جامعة أبو بكر بلقايد – تلمسان

Université Abou Bakr Belkaïd de Tlemcen

Faculté de Technologie

Département de Génie Biomédical

MEMOIRE DE PROJET DE FIN D'ETUDES

pour l'obtention du Diplôme de

MASTER en GENIE BIOMEDICAL

Spécialité : **Informatique Biomédicale**

présenté par : GHEZALI Waffa

Automated histopathological diagnosis of childhood medulloblastoma using Convolutional Neural Networks

Soutenu le 01 juillet 2021 devant le Jury

Mme	KAZI TANI Lamia Fatiha	*MCB*	Université de Tlemcen	Présidente
Mme	SETTOUTI Nesma	*MCA*	Université de Tlemcen	Encadreur
M.	BEHADADA Omar	*MCB*	Université de Tlemcen	Examinateur

Année universitaire 2020-2021

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ ABOU BEKR BELKAID
FACULTÉ DE TECHNOLOGIE
DÉPARTEMENT DE GÉNIE BIOMÉDICAL

MÉMOIRE DE FIN D'ÉTUDES

pour obtenir le grade de

MASTER EN GÉNIE BIOMÉDICAL

Spécialité : **Informatique Biomédicale**

présenté et soutenu publiquement
par

GHEZALI Waffa

le 01 Juillet 2021

Titre:

Automated histopathological diagnosis of childhood medulloblastoma using convolutional neural networks

Jury

Présidente du jury.	Dr. DALI YOUCEF Lamia Fatiha,	MCB UABB Tlemcen
Examinateur.	Dr. BEHADADA Omar,	MCB UABB Tlemcen
Directrice de mémoire.	Dr. SETTOUTI Nesma,	MCA UABB Tlemcen

I dedicate this work:

To the dearest and the jewel of my life, my parents who have been waiting for this day with great anticipation,

To my beloved brothers kheir Eddine, Abd Elaziz, Nassib El Mokhtar,

To Ms. Settouti Nesma my great and wonderful teacher,

To my cousins Asma, Rayane Hiba Elrahmane, Hadjer, Fatima, Imene, Khansaa,

To my aunts and my uncles,

To all family of Ghezali and Seksaf,

To whom I consider Expensive, not just friends Malik, Saadia, Ikram, Chahra, Marwa, Setti, Ibtissem,

To everyone who has supported me in my field of study,

A special dedication to Mr. Gueddim Larbi and my grandparent's soul.

Waffa

Acknowledgments

This document presents the work carried out as part of my final Master's degree project at the Department of Biomedical Engineering, Faculty of Technology, University of Tlemcen.

My thanks go first to ALMIGHTY GOD for the willpower, health and patience He has given me over all these long years.

I would to thank my supervisor **Ms. SETTOUTI Nesma** Associate Professor at the Biomedical Engineering Laboratory (GBM), University of Tlemcen, for her support, her advice and, her supervision which were beneficial to me in carrying out this work.

I would then like to express my sincere gratitude to **Ms. DALI YOUCEF Lamia Fatiha** and **M. BEHADADA Omar** for honoring us by their presence as Defense Jury, could you find here the expression of my gratitude and my respect.

A special thanks to **M. BECHAR Mohammed El Amine** for his advice and help in this project.

I also want to thank my parents for all the moral, financial, psychological support and especially for their love.

Finally, my thanks go to all those who have contributed to the accomplishment of this modest work.

Résumé

La classification automatisée a envahi le monde de la santé, mais certaines maladies sont encore diagnostiquées manuellement comme c'est le cas du cancer du médulloblastome (MB) qui est une tumeur maligne à haut risque du système nerveux central (SNC). Cette tumeur a de nombreux types et est diagnostiquée par biopsie en examinant des images histologiques, ce qui demande beaucoup d'efforts et de temps. Dans notre projet de fin d'études de Master, nous proposons d'effectuer une classification automatisée du médulloblastome de l'enfance basée sur les réseaux de neurones convolutifs CNN en exploitant les connaissances apprises à travers différentes bases (medical PatchCamelyon (PCam) ou Nature ImageNet) avec apprentissage par transfert. Une étude comparative est réalisée en suivant deux stratégies : extraction de caractéristiques du réseau et apprentissage partiel du réseau, nous utilisons les architectures les plus courantes à savoir VGG-16, VGG-19, ResNet-50, Inception V3, CancerNet. Les expérimentations réalisées prouvent que la similitude joue un rôle important lorsqu'il s'agit d'approche d'extraction de caractéristiques du réseau ou d'apprentissage partiel du réseau. En effet, lorsqu'il y a une forte similarité entre les deux données, l'extraction de caractéristiques du réseau est la meilleure stratégie, sinon un réglage fin partiel est plus adapté. Dans notre cas d'application, les résultats démontrent que le processus d'apprentissage partiel du réseau en utilisant Inception V3 pré-entraîné sur ImageNet a obtenu les meilleures performances avec une précision de 99,16%.

Mots clés

Medulloblastome, réseau de neurone convolutifs, apprentissage par transfert, apprentissage partiel, extraction de caractéristiques profondes, histologie, classification.

Abstract

The automated classification invaded the world of health, but some diseases are still diagnosed manually as is the case of medulloblastoma (MB) cancer, which is a high-risk malignant tumor in the Central Nervous System (CNS). This tumor has many types and is diagnosed by biopsy by examining histological images, which takes a lot of effort and time. In our master's project, we propose to perform an automated Childhood medulloblastoma classification based on Convolutional Neural Networks (CNN) by exploiting the knowledge learned through different bases (medical PatchCamelyon (PCam) or Nature ImageNet) with transfer learning. A comparative study is carried out by following two strategies: deep feature extractor and partial fine-tuning, we applied the most popular architectures namely VGG-16, VGG-19, ResNet-50, Inception V3, CancerNet. Experiments prove that there is an important role of similarity when dealing with deep features extractor or partial fine-tuning. Indeed, when there is a high similarity between the two data, the deep features extractor is the best, otherwise, partial fine-tuning is more suitable. In our case application, results demonstrate that the process of partial fine-tuning using inception V3 pre-trained on ImageNet achieved the best result with an accuracy of 99.16%.

Keywords

Medulloblastoma, convolutional neural network CNN, transfer learning, partial fine-tuning, deep features extractor, histology, classification.

ملخص

بالرغم من ان التصنيف الآلي غزا عالم الصحة، إلا انه لا تزال بعض الأمراض تُشخص يدويًا كما هو الحال في سرطان الورم الأرومي النخاعي والذي يعتبر ورم خبيث شديد الخطورة يكون على مستوى الجهاز العصبي المركزي. هذا الورم له أنواع عديدة ويتم تشخيصه بالخزعة عن طريق فحص الصور النسيجية، الأمر الذي يتطلب الكثير من الجهد والوقت. في مشروع الماجستير الخاص بنا، نقترح إجراء تصنيف آلي للورم الأرومي النخاعي في الطفولة بناءً على الشبكات العصبية التلافيفية من خلال استغلال المعرفة المكتسبة من خلال قواعد الطبيعية) باستعمال التعلم بالنقل. يتم إجراء دراسة ImageNet الطبية أو PCam) مختلفة مقارنة باتباع استراتيجيتين: استخراج خصائص الشبكة والتعلم الجزئي للشبكة، نستخدم أكثر و inception V3 و ResNet 50 و VGG19 و VGG16 البنى شعبية وهي CancerNet .

تظهر النتائج أن التشابه بين مجموعة البيانات التي تم اختبارها مسبقًا والبيانات التي لدينا دور مهم في كل من هذين النهجين. في مستخرج الميزات العميقة، عندما يكون التشابه مرتفعًا جدًا، تكون النتيجة أفضل. في الضبط الجزئي، عندما يكون هناك تشابه ضعيف، تكون النتائج جيدة. في الأخيرً، اعتمدنا عملية الضبط الجزئي باستخدام inception V3 التي تم اختبارها مسبقًا على ImageNet لأنها تحقق أفضل نتيجة بدقة 99.16٪.

كلمات البحث

الورم الأرومي النخاعي، شبكة الخلايا العصبية التلافيفية، التعلم بالنقل، التعلم الجزئي، استخراج السمات العميقة، علم الأنسجة، التصنيف.

Contents

List of Figures

List of Tables

Glossary

ADC: apparent diffusion coefficient
B-MB: Classic biphasic medulloblastoma
CCBV: corrected cerebral blood volume
CMB: Medulloblastoma classic
CNN: Convolutional Neural Network
CNS: Central Nervous System
CONV: Convolutional layers
CSF : Cerebro-Spinal Fluid
DMG: midline glioma of points
DN: Desmoplastic/nodular medulloblastoma
DSC: dynamic susceptibility contrast imaging
DWI: diffusion-weighted imaging
EP: ependymoma
ESD: ensemble subspace discriminant
GLCM: gray level co-occurrence matrix
GNB: Ganglioneuroblastoma
GNRC: Guwahati Neurological Research Center
GRLN: gray level path length matrix
H&E: hematoxylin and eosin
HOG: histogram of oriented gradient
K2: leakage coefficient
KNN: K-Nearest Neighbor
LBP: local binary model
LCA: Large cell/anaplastic medulloblastoma
LD: linear discriminant
LDA: linear discriminant analysis
LR: logistic regression
MB : medulloblastoma
MBEN: Medulloblastoma with extensive nodularity
MRI: Magnetic Resonance Imaging
NOS: Medulloblastoma non-specified
PA: pilocytic astrocytoma
PCA: principal component analysis
PCam: PatchCamelyon
QD: quadratic discriminant
RELU: Rectified Linear Units
ROI: region of interest
SHH: Sonic hedgehog

SVM: support vector machine
UCBV: uncorrected cerebral blood volume
WHO : world health organization
WNT: Wingless

Introduction

Medulloblastoma is one of the most common cancers of central nervous system tumors which have several types. The diagnostic of this malignancy tumor is based on some examinations and much more on the biopsy through histopathological images. However, this diagnosis requires hard and tedious work, which makes the diagnosis burdensome for experts. The solution to this problem is to propose an automated histopathological diagnosis of childhood medulloblastoma using machine learning tools.

One of the new techniques of machine learning is deep learning. Lately, with the development of machine learning methods and the contribution of Deep Learning this has allowed great advances in computer vision, therefore, it is important to experiment and test the application of methods based on convolutional neural networks. CNNs is a neural network based on convolutional layers and is formed by a succession of blocks. There are several architectures with different number of blocks that varied from one to another.

Learning CNNs requires a large database with thousands and millions of data, but we can use it with small data by resorting to transfer learning. It has three approaches, two of them (partial fine-tuning and deep features extractor) are designated for the small data. These transfer learning approaches helps to reuse image representations learned from a source task and a dataset with a large amount tagged into a second dataset which is small, like our case, in which we have a database (Childhood Medulloblastoma) which is relatively small (as is the case in most medical data).

In this master project, we investigate which approach and which architecture and knowledge (through images of nature / medical) will be the most relevant for the recognition of MB. A comparative study is conducted with some of the most famous architectures like VGG-16, VGG-19, ResNet-50, Inception V3. These models are pre-trained on ImageNet (straight-forward natural-image classification datasets) and the PCam (histological images benchmark) using partial fine-tuning and deep features extractor strategies. Through experimentations, we will demonstrate what strategy and model are the most suitable for our case application.

Our work is divided into three chapters :

- In the first one, we will talk about the medulloblastoma, its symptoms, its types, the diagnostics...

- In the second chapter, we will cite the different recent works related to medulloblastoma classification in the machine learning domain.
- In the last chapter, we will define our process and cite its details, its results.

Chapter 1

Medulloblastoma - Overview

1 Definition of Medulloblastoma "MB"

Medulloblastoma is an embryonic cancerous primary neuroectodermal [1] malignancy of the fourth degree according to the world health organization (WHO) [21] of the brain that develops in the middle of the cerebellum near the fourth ventricle at the level of the vermis as designated in Figure 1.1.

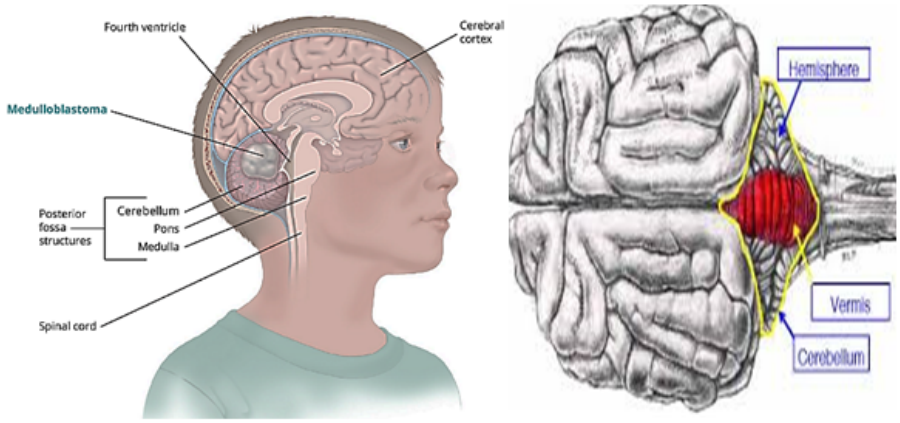

Figure 1.1: Anatomy of the inside of the brain [1].

It is a rapidly growing invasive tumor and can metastasize throughout the central nervous system via the Cerebro-Spinal Fluid (CSF). It is the most common and best-known accounting for 20% of Central Nervous System tumors and half of the posterior fossa tumors [2]. It is more common in children under 16 years and can be seen in older adolescents, less often, in adults.

In 1925, the term "medulloblastoma" was first described by "Percival Bailey", who was working in conjunction with "Harvey Cushing" [22].

The cause of medulloblastoma is not known, the hypothesis is that it could be the result of an error in the early development of brain cells, but nothing is

[1] malignant tumors thought to arise from the central or peripheral nervous system.

certain.

- MB signs and symptoms are:
  - Headache.
  - Nausea and vomiting, often worse in the morning.
  - Fatigue or changes in activity levels.
  - Dizziness.
  - Loss of balance, clumsiness.
  - Difficulty writing.
  - Vision impairment.
- If the tumor has spread to the spinal cord:
  - Back pain,
  - Difficulty walking,
  - Problems with urination or changes in bowel function.

