



جامعة العربي التبسي - تبسة
Université Larbi Tébessi - Tébessa

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la
recherche scientifique

Université Larbi Tébessi - Tébessa



كلية العلوم الدقيقة وعلوم الطبيعة والبيئة
FACULTÉ DES SCIENCES EXACTES
ET DES SCIENCES DE LA NATURE ET DE LA VIE

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département : Mathématiques et Informatique

Mémoire de fin d'études
Pour l'obtention du diplôme de *MASTER*
Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Systèmes d'informations

Thème

**Deep learning pour la mise en évidence
du diagnostic en anatomie pathologique
à base d'immunohistochimie**

Présenté Par :
Benmabrouk Yasmina

Devant le jury :

Dr.Haouam.Y	MCB	Université Larbi Tébessi	Président
Dr.Zemmar.A	MAA	Université Larbi Tébessi	Examineur
Dr.Gasmi.M	MCB	Université Larbi Tébessi	Encadreur
Dr.Bendib.I	MCB	Université Larbi Tébessi	Co-Encadreur

Date de soutenance : 28/06/2020

REMERCIEMENT

Tout d'abord je tiens à remercier ALLAH, le tout puissant, qui m'a donné la force, l'intelligence et la patience d'accomplir ce modeste travail.

Je voudrais exprimer mes plus profonds remerciements à mes encadreurs Dr.Gasmi Mohamed et Dr. Bendib Issam pour leur disponibilité et leur aide et conseils pendant toute la préparation de ce travail.

Je tiens aussi à remercier tous les membres de jury : Dr.Haouam Med Yassine et Dr.Zemmar Ammar, pour leur disponibilité et acceptation d'examiner et de rapporter mon travail.

Je remercie ainsi que tous les enseignants de département Mathématique et informatique.

Je remercie mes chers parents qui m'ont toujours soutenu et encouragé, sans eux je n'y serai jamais arrivé. Je remercie aussi mes chères sœurs Chayma et Zineb, pour leur soutien et leur aide à comprendre dans le domaine médical d'application de ce travail.

Je remercie mes amies Safa, Amani, Khaoula, Chayma et ma chère cousine Hiba qui ont toujours été là pour moi. Leur soutien inconditionnel et leurs encouragements ont été d'une grande aide.

Enfin, Que tous ceux qui directement ou indirectement m'ont apporté leur aide, trouvent ici l'expression de mes sincères remerciements.

Dédicace

*À mes anges : **mes parents**, pour leur encouragement, et leur soutien, pour tous les sacrifices qu'ils me consentent, et leur amour infini. Aucune dédicace ne saurait exprimer mes sentiments, que Dieu les préserve et leur procure santé et longue vie.*

*À mes chères sœurs : **Chayma** et **Zineb**, je vous souhaite une vie pleine de joie, de bonheur et de réussite...♥.*

المخلص

في المجال الطبي، تعد مرحلة التشخيص هي المرحلة الأكثر أهمية، لأن عملية العلاج بأكملها تستند على هذه الخطوة. تتطلب أمراض الأورام مثل سرطان الثدي دراسة تشريحية دقيقة مصحوبة بفحوصات إضافية مثل تلوين الكيمياء المناعية، والتي يتمثل عيبها الرئيسي في المعالجة اليدوية للنتائج التي تتطلب الكثير من الوقت والجهد، كما انها غير دقيقة بما فيه الكفاية، مما يؤثر سلبا على الخطة العلاجية.

في هذا العمل هدفنا هو تسريع وإعطاء المزيد من الدقة في عملية الكيمياء المناعية عن طريق المعالجة الأوتوماتيكية او الالية. وتستند هذه الأخيرة على تصنيف شرائح الفحص التشريحي لمختلف عينات سرطان الثدي، باستخدام الشبكات العصبونية الالتفافية ونقل التعلم، مما ساهم في تحقيق نموذجنا نسبة دقة تقدر ب 97.27%.

الكلمات المفتاحية: التعلم العميق، علم التشريح المرضي، علم الكيمياء المناعية، سرطان الثدي، الشبكات العصبونية الالتفافية.

ABSTRACT

In the medical field, the diagnostic phase is the most important since the whole treatment process is based on this step. Oncological diseases, such as breast cancer require a precise anatomopathological study accompanied by additional examinations, such as immunohistochemical staining. Unfortunately, its major drawback is the loss of time and effort in the manual processing of the results, which has a negative influence on the treatment.

In this work our objective is to accelerate and precise the process of immunohistochemistry by an automatic treatment. Which is based on the classification of anatomopathological examination slides of the breast cancer process, using convolutional neural networks and transfer learning, thanks to this our model achieves 97.27% accuracy.

Keywords: Deep learning, pathological anatomy, immunohistochemistry, breast cancer, CNN

RÉSUMÉ

Dans le domaine médical, la phase de diagnostic est la plus importante, car l'ensemble du processus de traitement sera basé sur cette étape. Les maladies oncologiques comme le cancer du sein nécessitent une étude anatomopathologique précise accompagnée par des examens complémentaires tels que la coloration immunohistochimiques dont l'inconvénient majeur est la perte du temps et d'efforts dans le traitement manuel des résultats obtenus, ce qui influence négativement sur l'étape de traitement.

Dans ce travail notre objectif est d'accélérer et donner plus de précision dans le processus d'immunohistochimie par un traitement automatique. Ce dernier est basé sur la classification des lames d'examen anatomopathologique du processus cancéreux mammaire, en utilisant les réseaux de neurones convolutionnels et l'apprentissage par transfert, grâce à cela notre modèle atteint la précision de 97,27%.

Les mots clés: Deep learning, anatomie pathologique, immunohistochimie, CNN, cancer du sein

Table des matières

Introduction générale	1
Chapitre 1: L'anatomie pathologique et le cancer du sein	3
1. Cancer du sein	3
1.1. Description	3
1.2. Les causes	4
1.3. Symptômes et Complications	5
1.4. Diagnostic	5
1.5. Traitement et Prévention	6
2. L'étude anatomo-pathologique	7
2.1. Définitions et généralités	7
2.1.1. Anatomie pathologie	7
2.1.2. L'histologie.....	7
2.1.3. La cytologie	8
2.1.4. La lésion	8
2.1.5. L'étude anatomo-pathologique du cancer du sein	8
2.2. Les prélèvements en pathologie mammaire	8
2.3. Prélèvements, gestion et interprétation	9
2.4. Les modalités d'étude anatomopathologique des cancers du sein	10
2.4.1. Etude morphologique des tissus et des cellules	10
2.4.2. Etude du génotype des cellules	10
2.4.3. Immunohistochimie.....	10
Chapitre 2: L'apprentissage Profond (Deep Learning)	14
1- L'apprentissage profond	14
1.1. L'apprentissage profond et les approches d'apprentissages	14
1.2. Historique de l'apprentissage profond	15
1.3. Fonctionnement de deep learning	16
1.4. Les architectures de l'apprentissage profond	16
1.4.1. Les réseaux de neurones récurrents (RNN)	16
1.4.2. Le réseau de neurones convolutionnel (CNN)	17
1.5. Transfer Learning : (l'apprentissage par transfert)	19
1.6. Les modèles du CNN	21
2. La classification	25
2.1. Les méthodes de classification	25
2.2. La classification des images	26
3. Travaux connexes	27
3.1. Classification par apprentissage automatique	27
3.2. Classification par apprentissage profond	28
Chapitre 3 : Conception et implémentation d'un modèle par le transfer learning	30
1. Base de données	30
1.1. Base de données1.....	30
1.2. Base de données2.	30
2. Prétraitement de base de données	32
3. Le modèle pré-entraîné	34
3.1. SE-ResNet	34
3.2. Small SE-ResNet	34
3.3. Chargement du modèle pré-entraîné	35
4. Modèle de transfert proposé	36

4.1.	La classification avec un modèle CNN près-entraînés	36
4.2.	Méthode 1	38
4.3.	Méthode 2	41
4.4.	Méthode 3	44
4.5.	Méthode 4	45
4.6.	Méthode 5	46
5.	Discussion	48
6.	Génération automatique de légende.....	50
	Conclusion générale	53
	Bibliographie	55
	Annexe.....	58

Liste des abréviations

ANN	Artificial Neurone Network
BoVW	Bag Of Visual Words
CAD	Computer-aided Diagnosis
CNN	Convolution Neurones Networks
GLCM	Gray Level Co-occurrence Matrix
ILSVRC	ImageNet Large-Scale Visual Recognition Challenge
LBP	Local Binary Patterning
LPQ	Local Phase Quantization
PCA	Principal Component Analysis.
PFTAS	Parameter Free TAS
ReLU	Rectified Linear Unit
ResNet	Residual neural Network
RF	Random Forest
RNN	Recurrent Neural Networks
SENet	Squeeze-and-Excitation Network
SIFT	Scale Invariant Feature Extraction
SVM	Support Vector Machine
TAS	Threshold Adjacency Statistics
VGG	Visual Geometry Group

LISTE DES FIGURES

Figure 1: Photographies de quelques étapes de l'immunohistochimie.....	13
Figure 2: Parcours de filtre sur une image d'entrée	18
Figure 3: Max Pooling avec un filtre 2x2 et un pas de 2.....	18
Figure 4: Couche de correction Relu	19
Figure 5: Réutilisation de la couche pré-entraînée.....	20
Figure 6: L'apprentissage résiduel.....	23
Figure 7: Module SE-Inception (gauche) et Unité SE-ResNet (droite)	24
Figure 8: Un bloc SE effectue un recalibrage de la carte des caractéristiques.....	24
Figure 9: Architecture de SE Block	25
Figure 10: Classification d'images.a Méthode de classification traditionnelle.b Apprentissage profond.....	26
Figure 11: Échantillons d'images de la base de données BreaKHis	31
Figure 12: Échantillons d'images de la base de données (E-cadhérine positive).....	31
Figure 13: Échantillons d'images de la base de données (E-cadhérine négative).....	32
Figure 14: Génération de nouvelles images à partir de celles existantes.....	32
Figure 15: Les arguments de « ImageDataGenerator » utilisés	33
Figure 16: Partie de code d'augmentation de données.....	33
Figure 17: Architecture de module Small SE-ResNet	35
Figure 18: L'architecture BHCNet-3 pour la classification bénigne et maligne des images histopathologies du cancer du sein.....	36
Figure 19: Graphe de précision et de perte de la méthode 0.....	37
Figure 20: Architecture du transfert d'apprentissage avec la méthode 1.....	38
Figure 21: Geler les couches superficielles du modèle.....	39
Figure 22: Graphe de précision et de perte de la méthode 1 pour batch_size=64, epochs=100.	39
Figure 23: Graphe de précision et de perte de la méthode 1 pour batch_size=128, epochs=100.	40
Figure 24: Graphe de précision et de perte de la méthode 1 pour batch_size=16, epochs=50.....	40
Figure 25 : Changement des deux dernières couches du modèle BHCNet	41
Figure 26 : Architecture du transfert d'apprentissage avec la méthode 2.a	41
Figure 27 : Changement des couches.....	42
Figure 28: Graphe de précision et de perte de la méthode 2.a.....	43
Figure 29: Architecture du transfert d'apprentissage dans le cas de geler les couches entre 4 et 10 avec la méthode 2.b	43
Figure 30: Graphe de précision et de perte de la méthode 2.b.	44
Figure 31: Architecture du transfert d'apprentissage avec la méthode 3	44
Figure 32: Graphe de précision et de perte de la méthode 3.....	45
Figure 33: Graphe de précision et de perte de la méthode 4.....	46
Figure 34: Architecture du transfert d'apprentissage avec la méthode 5.....	46
Figure 35: Graphe de précision et de perte du modèle de cas 02 de la méthode 5... ..	47
Figure 36: Prédit: E-cadhérine positive (91.76%). Original : E-cadhérine positive.....	50
Figure 37: Prédit: E-cadhérine négative (88.68%). Original : E-cadhérine négative	50
Figure 38: Prédit: E-cadhérine positive (91.88%). Original : E-cadhérine positive	51

Figure 39: Prédit: E-cadhérine négative (82.38%). Original : Non décidable 51
Figure 40: Prédit: E-cadhérine négative (88.61%). Original : E-cadhérine négative.....51
Figure 41: Prédit: E-cadhérine négative (84.74%). Original : E-cadhérine négative.....52

LISTE DES TABLEAUX

Table 1: La division de notre data set	33
Table 2: Taux d'apprentissage pour la classification avec le modèle BHCNet sans aucune modification	36
Table 3: Taux d'apprentissage en changeant batch_size et epochs pour une classification avec le modèle BHCNet.....	37
Table 4: Taux d'apprentissage en changeant batch_size et epochs pour une classification avec la méthode 1	39
Table 5: Taux d'apprentissage en changeant batch_size et epochs pour la classification avec la méthode 2.a	42
Table 6: Taux d'apprentissage en changeant batch_size et epochs pour la classification avec la méthode 2.....	44
Table 7: Taux d'apprentissage en changeant batch_size et epochs pour la classification avec La méthode 3	45
Table 8: Taux d'apprentissage en changeant le nombre de neurone des couches « fully connected » avec la méthode 5.....	47
Table 9: Les meilleurs taux de précision obtenus de chaque méthode	50

INTRODUCTION GÉNÉRALE

Le cancer constitue un problème majeur de santé publique dans le monde, présentant des taux de mortalité élevés. Surtout dans la population féminine, le cancer du sein constitue l'affection tumorale maligne la plus fréquemment rencontrée chez la femme et la plus meurtrière. Son diagnostic repose toujours sur le trépied histo-radio-clinique et sa prise en charge doit absolument être multidisciplinaire.

Concernant la thérapeutique, le choix du traitement adjuvant, après l'exérèse d'un cancer du sein, repose sur un ensemble de critères pronostiques tels que l'âge de la patiente, la taille de la tumeur, son grade, son stade, la présence de métastases [1]...etc. Mais le plus important c'est l'étude anatomopathologique qui confirme la classification TNM (Tumeur, Nœuds et Métastases) du cancer et dans un second rôle aide les oncologues dans leurs choix thérapeutiques par la thérapie ciblée selon les résultats immunohistochimiques.

L'immunohistochimie (IHC) a apporté un confort appréciable aussi bien pour confirmer le diagnostic de malignité, vu qu'elle donne une information d'ordre pronostique et thérapeutique. Les marqueurs immunohistochimiques tels que « E-cadhérine » sont utilisés selon des directives spécifiques où l'intensité des taches et le nombre de cellules colorées positivement sont exprimés en pourcentage de toutes les cellules malignes. L'interprétation des colorations IHC est rendue par un pathologiste formé à l'aide d'une méthode manuelle, qui consiste à compter chaque cellule colorée positivement et négativement au microscope [2]. Face à cette tâche intimidante et au manque de temps, certains pathologistes reviennent à une estimation visuelle de ce rapport. Comme prévu, l'estimation manuelle du rapport souffre d'une mauvaise reproductibilité même entre les mains de pathologistes experts.

Il est démontré que l'utilisation du diagnostic assisté par ordinateur (CAD) pour classer automatiquement les images histopathologiques peut non seulement améliorer l'efficacité du diagnostic, mais également fournir aux médecins des résultats de diagnostic plus objectifs et plus précis. Les techniques d'apprentissage profond, telles que les réseaux neuronaux convolutionnels (CNN), ont mis en évidence un grand succès dans la détection des cellules mitotiques à partir d'images histopathologiques colorées à l'hématoxyline et à l'éosine [3]. Cela constitue une précieuse source d'inspiration pour le développement d'un algorithme de

Introduction générale

classification des tissus cancéreux à partir d'images histopathologiques colorées avec le « E-cadhérine »

Notre but dans ce projet est de classifier des images histopathologiques colorées avec le E-cadhérine afin d'aider le pathologiste de préciser la sensibilité des tissus cancéreux aux anticorps utilisés afin de faciliter le travail aux oncologues pour l'utilisation du traitement le plus approprié.

Dans notre travail on a utilisé l'apprentissage par transfert pour créer un modèle de classification dédié aux lames anapaths des tissus mammaires cancéreux traités par des anticorps E-cadhérine. Le modèle d'apprentissage par transfert est créé à partir d'un modèle pré-entraîné sur une base de données d'images histopathologies réelles (BreakHis). Le nouveau modèle est à leur tours ré-entraîne sur une dataset traité par E-cadhérine proposé et validé par des pathologistes de CHU Constantine.

Pour ce faire, on a organisé notre mémoire en trois chapitres :

- Dans le premier chapitre nous avons consacré le domaine d'application, pour expliquer grosso modo le cancer du sein, ces causes, son diagnostic et traitement ainsi que l'étude anatomo-pathologique et le protocole d'immunohistochimie.
- Le deuxième chapitre a été consacré à la description de fonctionnement du Deep Learning et les réseaux de neurones convolutionnels. Nous avons parlé des différents types de classifications, ainsi que quelques travaux connexes.
- Dans le troisième chapitre, nous avons montré la partie expérimentale de notre travail. Nous avons détaillé l'architecture et les paramètres choisis et discuter les différents résultats obtenus.

