandon N, Rastogi AR. Deep Learning Application in Classification of Brain Metastases: Sensor Usage in Medical Diagnosis for Next Gen Healthcare. In: Nandan Mohanty S, Chatterjee JM, Satpathy S, editors. *Internet of Things and Its Applications [Internet]*. Cham: Springer International Publishing; 2022. p. 117–35. Available from: [https://doi.org/10.1007/978-3-030-77528-5\\_6](https://doi.org/10.1007/978-3-030-77528-5_6)
109. Dikaios N. Deep learning magnetic resonance spectroscopy fingerprints of brain tumours using quantum mechanically synthesised data. *NMR Biomed*. 2021 Apr;34(4):e4479.
110. Naeem A, Anees T, Naqvi RA, Loh WK. A Comprehensive Analysis of Recent Deep and Federated-Learning-Based Methodologies for Brain Tumor Diagnosis. *Journal of Personalized Medicine [Internet]*. 2022;12(2). Available from: <https://www.mdpi.com/2075-4426/12/2/275>
111. Stadlbauer A, Heinz G, Marhold F, Meyer-Bäse A, Ganslandt O, Buchfelder M, *et al.* Differentiation of Glioblastoma and Brain Metastases by MRI-Based Oxygen Metabolomic Radiomics and Deep Learning. *Metabolites [Internet]*. 2022;12(12). Available from: <https://www.mdpi.com/2218-1989/12/12/1264>

112. Avants BB, Tustison NJ, Johnson HJ. Advanced Normalization Tools [Internet]. Available from: <http://stnava.github.io/ANTs/>
113. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, *et al.* N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. 2010 Jun;29(6):1310–20.
114. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002 Nov;17(3):143–55.
115. Isensee F, Schell M, Pflueger I, Brugnara G, Bonekamp D, Neuberger U, *et al.* Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*. 2019;40(17):4952–64.
116. Reinhold JC, Dewey BE, Carass A, Prince JL. Evaluating the Impact of Intensity Normalization on MR Image Synthesis. *Proc SPIE Int Soc Opt Eng*. 2019 Mar;10949:109493H.
117. Carré A, Klausner G, Edjlali M, Lerousseau M, Briend-Diop J, Sun R, *et al.* Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. *Sci Rep*. 2020 Jul 23;10(1):12340.
118. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, *et al.* Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res*. 2017 Nov 1;77(21):e104–7.
119. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative [Internet]. Available from: <https://doi.org/10.48550/arXiv.1612.07003>
120. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, *et al.* The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020 May;295(2):328–38.
121. Pasvolsky J, Stevanovic R, Condello A, Tong H. Feature Selection Notebook [Internet]. D-Wave; 2021. Available from: <https://github.com/dwave-examples/feature-selection-notebook>
122. Bergholm V, Izaac J, Schuld M, Gogolin C, Ahmed S, Ajith V, *et al.* PennyLane: Automatic differentiation of hybrid quantum-classical computations [Internet]. arXiv; 2018. Available from: <https://arxiv.org/abs/1811.04968>
123. Farhi E, Neven H. Classification with Quantum Neural Networks on Near Term Processors [Internet]. arXiv; 2018. Available from: <https://arxiv.org/abs/1802.06002>

124. Schuld M, Bocharov A, Svore KM, Wiebe N. Circuit-centric quantum classifiers. *Physical Review A* [Internet]. 2020 Mar;101(3). Available from: [https://doi.org/10.1103%2Fphysreva.101.032308](https://doi.org/10.1103/PhysRevA.101.032308)
125. Schuld M. Variational Classifier [Internet]. PennyLane; 2021. Available from: [https://github.com/PennyLaneAI/qml/blob/master/demonstrations/tutorial\\_variational\\_classifier.py](https://github.com/PennyLaneAI/qml/blob/master/demonstrations/tutorial_variational_classifier.py)
126. Möttönen M, Vartiainen JJ, Bergholm V, Salomaa MM. Transformation of Quantum States Using Uniformly Controlled Rotations. *Quantum Info Comput.* 2005 Sep;5(6):467–73.
127. Shapley LS. Notes on the n-Person Game -- II: The Value of an n-Person Game [Internet]. Santa Monica, Calif.: RAND Corporation; 1951. Available from: [https://www.rand.org/content/dam/rand/pubs/research\\_memoranda/2008/RM670.pdf](https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM670.pdf)
128. The Shapley Value: Essays in Honor of Lloyd S. Shapley. Cambridge University Press; 1988.
129. Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, *et al.*, editors. *Advances in Neural Information Processing Systems 30* [Internet]. Curran Associates, Inc.; 2017. p. 4765–74. Available from: <http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>
130. Tsai CF, Hsiao YC. Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. *Decision Support Systems.* 2010;50(1):258–69.
131. Rokach L, Chizi B, Maimon O. Feature Selection by Combining Multiple Methods. In: Last M, Szczepaniak PS, Volkovich Z, Kandel A, editors. *Advances in Web Intelligence and Data Mining* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006. p. 295–304. Available from: [https://doi.org/10.1007/3-540-33880-2\\_30](https://doi.org/10.1007/3-540-33880-2_30)
132. de Causans A, Carré A, Roux A, Tauziède-Espariat A, Ammari S, Dezamis E, *et al.* Development of a Machine Learning Classifier Based on Radiomic Features Extracted From Post-Contrast 3D T1-Weighted MR Images to Distinguish Glioblastoma From Solitary Brain Metastasis. *Front Oncol.* 2021;11:638262.
133. Lee J, Steinmann A, Ding Y, Lee H, Owens C, Wang J, *et al.* Radiomics feature robustness as measured using an MRI phantom. *Scientific Reports.* 2021 Feb 17;11(1):3973.