2 Diagnosis and treatment of medulloblastoma

The diagnosis of this tumor is based on medical history, neurological examinations, Magnetic Resonance Imaging (MRI), and much more on the biopsy which is an integral part of the prognosis because it allows us to know the different characteristics of the tumor such as dimensions, location, the form, etc.

The treatment of medulloblastoma includes surgery, chemotherapy, and radiotherapy but long-term monitoring has shown that such treatment leads to mainly neurocognitive impairment, neuropathy, endocrinopathy, delayed bone growth, impaired motor function, hearing loss, and secondary malignancy.

These effects are dose-related to radiation therapy and the age of the patient (overall survival is 5 years with a percentage of 75%) [2].

3 Risk stratification of medulloblastoma

The risks of medulloblastoma are broken down into two categories, average (standard) and high risk [23].

- The MB is considered average risk if all the following conditions are true:
  - The child is over 3 years old.
  - The location of tumor at the back of the brain.
  - The cancer has not spread to other parts of the body.
  - Do not have molecular features.

  - The rest of the tumor after surgery is less than 1.5 cm^3.

- The MB is considered high risk if one of the following conditions is met:
  - The age of child is less than three years.
  - The tumor location is not at the very back of the brain.
  - The cancer has spread to other parts of the body.
  - Molecular features will be found.
  - The rest of the tumor after surgery is more than 1.5 cm^3.

4 Classification and types of medulloblastoma

Medulloblastoma is classify and describe based on certain characteristics such as histology, molecular features, and metastasis as presented in Figure 1.2.

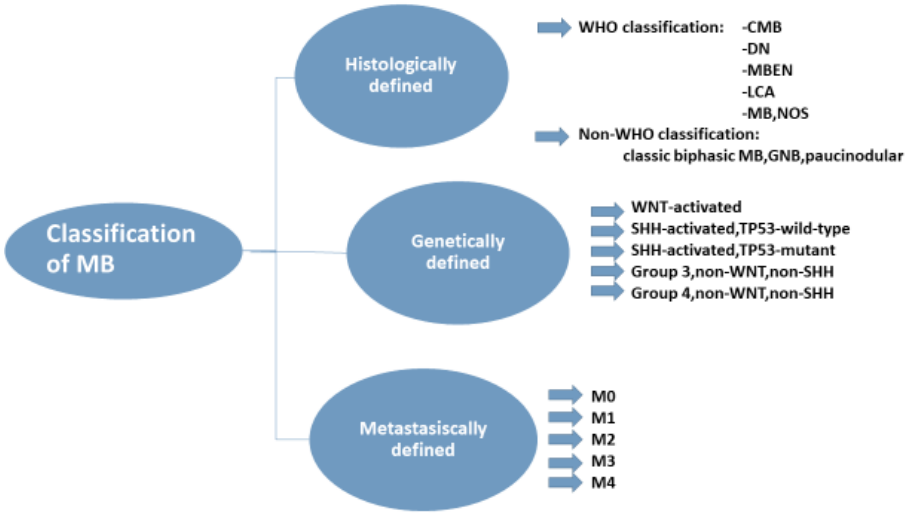

Figure 1.2: The classification and subtypes of medulloblastoma.

4.1 Histological classification

The histological classification of medulloblastoma using microscopic features (size, shape, ...) shows that there are several types, some are defined by the WHO and others not [3].

The sub-groups defined by the WHO are:

- Medulloblastoma classic (CMB): it is the most common of MB accounting for 72% [2]. It is undifferentiated cells with a hyperchromatic nucleus, (Homer wright rosette) suggesting neuronal differentiation, blue tumor as it is represented in Figure 1.3. CMB is characterized by:

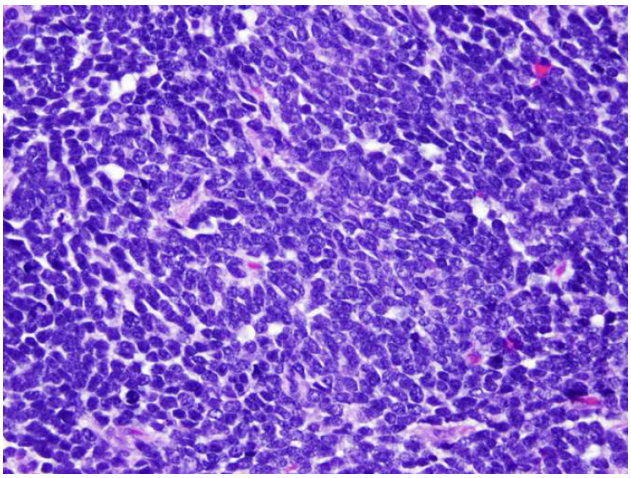

Figure 1.3: Classic MB [2].

  - Diffuse masses of small undifferentiated oval or round cells.
  - Some medulloblastomas exhibit neuronal, glial, and other differentiation.
  - Nuclear molding due to the compactness inside the cell.
  - Neuronal differentiation is manifested by the formation of neuropils and rosettes.
  - Rosettes are groups of tumor cells arranged in a circle around a fibrillar center.
  - Mature neurons can also be found, but infrequently. Glial differentiation in some tumors results in GFAP-positive cells. There may also be differentiation along the oligodendroglial or ependymal lines

- Desmoplastic/nodular medulloblastoma (DN): Classic + Glomeruli (pale collagen island): The least aggressive. Pale nodules are made up of uniform, round, or spindle-shaped neural cells that are not as mitotically active as the surrounding, darker tumor. They are present in a fine fibrillar base poor in reticulin. Figure 1.4 represent the DN case.

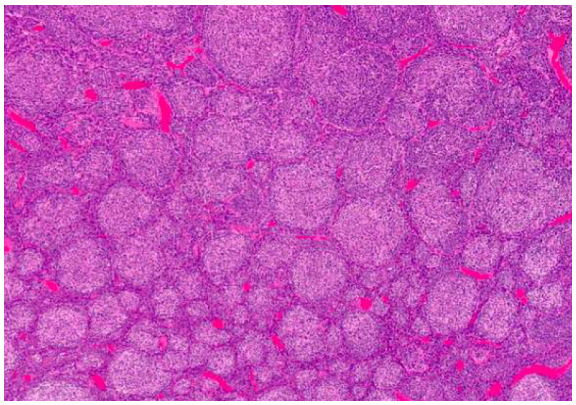

Figure 1.4: Desmoplastic/nodular MB [3].

- Medulloblastoma with extensive nodularity (MBEN): it is tumor-specific for DN that shows more nodular regions within the tissue.it characterized by nodules that tend to be irregular and merge into numerous sets. The MBEN case is represented in Figure 1.5.

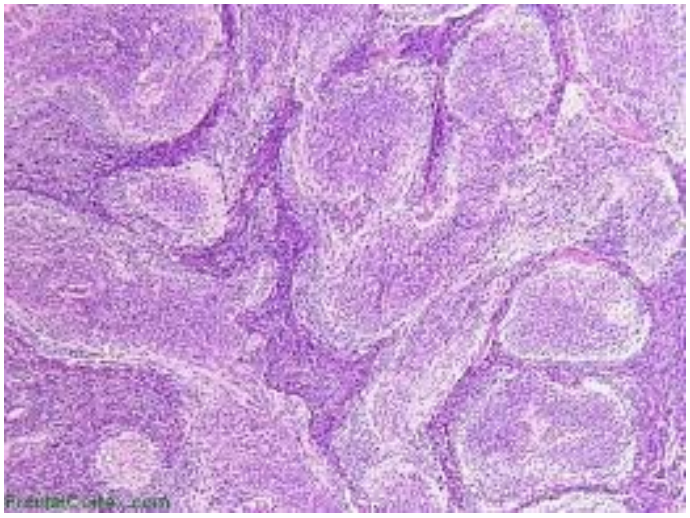

Figure 1.5: MB with extensive nodularity [2].

- Large cell/anaplastic medulloblastoma (LCA): early seed, high mitosis: poor prognosis It has tumor cells with large nuclei with anaplastic which is made up of cells that are irregularly shaped. LCA is composed of very atypical layers of cells. It is characterized by a focus of necrosis is present just above the center of Figure 1.6, and an irregular hyperchromatic nucleus with a nuclear cast and little cytoplasm.

  The anaplasia was defined by increased cell size, increased cytologic pleomorphism, necessary envelopment molding, frequent mitotic activity, and frequent apoptotic bodies.

- Medulloblastoma non-specified (NOS).

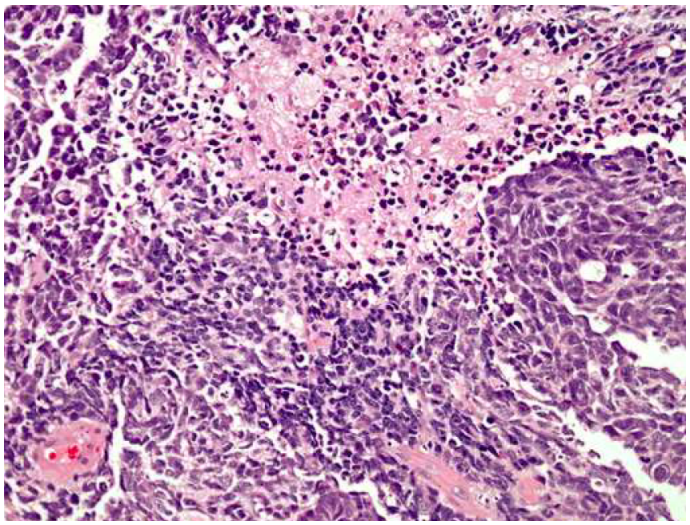

Figure 1.6: Large cell/anaplastic MB [2].

The sub-groups not defined by the WHO are:

- Classic biphasic medulloblastoma (B-MB): The term biphasic medulloblastoma is coined by Ellison et al. in 2007 [24]. Classic biphasic MB is characterized by an alternating sheet of tumor cells with nodules of neurolytic differentiation, these nodules are more irregular in their contour than those encountered in the desmoplastic/nodular variant.

  The difference between the DN variant and B-Mb is that in the DN variant, the nodules are outlined by reticulin but in M-MB are not outlined by reticulin.

- Ganglioneuroblastoma (GNB): are tumors that develop in nerve tissue from ganglia (a mass of nerve cells). Some intermediate tumors, ones that have both malignant and benign cells, or both cancerous and noncancerous cells, respectively. These tumors are rare and primarily occur in children under the age of 5 [25].

- Paucinodular: The paucinodular D/N medulloblastoma displays scattered small nodules amid widespread desmoplasia. It contained just a few synaptophysin-immunopositive reticulin-negative nodules that showed virtually no decrease in nuclear: cytoplasmic ratio.

4.2 Genetical classification or Molecular Subgroups:

Molecular subgroups are genetically derived from a study and the understanding of the three syndromes associated with medulloblastoma: Gorlin syndrome (mutation in the PTCH1, PTCH2, or SUFU), Turcot syndrome (mutation in the adenomatous Polyposis Coli APC), and Fraumeni syndrome (germline P53 mutation) [2,26]; representing:

- Wingless (WNT) activated:

List 10-15 % of MB [2]. It has a classic morphology and tends to occur along the mid-line of the cerebellum. It characterized by:

  - Expression of WNT pathway genes contains a mutation in exon 3 of the CTTNB1 gene.
  - Exhibit loss/ partial loss of chromosome 6.

  The APC can be identified by including CTTNB1, TP53, SMARCA4, KMT2D, or DDX3X mutation.

- Sonic hedgehog (SHH) activated:

Is counted for 30 % of MB [2] tend to be found in the cerebellar hemisphere, but also can arise in the cerebellar vermis. SHH MB is characterized by activation of SHH pathway transcriptional programs and recurrent mutations in SHH pathway genes including PTCH1, SMO, and SUFU.

- Group 3 (non-WNT, non-SHH) / Group 4 (non-WNT, non-SHH) :

Both group 3 and group 4 represent in the mid-line filling the fourth ventricle. They are tumors associated with metastasis, the first has a poor prognosis, and the second intermediate.
Group 3 are tumors exceedingly rare which represent in the pediatric age spectrum.
Group 4 are typical tumors that present in older children.

4.3 Metastasis classification

The classification of MB based on metastasis or spread of disease shows the following stages [3]:

- M0: Tumor without spreading the disease.
- M1: Tumor cells found in the CSF.
- M2: the presence of evidence of disease spread in the brain (intra-cranial).
- M3: The spread of the tumor to the spine.
- M4: The tumor has spread outside the CNS. Common sites of metastasis include the bones, lungs, and liver.

5 Aim of this work

Identification of childhood medulloblastoma and its appropriate subtype from a biopsied tissue sample of the child's tumor is integral to prognosis. Survival rates can increase with early diagnosis. Indeed, the extent of patient survival is highly dependent on the prior identification and treatment of pediatric medulloblastoma. Diagnosis depends on qualitative visual examination of histological

slides of biopsy samples by clinical experts.

Algorithms and computer techniques have been used to assist pathologists and physicians with various methodologies. Clinical diagnosis was studied to better understand malignancy and the phases a patient undergoes when diagnosed with infantile medulloblastoma.

The sensitivity of the treatment is not systematic, and the identification of this tumor requires a great deal of work and image analysis time for each patient.

In this master project, we propose a decision support system for automated classification of normal and abnormal biopsy specimens of childhood medulloblastoma tumors, the most common childhood brain tumor.

The proposed application is based on the classification of histopathological images of pediatric medulloblastoma through deep learning methods on patient data collected from medical centers around the world from the IEEE Dataset repository [27].

6 Conclusion

MB is a tumor that contains several subgroups which are defined based on different characteristics. in our project, we will propose an automated classification of medulloblastoma between normal and abnormal biopsy samples after recognizing the studies and work that are done on medulloblastoma which will be cited in the following chapters.

Chapter 2

Related works

In this chapter, we will review the related works that have been done on the automated classification of medulloblastoma. We will classify these works into two categories according to the type of images used (MRI and histological) at the end we will synthesize these different works.

1 Associated work

1.1 Approaches based MRI images

In 2020, a study based on machine learning and images obtained by MRI is conducted to compare and classify three brain tumors which are: medulloblastoma, ependymoma (EP), and astrocytoma to obtain at the end the grade of the tumor (low or high) or type of tumor. the process of this study is represented in Figure2.1.