L'anatomie Pathologique Et Le Cancer Du Sein

Introduction

Le cancer du sein est le premier cancer chez la femme et représente la première cause de mortalité dans le monde. Il est extrêmement rare chez l'homme. Cette pathologie peut survenir à tout âge. Les stratégies thérapeutiques de cette maladie sont délivrées sur la base des facteurs pronostiques et prédictifs de la réponse thérapeutique. Elles sont guidées par l'utilisation de la classification clinique et anatomopathologique.

L'utilisation des outils de diagnostic permet de détecter ces cancers. La détermination de ces lésions se fait par une étude histologique qui correspond à l'examen macroscopique et microscopique des prélèvements (biopsie, mastectomie, tumorectomie). Des études immunohistochimiques sont nécessaires dans le but d'apporter des compléments d'information plus précis sur la nature de la lésion.

Dans ce premier chapitre, nous allons expliquer grosso modo le cancer du sein, ces causes, son diagnostic et traitement ainsi que l'étude anatomo-pathologique.

1. Cancer du sein

1.1.Description

Le sein féminin est composé de glandes mammaires et de canaux galactophores (servant à acheminer le lait vers le mamelon) qui sont entourés de tissus graisseux et de tissu conjonctif. Une croissance anarchique des cellules dans l'un ou l'autre de ces tissus peut provoquer le cancer du sein.[4]

Le cancer est une maladie génétique de la cellule animale, qui a pour origine une modification quantitative et/ou qualitative des gènes, aboutissant au dysfonctionnement des mécanismes de régulation cellulaire, qui se traduit par la transformation d'une cellule normale en cellule tumorale maligne [5] .

Le cancer du sein est la forme de cancer féminin la plus répandue. Il se produit souvent à un âge plus précoce que les autres types de tumeurs. La probabilité de cancer du sein augmente de façon spectaculaire avec l'âge : 13 femmes sur 1 000 femmes quadragénaires ; 23 femmes sur 1 000 femmes quinquagénaires ; 29 femmes sur 1 000 femmes sexagénaires ; 31 femmes sur 1 000 femmes septuagénaires.[4]

Chapitre 1 : l'anatomie pathologique et le cancer du sein

Les cas de cancer du sein masculin sont rares et représentent moins de 1 % de tous les cas. Par contre, le cancer du sein féminin occupe le 2e rang parmi les causes de décès lié au cancer [6].

1.2. Les causes

Les cellules de l'organisme se reproduisent pendant toute la durée de la vie pour remplacer de manière ordonnée les cellules des tissus qui se sont détériorés. Le cancer du sein, comme toutes les formes de cancer, se produit lorsque ce processus est compromis et que les cellules démarrent à se diviser à une cadence inhabituellement rapide. Il n'existe pas une raison unique ni un seul déclencheur du cancer du sein. En fait, certains facteurs de risque qui accroissent les chances de cette maladie ont été définies: [4]

- L'âge – le cancer du sein atteint plus fréquemment les femmes de plus de 50 ans;
- Les antécédents familiaux – si la mère ou la sœur d'une femme a été frappée de cette affection avant la ménopause, 1 ou 2 gènes liés au cancer du sein pourraient être en cause;
- Des antécédents de cancer du sein;
- Des antécédents familiaux de cancer des ovaires;
- L'âge au moment de la première grossesse – les femmes qui n'ont pas eu d'enfant ou celles qui ont donné naissance à leur premier enfant quand elles avaient plus de 30 ans;
- L'âge auquel les menstruations se sont établies – des menstruations survenant à un jeune âge (soit moins de 12 ans);
- L'apparition tardive de la ménopause – l'interruption du cycle menstruel après l'âge 55 ans accroît le risque de cancer du sein ;
- Le tabagisme – des études récentes semblent indiquer que le risque de cancer du sein avant l'âge de 50 ans est 70 % plus élevé pour les femmes qui commencent à fumer de façon régulière dans les 5 ans qui suivent l'apparition des menstruations que pour les non-fumeuses ;
- Une radiothérapie au thorax, surtout si elle a été reçue avant l'âge de 30 ans;
- La consommation d'alcool joue un rôle multiplicatif dans le risque d'un cancer du sein;
- Une hormonothérapie substitutive (œstrogènes et progestérone) accroît légèrement le risque après 5 ans de traitement ;
- Les contraceptifs oraux augmentent légèrement le risque s'ils sont pris pendant plusieurs années ;
- Une obésité qui s'accompagne d'une consommation excessive de calories et de matières grasses.

Chapitre 1 : l'anatomie pathologique et le cancer du sein

L'augmentation du risque de cancer du sein liée aux facteurs précédemment évoqués est toujours assez faible sur la stratégie statistique. En fait, pour la majorité des femmes, l'unique facteur de crainte est leur âge, quand elles ont plus tout juste la cinquantaine. Toute crainte à ce propos doit être évoquée avec le médecin.

1.3.Symptômes et Complications

L'apparition d'une bosse ou d'une masse dans un sein est découverte 9 fois sur 10 au cours d'un auto-examen. Il n'y a généralement aucune douleur, mais la masse peut provoquer une sensation inhabituelle à l'endroit où elle est située.

Lorsqu'une tumeur apparaît dans les canaux galactophores, la taille ou la forme du sein peut changer. En outre, le mamelon peut entrer vers l'intérieur ou la peau peut se rétracter, entraînant ainsi la formation d'une fossette. L'apparition d'une bosse ou d'une enflure dans l'aisselle, ou encore d'une rougeur ou d'un gonflement du sein sont d'autres symptômes.

Un sein sur lequel apparaît une éruption cutanée rougeâtre doit être évalué pour éliminer la présence d'un cancer du sein, même en l'absence d'une bosse.[4]

Ces transformations peuvent constituer des signes précoces du cancer du sein, mais elles peuvent aussi indiquer un autre trouble non cancéreux. En fait, environ 8 masses sur 10 sont non cancéreuses. Il faut toutefois consulter un médecin pour repérer précisément la cause exacte de la bosse.

La complication la plus grave du cancer du sein est la formation de métastases. Les métastases apparaissent lorsque certaines des cellules d'une tumeur se détachent et se déplacent vers d'autres parties du corps, en passant par les vaisseaux sanguins ou les vaisseaux lymphatiques, pour envahir les tissus à des endroits parfois éloignés. Lorsque des métastases sont formées à partir de cellules du cancer du sein, elles se situent le plus souvent dans les ganglions lymphatiques, les poumons, le foie, les os, le cerveau et la peau. Plusieurs années peuvent s'écouler – même après le diagnostic et le traitement du cancer du sein – avant l'apparition du cancer disséminé à partir de la tumeur initiale.

Au moment où les métastases sont découvertes, le cancer s'est probablement déjà disséminé à d'autres endroits, même si ces tumeurs ne sont pas dépistées.[4]

1.4.Diagnostic

Il faut toujours faire examiner par votre médecin une masse suspecte dans un sein. En plus d'un examen physique complet, le médecin voudra également procéder à une mammographie, c'est-à-dire à une radiographie à faible dose qui peut révéler la présence d'anomalies dans le

Chapitre 1 : l'anatomie pathologique et le cancer du sein

sein. Une échographie peut parfois être demandée afin d'établir si la bosse est un kyste rempli de liquide ou une masse solide. Une biopsie pourrait être nécessaire, même si la bosse décelable au toucher n'apparaît pas sur la mammographie.

Du liquide peut être prélevé dans un kyste à l'aide d'une aiguille afin de déterminer s'il contient des cellules cancéreuses. Une masse solide doit être analysée au microscope. À cette fin, il faut réaliser une biopsie à l'aiguille, une biopsie au trocart ou une biopsie chirurgicale.[4]

Dans le cas de la biopsie à l'aiguille, une aiguille fine est introduite dans la bosse pour prélever des cellules. La biopsie au trocart consiste à retirer un échantillon de tissu et de cellules mammaires afin de l'examiner. La biopsie chirurgicale consiste à enlever toute la masse ainsi qu'une certaine quantité de tissu sain avoisinant. Cette intervention est réalisée sous anesthésie locale ou générale.[4]

Les cellules cancéreuses peuvent être soumises à un examen visant à déceler la présence de récepteurs des œstrogènes ou de la progestérone. En outre, on peut procéder à d'autres examens – y compris la radiographie, la scintigraphie osseuse, l'échographie, la tomodensitométrie ou l'imagerie par résonance magnétique – afin de vérifier si le cancer s'est disséminé à d'autres organes ou tissus.

1.5.Traitement et Prévention

Le dépistage précoce du cancer du sein réduit la probabilité de dissémination du cancer et augmente les chances de guérison complète.

Le traitement du cancer du sein dépend de circonstances personnelles comme la vitesse de croissance de la tumeur, la réponse au traitement et l'absence ou la présence de disséminations. Les options thérapeutiques comprennent l'intervention chirurgicale, la radiothérapie, la chimiothérapie et le traitement médicamenteux (y compris l'hormonothérapie et la thérapie biologique).[4]

- L'intervention chirurgicale consiste à enlever une partie ou la totalité du sein. L'ablation d'une petite partie du sein (c'est-à-dire la tumeur et quelques tissus avoisinants) est appelée tumorectomie du sein, ou mastectomie partielle si une partie plus importante du sein est enlevée. Une mastectomie simple consiste à enlever le sein en entier, et une mastectomie radicale inclut également les tissus et les muscles thoraciques sous-jacents. Les ganglions lymphatiques dans les aisselles peuvent aussi être enlevés. Bon nombre de femmes ont l'option d'avoir une reconstruction du sein, soit au moment de l'intervention chirurgicale ou plus tard.[4]

Chapitre 1 : l'anatomie pathologique et le cancer du sein

- La chimiothérapie est un traitement général (appelé traitement systémique). Elle regroupe des médicaments dits anti-tumoraux administrés par injection, qui agissent sur les mécanismes de la division cellulaire et autres mécanismes. Elle nuit à la croissance des cellules cancéreuses, mais elle perturbe aussi les cellules saines.
- Le traitement hormonal (par ex. le tamoxifène*, ou la classe de médicaments appelés inhibiteurs de l'aromatase) C'est un traitement qui consiste à cibler les récepteurs hormonaux afin d'empêcher l'action stimulante des hormones féminines sur les cellules cancéreuses. Il peut être utilisé pendant une période de temps pouvant aller jusqu'à 5 ans pour les femmes post ménopausées dont le cancer possède des récepteurs hormonaux. Dans de rares cas, une thérapie générale est administrée avant l'intervention chirurgicale. Cette méthode est désignée thérapie néo adjuvante. On utilise ce type de traitement pour améliorer les chances de pouvoir pratiquer une tumorectomie plutôt qu'une mastectomie ou pour contrôler un cancer qui touche une portion importante du sein
- Une radiothérapie est souvent administrée après une tumorectomie ou une mastectomie partielle afin de détruire les cellules cancéreuses en les empêchant de se multiplier. Elle consiste à diriger des rayonnements ionisés de façon localisée et ciblée sur la zone à traiter.
- La thérapie biologique nuit à la croissance des cellules cancéreuses et aide l'organisme à détruire ces cellules. En règle générale, elle est employée pour traiter un cancer du sein lié à un excès de la protéine appelée HER2.

2. L'étude anatomo-pathologique

2.1. Définitions et généralités

2.1.1. Anatomie pathologie

L'anatomie pathologique (ou anatomo-pathologie) est la discipline médicale qui permet la reconnaissance des anomalies des cellules et des tissus d'un organisme, appelées lésions, pour effectuer le diagnostic des maladies, porter un pronostic et, plus généralement, en comprendre les causes et les mécanismes.[7]

2.1.2. L'histologie

Autrefois appelée **anatomie microscopique**, est la branche de la biologie et de la médecine qui étudie les tissus biologiques. Elle a pour objectif d'explorer la structure des organismes vivants, les rapports constitutifs et fonctionnels entre leurs éléments fonctionnels,

Chapitre 1 : l'anatomie pathologique et le cancer du sein

ainsi que le renouvellement des tissus. Elle participe à l'exploration des processus pathologiques et de leurs effets. [7]

2.1.3. La cytologie

Est l'étude des cellules isolées. Elle se différencie de l'histologie, qui est l'étude morphologique des tissus biologiques. Il s'agit de l'étude des cellules normales ou pathologiques (cytopathologie), ainsi que de leur aspect morphologique ou biochimique.[7]

2.1.4. La lésion

On peut définir la lésion comme une altération morphologique décelable par un quelconque moyen d'observation. Elle peut trouver comme une conséquence de processus pathologique [7]

2.2. L'étude anatomo-pathologique du cancer du sein

L'étude anatomo-pathologique des cancers du sein est restée longtemps limitée à leur étude morphologique classique à l'œil (macroscopie) et au microscope. Elle bénéficie maintenant de techniques supplémentaires telles que l'immunohistochimie et la biologie moléculaire. Les résultats obtenus par ses différentes techniques sont complémentaires et permettent de caractériser les cancers selon des classifications internationales validées par consensus professionnel et en permanence actualisées.

Le diagnostic du pathologiste apporte des informations déterminantes qui conditionnent les modalités de prise en charge thérapeutique du cancer et de son suivi

2.3. Les prélèvements en pathologie mammaire

L'anatomie pathologique est basée sur l'étude morphologique conventionnelle des prélèvements menés principalement au microscope optique.

Il existe 3 types de prélèvements en pathologie mammaire : [8]

- a. La cytoponction** : ponction à l'aiguille d'une zone suspecte du sein, repérée cliniquement et/ou par mammographie, pour recueillir des cellules.
- b. La pièce opératoire** : obtenue par exérèse chirurgicale de la tumeur
- c. La biopsie** : Une biopsie est un examen qui consiste à prélever un fragment de tissu ou d'organe, au moyen d'une aiguille ou par frottis. Le morceau retiré fait l'objet d'un examen microscopique et/ou d'une analyse biochimique pour déceler diverses anomalies ou

maladies. La biopsie permet aussi d'évaluer la progression de la maladie si celle-ci est connue et de décider un traitement adapté.

Elle permet également de suivre l'évolution de la maladie, de l'évaluer, de savoir si elle existe et de prendre les mesures nécessaires au traitement.

2.4. Prélèvements, gestion et interprétation

Les étapes de l'examen anatomopathologique sont en fonction du type de prélèvement. Ces prélèvements peuvent être variés, allant des seules cellules (cytologie) à la pièce d'amputation mammaire, chacun nécessitant une gestion technique particulière dont le bon respect conditionne la fiabilité de son interprétation microscopique.[9]

Les divers prélèvements sont soumis aux techniques histologiques usuelles, qui comportent trois temps :

-La fixation : les pièces reçues sont fixées dans un liquide fixateur (le plus recommandé est formol dilué à 10%). La durée de fixation dépend de la taille du prélèvement de 24h à 72h. Il faudrait utiliser une quantité suffisante de fixateur afin qu'il soit 10 fois plus volumineux que la pièce. La fixation permet de conserver l'architecture tissulaire et cellulaire de façon aussi proche que possible de leur aspect à l'état vivant. C'est une étape primordiale car elle détermine la qualité de tout le processus, la fiabilité du diagnostic et la rapidité du résultat. [9]

- Le découpage des pièces opératoires doit être très précis :

- . En ce qui concerne le curage ganglionnaire axillaire, une dissection soignée de la graisse après une fixation correcte permet l'identification d'une quinzaine de ganglions en moyenne par curage, et l'inclusion en totalité de chacun selon des tranches macroscopiquement sériées.
- . Sur la pièce de tumorectomie, plusieurs tranches sont prélevées sur la tumeur, et le tissu péri-tumoral est examiné pour apprécier les limites de l'exérèse.
- . Au niveau de la pièce de mastectomie, les prélèvements effectués sur le mamelon et tous les quadrants permettent d'apprécier la multicentricité du cancer et les lésions bénignes associées.
- . Les pièces d'exérèse partielle sont coupées en coupes macroscopiquement sériées de 2 mm environ. Ces coupes perpendiculaires à l'axe de la pièce sont numérotées à partir d'un repère placé par le chirurgien du côté du mamelon, afin de mieux localiser les éventuels foyers tumoraux décelés.[9]

- L'inclusion des fragments tissulaires en paraffine permet la préparation des lames soumises à l'examen microscopique.

