KINGDOM OF MOROCCO
MOHAMMED V UNIVERSITY OF RABAT
FACULTY OF MEDECINE
AND PHARMACY
RABAT

Year : 2020 **Thesis N°: 77**

Image classification of gliomas using machine learning: A RADIOMICS APPROACH

THESIS

Publicly submitted and defended on the : / /2020

BY

Mr. Mohamed Sobhi JABAL
Born on May 21st, 1991 in Rabat

FOR THE DEGREE

Doctor of Medicine

Key Words : Glioma; Neuroimaging; Radiomics; Artificial Intelligence

Jury Members:

Mr. Azeddine IBRAHIMI Professor of Medical Biotechnology	**President**
Mrs. Mahjouba BOUTARBOUCH Professor of Neuroanatomy	**Director**
Mr. Amine BENKABBOU Professor of General Surgery	**member**
Mr. Mohammed Anass Majbar Professor of General Surgery	**member**

سبحانك لا علم لنا إلا ما علمتنا
إنك أنت العليم الحكيم

سورة البقرة: الآية: 31

صدق الله العظيم

UNIVERSITE MOHAMMED V

FACULTE DE MEDECINE ET DE PHARMACIE
RABAT

DOYENS HONORAIRES :

1962 – 1969: Professeur Abdelmalek FARAJ
1969 – 1974: Professeur Abdellatif BERBICH
1974 – 1981: Professeur Bachir LAZRAK
1981 – 1989: Professeur Taieb CHKILI
1989 – 1997: Professeur Mohamed Tahar ALAOUI
1997 – 2003: Professeur Abdelmajid BELMAHI
2003 - 2013: Professeur Najia HAJJAJ – HASSOUNI

ADMINISTRATION :

Doyen **Professeur Mohamed ADNAOUI**

Vice-*Doyen chargé des Affaires Académiques et Estudiantines*
Professeur Brahim LEKEHAL

Vice-Doyen chargé de la Recherche et de la Coopération
Professeur Toufiq DAKKA

Vice-Doyen chargé des Affaires Spécifiques à la Pharmacie
Professeur Younes RAHALI

Secrétaire Général
Mr. Mohamed KARRA

**** Enseignants Militaires***

1 - ENSEIGNANTS-CHERCHEURS MEDECINS ET PHARMACIENS

PROFESSEURS DE L'ENSEIGNEMENT SUPERIEUR :

Décembre 1984

Pr. MAAOUNI Abdelaziz	Médecine Interne – ***Clinique Royale***
Pr. MAAZOUZI Ahmed Wajdi	Anesthésie -Réanimation
Pr. SETTAF Abdellatif	Pathologie Chirurgicale

Décembre 1989

Pr. ADNAOUI Mohamed	Médecine Interne –***Doyen de la FMPR***
Pr. OUAZZANI Taïbi Mohamed Réda	Neurologie

Janvier et Novembre 1990

Pr. KHARBACH Aîcha	Gynécologie -Obstétrique
Pr. TAZI Saoud Anas	Anesthésie Réanimation

Février Avril Juillet et Décembre 1991

Pr. AZZOUZI Abderrahim	Anesthésie Réanimation- ***Doyen de FMPO***
Pr. BAYAHIA Rabéa	Néphrologie
Pr. BELKOUCHI Abdelkader	Chirurgie Générale
Pr. BENCHEKROUN Belabbes Abdellatif	Chirurgie Générale
Pr. BENSOUDA Yahia	Pharmacie galénique
Pr. BERRAHO Amina	Ophtalmologie
Pr. BEZAD Rachid	Gynécologie Obstétrique ***Méd. Chef Maternité des Orangers***
Pr. CHERRAH Yahia	Pharmacologie
Pr. CHOKAIRI Omar	Histologie Embryologie
Pr. KHATTAB Mohamed	Pédiatrie
Pr. SOULAYMANI Rachida	Pharmacologie- ***Dir. du Centre National PV Rabat***
Pr. TAOUFIK Jamal	Chimie thérapeutique

Décembre 1992

Pr. AHALLAT Mohamed	Chirurgie Générale ***Doyen de FMPT***
Pr. BENSOUDA Adil	Anesthésie Réanimation
Pr. CHAHED OUAZZANI Laaziza	Gastro-Entérologie
Pr. CHRAIBI Chafiq	Gynécologie Obstétrique
Pr. EL OUAHABI Abdessamad	Neurochirurgie
Pr. FELLAT Rokaya	Cardiologie
Pr. JIDDANE Mohamed	Anatomie
Pr. TAGHY Ahmed	Chirurgie Générale
Pr. ZOUHDI Mimoun	Microbiologie

**** Enseignants Militaires***

Mars 1994

Pr. BENJAAFAR Noureddine	Radiothérapie
Pr. BEN RAIS Nozha	Biophysique
Pr. CAOUI Malika	Biophysique
Pr. CHRAIBI Abdelmjid	Endocrinologie et Maladies Métaboliques ***Doyen de la FMPA***
Pr. EL AMRANI Sabah	Gynécologie Obstétrique
Pr. ERROUGANI Abdelkader	Chirurgie Générale – ***Directeur du CHIS***
Pr. ESSAKALI Malika	Immunologie
Pr. ETTAYEBI Fouad	Chirurgie Pédiatrique
Pr. IFRINE Lahssan	Chirurgie Générale
Pr. RHRAB Brahim	Gynécologie –Obstétrique
Pr. SENOUCI Karima	Dermatologie

Mars 1994

Pr. ABBAR Mohamed*	Urologie ***Inspecteur du SSM***
Pr. BENTAHILA Abdelali	Pédiatrie
Pr. BERRADA Mohamed Saleh	Traumatologie – Orthopédie
Pr. CHERKAOUI Lalla Ouafae	Ophtalmologie
Pr. LAKHDAR Amina	Gynécologie Obstétrique
Pr. MOUANE Nezha	Pédiatrie

Mars 1995

Pr. ABOUQUAL Redouane	Réanimation Médicale
Pr. AMRAOUI Mohamed	Chirurgie Générale
Pr. BAIDADA Abdelaziz	Gynécologie Obstétrique
Pr. BARGACH Samir	Gynécologie Obstétrique
Pr. EL MESNAOUI Abbes	Chirurgie Générale
Pr. ESSAKALI HOUSSYNI Leila	Oto-Rhino-Laryngologie
Pr. IBEN ATTYA ANDALOUSSI Ahmed	Urologie
Pr. OUAZZANI CHAHDI Bahia	Ophtalmologie
Pr. SEFIANI Abdelaziz	Génétique
Pr. ZEGGWAGH Amine Ali	Réanimation Médicale

Décembre 1996

Pr. BELKACEM Rachid	Chirurgie Pédiatrie
Pr. BOULANOUAR Abdelkrim	Ophtalmologie
Pr. EL ALAMI EL FARICHA EL Hassan	Chirurgie Générale
Pr. GAOUZI Ahmed	Pédiatrie
Pr. OUZEDDOUN Naima	Néphrologie
Pr. ZBIR EL Mehdi*	Cardiologie ***Directeur HMI Mohammed V***

*** Enseignants Militaires**

Novembre 1997

Pr. ALAMI Mohamed Hassan	Gynécologie-Obstétrique
Pr. BIROUK Nazha	Neurologie
Pr. FELLAT Nadia	Cardiologie
Pr. KADDOURI Noureddine	Chirurgie Pédiatrique
Pr. KOUTANI Abdellatif	Urologie
Pr. LAHLOU Mohamed Khalid	Chirurgie Générale
Pr. MAHRAOUI CHAFIQ	Pédiatrie
Pr. TOUFIQ Jallal	Psychiatrie ***Directeur Hôp.Ar-razi Salé***
Pr. YOUSFI MALKI Mounia	Gynécologie Obstétrique

Novembre 1998

Pr. BENOMAR ALI	Neurologie ***Doyen de la FMP Abulcassis***
Pr. BOUGTAB	Abdesslam Chirurgie Générale
Pr. ER RIHANI Hassan	Oncologie Médicale
Pr. BENKIRANE Majid*	Hématologie

Janvier 2000

Pr. ABID Ahmed*	Pneumo-phtisiologie
Pr. AIT OUAMAR Hassan	Pédiatrie
Pr. BENJELLOUN Dakhama Badr.Sououd	Pédiatrie
Pr. BOURKADI Jamal-Eddine	Pneumo-phtisiologie ***Directeur Hôp. My Youssef***
Pr. CHARIF CHEFCHAOUNI Al Montacer	Chirurgie Générale
Pr. ECHARRAB El Mahjoub	Chirurgie Générale
Pr. EL FTOUH Mustapha	Pneumo-phtisiologie
Pr. EL MOSTARCHID Brahim*	Neurochirurgie
Pr. TACHINANTE Rajae	Anesthésie-Réanimation
Pr. TAZI MEZALEK Zoubida	Médecine Interne

Novembre 2000

Pr. AIDI Saadia	Neurologie
Pr. AJANA Fatima Zohra	Gastro-Entérologie
Pr. BENAMR Said	Chirurgie Générale
Pr. CHERTI Mohammed	Cardiologie
Pr. ECH-CHERIF EL KETTANI Selma	Anesthésie-Réanimation
Pr. EL HASSANI Amine	Pédiatrie - ***Directeur Hôp.Cheikh Zaid***
Pr. EL KHADER Khalid	Urologie
Pr. GHARBI Mohamed El Hassan	Endocrinologie et Maladies Métaboliques
Pr. MDAGHRI ALAOUI Asmae	Pédiatrie

*** Enseignants Militaires**

Décembre 2001

Pr. BALKHI Hicham*	Anesthésie-Réanimation
Pr. BENABDELJLIL Maria	Neurologie
Pr. BENAMAR Loubna	Néphrologie
Pr. BENAMOR Jouda	Pneumo-phtisiologie
Pr. BENELBARHDADI Imane	Gastro-Entérologie
Pr. BENNANI Rajae	Cardiologie
Pr. BENOUACHANE Thami	Pédiatrie
Pr. BEZZA Ahmed*	Rhumatologie
Pr. BOUCHIKHI IDRISSI Med Larbi	Anatomie
Pr. BOUMDIN El Hassane*	Radiologie
Pr. CHAT Latifa	Radiologie
Pr. DAALI Mustapha*	Chirurgie Générale
Pr. EL HIJRI Ahmed	Anesthésie-Réanimation
Pr. EL MAAQILI Moulay Rachid	Neuro-Chirurgie
Pr. EL MADHI Tarik	Chirurgie-Pédiatrique
Pr. EL OUNANI Mohamed	Chirurgie Générale
Pr. ETTAIR Said	Pédiatrie - ***Directeur Hôp. Univ. Cheikh Khalifa***
Pr. GAZZAZ Miloudi*	Neuro-Chirurgie
Pr. HRORA Abdelmalek	Chirurgie Générale ***Directeur Hôpital Ibn Sina***
Pr. KABIRI EL Hassane*	Chirurgie Thoracique
Pr. LAMRANI Moulay Omar	Traumatologie Orthopédie
Pr. LEKEHAL Brahim	Chirurgie Vasculaire Périphérique **V-*D chargé Aff Acad. Est*.**
Pr. MEDARHRI Jalil	Chirurgie Générale
Pr. MIKDAME Mohammed*	Hématologie Clinique
Pr. MOHSINE Raouf	Chirurgie Générale
Pr. NOUINI Yassine	Urologie
Pr. SABBAH Farid	Chirurgie Générale
Pr. SEFIANI Yasser	Chirurgie Vasculaire Périphérique
Pr. TAOUFIQ BENCHEKROUN Soumia	Pédiatrie

Décembre 2002

Pr. AL BOUZIDI Abderrahmane*	Anatomie Pathologique
Pr. AMEUR Ahmed *	Urologie
Pr. AMRI Rachida	Cardiologie
Pr. AOURARH Aziz*	Gastro-Entérologie ***Dir.-Adj. HMI Mohammed V***
Pr. BAMOU Youssef *	Biochimie-Chimie
Pr. BELMEJDOUB Ghizlene*	Endocrinologie et Maladies Métaboliques
Pr. BENZEKRI Laila	Dermatologie
Pr. BENZZOUBEIR Nadia	Gastro-Entérologie
Pr. BERNOUSSI Zakiya	Anatomie Pathologique

*** Enseignants Militaires**

Pr. CHOHO Abdelkrim *	Chirurgie Générale
Pr. CHKIRATE Bouchra	Pédiatrie
Pr. EL ALAMI EL Fellous Sidi Zouhair	Chirurgie Pédiatrique
Pr. EL HAOURI Mohamed *	Dermatologie
Pr. FILALI ADIB Abdelhai	Gynécologie Obstétrique
Pr. HAJJI Zakia	Ophtalmologie
Pr. JAAFAR Abdeloihab*	Traumatologie Orthopédie
Pr. KRIOUILE Yamina	Pédiatrie
Pr. MOUSSAOUI RAHALI Driss*	Gynécologie Obstétrique
Pr. OUJILAL Abdelilah	Oto-Rhino-Laryngologie
Pr. RAISS Mohamed	Chirurgie Générale
Pr. SIAH Samir *	Anesthésie Réanimation
Pr. THIMOU Amal	Pédiatrie
Pr. ZENTAR Aziz*	Chirurgie Générale

Janvier 2004

Pr. ABDELLAH El Hassan	Ophtalmologie
Pr. AMRANI Mariam	Anatomie Pathologique
Pr. BENBOUZID Mohammed Anas	Oto-Rhino-Laryngologie
Pr. BENKIRANE Ahmed*	Gastro-Entérologie
Pr. BOULAADAS Malik	Stomatologie et Chirurgie Maxillo-faciale
Pr. BOURAZZA Ahmed*	Neurologie
Pr. CHAGAR Belkacem*	Traumatologie Orthopédie
Pr. CHERRADI Nadia	Anatomie Pathologique
Pr. EL FENNI Jamal*	Radiologie
Pr. EL HANCHI ZAKI	Gynécologie Obstétrique
Pr. EL KHORASSANI Mohamed	Pédiatrie
Pr. HACHI Hafid	Chirurgie Générale
Pr. JABOUIRIK Fatima	Pédiatrie
Pr. KHARMAZ Mohamed	Traumatologie Orthopédie
Pr. MOUGHIL Said	Chirurgie Cardio-Vasculaire
Pr. OUBAAZ Abdelbarre *	Ophtalmologie
Pr. TARIB Abdelilah*	Pharmacie Clinique
Pr. TIJAMI Fouad	Chirurgie Générale
Pr. ZARZUR Jamila	Cardiologie

Janvier 2005

Pr. ABBASSI Abdellah	Chirurgie Réparatrice et Plastique
Pr. ALLALI Fadoua	Rhumatologie
Pr. AMAZOUZI Abdellah	Ophtalmologie
Pr. BAHIRI Rachid	Rhumatologie
Pr. BARKAT Amina	Pédiatrie

Directeur Hôp. Al Ayachi Salé

*** Enseignants Militaires**

Pr. BENYASS Aatif	Cardiologie
Pr. DOUDOUH Abderrahim*	Biophysique
Pr. HAJJI Leila	Cardiologie *(mise en disponibilité)*
Pr. HESSISSEN Leila	Pédiatrie
Pr. JIDAL Mohamed*	Radiologie
Pr. LAAROUSSI Mohamed	Chirurgie Cardio-vasculaire
Pr. LYAGOUBI Mohammed	Parasitologie
Pr. SBIHI Souad	Histo-Embryologie Cytogénétique
Pr. ZERAIDI Najia	Gynécologie Obstétrique

AVRIL 2006

Pr. ACHEMLAL Lahsen*	Rhumatologie
Pr. BELMEKKI Abdelkader*	Hématologie
Pr. BENCHEIKH Razika	O.R.L
Pr. BIYI Abdelhamid*	Biophysique
Pr. BOUHAFS Mohamed El Amine	Chirurgie - Pédiatrique
Pr. BOULAHYA Abdellatif*	Chirurgie Cardio – Vasculaire. ***Directeur Hôpital Ibn Sina Marr.***
Pr. CHENGUETI ANSARI Anas	Gynécologie Obstétrique
Pr. DOGHMI Nawal	Cardiologie
Pr. FELLAT Ibtissam	Cardiologie
Pr. FAROUDY Mamoun	Anesthésie Réanimation
Pr. HARMOUCHE Hicham	Médecine Interne
Pr. IDRISS LAHLOU Amine*	Microbiologie
Pr. JROUNDI Laila	Radiologie
Pr. KARMOUNI Tariq	Urologie
Pr. KILI Amina	Pédiatrie
Pr. KISRA Hassan	Psychiatrie
Pr. KISRA Mounir	Chirurgie – Pédiatrique
Pr. LAATIRIS Abdelkader*	Pharmacie Galénique
Pr. LMIMOUNI Badreddine*	Parasitologie
Pr. MANSOURI Hamid*	Radiothérapie
Pr. OUANASS Abderrazzak	Psychiatrie
Pr. SAFI Soumaya*	Endocrinologie
Pr. SOUALHI Mouna	Pneumo – Phtisiologie
Pr. TELLAL Saida*	Biochimie
Pr. ZAHRAOUI Rachida	Pneumo – Phtisiologie

Octobre 2007

Pr. ABIDI Khalid	Réanimation médicale
Pr. ACHACHI Leila	Pneumo phtisiologie
Pr. ACHOUR Abdessamad*	Chirurgie générale

*** Enseignants Militaires**

Pr. AIT HOUSSA Mahdi *	Chirurgie cardio vasculaire
Pr. AMHAJJI Larbi *	Traumatologie orthopédie
Pr. AOUFI Sarra	Parasitologie
Pr. BAITE Abdelouahed *	Anesthésie réanimation
Pr. BALOUCH Lhousaine *	Biochimie-chimie
Pr. BENZIANE Hamid *	Pharmacie clinique
Pr. BOUTIMZINE Nourdine	Ophtalmologie
Pr. CHERKAOUI Naoual *	Pharmacie galénique
Pr. EHIRCHIOU Abdelkader *	Chirurgie générale
Pr. EL BEKKALI Youssef *	Chirurgie cardio-vasculaire
Pr. EL ABSI Mohamed	Chirurgie générale
Pr. EL MOUSSAOUI Rachid	Anesthésie réanimation
Pr. EL OMARI Fatima	Psychiatrie
Pr. GHARIB Noureddine	Chirurgie plastique et réparatrice
Pr. HADADI Khalid *	Radiothérapie
Pr. ICHOU Mohamed *	Oncologie médicale
Pr. ISMAILI Nadia	Dermatologie
Pr. KEBDANI Tayeb	Radiothérapie
Pr. LOUZI Lhoussain *	Microbiologie
Pr. MADANI Naoufel	Réanimation médicale
Pr. MAHI Mohamed *	Radiologie
Pr. MARC Karima	Pneumo phtisiologie
Pr. MASRAR Azlarab	Hématologie biologique
Pr. MRANI Saad *	Virologie
Pr. OUZZIF Ez zohra *	Biochimie-chimie
Pr. RABHI Monsef *	Médecine interne
Pr. RADOUANE Bouchaib*	Radiologie
Pr. SEFFAR Myriame	Microbiologie
Pr. SEKHSOKH Yessine *	Microbiologie
Pr. SIFAT Hassan *	Radiothérapie
Pr. TABERKANET Mustafa *	Chirurgie vasculaire périphérique
Pr. TACHFOUTI Samira	Ophtalmologie
Pr. TAJDINE Mohammed Tariq*	Chirurgie générale
Pr. TANANE Mansour *	Traumatologie-orthopédie
Pr. TLIGUI Houssain	Parasitologie
Pr. TOUATI Zakia	Cardiologie

Mars 2009

Pr. ABOUZAHIR Ali *	Médecine interne
Pr. AGADR Aomar *	Pédiatrie
Pr. AIT ALI Abdelmounaim *	Chirurgie Générale
Pr. AKHADDAR Ali *	Neuro-chirurgie

*** Enseignants Militaires**

Pr. ALLALI Nazik	Radiologie
Pr. AMINE Bouchra	Rhumatologie
Pr. ARKHA Yassir	Neuro-chirurgie ***Directeur Hôp.des Spécialités***
Pr. BELYAMANI Lahcen *	Anesthésie Réanimation
Pr. BJIJOU Younes	Anatomie
Pr. BOUHSAIN Sanae *	Biochimie-chimie
Pr. BOUI Mohammed *	Dermatologie
Pr. BOUNAIM Ahmed *	Chirurgie Générale
Pr. BOUSSOUGA Mostapha *	Traumatologie-orthopédie
Pr. CHTATA Hassan Toufik *	Chirurgie Vasculaire Périphérique
Pr. DOGHMI Kamal *	Hématologie clinique
Pr. EL MALKI Hadj Omar	Chirurgie Générale
Pr. EL OUENNASS Mostapha*	Microbiologie
Pr. ENNIBI Khalid *	Médecine interne
Pr. FATHI Khalid	Gynécologie obstétrique
Pr. HASSIKOU Hasna *	Rhumatologie
Pr. KABBAJ Nawal	Gastro-entérologie
Pr. KABIRI Meryem	Pédiatrie
Pr. KARBOUBI Lamya	Pédiatrie
Pr. LAMSAOURI Jamal *	Chimie Thérapeutique
Pr. MARMADE Lahcen	Chirurgie Cardio-vasculaire
Pr. MESKINI Toufik	Pédiatrie
Pr. MESSAOUDI Nezha *	Hématologie biologique
Pr. MSSROURI Rahal	Chirurgie Générale
Pr. NASSAR Ittimade	Radiologie
Pr. OUKERRAJ Latifa	Cardiologie
Pr. RHORFI Ismail Abderrahmani *	Pneumo-Phtisiologie

Octobre 2010

Pr. ALILOU Mustapha	Anesthésie réanimation
Pr. AMEZIANE Taoufiq*	Médecine Interne ***Directeur ERSSM***
Pr. BELAGUID Abdelaziz	Physiologie
Pr. CHADLI Mariama*	Microbiologie
Pr. CHEMSI Mohamed*	Médecine Aéronautique
Pr. DAMI Abdellah*	Biochimie- Chimie
Pr. DARBI Abdellatif*	Radiologie
Pr. DENDANE Mohammed Anouar	Chirurgie Pédiatrique
Pr. EL HAFIDI Naima	Pédiatrie
Pr. EL KHARRAS Abdennasser*	Radiologie
Pr. EL MAZOUZ Samir	Chirurgie Plastique et Réparatrice

*** Enseignants Militaires**

Pr. EL SAYEGH Hachem	Urologie
Pr. ERRABIH Ikram	Gastro-Entérologie
Pr. LAMALMI Najat	Anatomie Pathologique
Pr. MOSADIK Ahlam	Anesthésie Réanimation
Pr. MOUJAHID Mountassir*	Chirurgie Générale
Pr. NAZIH Mouna*	Hématologie
Pr. ZOUAIDIA Fouad	Anatomie Pathologique

Decembre 2010

Pr. ZNATI Kaoutar	Anatomie Pathologique

Mai 2012

Pr. AMRANI Abdelouahed	Chirurgie pédiatrique
Pr. ABOUELALAA Khalil *	Anesthésie Réanimation
Pr. BENCHEBBA Driss *	Traumatologie-orthopédie
Pr. DRISSI Mohamed *	Anesthésie Réanimation
Pr. EL ALAOUI MHAMDI Mouna	Chirurgie Générale
Pr. EL OUAZZANI Hanane *	Pneumophtisiologie
Pr. ER-RAJI Mounir	Chirurgie Pédiatrique
Pr. JAHID Ahmed	Anatomie Pathologique
Pr. RAISSOUNI Maha *	Cardiologie

Février 2013

Pr. AHID Samir	Pharmacologie
Pr. AIT EL CADI Mina	Toxicologie
Pr. AMRANI HANCHI Laila	Gastro-Entérologie
Pr. AMOR Mourad	Anesthésie Réanimation
Pr. AWAB Almahdi	Anesthésie Réanimation
Pr. BELAYACHI Jihane	Réanimation Médicale
Pr. BELKHADIR Zakaria Houssain	Anesthésie Réanimation
Pr. BENCHEKROUN Laila	Biochimie-Chimie
Pr. BENKIRANE Souad	Hématologie
Pr. BENNANA Ahmed*	Informatique Pharmaceutique
Pr. BENSGHIR Mustapha *	Anesthésie Réanimation
Pr. BENYAHIA Mohammed *	Néphrologie
Pr. BOUATIA Mustapha	Chimie Analytique et Bromatologie
Pr. BOUABID Ahmed Salim*	Traumatologie orthopédie
Pr. BOUTARBOUCH Mahjouba	Anatomie
Pr. CHAIB Ali *	Cardiologie
Pr. DENDANE Tarek	Réanimation Médicale