134. Priya S, Liu Y, Ward C, Le NH, Soni N, Pillenahalli Maheshwarappa R, *et al.* Machine learning based differentiation of glioblastoma from brain metastasis using MRI derived radiomics. *Scientific Reports*. 2021 May 18;11(1):10478.
135. Smedby KE, Brandt L, Bäcklund ML, Blomqvist P. Brain metastases admissions in Sweden between 1987 and 2006. *Br J Cancer*. 2009 Dec 1;101(11):1919–24.
136. Schouten LJ, Rutten J, Huveneers HAM, Twijnstra A. Incidence of brain metastases in a cohort of patients with carcinoma of the breast, colon, kidney, and lung and melanoma. *Cancer*. 2002 May 15;94(10):2698–705.
137. Vogelbaum MA, Brown PD, Messersmith H, Brastianos PK, Burri S, Cahill D, *et al.* Treatment for Brain Metastases: ASCO-SNO-ASTRO Guideline. *Journal of Clinical Oncology*. 2022;40(5):492–516.
138. Gondi V, Bauman G, Bradfield L, Burri SH, Cabrera AR, Cunningham DA, *et al.* Radiation Therapy for Brain Metastases: An ASTRO Clinical Practice Guideline. *Pract Radiat Oncol*. 2022 Aug;12(4):265–82.
139. Li Y, Jin G, Su D. Comparison of Gadolinium-enhanced MRI and 18FDG PET/PET-CT for the diagnosis of brain metastases in lung cancer patients: A meta-analysis of 5 prospective studies. *Oncotarget*. 2017 May 30;8(22):35743–9.
140. Cho SJ, Sunwoo L, Baik SH, Bae YJ, Choi BS, Kim JH. Brain metastasis detection using machine learning: a systematic review and meta-analysis. *Neuro Oncol*. 2021 Feb 25;23(2):214–25.
141. Kato S, Amemiya S, Takao H, Yamashita H, Sakamoto N, Miki S, *et al.* Computer-aided detection improves brain metastasis identification on non-enhanced CT in less experienced radiologists. *Acta Radiol*. 2023 May;64(5):1958–65.
142. Bauer S, Fejes T, Reyes M. A Skull-Stripping Filter for ITK. 2013 Jan; Available from: <http://hdl.handle.net/10380/3353>
143. Zunair H, Rahman A, Mohammed N, Cohen JP. Uniformizing Techniques to Process CT Scans with 3D CNNs for Tuberculosis Prediction. In: Rekik I, Adeli E, Park SH, Valdés Hernández M del C, editors. *Predictive Intelligence in Medicine*. Cham: Springer International Publishing; 2020. p. 156–68.
144. Yoo SK, Kim TH, Chun J, Choi BS, Kim H, Yang S, *et al.* Deep-Learning-Based Automatic Detection and Segmentation of Brain Metastases with Small Volume for Stereotactic Ablative Radiotherapy. *Cancers (Basel)*. 2022 May 23;14(10).

145. Xue J, Wang B, Ming Y, Liu X, Jiang Z, Wang C, *et al.* Deep learning-based detection and segmentation-assisted management of brain metastases. *Neuro Oncol.* 2020 Apr 15;22(4):505–14.
146. Pflüger I, Wald T, Isensee F, Schell M, Meredig H, Schlamp K, *et al.* Automated detection and quantification of brain metastases on clinical MRI data using artificial neural networks. *Neurooncol Adv.* 2022 Dec;4(1):vdac138.
147. Liang Y, Lee K, Bovi JA, Palmer JD, Brown PD, Gondi V, *et al.* Deep Learning-Based Automatic Detection of Brain Metastases in Heterogenous Multi-Institutional Magnetic Resonance Imaging Sets: An Exploratory Analysis of NRG-CC001. *Int J Radiat Oncol Biol Phys.* 2022 Nov 1;114(3):529–36.
148. Kikuchi Y, Togao O, Kikuchi K, Momosaka D, Obara M, Van Cauteren M, *et al.* A deep convolutional neural network-based automatic detection of brain metastases with and without blood vessel suppression. *Eur Radiol.* 2022 May;32(5):2998–3005.
149. Huang Y, Bert C, Sommer P, Frey B, Gaipl U, Distel LV, *et al.* Deep learning for brain metastasis detection and segmentation in longitudinal MRI data. *Med Phys.* 2022 Sep;49(9):5773–86.
150. Chartrand G, Emiliani RD, Pawlowski SA, Markel DA, Bahig H, Cengarle-Samak A, *et al.* Automated Detection of Brain Metastases on T1-Weighted MRI Using a Convolutional Neural Network: Impact of Volume Aware Loss and Sampling Strategy. *J Magn Reson Imaging.* 2022 Dec;56(6):1885–98.
151. Amemiya S, Takao H, Kato S, Yamashita H, Sakamoto N, Abe O. Feature-fusion improves MRI single-shot deep learning detection of small brain metastases. *J Neuroimaging.* 2022 Jan;32(1):111–9.