2.5. Les modalités d'étude anatomopathologique des cancers du sein

2.5.1. Etude morphologique des tissus et des cellules

- **Examen extemporané** : Le pathologiste réalise un examen au microscope d'un prélèvement pendant l'intervention chirurgicale. Il peut ainsi donner au chirurgien des informations sur la nature de la tumeur, son extension, l'envahissement ou non du ganglion sentinelle préalablement repéré. Ces éléments conditionnent la suite du geste opératoire.
- **Examen macroscopique de la pièce opératoire** : Examen à l'œil nu et à la loupe de la pièce opératoire qui décrit l'aspect des tissus examinés, leurs dimensions...
- **Examen microscopique** : observation au microscope de l'aspect et de la forme des cellules et de leur organisation architecturale afin de déterminer leur nature et de contribuer à la classification du cancer.[4]

2.5.2. Etude du génotype des cellules

Elle étudie par biologie moléculaire le génome de la cellule afin d'identifier des anomalies telles que des mutations ou des amplifications de gènes. Elle permet de préciser le type et le sous type du cancer.

Ces études de phénotype et de génotype des tumeurs complètent et affinent la classification du cancer et apportent de précieuses informations pronostiques et décisives pour la prise en charge thérapeutique initiale (chimiothérapies, traitements adjuvants...). Elles ont un intérêt dans le suivi secondaire de l'efficacité des traitements.

Les échantillons tissulaires sont conservés après analyse au sein de bibliothèques, selon les normes réglementaires en vigueur, offrant ainsi la possibilité de nouvelle analyse ultérieure pour évaluer par exemple la sensibilité ou non de la tumeur à une nouvelle thérapeutique.[4]

2.5.3. Immunohistochimie

Cette technique permet d'identifier et de localiser des protéines spécifiques par la détection d'antigène au moyen d'anticorps sur les cellules d'une coupe de tissu (est préalablement fixée au formol et incluse en paraffine) [10] .

L'immunohistochimie exploite le fait qu'un anticorps se lie spécifiquement à des antigènes dans les tissus. Elle s'effectue sur des lames non colorées (lames « blanches ») après déparaffinage éventuel. Un couple anticorps-antigène peut être visualisé de plusieurs façons. Dans la plupart des cas, un anticorps est conjugué à une enzyme (ex : peroxydase) qui peut catalyser une réaction de production de couleur. Alternativement, les anticorps peuvent aussi être marqués par un fluorophore (immunofluorescence).

Chapitre 1 : l'anatomie pathologique et le cancer du sein

L'immunohistochimie est devenue une technique cruciale et largement utilisée dans les diagnostics cliniques de maladies : elle permet d'orienter un diagnostic en déterminant la famille de cellule à laquelle la tumeur appartient (cellule épithéliale, musculaire, nerveuse,...) et, de manière plus récente, apporte des informations sur la sensibilité des certaines tumeurs à l'un ou l'autre traitement (expression des récepteurs hormonaux, des récepteurs à l'EGF,...). L'étude immunohistochimique permet d'établir une classification de biologie moléculaire des cancers du sein. Les anticorps utilisés dans le cas du cancer du sein sont :

- **E-cadhérine** : qui est le sujet de notre étude, présentant une molécule transmembranaire d'adhésion cellulaire Ca^{2+} -dépendante. Elle joue un rôle important dans la croissance, le développement et l'adhésion intercellulaire des cellules épithéliales. La plupart des tumeurs ont une architecture anormale et toute perte d'adhésion résultante est probablement une étape importante dans le développement de l'invasion locale.

L'E-cadhérine pourrait avoir un rôle dans la progression néoplasique, en particulier en tant que suppresseur de l'invasion. Par exemple l'expression réduite ou l'absence d'E-cadhérine en plus d'alpha, bêta et gamma-caténines dans les carcinomes primitifs du sein a été observée, et ces quatre protéines sont associées au développement de métastases. L'utilité immunohistochimique principale de l'E-cadhérine est de mettre en évidence l'expression différentielle de cette protéine dans les carcinomes lobulaires et canaux. [11]

Il existe d'autres anticorps utiliser pour préciser le diagnostic du cancer du sien telle que :

- **Ki-67** : marqueurs de prolifération.
- **Her2** : marqueurs cytoplasmique.
- **RO** : Récepteur oestrogénique.
- **-RP** : Récepteur progestatif.

Les limitations de la technique sont d'ordre purement pratique, d'où l'importance cruciale d'une formation adéquate des techniciens dans la manipulation et l'utilisation des anticorps d'une part, et d'une expérience minimale du pathologiste dans l'interprétation des images. D'autre part. L'immunohistochimie ayant un prix, il est indispensable de faire le bon choix parmi les très nombreux anticorps disponibles, leur importance diagnostique étant fort variable.

2.5.3.1. Le protocole d'immunohistochimie

Après la préparation des coupes sur des lames silanisées, cette technique se déroule selon le protocole suivant (Figure.1) [10] :

a. Déparaffinage

Cette étape consiste en l'incubation des lames dans une étuve à 62°C pendant 24 heures. Après, les lames sont plongées dans trois bains de xylène pendant 05 minutes chacun. Ensuite, elles sont émergées successivement dans trois bains d'alcool durant 05 minutes chacun. Suivi d'un rinçage par l'eau distillée, ce qui permet de débarrasser les fragments de la paraffine et de les réhydrater.

b. Démasquage thermique

- Les lames sont placées dans un portoir qui est transféré dans un bac de solution de démasquage préchauffée + 250 ml d'eau distillée. Puis mettre les bacs dans un bain marie à 95- 97° C pendant 60 min.

L'intérêt de cette étape est de rompre les liaisons moléculaires créées par le fixateur et donc de restaurer les déterminants antigéniques qui ont été masqués par la fixation au formol.

- Retirer ensuite du bain marie et laisser refroidir dans la même solution pendant 20min.
- Tremper les lames dans l'eau distillée pendant 5 min. Le but étant d'éliminer toute trace de la solution de démasquage.
- Essuyer les lames et encercler les coupes histologiques par Dako pen (pour diminuer la quantité des solutions utilisées et de traiter aisément plus d'une coupe sur une lame) et remettre dans l'eau distillé pour que les lames restent humides.

c. Blocage de peroxydase endogène

Elle consiste à verser quelques gouttes de peroxydase sur toutes les coupes. Pour Her2, laisser réagir 10 min, par contre les autres anticorps on laisse 5 min afin d'éviter le bruit de fond.

- d.** Après, les lames sont rincées à l'eau distillée dans un bac pendant 5 min, ensuite sont immergées dans le PBS (tampon phosphate salin) (50ml du produit + 950ml eau distillée) pendant 5 min pour éliminer la peroxydase endogène. Egoutter les lames verticalement.

e. Marquage à Anti corps primaire :

Quelques gouttes d'Anti corps primaire (environ 100 ml) (RO, RP, Her2, Ki67) sont versées sur les coupes séparément, on laisse incuber 30 min et les lames doivent être recouvertes. Suivi par rinçage dans 2 bains de PBS (5 min chacun) ou simple rinçage.

f. Révélation par le système (HRP)

Chapitre 1 : l'anatomie pathologique et le cancer du sein

Quelques gouttes d'Anticorps secondaire sont ajoutées, les lames doivent être recouvertes et laisser incuber pendant 30 min. Ensuite, les lames sont plongées dans un bain de PBS pendant 5 min.

g. Application du chromogène DAB

Appliquer suffisamment en solution de la préparation suivante (05 ml de diluant + 5 gouttes de DAB) puis recouvrir les éléments histologiques, laisser incuber 5 à 10 min et puis rinçage à eau distillée

h. Contre coloration et montage

Les lames sont immergées dans un bain d'Hématoxyline de Mayer qui colore intensément les noyaux en bleu violacé mais aussi les cytoplasmes et le tissu de soutien (de façon moins intense) pendant quelques seconds suivis d'un rinçage à l'eau distillée. Puis on procède au montage lame-lamelle.



Figure 1: Photographies de quelques étapes de l'immunohistochimie[10]

Conclusion :

Dans ce chapitre, nous avons essayé de clarifier le concept d'anatomie pathologique, et nous sommes allés un peu plus loin sur ce concept afin de comprendre comment le diagnostic du cancer du sein est basé sur cette science. Nous avons essayé de représenter certains des concepts sous le titre d'immunohistochimie qui représente une technique cruciale permet au pathologiste de préciser son diagnostic au maximum, ce qui à son tour, donne au clinicien la possibilité de choisir le traitement optimal pour son patient.

L'apprentissage Profond (Deep Learning)

Introduction

Plusieurs inventions ont été inspirées par la nature. Il semble donc logique de s'inspirer de l'architecture du cerveau pour construire une machine intelligente. C'est l'idée clé qui a déclenché les réseaux de neurones artificiels, qui sont au cœur de l'apprentissage profond. Ils sont puissants et évolutifs, ce qui les rend idéaux pour les tâches d'apprentissage automatique volumineuses et très complexes, telles que la classification de milliards d'images (par exemple, Google Images), la reconnaissance vocale, la recommandation des meilleures vidéos (par exemple, YouTube)...

Dans ce chapitre nous allons présenter tout d'abord le fonctionnement de l'apprentissage profond et leurs différentes architectures, ainsi que les différents techniques dédiées la classification des images.

1. L'apprentissage profond

1.1.L'apprentissage profond et les approches d'apprentissages

L'apprentissage profond est un sous type de l'apprentissage automatique qui vise à former et à enseigner la machine à exécuter des fonctions humaines telles que la distinction d'objets visuels et l'identification du son et de l'image. Au lieu d'organiser les données, l'apprentissage profond définit ses propres paramètres de base qui permettent à la machine d'apprendre de manière indépendante.

Le deep learning, ou apprentissage profond, est une partie de la grande famille de l'intelligence artificielle. Plus précisément, il constitue un sous-ensemble de la machine learning, et fait appel à des types particuliers de réseaux de neurones artificiels. Il présente donc des caractéristiques similaires à ces techniques, principalement la capacité d'apprentissage de façon autonome. Pour cela on a trois approches d'apprentissage selon le cas d'utilisation :

- Apprentissage supervisé
- Apprentissage non supervisé
- Apprentissage par renforcement

Chapitre 2 : L'apprentissage profond (Deep learning)

- **Apprentissage supervisé** : est une approche dans laquelle le programme informatique apprend des données étiquetées qui lui sont transmises par l'utilisateur, puis utilise cet apprentissage pour classer une nouvelle observation. Cet ensemble de données peut simplement être bi-classe ou multiclasse.
- **Apprentissage non supervisé** : consiste à concevoir un modèle structurant l'information. Dans cet approche les données sont des données non étiquetées et on cherche à trouver les classes des données. Des différentes tâches sont associées à l'apprentissage non supervisé, tel que : Clustering (segmentation), règle d'association, réduction de dimension.
- **Apprentissage par renforcement** : consiste à apprendre des actions qui doivent être effectuées dans un environnement en évolution afin d'obtenir la récompense maximale.

1.2. Historique de l'apprentissage profond

L'apprentissage profond est basé sur le concept de réseaux de neurones artificiels. Et donc, notre brève historique d'apprentissage profond doit commencer par ces réseaux de neurones.

-1943: Warren McCulloch et Walter Pitts créent un modèle de calcul pour les réseaux de neurones basé sur les mathématiques et les algorithmes appelé logique à seuil.

-1958: Frank Rosenblatt crée le perceptron, un algorithme de reconnaissance de formes basé sur un réseau de neurone informatique à deux couches utilisant une simple addition et soustraction. Il a également proposé des couches supplémentaires avec des notations mathématiques, mais celles-ci ne seront pas réalisées avant 1975.

-1980: Kunihiko Fukushima propose le Neoconitron, un réseau de neurone artificiel multicouche hiérarchisé qui a été utilisé pour la reconnaissance de l'écriture manuscrite et d'autres problèmes de reconnaissance de formes.

-1989: Les scientifiques ont pu créer des algorithmes qui utilisaient des réseaux de neurones profonds, mais les temps de formation des systèmes étaient mesurés en jours, ce qui les rend peu pratiques pour une utilisation dans le monde réel.

-Milieu des années 2000: Le terme «apprentissage profond» commence à gagner en popularité après qu'un article de Geoffrey Hinton et Ruslan Salakhutdinov a montré comment un réseau de neurone à plusieurs couches pouvait être pré-entraîné une couche à la fois.

-2012: Les algorithmes de reconnaissance de forme artificielle permettent d'atteindre des performances de niveau humain sur certaines tâches. Plus l'algorithme d'apprentissage profond de Google "découvre les chats".

-2015: Facebook met en œuvre la technologie d'apprentissage profond - appelée DeepFace -

Chapitre 2 : L'apprentissage profond (Deep learning)

pour étiqueter et identifier automatiquement les utilisateurs de Facebook sur les photos.
-2016: l'algorithme AlphaGo de Google DeepMind maîtrise l'art du jeu de plateau complexe Go et bat le joueur de go professionnel Lee Sedol lors d'un tournoi très médiatisé à Séoul. [12]

1.3.Fonctionnement de deep learning

L'apprentissage profond est basé sur un réseau de neurones artificiels inspiré du cerveau humain. Ce réseau est composé de plusieurs «couches» de neurones, dont les données sont transmises à une couche d'entrée et sont ensuite transférées vers une ou plusieurs couches de traitement - appelées couches cachées - jusqu'à la couche de sortie.

Chaque couche est composée de neurones interconnectés entre eux. Un neurone peut ou non être connecté à tous les neurones des couches adjacentes. Chaque connexion entre deux neurones est associée à un «poids». Lorsque les données sont introduites dans le modèle, elles traversent chaque couche et sont transformées en représentations de plus en plus abstraites. Chaque couche utilise les informations de la couche précédente pour apprendre à détecter des caractéristiques spécifiques. Le modèle associera ensuite des caractéristiques à un objet spécifique en ajustant les poids des connexions neuronales associées à ces caractéristiques.

Par exemple, si le modèle essaie de détecter une balle dans une image, les poids des neurones associés aux cercles seront favorisés. Cet ajustement est effectué à l'aide d'une méthode appelée «rétropropagation». Il fonctionne en calculant l'erreur entre la sortie du modèle et l'entrée. Les poids sont ensuite modifiés pour minimiser cette erreur, en commençant par la couche de sortie et en revenant à la couche d'entrée. En fournissant suffisamment d'exemples au modèle, il peut ajuster les poids jusqu'à ce qu'il puisse détecter avec précision les objets souhaités. Ces exemples n'ont pas besoin de dire au modèle où se trouve un objet. Ils ont juste besoin de lui dire qu'un objet est là, et il apprendra à le détecter par lui-même, tout comme un humain apprendrait à reconnaître un objet s'il est montré plusieurs fois.

1.4.Les architectures de l'apprentissage profond

Il existe plusieurs types d'apprentissage profond, parmi eux le réseau de neurones convolutionnel (CNN) et le réseau de neurones récurrents (RNN).

1.4.1. Les réseaux de neurones récurrents (RNN)

C'est un type de réseau de neurones où la sortie de l'étape précédente est alimentée en entrée de l'étape en cours.

Chapitre 2 : L'apprentissage profond (Deep learning)

Les réseaux de neurones récurrents (RNN) sont conçus pour reconnaître des modèles qui contiennent des séquences de données, telles que du texte, des génomes.... Les algorithmes de ce type prennent en compte le temps et la séquence, ils ont donc une dimension temporelle, ce qui signifie que leur entrée n'est pas seulement l'exemple actuel ils voient, mais aussi ce qu'ils ont perçu précédemment dans le temps. Ainsi, on dit souvent que les réseaux récurrents ont de la mémoire.

Les RNN sont principalement utilisés pour :

- **Classification des séquences** : Classification des sentiments et classification vidéo
- **Étiquetage de séquence** : Partie du balisage vocal et reconnaissance des entités nommées
- **Génération de séquences** : Traduction automatique ...

1.4.2. Le réseau de neurones convolutionnel (CNN)

Les réseaux de neurones convolutionnels est une amélioration des réseaux de neurones artificiels (ANN) traditionnel, qui comprend généralement des couches convolutionnelles, des couches de pool et des couches entièrement connectées. Le CNN peut être divisé en deux parties: une partie d'extraction d'entités (couches de convolution et couches de regroupement) et une partie de classification (couches entièrement connectées). L'image est d'abord passée à travers une série de convolution, regroupant des couches pour l'extraction d'entités, puis passer à travers des couches entièrement connectées pour la classification.

a- Couche convolution (Convolution layer): c'est l'élément clé des réseaux de CNN. Son but est d'identifier un ensemble de caractéristiques (*features*) d'une image reçue en entrée. Pour cela, nous effectuons un filtrage par convolution: dans le principe est de "faire glisser" la fenêtre représentant le filtre sur l'image et de calculer le produit de convolution entre le filtre et chaque partie de l'image balayée. Donc cette couche reçoit plusieurs images en entrée et utilise chaque filtre pour calculer la convolution de chaque image. Les filtres correspondent exactement aux caractéristiques que nous voulons trouver dans l'image. Pour chaque paire (image, filtre), nous obtenons une carte d'activation « feature map » (comme la figure 2), qui nous indique où le caractéristique se trouve dans l'image: plus la valeur est élevée, plus il y a de positions dans l'image qui correspondent au filtre.