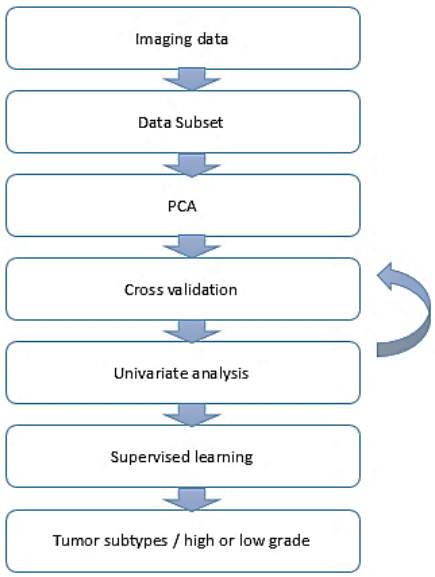

Figure 2.1: Data processing on MRI images [4].

Grist et al. [4] worked with the hypothesis that "the combination of perfusion data from the region of interest (ROI), the apparent diffusion coefficient (ADC) and whole brain, provided considerable precision in distinguishing tumor type or their grade". This study was carried out on 49 patients from four clinical sites in the U.K with medulloblastoma high grade and other tumors low grade with a performed protocol of MRI at 3 or 1.5T including T1-weighted, T2-weighted, T2-FLAIR, and T1-post contrast. Authors used features like tumor volume (in cm, calculated from the masking T2-ROI), ADC (calculated from diffusion-weighted imaging), corrected cerebral blood volume (CCBV), uncorrected cerebral blood volume (UCBV), leakage coefficient (K2). These last three features were extracted using one of the advanced techniques of MRI which is dynamic susceptibility contrast imaging (DSC), the other technique is diffusion-weighted imaging (DWI) which was used to remove the signal from the static water compartment in the brain. All these features were reduced by using principal component analysis (PCA) and its tested by using a Shapiro-milk test under R and ANOVA/KRUSKAL-Wallace and Tukey for the evaluation. The authors used two oversampling methods for increasing data size: data replication and SMOTE.

The experimentation analysis was carried out under MATLAB and the training was done under ORANGE toolbox on python using five classifiers AdaBoost, Random Forest, K-Nearest Neighbor (KNN), and neuron network with a single layer. The F-statistic (sensibility and specificity) and precision were calculated to evaluate the classifiers after a three-fold cross-validation stratification. Between each fold the univariate analysis was applied. They concluded that the combination between the classifier AdaBoost and univariate analysis was the best with a high accuracy.

In the same year, another study also based on MRI images was carried out by Quon et al. [28] for the detection and classification of four tumors of the posterior fossa: midline glioma of points (DMG), MB, pilocytic astrocytoma (PA), EP using deep learning with ResNext-50-32-4d architecture. This study was carried out on 805 patients from five different institutional with a performed protocol of MRI at 3 or 1.5 including T2 FSE, T2 PROPELLER, T2 BLADE, T2 drive sense, T1 MPRAGE, T1 BRAVO, T1 fast-spoiled gradient recalled, T1 spoiled gradient-echo and T1 spin-echo. The comparison of the performance of this model and the average of radiologist prove that the model is the performed because of their high accuracy and f-score than accuracy biologist.

1.2 Approaches based histological images

In 2015, Cruz-Roa et al. [5] proposed an automated classification of anaplastic and non-anaplastic MB using the process shown in Figure 2.2.

The data used are images from WSI. The extraction of the characteristics was done using the two Convolutional Neuronal Networks architectures, VGG-16 and IBCa pre-trained models on both of natural and histological images. Then the use of the softmax to make the classification. Accuracy, sensitivity, and speci-

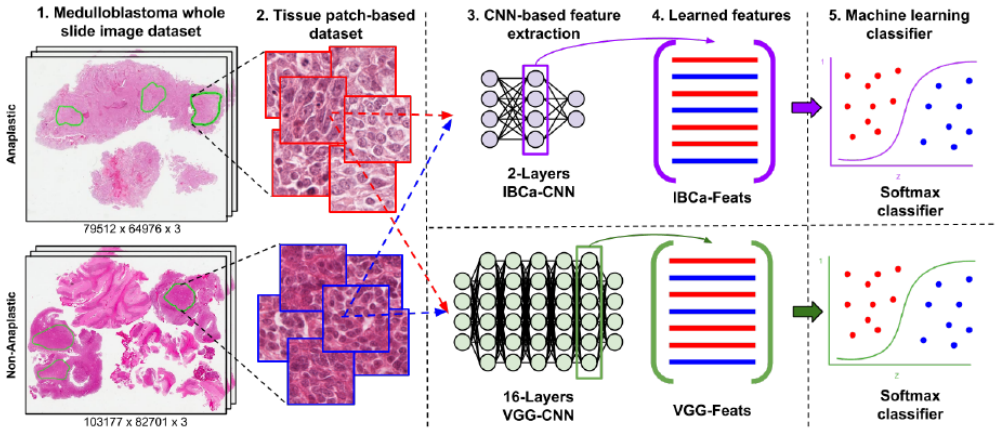

Figure 2.2: The work-flow process adopted in [5].

ficity are calculated for the measurement of classifier evaluation. The use of pre-trained models on histological images gives the best result. The best performance, whatever with color or grayscale images or with pre-trained model on natural or histological images, is achieved by IBCa architecture.

In 2018, Das et al. [6] have done a study about the automated classification of CMB based on texture features with normal and abnormal biopsy images going through four phases represented in Figure 2.3, which are: Database Generation, Pre-processing, features extraction, and classification.

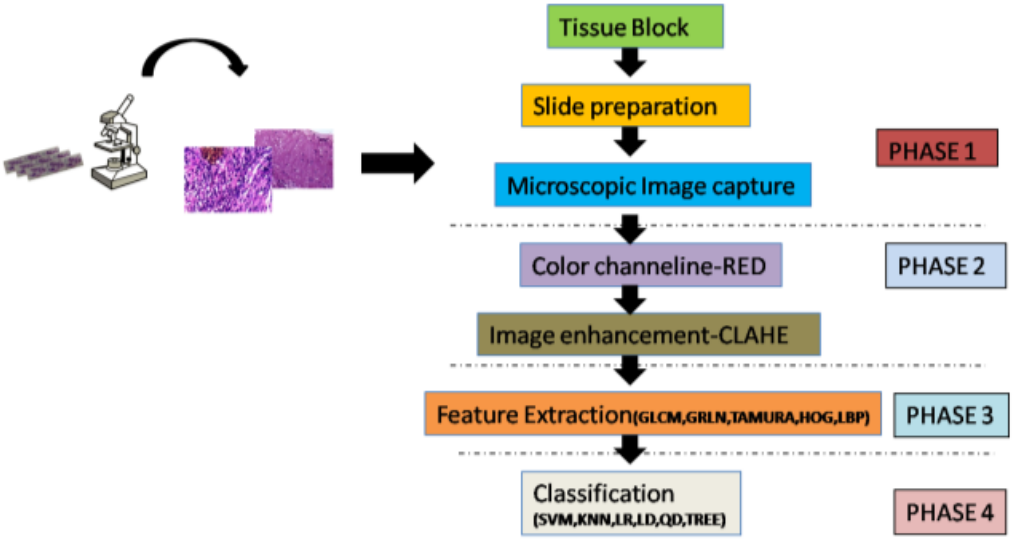

Figure 2.3: The diagram of the study [6] using texture features.

The first phase contains three sub-steps, the first is the tissue blocks which were obtained from GMCH, Guwahati. in the second step, the slide preparation

was carried out at Ayursundra Healthcare Pvt by staining the slides with hematoxylin and eosin (H&E) to mark the cytoplasm in pink and the nucleus in blue. In the last sub-step, the observation of the slides under a microscope was done in the neurological research center of Guwahati (GNRC) to obtain 80 normal and other abnormal images with a resolution of 512348 and 10 magnification. The second phase breaks down into two sub-stages.

The first is the choice of the red channel after training all standard and available canalizations. This channel was chosen because it clearly distinguishes the nucleus and the cytoplasm. The second is image enhancement to have colors with uniform distribution, adaptive histogram equalization limited to contrast CLAHE was chosen for enhancement after image conversion to gray-scale. The five characteristics were chosen in the third phase, namely: the gray level co-occurrence matrix (GLCM) (describes the probability of the passage in a certain direction of a pixel i to j), the matrix of the gray level path length (GRLN) (gives the connected length in a defined direction of a particular pixel), the histogram of oriented gradient (HOG) (the characterization of the appearance and shape of a local object), the local binary model (LBP) (the encoding of the information of local texture), and three characteristics of Tamura (coarseness, contrast, regularity). The last phase is a classification under MATLAB 2016b with six classifiers. The Tree-based classification is applied to train the model, logistic regression (LR) to fit data model using linear regression, linear discriminant (LD), and quadratic discriminant (QD) to calculate for each class the sample mean and the sample covariance, support vector machine (SVM) and KNN (K=10) to obtain a class.

Five cross-validations are used in the training and test such as 95% for train and 5% for test. The performance of the classifier was evaluated by accuracy. The LR was performed using each feature alone but Trees, SVM, and KNN had the best and the high performance with an accuracy = 100% using all features.

Later, Das et al. [29] adjust their previous work by adding the notion of PCA. PCA shows the forte correlation between the LBP, GRLM, GLCM, and Tamura. the evaluation of experiences (each one alone, the two best LBP+GRLM, the three best LBP+GRLM+GLCM, the four best LBP+GRLM +GLCM +Tamura, all the five) proves that the combination between the four best features had the better accuracy.

Another study was done by the same authors in [7] to classify CMB – LCA – MBEN – DN - Normal cases using three types of characteristics as texture, color, morphology. The number of these characteristics was 259, they were reduced using PCA. The whole process of this study is represented in Figure2.4. This study was done on MATLAB 2016b also, obtaining the results which prove the performance of PCA and the forte correlation between the color and shape. The evaluation of experiences (texture + color, texture + shape, color + shape) proof that the combination between these two features gives the best.

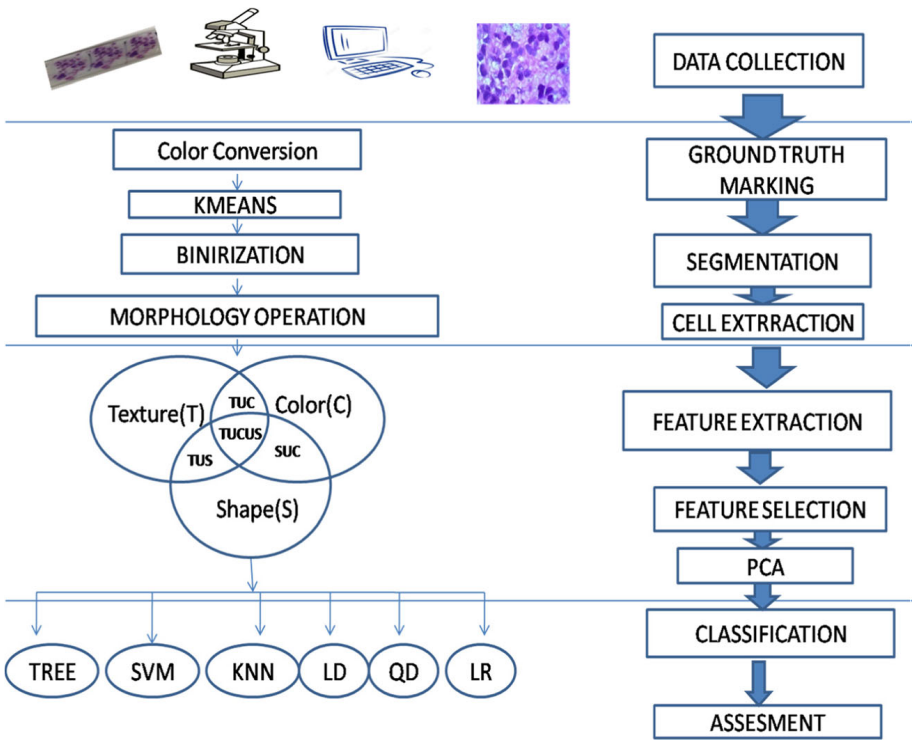

Figure 2.4: The process proposed by Das et al. using three features types [7].

In 2020, the same authors proposed in [8] an automated classification using the reduced feature set based on the multivariate variance analysis MANOVA. The classification was done at two levels, the first binary (malign or not) the second multi-class (the MB subtypes). The process of this work is shown in Figure2.5. The images were collected from the postoperative tissue blocks and were observed and marked for ground truth with a Leica 1CC50 HD microscope with a JPEG storage format. The characteristics used are texture (GLCM, GRLM, LBP, Tamura characteristics), color (intensity histogram, color chromatic moment, color correlogram), and morphology (Area, Euler Number, Orientation, Extent, Perimeter, Filled Area, Solidity, Eccentricity, major axis length, Equiv Diameter, and minor axis length). The images have been segmented using K-means and MANOVA for the verification of the equality of the variance between the groups (The experience was done first using PCA, thereafter with MANOVA). In the end, the classification is applied by SVM by adding the strategy "one against all" for the passage from SVM-binary to SVM-multi-classes with 75% of the data for learning and 25% for the test. the performance of this classifier was measured by accuracy, precision, recall, and Kappa. The results show that MANOVA is better than PCA because it can identify the significant features without any transformation of data assuring good performance.