*** Enseignants Militaires**

Pr. DINI Nouzha *	Pédiatrie
Pr. ECH-CHERIF EL KETTANI Mohamed Ali	Anesthésie Réanimation
Pr. ECH-CHERIF EL KETTANI Najwa	Radiologie
Pr. ELFATEMI Nizare	Neuro-chirurgie
Pr. EL GUERROUJ Hasnae	Médecine Nucléaire
Pr. EL HARTI Jaouad	Chimie Thérapeutique
Pr. EL JAOUDI Rachid *	Toxicologie
Pr. EL KABABRI Maria	Pédiatrie
Pr. EL KHANNOUSSI Basma	Anatomie Pathologique
Pr. EL KHLOUFI Samir	Anatomie
Pr. EL KORAICHI Alae	Anesthésie Réanimation
Pr. EN-NOUALI Hassane *	Radiologie
Pr. ERRGUIG Laila	Physiologie
Pr. FIKRI Meryem	Radiologie
Pr. GHFIR Imade	Médecine Nucléaire
Pr. IMANE Zineb	Pédiatrie
Pr. IRAQI Hind	Endocrinologie et maladies métaboliques
Pr. KABBAJ Hakima	Microbiologie
Pr. KADIRI Mohamed *	Psychiatrie
Pr. LATIB Rachida	Radiologie
Pr. MAAMAR Mouna Fatima Zahra	Médecine Interne
Pr. MEDDAH Bouchra	Pharmacologie
Pr. MELHAOUI Adyl	Neuro-chirurgie
Pr. MRABTI Hind	Oncologie Médicale
Pr. NEJJARI Rachid	Pharmacognosie
Pr. OUBEJJA Houda	Chirugie Pédiatrique
Pr. OUKABLI Mohamed *	Anatomie Pathologique
Pr. RAHALI Younes	Pharmacie Galénique ***Vice-Doyen à la Pharmacie***
Pr. RATBI Ilham	Génétique
Pr. RAHMANI Mounia	Neurologie
Pr. REDA Karim *	Ophtalmologie
Pr. REGRAGUI Wafa	Neurologie
Pr. RKAIN Hanan	Physiologie
Pr. ROSTOM Samira	Rhumatologie
Pr. ROUAS Lamiaa	Anatomie Pathologique
Pr. ROUIBAA Fedoua *	Gastro-Entérologie
Pr SALIHOUN Mouna	Gastro-Entérologie
Pr. SAYAH Rochde	Chirurgie Cardio-Vasculaire
Pr. SEDDIK Hassan *	Gastro-Entérologie
Pr. ZERHOUNI Hicham	Chirurgie Pédiatrique
Pr. ZINE Ali *	Traumatologie Orthopédie

*** Enseignants Militaires**

AVRIL 2013

Pr. EL KHATIB MOHAMED KARIM *	Stomatologie et Chirurgie Maxillo-faciale

MARS 2014

Pr. ACHIR Abdellah	Chirurgie Thoracique
Pr. BENCHAKROUN Mohammed *	Traumatologie- Orthopédie
Pr. BOUCHIKH Mohammed	Chirurgie Thoracique
Pr. EL KABBAJ Driss *	Néphrologie
Pr. EL MACHTANI IDRISSI Samira *	Biochimie-Chimie
Pr. HARDIZI Houyam	Histologie- Embryologie-Cytogénétique
Pr. HASSANI Amale *	Pédiatrie
Pr. HERRAK Laila	Pneumologie
Pr. JANANE Abdellah *	Urologie
Pr. JEAIDI Anass *	Hématologie Biologique
Pr. KOUACH Jaouad*	Génycologie-Obstétrique
Pr. LEMNOUER Abdelhay*	Microbiologie
Pr. MAKRAM Sanaa *	Pharmacologie
Pr. OULAHYANE Rachid*	Chirurgie Pédiatrique
Pr. RHISSASSI Mohamed Jaafar	CCV
Pr. SEKKACH Youssef*	Médecine Interne
Pr. TAZI MOUKHA Zakia	Génécologie-Obstétrique

DECEMBRE 2014

Pr. ABILKACEM Rachid*	Pédiatrie
Pr. AIT BOUGHIMA Fadila	Médecine Légale
Pr. BEKKALI Hicham *	Anesthésie-Réanimation
Pr. BENAZZOU Salma	Chirurgie Maxillo-Faciale
Pr. BOUABDELLAH Mounya	Biochimie-Chimie
Pr. BOUCHRIK Mourad*	Parasitologie
Pr. DERRAJI Soufiane*	Pharmacie Clinique
Pr. DOBLALI Taoufik	Microbiologie
Pr. EL AYOUBI EL IDRISSI Ali	Anatomie
Pr. EL GHADBANE Abdedaim Hatim*	Anesthésie-Réanimation
Pr. EL MARJANY Mohammed*	Radiothérapie
Pr. FEJJAL Nawfal	Chirurgie Réparatrice et Plastique
Pr. JAHIDI Mohamed*	O.R.L
Pr. LAKHAL Zouhair*	Cardiologie
Pr. OUDGHIRI NEZHA	Anesthésie-Réanimation
Pr. RAMI Mohamed	Chirurgie Pédiatrique
Pr. SABIR Maria	Psychiatrie
Pr. SBAI IDRISSI Karim*	Médecine préventive, santé publique et Hyg.

*** Enseignants Militaires**

AOUT 2015

Pr. MEZIANE Meryem	Dermatologie
Pr. TAHIRI Latifa	Rhumatologie

PROFESSEURS AGREGES :

JANVIER 2016

Pr. BENKABBOU Amine	Chirurgie Générale
Pr. EL ASRI Fouad*	Ophtalmologie
Pr. ERRAMI Noureddine*	O.R.L
Pr. NITASSI Sophia	O.R.L

JUIN 2017

Pr. ABBI Rachid*	Microbiologie
Pr. ASFALOU Ilyasse*	Cardiologie
Pr. BOUAYTI El Arbi*	Médecine préventive, santé publique et Hyg.
Pr. BOUTAYEB Saber	Oncologie Médicale
Pr. EL GHISSASSI Ibrahim	Oncologie Médicale
Pr. HAFIDI Jawad	Anatomie
Pr. OURAINI Saloua*	O.R.L
Pr. RAZINE Rachid	Médecine préventive, santé publique et Hyg.
Pr. ZRARA Abdelhamid*	Immunologie

NOVEMBRE 2018

Pr. AMELLAL Mina	Anatomie
Pr. SOULY Karim	Microbiologie
Pr. TAHRI Rajae	Histologie-Embryologie-Cytogénétique

NOVEMBRE 2019

Pr. AATIF Taoufiq *	Néphrologie
Pr. ACHBOUK Abdelhafid *	Chirurgie Réparatrice et Plastique
Pr. ANDALOUSSI SAGHIR Khalid *	Radiothérapie
Pr. BABA HABIB Moulay Abdellah *	Gynécologie-obstétrique
Pr. BASSIR RIDA ALLAH	Anatomie
Pr. BOUATTAR TARIK	Néphrologie
Pr. BOUFETTAL MONSEF	Anatomie
Pr. BOUCHENTOUF Sidi Mohammed *	Chirurgie Générale
Pr. BOUZELMAT Hicham *	Cardiologie
Pr. BOUKHRIS Jalal *	Traumatologie-orthopédie

*** Enseignants Militaires**

Pr. CHAFRY Bouchaib *	Traumatologie-orthopédie
Pr. CHAHDI Hafsa *	Anatolmie Pathologique
Pr. CHERIF EL ASRI Abad *	Neurochirugie
Pr. DAMIRI Amal *	Anatolmie Pathologique
Pr. DOGHMI Nawfal *	Anesthésie-réanimation
Pr. ELALAOUI Sidi-Yassir	Pharmacie Galénique
Pr. EL ANNAZ Hicham *	Virologie
Pr. EL HASSANI Moulay EL Mehdi *	Gynécologie-obstétrique
Pr. EL HJOUJI Aabderrahman *	Chirurgie Générale
Pr. EL KAOUI Hakim *	Chirurgie Générale
Pr. EL WALI Abderrahman *	Anesthésie-réanimation
Pr. EN-NAFAA Issam *	Radiologie
Pr. HAMAMA Jalal *	Stomatologie et Chirurgie Maxillo-faciale
Pr. HEMMAOUI Bouchaib *	O.R.L
Pr. HJIRA Naoufal *	Dermatologie
Pr. JIRA Mohamed *	Médecine Interne
Pr. JNIENE Asmaa	Physiologie
Pr. LARAQUI Hicham *	Chirurgie Générale
Pr. MAHFOUD Tarik *	Oncologie Médicale
Pr. MEZIANE Mohammed *	Anesthésie-réanimation
Pr. MOUTAKI ALLAH Younes *	Chirurgie Cardio-vasculaire
Pr. MOUZARI Yassine *	Ophtalmologie
Pr. NAOUI Hafida *	Parasitologie-Mycologie
Pr. OBTEL Majdouline	Médecine préventive, santé publique et Hyg.
Pr. OURRAI Abdelhakim *	Pédiatrie
Pr. SAOUAB Rachida *	Radiologie
Pr. SBITTI Yassir *	Oncologie Médicale
Pr. ZADDOUG Omar *	Traumatologie Orthopédie
Pr. ZIDOUH Saad *	Anesthésie-réanimation

*** Enseignants Militaires**

2 - ENSEIGNANTS-CHERCHEURS SCIENTIFIQUES

PROFESSEURS/Prs. HABILITES

Pr. ABOUDRAR Saadia	Physiologie
Pr. ALAMI OUHABI Naima	Biochimie-chimie
Pr. ALAOUI KATIM	Pharmacologie
Pr. ALAOUI SLIMANI Lalla Naïma	Histologie-Embryologie
Pr. ANSAR M'hammed	Chimie Organique et Pharmacie Chimique
Pr .BARKIYOU Malika	Histologie-Embryologie
Pr. BOUHOUCHE Ahmed	Génétique Humaine
Pr. BOUKLOUZE Abdelaziz	Applications Pharmaceutiques
Pr. CHAHED OUAZZANI Lalla Chadia	Biochimie-chimie
Pr. DAKKA Taoufiq	Physiologie
Pr. FAOUZI Moulay El Abbes	Pharmacologie
Pr. IBRAHIMI Azeddine	Biologie moléculaire/Biotechnologie
Pr. KHANFRI Jamal Eddine	Biologie
Pr. OULAD BOUYAHYA IDRISSI Med	Chimie Organique
Pr. REDHA Ahlam	Chimie
Pr. TOUATI Driss	Pharmacognosie
Pr. YAGOUBI Maamar	Environnement,Eau et Hygiène
Pr. ZAHIDI Ahmed	Pharmacologie

Mise à jour le 11/06/2020
KHALED Abdellah
Chef du Service des Ressources Humaines
FMPR

*** Enseignants Militaires**

Dédicaces

A dieu

Aux âmes de mes grands-mères et grands-pères.

A mes chers parents

Les deux êtres les plus chers du monde, pour leur amour inconditionnel, leur soutien constant et patience depuis mon enfance.

A Mon père

Ma principale inspiration et l'homme qui m'a appris mes principes, qui m'a tout donné, qui m'a protégé et porté de l'aide pas à pas depuis toujours.

A Ma mère

Mon premier amour, l'épaule qui m'a toujours soutenue, rassurée, écouté et comblé avec sa tendresse dans tous mes états,

A ma chère sœur Aya

Le pilier de la famille pour son soutien direct et indirect, merci de m'avoir accompagné durant mon enfance.

A ma chère sœur Zouha

La fontaine de joie de la famille qui colorie notre vie.

À ma chère sœur Rama

Le génie de la famille, qui a toujours été là pour moi, qui m'inspire, me motive, m'encourage et me pousse à toujours aller plus loin.

A ma tante Lalla Nouzha

Qui est pour moi comme une deuxième mère

A mon oncle Lahoucine

Qui m'a toujours inspiré par son esprit et orienté par ses conseils

Mes plus chères tantes Lalla Fatiha, Lalla Khadija , Lalla Oum Keltoum, et Lalla Sanae d'avoir été là pour moi.

A mes chers tentes et oncles, Ghazwa, Mohammad, et Wasef

Pour tous les beaux souvenirs et leur support

A mon cousin et meilleur ami Moulay Zaidane

A mon cousin et meilleur ami d'enfance et d'adolescence Soubhi

A mes cousins

Pour leur amitié, leurs encouragements et les bons souvenirs.

A mes meilleurs amis Azeddine, Yassir, Oussama,

Qui étaient les frères que je n'ai jamais eu qui m'accompagnent dans ce voyage sur terre

A ma meilleure amie Sara,

A mes meilleurs amis

Yaman, abdallah, Naser, Faten

A ma famille, amis et mentors

Je leur dois tout ce que je suis devenu. Il est principalement grâce à dieu et à eux que je suis là où je suis dans ma vie.

Remerciements

A mon maître et président de thèse

Mr. Azzedine Ibrahimi

Professor of Medical Biotechnology

Je vous remercie de l'honneur que vous m'avez fait en acceptant de présider ce jury. Je suis très redevable au Pr. Azeddine Ibrahimi, président du laboratoire de biotechnologie de Rabat, UMV.

Je vous remercie pour votre gentillesse, vos qualités humaines, et votre modestie qui n'a d'égal que votre compétence.

A mon maître et rapporteur de thèse

Mrs. Mahjouba BOUTARBOUCH

Professor of Neuroanatomy

Un merci spécial au merveilleux mentor et neurochirurgien, le Pr. Mahjouba Boutarbouch, Professeur de Neuroanatomie. Merci pour votre soutien et votre support sans fin, ainsi que votre supervision et vos encouragements incessants tout au long de mon parcours académique.

A mon maître et juge de thèse

Mr. Amine BENKABBOU

Professor of General Surgery

Pour le grand honneur que vous m'avez fait en acceptant de juger ce travail, et pour vos encouragements à l'égard de mes projets.

A mon maître et juge de thèse

Mr. Mohammed Anass Majbar

Professor of General Surgery

Nous étions énormément marqués par votre sérieux, votre compétence et votre culture scientifique, vous êtes pour nous un exemple à suivre.

Nous vous remercions du grand honneur que vous nous faites en acceptant de juger ce travail.

Enfin, je tiens à remercier le corps enseignant et administratif de la faculté de médecine et de pharmacie de Rabat pour tout l'enseignement que j'ai reçu et pour m'avoir donné l'opportunité de vivre une expérience inoubliable.

Abbreviations

Abbreviations

Adam : Adaptive Moment Estimation

ATRX : Alpha Thalassemia/Mental Retardation Syndrome X-Linked

AUC : Area Under the Curve

BraTS : Brain Tumor Segmentation

CNN : Convolutional Neural Network

CNS : Central Nervous System

CT : Computed Tomography

DCNN : Deep Convolutional Neural Network

DL : Deep Learning

ET : Enhanced Tumor

FLAIR : Fluid Attenuated Inversion Recovery

FN : False Negative

FP : False Positive

GBM : Glioblastoma Multiforme

GDC : Genomic Data Commons

GLCM : Grey-Level Co-occurrence Matrix

GLRLM : Grey-Level Run-Length Matrix

HGG : High-Grade Glioma

IARC : International Agency for Research on Cancer

IDH : Isocitrate Dehydrogenase

KNN : K-Nearest Neighbour

LFS : Li-Fraumeni

LGG : Low-Grade Glioma

LR : Logistic Regression

MGMT : O-6-methylguanine-DNA methyltransferase

ML : Machine Learning

MR : Magnetic Resonance

MRI : Magnetic Resonance Imaging

MRS : Magnetic Resonance Spectroscopy

NAA : N-acetylaspartate

NaN : Not a Number

NET : Non Enhanced Tumor

NIfTI : Neuroimaging Informatics Technology Initiative

OS : Overall survival

PCA : Principal Component Analysis

PFS : Progression Free Survival

PPV : Positive Predictive Value

ReLU : Rectified Linear Unit

RF : Random Forest

RGB : Red Green Blue

ROC : Receiver operating characteristic

SEGA : Subependymal Giant-cell Astrocytoma

SVD : Single Value Decomposition

TC **:** Tumor Core

TCGA **:** The Cancer Genome Atlas

TCIA **:** The Cancer Imaging Archive

TE **:** Echo Time

TN **:** True Negative

TP **:** True Positive

TR **:** Repetition Time

VGG **:** Visual Geometry Group

VOI **:** Volume of Interest

WHO **:** World Health Organization

WT **:** Whole Tumor

2D **:** Two Dimensional

3D **:** Three Dimensional

Table of Figures

TABLE OF FIGURES

Table of Content

TABLE OF CONTENT

Introduction

1 Introduction

1.1 Gliomas

Overview

The term Glioma refers to tumors stemming generally from the brain that are classically thought to arise from Central Nervous System (CNS) cells called “glial”, which provide mechanical, functional, and nutrition support to neurons. Gliomas are a broad heterogeneous group of tumors, encompassing astrocytomas, glioblastomas (astrocytoma grade IV), oligodendrogliomas, mixed gliomas (oligoastrocytomas), ependymomas, and other few remaining rare types.

Tumorigenesis

Two different hypotheses addressed the origin of gliomas (figure 1), a classical hypothesis postulating that cancer cells arise from normal mature glial cells with accumulating alterations, leading to their dedifferentiation and transformation into neoplasms. The more recent hypothesis suggests that cancer cells originate directly from stem cells, or progenitor undifferentiated cells [1-4].

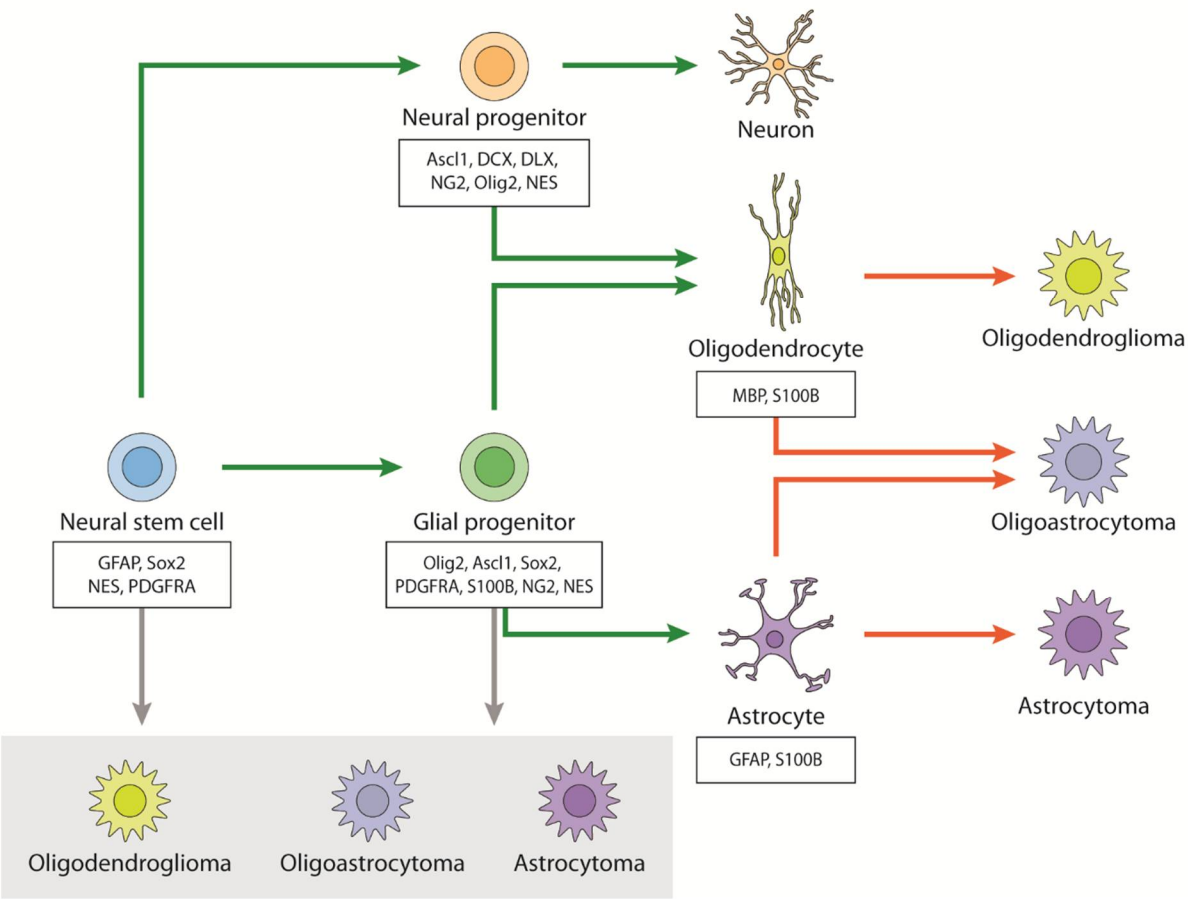


Figure 1: Differentiation process of neural stem cells and origin of gliomas. Normal differentiation (green arrows), classical hypothesis (orange arrows), most recent hypothesis (grey arrows) [2]

Epidemiology

Brain and Central Nervous System (CNS) neoplasms in general are highly heterogeneous and variable, they are globally ranked 20th in incidence rate at 1.6% and 13th in mortality rate at 2.5% amongst all cancers (figure 2, figure 3) [6]. In Morocco, according to the cancer registry of the greater Casablanca region for the period 2008 - 2012, central nervous system tumors have an incidence rate of 1.2% [7]. Distribution of brain cancers in Rabat from 2006 to

2008, is 2.7% in men, in contrast to 1.3% in women [8]. They generally range from benign to malignant tumors and can be organized into two major groups based on their site: primary brain tumors originate in the brain, and secondary or metastatic. Malignant brain tumors are relatively rare, though they are responsible for a disproportionately high mortality and morbidity rates compared to other cancers [9]. This could be due in part to the reduced effect of systemic chemotherapy owing to the presence of the blood-brain barrier. In addition to the reserved prognosis of malignant gliomas attributed to their infiltrative nature [10].

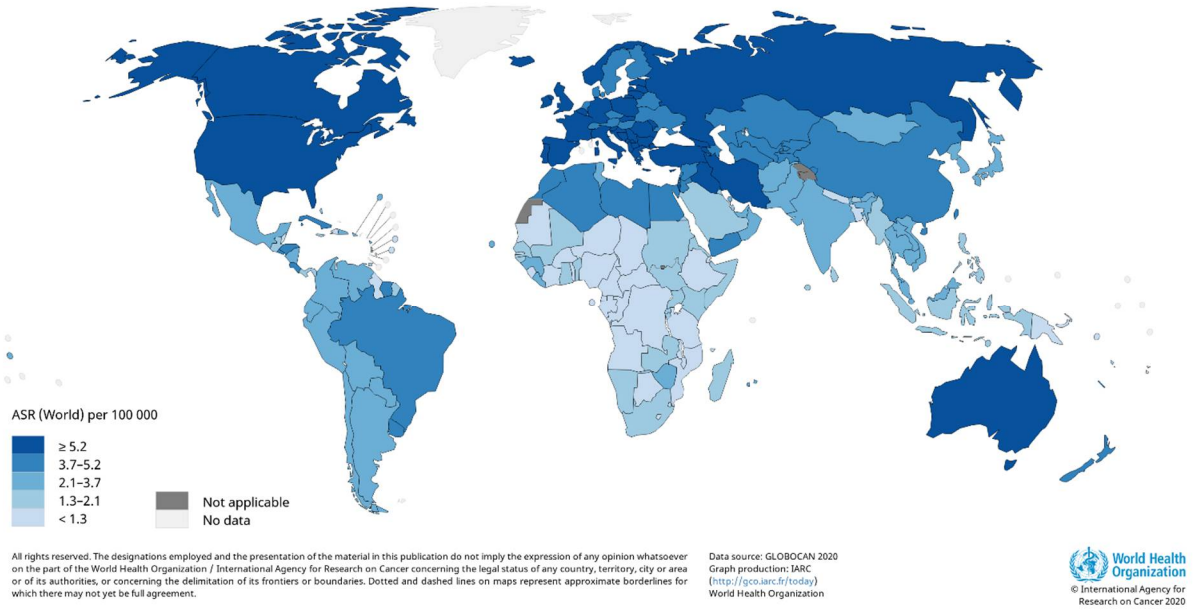


Figure 2: Estimated age-standardized incidence rates (World) in 2020, brain, central nervous system, both sexes, all ages. [6]

Gliomas are recognized to be the most common primary brain tumors, representing globally almost 2% of human malignancies, and are responsible for 2.3% of cancer deaths [11, 12].