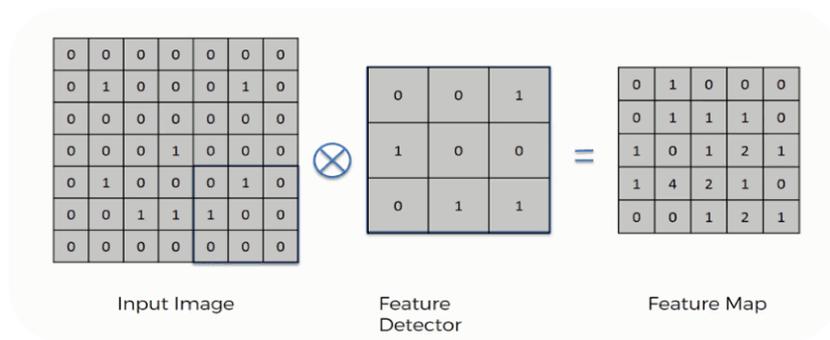


Figure 2 : Parcours de filtre sur une image d'entrée[13]

- b- la couche de groupement (*pooling layer*):** il reçoit plusieurs « feature map » en entrée et applique une opération de pooling (subsampling) à chaque feature map. Elle permet la réduction de la taille de l'image en tenant compte de ses caractéristiques importantes. Le principe est de couper l'image en cellules régulières, puis nous conservons la valeur maximale dans chaque cellule par rapport au filtre utilisé (2 x 2 pixels, 3x3...). Donc des petites cellules carrées sont souvent utilisées afin de ne pas perdre trop d'informations.

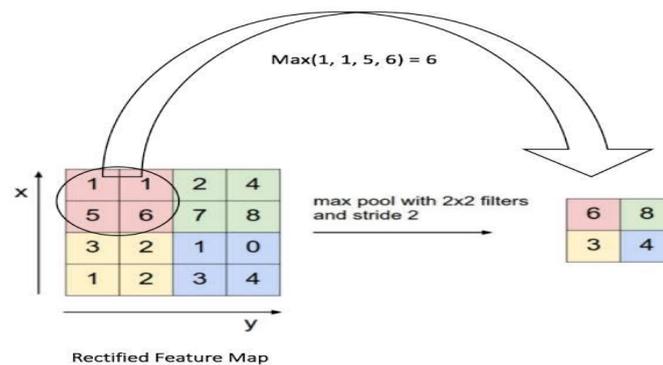


Figure 3 : Max Pooling avec un filtre 2x2 et un pas de 2.[14]

Nous obtenons le même nombre de « features map » que l'entrée, mais elles sont beaucoup plus petites. Cela permet d'accélérer non seulement les calculs, mais d'éviter également le problème du sur-ajustement. Il existe trois principaux types de pooling : max et min et moyenne. Comme son nom l'indique, pooling maximal est basé sur la détection de la valeur maximale dans la région sélectionnée (vois la figure 3), le pooling minimal sur la détection de la valeur minimale dans la région sélectionnée et le pooling moyenne (average pooling) basé sur la détection de la valeur moyenne dans la région sélectionnée.

- c- La couche de correction ReLU :** représente la fonction réelle non-linéaire permet de remplacer toutes les valeurs négatives reçues en entrées par des zéros. (figure.4). Elle joue le rôle de fonction d'activation. $\text{ReLU}(x)=\max(0,x)$.

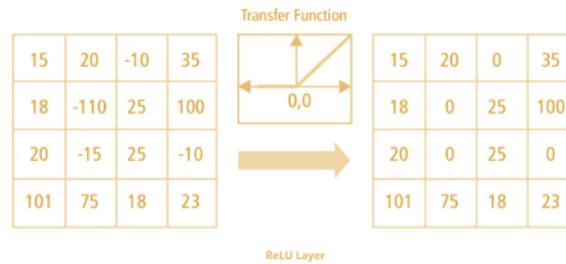


Figure 4 : Couche de correction Relu [14]

- d- **La couche entièrement connectée (fully-connected)** : permet de classifier l'image en entrée du réseau dans une étiquette .elle reçoit un vecteur en entrée (représente le résultat d'aplatir de notre matrice) et produit un nouveau vecteur en sortie. Chaque élément du vecteur représentant une probabilité qu'une certaine caractéristique appartienne à une étiquette.
- e- **Fonction de perte (loss function)**: la dernière chose que nous devons définir est la fonction de perte. La fonction de perte peut refléter la qualité de la prédiction du modèle, qui nous indique comment le réseau de neurones effectue une tâche particulière. Au cours de l'entraînement du réseau, le réseau génère des valeurs prédites à travers chaque couche d'opérations, puis utilise la fonction de perte pour calculer la différence entre la valeur prédite et la valeur réelle. L'entraînement du réseau de neurone vise à réduire cette différence (perte). La perte doit être aussi faible que possible, si la fonction de perte est grande, les performances de notre réseau de neurones ne sont pas bonnes.

1.5.Transfer Learning : (l'apprentissage par transfert)

L'apprentissage par transfert est une technique par laquelle un modèle de réseau de neurones est d'abord entraîné sur un problème similaire au problème en cours de résolution. Une ou plusieurs couches du modèle entraîné sont ensuite utilisées dans un nouveau modèle entraîné sur le problème d'intérêt. Il a été prouvé que l'apprentissage par transfert peut mieux fonctionner dans de nombreux cas avec un ensemble de données d'entraînement plus petit. [15]

Les CNN utilisés pour la classification des images se composent de deux parties: les couches de convolution, qui sont utilisées pour l'extraction des caractéristiques, et les couches de connexion complète, qui sont utilisées pour la classification. Par conséquent, nous pouvons utiliser directement le réseau de neurone pré-entraîné pour extraire des caractéristiques de l'image, puis le vecteur de caractéristique extrait est utilisé comme entrée pour former une nouvelle couche entièrement connectée afin de résoudre d'autres problèmes de classification. [15]

Chapitre 2 : L'apprentissage profond (Deep learning)

Par exemple, nous pouvons en apprendre davantage sur un ensemble de différentes catégories visuelles, y compris les animaux, les plantes, les véhicules et les objets du quotidien. Puis en apprendre davantage sur un ensemble différent de catégories visuelles, pour classer des types spécifiques de véhicules. Ces tâches sont très similaires, vous devez donc essayer de réutiliser des parties du premier réseau (voir Figure 5).

La couche de sortie du modèle d'origine doit généralement être remplacée car elle n'est probablement pas du tout utile pour la nouvelle tâche, et elle peut même ne pas avoir le bon nombre de sorties pour la nouvelle tâche.

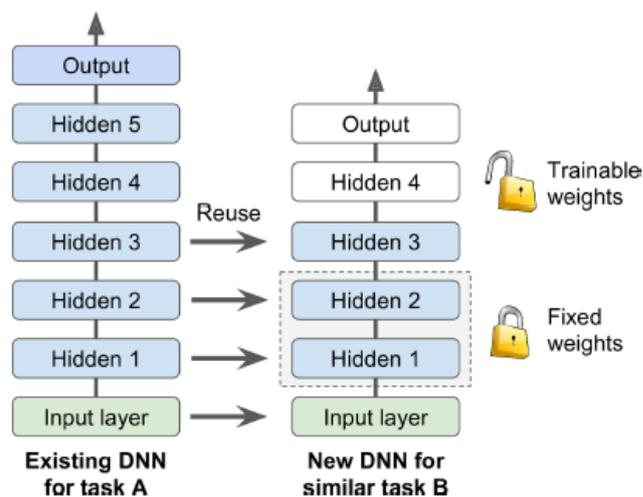


Figure 5: Réutilisation de la couche pré-entraînée [16]

Le modèle pré-entraîné ou la partie souhaitée du modèle peut être intégré directement dans un nouveau modèle de réseau de neurone. Dans cette utilisation, les poids des pré-entraînés peuvent être gelés (frozen) afin qu'ils ne soient pas mis à jour lors de l'apprentissage du nouveau modèle (le gradient descente ne les modifie pas). Ensuite on essaye de dégeler « unfreeze » une ou deux des couches cachées supérieures pour laisser la rétropropagation les modifier et voir si les performances s'améliorent. Plus vous avez de données d'entraînement, plus vous pouvez dégeler de couches. Alternativement, les poids peuvent être mis à jour pendant l'entraînement du nouveau modèle, peut-être avec un taux d'apprentissage inférieur, permettant au modèle pré-entraîné d'agir comme un schéma d'initialisation de poids lors de l'entraînement du nouveau modèle. [16]

Nous pouvons résumer certains de ces modèles d'utilisation comme suit:

Chapitre 2 : L'apprentissage profond (Deep learning)

- **Classificateur:** Le modèle pré-entraîné est utilisé directement pour classer les nouvelles images.
- **Extracteur de caractéristiques intégré:** le modèle pré-entraîné ou une partie du modèle, est intégré dans un nouveau modèle, mais les couches du modèle pré-entraîné sont gelées pendant l'entraînement.
- **Initialisation du poids:** le modèle pré-entraîné ou une partie du modèle, est intégré dans un nouveau modèle, et les couches du modèle pré-entraîné sont entraînées de concert avec le nouveau modèle.

Chaque approche peut être efficace et permet de gagner un long temps dans le développement et l'entraînement d'un modèle de réseau CNN.

Les modèles pré-entraînés sont généralement partagés sous la forme de millions de paramètres/poids que le modèle a atteints. Ils sont disponibles pour tout le monde à utiliser par différents moyens. La célèbre bibliothèque Python d'apprentissage profond '*keras*', fournit une interface pour télécharger certains modèles populaires. Vous pouvez également accéder à des modèles pré-entraînés à partir du Web, car la plupart d'entre eux sont open source.

1.6. Les modèles du CNN

Les architectures CNN empilent quelques couches convolutionnelles (chacune généralement suivie d'une couche ReLU), puis une couche de regroupement, puis quelques autres couches convolutionnelles (+ ReLU), puis une autre couche de regroupement, etc. L'image devient de plus en plus petite au fur et à mesure qu'elle progresse à travers le réseau, mais elle devient généralement de plus en plus profonde (c'est-à-dire avec plus de cartes de caractéristiques) grâce aux couches de convolutions. Au sommet de la pile, un réseau de neurones « feedforward » régulier est ajouté, composé de quelques couches entièrement connectées (+ ReLU), et la couche finale génère la prédiction (par exemple, une couche softmax qui génère des probabilités de classe estimées).

Au fil du temps, des variantes de cette architecture fondamentale ont été développées, conduisant à des avancées étonnantes dans le domaine.

Avant de détailler les architectures de CNN, nous présenterons d'abord l'un des ensembles de données les plus importants dans la classification d'images ImageNet [17]. ImageNet est un ensemble de données de plus de 15 millions d'images haute résolution catégorisées appartenant à environ 22 000 catégories. Les images ont été collectées sur le Web

Chapitre 2 : L'apprentissage profond (Deep learning)

et marquées par des étiqueteurs humains à l'aide de l'outil « *Amazons Mechanical Turk crowd-sourcing* ».

À partir de 2010, dans le cadre du *PascalVisual Object Challenge*, un concours annuel appelé *ImageNet Large-Scale Visual Recognition Challenge* (ILSVRC) a été organisé. ILSVRC utilise un sous-ensemble d'ImageNet avec environ 1000 images dans chacune des 1000 catégories. Au total, il y a environ 1,2 million d'images de formation, 50 000 images de validation et 150 000 images de test.

- **AlexNet** : l'architecture AlexNet [18] a remporté le défi ImageNet ILSVRC 2012 par une large marge: elle a atteint 17% de taux d'erreur dans le top 5 tandis que la deuxième meilleure n'a atteint que 26%! Il a été développé par Alex Krizhevsky (d'où son nom), Ilya Sutskever et Geoffrey Hinton. Il a été le premier à empiler des couches convolutionnelles directement les unes sur les autres, au lieu d'empiler une couche de regroupement au-dessus de chaque couche convolutionnelles.
- **GoogLeNet** : l'architecture GoogLeNet a été développée par Christian Szegedy et al de Google Research, [19] et il a remporté le défi ILSVRC 2014 en poussant le taux d'erreur du top 5 en dessous de 7%. Cette excellente performance provient en grande partie du fait que le réseau était beaucoup plus profond que les CNN précédents. Cela a été rendu possible par des sous-réseaux appelés modules de création (inception module), qui permet à GoogLeNet d'utiliser les paramètres beaucoup plus efficacement que les architectures précédentes: GoogLeNet a en fait 10 fois moins de paramètres qu'AlexNet (environ 6 millions au lieu de 60 millions).
- **VGGNet** : le finaliste du défi ILSVRC 2014 était VGGNet [20] développé par K. Simonyan et A. Zisserman. Il avait une architecture très simple et classique, avec 2 ou 3 couches convolutionnelles, une couche de regroupement, puis à nouveau 2 ou 3 couches convolutionnelles, une couche de regroupement, etc. (avec un total de seulement 16 couches convolutionnelles), plus un réseau dense final avec 2 couches cachées et la couche de sortie. Il n'utilisait que des filtres 3×3 , mais de nombreux filtres.
- **ResNet** : le défi ILSVRC 2015 a été remporté en utilisant un réseau résiduel (ou ResNet), développé par Kaiming He et al [21] qui a réalisé un taux d'erreur top-5 incroyable sous 3,6%, en utilisant un CNN extrêmement profond composé de 152 couches. Cela confirme la tendance générale: les modèles deviennent de plus en plus profonds, avec de moins en moins de paramètres. La clé pour pouvoir former un réseau aussi profond est d'utiliser des connexions de saut (également appelées connexions de raccourci): le signal entrant dans

Chapitre 2 : L'apprentissage profond (Deep learning)

une couche est également ajouté à la sortie d'une couche située un peu plus haut dans la pile.

Lors de l'entraînement d'un réseau de neurone, l'objectif est de lui faire modéliser une fonction cible $h(x)$. Si vous ajoutez l'entrée x à la sortie du réseau (c'est-à-dire que vous ajoutez une connexion de saut), le réseau sera forcé de modéliser $f(x) = h(x) - x$ plutôt que $h(x)$. C'est ce qu'on appelle l'apprentissage résiduel (voir la figure 6).

Lorsque l'initialisation d'un réseau de neurones normal, ses poids sont proches de zéro, donc le réseau ne sort que des valeurs proches de zéro. Si vous ajoutez une connexion de saut, le réseau résultant sort simplement une copie de ses entrées; en d'autres termes, il modélise initialement la fonction d'identité. Si la fonction cible est assez proche de la fonction identité (ce qui est souvent le cas), cela accélérera considérablement l'entraînement.

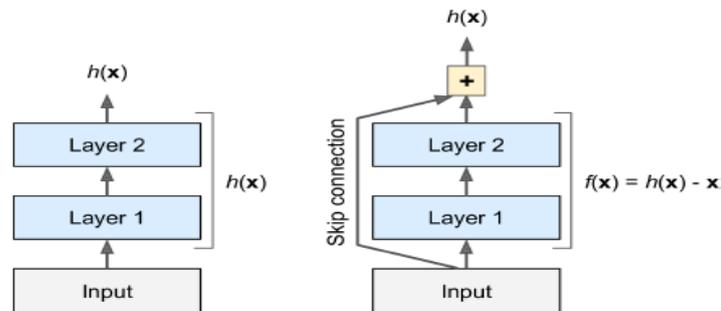


Figure 6: l'apprentissage résiduel[21]

- **SENet** : l'architecture gagnante du défi ILSVRC 2017 était le « *Squeeze-and-Excitation Network* » (SENet) [22]. Cette architecture étend les architectures existantes telles que les réseaux de création (*inception*) ou ResNets et améliore leurs performances. Cela a permis à SENet de remporter la compétition avec un taux d'erreur étonnant de 2,25% dans le top 5! Les versions étendues des réseaux de création et ResNet sont appelées SE-Inception et SE-ResNet respectivement. Le coup de pouce vient du fait qu'un SENet ajoute un petit réseau de neurone, appelé SE Block, à chaque unité de l'architecture d'origine (c'est-à-dire, chaque module de démarrage ou chaque unité résiduelle), comme le montre la figure 7.

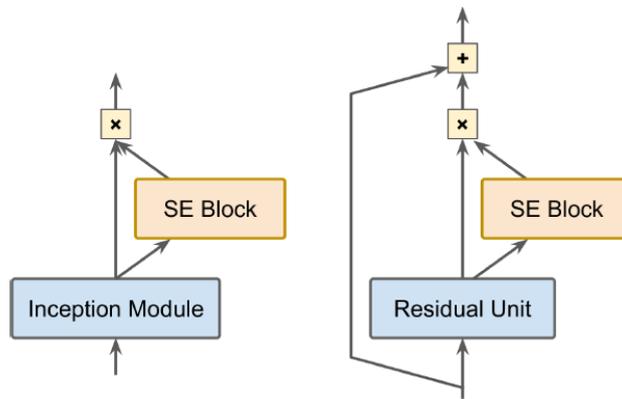


Figure 7: Module SE-Inception (gauche) et Unité SE-ResNet (droite) [22]

Un bloc SE analyse la sortie de l'unité à laquelle il est attaché, et il apprend quelles caractéristiques sont généralement les plus actives ensemble. Il utilise ensuite ces informations pour recalibrer les cartes d'entités, comme le montre la figure 8.