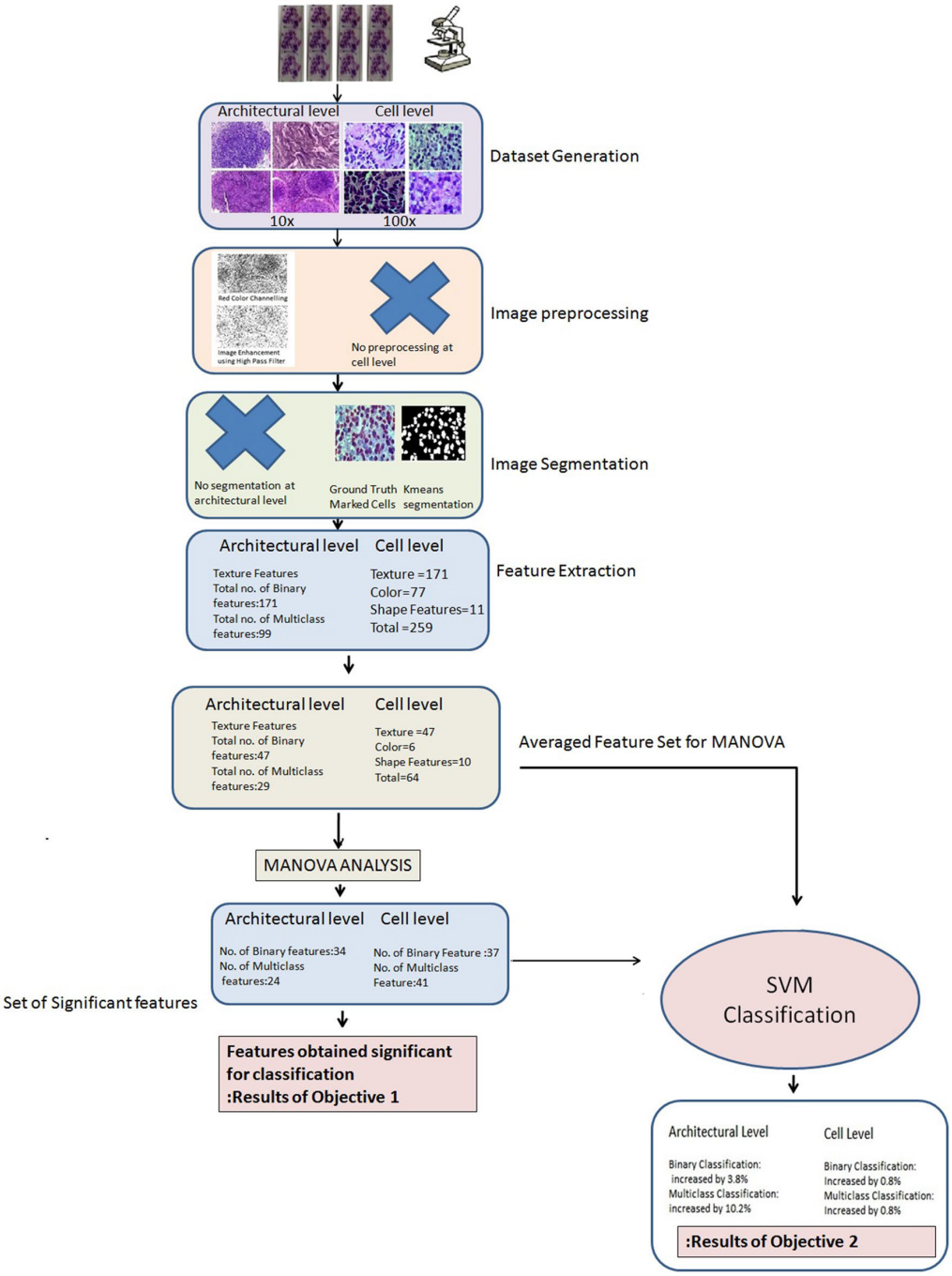

Figure 2.5: The proposed process by Das et al. using MANOVA [8].

In 2021, the same database used by Das et al. was applied in by O. Attallah [9] following the process shown in Figure2.6. In this study, the data was resized and augmented by acquiring the images in several directions. By next, feature extraction using three CNNs architectures: DenseNet-201, ResNet-50 and MobileNet. Then, the use of the method discrete wavelet transforms (DWT) for the time calculation. The PCA for dimension reduction is applied afterward. In the end, the classification with linear discriminant analysis (LDA) and ensemble subspace discriminant (ESD). The choice of this process has been through 4 scenarios based

on the best performance. The first scenario is using of the three deep learning CNNs. The second, extraction of spatial features from the 3 CNNs and classify them with SVM, cubic SVM, LDA, and KNN, and ESD classifiers. The third scenario concerned the extraction from the spatial DL features of the time-frequency features to form three spatial-time-frequency DL features sets and classify them with the same classifiers used in the second scenario. In the last scenario, PCA and DWT were applied on the three spatial-time-frequency DL feature sets.

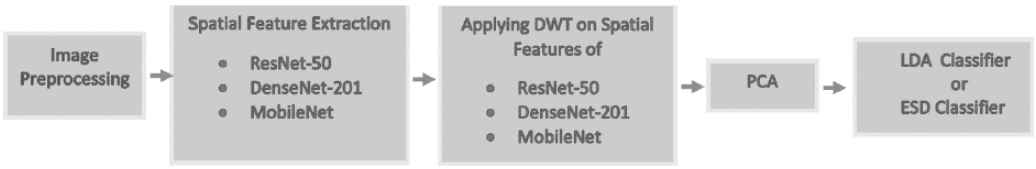

Figure 2.6: the proposed process of O. Attallah study [9].

In the same year, Bengs et al. [10] proposed the classification of the two subtypes of MB (CMB, and DN) using deep transfer learning with multi-scale-efficient Nets (end-to-end approach). In this study, 161 cases were analyzed with 103 CMB and 58 DN. The images were stained with hematoxylin and eosin. The labeling of these images is validated by experts in whole slide image WSI with the possibility of the presence of several cancerous areas in a single WSI image with a resolution of 2000x2000 and magnification of 200x. In this proposal, Bengs et al. applied the efficientNet from B0 to B5 with AlexNet, VGG16, and ResNet for input resolution impact assessment. Data is divided into 10 sub-blocks (cross-validation) in which each subset contains five cases of CMB and 2 cases of DN. A weighting of the loss of individual classes inversely proportional to the samples of each class was made to solve the problem of unbalanced classification. To increase the size of the data, the images were rotated vertically, horizontally, and randomly. The performance was measured by F-score, sensibility, specificity, area under the receiver operating curve AUC, confidence interval CI of 95%, bias-corrected, accelerated bootstrapping with 10000 bootstraps, permutation test with 10000 samples. The process of this study is in Figure2.7 and the results show that the performance of EfficientNet-B5 with the high resolution gives the best performance compared to the tree classic deep performance used in this experience.

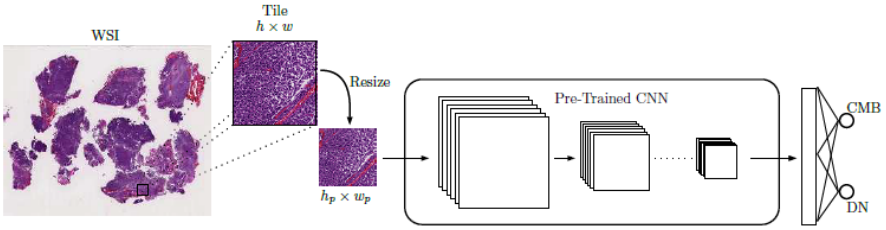

Figure 2.7: The workflow process proposed by Bengs et al. [10].

2 Synthesis

In these recent works, two approaches were used based on the type of image, MRI and histological, but we will focus on the histological approach because the gold standard for any tumor classification is the biopsy.

We report several works (see Table 2.1) that used features with different levels of features, high-level features, advanced latent features, and deep level features. In the advanced level features, MANOVA and PCA were used for dimensions reducing to obtain performance better than using all features with classic classifiers.

IMG TYPE	FEATURES	CLASSIFIERS	CLASSES
MRI [4]	ACP, univariate with ADC, CCBV, UCBV, K2	AdaBoost, RandomForest, KNN, RN	MB, EP, Astrocytoma
MRI [28]	/	ResNet 50-32-4d	MB, DMC, PA, EP
Hist [5]	VGG-16, IBCa	softmax	anaplasic, non anaplasic
Hist [6]	GLCM, GRLN, HOG,Tamura	SVM, KNN, Trees, LR, LD, QD	Normal, CMB
Hist [29]	PCA with GLCM, GRLN, HOG,Tamura	SVM, KNN, Trees, LD, QD	CMB, LCA, MBEN, DN, Normal
Hist [8]	PCA with texture, color, morphology	SVM, KNN, Trees, LR, LD, QD	CMB, LCA, MBEN, DN, Normal
Hist [7]	with MANOVA	SVM	CMB, LCA, MBEN, DN, Normal
Hist [9]	PCA with denseNet-201, mobileNet and ResNet-50	LDA, ESD	CMB, LCA, MBEN, DN, Normal
Hist [10]	/	AlexNet, ResNet 50, EfficientNet,VGG 16	CMB, DN

Table 2.1: List of related works with details of features and classifiers applied.

For the deep level features, a hybrid approach was applied where CNNs pretrained models were used to extract features and different classifiers to obtain tumor class. In these hybrid approaches, the process goes through three stages, first, feature extraction, second, feature selection or dimension reduction step, and finally, the classification step.

So, the transfer learning is used in some works to extract features in which a model that was trained on images of natures (ImageNet) which have a very low similarity with our small medulloblastoma dataset, and histological images (just in one work).

To work with CNN which is current and trending is good idea, but must be careful on the choice of the technique and approach. Since we have a small dataset we must go into two approaches of transfer learning, the partial fine tuning or deep features extractor. The partial fine tuning, if the models are trained on dataset not similar with our data. And the deep features extractor, if we have pretrained models on dataset very similar to our data.

The choice of the process is clear now, but there are different architectures and models that it is important to make the right choice. That what we will see in the next chapter.

3 Conclusion

In this chapter, we have cited the different works on the classification of medulloblastoma. The success of deep learning methods in achieving image classification has gained a lot of attention in recent years. Recent study has naturally proposed some works based on deep learning architectures for the classification of medulloblastoma. In the next chapter, we will show our proposition to classify normal and abnormal medulloblastoma cases.

Chapter 3

Transfer learning comparative study

In this chapter, for the classification of MB images, we will carry out a comparative study with two transfer learning approaches (partial fine-tuning, and deep features extractor) using different Convolution Neuronal Networks architecture as VGG-16, VGG-19, ResNet-50, and Inception V3 pre-trained models on ImageNet (straight-forward natural-image classification datasets) and the CancerNet pre-trained on PCam (histological images benchmark).

The idea behind this experimentation is to investigate which from the partial fine-tuning and deep features using pre-trained models on ImageNet or PCam are the most suitable for the Childhood MB Dataset used in our application.

1 Childhood MB Dataset Description

All of Guwahati Medical College, GMCH Hospital and GNRC had been collaborating to collect the childhood MB dataset. This data was obtained from patients under 15 years of age who were identified in GMCH. Samples were taken from tissue blocks and used as a component of the postoperative process. These blocks of tissues were stained, at Ayursundra Pvt, using H & E with the help of a local medical professional.

Subsequently, a qualified pathologist at the pathology department of the GNRC observed the scans of the slide and the region of interest for ground truth. Pictures of the region of interest were captured using a Leica 1CC50 HD microscope and were saved in JPG format for the four subtypes of MB tumors. The dataset contains 204 and 258 pictures with magnification 10x and 100x [see Figure 3.1 , 3.2] from 15 patients [9,27].

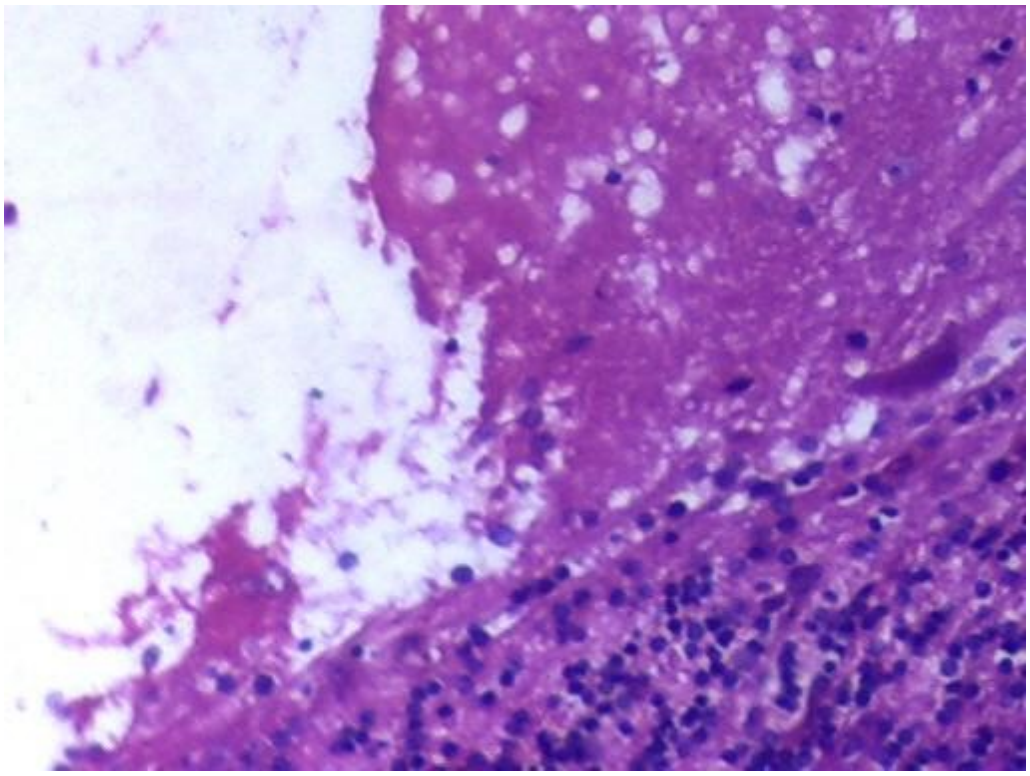

Figure 3.1: Normal medulloblastoma with magnification 10x.