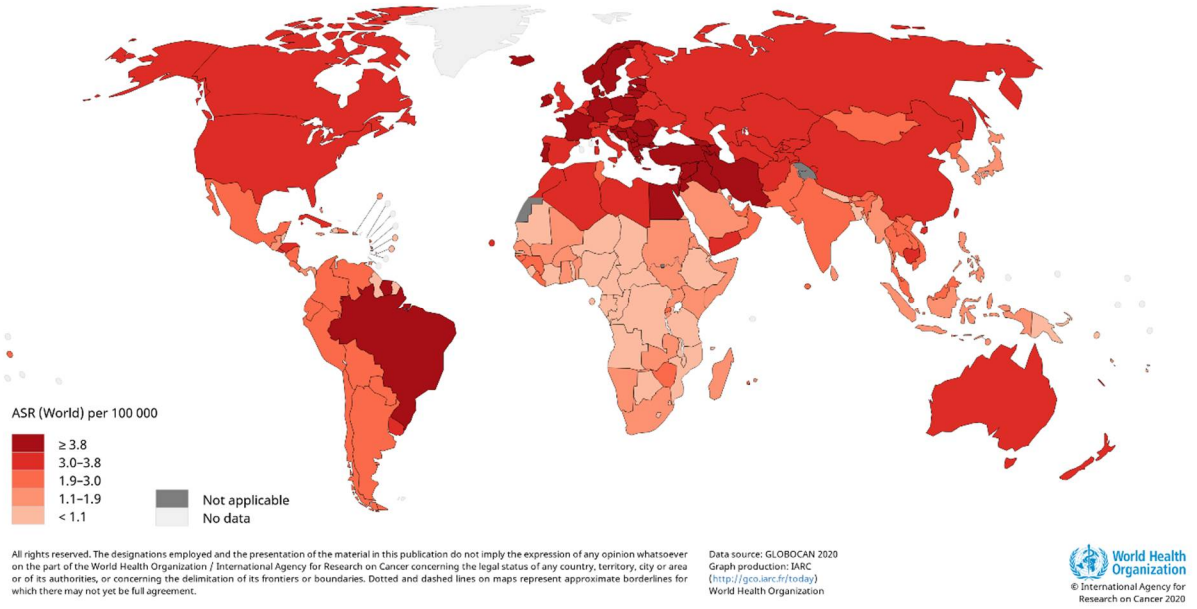


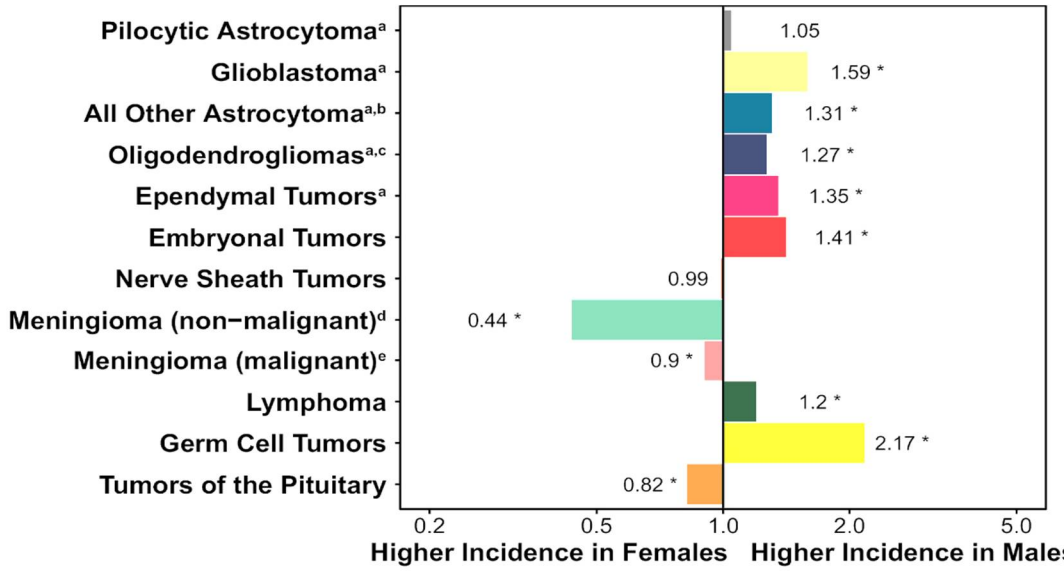
Figure 3: Estimated age-standardized mortality rates (World) in 2020, brain, central nervous system, both sexes, all ages [6].

80% of malignant brain tumors are gliomas, however some gliomas are benign. Glioblastoma, which accounts for 14.5% of all CNS tumors and 48.6% of malignant CNS tumors, is the most frequent primary malignant CNS tumor in the United States [12]. Gliomas have varying incidence rates depending on tumor grade, patient's gender, age, ethnicity, and country. The rates tend to be higher in older patients, male gender, white, and non-Hispanic ethnicity [12]. Glioblastoma multiforme (GBM), the most common type of gliomas in adults, has an incidence rate ranging from 0.6 to 3.7 per 100,000 persons a year adjusting to age and reporting geographic region [13].

People between 75 and 84 years of age are vulnerable to the highest occurrence of glioblastoma and anaplastic astrocytomas, while oligodendrogliomas are most frequently found between 35 and 44 years of age [14].

Overall, in contrast to benign tumors such as meningiomas, which are more noticeable in women, gliomas are more prevalent in males. And compared to other races, primary brain tumors are more common in whites [12].

The incidence of low-grade glioma does not differ greatly between males and females, while, regardless of age and region, while the high-grade tumors are seen more commonly in males, (figure 4) with a female to male ratio of 1:1.6 in glioblastoma [12].



** Incidence Rate is significantly different between groups at the p<0.05 level.*
a. All or some of this histology are included in the CBTRUS definition of gliomas, including ICD−O−3 histology codes 9380−9384 and 9391−9460, (Table 2).
b. ICD−O−3 Histology Codes: 9381, 9384, 9424, 9400, 9401, 9410, 9411, and 9420, all ICD−O−3 behavior codes.
c. ICD−O−3 Histology and Behavior Codes: 9450/3, 9451/3, and 9460/3.
d. ICD−O−3 Histology and Behavior Codes: 9530/0, 9530/1, 9531/0, 9532/0, 9533/0, 9534/0, 9537/0, 9538/1, and 9539/1.
e. ICD−O−3 Histology and Behavior Codes: 9530/3, 9538/3, and 9539/3.

Figure 4: Incidence Rate Ratios by Sex for Selected Primary Brain and Other CNS Tumor Histologies, CBTRUS Statistical Report: US Cancer Statistics - NPCR and SEER, 2013–2017. [2]

Anatomic Distribution

Regarding glioma's general localization in the brain, 23.6 percent are usually found in the frontal lobe, 17.4 percent in the temporal lobe, 10.6 percent in the parietal lobe, and 2.8 percent in the occipital lobe, the vast majority of gliomas occur in the cerebrum. It can occasionally be located in the brain stem, cerebellum and spinal cord as well [14].

Prognosis

Up until today, GBM universally has a very poor prognosis, with an average survival time of around 12-18 months and only 5% of patients surpass 5-year survival following diagnosis [13]. Lower grades of malignancy can typically degenerate into higher grades and eventually become secondary glioblastoma (grade IV), however approximately 90% of glioblastomas occur de novo as primary glioblastoma [15].

Risk Factors

Few risk factors have been linked to gliomas. There has been suggested evidence from epidemiological studies for positive associations with environmental ionizing radiation, and radiotherapy and negative associations with atopic conditions [15, 16].

The main risk factors found to date include ionizing radiation and inherited genetic anomalies, while asthma and atopic disorders have a preventive impact. Scientific literature on the correlation between non-ionizing radiation from cell phone usage and risk of glioma remains inconclusive to this date [17].

Therefore, more research across the globe is needed in the future, to better assess and understand the external risk factors and their contribution to the epidemiological pattern.

Genetic Risk Factors

Gliomas are sporadic for the most part, nonetheless, there have been multiple rare genetic disorders that were linked with a significant association to gliomas, these familial syndromes represent less than 5 percent of gliomas [13, 18]. To cite a few: phakomatosis syndromes (like Neurofibromatosis 1 and Tubular sclerosis), Li-Fraumeni (LFS), and enchondromatosis [19].

Advancements in genetic technology has facilitated rapid whole genome sequencing, five genome-wide association studies have detected seven genomic variants that increases glioma risk. Four of which (i.e. EGFR, TERT, TP53, RTEL1) are linked with higher risk for all types of glioma, whereas the other three (CDKN2B, CCDC26, PHLDB1) increase the risk solely for specific tumor grade or subtype.

Ionizing Radiation

There have been numerous studies looking into the environmental risk factors for developing a glioma, nonetheless little strong evidence has been found.

The main well-founded risk factor for brain tumors is exposure to high or moderate doses of ionizing radiation in medical or environmental settings. This has been verified and established in many investigations of atomic bomb survivors and radiation therapy for children with medical malignant conditions [15, 20].

Raising awareness amongst physicians, about the ALARA principle ("as low as reasonably achievable") in the utilization of CT scans pediatric population is a very important aspect to reduce radiation exposure [21].

Allergies

Allergies have been reported to protect against several forms of cancer, including glioma. [22]

It has been proposed that this effect could be attributed to improved monitoring of the innate immune system in those with allergies, but this possible mechanism has not been definitively proved. Moreover, while most findings have shown a correlation between allergies and atopic diseases with decreased risk of glioma, several studies have reported the opposite result. [22-24]

The original approach to the investigation of the relationship between allergies and glioma risk included a self-reported background review.

Analyzed data obtained as part of the INTERPHONE case-control analysis revealed a decline in glioma risk where some history of allergy has been identified. A meta-analysis of 12 studies conducted between 1990 and 2009 involving 61,090 participants with 6,408 of whom had gliomas found a decrease in the risk of glioma associated with allergic conditions [25].

Non-ionizing Radiation: Cellular Phones

There has been a great number of studies that addressed the rise of brain tumors incidence with the accelerated cell phone usage in the recent times with worries about the effects that radiofrequency and electromagnetic fields may have.

The International Agency for Research on Cancer (IARC) thoroughly evaluated the findings of a review of studies published up until 2011, and classified radiofrequency fields as a possible (IARC group 2B) carcinogen ("possibly carcinogenic to humans"). However, the evidence remains very limited to conclude a possible risk increase of brain tumors from using cellular phones [26].

INTERPHONE, an international study with 13 participating countries, the largest case-control study done to date, looked at whether mobile phone use increased the risk of tumor genesis over 10 or more years. The study found no link between risk of glioma and cell phone use, there were suggestions of an increased risk at the highest exposure levels, for ipsilateral use (tumor on same side of head as preferred phone location) and in the temporal lobe for heavy cell phone users. Similar conclusions were drawn in the CERENAT study [27, 28].

Molecular and Genetic Profile

Several factors affect survival rates such as tumor grade, age, and the presence of specific molecular and genetic biomarkers. In high-grade glioma, molecular markers such as 1p/19q co-deletion, IDH genetic mutations, and MGMT methylation are indicative of a favorable prognosis and can guide the treatment decision making process.

IDH Mutation

IDH1 and IDH2 genes are the most commonly mutated genes in gliomas, and IDH1 are found mutations in approximately 80% of gliomas grade II and III [22], and 12% of glioblastomas are associated with IDH mutations. Isocitrate dehydrogenase (IDH) is an enzyme that exists in the body through three

isoforms IDH1, IDH2, IDH3. The mutations of IDH in gliomas, can be the reason behind deformed enzymes [29]. They are found mostly on the codon 132 and 172 for IDH1 and IDH2 mutations respectively. these mutations are of valuable as an indicator of better prognosis for an overall survival time [30, 31]. They are easier to spot in the youth compared with the non-mutated wild type that are widespread amongst the elderly [29].

1P/19q Co-deletions

1p/19q co-deletion stands for the combined loss of genetic material from the short arm chromosome 1 (1p) and the long arm of chromosome 19 (19q). This chromosomal co-deletion has been found to be a positive prognostic marker [29]. and a key indicator of longer overall survival (OS) and progression free survival (PFS) regardless of the treatment protocol received [4], [5]. The 2016 WHO classification of central nervous system tumors highlighted the essential role of 1p19q co-deletion as a pathognomonic biomarker characteristic of oligodendrogliomas [13]. Virtually, all 1p/19q co-deleted oligodendrogliomas have accompanying mutation in isocitrate dehydrogenase, IDH1 or IDH2 [32].

MGMT Methylation

In order to neutralize the impact of alkylating chemotherapy, we use a DNA repair enzyme, like temozolomide, delivered to attack cancer cell's DNA. That type of enzymes is encoded by the MGMT (O-6-methylguanine-DNA methyltransferase) gene.

The expression of this gene activates the DNA repair process, thus causing ineffective chemotherapy. To avoid this situation and stop this gene expression we use DNA methylation. The methylation of this gene promoter restrains the

transcription, thus decreasing its levels, and turns off the repair mechanism. It is present in approximately 80% of low-grade gliomas and in 35% to 45% of high-grade ones [29, 33, 34].

It is a biomarker indicator of gliomas' positive prognosis and response to chemotherapy [35]. However, its effect on overall survival in low-grade gliomas is not very evident, presumably due to the coexistence of other positive prognostic biomarkers [36].

ATRX mutation

Alpha Thalassemia/Mental Retardation Syndrome X-Linked (ATRX) mutations have been associated with in telomere maintenance dysregulation and are considered important in classifying gliomas and their prognosis.

Mutations in the ATRX gene, situated on chromosome Xq21.1, cause a dysfunction in protein translation, genetic instability, and telomere dysregulation [29]. These mutations are found in around 75% of grade II and III astrocytomas and in secondary Glioblastomas, but is less common in primary Glioblastomas and oligodendrogliomas [37, 38].

ATRX mutations commonly accompany TP53 and IDH1 mutations and are almost mutually exclusive to 1p/19q co-deletion; tumors with ATRX and IDH mutations are seen usually in young adults with a notably longer overall survival [29].

TERT mutations

We often associate TERT promoter mutations in IDH-mutated gliomas with a good prognosis, however it has been proved to have a relatively poor one in wild-type IDH or for Glioblastoma with unmethylated MGMT [39]. In

primary Glioblastoma, they are very widespread with a rate of 74% to 78% in oligodendrogliomas, 70% to 84% in Glioblastomas, 25% to 50 in oligoastrocytomas and 10% to 25% in astrocytomas [37].

TP53 Mutations

Placed on chromosome 17p13, TP53 encodes for the tumor suppressor protein p53. This gene is commonly deregulated in cancer. Its deregulations are barely seen in oligodendrogliomas and tumors of grade I in WHO's grading system of tumors. Although, its prevalence keeps on increasing and is very common in astrocytomas (WHO grade II / III), with a rate of 94% [39]. Regarding the glioma prognosis of TP53 it is still an ongoing subject of research [38].

1.2 Classification

Historical Perspective

There have been many attempts to classify gliomas through the last two centuries. Following Zulch and colleagues publication the first edition of the WHO classification in 1979, another grading system, the St. Anne – Mayo grading system published by Daumas Duport et al. in 1988, became popular [40]. It relied on four histological parameters and criteria for grading astrocytic tumors; Nuclear Atypia, Mitoses, Endothelial Proliferation, and Necrosis.

Afterwards, the WHO released numerous revisions and updates, with their second edition published in 1993, third edition in 2000, and the fourth one in 2007. They mainly depended on histological features to inform their classification of tumors according to their malignancy and aggressiveness grade (from I to IV) and into different types primarily astrocytoma, oligodendroglioma, and mixed oligoastrocytoma (figure 5) [41].

It was not until 2016, that the WHO classification introduced the new concept of layered classification combining histology and molecular genetics. Where the World Health Organization updated the classification in accordance with the progress made in molecular biology, genetics, and the advances of epigenetic profiling, the revised guidelines integrate molecular criteria alongside histological criteria with the goal of obtaining more precise diagnosis, management and prognosis. (figure 6) [32]

It is still the classification used till this day, meanwhile, with the rapid acceleration of progress in research and technology, cIMPACT-NOW (the Consortium to Inform Molecular and Practical Approaches to CNS Tumor Taxonomy) began since 2016 to put forward evaluations and recommendations of proposed changes for the soon expected upcoming sixth edition of the WHO CNS tumor classification [42].

Grading and Subtypes

According to the 2007 World Health Organization (WHO) histopathological standards, these neoplasms are subclassified into numerous types and four cancer grades. The most common gliomas are glioblastomas, astrocytomas (WHO grades I–IV), and oligodendroglioma (WHO grades II–III).

Low-grade gliomas are characterized by longer patient survival and more treatment options compared to high-grade gliomas. Grade I gliomas are regarded as benign tumors, grade II have a relatively good prognostic profile with more than 50% with five-year survival, however high-grade gliomas have a much poorer prognosis, [43, 44] translated into a shorter overall survival with a more somber outcome. [32, 45]

Grade III gliomas have a greater propensity to diffuse and spread outside their macroscopic margins and are thus harder to be completely resected surgically. [46]

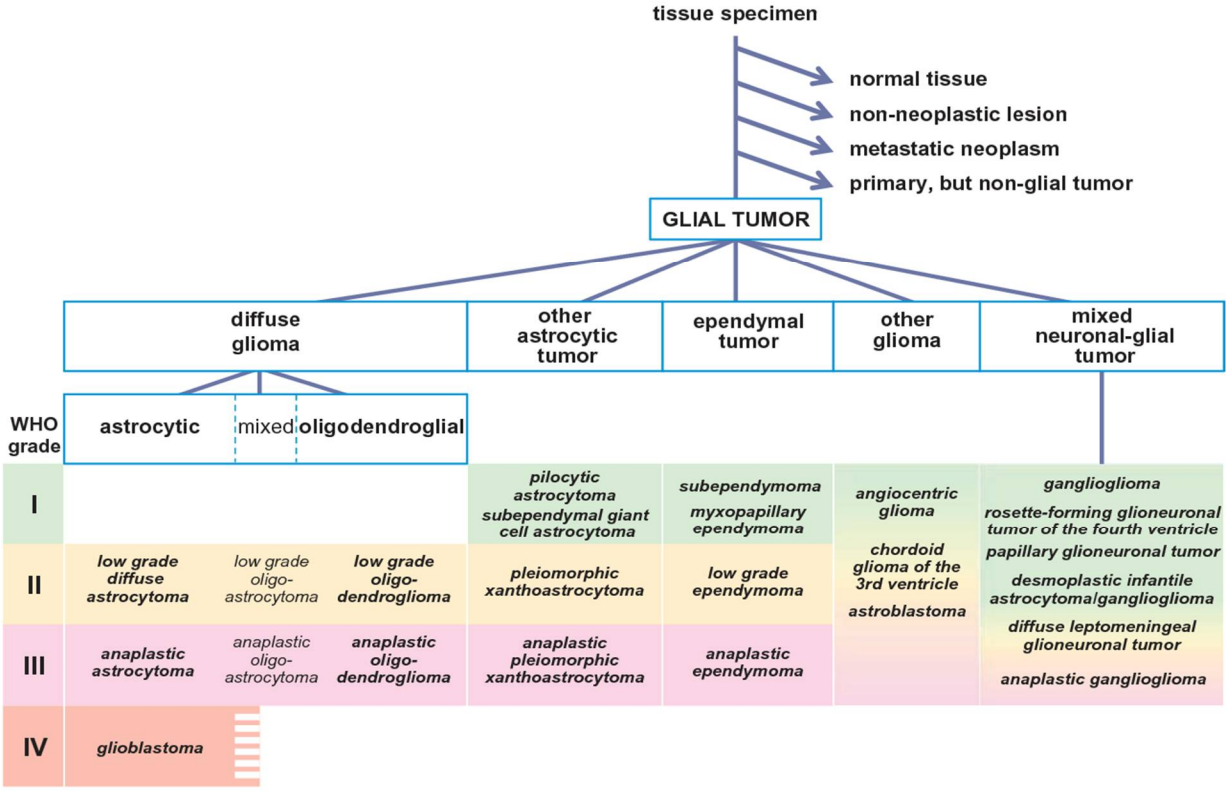


Figure 5: WHO Histological Diagnosis of Glial Tumors - Decision Tree 2007. [41]

Specific histopathological features, including cytological and nuclear atypia, anaplasia, mitosis, necrosis, and proliferation of micro vessels have been incorporated into the WHO classification. These characteristics are typical of gliomas grades III / IV. Grade II tumors are characterized by cytological atypia alone, whereas grade I tumors do not include any of the irregularities mentioned above [41].

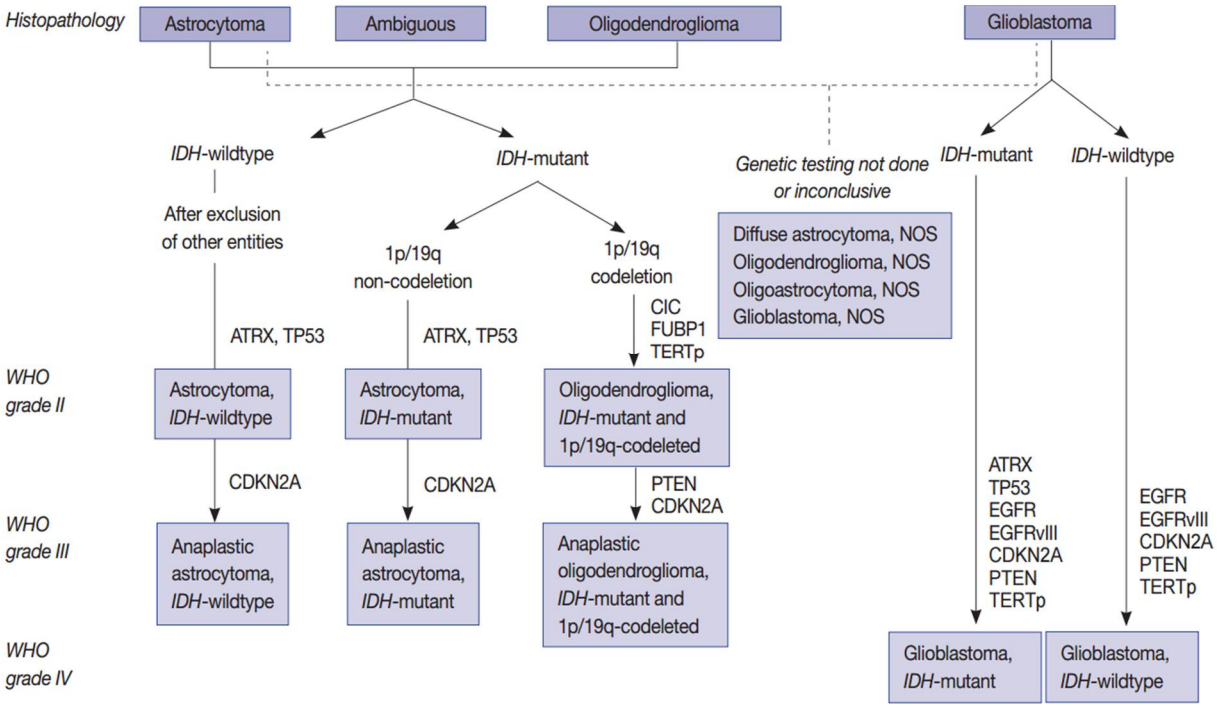


Figure 6: 2016 Classification of adult diffuse glioma according to the status of key genes. [39]

Low-Grade Glioma (LGG)

Low-grade glioma (LGG) is characterized by having more indolent course and longer-term survival in comparison with high-grade gliomas. Low-grade astrocytic tumors include diffuse astrocytoma, pilomyxoid astrocytoma, and pleomorphic xanthoastrocytoma (grade II), as well as subependymal giant cell astrocytoma and pilocytic astrocytoma (grade I tumors). Low-grade oligodendrogliomas include oligodendrogliomas and oligoastrocytomas (grade II). [41] Low-grade glioma (LGGs) is primarily a young-adult brain tumor and appear starting in the 20s to peak located in the third and fourth decades of life. [48] Despite their overall better prognosis, discrepancies still exist between subgroups.

Treatment options include surgery, radiation therapy, chemotherapy, or combined approaches. The management is individualized based on tumor location, histology, molecular profile, and patient characteristics. Moreover, in this type of brain tumor with a relatively good prognosis and prolonged survival, the potential benefits of treatment must be carefully weighed against potential treatment-related risks. [49]

Glioblastoma Multiforme (GBM)

On the other hand, Glioblastoma multiforme (GBM) is a high-grade, grade IV glioma. It is the most aggressive and common malignant primary brain tumor in adults (45.2%); and still to this day carries a very poor prognosis. [50] It is associated with a median survival of 15 months. [51] Glioblastoma multiforme are divided into two subtypes: primary de novo and secondary devolved from astrocytoma. The two groups have distinct genetic pathways, epidemiologic settings, and different outcomes. [50] Glioblastoma is essentially an older adult's tumor with a median age of diagnosis of 64 years with a peak at 75-84 years of age, and a decline following age 85. Glioblastoma multiforme is more common in men, and has high incidence in white population, followed by blacks, Asian/Pacific Islanders and American Indian/Alaska Native. [12] With regard to localization sites, glioblastoma is most often seen in the supratentorial region, and rarely, it has been reported in the cerebellar fossa and spinal cord, in which case it was described often in younger patients. [52, 53]

1.3 Radiological Diagnosis

Neuroimaging is key in the management of intracranial tumors, from diagnosis to surveillance and management follow up. The crucial role of radiology in brain tumors comes from the inaccessibility and to the stakes and

risks that invasive open neurosurgery entails on cognitive functions and quality of life. Magnetic Resonance Imaging has progressively dethroned enhanced CT-scan in the management mainstream of brain tumors, especially gliomas. [54]

MRI

Magnetic Resonance Imaging is technique that made possible the high-resolution imaging of soft tissue thanks to the properties of hydrogen nuclei found in water molecules in all our cells.

A magnetic field is rapidly applied causing the alignment of the protons in cellular water, followed by a radio frequency to disturb the alignment of these protons.

The protons properties lead them to return to their equilibrium position, and as they return to their basic position, they release energy in form of radio-frequency waves meanwhile. After the signal is detected and measured after echo time (TE). The signal is translated from frequency of the waves to intensity levels translating to grayscale voxels in the image.

The intensities of each tissue on the image are different because the longitudinal relaxation times and transverse (T1 and T2) protons from different tissues are different. By varying the time repetition rate (TR) between magnetic field disturbances and TE, different types of images can be created.

Currently, MR imaging provides the ability to evaluate features of brain tumors that indicate a particular tumor grade, as studies have correlated specific characteristics with high-grade gliomas like GBM with overall survival.

Amongst these specific features we find shape, intensity, volume, contrast enhancement texture, infiltration, diffusivity, beside evidence of core necrosis and blood volume proliferation [54].