Par exemple, un bloc SE apprend que la bouche, le nez et les yeux apparaissent généralement ensemble sur les images: si vous voyez une bouche et un nez, vous devez également vous attendre à voir des yeux. Donc, si un bloc SE voit une forte activation dans les cartes des fonctionnalités de la bouche et du nez, mais seulement une activation légère dans la carte des fonctionnalités de l'œil, il augmentera la carte des fonctionnalités de l'œil (plus précisément, cela réduira les cartes des fonctionnalités non pertinentes). Si les yeux étaient quelque peu confondus avec autre chose, ce recalibrage de la carte des fonctionnalités aidera à résoudre l'ambiguïté.

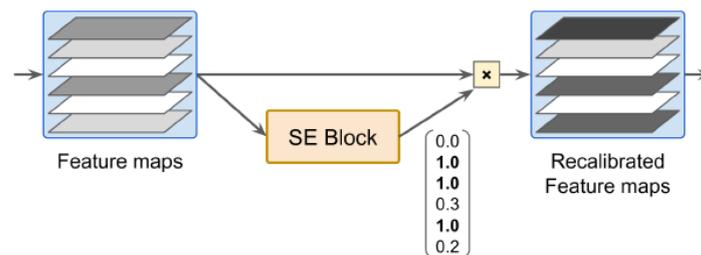


Figure 8: Un bloc SE effectue un recalibrage de la carte des caractéristiques [22]

Un bloc SE est composé de seulement trois couches: une couche de mise en commun moyenne globale, une couche dense cachée utilisant la fonction d'activation ReLU et une couche de sortie dense utilisant la fonction d'activation sigmoïde (voir Figure 9).

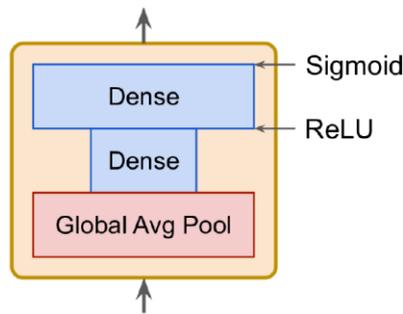


Figure 9: Architecture de SE Block[22]

La couche de regroupement moyenne globale calcule l'activation moyenne pour chaque mappe d'entités: par exemple, si son entrée contient 256 mappes d'entités, elle affichera 256 nombres représentant le niveau global de réponse pour chaque filtre.

La couche suivante est l'endroit où se produit la compression (*squeeze*): cette couche a beaucoup moins de 256 neurones, généralement 16 fois moins que le nombre de cartes d'entités (par exemple, 16 neurones), de sorte que les 256 nombres sont compressés dans un petit vecteur (par exemple, 16 dimensions). Il s'agit d'une représentation vectorielle de faible dimension (c'est-à-dire une incorporation) de la distribution des réponses des caractéristiques. Cette étape de goulot d'étranglement (*bottleneck*) oblige le bloc SE à apprendre une représentation générale des combinaisons de fonctionnalités.

Enfin, la couche de sortie prend l'incorporation et génère un vecteur de recalibrage contenant un nombre par mappage d'entités (par exemple, 256), chacun entre 0 et 1. Les maps de caractéristiques sont ensuite multipliées par ce vecteur de recalibrage, donc les caractéristiques non pertinentes (avec un faible score de recalibrage) sont réduits alors que les fonctionnalités pertinentes (avec un score de recalibrage proche de 1) sont laissées seules.

2. La classification

2.1. Les méthodes de classification

La classification c'est une méthode qui permet de déterminer et définir la catégorie des nouvelles observations à partir des données existant (training data). On trouve deux types de classification :

- **Classification binaire** : où les données entrant sont classées dans l'une des deux catégories possibles,

Chapitre 2 : L'apprentissage profond (Deep learning)

- **Classification multiclasse** : dont les données sont divisées en plusieurs catégories possibles supérieures à deux. Ce type de classification est très utile pour la grande masse de données.

2.2.La classification des images

La classification des images consiste à attribuer une ou plusieurs étiquettes à une image, qui est l'un des problèmes les plus fondamentaux de la vision par ordinateur et de la reconnaissance des formes. Elle a un large éventail d'application, par exemple : la récupération d'images et de vidéos, surveillance vidéo, et la biométrie...

En général, le processus de classification d'image consiste à extraire des caractéristiques d'images puis à les classer. Les méthodes de classification traditionnelles utilisent des fonctionnalités de bas niveau ou de niveau intermédiaire pour représenter une image. Les fonctionnalités de bas niveau sont généralement basées sur des informations de densité, de couleur, de texture, de forme et de position en niveaux de gris, qui sont définis par l'homme (également appelées fonctionnalités fabriquées à la main, « hand-crafted features »).

Les fonctionnalités de niveau intermédiaire, ainsi que les fonctionnalités basées sur l'apprentissage, sont généralement distillées par des algorithmes de sac de mots visuels (BoVW) [23], [24] qui sont efficaces et populaires dans le cadre de la classification ou de la récupération d'images au cours des dernières années. En vision par ordinateur, après l'extraction des caractéristiques, un classificateur (par exemple SVM[25], forêt aléatoire [26], etc.) est généralement utilisé pour attribuer l'étiquette aux différents types d'objets. La classification traditionnelle des images est illustrée à la figure 10.a.

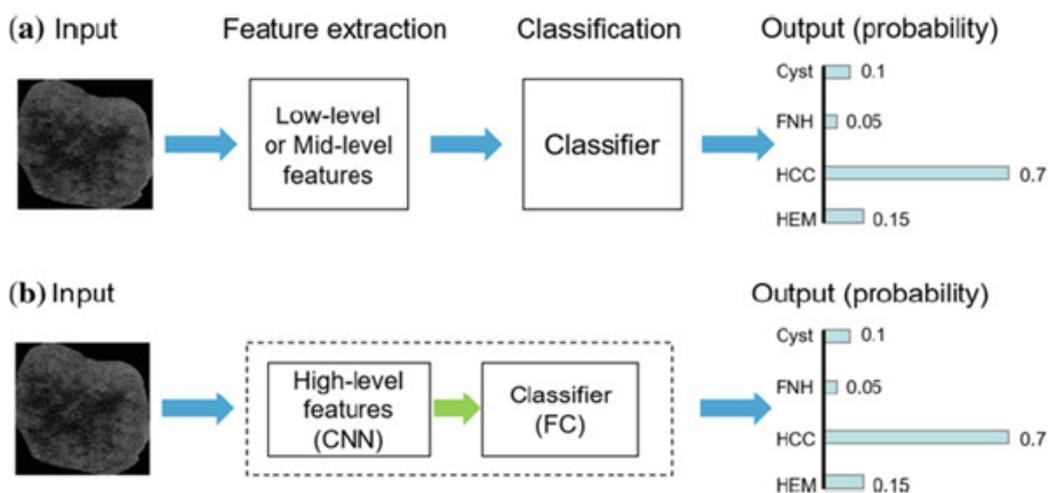


Figure 10: Classification d'images. a Méthode de classification traditionnelle. b Apprentissage profond[27]

Chapitre 2 : L'apprentissage profond (Deep learning)

Différente de la méthode de classification d'image traditionnelle, la méthode d'apprentissage profond combine le processus d'extraction et de classification des caractéristiques de l'image sur un même réseau. Le processus de classification en apprentissage profond est illustré à la figure 10.b.

La représentation des fonctionnalités de haut niveau du deep learning s'est avérée supérieure aux fonctionnalités de bas niveau et aux fonctionnalités de niveau intermédiaire réalisées à la main, et a obtenu de bons résultats en termes de reconnaissance et de classification d'images. Ce concept est à la base du modèle d'apprentissage profond, qui est composé de nombreuses couches (telles que les couches convolutionnelles et les couches entièrement connectées) qui transforme les données d'entrée (par exemple : les images) en sorties (le résultat de la classification) tout en apprenant de plus en plus haut de niveau supérieur [27].

3. Travaux connexes

Au cours de la dernière décennie, des efforts importants ont été déployés pour la reconnaissance du cancer du sein à partir d'images histologiques, où la plupart des efforts sont déployés pour classer les deux types fondamentaux de cancer du sein (bénin et malin) à l'aide du diagnostic assisté par ordinateur (CAD). Ces recherches contiennent principalement deux catégories: 1) des algorithmes de classification basés sur l'ingénierie des caractéristiques humaines et l'apprentissage automatique classique ; 2) les récentes méthodes basées sur l'apprentissage profond.

3.1. Classification par apprentissage automatique

Avant la révolution de l'apprentissage profond, des approches d'apprentissage automatique, y compris la machine à vecteur de support (SVM), l'analyse des composants principaux (PCA) et la forêt aléatoire (RF) ont été utilisées pour étudier les données dont les caractéristiques ont été extraites avec une extraction de caractéristiques invariantes à l'échelle (SIFT), locale configuration binaire (LBP), quantification de phase locale (LPQ), matrice de cooccurrence de niveau de gris (GLCM), statistiques de contiguïté de seuil (TAS) et TAS sans paramètre (PFTAS).

Kowal *et al.*[28] ont utilisé différents algorithmes de regroupement pour la segmentation des noyaux (y compris les moyennes K, les moyennes floues, les réseaux de neurones d'apprentissage compétitifs et les modèles de mélange gaussiens) sur des images microscopiques de biopsie. En 2016, l'une des bases de données les plus populaires pour les problèmes de classification de cancer du sein a été publiée et un groupe de recherche a signalé une précision d'environ 85,1% en utilisant les fonctionnalités SVM et PFTAS pour l'analyse au

Chapitre 2 : L'apprentissage profond (Deep learning)

niveau du patient [29], ce qui était la précision de reconnaissance la plus élevée à l'époque. Wang et al. [30] ont proposé un cadre pour la segmentation des noyaux cellulaires et la classification des images pathologiques de cancer du sein. Pour l'étape de classification, 4 entités basées sur la forme et 138 entités texturales basées sur les espaces colorimétriques sont extraites. L'ensemble des caractéristiques optimales est obtenu par SVM avec un algorithme génétique d'agent en chaîne. La méthode proposée a obtenu une performance prometteuse sur 68 images d'histopathologie des cellules mammaires. Gupta et al. [31] ont proposé un cadre sur plusieurs grossissements pour la classification des images histopathologies de cancer du sein. Les auteurs ont utilisé des caractéristiques et des classificateurs de couleur-texture communs pour démontrer que certaines de ces caractéristiques et classificateurs étaient effectivement efficaces.

Il convient de noter que les méthodes de classification ci-dessus n'ont pas de norme de comparaison unifiée et qu'il n'y a pas de comparabilité entre la métrique de précision. Plus important encore, ces algorithmes sont basés sur une méthode d'extraction manuelle des caractéristiques, qui non seulement nécessite des connaissances du domaine, mais consomme également beaucoup de temps et d'énergie pour terminer, en outre, le problème clé est que l'extraction de fonctionnalités discriminantes de haute qualité reste difficile

3.2. Classification par apprentissage profond

De nombreux travaux ont déjà été publiés qui traitent de la reconnaissance du cancer du sein en utilisant des approches DL, où les variantes CNN sont appliquées pour la classification. Quelques-unes de ces expériences sont menées avec l'ensemble de données « BreakHis » [29] Spanhol et al [32, 33] a employé AlexNet pour extraire la caractéristique profonde et a combiné différentes stratégies de fusion de caractéristiques pour la reconnaissance du cancer du sein. Les performances du modèle proposé sont bien meilleures que celles traditionnelles

Wei et al. [34] ont proposé une nouvelle méthode basée sur le CNN profond (nommé BiCNN) pour traiter la classification des images pathologiques de cancer de sein à deux classes. Ce modèle considérait les étiquettes de classe et de sous-classe de cancer du sein comme une connaissance préalable, ce qui pourrait restreindre la distance des caractéristiques de différentes images pathologiques de cancer du sein. Garud et al. [35] ont présenté un modèle de classification basé sur l'architecture GoogLeNet pour le diagnostic des échantillons de cellules à l'aide de leurs vues microscopiques à fort grossissement.

Han et al.[36] a utilisé GoogLeNet comme réseau de base et a proposé une méthode de classification multiple de cancer du sein à partir d'images histopathologies. Le modèle structuré

Chapitre 2 : L'apprentissage profond (Deep learning)

avait obtenu des performances remarquables sur un ensemble de données à grande échelle, qui était un outil potentiel pour la multi-classification du cancer du sein en milieu clinique. Inspiré par les modules de création (inception module), Akbar et al. [37] ont proposé une technique de régularisation nommée module de transition, qui capture les filtres à plusieurs échelles, puis les réduit par regroupement moyen mondial pour faciliter la réduction de la taille du réseau des couches convolutives aux couches FC. Zhi et al. [38] ont étudié l'utilisation de l'apprentissage par transfert sur des réseaux de neurones convolutifs (VGGNet et le modèle personnalisé) pour diagnostiquer le cancer du sein à partir d'images histopathologies. Zhang et al. [39] proposent une méthode pour utiliser la connexion par saut dans Resnet afin de résoudre les problèmes d'optimisation lorsque le réseau devient plus profond. Vo et al. [40] appliquent les techniques d'augmentation comme rotation, coupe, transformation d'image pour augmenter les données d'entraînement avant d'extraire la fonction profonde du modèle Inception-ResNet-v2 afin d'éviter le sur-ajustement.

Dans [41] les auteurs conçoivent un nouveau réseau CNN utilisant un module Small SE-ResNet basé sur la combinaison du module résiduel et de bloc Squeeze -et- d'excitation. Par rapport au module « bottleneck SE-ResNet » et du module de base SE-ResNet, les paramètres du module small SE-ResNet sont réduits à 29,4% et 33,3%, respectivement. Les résultats montrent que ce modèle atteint la précision entre 98,87% et 99,34% pour la classification binaire et atteint la précision entre 90,66% et 93,81% pour la classification multi-classes.

Conclusion

Dans ce chapitre on a expliqué la notion du deep learning et les réseaux de neurones convolutionnels et ces différents modèles. On a parlé de l'apprentissage par transfert et la classification des images, ainsi que quelques travaux connexes réalisés pour le diagnostic du cancer du sein.

Dans le chapitre suivant, on va présenter la partie expérimentale de notre travail : l'architecture et les paramètres choisis ainsi que la discussion des différents résultats obtenus.

Conception Et Implémentation D'un Modèle Par Le Transfer Learning

Introduction

À l'heure actuelle, la classification automatique des images pathologiques du cancer du sein à base des réseaux de neurones convolutionnels reste un problème de recherche très important. Avoir un taux de précision très élevé nécessite la conception d'un modèle très approfondi, mais l'inconvénient c'est le nombre de paramètres important d'un côté et le sur-ajustement (over-fitting) d'un autre côté.

Pour répondre à ces problèmes plusieurs travaux dans la littérature sont réalisés. Parmi eux le modèle SE-Small-ResNet choisis dans notre travail comme modèle pré-entraîné. Ce modèle a pu répondre aux problèmes cités précédemment par la conception d'un modèle très approfondi avec réduction très importante des paramètres.

Dans ce chapitre, nous allons montrer la technique de transfert learning choisis pour répondre à notre problématique d'étudier la sensibilité de la pathologie cancéreuse contre un anticorps E-cadhérine, en se basant sur le module de Small SE-ResNet.

1. Base de données

Dans notre travail, nous avons utilisé deux ensembles de données, le premier est une base de données open source utilisé pour le modèle pré-entraîné et une deuxième base de données collectée à l'aide des pathologistes du CHU Constantine pour entraîner notre modèle de transfert learning.