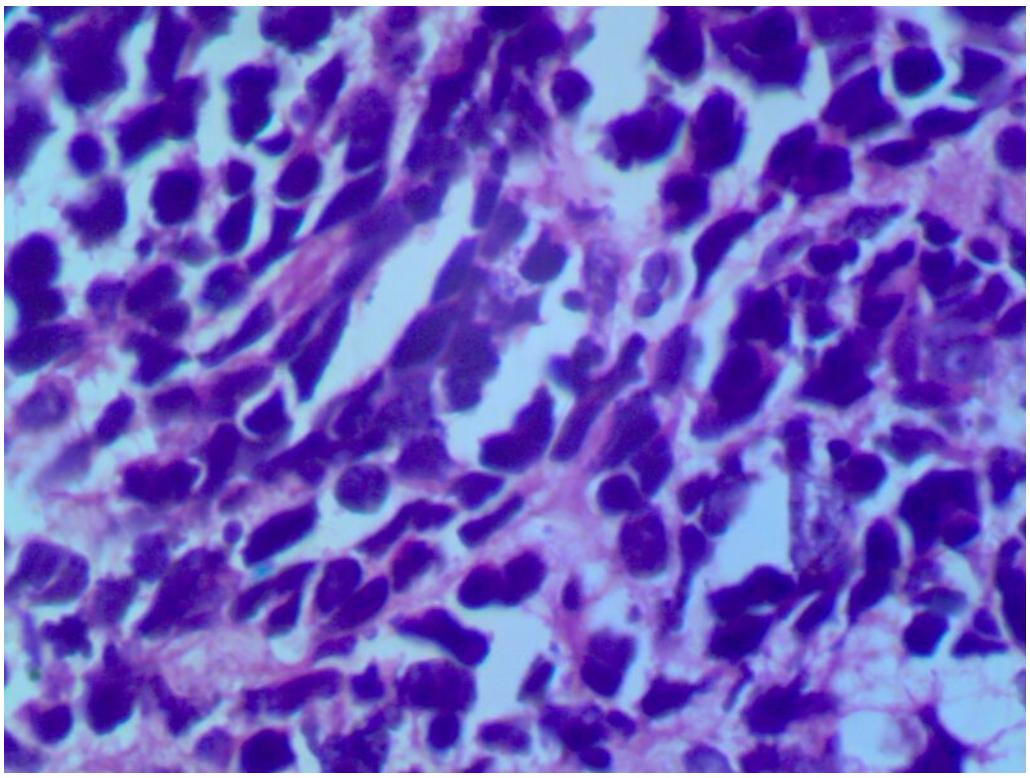

Figure 3.2: Abnormal medulloblastoma with magnification 100x.

2 Convolutional Neural Network

CNN is a type of artificial neural network with convolutional operations; it is made up of neurons whose parameters take the form of weights and biases that can be changed or adjusted during the learning operation.

CNN is formed by a succession of processing blocks which are represented in Figure 3.3, to extract the characteristics discriminating the class of membership of the image of others [11]. A treatment block consists of one to several:

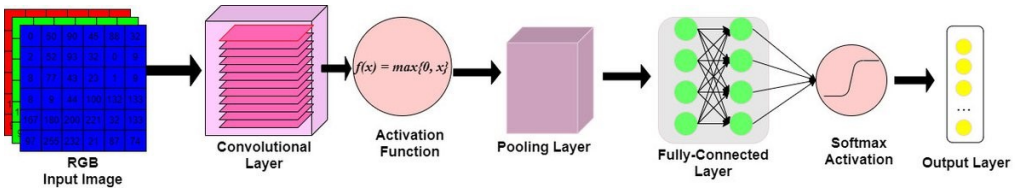

Figure 3.3: CNN architecture [11].

- Convolutional layers (CONV): Its goal is to identify the presence of a set of features in the images received as input. For this, convolution filtering is carried out: the principle is to "drag" a window representing the filter onto the image, and to calculate the convolution product between the filter and each portion of the scanned image (Figure 3.4). The output of this operation is called feature map or also activation map.

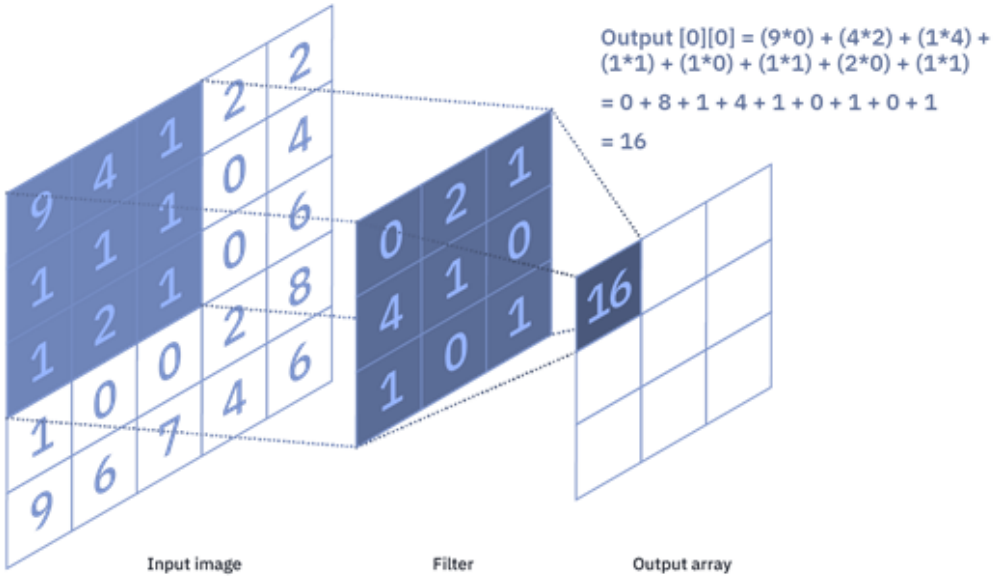

Figure 3.4: The convolutional layer of CNN [12].

- Activation layer (ReLU): The correction or activation layer is the application of a non-linear function to the characteristic maps output from the convolution layer. The ReLU function ($ReLU(x) = max(0, x)$) (Rectified Linear

Units) remove every negative value from convolution results and replace it with zeros as shown in Figure 3.5. it makes it easier to extract complex features that cannot be modeled by a linear combination of a regression algorithm.

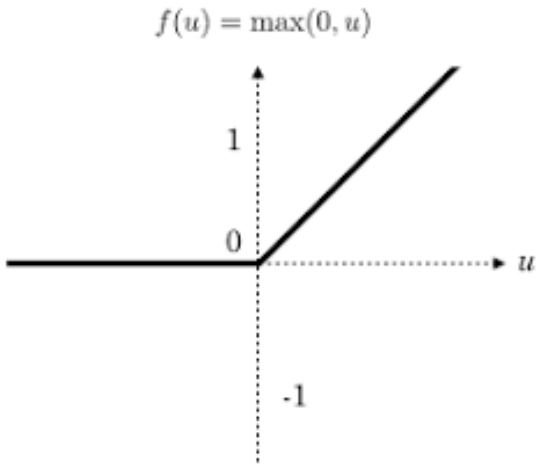

Figure 3.5: The RELU function.

- Pooling layer: In this layer, a downsampling operation is applied after a convolutional layer (see Figure 3.6). In particular, the most popular types of pooling are max and average pooling, where maximum and average values are taken, respectively. The principle of the max and average pool is represented in Figure 3.7.

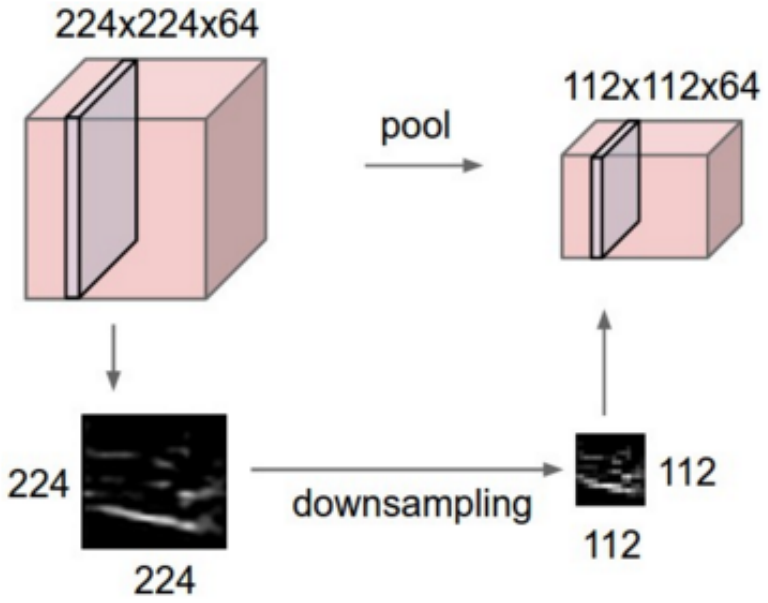

Figure 3.6: The size of image before and after using pooling layer.

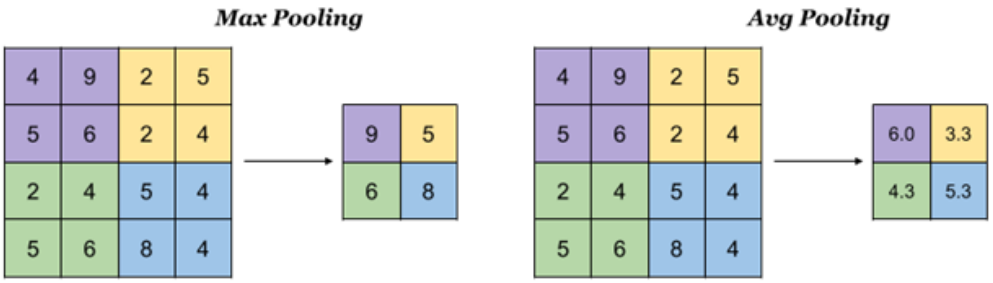

Figure 3.7: The average and the max pool layer [13].

- Fully-Connected Layer: This layer is applying to a previously flattened input (It simply consists of taking all previously calculated matrices and put them into a single list or a vector, in order to exploit them in the input layer of a neural network) where each input is connected to all neurons as represented in Figure 3.8. Fully connected layers are typically present at the end of CNN architectures and can be used to optimize goals such as class scores.

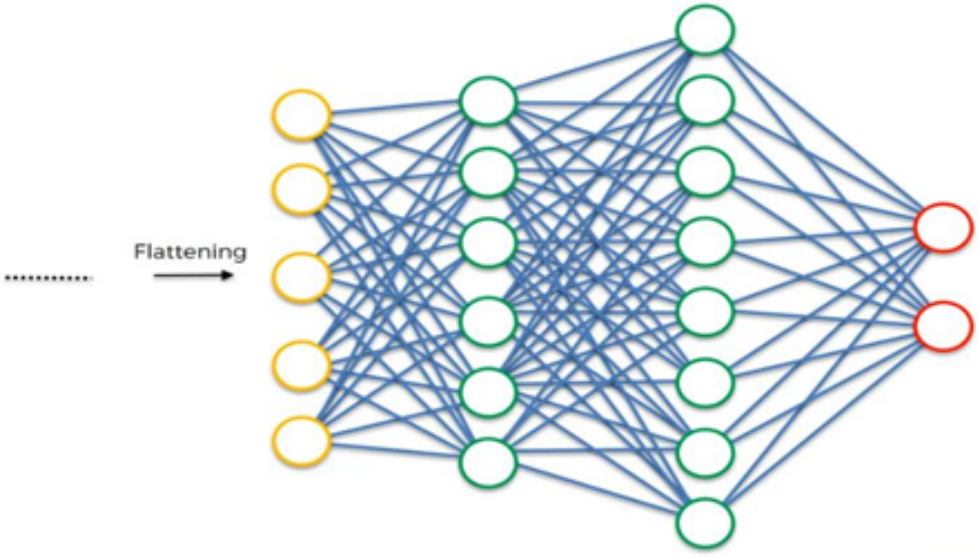

Figure 3.8: The fully connected layer.

- The activation function used for prediction: selecting the right activation function depends on the type of classification problem, and two common functions are:
  - Sigmoid is generally used for binary classification problems, as it is a logistic function
  - Softmax ensures that the sum of values in the output layer sum to 1 and can be used for both binary and multi-class classification problems.

3 Transfer learning

With the contribution of Deep Learning to allowed great advances in computer vision, it is worth experimenting and tests the application of CNNs. However, to achieve satisfactory performances these models rely on huge datasets which in some contexts, it is very difficult, and sometimes impossible to get or to compute.

Transfer learning is a deep learning technique that makes it possible to re-use general models already trained (a model created by someone else to solve a similar problem) to another one. Transfer learning is applied when the source task's model has been trained on a vastly bigger training set than could be acquired for the destination task. In transfer learning, the trained model is compiled as "weights" of the network, then repurpose the learned features, or transfer to a second network.

There are three approaches of transfer learning, which are: total fine-tuning, partial fine-tuning, and deep features extractor [11,30].

3.1 Total fine-tuning

Total fine-tuning is used when the size of the data of our task is large and its similarity with the data pre-trained is very high [30]. We replace the last fully connected layer of the pre-trained network with a classifier adapted to the new problem (SVM, logistic regression...) and initialized randomly. All layers are then trained on the new images. It is used when the new collection of images is large: in this case, we can afford to train the whole network without running the risk of over-fitting. Moreover, since the parameters of all layers (except the last one) are initially those of the pre-trained network, the learning phase will be done faster than if the initialization had been random.

3.2 Partial fine tuning

Partial fine-tuning is used when the size of the data of our task is small and its similarity with the data pre-trained is very low [30]. In this approach, we use the pre-trained model of 'n' layers with random initialization of all its weights. Have freeze a number 'k' of initial layers of the pre-trained model then we train this model again on our dataset with the remaining '$n - k$' layers. More precisely, the last fully connected layer is replaced by the new randomly initialized classifier, and the parameters of some layers of the pre-trained network are fixed. Thus, in addition to the classifier, we train on the new images the non-fixed layers, which generally correspond to the highest ones of the network.

3.3 Deep features extractor

Deep features extractor is used when the size of the data of our task is small and its similarity with the data pre-trained is very high. In this approach, we don't need to retrain the model because we have the same task. Here we remove the

last fully connected layer and fix all the other parameters. This truncated network will then compute the representation of each input image from the features already learned during the pre-training. We then train a randomly initialized classifier on these representations to solve the new problem [30].

4 Material and methods

In this master project, we will work with two transfer learning approaches such as partial fine-tuning and deep features extractor which correspond the best to our dataset classification. In each approach we will use two different pre-trained models, models trained on ImageNet (straight-forward natural-image classification datasets) and on PCam (histological images benchmark).

4.1 Pre-trained models on ImageNet dataset for nature image classification

ImageNet [31] is a large and an organized image database according to the WordNet which contains 14197122 (more than 14 million) images organized and labeled in a hierarchy (1000 classes, a sample is presented in Figure 3.9). Each node of the hierarchy is depicted by hundreds and thousands of images. It was designed by academics for research in computer vision and it has an important role in the advancement of this domain through various challenges.