Prolonged transverse relaxation times due to an increase of tissue water and ultrastructure translate into a tumor hyperintensity on both spin echo and fluid-attenuated inversion-recovery (FLAIR) of T2 weighted sequence. Also, foci of signal dropout may reflect areas of calcification or hemosiderin [54].

Abnormal contrast enhancement product gets accumulated in the interstitium after administrating intravenous gadolinium, due to a lower impermeability of blood-brain barrier in relation with the neovascularization and necrosis. However, relying solely on contrast enhancement to discriminate between high-grade glioma and low-grade glioma is insufficient [54].

Radiomics techniques covers many of the described features on MRI, they allow to enrich the image with information by extracting quantitative and semi-quantitative image features aimed to establish the relation between them and the clinical sought-after result [55].

According to one of the studies on patient survival's endpoint and features extracted from MRI, a longer survival is related to non-contrast enhancing tumor while edema and multifocality showed restricted prognostic signs in high-grade gliomas [54, 56].

MR Spectroscopy

Magnetic Resonance Spectroscopy (MRS) In a selected volume of brain tissue, of measuring the concentration of chemical metabolites. The decrease in the peak ratio of N-acetylaspartate (NAA) and NAA / Creatine (Cr) observed in

brain tumors indicates a decrease in neuronal viability and amount. An increase in peak Choline (Cho) or an increase in the ratios of Cho / NAA or Cho / Cr is associated with an increase in turnover cell membranes due to prolife and higher cell density. Recently, multiple studies are exploring the applications of MRS in Glioma and are targeting molecular biomarker such as IDH using MR Spectroscopy with fairly good prediction outcomes [57].

In another perspective, radiology imaging and MRI more specifically enables grading and organizing tumors into subgroups. In fact, subgroup distinction is very essential owing to the extremely heterogeneity of glioma tumors. This high variability requires therefore, an accurate classification system in order to draw predictions of tumor nature, tumor behavior, and therefore tailor adequate treatment. We can cite the Brain-Grid classification system that uses a standardized anatomical grid system to compare morphological MRI with segmented white matter atlas [58]. The systems allow to predict and inspect interesting biological information on tumor dynamic such as its extension, speed, and preferential direction of progression. Thanks to this system we could study prospective and retrospective cohorts of patients in a very easy way and achieve a better grasp of glioma behavior [58].

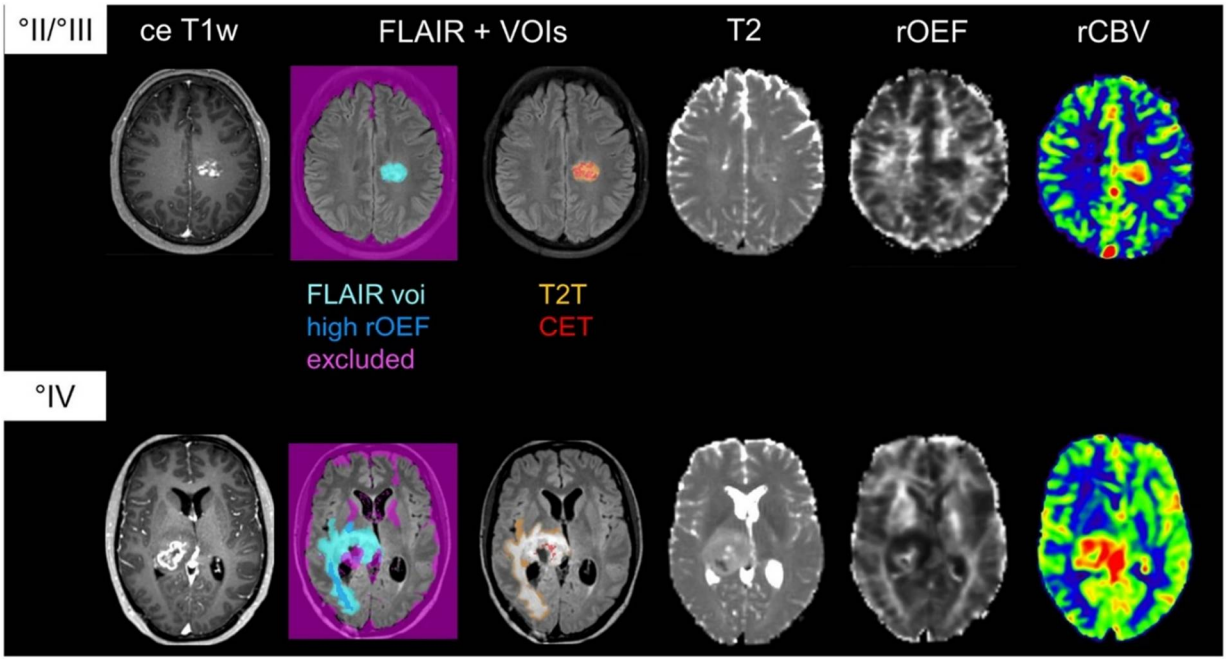


Figure 7: Multiparametric MRI-based differentiation of WHO grade II/III gliomas and WHO grade IV. [D]

1.4 Machine Learning and Deep Learning

Machine Learning is an essential part of Artificial Intelligence (figure 8), consisting of statistical models and algorithms that learns from data to reach a correct answer. It is divided into two branches: Supervised learning, where the data getting fed to the algorithms are already labeled with correct answers. The second branch is unsupervised learning where the algorithm seeks to find similarities between the data and group them into labelled subgroups.

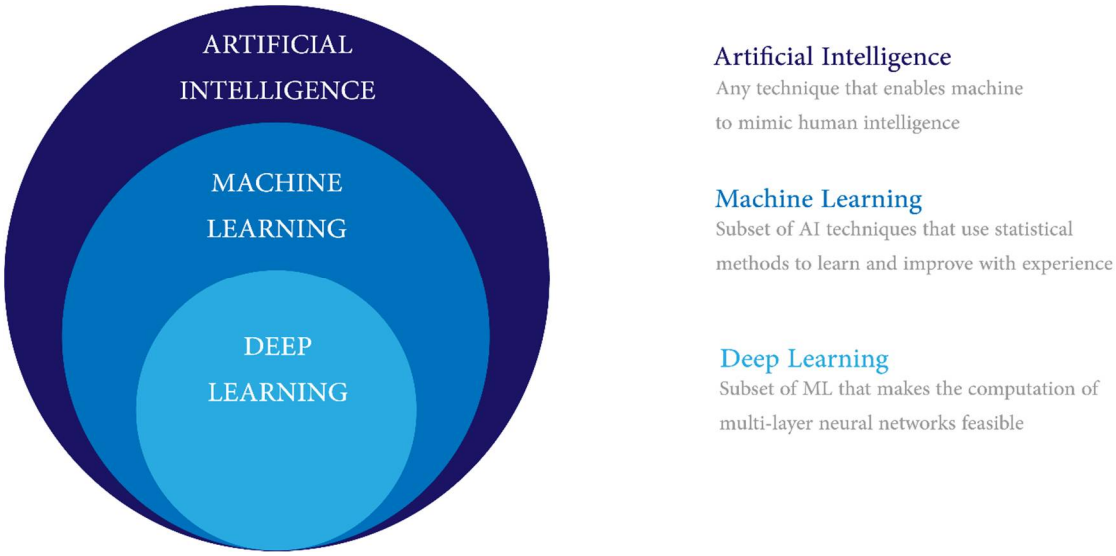


Figure 8: Subsets of Artificial Intelligence

Deep learning (DL) is a novel subset of Machine Learning techniques (figure 8) that revolutionized image interpretation. DL works on extracting immense amounts of detailed characteristics and features of objects and images, and recognizes the label based on the specific pattern of said image or object [59, 60]. Undoubtedly this underlines a great advantage due to the ability of analyzing in a framed period of time a large amount of information, and in a detailed manner. Taking in consideration CNNs technology, applied in such a field, deep convolutional neural networks (DCNNs) is an outstanding tool in object recognition in medical imaging with prior training and implementation of the program of large amounts of data [61]. Such realization is due to the ability to superimpose layers of data guided by hierarchical feature representations.

Thus, human participation has been downplayed in the purpose of brain tumor image studying, though implication in clinical decision making is somehow limited due to limited private medical data. Some studies have used

DCNN and machine learning for classification of gliomas and achieved good accuracy [62], other studies used machine learning to classify tumors based on the features extracted after image segmentation [63]. Convolutional neural network (CNN) techniques of segmentation have re-emerged as a powerful tool for image classification and object detection in the medical fields as in every other field, since its introduction in the 80', inspired by Fukushima and LeCun (figure 9). Thanks to the recent advances achieved in GPU computing, and the rising training datasets have led CNN to become a new standard of segmentation exploration in many medical fields, in particular brain tumor exploration. Furthermore, deep learning rising success can be linked to CNNs (21,22)

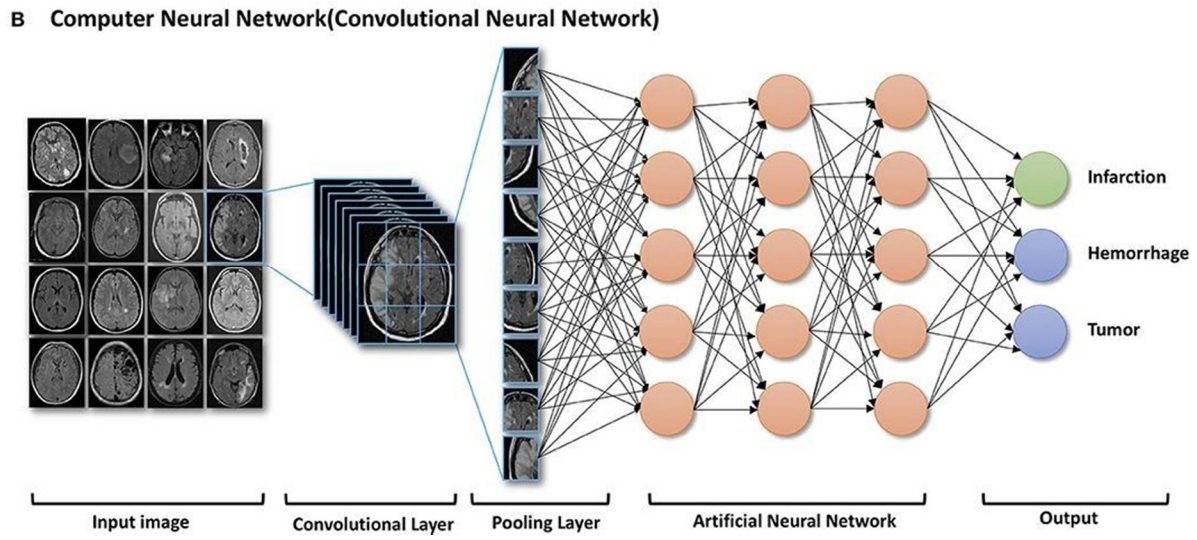


Figure 9: Convolutional neural network architecture. [H]

1.5 Radiomics

Radiomics, a term coined by LAMBIN et al. [2012] [64], is recently emerging field that studies quantitively and in a non-invasive manner the characteristics tumors [65].Radiomics allows to explore the tumors by extracting a large number of quantitative features from medical images. The hypothesis is that these clues can synthesize relevant information that is not easily located by radiologists and clinicians. The medical images used can be of anatomical, functional or molecular origin. This discipline is defined in particular in oncology. All sources of imagery information as well as information clinics can be combined to better characterize a tumor. The aim is to put in sets up personalized medicine aimed at offering each patient a specific treatment and thus improving their response to therapy and prolonging their survival time.

In oncology, the major interests of radiomic analyzes are linked to:

- The non-invasiveness nature of imaging.
- The representation of the tumor as a whole, in contrast to local biopsies which represent only a portion of the tumor.
- The possibility of analyzing the tumor several times and thus monitoring and predicting its progression or recovery over time.
- The relatively low cost of applying radiomics which could be helpful in low resources settings, and in regions with limited access to high end medical equipment such as gene sequencing.

Given the attractive and ambitious promises of radiomics, the number of publications using radiomics increased exponentially since its emergence in 2012.

Radiomics research shows that these methods can be applied on different types of tumors and in different parts of the body. Hence, tumors can be quantified, analyzed constructing characteristic signatures representing imaging biomarkers that help in clinical decision making [66-71].

Clinical images contain diverse types of information that can be extracted in multiple forms. In radiology lexicon, qualitative semantic characteristics are widely used to identify lesions [72], while quantitative characteristics consists of different analyzable descriptors describing the lesion in various degrees of complexity. They can describe in a lower degree: the morphology of the lesion with the histogram of its image voxels' intensity, and in a higher one: the texture of the lesion and the way that voxels' values are disposed in the space.

These quantitative characteristics are generally extracted by the means of computer algorithms [73]. They can either be derived solely from the clinical image or with the application of a number of masks and transformation (e.g., wavelet transform).

In more detail, there are different categories for these features, for a certain volume of interest (COI) we note some of them such as the shape features that cover the shape description, and the geometrical properties like the volume, the number of voxels the maximum diameter in an orthogonal base, its surface area, the tumor compactness and sphericity. For instance, as an application we can use the surface-to-volume ratio to distinguish between a round and a speculated tumor even if they have the same volume.

Regarding the characteristics of first degree, they take into consideration voxel values individually without covering their relation in space. They are represented with statistical properties of all voxels' intensities of the image like the mean, median, range, skewness (asymmetry), kurtosis (flatness), uniformity, randomness (entropy).

In what concerns second degree features or also known as textural features [74, 75], they are built with studying interrelation between voxels and the statistics that lays behind it [76] that can be translated to the measure of heterogeneity in the tumor.

A way of extracting these second order features is using the grey-level co-occurrence matrix (GLCM) that quantifies the occurrences of neighbored pair of voxels along a certain axis and at a certain distance. Another way is by using the Grey-level run-length matrix (GLRLM), that quantifies the occurrences of neighbored voxels with the same intensity along a certain axis [77].

Furthermore, higher-order features are the result of applying mathematical filter transformations and statistical approach, like fractal analysis, wavelet transform, and Laplacian transforms of Gaussian-filtered images, to reduce or suppress noise, detect recurrent pattern or certain structure. The predication of patients' reaction to chemotherapy when they suffer from glioblastoma or their survival has been a subject of several studies who adapt a large number of radiomic features derived from MRIs. Other studies used features obtained from functional imaging to predict prognosis [66-68, 71, 78, 79].

This new field continues to evolve and is being associated with other areas like genomics, creating a new field named radiogenomics, in order to benefit from more diverse and abundant high dimensional information on genes and imaging while treating and studying tumors [80, 81]. This gives place to new miscellaneous fields of multi-omics with a promising future of novel discoveries and breakthroughs.

Machine learning algorithms begin to take an important role in precision medicine during this modern age. Their application in radiomics are revolutionizing patient care in a considerable way and giving place for efficient and accurate computer-assisted diagnosis.

Radiomics could help enhance and complement tumor biopsies through ‘virtual biopsies’ at multiple sites and time points, especially in heterogeneous tumors and quickly degenerating tumors. Also, it can potentially assist in the detection of clinically relevant biomarkers, through correlating data from the anatomic to the molecular scale (figure 10).

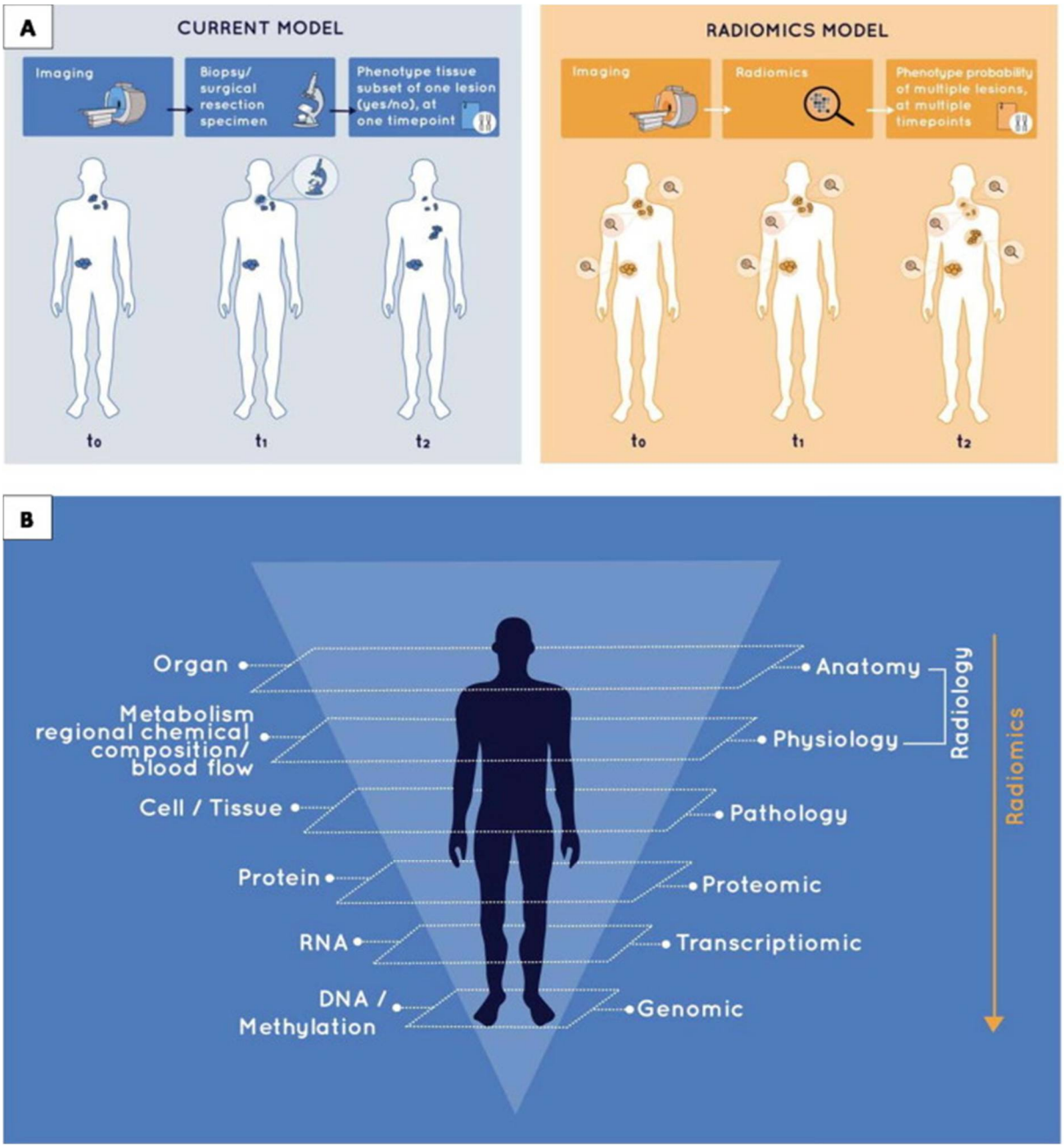


Figure 10: Radiomics as an efficient identification and qualification of biomarkers.

1.6 Objective

The classification of gliomas depends mainly on the grading and the molecular profile of the tumor. The objective of this thesis is to explore the potential of radiomics combined with AI approaches in glioma grading and molecular profiling, using state of the art machine learning algorithms. First, our aim is to compare radiomic feature-based with image-based classification methods of glioma to distinguish between MR images of low-grade gliomas and high-grade glioblastoma using machine learning models and deep convolutional neural networks. Second, to assess the radiomics capacity of predicting tumor molecular profile, through extracting radiomics features from low-grade glioma images and applying a machine learning model to predict IDH mutation status of the tumors.

The ultimate purpose is to reach a potential future these tools where can help realize a more accurate diagnosis, complementing and reinforcing histological confirmation, especially in heterogeneous tumors, as well as to discover novel biomarkers conducive to establishing a new era of precision and personalized medicine.

The above aspects were assessed as follows:

- Chapter 1: Radiomics' ability to help predict tumor grade, differentiating Glioblastoma (Grade IV) from low grade glioma (Grade II-III).
- Chapter 2: Radiomics' ability to help with molecular biomarker classification, predicting IDH mutation status in low grade gliomas.

Chapter I: LGG / GBM Differentiation

2 Chapter I: LGG / GBM Differentiation

2.1 Materials and methods

2.1.1 Dataset

We use in our research pre-operative scans with their automatically generated and manually corrected segmentation labels from The Cancer Genome Atlas (TCGA),

The Cancer Genome Atlas (TCGA), joint effort between the National Cancer Institute (part of the National Institutes of Health (NIH) of the USA) and the National Human Genome Research Institute, a landmark cancer genomics program that started in 2006, and characterized over 20,000 primary cancer and matched normal samples spanning 33 cancer types. It is considered the world's largest catalogue of genomic information related to cancer. During the last decade, TCGA generated over 2.5 petabytes of genomic, epigenomic, transcriptomic, and proteomic data. The data, which has already led to improvements in our ability to diagnose, treat, and prevent cancer, will remain publicly available for anyone in the research community to use [--].

TCGA data can be accessed through the Genomic Data Commons (GDC) Data Portal (figure 11). In our case study, we are particularly interested in the TCGA public datasets of Glioblastoma Multiforme (GBM) and Low-Grade Glioma (LGG) gliomas that we will refer to later with TCGA-GBM and TCGA-LGG respectively. Both of these datasets can be found on The Cancer Imaging Archive's (TCIA) website, an open-source database where abundant number of medical images for research can be found. (28 - 29)

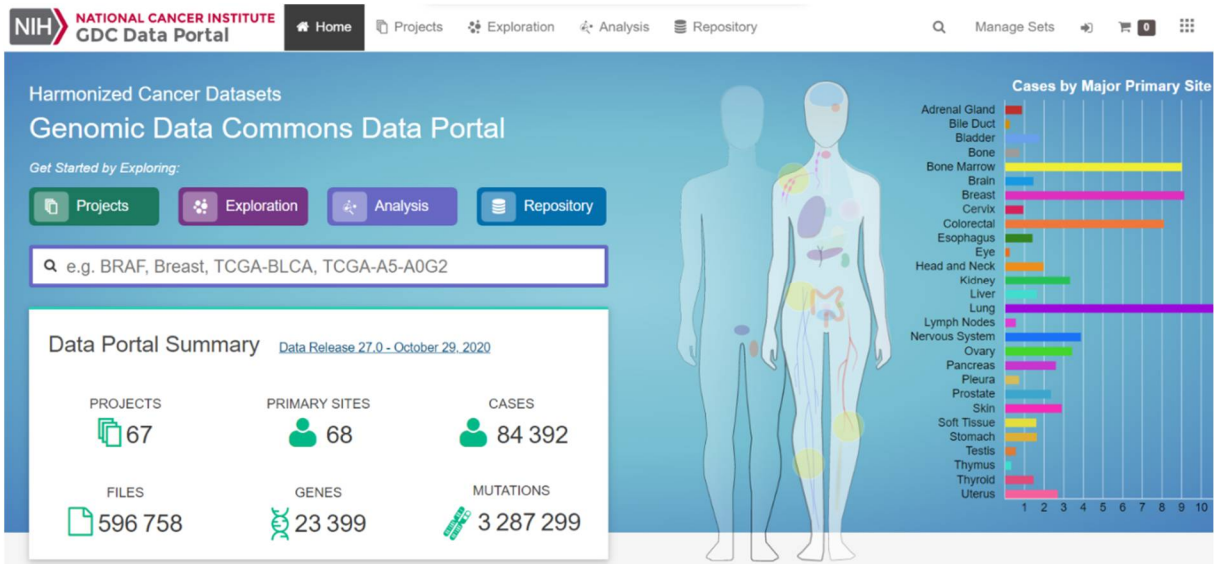


Figure 11: Genome Data Commons (GDC) Data Portal webpage: https://portal.gdc.cancer.gov/

Each of the datasets contains preprocessed, skull-stripped and co-registered preoperative multimodal (i.e. T1, T1-Gd, T2, T2-FLAIR) MRI scans for the corresponding type of tumor in NifTI format. The segmentation labeling was made using the winning algorithm of the International multimodal Brain Tumor Segmentation challenge (BraTS 2015) that uses a generative and discriminative approach, and then validated and corrected by an expert board-certified neuroradiologist (figure 12).

Radiomic features were then extracted from the final scans and labels and were made available on TCIA's website along with them. Some of the derived features are textural, morphological, volumetric and histogram-based parameters, intensity, spatial and diffusion characteristics drawn from models of glioma growth. In addition, in some of the computational challenges they could be considered as an evaluation database for the proposed solutions.

Regarding our database composition, it includes 65 patients in TCGA-LGG and 102 patients in TCGA-GBM and an overall number of patients of 167. As for the scans, for each patient a folder of 4 types of MRI scans and a segmented MRI scan by Gistr Boost are available as well as a manually revised segmentation for some. The features database on the other hand contains LGG and GBM patients' IDs, the scan's date and 724 radiomic features.

In both datasets of images and features, GBM tumor is considered as the positive class and LGG as the negative one.