- **Base de données 1 :**

C'est une base de données des images histopathologiques du cancer du sein appelée *BreakHis*, qui est composée de 9109 images microscopiques de lames de biopsie chirurgicale des tumeurs du sein, colorées à l'hématoxyline-éosine (HE). Les images sont divisées en tumeurs bénignes et malignes, qui ont été recueillies auprès de 82 patientes en utilisant différents facteurs d'agrandissement (figure 1) : 40, 100, 200 et 400. Cette base de données a été construite en collaboration avec le P&D 1 - Anatomie Pathologique et Cytopathologie, Paraná, Brésil.[29]

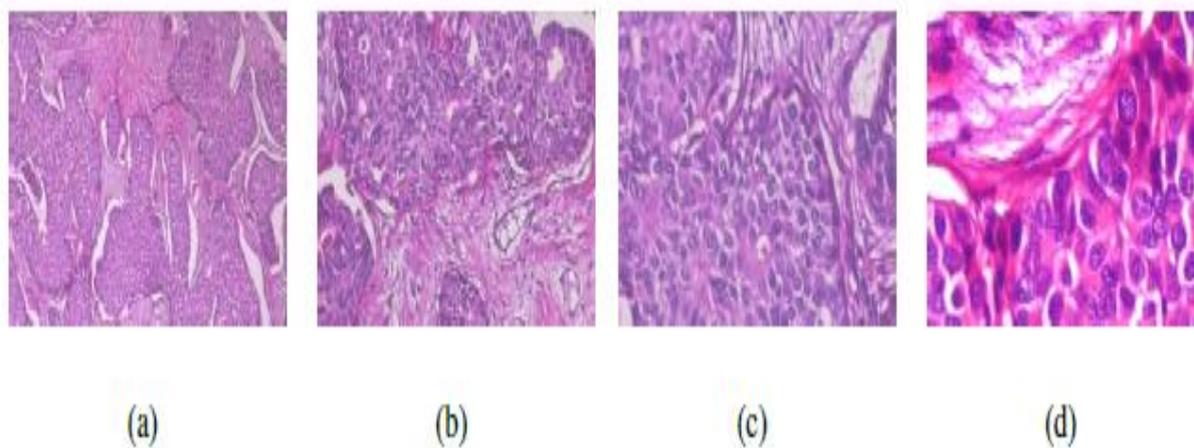


Figure 11: Échantillons d'images de la base de données BreKHis[29]

• Base de données 2

C'est une collection d'images histopathologiques prises à partir des échantillons réels des personnes atteintes du cancer du sein. Les échantillons sont marqués par l'anticorps « E-cadhérine ». L'auteur de l'ensemble de données utilise des lames de biopsie des tissus mammaires pour générer ces échantillons. Ils sont étiquetés par les pathologistes du laboratoire du CHU Constantine, et évalués par immunohistochemistry selon le protocole cité dans le chapitre 1.

L'ensemble des images contient 177 images, divisé en deux catégories :

- E-cadhérine positive : 107 images (figure 12)
- E-cadhérine négatives : 70 images (figure 13)

Dans chaque classe les images ont de type « .png », et nommer de forme synchrone (pos1, pos2,..., neg1, neg2....)

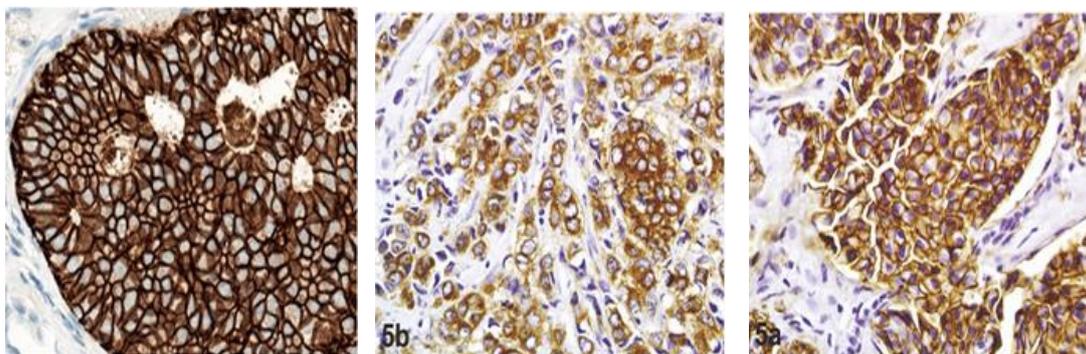


Figure 12 : Échantillons d'images de la base de données (E-cadhérine positive)

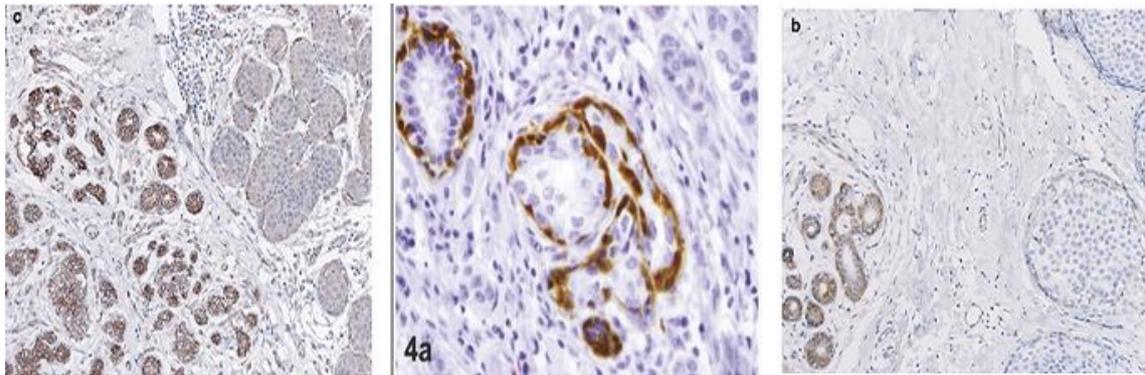


Figure 13: Échantillons d'images de la base de données (E-cadhérine négative)

2. Prétraitement de base de données

Le réseau nécessite des images d'entrée de taille 224 x 224 x 3, mais les images utilisées ont des tailles différentes. Nous devons donc redimensionner les images de l'apprentissage et faire une augmentation de données pour enrichir notre base de données et améliorer le résultat. Autrement dit, effectués des transformations aléatoires sur les images d'entrée. Cela permet d'empêcher le réseau de mémoriser la position et l'orientation exactes des objets, ou encore en générant des échantillons artificiels pour une ou plusieurs classes du jeu de données en **rotation**, **zoom**, **mise en miroir**, **flou** sur l'ensemble original d'images.

L'augmentation de données est faite grâce à la classe prédéfinie de Keras « ImageDataGenerator » qui permet d'appliquer des transformations géométriques, des effets de couleur, de luminosité, de résolution... etc. Il suffit de les parcourir et de choisir les plus pertinentes pour l'étude. (Figure 14)

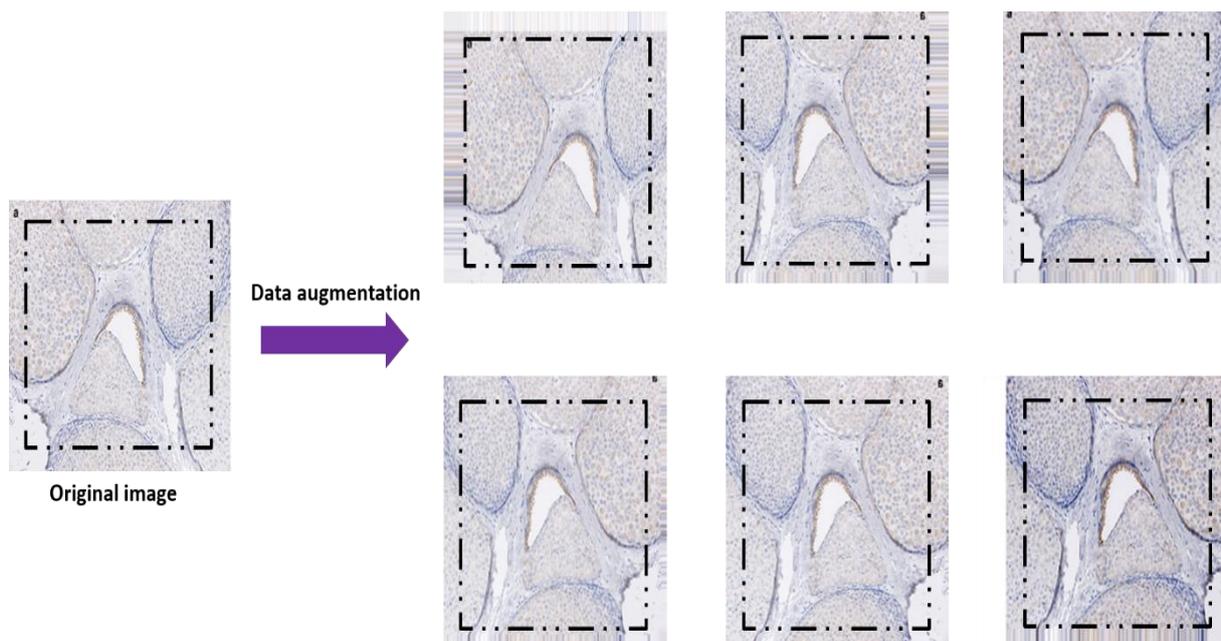


Figure 14: Génération de nouvelles images à partir de celles existantes.

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

Dans notre travail nous avons appliqué les transformations comme suis :

```
#data aug
datagen= ImageDataGenerator(
    featurewise_center=False,
    samplewise_center=False,
    featurewise_std_normalization=False,
    samplewise_std_normalization=False,
    zca_whitening=False,
    rotation_range=0,
    width_shift_range=0.1,
    height_shift_range=0.1,
    horizontal_flip=True,
    vertical_flip=False)
```

Figure 15: les arguments de « ImageDataGenerator » utilisés.

Différentes méthodes sont disponibles pour augmenter les images à partir d'un data frame ou alors directement du répertoire contenant les images. Nous avons utilisé de la méthode « datagen.flow » (figure 16):

```
x=img_to_array(img)
x=x.reshape((1,)+x.shape)
i=0
for batch in datagen.flow(x, batch_size=1, save_to_dir=path1,
                          save_prefix=y, save_format='png'):

    i+=1
    if i>5:
        break
```

Figure 16 : partie de code d'augmentation de données.

Nous avons augmenté le nombre total des images de 177 à 1217 images, et les divisées en deux : 70% pour les données d'entraînement et 30% pour les données de test. (Table 1)

Table 1 : la division de notre base de données

	Notre base de données	
	Sans data augmentation	Avec data augmentation
Train data	123	851
Test data	54	366
Total data	177	1217

3. Le modèle pré-entraîné

Le modèle pré-entraîné utilisé dans cette étude est une combinaison de bloc Squeeze-and-Excitation (bloc SE) et le bloc résiduel de ResNet.

3.1. SE-ResNet

Le modèle SE-ResNet est une combinaison de bloc Squeeze-and-Excitation (bloc SE) et le bloc résiduel de ResNet.

Les blocs « SE » exploitent des informations globales pour sélectionner les caractéristiques informatives en supprimant les moins utiles.

- L'opération de compression utilise la mise en commun moyenne globale (global average pooling) de canaux. Elle rassemble des données et les préserve dans une dimension inférieure.
- L'opération d'excitation sert à décider lesquelles des cartes d'entités sont importantes.

La connexion de raccourci (SR) est la connexion qui saute une ou plusieurs couches et permet aux gradients de se propager davantage et de permettre un apprentissage efficace des réseaux très profonds. Ce dernier garantit que le gradient est toujours supérieur ou égal à 1 dans la rétropropagation, ce qui évite le problème de disparition du gradient de CNN.

Après avoir formé un lot (batch) d'images par époque (epoch), la fonction de coût calcule la distance entre la prédiction et les résultats cibles et obtient une valeur de perte (loss) pour la mise à jour du poids CNN par la rétropropagation.

3.2. Small SE-ResNet

Un module SE-ResNet conçu pour réduire les paramètres du réseau appelé le module Small SE-ResNet. C'est une amélioration de la combinaison du module résiduel et du bloc Squeeze-and-Excitation, et atteint les mêmes performances avec moins de paramètres. [41]

Dans le module SE-ResNet, il y a deux convolutions 1×3 et 3×1 consécutives avec normalisation par lots, et ReLU précédant la convolution, puis combinées avec le bloc SE:

conv 1×3 — conv 3×1 — conv 1×3 — conv 3×1 - bloc SE (figure 17).

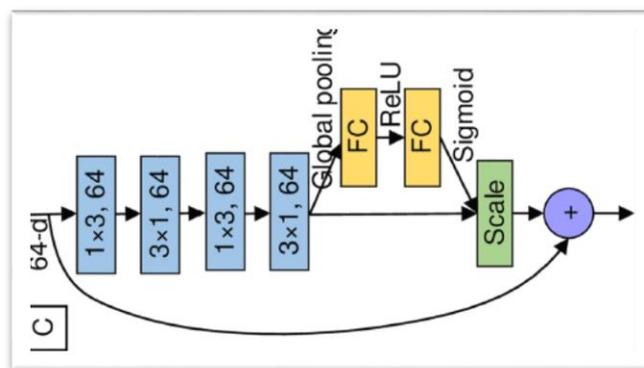


Figure 17: Architecture de module Small SE-ResNet [41]

Nous considérons uniquement le nombre total de paramètres dans les couches convolutionnelles. Le nombre total de paramètres pour une couche convolutionnelle est:

$$T(conv) = \text{taille de kernel} * \text{nombre de kernel} = (C * H * W) * K \quad [41]$$

Où $C \times H \times W$ est la taille du noyau et K est le nombre de noyaux. Ensuite, le nombre de paramètres pour le module Small SE-ResNet pourrait être obtenu par la formule suivante:

$$T(\text{small}) = (64 \times 1 \times 3 \times 64) \times 2 + (64 \times 3 \times 1 \times 64) \times 2 = 12 \times 64^2. [41]$$

3.3. Chargement du modèle pré-entraîné

Le réseau BHCNet (*breast cancer histopathology image classification network*) est une nouvelle architecture CNN pour la classification des images d'histopathologie du cancer du sein en utilisant le petit module SE-ResNet. Il comprend une couche convolutionnelle simple, trois blocs SE-ResNet et une couche entièrement connectée. Chaque bloc SE-ResNet est empilé par N petits modules SE-ResNet, désignés par BHCNet- N . Lorsque $N=3$, l'architecture BHCNet est représentée sur la figure 2. Le modèle BHCNet-3 a 198k paramètres et la taille du modèle est juste 2.1Mb. [41]

Ce modèle a été testé sur l'ensemble de données *BreakHis* pour la classification binaire de cancers du sein (bénigne et maligne) avec des résultats expérimentaux compétitifs (précision entre 98,87% et 99,34% pour la classification binaire.) [41]

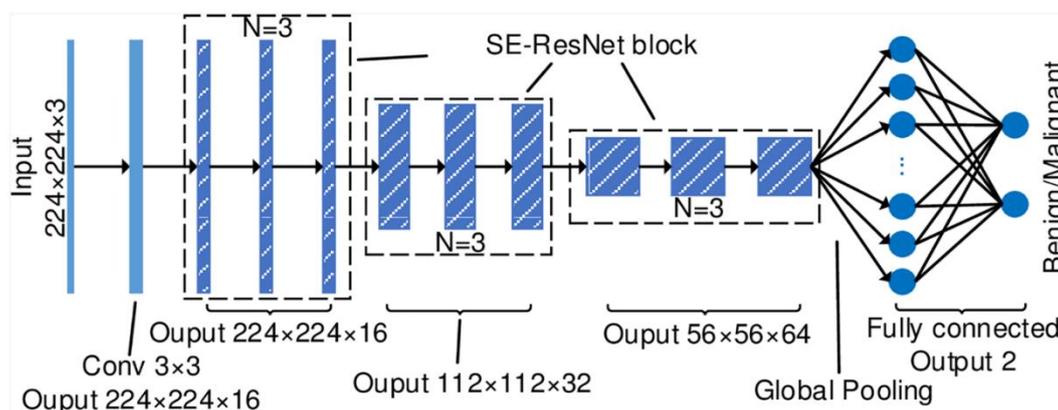


Figure 18: L'architecture BHCNet-3 pour la classification bénigne et maligne des images histopathologies du cancer du sein.[41]

4. Modèle de transfert proposé

4.1. La classification avec un modèle CNN près-entraînés

La première chose à faire avant d'entamer l'utilisation des techniques de transfer learning sur le modèle pré-entraîné que ce soit par suppression ou ajout des nouvelles couches ou par entraîner ou fixer d'autres couches est d'entraîner le modèle eu même par la nouvelle base de données.

Pour ce faire nous avons chargé le model BHCNet a l'aide de classe de *Keras* « *load_model* » et nous avons lancé l'apprentissage sans aucune modification avec l'utilisation de la descente de gradient stochastique comme algorithme d'optimisation pour corriger les prédictions et guider le réseau vers des poids précis.

Voici donc un tableau qui va décrire les résultats obtenus avec les deux base de données utilisés :

Table 2: Taux d'apprentissage pour la classification avec le modèle BHCNet sans aucune modification

Base de données utilises	Taux d'apprentissage
Base de données1 (BreakHis)	97,87%
Base de données2 (notre base de données)	87,70%

Nous pouvons voir que le taux obtenus pour notre base de données est moins précis par rapport au taux obtenus pour la base de données BreakHis. Donc la première chose à faire est de jouer avec les options du training qui sont : *batch_size* et *epochs*.

- **Le *batch_size*** : est un hyper-paramètre de descente de gradient qui contrôle le nombre d'échantillons d'entraînement à traiter avant la mise à jour des paramètres internes du modèle.

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

- **Les epochs** : Le nombre d'époques est un hyper paramètre de descente de gradient qui contrôle le nombre de passages complets dans l'ensemble de données d'apprentissage.