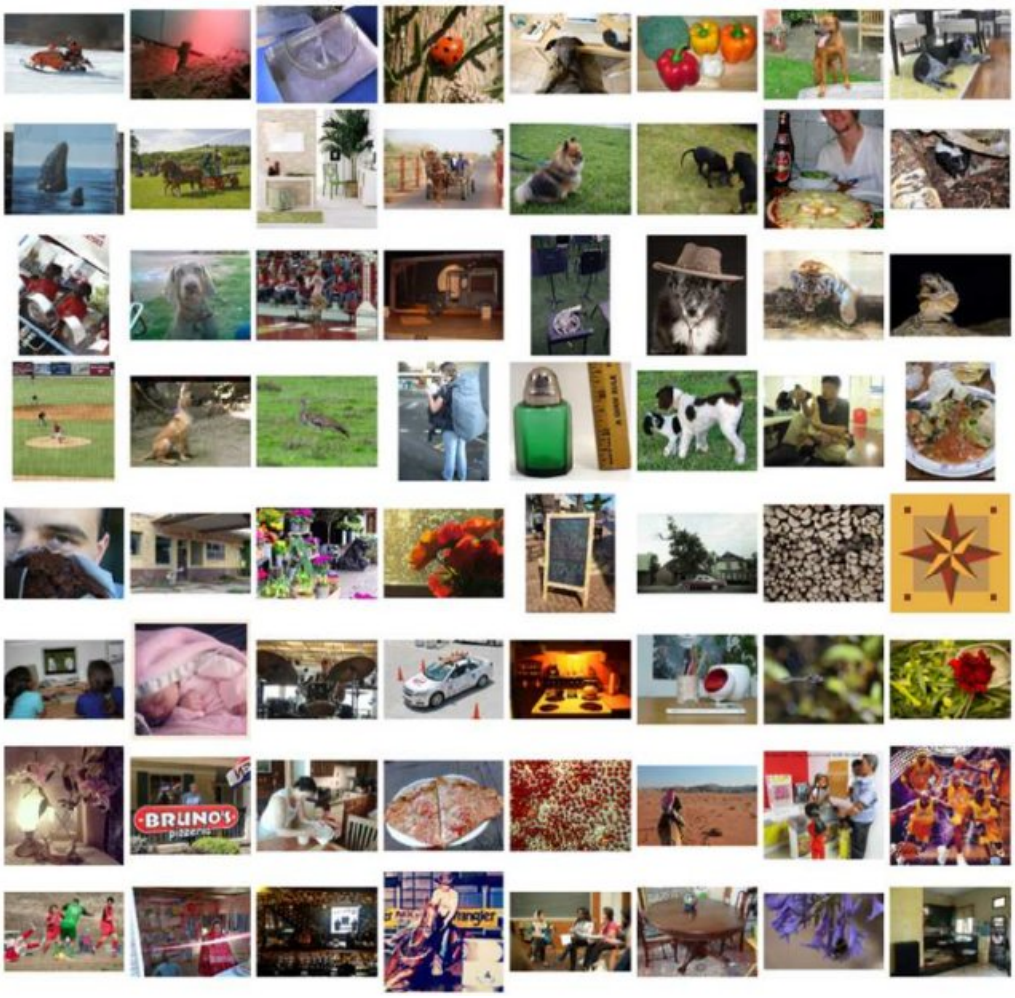

Figure 3.9: ImageNet samples [14]

There are several and variant architectures of CNN, each has its own structure which is completely different from the other or which is an improvement on a previous architecture. In our work, we are interested in those who have proven themselves in science competitions and who have given a boost to the field of deep learning around the world [32] such as VGG-16, VGG-19, ResNet-50, Inception V3.

VGG-16

VGG-16 is a CNN model that contain 16 layers of convolution as represented in Figure 3.10. proposed by Simonyan and Zisserman [16], it gained notoriety by winning the ILSVRC (ImageNet Large Scale Visual Recognition Challenge) competition in 2014 [1]. It marked a progression compared to previous models by proposing, in the convolution layers, convolution kernels of smaller dimensions (3×3) than what had been done until then (the AlexeNet [33]). The model has been trained over weeks using state of the art graphics cards.

[1] https://image-net.org/challenges/LSVRC/2014/

As we can see in Figure 3.10, the architecture of VGG-16 is clear and easy to understand which is also a strength of this model.

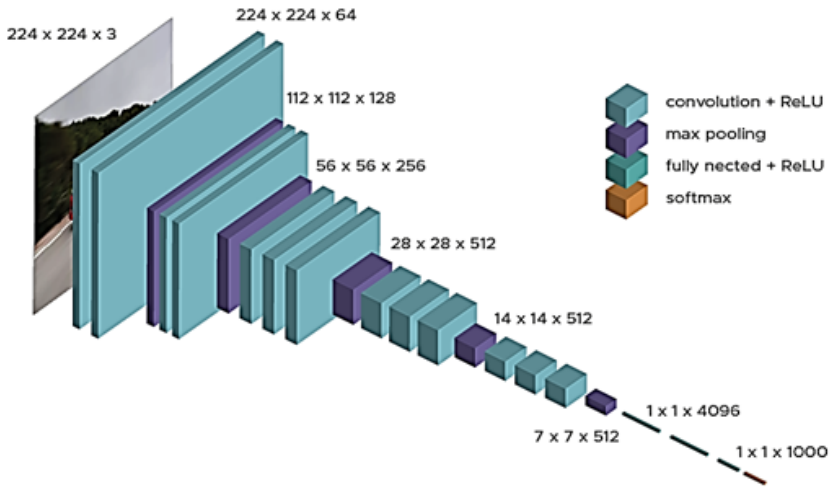

Figure 3.10: The structure of VGG-16 [15].

VGG-19

VGG-19 [16] is the improvement of VGG-16 which goes from 16 layers of convolutions to 19 layers (as represented in Figure 3.11). The 19 weight layers consist of 16 convolutional layers with 3 fully connected layers and the same 5 pooling layers.

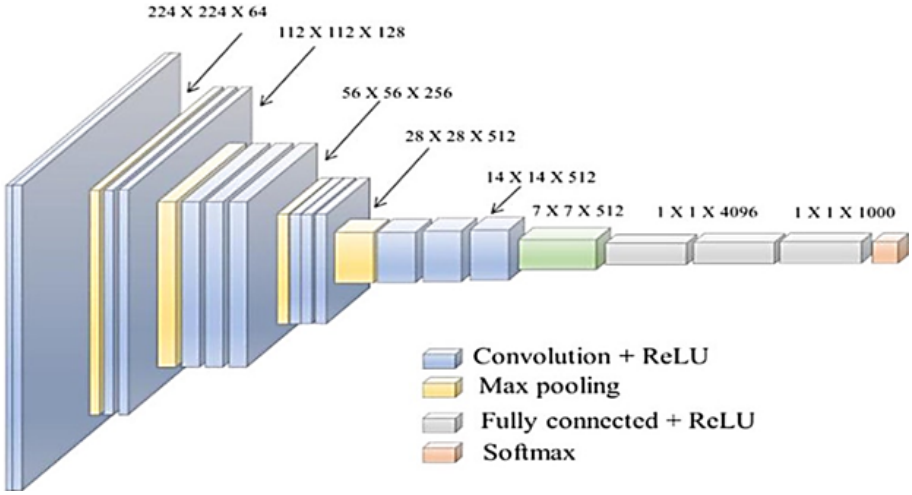

Figure 3.11: The structure of VGG-19 [16]

Inception (GoogleNet) V3

Inception's deep convolutional architecture was introduced as GoogLeNet (Szegedy et al. 2015a [34]), named also Inception V1. Later, the Inception architecture was

refined in various ways, first by the introduction of batch normalization (Ioffe and Szegedy 2015 [35]) (Inception V2). Later, by additional factorization ideas in the third iteration (Szegedy et al. 2015b [36]) which is known as Inception V3 [17].

In GoogLeNet / Inception V1, auxiliary classifiers are used to have a deeper network. In Inception V3, the auxiliary classifier is used as a regularizer. In fact, in deep learning the modules are still intuitive. Inception V3 architecture use Label Smoothing, 7×7 convolutions, and auxiliary classifier with the use of batch normalization for layers [17]. The structure is represented in Figure 3.12. With 42 layers deep, the computational cost is only about 2.5 more than that of GoogLeNet, and much more efficient than that of VGG-Net [16].

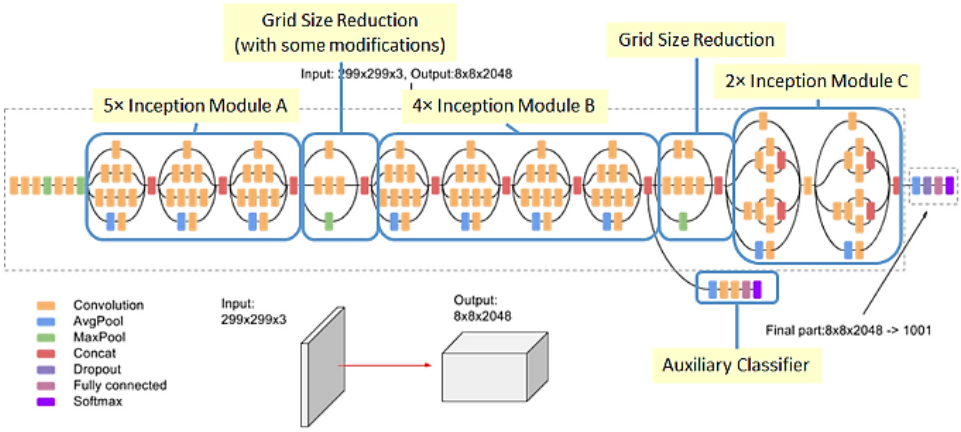

Figure 3.12: The structure of Inception V3 [17].

ResNet-50

The ResNet architecture, developed by He et al. [37], marked a milestone in the introduction of an exotic type of architecture based on "modules", or as it is now called, "networks within networks". These networks introduced the concept of "residual connections". There are variations of ResNet with different number of layers, but the most used is ResNet-50 the winner of ImageNet challenge 2015 [2], which consists of 50 layers (Figure3.13).

[2]https://image-net.org/challenges/LSVRC/2015/

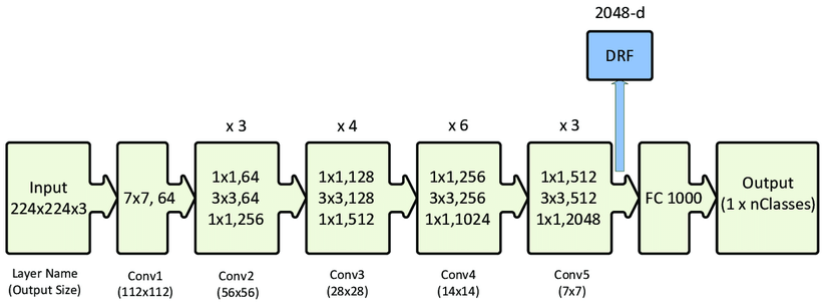

Figure 3.13: The structure of ResNet-50 [18].

It is remarkable that although it has a lot more layers than the VGG, it needs a lot less memory, almost 5 times less. This is because this network, instead of dense layers in the classification stage, uses a type of layer called GlobalAveragePooling, which converts the 2D activity maps from the last layer of the feature extraction stage into a vector of n-classes used to calculate the probability of belonging to each class.

4.2 Pre-trained models on PCam dataset for histological image classification

PCam is derived by Veeling et al. [19] from the Camelyon16 Challenge [38], which contains 400 H&E stained WSIs of sentinel lymph node sections. The slides were acquired and digitized at 2 different centers using a 40x objective (resultant pixel resolution of 0.243 microns).

The PCam benchmark dataset contains 327680 color histological images with a resolution of 96×96 px (see Figure 3.14). This database consists of two classes in which each image is labeled indicating the presence of metastatic tissue.

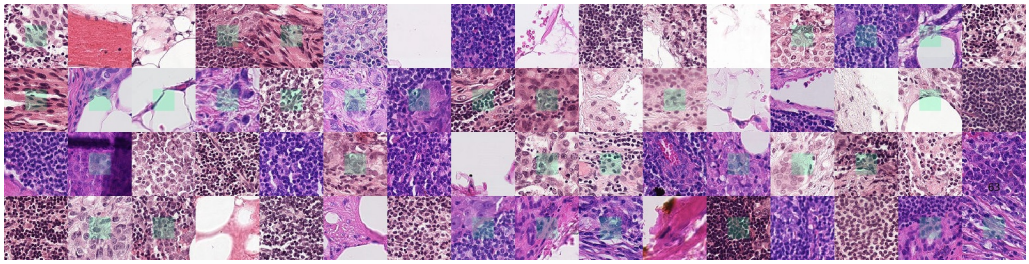

Figure 3.14: Samples example from PCam. Green boxes indicate tumor tissue in center region, which dictates a positive label [19].

Its packs the clinically relevant task of metastasis detection into a straightforward binary image classification task, akin to CIFAR-10 and MNIST, it should be noted that PCam benchmark is bigger than CIFAR10 and smaller than ImageNet.

CancerNet

CancerNet is a CNN model Similar to VGGNet. It uses exclusively 3×3 CONV filters, however, unlike VGG, it uses depth-wise separable convolution rather than standard convolution layers (see Figure 3.15). The depth-wise separable convolution is appreciated for its efficiency in addition it requires less memory and less computation [39]. This model was firstly used by Adrian Rosebrock on Breast cancer images [40] and some modifications have been taken over by Soumya Ranjan Behera on Histopathologic Cancer Detection in the Kaggle Competition [20] for training the model on PCam benchmark dataset.

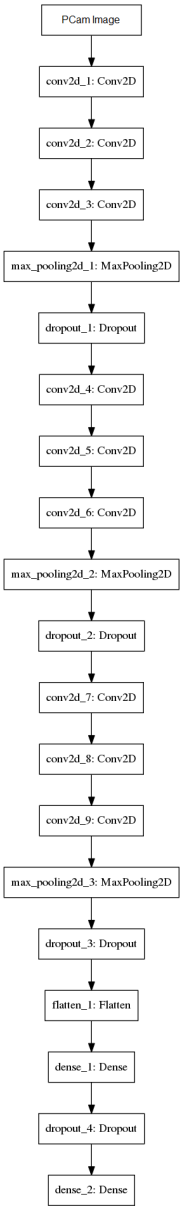

Figure 3.15: CancerNet layers used by Soumya Ranjan Behera on Histopathologic Cancer Detection Kaggle Competition [20].