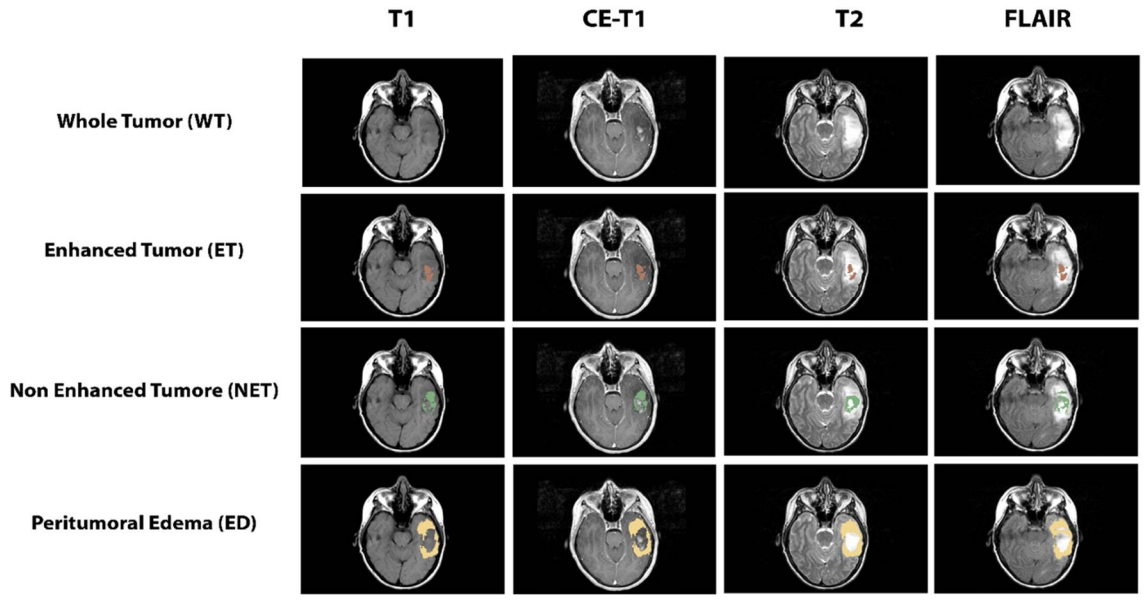


Figure 12: Visualization of an example case of segmentation images used of one of the included tumors.

2.1.2 Preprocessing

Before being able to feed data to our machine learning algorithms, it should be pre-processed, treated and made adapted to the case of usage. Also, we need to divide each dataset to two sets: training and test. It is worth noting that a third set for validation isn't mandatory in our case, since only model comparison was carried out.

The split of both databases was made in a way that the training set includes 80% of the original points, and the test set 20%.

2.1.2.1 Features preprocessing

For the features dataset, the first step of preprocessing is to clean our database. This includes removing the features that are uninteresting for training our model. This includes columns with very few data, very low variance or irrelevant to the target we want to predict. As shown on the graph (figure 13), two clusters of features have a percentage of Not a Number (NaN) higher than 80%, therefore they bring too little information to the model. The number of columns omitted during this phase is 21, leaving us with 703 features. The next step would be to replace the missing data of some entries (corresponding to the columns with 20% of non-available data) with an adapted value. In most of the times, the missing values are replaced by either 0, the mean of the column or its median. We chose the median of the columns as a substituting value, after comparing the performance of the models according to the chosen value. In fact, the median is the one that is least affected by outliers, and therefore the safest choice to make.

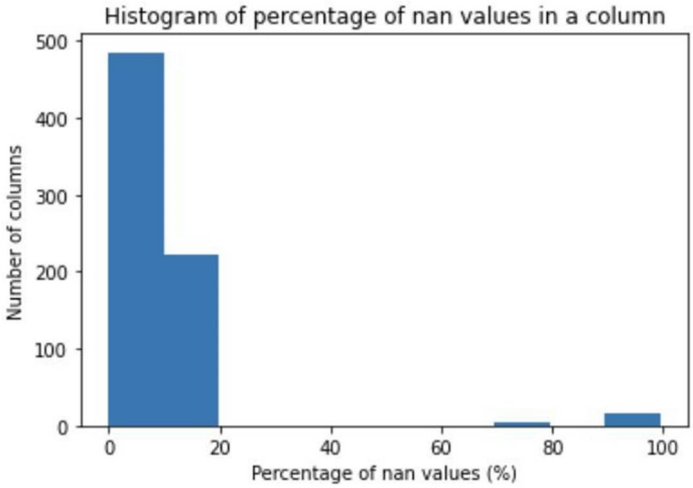


Figure 13: Histogram of the percentages of nan values in image dataset's columns

After cleaning our data, we proceed to normalizing it. Since the features' ranges of value can be very distinct, this step of preprocessing is mandatory. As features' importance can be influenced by the range of the variable, all variables' ranges should be rescaled to [0,1]. We chose her the range [0,1] over others, especially [-1,1], to keep the variable positive like in the original dataset. The scaling used is the Min Max Scaling is carried out according to the following equation:

$$\text{Xi} = \frac{Xi - min(X)}{max(X) - min(X)}$$

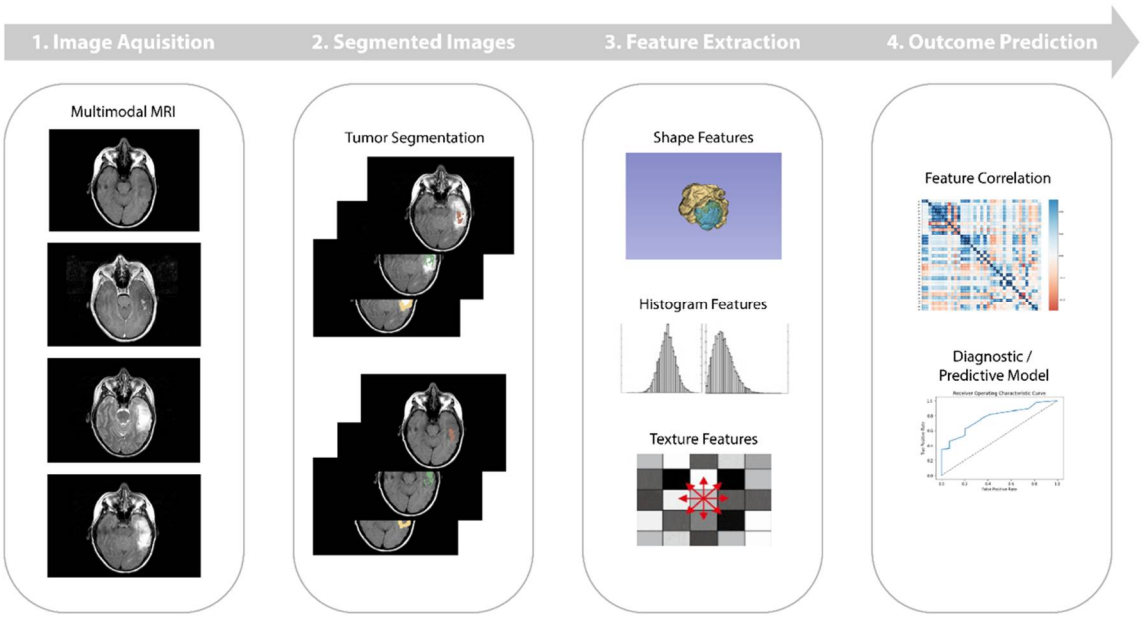


Figure 14: Schematic illustration of pipeline used for radiomics analysis.

2.1.2.2 Image preparation

In order to avoid data leakage between training and testing dataset, we carried out the data split using the patients' IDs instead of using directly the data points.

Moreover, since the MRI images are all (240 x 240) 2D black and white images, we selected 3 slices of the 3D MRIs from different brain areas in order to maximize the tumor coverage, and then concatenated the three to obtain an image of shape (240, 240, 3). The latter can be seen as the representation of a colored image in RGB.

2.1.3 Machine learning algorithms

2.1.3.1 Feature based classification

Given the limited data points available in our dataset of features, we considered trying to use classification methods adapted to this case. The classification algorithms used are:

Decision Tree, Random Forest and XGboost. In the following sections we'll be explaining briefly each of the mentioned algorithms, as well as presenting the parameters and the approach used for each.

Decision Tree

A decision tree is a type of supervised model that can be represented with a tree like graph. Each node of the tree consists of a binary evaluation of one of the features, depending on its value we select the corresponding branch, in which we continue on evaluating other features until reaching a leaf with no further descendants. Once we reach the end of the tree, the class of the input is determined.

A good decision tree is determined by the features used in its construction, in other words by the evaluation's criteria. Its performance varies significantly when changing the input, that's why we risk to overfit easily when using this kind of estimators.

Random Forest

RF is another classification model based on ensemble learning. An ensemble model or algorithm uses a combination of multiple individual machine learning models, in order to achieve better predictive performances. Random forests use the decision trees, as base models, that are chosen using the bagging

method. The bagging method aims to lower a model variance, by training a group of models, each on a subset of the training database, and then aggregating at the end their predictions to give one final label: the prediction of the ensemble model.

The feature selection on the level of each tree, is based on either the entropy or the Gini impurity.

XGBoost

XGBoost is an Extreme Gradient Boost Machine. As its name suggests, it is based on the boosting mechanism. Contrary to the bagging method, in boosting the trees are built subsequently and not in parallel, where each tree learns from the errors of its predecessor.

In this algorithm, the problem of finding the best split is still present and is essential to obtain the best performances (30). However, XGBoost use of multiple regularization algorithms to avoid overfitting and handle sparsity (31), makes it a very robust model.

It is worth noting that considering the few samples in our dataset, a deep neural network approach cannot be conceivable.

2.1.3.2 Image based classification

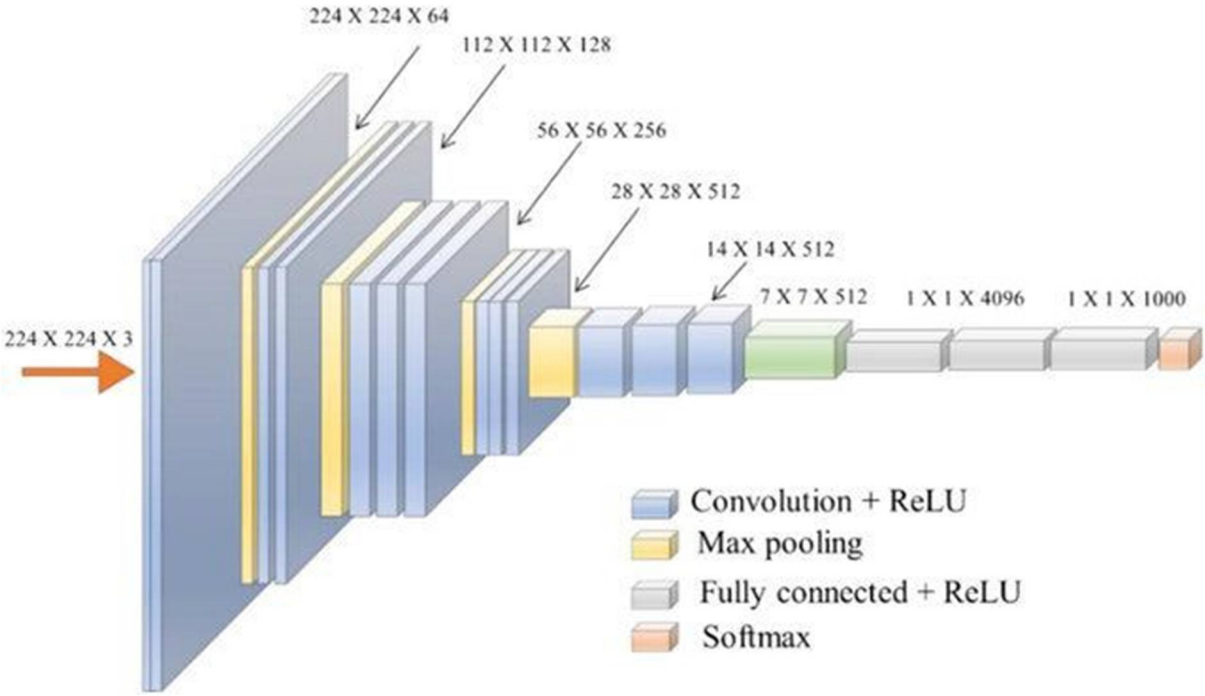


Figure 15: Architecture of a VGG-16 network. [I]

VGG-16, also called OxfordNet, (named after the Visual Geometry Group at University of Oxford) is a 16-layer Convolutional Neural Network model that was first introduced by K. Simonyan and A. Zisserman (32) (figure 15).

The network's architecture consists of multiple convolutional and max-pooling layers that use ReLU activation function, starting with an initial input convolutional layer of a 3-channel image with dimensions 224×224. It includes multiple 3×3 filters in a serialized fashion. The idea behind it is that the multiple non-linear layers enhances the depth of ConvNet. Thus, with an increase in depth, the ability to learn hidden features increases at a lower cost.

In order to adapt the VGG architecture to our classification task, we built a new model shell with some additional layers. The architecture of the used model is a concatenation of pre-trained VGG model called VGG-16, available in Keras package, a flatten layer that convert the output of VGG into a 1-dimensional vector, a dense layer with smaller dimension, a dropout function that randomly sets some of the input entries to 0 in the aim of avoiding overfitting, and lastly a softmax layer that returns a probability of belonging to a class (figure 16).

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vgg16 (Model)                (None, 7, 7, 512)         14714688
_________________________________________________________________
flatten (Flatten)            (None, 25088)             0
_________________________________________________________________
dense (Dense)                (None, 256)               6422784
_________________________________________________________________
dropout (Dropout)            (None, 256)               0
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 257
=================================================================
Total params: 21,137,729
Trainable params: 21,137,729
Non-trainable params: 0
_________________________________________________________________
```

Figure 16: Employed VGG-16 model architecture for image-based LGG vs GBM classification

2.1.4 Model selection

For each of the presented models, there are a number of hyperparameters that need to be optimized to reach higher performance. The method we opted for to choose the feature-based models' hyperparameters is the Grid Search. We determined a set of candidate values for the hyperparameters of interest, then by the means of the grid search algorithm of the package Sklearn, an optimal combination of values is returned.

As for the image-based classification, due to the limited computational resources, only two models were tested, with a manual grid search on the learning rate, only the best model was kept.

2.1.5 Model training

2.1.5.1 Feature based classification

Using the selected model in the previous section and its optimized hyperparameters we train the model on the training set before we test it and evaluate its performances.

2.1.5.2 Image based classification

To train the image-model we used Adaptive Moment Optimization (Adam) of the Keras package as an optimization algorithm, with a learning rate of 1e-6 and the binary cross entropy as the loss function (error function).

The model was trained on 60 epochs. The number of epochs was determined using the early stopping mechanism that consists of stopping the training if the model fails to reach better performance after 20 epochs from the best it reached from the beginning of the training. Also, in order to keep hold of the information that the pre-trained model bears, we froze the VGG layers to keep the weights unchanged.

2.1.6 Evaluation Metrics

2.1.6.1 Feature based classification

Before introducing the metrics used to evaluate the performance of our models, below are some standard annotations used in the metrics' expressions. When an input is fed to a classifier, the prediction of the model is considered

one of the following:

TP (true positive): Classified as positive (class 1) and its true class is positive.

TN (true negative): Classified as negative (class 0) and its true class is negative.

FP (false positive): Classified as positive (class 1) while its true class is negative.

FN (false negative): Classified as negative (class 0) while its true class is positive.

Considering the positive class P, one of the classifier's goals is to detect all entries that belong to the class P. The metrics indicate how many of the found were correct hits, how many of the really positive entries were found, the rate of correct classifications in all, respectively.

The basic metrics that were used are described below:

- Precision (Positive Predative Value – PPV)

The ratio of entries of true class positive that were predicted positive.

$$\text{precision (PPV)} = \frac{\text{TP}}{\text{TP+FP}}$$

- Recall (Sensitivity)

  The ratio of entries predicted positive amongst all true positives.

$$\text{recall (Sensitivity)} = \frac{\text{TP}}{\text{TP+FN}}$$

- Accuracy

The ratio of correct predictions that were mage amongst all predictions.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}$$

- F1-score

A quantitative metric of the model's performance combining both the precision and the recall.

$$\text{F1-score} = 2 * \frac{precision * recall}{\text{precision+recall}}$$

Among the other type of evaluation methods like the ROC curve (Receiver Operating Characteristic curve). The ROC curve (figure 17) is the plot of true positive rate and false positive rate while varying the prediction threshold (a predicted value above the threshold value is considered positive and vice-versa). The Area Under the Curve of ROC (ROC AUC), is the derived metric from the curve used to evaluate the models. It is different from other metrics in terms that it doesn't depend on the threshold used to determine the class, making it an intrinsic value and it also represent the discriminative power of our model.

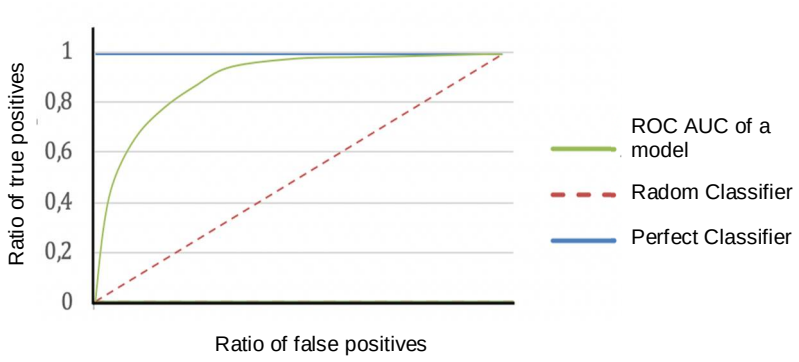


Figure 17: An example of ROC curves for random classifier, perfect classifier and an average classifier

2.1.6.2 Image based classification

To evaluate the convolutional neural network performance, we used the most common metrics: Accuracy and Loss function (error function). The two metrics are calculated on the training and validation set at the end of each epoch. During the training phase, we search to maximize the accuracy on the validation set and minimize the loss. To get a better idea of the model's evolution throughout the epochs, we plot the evolution of these metrics.

2.2 Results

2.2.1 Feature based classification

Before using machine algorithms, we tried to see if there is a direct correlation between any of the features and the class of the tumor. Therefore, we calculated the correlation matrix that consists of the correlation coefficient between each two features. The matrix is represented in the figure below (figure 18) using a heat map.

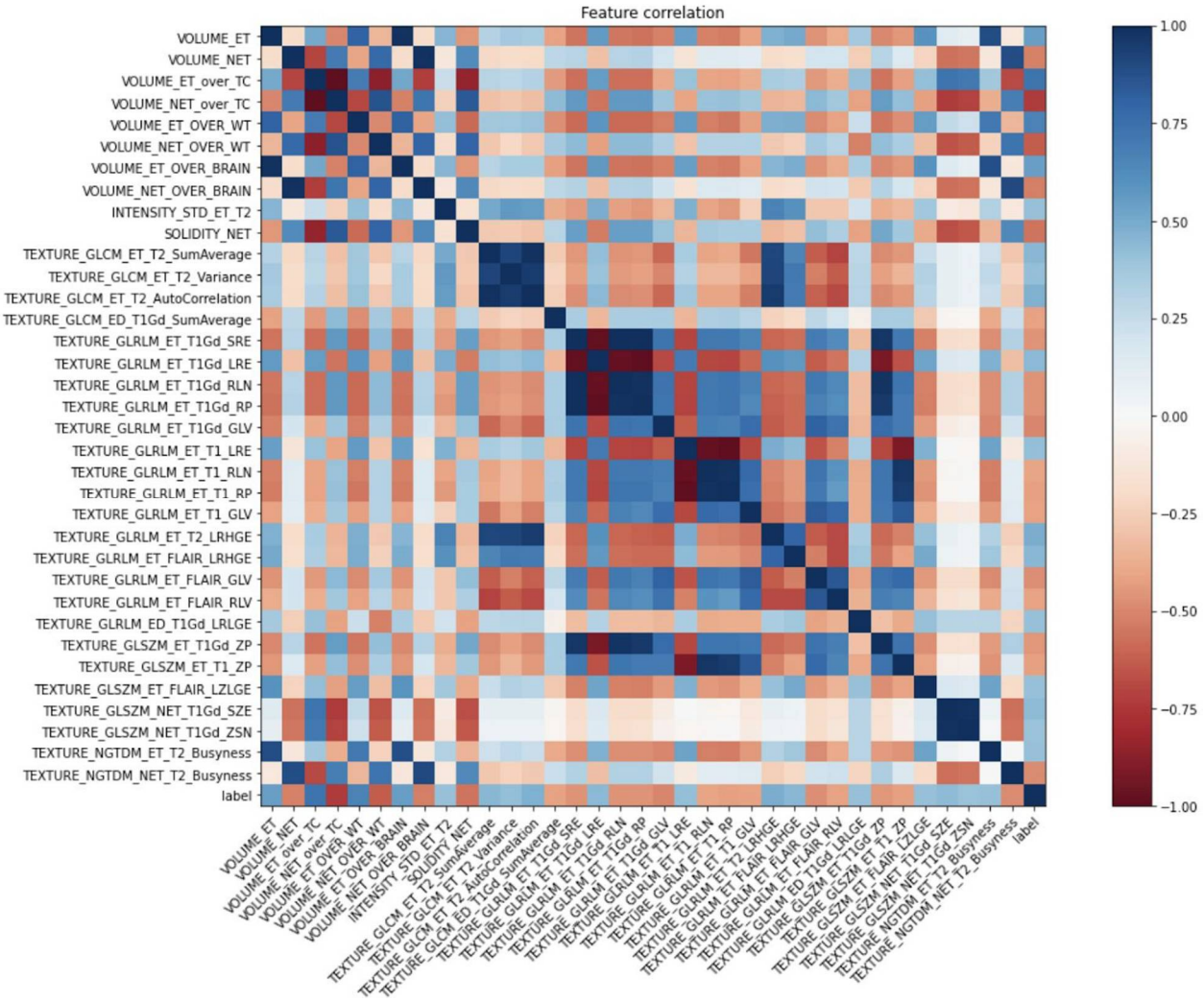


Figure 18: Heatmap of feature-feature correlation matrix for LGG vs GBM classification

We notice that there is no strong correlation between the features and the label or the class of tumor, as none of the boxes in the last column are colored with dark red or blue. Except for the volume of non-enhanced tumor (NET) over tumor core (TC) 'VOLUME_NET_OVER_TC' that is inversely correlated to the positive class of label, which is in this case the GBM tumor.

To select the best machine learning model. The result the end of the Grid search, the RF hyperparameters that were selected are:

- Criterion for feature selection: entropy
- Maximum depth of each tree: 7
- Maximum number of candidate features: sqrt (number of features)
- Number of estimators: 200

We also used it on the XGBoost classifier. The model used for final predictions has the following hyperparameters:

- Learning rate: 0.01
- Maximum depth of each tree: 6
- Number of estimators: 5
- Subsample ratio of the training prior to growing trees: 0.8

Second, after running the described above classification models on the TCGA features dataset, the following results were obtained. Below are the confusion matrices for each model that shows the number of TP, TN, FP, FN of the test (figure 19).

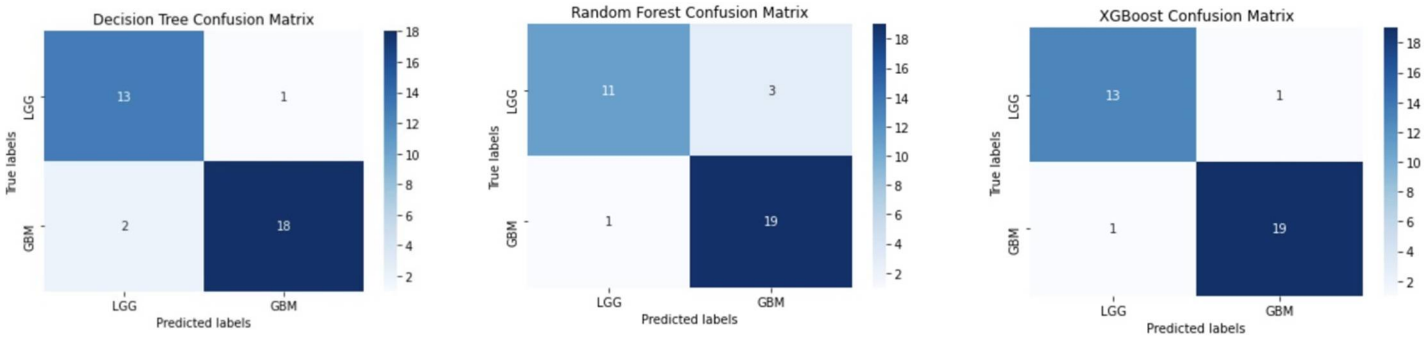


Figure 19: Heatmap of confusion matrix of the Decision Tree, Random Forest and XGBoost. (from left to right) for LGG vs GBM classification

As described on the confusion matrix heatmaps, only 1 LGG patient was classified as GBM using the decision tree, also only 3 and 1 using Random Forest and XGBoost respectively. In the same way, 2 of the GBM patients were classified as LGG for Decision Tree and only 1 for both RF and XGBoost. All the other patients were well-classified.

Using these values, we were able to calculate the different evaluation and performance metrics for each model, as shown in the figures below.

Decision Tree

	precision	recall	f1-score	accuracy	ROC-AUC
LGG	0.87	0.93	0.90	0.91	0.91
GBM	0.95	0.90	0.92	0.91	0.91

Random Forest

	precision	recall	f1-score	accuracy	ROC-AUC
LGG	0.92	0.79	0.85	0.88	0.97
GBM	0.86	0.95	0.90	0.88	0.97

XGBoost

	precision	recall	f1-score	accuracy	ROC-AUC
LGG	0.93	0.93	0.93	0.94	0.95
GBM	0.95	0.95	0.95	0.94	0.95

Figure 20: Table of metrics of the three models

In the tables above (figure 20), for each class we precise the precision, the recall, the f1-score, the accuracy and the ROC-AUC, according to the algorithm used for prediction.

Furthermore, thanks to the model implementation in the package sklearn, we can determine for each algorithm the features with higher influence on predictions, a criterion known as feature's importance. For the decision tree and XGBoost all features with non-null importance are shown, as for the random forest we kept only those with an importance indicator bigger than 0.1.

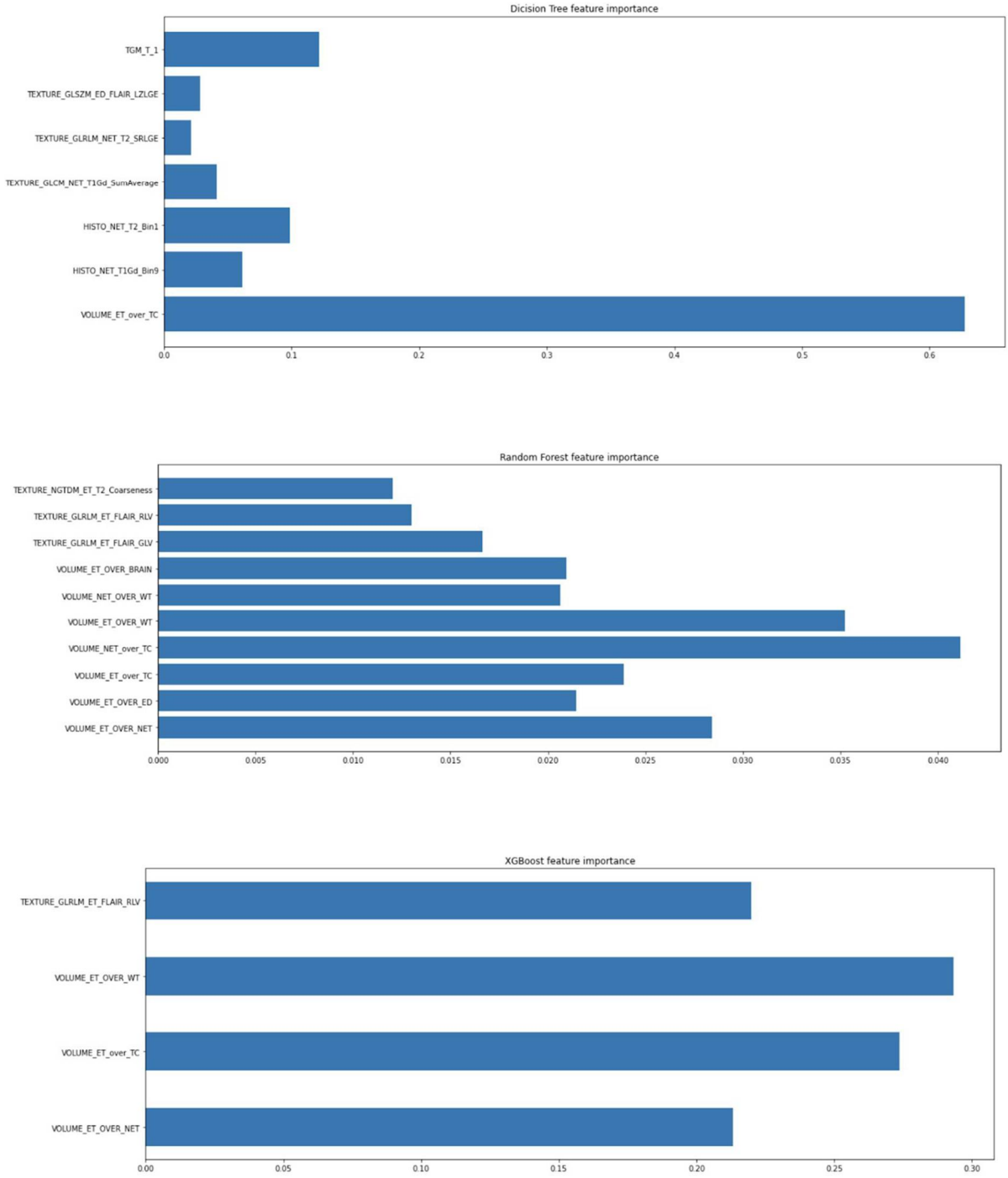


Figure 21: Feature importance graph for LGG vs GBM classification models

As shown in the figure above (figure 21), the volume of the enhanced tumor (ET) over tumor core (TC) ‘VOLUME_ET_OVER_TC’ figures as the most important feature for decision tree, the volume of non-enhanced tumor (NET) over TC ‘VOLUME_NET_OVER_TC’, then the volume of ET over the whole tumor (WT) ‘VOLUME_ET_OVER_WT’ for the random forest and finally the volume of ET over WT ‘VOLUME_ET_OVER_WT’ and the volume of ET over TC ‘VOLUME_ET_OVER_TC’ for XGBoost.

We notice that, out of the three models, the random forests accord the higher importance to the feature most correlated with the tumor class : the volume of NET over TC ‘VOLUME_NET_OVER_TC’.

It is also worth mentioning that the features of volume ratios between the segmented parts of the tumor are the one to figure on the top when measuring the features’ importance, thus we can assume that they can be very good indicators for the tumor’s type.

2.2.2 Image based classification

As illustrated on (figure 24), the value of accuracy keeps on going up with the number of epochs, as the loss keeps on going down approaching 0. Using the early stopping method with a patience of 20 epochs, the model trained on 60 epochs (figure 22 and 23) with a run time of 8 hours and 20 minutes.