Voici les résultats obtenu après la modification des valeurs de ces options :

Table 3 : Taux d'apprentissage en changeant batch_size et epochs pour une classification avec le modèle BHCNet

Taux d'apprentissage			
	batch_size=32, epochs= 100	batch_size=64, epochs=100	batch_size=128, epochs=200
Accuracy	88,52%	87,90%	87,70%

D'après le tableau 3, au premier temps nous avons utilisé 32 batchs et 100 epochs, le taux de précision obtenu dans ce cas est plus élevé que le taux obtenus précédemment, mais dans le cas de lever le nombre de batch_size et le nombre des époques, le taux est diminuer.

Nous pouvons voir un graphe qui montre les métriques d'apprentissage et les met à jour automatiquement après chaque époque. La bibliothèque « *Matplotlib* » propose de nombreux outils différents pour aider dans ce processus de visualisation.

À partir de graphe de précision et loss dans les figures 19, nous pouvons voir que le modèle ne pourrait probablement pas être entraîné un peu plus car la tendance à la précision sur les deux ensembles de données est stabilisée après une augmentation.

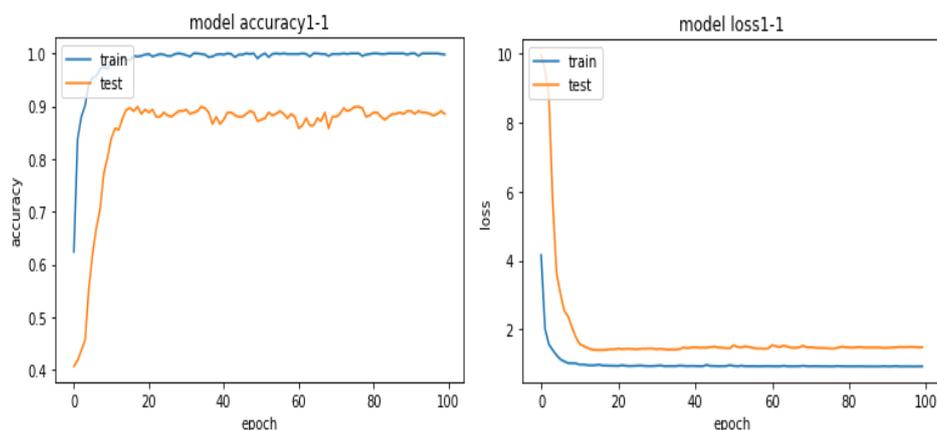


Figure 19: graphe de précision et de perte de la méthode 0

Nous veu maintenant utiliser le Transfer learning, qui permet de réutiliser le modèle pré-entraîné sur un grand ensemble de données et ajuster ses paramètres pour mieux s'adapter à notre ensemble de données.

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

Pour déterminer si l'utilisation de l'apprentissage par transfert rend le réseau pré-entraîné plus efficaces pour nos données, nous effectuons différents types d'expériences :

- Fixer les poids de toutes les couches du modèle pré-entraîné, et entraîner la dernière couche (*dense_2*)
- Fixer les poids de toutes les couches supérieures de modèle pré-entraîné, et remplacer les deux dernières couches du modèle (*dense_2*, *flatten_2*) par des nouvelles couches entraînables.
- Fixer un ensemble des couches superficielles, et entraîner le reste du modèle.

4.2. Méthode 1

Le modèle pré-entraîné contient 219 couches initialisées par des poids optimisés pour l'ensemble de données BreakHis.

Pour conserver et profiter des paramètres et des poids du modèle pré-entraîné, la première idée été de geler les trois blocs SE-ResNet en gardant les poids initiaux du modèles pré-entraîné et appliquant l'apprentissage seulement sur la couche entièrement connecté comme présenté dans la figure 20.

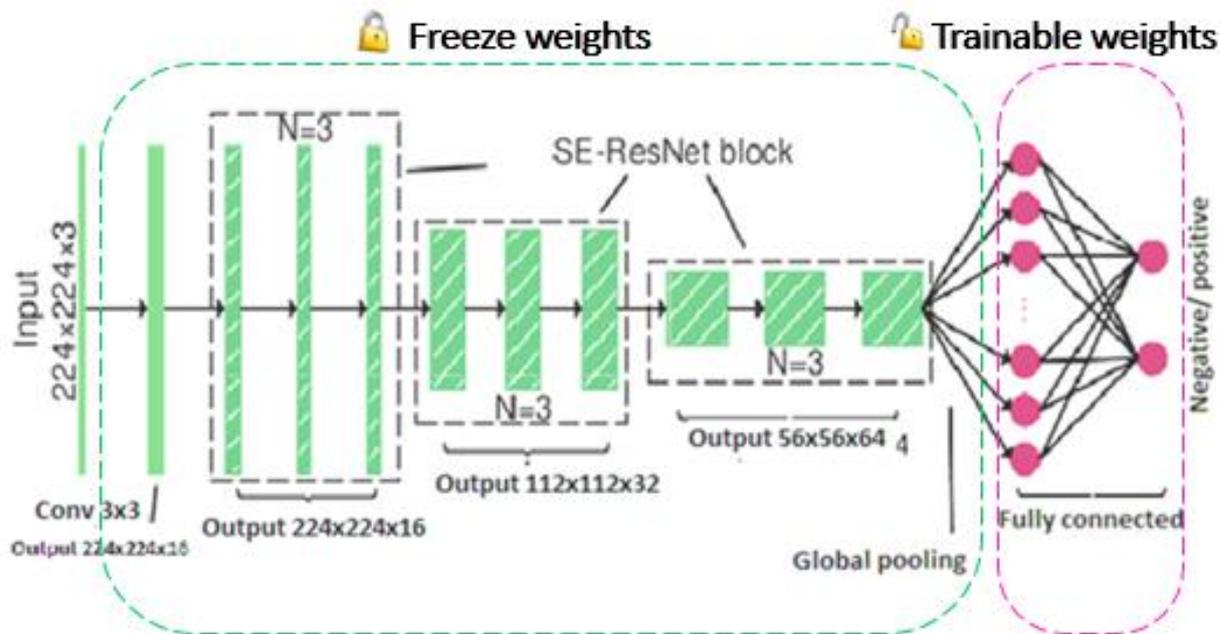


Figure 20: Architecture du transfert d'apprentissage avec la méthode 1

La figure 21 montre les instructions utilisés pour freezer les tous les couche sauf la couches fully connected.

```

for layer in new_model.layers[:-1]:
    layer.trainable = False

new_model.compile(optimizer='SGD',
                  loss='binary_crossentropy',
                  metrics=['accuracy'])
    
```

Figure 21: geler les couches superficielles du modèle.

Après l'application du transfert du modèle par fixation de quelques couches et l'entraînement des autres, nous avons lancé l'apprentissage par plusieurs valeurs du batch-size et par différent nombre d'époques. Les résultats obtenus sont mentionnés dans le tableau 4 où nous avons remarqué que la précision est restée la même et les valeurs d'options utilisées non pas influencé sur le taux de précision.

Table 4: Taux d'apprentissage en changeant batch_size et epochs pour une classification avec la methode 1

Taux d'apprentissage (Méthode 1)			
	batch_size=64 epochs= 100	batch_size=128 epochs=100	batch_size=16 epochs=50
Accuracy	40.71%	40.71%	40.71%

L'apprentissage de modèle de la première méthode et leur précision selon les différentes valeurs de batch et époques sont présentés dans les graphes suivants (figure22, 23, 24).

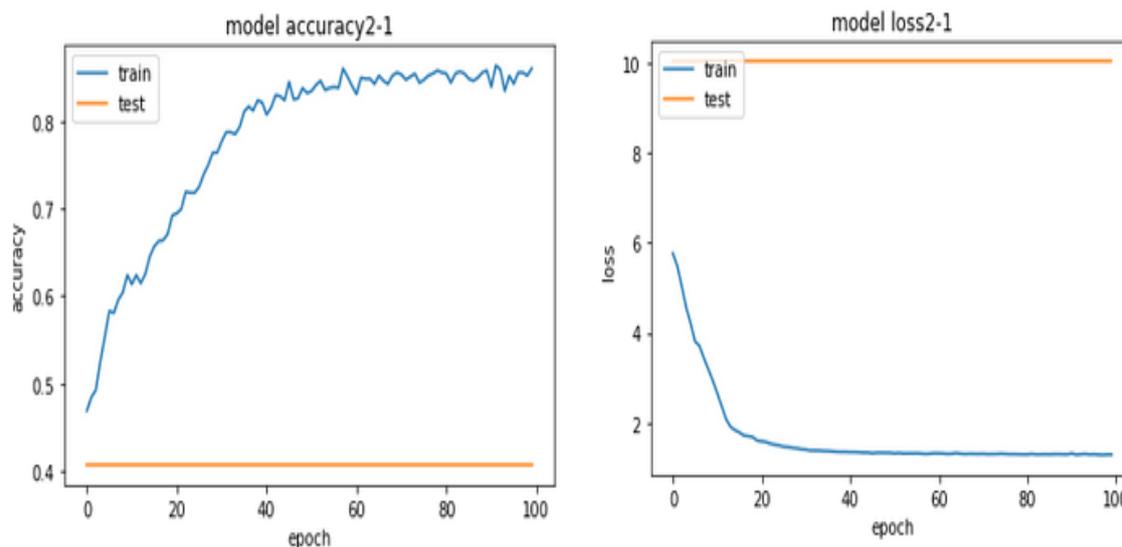


Figure 22 graphe de précision et de perte de la méthode 1 pour batch_size=64, epochs=100.

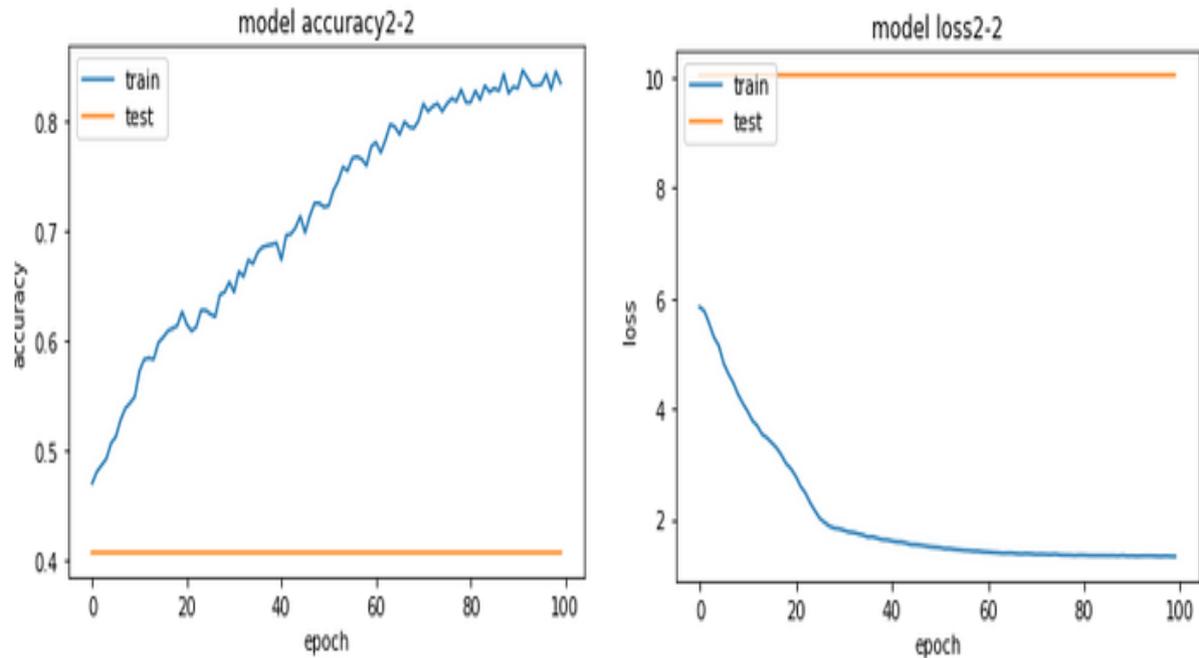


Figure 23: graphe de précision et de perte de la méthode 1 pour $batch_size=128$, $epochs=100$.

D'après la figure 22 et 23 nous remarquons que le modèle peut probablement accepter plus d'entraînement, autrement dit, nous pouvons augmenter soit le nombre d'épochs soit le $batch_size$, par contre la figure 24 montre que l'apprentissage est stable et il n'y a pas une nécessité pour modifier les options d'apprentissage et donc la modification doit être établie dans le modèle de transfert en même

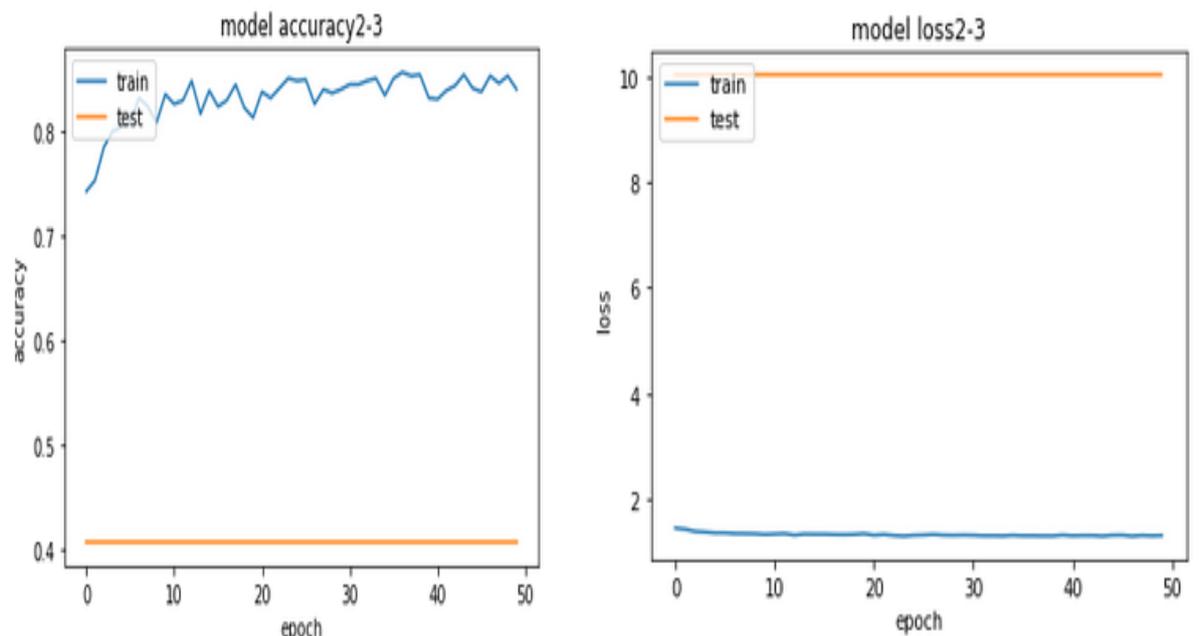


Figure 24 graphe de précision et de perte de la méthode 1 pour $batch_size=16$, $epochs=50$.

4.3. Méthode 2

Les faibles résultats de la première méthode nous ont ramené à penser à introduire une modification dans la structure du modèle pré-entraîné. L'idée est de supprimer des couches et ajouter d'autres puis jouer sur le choix de fixation et l'entraînement des nouvelles couches ajoutées.

- a. La première idée est de supprimer les deux dernières couches du modèle pré entraîné, et les remplace par une nouvelle « *flatten layer* » et deux couches *fully connected* a deux neurones. Les résultats obtenus sont similaires aux résultats de la première méthode. A partir des résultats de la méthode 0 et la méthode 1 nous avons constaté que l'ajout de plusieurs couches entièrement connecté rend le modèle de transfert plus efficace (figure 25).