5 Results and discussion

5.1 Programming environment

All the experiments were performed with google colaboratory, applied also Colab[3] or google colab, which is a Google research project that allows the researchers to write and execute Python in their browser, with zero configuration required, Free access to GPUs, Easy sharing. Colab is based on a Jupyter notebook environment that requires no setup to use and runs entirely in the cloud.

5.2 Database preparation

We split our data by distributing the images into three folders, train, test, and validation. The first folder contains 240 images split into two other folders, normal and abnormal. The test folder contains 120 images split into two folders, normal and abnormal. The validation contains two images, the first image represents normal MB, and the second represents abnormal MB.

5.3 Classification process

The experimentation is based on a comparative study of two transfer learning strategies with various CNN architecture trained on nature image classification (ImageNet) and histological image classification (PCam). The idea behind this experimentation is to investigate which from the partial fine-tuning and deep features using pre-trained models on ImageNet or PCam a dataset on the same domain gives the best results on the Childhood MB Dataset used in our application.

- In partial fine-tuning strategy, first with ImageNet dataset, we need to freeze some layers and retrain them again with our dataset. In the VGG-16 and VGG-19, we had to freeze the two last layers. However, in ResNet-50 and inception, we had freeze layers from the layer 171 and 305 respectively. Second, for the fine-tuning with PCam dataset, we load all of VGG-16, VGG-19, ResNet-50, Inception V3 pre-trained models without weights. Next, we add the weights obtained after the training of CancerNet on PCam dataset available on Kaggle [4] to these models. We freeze the same layers as partial fine-tuning with ImageNet dataset in all of the four architectures.

- In deep features extractor strategy, the models are used to extract features from ImageNet dataset. Then transfer them to another classifier. Here, we use the SVM [41] and the Random forest [42] for the classification. For PCam dataset, we load the models with CancerNet weights and without the top layer. We use the models to extract features and transfer them to the SVM and Random Forest classifier for performance study.

[3]https://colab.research.google.com/

[4]https://www.kaggle.com/soumya044/histopathologic-cancer-detection

5.4 Results

The evaluation of models was measured by the accuracy. The accuracy was calculated as: Accuracy = (true positive + true negative)/number of samples *100 The results of the two approaches using pre-trained models on ImageNet dataset are in Table 3.1, and pre-trained models on PCam dataset are in Table 3.2.

- In the deep features extractor strategy, the pre-trained models on PCam dataset give the best result compared to the pre-trained models on ImageNet. This can be explained by the fact that the PCam dataset (histological lymph node sections images) have very high similarity with our dataset in contrary of ImageNet dataset (Nature images) which have a poor similarity.

  Nevertheless, the results of the deep features extractor with the pre-trained model on PCam must better than that we have obtained, it needs more parametrization. To obtain the best parameters, more experimentation should be have done. The SVM and random forest classifiers are at the same level, both give a high accuracy of 92.5%.

- On the other side, in the partial fine-tuning strategy, the pre-trained models on ImageNet give the best result in comparison with pre-trained on PCam dataset. Here the notion of similarity changes, because partial fine-tuning is used when the models were trained on a dataset not similar to our dataset. The partial fine-tuning allows us to re-use image representations learned from a source task and a dataset with a large quantity labeled on a second dataset which is different. The partial fine-tuning is intended for databases that are different from the pre-trained databases. The Inception V3 model outperforms the other architectures for partial fine-tuning on ImageNet. It gives the best accuracy at all (99.16%) (the graphs of loss and accuracy are represented in Figures 3.16,3.17).

	PARTIAL	DEEP FEATURES EXTRACTOR	
MODELS	FINE-TUNING (%)	**SVM** (%)	**Random Forest** (%)
VGG-16	94.16	87.5	85.83
VGG-19	93.33	90.83	87.5
ResNet-50	87.50	80	75
Inception V3	99.16	88.33	86.66

Table 3.1: Performances of pre-trained models on ImageNet.

	PARTIAL	DEEP FEATURES EXTRACTOR	
MODELS	FINE-TUNING (%)	**SVM** (%)	**Random Forest** (%)
VGG-16	80	90	88.33
VGG-19	81.66	92.5	92.5
ResNet-50	84.16	80	85.83
Inception V3	81.66	80	90

Table 3.2: Performances of pre-trained models on PCam.

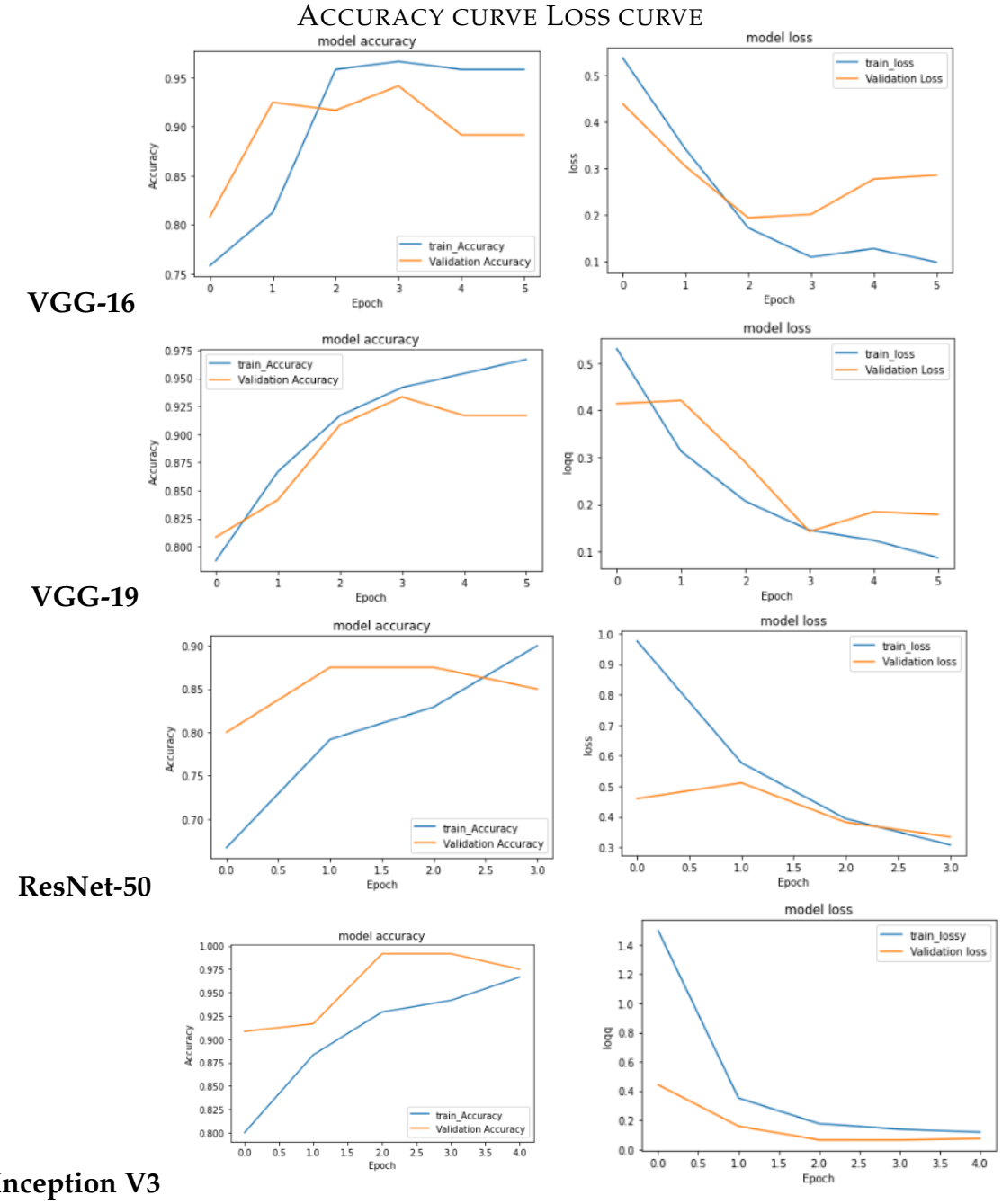

Figure 3.16: Performances of partial fine tuning with VGG-16, VGG-19, ResNet-50, Inception V3 pre-trained models on ImageNet dataset.

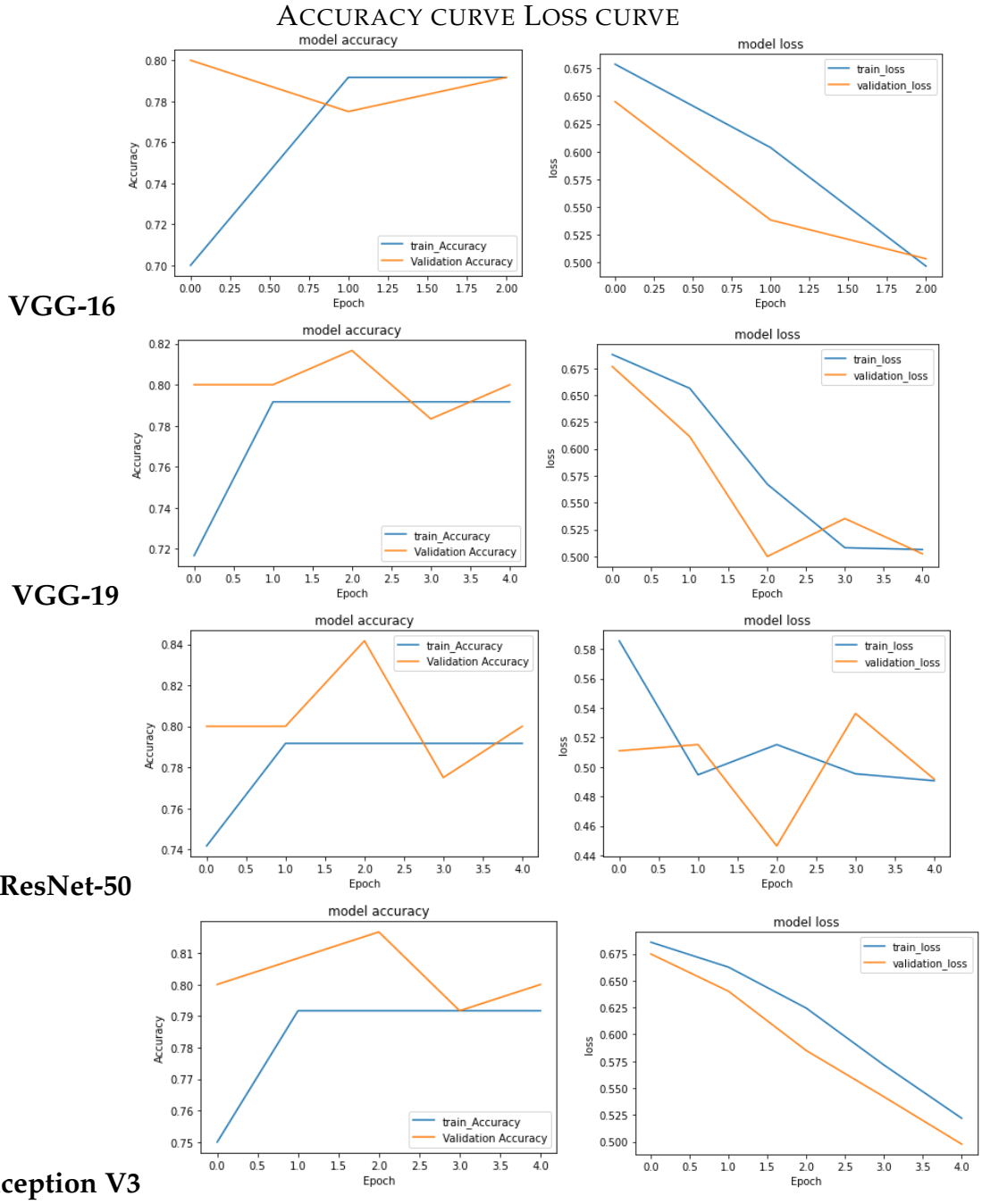

Figure 3.17: Performances of partial fine tuning with VGG-16, VGG-19, ResNet-50, Inception V3 pre-trained models on PCam dataset.

In the biomedical domain, there is a huge lack of big data, therefore, transfer learning via deep features extractor or partial fine-tuning is the best option for CNN application on a small dataset as is the case of our Childhood MB dataset. The idea behind this experimentation is to investigate which from the partial fine-tuning and deep features using pre-trained models on ImageNet (straightforward natural-image classification datasets) or PCam (histological images benchmark) are the most suitable for our case application.

In the related works cited, we found only the work of O. Attallah [9] and Bengs et al. [10] with deep features extractor using a pre-trained model on ImageNet, and Cruz-Roa et al. [5] with two different pre-trained models, one on ImageNet and the other on histological data. The comparative study conducted in this master project, quest on what is the best choice of transfer learning strategy to use and if the similarity between the pre-trained data and target data is important for best performances. Experiments prove that there is an important role of similarity when dealing with deep features extractor or partial fine-tuning. We conclude that when there is a high similarity between the two data, the deep features extractor is the best, when none, partial fine-tuning is more suitable.

6 Conclusion

For the automatic recognition of Childhood MB, we used in this work a small database, which forces us when applying convolutional neural networks models to use transfer learning approaches such as the deep features extractor or the partial fine-tuning. After the comparison study realized, experiments reveal that the deep features extractor is possible when the similarity condition is satisfied between the data we have and the data the model has trained. Otherwise, we must move to partial fine-tuning with no similarity between the two datasets and vice versa.

Conclusion

In this Master project, we set ourselves the objective of answering the question "*how to avoid the hard work and to gain time for medulloblastoma classification task*". We assumed that the solution is in machine learning approaches. We had then proposed an automated classification of the childhood medulloblastoma dataset based on CNNs. The size of our dataset forced us to deal with two approaches of transfer learning, partial fine-tuning or deep features extractor. The partial fine-tuning is used when we have models pre-trained on a dataset not similar to our dataset case application, the deep features extractor is used when there is a high similarity between our data and the dataset on which the model is trained.