```
Epoch 1/60
100/100 [==============================] - 592s 6s/step - loss: 8.0118 - accuracy: 0.5447 - val_loss: 1.9007 - val_ac
curacy: 0.5980
Epoch 2/60
100/100 [==============================] - 469s 5s/step - loss: 2.6484 - accuracy: 0.5711 - val_loss: 1.0934 - val_ac
curacy: 0.6281
Epoch 3/60
100/100 [==============================] - 556s 6s/step - loss: 1.0915 - accuracy: 0.6616 - val_loss: 0.8594 - val_ac
curacy: 0.6633
Epoch 4/60
100/100 [==============================] - 528s 5s/step - loss: 1.0920 - accuracy: 0.6453 - val_loss: 0.7331 - val_ac
curacy: 0.6884
Epoch 5/60
100/100 [==============================] - 620s 6s/step - loss: 0.9124 - accuracy: 0.6755 - val_loss: 0.8041 - val_ac
curacy: 0.7035
Epoch 6/60
100/100 [==============================] - 612s 6s/step - loss: 0.8269 - accuracy: 0.6956 - val_loss: 0.9482 - val_ac
curacy: 0.6985
Epoch 7/60
100/100 [==============================] - 688s 7s/step - loss: 0.5798 - accuracy: 0.7233 - val_loss: 0.6273 - val_ac
curacy: 0.7186
Epoch 8/60
100/100 [==============================] - 646s 6s/step - loss: 0.5180 - accuracy: 0.7459 - val_loss: 0.6354 - val_ac
curacy: 0.7186
Epoch 9/60
100/100 [==============================] - 562s 6s/step - loss: 0.4451 - accuracy: 0.7836 - val_loss: 0.6812 - val_ac
curacy: 0.7236
Epoch 10/60
100/100 [==============================] - 577s 6s/step - loss: 0.4098 - accuracy: 0.8063 - val_loss: 0.7597 - val_ac
curacy: 0.7286
```

Figure 22: Training log: starting point.

```
Epoch 51/60
100/100 [==============================] - 420s 4s/step - loss: 0.2901 - accuracy: 0.8881 - val_loss: 0.5187 - val_ac
curacy: 0.8040
Epoch 52/60
100/100 [==============================] - 426s 4s/step - loss: 0.2523 - accuracy: 0.8994 - val_loss: 0.3576 - val_ac
curacy: 0.8643
Epoch 53/60
100/100 [==============================] - 430s 4s/step - loss: 0.2190 - accuracy: 0.9157 - val_loss: 0.4689 - val_ac
curacy: 0.8040
Epoch 54/60
100/100 [==============================] - 426s 4s/step - loss: 0.1697 - accuracy: 0.9472 - val_loss: 0.3642 - val_ac
curacy: 0.8543
Epoch 55/60
100/100 [==============================] - 448s 4s/step - loss: 0.1886 - accuracy: 0.9220 - val_loss: 0.3650 - val_ac
curacy: 0.8492
Epoch 56/60
100/100 [==============================] - 491s 5s/step - loss: 0.1087 - accuracy: 0.9686 - val_loss: 0.5747 - val_ac
curacy: 0.8342
Epoch 57/60
100/100 [==============================] - 493s 5s/step - loss: 0.0851 - accuracy: 0.9774 - val_loss: 0.3566 - val_ac
curacy: 0.8945
Epoch 58/60
100/100 [==============================] - 478s 5s/step - loss: 0.0594 - accuracy: 0.9899 - val_loss: 0.3423 - val_ac
curacy: 0.8995
Epoch 59/60
100/100 [==============================] - 498s 5s/step - loss: 0.1694 - accuracy: 0.9560 - val_loss: 0.4689 - val_ac
curacy: 0.8744
Epoch 60/60
100/100 [==============================] - 475s 5s/step - loss: 0.1372 - accuracy: 0.9572 - val_loss: 0.4591 - val_ac
curacy: 0.8643
```

Figure 23: Training log: end point.

The metrics attained by the end of the training are:

- Training loss: 0.14
- Training accuracy: 0.96
- Validation loss: val_loss: 0.46
- Validation accuracy: val_accuracy: 0.86

Where loss and accuracy are calculated on the training set and val_loss as well as val_accuracy on the validation set. Although, the form of the accuracy curve on figure 24 is a bit unusual, as it goes down abruptly after a certain number of epochs. This could be explained by the fact that the learning rate may become so big once the accuracy goes above a certain value, and therefore diverges away from the optimum. Such a problem could be addressed using an adaptive learning rate that gets smaller by the time the model's performance becomes higher.

Other tests were carried out using a smaller learning rate, though they led to a significantly longer run time and lower accuracy.

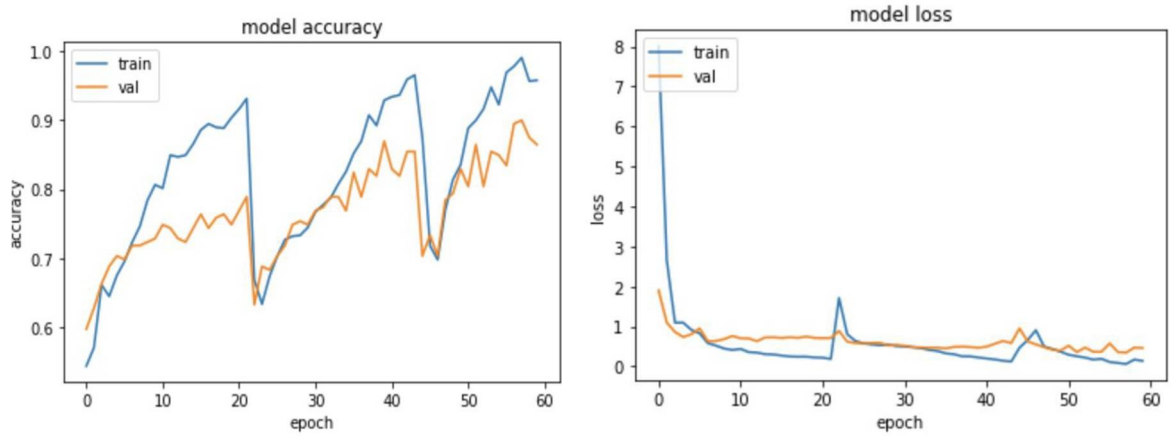


Figure 24: The evolution of the image-based model's accuracy and loss through the learning

epochs

2.3 Discussion

In this study, two approaches were developed, optimized and tested on patients' MR images to differentiate between LGG and GBM.

The first approach was based on high dimensional radiomic features extracted from patients' MR scans that we analyzed and employed with state-of-the-art machine learning algorithms. Our radiomic method was evaluated on 3 types of classifiers; a decision tree model, a random forest and XGBoost, we achieved a mean value of AUC and Accuracy of 0.91 and 0.94 respectively. In particular, the XGBoost model showed the best overall performance with an AUC of 0.95 and an accuracy of 0.94 on the validation set.

We compared the radiomics method with an image-based approach, we put in place a pretrained convolutional neural network for image recognition. By the means of this approach, we achieved an accuracy on validation set of 0.86 trained on 60 epochs.

Both of the models achieved high performances which shows the value of machine learning methods in distinguishing with a high confidence and in an automatic way between LGG and GBM MR images. Moreover, we perceived that radiomics bring an added value and more precise predictions since the radiomic-based XGBoost model outperformed the VGG16 image-based neural network. This shows promising potential for radiomics applications in computer-aided tumor grading. And it would be more clinically relevant to reproduce these results for differentiating between images of glioma grade II and III.

Chapter II: IDH mutation status prediction in LGG

3 Chapter II: IDH mutation status prediction in LGG

3.1 Materials and methods

3.1.1 Dataset

For studying IDH mutation and LGG tumor, we included in the study 65 LGG patients' scans available on the TCIA website [82]. We employed a database of segmented and original imaged of TCGA-LGG patients in Dicom file format for medical imaging. [83] The patients MRI scans include different modalities (i.e. T1, T1-Gd, T2, T2-FLAIR) and for each modality multiple segmentation masks for the core tumor, necrosis, edema and enhanced tumor. This different MR images with their corresponding masks were used for feature extraction, a step we will describe in more detail in the next section.

In order to gather information on IDH mutation status, we used a collected clinical data that includes a number of clinical information on the patient like gender, age, tumor site, medical history and genetic background, from which we extracted the IDH mutation status for each patient. The database gathers information on all LGG patients, we chose from them only the ones for which we also have the MRI scans. It is also available on TCIA's public website along with the TCGA-LGG database [84].

3.1.2 Feature extraction

Here again, radiomics features are used to predict the IDH mutation, only this time we extracted our own radiomic features from the segmented images and used the result of 115 features as an entry for our artificial models. Each entry corresponds to an image with a different type of mask for a different part of the tumor: Non enhanced tumor core, enhanced tumor, and edema.

The radiomic feature extraction was done using PyRadiomics package. We started by converting the images and the segmentation masks to the supported format in PyRadiomics (dicom slices to 3D NIfTI volumes), associating the segmentation masks to the corresponding images, resampling segmentation masks to the reference image to them to have the same the origin and spacing in space and finally extracting the radiomics features. For each image, we dispose of multiple sets of radiomics features, including shape feature as well as first order histogram and second order texture features. Below is the description of the extracted features:

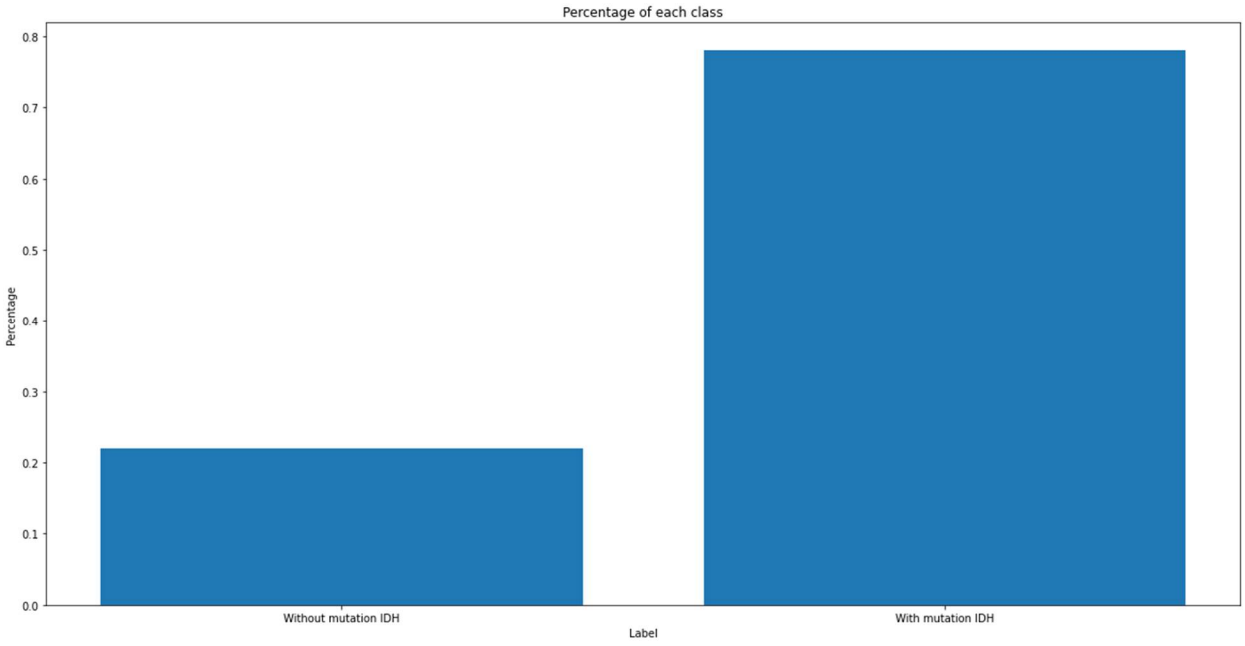
- Shape features: 14
- First order features: 18
- Glcm features: 24
- Gldm features: 14
- Glrlm features: 16
- Glszm features: 16
- Ngtdm features: 5

More statistical details on the extracted radiomic features are available in Annexe 2 (Table 2).

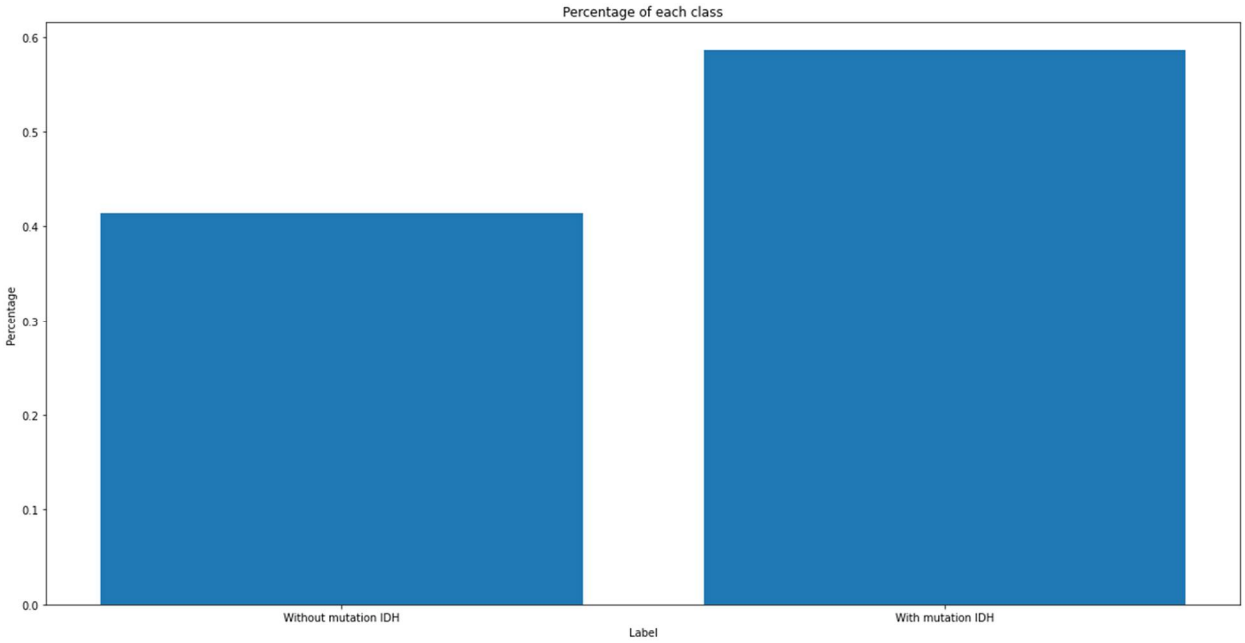
3.1.3 Feature Preprocessing

Before we train our models, data need to be preprocessed and cleaned to extract maximum information of it. This includes purifying our database of any irrelevant data, detecting and handling missing values, scaling columns to bring values to the same range, sampling our data as well as reducing database's dimensionality.

In the case of predicting IDH mutation, we started by omitting nonnumerical values, since we are only interested in quantitative features of the images database leaving us with 111 radiomic features. Afterwards, we standardized the columns bringing all values around the center of 0 with an overall standard deviation of 1. To select the pertinent features, we used first the lasso regression which is merely a linear regression with L1 regularization. It consists of shrinking the weight of different features to 0 when the latter have a very low impact on the prediction. After applying this method of feature selection, we ended up with 36 out of 111 that we use for our prediction. Lastly, before starting to develop our machine learning algorithms, we observed that our data is significantly imbalanced (figure 25 - A) as only 14 out of 65 patients don’t have mutation of the enzyme IDH against 51 out of 65 who have the mutation. To overcome this problem we considered different solutions, such as using metrics that are less sensitive to imbalanced data, associating a greater weight to the disadvantaged class and under sampling the advantaged class. Hence, we randomly omit 31 positive patients (with IDH mutation) with the purpose of obtaining a more balanced dataset that results in 20 positive patients versus 14 negative ones (figure 25 - B).



(A) Percentage of the number of patients per class before under sampling



(B) Percentage of the number of patients per class after under sampling

Figure 25: Percentage of the number of patients per class for IDH mutation prediction

We studied lastly the correlation in-between features and between the features and the target endpoint (which is the IDH mutation status). For that we plotted the correlation matrix as a heat map. Followed by Principal Component Analysis (PCA) or Single Value Decomposition (SVD) algorithms for dimensionality reduction that allow at the same to surpass the problem of correlated features as well as reduce the dimension of our dataset.

3.1.4 Machine learning algorithms

Given the limited data points available in our dataset of features, we considered using classification methods that are adapted to this case. These classification algorithms are:

Logistic Regression, Random Forest (RF), XGboost, K-Nearest Neighbour (KNN).

In the following, we describe two of the used models : logistic regression and KNN, while the other models were explained in the previous sections.

Logistic Regression

Logistic regression is a statistical model that basically uses the logistics function (or logit) as a link function. It is usually used as a regression analysis to when the assessed variable data are binary.

K-Nearest Neighbor

The KNN model is one of the simplest and yet most used models in machine learning. As the name suggests it, the idea behind its algorithm is to classify a new input based on its k nearest neighbors. This model can be used for classification tasks. In order to estimate the output for a new entry, the KNN algorithm computes the majoritarian class among the k nearest neighbors in the feature space using a chosen distance. Thus, the training phase only consists of storing the training examples and their labels.

Before training our final model, we started by optimizing the model, comparing the models to each other and selecting select the one with the best performances. In the following section, we describe more in detail this step.

3.1.5 Model selection

To select our model, we constructed a pipeline that gathers a part of the processing as well as state of the art models previously presented. In the pipeline, we inserted the PCA and SVD algorithms, standardization functions as well as the different classification algorithms.

For each of component of the pipeline, we set a number of candidate values for each hyperparameter that needs to be optimized. Here again we used a Grid Search that tests the different combinations of these hyperparameters until the highest performances are reached. In addition, we set the accuracy as a scoring metric that the Grid Search aims to optimize and chooses accordingly the best model.

3.1.6 Model training

After comparing the different models, we chose the one that is the most adapted to our application. We train this model using the fine-tuned hyperparameters on the training set and evaluate it later on the test set.

3.2 Results

Here again, we looked for any correlation between the feature or between the features and IDH mutation status. For that reason, we plotted the correlation matrix as a heat map in the following figure (figure 26). We noticed that no feature is significantly correlated to our target. However, some of them are inter-correlated and thus need to be orthogonalized.

By the end of the Grid search the SVD algorithm was selected for the dimensionality reduction algorithm with 35 components. As for the model, the XGBoost model was opted for with the following hypermeters:

- Learning rate : 0.01
- Max depth of each tree : 1
- Number of estimators : 100

In order to confirm the grid search's result and compare the different models, we illustrated the value of ROC AUC for each model (figure 27).

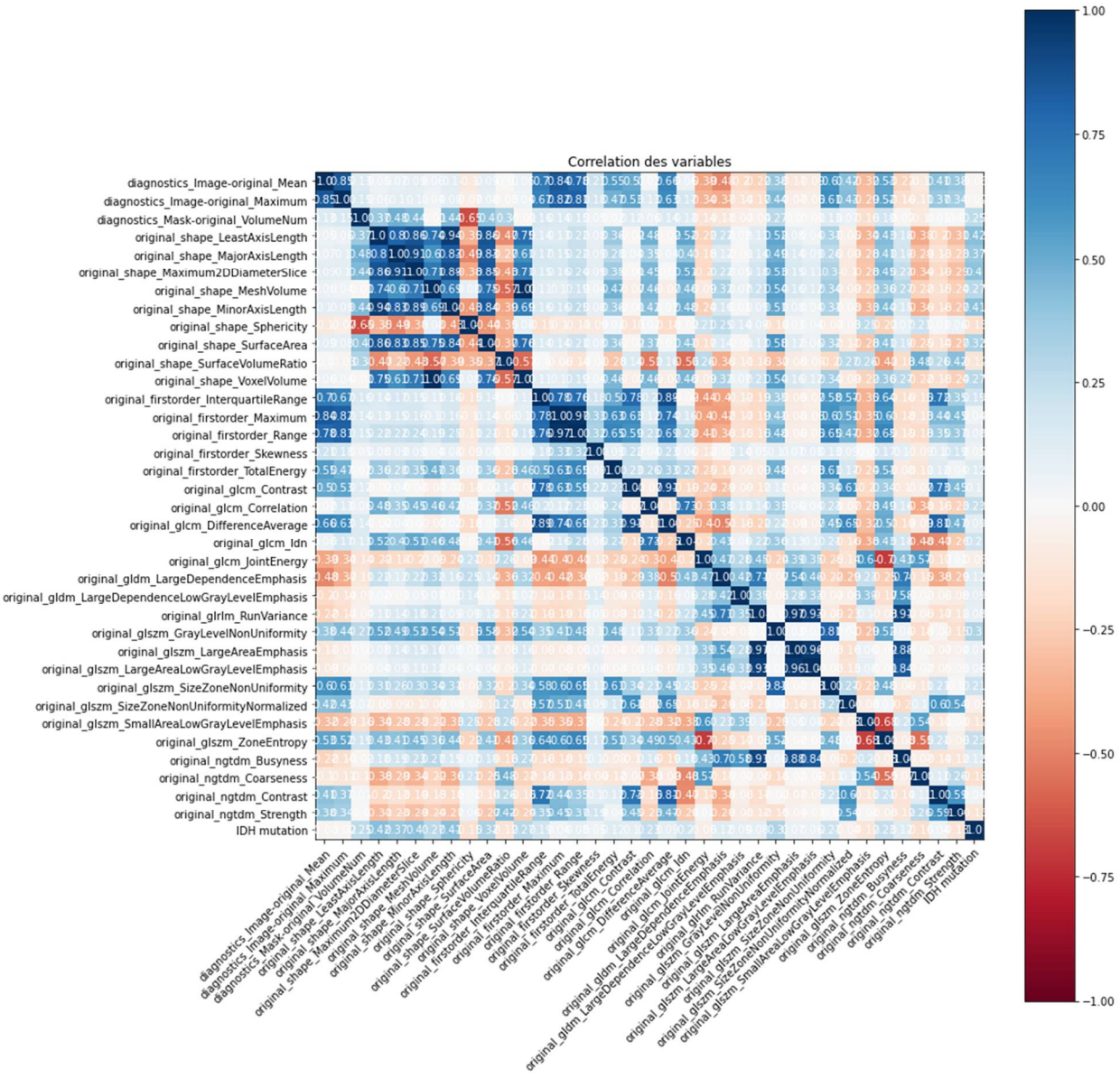


Figure 26: Heatmap of the correlation between radiomic features for IDH mutation prediction

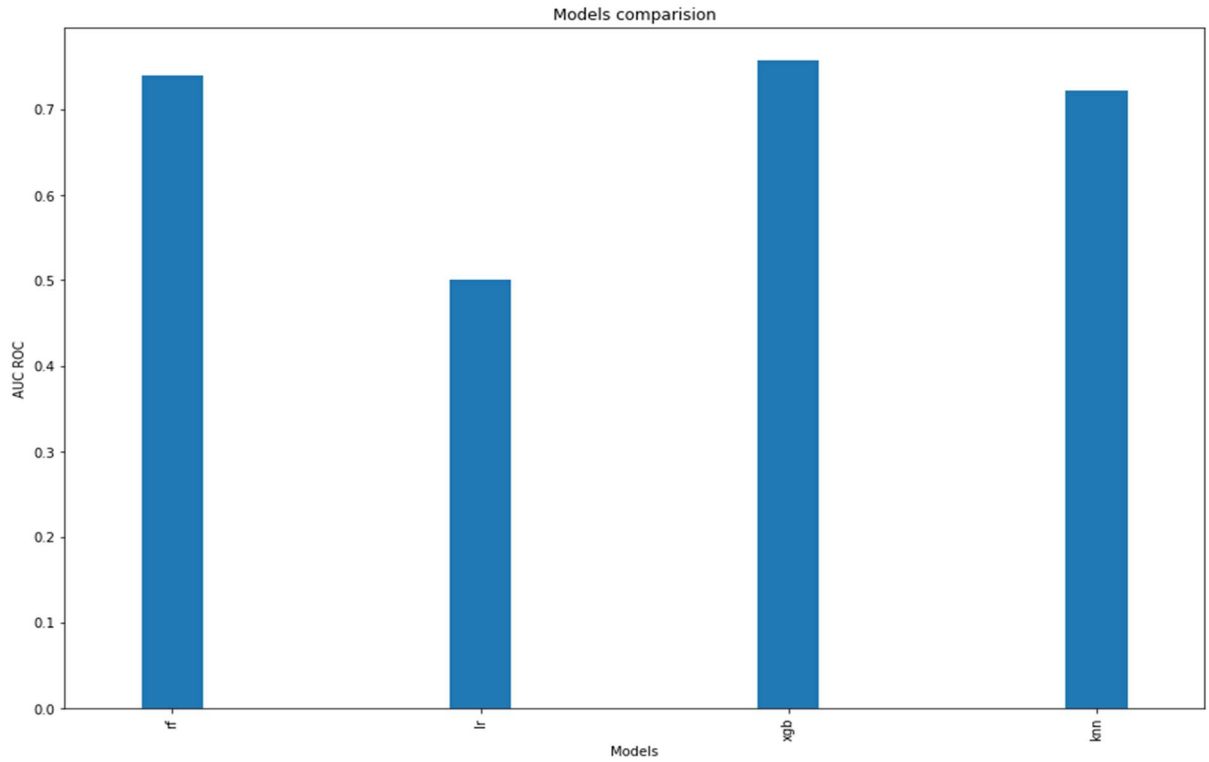


Figure 27: Comparison of different models based on ROC AUC score for IDH mutation prediction

For the prediction of IDH mutation status, we found the results illustrated in the table below (figure 28). We noticed that the machine learning algorithm's precision reaches a value of 0.85 making it a reliable method to predict people with IDH mutation. However, the recall had a value of only 0.61 meaning that the algorithm misses on some of the positive patients. Overall, we reached a fairly good value of 0.72 and 0.77 for f1 score and ROC AUC respectively. Regarding the accuracy, a relatively low value of 0.69 is found. This could be explained with the small number of entry data as well as the high imbalance between the two classes in number making it hard on the model to extract information and learn to predict accurately both classes.

Precision	recall	f1-score	accuracy	ROC AUC
0.85	0.63	0.72	0.69	0.77

Figure 28: Table of metrics of the selected model XGBoost for IDH mutation prediction

For illustration purposes, we plot the ROC curve in the figure below (figure 29). We notice that the curve is above the identity line that represent the ROC curve of a random classifier. The curve confirms the found value of ROC AUC and shows that our classifier has learnt some information of our database.

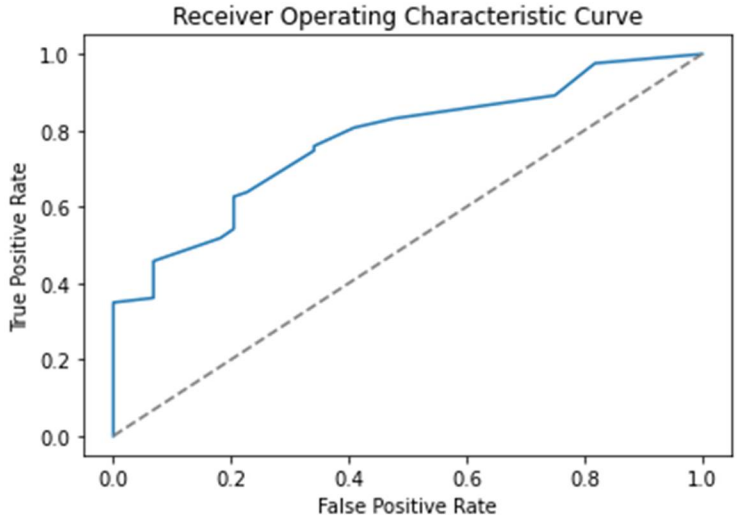


Figure 29: Receiver Operating Characteristic curve of best model for IDH mutation prediction

3.3 Discussion

We explore in this part of study the extraction of radiomic features and their usefulness in detecting IDH mutation status in LGG patients. We evaluated different types of machine learning models: Random Forest, XGBoost, Logistic Regression and K-Nearest Neighbor. The model with the best performances was XGBoost with an accuracy and AUC of 0.69 and 0.77 respectively.

With an AUC value of 0.77, our model performed fairly well in predicting IDH mutation in MR images of LGG patients.

To compare with other studies in the literature, we conducted a limited-scope review of the literature, summarized in (Table. 1) of (Annex. I), on the combined application of radiomics and artificial intelligence in IDH status prediction in Gliomas.

Out of the 15 papers that aimed to detect IDH mutation, some of the studies had higher performance and AUC values ranging (79% to 97%), which could be explained by their significantly larger image dataset and higher number of included patients compared to the 34 patients in our study (after under-sampling). Almost all the other studies used images of more than 100 patients, thus, their results are expected to outperform our proposed radiomic-based machine learning algorithms. As the higher the number of cases the model trains on, the better its performance becomes.

Another factor that could be in play is presented by Ren Y et al. [--] study with a slightly comparable number of patients = 56 to ours states results with an AUC of 0.93, the difference in performance may be attributed to the higher resolution of the images used in the study which were obtained with a 3-Tesla MRI machine.

Therefore, we recommend re-conducting our study using a higher resolution MRI or a larger number of patients.

It is worth noting that the accuracy of our model at 69% was slightly lower than human neuroradiologists with 69% and 72% respectively. PPV (Precision) in our model was higher than traditional neuroradiologist prediction at 85% and 17% respectively.

In addition, the significant imbalance between the two classes and the lack of enough examples for our both classes was problematic. And compelled us to under-sample the IDH positive class to correct the problem, reducing drastically the total number of patients from 65 to 34.

Conclusion

4 Conclusion

While High Grade Gliomas remain a major healthcare challenge in our current time, the diagnostic and analytic tools have become very democratized, globally available and open source for scientists worldwide. We have been able in the study to apply feature-based and image-based classifiers with high accuracy for purpose of differentiation between MR images of low-grade glioma and glioblastoma.

We also succeeded in using state of the art tools for radiomic feature extraction and reached fairly satisfying results in predicting the IDH mutation status for LGG patients given the scarcity of the data and the dataset imbalance. Though, it is worth noting that these methods and results are still subject to improvement, once more data is available.

This study serves as a foundational step to bridge the results with other types of “omics” data, and gives a glimpse of confirmation of the utility, potential, and promising future of integrated diagnostics as well as virtual digital biopsies. This could be made possible by merging and leveraging radiomics with machine learning tools to help make brain tumors diagnosis more accurate, precise and individualized.

More steps could be taken to discover new biomarker correlates, in order to accelerate the development of novel therapies for CNS neoplasms and facilitate reaching more precise and faster access to gold standard diagnosis, and hopefully arrive at a point in the future where we can noninvasively inform eventual therapeutic venues.

Résumés

Résumé

Titre : Classification d'images des gliomes à l'aide de l'apprentissage automatique.

Auteur : Mohamed Sobhi JABAL

Rapporteur : Pr. Mahjouba BOUTARBOUCH

Mots clés : Gliome – Neuroimagerie – Radiomique – Intelligence Artificielle.

Les gliomes sont les tumeurs cérébrales primaires les plus fréquentes, avec une histologie et profils génomique hétérogènes. Elles sont classifiées a des tumeurs bénignes et malignes selon leur grade de I à IV.

Cette étude vise à appliquer différents outils d'apprentissage automatique pour le classement et la classification des gliomes en mettant l'accent sur l'approche radiomique dans l'imagerie par résonance magnétique cérébrale des adultes humains.

Expérimentalement in silico, nous avons performé différents modèles d'apprentissage automatique pour la classification des gliomes. La performance a été mesurée à travers plusieurs paramètres, principalement le taux de précision et l'aire sous la courbe (AUC) de la caractéristique de fonctionnement du récepteur (ROC). Les données d'imagerie ont été obtenues de The Cancer Imaging Archive (TCIA), en relation avec des patients de The Cancer Genome Atlas (TCGA), correspondant à des images IRM cérébrales de patients diagnostiqués avec un gliome.

Notre approche et les modèles d'apprentissage automatique appliqués ont réussi à classifier correctement les images RM de différents grades de gliomes et à atteindre des taux de précision élevés sur l'ensemble de données de validation.

En conclusion, l'apprentissage automatique est un outil prometteur émergent pour aider au diagnostic ainsi qu'à la prédiction du pronostic clinique des gliomes, contribuant ainsi à la transition vers une ère de médecine de précision dans les soins de santé.

Abstract

Title: Image Classification of Gliomas using machine learning.

Author: Mohamed Sobhi JABAL

Director: Pr. Mahjouba BOUTARBOUCH

Keywords: Glioma – Neuroimaging – Radiomics – Artificial Intelligence.

Gliomas are the most common primary brain tumors, with heterogeneous histology and genomic profiles. They are classified into benign and malignant tumors according to their grade from I to IV.

This study aims to apply different machine learning tools for grading and classification of gliomas with a focus on radiomics approach in brain MR imaging of human adults.

Experimentally in silico, we have performed different machine learning models for classifying gliomas. The performance was measured through multiple metrics, mainly the Accuracy rate and the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC). Imaging data was obtained from The Cancer Imaging Archive (TCIA), in relation to patients of The Cancer Genome Atlas (TCGA), corresponding to brain MR images of patients diagnosed with glioma.

Our approach and applied machine learning models succeeded in the correctly classifying MR images of different gliomas grades and achieved high accuracy rates on the validation dataset.

In conclusion, machine learning is an emerging promising tool for assisting the diagnosis as well as the clinical prognosis prediction of gliomas, helping with the transition to a precision medicine era in healthcare.

ملخص

العنوان: تصنيف الصور للأورام الدبقية باستخدام التعلم الآلي.

تأليف: محمد صبحي جبل

الاشراف: الأستاذة محجوبة بوطربوش

الكلمات المفتاحية: الورم الدبقي - التصوير العصبي - الراديوميكس - الذكاء الاصطناعي.

الأورام الدبقية هي أورام الدماغ الأولية الأكثر شيوعًا ، ذات أنسجة و خواص جينية مختلفة. تقسم الى أورام حميدة وخبيثة حسب تصنيفها من I الى IV.

تهدف هذه الدراسة إلى تطبيق أدوات مختلفة للتعلم الآلي لتصنيف وتمييز الأورام الدبقية مع التركيز على نهج الراديوميكس في تصوير الدماغ بالرنين المغناطيسي للبشر البالغين.

تجريبيًا في السيليكو ، أجرينا تطبيق لنماذج مختلفة من التعلم الآلي لتصنيف الأورام الدبقية. تم قياس الأداء من خلال مقاييس متعددة ، وبشكل أساسي معدل الدقة والمنطقة الواقعة تحت المنحنى (AUC) لخاصية تشغيل جهاز الاستقبال (ROC). تم الحصول على بيانات التصوير من أرشيف تصوير السرطان (TCIA) ، فيما يتعلق بمرضى أطلس جينوم السرطان (TCGA) ، المقابلة لصور الرنين المغناطيسي للدماغ لمرضى تم تشخيص إصابتهم بالورم الدبقي.

نجحت مقاربتنا ونماذج التعلم الآلي التطبيقية في تصنيف صور الرنين المغناطيسي لدرجات مختلفة من الأورام الدبقية بشكل صحيح وحققت معدلات دقة عالية في مجموعة بيانات التحقق.

في الاستنتاج ، التعلم الآلي تعد أداة واعدة ناشئة للمساعدة في التشخيص بالإضافة إلى التنبؤ السريري للأورام الدبقية ، مما يساعد في الانتقال إلى عصر الطب الدقيق في الرعاية الصحية.

Annexes

ANNEX. 1

PubMed search string:

("radiomic*") AND ("glioma"[Mesh]) AND ("artificial intelligence"[Mesh] OR "Diagnosis, Computer-Assisted"[mesh] OR "Image Interpretation, Computer-Assisted"[Mesh] OR "Image Processing, Computer-Assisted"[mesh]) AND ("Isocitrate Dehydrogenase"[Mesh] OR "IDH" OR "IDH1" OR "IDH2")

Results returned: 21 / Added through references pearl growing: 4 / Included: 20

Table. I: Review of literature on Radiomic based AI for IDH status prediction.

First Author	Pub Year	Disease	No. of patients	Imaging modality / sequences	Biological Correlate	Outcome	Findings	Classification Method
Peng H	2020	Glioma (grade II to IV)	105	MRI (T1ce - T2 - ASL)	IDH	Accuracy AUC	ACC = 0.77 AUC = 0.823	SVM
Sakai Y	2020	Glioma (grade II to IV)	100	MRI (FLAIR - DWI)	IDH1	Accuracy AUC	DWI: ACC = 0.90 AUC = 0.97 f1 = 0.75 FLAIR: ACC = 0.90 AUC = 0.95 f1 = 0.75	XGBoost

Niu L	2020	High-grade Glioma (III -IV)	182	MRI	IDH1	AUC	AUC = 0.86 ACC = 0.789 Sensitivity = 91.3% Specificity = 69.0% PPV = 0.700 NPV = 0.909	ROC
Kong Z	2020	Glioma (grade II - III) LGG	96	MRI (3D T1ce -T2 - SC T1ce)	1p/19q Co-deletion + IDH	Accuracy AUC	3D: IDH: AUC = 0.950	RF
Calabrese E	2020	GBM	199	MRI (T1 pre, T1 post, T2, T2/FLAIR, SWI, DWI, ASL, MD, AD, RD, and FA)	IDH ATRX chromosome 7/10 aneuploidies CDKN2	AUC Sensitivity Specificity	IDH: (sensitivity = 0.93, specificity = 0.88) AUC = 0.95	RF
Lo CM	2020	GBM	39 TCIA	MRI	IDH	Accuracy	Machine: ACC = 90% Human Neuroradiologists: ACC = 72%	logistic regression KNN SVM
Fukuma R	2019	grade II/III glioma	164	T1W T2W GdT1W FLAIR	IDH and TERT genotypes	Accuracy	accuracy = 63.1%	Feature extraction: CNN (AlexNet) support vector machine

Kim M	2020	LGGs	127	CE-T1WI FLAIR DWI DSC	IDH mutation tumor grade	AUC	IDH mutation AUC = 0.747 tumor grading AUC = 0.819	
Park CJ	2020	LGG	168	T1-weighted image with contrast enhancement, T2-weighted image, and FLAIR	(IDH) mutation status	AUC 95% confidence interval (CI)	DTI radiomics + conventional radiomics AUC = 0.900 95% CI = [0.855-0.945] DTI histogram parameters + conventional radiomics AUC = 0.869 95% CI = [0.816-0.922] conventional radiomics AUC = 0.835 95% CI = [0.773-0.896]	random forest
Tan Y	2019	astrocytomas (Grades II-IV)	105	1-weighted imaging (CE-T1WI) T2 fluid-attenuated inversion recovery (T2FLAIR) apparent diffusion coefficient (ADC)	IDH genotype	AUC Survival analysis: log rank test, C-index, and z-score test	radiomics models AUC = 0.888 clinico-radiological models AUC = 0.804 radiomics nomogram models AUC = 0.900 log rank test, $p < 0.001$ C-index = 0.762 and 0.687 z-score test, $p = 0.062$	SVM
Lee MH	2019	GBM	88	T2-weighted postcontrast 3D-T1-fast-field FLAIR with or without contrast	IDH1 mutation	accuracy prediction rate	prediction rate = 70.3%-87.3% accuracy = 66.3%-83.4%	K- Nearest Neighbors Support Vector Classification Decision Tree Random Forest AdaBoost Naive Bayes Linear Discriminant Analysis Gradient Boosting

Wu S	2019	diffuse gliomas	126	T1 T1-Gd T2 T2-FLAIR	IDH genotype	accuracy AUC	RF accuracy = 0.885 RF AUC = 0.931 NN accuracy = 0.829 NN AUC = 0.878	Random Forest (RF) neural network (NN) flexible discriminant analysis (FDA)
Ren Y	2019	LGG	57	3D arterial spin labeling (3D-ASL) T2 fluid-attenuated inversion recovery (T2 FLAIR) diffusion-weighted imaging (DWI).	IDH1(+) and ATRX (-)	accuracy AUC sensitivity specificity PPV NPV	IDH1 (+) accuracy: 94.74% IDH1 (+) AUC: 0.931 IDH1 (+) sensitivity : 100% IDH1 (+) specificity : 85.71% IDH1 (+) PPV : 92.31% IDH1 (+) NPV : 100%	SVM