We carry out a comparative study between partial fine-tuning and deep features extractor using the most popular architectures in literature as VGG-16, VGG-19, ResNet-50, Inception V3. The experimentation made is to evaluate the performances of the pre-trained models on both on ImageNet (straight-forward natural-image classification datasets) and the PCam dataset (histological images benchmark).

The experimentations conducted show that, even on partial fine-tuning or deep features extractor, the choice of the pre-trained model is very important.

The comparison study conducted make us concluded that there is a rule which is "*if the model was trained on data similar as which we have, the deep features extractor must be applied; If the model was trained on data similar as which we have, the partial fine-tuning must be applied; and vice versa.*" In our case application, the results show that the partial fine-tuning process using the Inception V3 pre-trained model on ImageNet achieved the best accuracy with 99.16%.

We had hoped that there would be more appreciable results even in deep features extractor and not just in the partial fine-tuning, this takes a lot of time and a lot of experimentation which we will be conducted in the future to reach a suitable setting. We have a great desire to use CanceNet and other architectures. Step-by-step and total fine-tuning are the two others CNN methods that we want to apply but the huge lack of big medical data binds us. We hope that will be resolved to dive into a new experience.

Bibliography

[1] Francisco Bautista, Victoria Fioravantti, Teresa de Rojas, Fernando Carceller, Luis Madero, Alvaro Lassaletta, and Lucas Moreno, "Medulloblastoma in children and adolescents: a systematic review of contemporary phase i and ii clinical trials and biology update," *Cancer medicine*, vol. 6, no. 11, pp. 2606–2624, Nov. 2017.

[2] Brent A. Orr, "Pathology, diagnostics, and classification of medulloblastoma," *Brain Pathology*, vol. 30, no. 3, pp. 664–678, 2020.

[3] Paul A Northcott, Adrian M Dubuc, Stefan Pfister, and Michael D Taylor, "Molecular subgroups of medulloblastoma," *Expert review of neurotherapeutics*, vol. 12, no. 7, pp. 871–884, July 2012.

[4] James T Grist, Stephanie Withey, Lesley MacPherson, Adam Oates, Stephen Powell, Jan Novak, Laurence Abernethy, Barry Pizer, Richard Grundy, Simon Bailey, Dipayan Mitra, Theodoros N Arvanitis, Dorothee P Auer, Shivaram Avula, and Andrew C Peet, "Distinguishing between paediatric brain tumour types using multi-parametric magnetic resonance imaging and machine learning: A multi-site study," *NeuroImage. Clinical*, vol. 25, pp. 102172–102172, 2020.

[5] Angel Cruz-Roa, John Arévalo, Alexander Judkins, Anant Madabhushi, and Fabio González, "A method for medulloblastoma tumor differentiation based on convolutional neural networks and transfer learning," in *11th International Symposium on Medical Information Processing and Analysis*, Eduardo Romero, Natasha Lepore, Juan D. García-Arteaga, and Jorge Brieva, Eds., Dec. 2015, vol. 9681 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 968103.

[6] Daisy Das, Lipi B.Mahanta, Shabnam Ahmed, Basanta Kr.Baishya, and Inamul Haque, "Automated classification of childhood brain tumoursbased on texture feature," *Songklanakarin J. Sci. Technol*, vol. 41, no. 5, pp. 1014–1020, 2019.

[7] Daisy Das, Lipi B. Mahanta, Shabnam Ahmed, Basanta Kr. Baishya, and Inamul Haque, "Study on contribution of biological interpretable and computer-aided features towards the classification of childhood medulloblastoma cells," *Journal of Medical Systems*, vol. 42, no. 8, pp. 151–, 2018.

[8] Daisy Das, Lipi B. Mahanta, Shabnam Ahmed, and Basanta K. Baishya, "A study on manova as an effective feature reduction technique in classification

of childhood medulloblastoma and its subtypes," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 9, no. 1, pp. 16–, 2020.

[9] Omneya Attallah, "Mb-ai-his: Histopathological diagnosis of pediatric medulloblastoma and its subtypes via ai," *Diagnostics*, vol. 11, no. 2, 2021.

[10] Marcel Bengs, Michael Bockmayr, Ulrich Schüller, and Alexander Schlaefer, "Medulloblastoma tumor classification using deep transfer learning with multi-scale EfficientNets," in *Medical Imaging 2021: Digital Pathology*, John E. Tomaszewski and Aaron D. Ward, Eds. International Society for Optics and Photonics, 2021, vol. 11603, pp. 70 – 75, SPIE.

[11] James McDermott, "Hands-on transfer learning with keras and the vgg16 model," learndatasci.com, March 20, 2021, Accessed June 20, 2021, https://www.learndatasci.com/tutorials/hands-on-transfer-learning-keras/.

[12] Qiuhong Ke, Jun Liu, Mohammed Bennamoun, Senjian An, Ferdous Sohel, and Farid Boussaid, "Chapter 5 - computer vision for human-machine interaction," in *Computer Vision for Assistive Healthcare*, Marco Leo and Giovanni Maria Farinella, Eds., Computer Vision and Pattern Recognition, pp. 127–145. Academic Press, 2018.

[13] Sabina Pokhrel, "Beginners guide to convolutional neural networks,learn about the important components that make up a convolutional neural network," towardsdatascience, September 19, 2019, Accessed 04 june 2021, https://towardsdatascience.com/beginners-guide-to-understanding-convolutional-neural-networks-ae9ed58bb17d.

[14] Khawaja Tehseen Ahmed, Aun Irtaza, and Muhammad Amjad Iqbal, "Fusion of local and global features for effective image extraction," *Applied Intelligence*, vol. 47, no. 2, pp. 526–543, 2017.

[15] Xiangyu Zhang, Jianhua Zou, Kaiming He, and Jian Sun, "Accelerating very deep convolutional networks for classification and detection," *CoRR*, vol. abs/1505.06798, 2015.

[16] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun, Eds., 2015.

[17] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.

[18] Ammar Mahmood, Ana Giraldo, Mohammed Bennamoun, Senjian An, Ferdous Sohel, Farid Boussaid, Renae Hovey, Robert Fisher, and Gary Kendrick, "Automatic hierarchical classification of kelps using deep residual features," *Sensors*, vol. 20, pp. 447, 01 2020.

[19] Bastiaan S. Veeling, Jasper Linmans, Jim Winkens, Taco Cohen, and Max Welling, "Rotation equivariant cnns for digital pathology," *CoRR*, vol. abs/1806.03962, 2018.

[20] Soumya Ranjan Behera, "Histopathologic cancer detection," Kaggle.com, May 15, 2019, Accessed 12 june 2021., www.kaggle.com/soumya044/histopathologic-cancer-detection.

[21] Anthony Joud, Olivier Klein, Pascal Chastagner, Pierre-Henri Pretat, Valérie Bernier-Chastagner, and Jean-Claude Marchal, "Pediatric medulloblastomas," Encyclopedia Neurochirurgica, May 12, 2015, Accessed in 04 April 2021., https://www.neurochirurgica.org/spip.php?article31&artpage=4-7.

[22] Lachi Pavan Kumar, Syed Fayaz Ahmed J Deepa, I Moinca, P Suresh, and K V J R Naidu, "Medulloblastoma: A common pediatric tumor: Prognostic factors and predictors of outcome," *Asian journal of neurosurgery*, vol. 10, no. 1, pp. 50–50, 2015.

[23] Jianzhong Yu, Rui Zhao, Wei Shi, and Hao Li, "Risk factors for the prognosis of pediatric medulloblastoma: a retrospective analysis of 40 cases," *Clinics (Sao Paulo, Brazil)*, vol. 72, no. 5, pp. 294–304, May 2017.

[24] C. Dunham, C. Foster, J. Triscott, and S. Dunn, "4. biphasic histology is associated with the non-wnt/shh molecular subtype of medulloblastoma," *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*, vol. 42, no. S2, pp. S2S3, 2015.

[25] S Alessi, M Grignani, and L Carone, "Ganglioneuroblastoma: Case report and review of the literature," *Journal of ultrasound*, vol. 14, no. 2, pp. 84–88, June 2011.

[26] Sarah E. S. Leary and James M. Olson, "The molecular classification of medulloblastoma: driving the next generation clinical trials," *Current opinion in pediatrics*, vol. 24, no. 1, pp. 33–39, Feb. 2012.

[27] Daisy Das and Lipi B. Mahanta, "Childhood medulloblastoma microscopic images," *IEEE DataPort*, 2020.

[28] J.L. Quon, W. Bala, L.C. Chen, J. Wright, L.H. Kim, M. Han, K. Shpanskaya, E.H. Lee, E. Tong, M. Iv, J. Seekins, M.P. Lungren, K.R.M. Braun, T.Y. Poussaint, S. Laughlin, M.D. Taylor, R.M. Lober, H. Vogel, P.G. Fisher, G.A. Grant, V. Ramaswamy, N.A. Vitanza, C.Y. Ho, M.S.B. Edwards, S.H. Cheshier, and K.W. Yeom, "Deep learning for pediatric posterior fossa tumor detection and classification: A multi-institutional study," *American Journal of Neuroradiology*, 2020.

[29] Daisy Das, Lipi B.Mahanta, Shabnam Ahmed, and Basanta Kr.Baishya, "Classification of childhood medulloblastoma into who-defined multiple subtypes based on textural analysis," *Journal of Microscopy*, vol. 279, no. 1, pp. 26–38, 2020.

[30] Dishashree Gupta, "Transfer learning and the art of using pre-trained models in deep learning," Analytics Vidhya, June 1, 2017, Accessed June 10, 2021., https://www.analyticsvidhya.com/blog/2017/06/transfer-learning-the-art-of-fine-tuning-a-pre-trained-model/.

[31] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[32] Trung Anh Dang, "Top 10 cnn architectures every machine learning engineer should know from 2012 to present," towardsdatascience, January 5, 2021 Accessed June 20, 2021, https://towardsdatascience.com/top-10-cnn-architectures-every-machine-learning-engineer-should-know-68e2b0e07201.

[33] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, Red Hook, NY, USA, 2012, NIPS'12, p. 10971105, Curran Associates Inc.

[34] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[35] *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, ICML'15. JMLR.org, 2015.

[36] *Rethinking the inception architecture for computer vision*, 2016.

[37] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[38] Babak Ehteshami Bejnordi, Mitko Veta, Paul Johannes van Diest, Bram van Ginneken, Nico Karssemeijer, Geert Litjens, Jeroen A. W. M. van der Laak, , and the CAMELYON16 Consortium, "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer," *JAMA*, vol. 318, no. 22, pp. 2199–2210, 12 2017.

[39] Zahid Younas Khan and Zhendong Niu, "Cnn with depthwise separable convolutions and combined kernels for rating prediction," *Expert Systems with Applications*, vol. 170, pp. 114528, 2021.

[40] Adrian Rosebrock, "Breast cancer classification with keras and deep learning," pyimagesearch, February 18, 2019, Accessed 20 june 2021, https://www.pyimagesearch.com/2019/02/18/breast-cancer-classification-with-keras-and-deep-learning/?5B8353D6-B267-2F42-8C58-DAE1ACC94500_kav_cup_C6FA3ED5_6D17_47D1_B6E2_F4B02CC905E0_.

[41] Vladimir Vapnik, Steven E. Golowich, and Alexander J. Smola, "Support vector method for function approximation, regression estimation and signal processing," in *Advances in Neural Information Processing Systems 9, NIPS, Denver, CO, USA, December 2-5, 1996*, 1996, pp. 281–287.

[42] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.

Abstract

The automated classification invaded the world of health, but some diseases are still diagnosed manually as is the case of medulloblastoma (MB) cancer, which is a high-risk malignant tumor in the Central Nervous System (CNS). This tumor has many types and is diagnosed by biopsy by examining histological images, which takes a lot of effort and time. In our master's project, we propose to perform an automated Childhood medulloblastoma classification based on Convolutional Neural Networks (CNN) by exploiting the knowledge learned through different bases (medical PatchCamelyon (PCam) or Nature ImageNet) with transfer learning. A comparative study is carried out by following two strategies: deep feature extractor and partial fine-tuning, we applied the most popular architectures namely VGG-16, VGG-19, ResNet-50, Inception V3, CancerNet. Experiments prove that there is an important role of similarity when dealing with deep features extractor or partial fine-tuning. Indeed, when there is a high similarity between the two data, the deep features extractor is the best, otherwise, partial fine-tuning is more suitable. In our case application, results demonstrate that the process of partial fine-tuning using inception V3 pre-trained on ImageNet achieved the best result with an accuracy of 99.16%.

Résumé

La classification automatisée a envahi le monde de la santé, mais certaines maladies sont encore diagnostiquées manuellement comme c'est le cas du cancer du médulloblastome (MB) qui est une tumeur maligne à haut risque du système nerveux central (SNC). Cette tumeur a de nombreux types et est diagnostiquée par biopsie en examinant des images histologiques, ce qui demande beaucoup d'efforts et de temps. Dans notre projet de fin d'études de Master, nous proposons d'effectuer une classification automatisée du médulloblastome de l'enfance basée sur les réseaux de neurones convolutifs (CNN) en exploitant les connaissances apprises à travers différentes bases (PatchCamelyon (PCam) médicale ou Nature ImageNet) avec apprentissage par transfert. Une étude comparative est réalisée en suivant deux stratégies : extraction de caractéristiques du réseau et apprentissage partiel du réseau, nous utilisons les architectures les plus courantes à savoir VGG-16, VGG-19, ResNet-50, Inception V3, CancerNet. Les expérimentations réalisées prouvent que la similitude joue un rôle important lorsqu'il s'agit d'approche d'extraction de caractéristiques du réseau ou d'apprentissage partiel du réseau. En effet, lorsqu'il y a une forte similarité entre les deux données, l'extraction de caractéristiques du réseau est la meilleure stratégie, sinon un réglage fin partiel est plus adapté. Dans notre cas d'application, les résultats démontrent que le processus d'apprentissage partiel du réseau en utilisant Inception V3 pré-entraîné sur ImageNet a obtenu les meilleures performances avec une précision de 99,16%.