Lohmann P	2018		84		IDH genotype			
Lu CF	2018	LGG and GBM	456	post-contrast T1-weighted (T1 + C), T2 FLAIR T2-weighted (T2W) optional diffusion-weighted (DWI) optional	IDH and 1p/19q classification of 5 molecular subtypes	Accuracy AUC	IDH accuracy = 87.7% IDH AUC = 0.922 1p/19q accuracy = 96.1% 1p/19q AUC = 0.975 subtypes based on MR phenotypes accuracy = 81.8% subtypes based also on histology diagnosis accuracy = 89.2%	SVM three-level binary classification model : first : GBM vs LGG second : IDH mutation vs wild type third : 1p/19q codelation vs no codelation forth: tumor subtype

Diamandis E	2018	Gliomas	65	3D T1 before and after contrast 3D FLAIR SPACE 2D T2	predicting the molecular subtype in gliomas	accuracy	accuracy = 93.8%	Random Forest
Zhang X	2018	LGG	103	T1 (before and after contrast-enhanced) T2 FLAIR	IDH TP53	Accuracy sensitivity specificity AUC	IDH test: (accuracy = 80.0%, AUC = 0.792)	SVM
Li Z	2017	LGG	151	MRI (T1CE, T2 FLAIR)	IDH1	AUC	normal radiomics AUC = 86% Deep learning radiomics AUC = 92% DL multimodal MRI AUC = 95%	SVM
Hsieh KL	2017	GBM	39	MRI (T1)	IDH	Accuracy agreement level (κ)	Accuracy = 85% agreement level (k) = 0.60	logistic regression
Yu J	2017	Glioma (grade II)	110	MRI (T2-FLAIR)	IDH1	accuracy sensitivity specificity AUC	accuracy = 0.80 sensitivity = 0.83 specificity = 0.74 AUC = 0.86.	SVM AdaBoost Leave-one-out cross-validation (LOOCV)

ANNEX. 2

Table. II: Extracted radiomic features by category using PyRadiomics.

Feature Name	mean	std	min	max
diagnostics_Image-original_Mean	248.4494803	159.5892585	29.30315749	910.2940996
diagnostics_Image-original_Maximum	3561.395991	2671.925392	559	17148
diagnostics_Mask-original_VoxelNum	44186.34716	78834.3904	28	654106
diagnostics_Mask-original_VolumeNum	12.27109267	18.99718837	1	199
original_shape_Elongation	0.714584905	0.15606918	0.168060814	0.985881112
original_shape_Flatness	0.540927796	0.157234319	0	0.913808567
original_shape_LeastAxisLength	32.87595609	15.87899962	0	69.7708518
original_shape_MajorAxisLength	60.87360898	25.95543926	6.097551088	151.2109362
original_shape_Maximum2DDiameterColumn	56.47687284	22.3396113	5.021929256	116.9539339
original_shape_Maximum2DDiameterRow	66.34556219	28.14875842	6.638253253	135.7013934
original_shape_Maximum2DDiameterSlice	62.98635053	28.22630369	4.997958817	132.5740794
original_shape_Maximum3DDiameter	71.49749987	27.75965075	7.455335757	144.4786222
original_shape_MeshVolume	44178.56531	47780.96337	7.005278648	221517.1143
original_shape_MinorAxisLength	42.72028271	18.60269948	4.165654926	90.29185154
original_shape_Sphericity	0.377073383	0.15681483	0.127216063	0.847518304
original_shape_SurfaceArea	16511.92598	15144.76666	89.36939232	72815.49947
original_shape_SurfaceVolumeRatio	0.772706883	0.979099481	0.142440324	19.86706457
original_shape_VoxelVolume	44583.56971	47933.62101	30.76828269	222731.543
original_firstorder_10Percentile	614.4293814	424.9791792	10	2729.423828
original_firstorder_90Percentile	951.2093501	646.3551843	103	5865
original_firstorder_Energy	56474248261	2.13723E+11	860544	2.43462E+12
original_firstorder_Entropy	3.962360888	1.04044945	1.428284892	7.334382865
original_firstorder_InterquartileRange	177.8100588	167.4180584	18	2663
original_firstorder_Kurtosis	4.432168549	3.184086526	1.44729988	39.53005776
original_firstorder_Maximum	1400.031674	1124.818501	195.6068115	6932
original_firstorder_MeanAbsoluteDeviation	106.8734744	91.41518204	11.78734667	1233.205765
original_firstorder_Mean	780.0264307	526.2433769	51.2431792	4066.778083
original_firstorder_Median	776.4638448	522.886274	32	3873
original_firstorder_Minimum	262.5445354	302.0400681	0	2070.700439
original_firstorder_Range	1137.487139	1027.260883	91	6701
original_firstorder_RobustMeanAbsoluteDeviation	74.53409712	68.72547058	7.763179418	1065.7785
original_firstorder_RootMeanSquared	794.0194504	534.5522454	62.35319313	4284.585348
original_firstorder_Skewness	0.004263293	0.82988342	-2.759198269	4.182743865
original_firstorder_TotalEnergy	37484502264	79777126513	1056395.263	6.93432E+11
original_firstorder_Uniformity	0.09910732	0.07186313	0.006952902	0.459204029
original_firstorder_Variance	30827.47554	81723.086	254.6381434	1818987.625

original_glcm_Autocorrelation	761.5263617	1256.043844	6.171436481	14261.07497
original_glcm_ClusterProminence	568672.2965	6399786.348	6.03361201	166822530.5
original_glcm_ClusterShade	1034.378194	8232.807268	-51502.5846	92024.70938
original_glcm_ClusterTendency	154.6193658	461.8176257	1.062773904	10589.80383
original_glcm_Contrast	31.99909024	66.14718041	0.315205409	830.8110915
original_glcm_Correlation	0.562621629	0.222350532	-0.33832236	0.978168547
original_glcm_DifferenceAverage	3.106185436	2.527832716	0.275530006	21.79271765
original_glcm_DifferenceEntropy	2.749703985	0.853468355	0.51294317	5.65257209
original_glcm_DifferenceVariance	13.69176884	27.43058684	0.143240816	292.2465825
original_glcm_Id	0.457906583	0.134849154	0.119044001	0.867943614
original_glcm_Idm	0.395747408	0.154875358	0.06027852	0.866104099
original_glcm_Idmn	0.986056282	0.0146329	0.872751048	0.999653169
original_glcm_Idn	0.932738915	0.030990021	0.764778787	0.988528692
original_glcm_Imc1	-0.203639945	0.116887726	-0.888855246	-0.039768054
original_glcm_Imc2	0.782998285	0.129632696	0.371783428	0.999513825
original_glcm_InverseVariance	0.341304145	0.100819261	0.063493445	0.570033197
original_glcm_JointAverage	21.83207386	15.94599268	2.476025596	108.3101903
original_glcm_JointEnergy	0.028392	0.044331071	0.000191124	0.428281156
original_glcm_JointEntropy	6.981737722	1.989896958	1.895669295	13.08023386
original_glcm_MCC	0.689138519	0.148197963	0.297369773	1
original_glcm_MaximumProbability	0.056589655	0.067532204	0.000802348	0.509709628
original_glcm_SumAverage	43.66414772	31.89198536	4.952051191	216.6203806
original_glcm_SumEntropy	4.646465689	1.157619453	1.238837678	8.290646066
original_glcm_SumSquares	46.654614	126.2334951	0.457937905	2767.899827
original_gldm_DependenceEntropy	6.939505761	0.928504553	3.40057884	9.038999997
original_gldm_DependenceNonUniformity	5511.21483	10274.52895	5.681818182	89536.02424
original_gldm_DependenceNonUniformityNormalized	0.147451665	0.088186317	0.042723782	0.709342561
original_gldm_DependenceVariance	8.784929353	7.568812693	0.14532872	48.32909562
original_gldm_GrayLevelNonUniformity	3390.837886	8988.736452	1.947368421	132126.3159
original_gldm_GrayLevelVariance	49.40329063	130.7401064	0.478982032	2910.339244
original_gldm_HighGrayLevelEmphasis	763.4665035	1229.922463	6.907407407	13353.20744
original_gldm_LargeDependenceEmphasis	39.94522587	42.3667282	1.529411765	411.0189067
original_gldm_LargeDependenceHighGrayLevelEmphasis	15919.01474	23079.85881	37.36111111	215107.4069
original_gldm_LargeDependenceLowGrayLevelEmphasis	1.261803851	6.410003786	0.000800174	110.9804392
original_gldm_LowGrayLevelEmphasis	0.018053904	0.041147524	0.000169035	0.45425416
original_gldm_SmallDependenceEmphasis	0.210124987	0.151222067	0.008779432	0.867647059
original_gldm_SmallDependenceHighGrayLevelEmphasis	229.0400311	510.7592381	0.142381137	7040.677392
original_gldm_SmallDependenceLowGrayLevelEmphasis	0.003639423	0.009289141	3.15E-05	0.151415466

original_glrlm_GrayLevelNonUniformity	2303.214393	3950.540171	1.764126533	43218.16307
original_glrlm_GrayLevelNonUniformityNormalized	0.094280044	0.065394147	0.006894519	0.368959446
original_glrlm_GrayLevelVariance	50.181056	130.1692842	0.653867039	2883.639223
original_glrlm_HighGrayLevelRunEmphasis	762.6864801	1226.572374	7.07353204	13252.64868
original_glrlm_LongRunEmphasis	2.045104841	1.892496077	1.042328833	40.98947897
original_glrlm_LongRunHighGrayLevelEmphasis	1148.427862	1658.217334	10.49448004	15047.82421
original_glrlm_LongRunLowGrayLevelEmphasis	0.053357958	0.22842912	0.00021432	4.243503222
original_glrlm_LowGrayLevelRunEmphasis	0.017936296	0.038400275	0.000173066	0.327262368
original_glrlm_RunEntropy	4.676671655	0.839320549	2.272831418	7.541857619
original_glrlm_RunLengthNonUniformity	27660.08807	51126.61609	15.07943144	462800.2947
original_glrlm_RunLengthNonUniformityNormalized	0.760372819	0.127371976	0.246917897	0.973179999
original_glrlm_RunPercentage	0.844814825	0.096333147	0.305983969	0.98654897
original_glrlm_RunVariance	0.449820023	1.068195853	0.013410001	24.19088084
original_glrlm_ShortRunEmphasis	0.882336846	0.074571142	0.483570796	0.989417792
original_glrlm_ShortRunHighGrayLevelEmphasis	704.9664379	1158.989436	6.331027891	12857.04809
original_glrlm_ShortRunLowGrayLevelEmphasis	0.015070264	0.030554767	0.000165177	0.298733176
original_glszm_GrayLevelNonUniformity	279.2861421	410.8680836	1.424242424	3292.470901
original_glszm_GrayLevelNonUniformityNormalized	0.070823554	0.045853971	0.006434415	0.278016529
original_glszm_GrayLevelVariance	60.93103411	125.3012687	1.382857143	2547.517108
original_glszm_HighGrayLevelZoneEmphasis	756.7074501	1202.754243	7.307692308	12063.25505
original_glszm_LargeAreaEmphasis	110482.1189	1208992.052	1.290322581	27517692.73
original_glszm_LargeAreaHighGrayLevelEmphasis	7287448.916	28029222.87	33.59090909	346761415.9
original_glszm_LargeAreaLowGrayLevelEmphasis	6737.087349	108113.2308	0.000653426	2808767.524
original_glszm_LowGrayLevelZoneEmphasis	0.021674583	0.039839352	0.000213794	0.414401709
original_glszm_SizeZoneNonUniformity	2825.568919	5938.335012	3	50584.03962
original_glszm_SizeZoneNonUniformityNormalized	0.332571499	0.099085445	0.10083752	0.825182102
original_glszm_SmallAreaEmphasis	0.585730377	0.093189228	0.2273421	0.927419355
original_glszm_SmallAreaHighGrayLevelEmphasis	483.4719354	826.6652023	3.546806891	9285.715963
original_glszm_SmallAreaLowGrayLevelEmphasis	0.012907899	0.022481497	0.00010782	0.217438568
original_glszm_ZoneEntropy	6.501460738	1.01415798	2.918295834	8.611237987
original_glszm_ZonePercentage	0.234076975	0.181701654	0.00461893	0.911764706
original_glszm_ZoneVariance	110145.9745	1207095.688	0.087408949	27470820.32
original_ngtdm_Busyness	7.451187924	27.48151335	0.007655148	407.2394252
original_ngtdm_Coarseness	0.008894638	0.031434735	2.85E-05	0.356839422
original_ngtdm_Complexity	5027.0002	13821.04244	5.702568536	136314.8029
original_ngtdm_Contrast	0.096976419	0.203836831	0.001845399	3.132964565
original_ngtdm_Strength	2.660417167	12.35353026	0.002736612	243.8516579

References

[1] Chen, J., R.M. McKay, and L.F. Parada, Malignant glioma: lessons from genomics, mouse models, and stem cells. Cell, 2012. 149(1): p. 36-47.

[2] Charles, N.A., et al., The brain tumor microenvironment. Glia, 2012. 60(3): p. 502-14.

[3] Dirks, P.B., Brain tumor stem cells: bringing order to the chaos of brain cancer. J Clin Oncol, 2008. 26(17): p. 2916-24.

[4] Nguyen, L.V., et al., Cancer stem cells: an evolving concept. Nat Rev Cancer, 2012. 12(2): p. 133-43.

[5] Céline S. Gonçalves, T.L.o., Ana Xavier-Magalhães, Marta Pojo and Bruno M. Costa (February 27th 2013). Mechanisms of Aggressiveness in Glioblastoma: Prognostic and Potential Therapeutic Insights, Evolution of the Molecular Biology of Brain Tumors and the Therapeutic Implications, Terry Lichtor, IntechOpen, DOI: 10.5772/52361. Available from: https://www.intechopen.com/books/evolution-of-the-molecular-biology-of-brain-tumors-and-the-therapeutic-implications/mechanisms-of-aggressiveness-in-glioblastoma-prognostic-and-potential-therapeutic-insights.

[6] Global cancer observatory: cancer today. Lyon, F., International Agency for Research on Cancer. https://gco.iarc.fr/today. Accessed November 20, 2020.

[7] Registre des Cancers de la région du grand Casablanca pour la période 2008 - 2012, E.h.w.c.m.s.m.u.f.R.p.

[8] Tazi, M.A., A. Er-Raki, and N. Benjaafar, Cancer incidence in Rabat, Morocco: 2006-2008. Ecancermedicalscience, 2013. 7: p. 338.

[9] National Cancer Institute Surveillance, E., and End Results Program. Cancer Stat Facts: Brain and Other Nervous System Cancer. http://seer.cancer.gov/statfacts/html/brain.html. Accessed 11, 30, 2020.

[10] Aldape, K., et al., Challenges to curing primary brain tumours. Nat Rev Clin Oncol, 2019. 16(8): p. 509-520.

[11] Bray, F., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, 2018. 68(6): p. 394-424.

[12] Ostrom, Q.T., et al., CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2013-2017. Neuro Oncol, 2020. 22(12 Suppl 2): p. iv1-iv96.

[13] Ostrom, Q.T., et al., Epidemiology of gliomas. Cancer Treat Res, 2015. 163: p. 1-14.

[14] Ohgaki, H. and P. Kleihues, The definition of primary and secondary glioblastoma. Clin Cancer Res, 2013. 19(4): p. 764-72.

[15] McNeill, K.A., Epidemiology of Brain Tumors. Neurol Clin, 2016. 34(4): p. 981-998.

[16] Butowski, N.A., Epidemiology and diagnosis of brain tumors. Continuum (Minneap Minn), 2015. 21(2 Neuro-oncology): p. 301-13.

[17] Ostrom, Q.T., et al., Epidemiology of Intracranial Gliomas. Prog Neurol Surg, 2018. 30: p. 1-11.

[18] Ranger, A.M., et al., Familial syndromes associated with intracranial tumours: a review. Childs Nerv Syst, 2014. 30(1): p. 47-64.

[19] Goodenberger, M.L. and R.B. Jenkins, Genetics of adult glioma. Cancer Genet, 2012. 205(12): p. 613-21.

[20] Braganza, M.Z., et al., Ionizing radiation and the risk of brain and central nervous system tumors: a systematic review. Neuro Oncol, 2012. 14(11): p. 1316-24.

[21] Shah, N.B. and S.L. Platt, ALARA: is there a cause for alarm? Reducing radiation risks from computed tomography scanning in children. Curr Opin Pediatr, 2008. 20(3): p. 243-7.

[22] Turner, M.C., et al., Allergy and brain tumors in the INTERPHONE study: pooled results from Australia, Canada, France, Israel, and New Zealand. Cancer Causes Control, 2013. 24(5): p. 949-60.

[23] Scheurer, M.E., et al., Effects of antihistamine and anti-inflammatory medication use on risk of specific glioma histologies. Int J Cancer, 2011. 129(9): p. 2290-6.

[24] Rittmeyer, D. and A. Lorentz, Relationship between allergy and cancer: an overview. Int Arch Allergy Immunol, 2012. 159(3): p. 216-25.

[25] Brain tumour risk in relation to mobile telephone use: results of the INTERPHONE international case-control study. Int J Epidemiol, 2010. 39(3): p. 675-94.

[26] Baan, R., et al., Carcinogenicity of radiofrequency electromagnetic fields. Lancet Oncol, 2011. 12(7): p. 624-6.

[27] Hardell, L., et al., Case-control study of the association between malignant brain tumours diagnosed between 2007 and 2009 and mobile and cordless phone use. Int J Oncol, 2013. 43(6): p. 1833-45.

[28] Coureau, G., et al., Mobile phone use and brain tumours in the CERENAT case-control study. Occup Environ Med, 2014. 71(7): p. 514-22.

[29] Siegal, T., Clinical Relevance of Prognostic and Predictive Molecular Markers in Gliomas. Adv Tech Stand Neurosurg, 2016(43): p. 91-108.

[30] Vigneswaran, K., S. Neill, and C.G. Hadjipanayis, Beyond the World Health Organization grading of infiltrating gliomas: advances in the molecular genetics of glioma classification. Ann Transl Med, 2015. 3(7): p. 95.

[31] Turkalp, Z., J. Karamchandani, and S. Das, IDH mutation in glioma: new insights and promises for the future. JAMA Neurol, 2014. 71(10): p. 1319-25.

[32] Louis, D.N., et al., The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol, 2016. 131(6): p. 803-20.

[33] Hegi, M.E., et al., MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med, 2005. 352(10): p. 997-1003.

[34] Cankovic, M., et al., The role of MGMT testing in clinical practice: a report of the association for molecular pathology. J Mol Diagn, 2013. 15(5): p. 539-55.

[35] Binabaj, M.M., et al., The prognostic value of MGMT promoter methylation in glioblastoma: A meta-analysis of clinical trials. J Cell Physiol, 2018. 233(1): p. 378-386.

[36] Ostrom, Q.T., et al., The epidemiology of glioma in adults: a "state of the science" review. Neuro Oncol, 2014. 16(7): p. 896-913.

[37] Pekmezci, M., et al., Adult infiltrating gliomas with WHO 2016 integrated diagnosis: additional prognostic roles of ATRX and TERT. Acta Neuropathol, 2017. 133(6): p. 1001-1016.

[38] Thakkar, J.P., et al., Epidemiologic and molecular prognostic review of glioblastoma. Cancer Epidemiol Biomarkers Prev, 2014. 23(10): p. 1985-96.

[39] Park, S.H., et al., Molecular Testing of Brain Tumor. J Pathol Transl Med, 2017. 51(3): p. 205-223.

[40] Daumas-Duport, C., et al., Grading of astrocytomas. A simple and reproducible method. Cancer, 1988. 62(10): p. 2152-65.

[41] Louis, D.N., et al., The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol, 2007. 114(2): p. 97-109.

[42] Louis, D.N., et al., cIMPACT-NOW: a practical summary of diagnostic points from Round 1 updates. Brain Pathol, 2019. 29(4): p. 469-472.

[43] McLendon, R.E. and E.C. Halperin, Is the long-term survival of patients with intracranial glioblastoma multiforme overstated? Cancer, 2003. 98(8): p. 1745-8.

[44] Law, M., et al., Low-grade gliomas: dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging--prediction of patient clinical response. Radiology, 2006. 238(2): p. 658-67.

[45] Crocetti, E., et al., Epidemiology of glial and non-glial brain tumours in Europe. Eur J Cancer, 2012. 48(10): p. 1532-42.

[46] Kelly, P.J., et al., Imaging-based stereotaxic serial biopsies in untreated intracranial glial neoplasms. J Neurosurg, 1987. 66(6): p. 865-74.

[47] Perry, A., & Wesseling, P. (2016). Histologic classification of gliomas. Handbook of clinical neurology, 134, 71-95

[48] Pouratian, N., et al., Surgery Insight: the role of surgery in the management of low-grade gliomas. Nat Clin Pract Neurol, 2007. 3(11): p. 628-39.

[49] Forst, D.A., et al., Low-grade gliomas. Oncologist, 2014. 19(4): p. 403-13.

[50] Kanderi, T. and V. Gupta, Glioblastoma Multiforme, in StatPearls. 2020, StatPearls Publishing Copyright © 2020, StatPearls Publishing LLC.: Treasure Island (FL).

[51] Stupp, R., et al., Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med, 2005. 352(10): p. 987-96.

[52] Adams, H., et al., Adult cerebellar glioblastoma: understanding survival and prognostic factors using a population-based database from 1973 to 2009. World Neurosurg, 2013. 80(6): p. e237-43.

[53] Engelhard, H.H., et al., Clinical presentation, histology, and treatment in 430 patients with primary tumors of the spinal cord, spinal meninges, or cauda equina. J Neurosurg Spine, 2010. 13(1): p. 67-77.

[54] Upadhyay, N. and A.D. Waldman, Conventional MRI evaluation of gliomas. Br J Radiol, 2011. 84 Spec No 2(Spec Iss 2): p. S107-11.

[55] Han, W., et al., Deep Transfer Learning and Radiomics Feature Prediction of Survival of Patients with High-Grade Gliomas. AJNR Am J Neuroradiol, 2020. 41(1): p. 40-48.

[56] Pope, W.B., et al., MR imaging correlates of survival in patients with high-grade gliomas. AJNR Am J Neuroradiol, 2005. 26(10): p. 2466-74.

[57] Tietze, A., et al., Noninvasive assessment of isocitrate dehydrogenase mutation status in cerebral gliomas by magnetic resonance spectroscopy in a clinical setting. J Neurosurg, 2018. 128(2): p. 391-398.

[58] Latini, F., et al., A novel radiological classification system for cerebral gliomas: The Brain-Grid. PLoS One, 2019. 14(1): p. e0211243.

[59] Samala, R.K., et al., Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms. Phys Med Biol, 2017. 62(23): p. 8894-8908.

[60] Banerjee, I., et al., Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma. Comput Med Imaging Graph, 2018. 65: p. 167-175.

[61] Krizhevsky, A., I. Sutskever, and G.E. Hinton, ImageNet classification with deep convolutional neural networks, in Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1. 2012, Curran Associates Inc.: Lake Tahoe, Nevada. p. 1097–1105.

[62] Yang, Y., et al., Glioma Grading on Conventional MR Images: A Deep Learning Study With Transfer Learning. Front Neurosci, 2018. 12: p. 804.

[63] Lotan, E., et al., State of the Art: Machine Learning Applications in Glioma Imaging. AJR Am J Roentgenol, 2019. 212(1): p. 26-37.

[64] Lambin, P., et al., Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer, 2012. 48(4): p. 441-6.

[65] Yip, S.S. and H.J. Aerts, Applications and limitations of radiomics. Phys Med Biol, 2016. 61(13): p. R150-66.

[66] Prasanna, P., et al., Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: Preliminary findings. Eur Radiol, 2017. 27(10): p. 4188-4197.

[67] Zhou, M., et al., Identifying spatial imaging biomarkers of glioblastoma multiforme for survival group prediction. J Magn Reson Imaging, 2017. 46(1): p. 115-123.

[68] Lao, J., et al., A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. Sci Rep, 2017. 7(1): p. 10353.

[69] Zacharaki, E.I., et al., Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med, 2009. 62(6): p. 1609-18.

[70] Aerts, H.J., et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun, 2014. 5: p. 4006.

[71] Kickingereder, P., et al., Large-scale Radiomic Profiling of Recurrent Glioblastoma Identifies an Imaging Predictor for Stratifying Anti-Angiogenic Treatment Response. Clin Cancer Res, 2016. 22(23): p. 5765-5771.

[72] Peeken, J.C., et al., Semantic imaging features predict disease progression and survival in glioblastoma multiforme patients. Strahlenther Onkol, 2018. 194(6): p. 580-590.

[73] Zwanenburg, A., et al., The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology, 2020. 295(2): p. 328-338.

[74] Parmar, C., et al., Machine Learning methods for Quantitative Radiomic Biomarkers. Sci Rep, 2015. 5: p. 13087.

[75] Peeken, J.C., et al., CT-based radiomic features predict tumor grading and have prognostic value in patients with soft tissue sarcomas treated with neoadjuvant radiation therapy. Radiother Oncol, 2019. 135: p. 187-196.

[76] Peeken, J.C., et al., Treatment-related features improve machine learning prediction of prognosis in soft tissue sarcoma patients. Strahlenther Onkol, 2018. 194(9): p. 824-834.

[77] Deist, T.M., et al., Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers. Med Phys, 2018. 45(7): p. 3449-3459.

[78] Lee, J., et al., Texture Feature Ratios from Relative CBV Maps of Perfusion MRI Are Associated with Patient Survival in Glioblastoma. AJNR Am J Neuroradiol, 2016. 37(1): p. 37-43.

[79] Ryu, Y.J., et al., Glioma: application of whole-tumor texture analysis of diffusion-weighted imaging for the evaluation of tumor heterogeneity. PLoS One, 2014. 9(9): p. e108335.

[80] Beig, N., et al., Radiogenomic analysis of hypoxia pathway is predictive of overall survival in Glioblastoma. Scientific Reports, 2018. 8(1): p. 7.

[81] Gutman, D.A., et al., Somatic mutations associated with MRI-derived volumetric features in glioblastoma. Neuroradiology, 2015. 57(12): p. 1227-37.

[82] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. T he Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository , Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057.

[83] Andrew Beers, Elizabeth Gerstner, Bruce Rosen, David Clunie, Steve Pieper, Andrey Fedorov, Jayashree Kalpathy-Cramer. (2018) DICOM-SEG Conversions for TCGA-LGG and TCGA-GBM Segmentation Datasets. The Cancer Imaging Archive. https://doi.org/10.7937/TCIA.2018.ow6ce3ml

[84] Pedano, N., Flanders, A. E., Scarpace, L., Mikkelsen, T., Eschbacher, J. M., Hermes, B., … Ostrom, Q. (2016). Radiology Data from The Cancer Genome Atlas Low Grade Glioma [TCGA-LGG] collection. The Cancer Imaging Archive. http://doi.org/10.7937/K9/TCIA.2016.L4LTD3TK

Serment d' H ippocrate

Au moment d'être admis à devenir membre de la profession médicale, je m'engage solennellement à consacrer ma vie au service de l'humanité.

- *Je traiterai mes maîtres avec le respect et la reconnaissance qui leur sont dus.*
- *Je pratiquerai ma profession avec conscience et dignité. La santé de mes malades sera mon premier but.*
- *Je ne trahirai pas les secrets qui me seront confiés.*
- *Je maintiendrai par tous les moyens en mon pouvoir l'honneur et les nobles traditions de la profession médicale.*
- *Les médecins seront mes frères.*
- *Aucune considération de religion, de nationalité, de race, aucune considération politique et sociale ne s'interposera entre mon devoir et mon patient.*
- *Je maintiendrai le respect de la vie humaine dès la conception.*
- *Même sous la menace, je n'userai pas de mes connaissances médicales d'une façon contraire aux lois de l'humanité.*
- *Je m'y engage librement et sur mon honneur.*

قسم أبقراط

بسم الله الرحمان الرحيم

أقسم بالله العظيم

في هذه اللحظة التي يتم فيها قبولي عضوا في المهنة الطبية أتعهد علانية:

- بأن أكرس حياتي لخدمة الإنسانية.
- وأن أحترم أساتذتي وأعترف لهم بالجميل الذي يستحقونه.
- وأن أمارس مهنتي بوازع من ضميري وشرفي جاعلا صحة مريضي هدفي الأول.
- وأن لا أفشي الأسرار المعهودة إلي.
- وأن أحافظ بكل ما لدي من وسائل على الشرف والتقاليد النبيلة لمهنة الطب.
- وأن أعتبر سائر الأطباء إخوة لي.
- وأن أقوم بواجبي نحو مرضاي بدون أي اعتبار ديني أو وطني أو عرقي أو سياسي أو اجتماعي.
- وأن أحافظ بكل حزم على احترام الحياة الإنسانية منذ نشأتها.
- وأن لا أستعمل معلوماتي الطبية بطريق يضر بحقوق الإنسان مهما لاقيت من تهديد.
- بكل هذا أتعهد عن كامل اختيار ومقسما بالله.

والله على ما أقول شهيد.

المملكة المغربية
جامعة محمد الخامس بالرباط
كلية الطب والصيدلة
الرباط

جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

سنـة : 2020 أطروحة رقم:

تصنيف صور الأورام الدبقية باستخدام التعلم الآلي: مقاربة بالراديوميكس

أطروحة

قدمت ونوقشت علانية يوم : / /2020

من طرف
السيد محمد صبحي جبل
المزداد في 21 ماي 1991 بالرباط

لـنـيـل شـهـادة

دكـتـور فـي الطب

الكلمات الأساسية : الورم الدبقي؛ التصوير العصبي؛ الراديوميكس؛ الذكاء الاصطناعي

أعضاء لجنة التحكيم:

السيد عز الدين إبراهيمي أستاذ في البيوتكنولوجيا الطبية	**رئيس**
السيدة محجوبة بوطربوش أستاذة في تشريح الأعصاب	**مشرف**
السيد أمين بن قبو أستاذ في الجراحة العامة	**عضو**
السيد محمد أنس مجبار أستاذ في الجراحة العامة	**عضو**