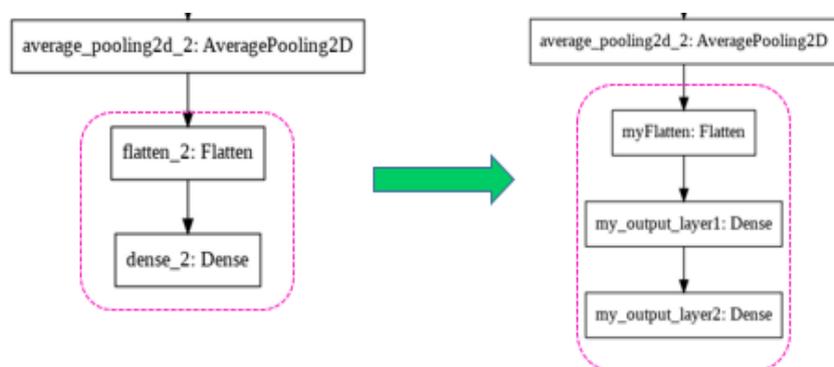


Figure 25 : Changement des deux dernières couches du modèle BHCNet

L'architecture de la méthode est montrée dans la figure 26 :

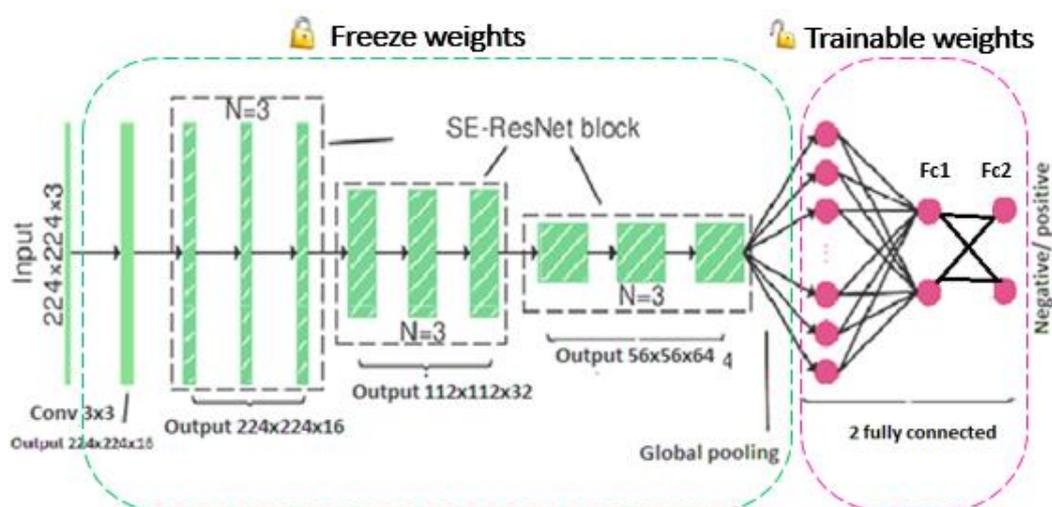


Figure 26: Architecture du transfert d'apprentissage avec la méthode 2.a

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

Il existe plusieurs méthodes de supprimer et ajouter des couches à un modèle, nous avons obtenus la sortie de la couche « *average_pooling2d_2* » puis nous avons ajouté un vecteur d'aplatissement « *myFlatten* » contient 512 neurones et deux couches entièrement connecté « *my_output_layer1* » et « *my_output_layer2* » respectivement contient chacune d'eux deux neurones. Les instructions sont exprimées dans la figure suivante :

```
last_layer=new_model.get_layer('average_pooling2d_2').output
x=Flatten(name='myFlatten')(last_layer)
out=Dense(2,activation="softmax",name='my_output_layer1')(x)
out2=Dense(2,activation="softmax",name='my_output_layer2')(out)
my_model=Model(inputs=input_shape,outputs=out2,name="yasmine_model")
```

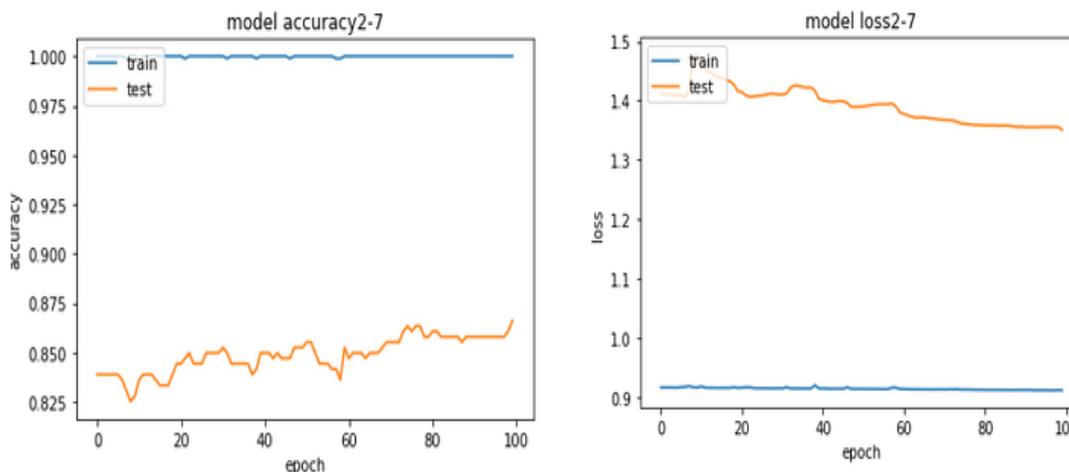
Figure 27 : changement des couches

Nous avons entraîné le modèle plusieurs fois en changeant le nombre d'époques et la taille du batch. Le taux augmente à chaque fois qu'on augmente le nombre d'époques et le batch_size. Selon le tableau 5 nous avons arrivé à un taux de précision de 86.61% pour 100 epochs..

Table 5 : Taux d'apprentissage en changeant batch_size et epochs pour la classification avec la methode 2.a

Taux d'apprentissage (Méthode 2.a)			
	batch_size=32 epochs=50	batch_size=64 epochs=100	batch_size=64 epochs=100
Accuracy	85,52%	84,15%	86,61%

Le graphe de la figure 28 montre que du taux de précision est important lors de l'entraînement mais instable lors de test. Cette interprétation est confirmée par le graphe de perte de cette méthode.



Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

Figure 28 : graphe de précision et de perte de la méthode 2.a

- b. Pour augmenter un peu plus le taux de précision exprimé dans la table 5, nous avons opté à geler quelques couches et entraîner les autres.

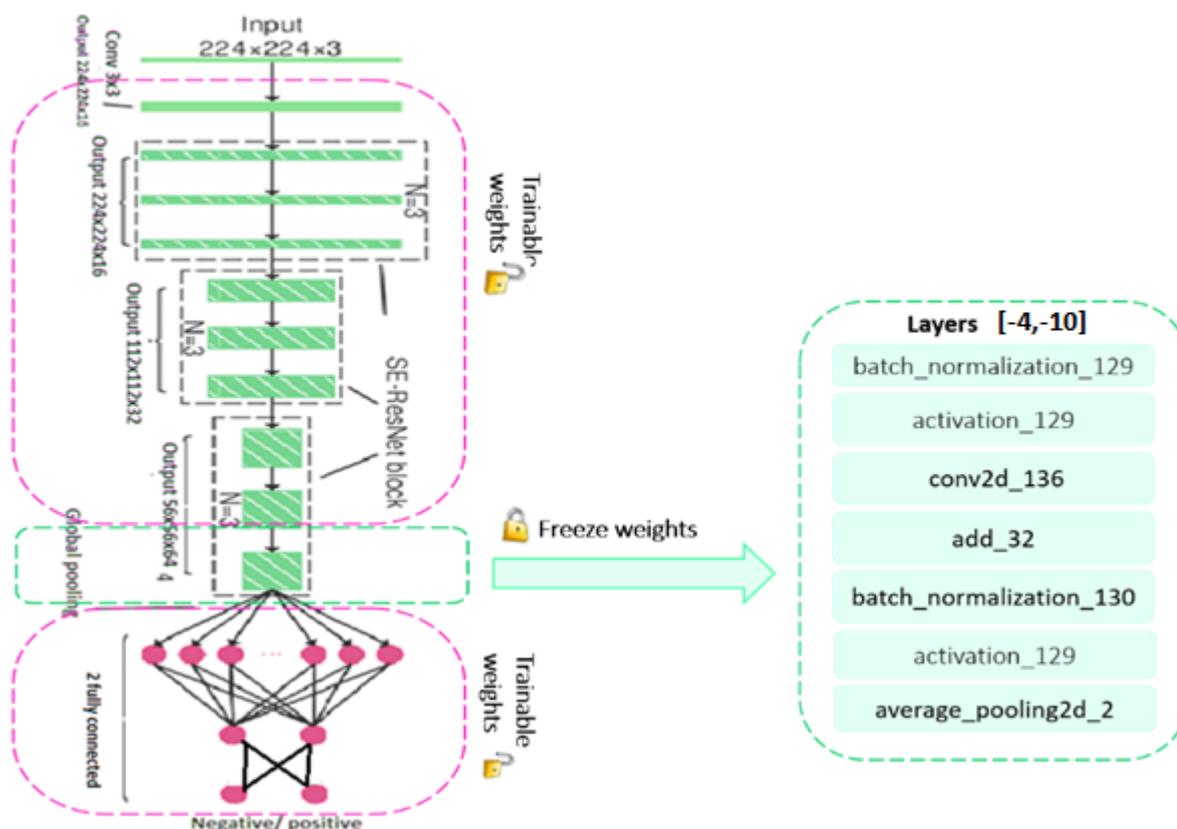


Figure 29: Architecture du transfert d'apprentissage dans le cas de geler les couches entre 4 et 10 avec la méthode 2.b

Après plusieurs jeux d'entraînement des couches du modèle, les meilleurs taux obtenus sont lors de la fixation de 3^{ème} module du troisième block du modèle pré-entraîné comme montrer dans la figure 29.

La table 6 montre les résultats obtenus lors d'une variation de nombre d'épochs et la taille de batch

Table 6: Taux d'apprentissage en changeant batch_size et epochs pour la classification avec la méthode 2

Taux d'apprentissage (Méthode 2)			
	batch_size=32 epochs= 100	batch_size=64 epochs=100	batch_size=128 epochs=200
Accuracy	89 ,07%	89 ,17%	89 ,89%

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

La précision d'apprentissage augmente avec le nombre de « epochs » ceci indique qu'après chaque « epoch » le modèle apprend de plus en plus (Figure 30)

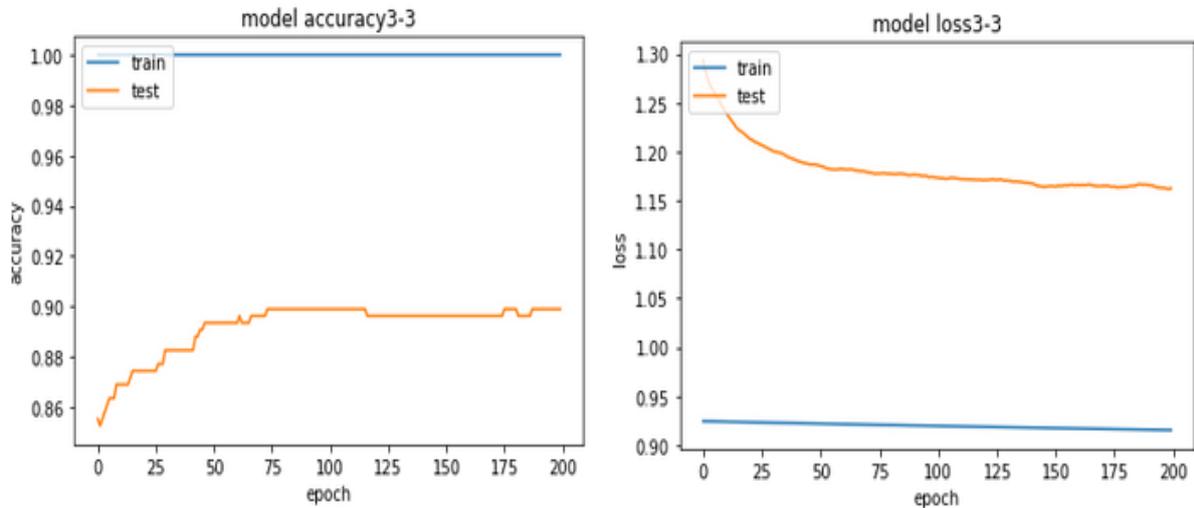


Figure 30 : graphe de précision et de perte de la méthode 2.b.

4.4. Méthode 3

Afin d'améliorer les résultats obtenus de la méthode précédente, nous avons ajouté une troisième couche entièrement connectée (figure 20) et appliqué le principe de la méthode 2.

- Alors nous avons enregistré le modèle pré entraîné de la méthode 2.b avec ces poids, par le biais de l'appel de la fonction « save » de la classe « Model » de Keras et le chargé par la fonction « load_model ».
- Ensuite nous avons ajouté une troisième couche entièrement connectée « my_output_layer3 » formé de deux neurone.et prend en entrée les poids de la deuxième couche entièrement connectée « my_output_layer2 » (figure 31)

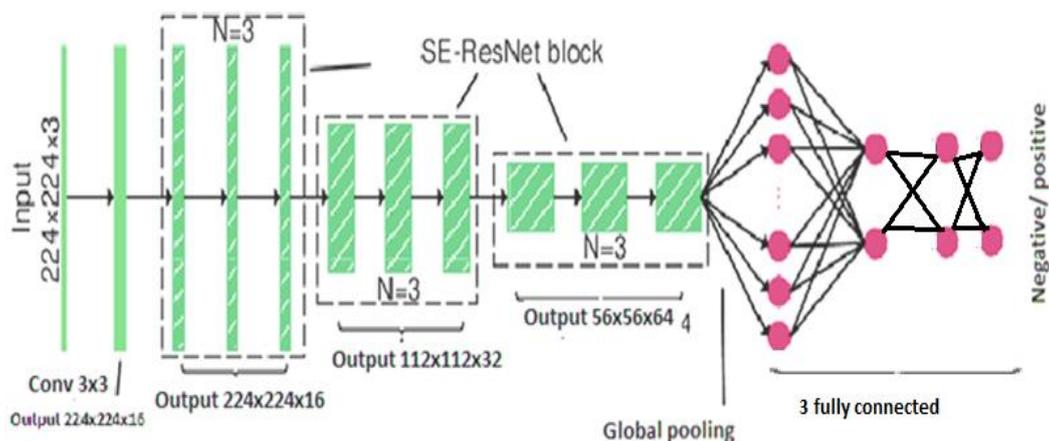


Figure 31: Architecture du transfert d'apprentissage avec la méthode 3

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

Cette fois ci nous avons fixé tous les modules du 3^{ème} bloc et entrainer le reste. Les résultats obtenus sont très intéressantes par rapport aux méthodes précédentes selon la table 7.

Le point remarqué lors de cette méthode est que le nombre d'épochs n'a pas la grande influence sur les résultats obtenus. La figure 32 interprète cette remarque.

Table 7 : Taux d'apprentissage en changeant batch_size et epochs pour la classification avec la méthode 3

Taux d'apprentissage (Méthode 3)				
	batch_size=32 epochs= 50	batch_size=32 epochs=100	batch_size=64 epochs=100	batch_size=128 epochs=200
Accuracy	92.62%	93.99%	91.80%	90.43%

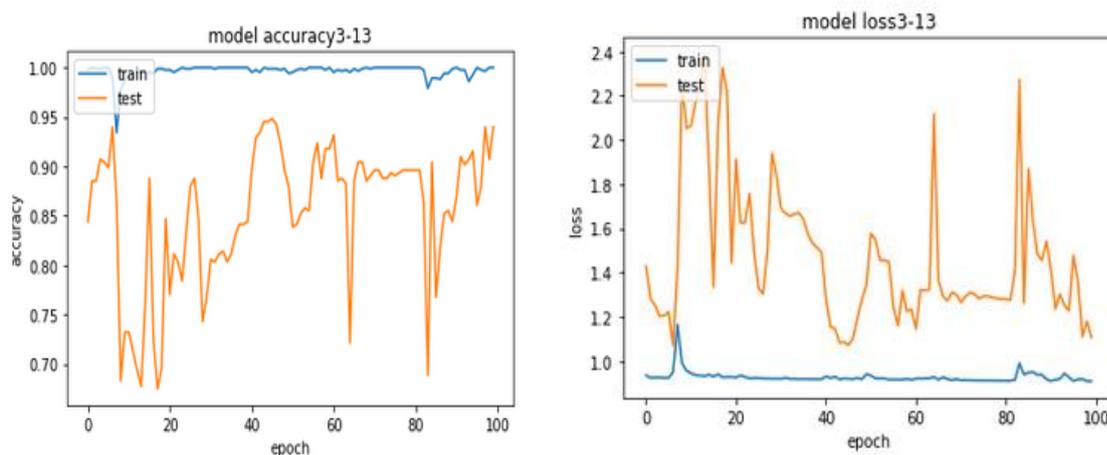


Figure 32: graphe de précision et de perte de la méthode 3.

Nous avons essayé d'ajouter une quatrième couche entièrement connecté, mais on n'a pas obtenu un taux élevé contrairement aux taux obtenus précédemment.

4.5. Méthode 4

Bien que le modèle de la méthode 3 nous a donné un taux acceptable de 93,99% avec 32 batch_size et 100 epochs. Notre objectif été d'atteindre le taux du modèle pré-entraîné avec la base de données « *BreakHis* »

Afin d'améliorer ce taux, nous avons utilisé le même modèle obtenue lors de la dernière méthode avec une modification au niveau de couches entraînées.

Pour ce faire, nous avons entraîné les couches gelées précédemment (les couches entre 210 et 216), et entraîné le modèle avec les mêmes paramètres d'apprentissage (batch_size=30,

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

epochs=100). Après la phase d'entraînement du modèle nous avons atteint le taux 95.90%. Le graphe d'accuracy est présenté dans la figure suivante :

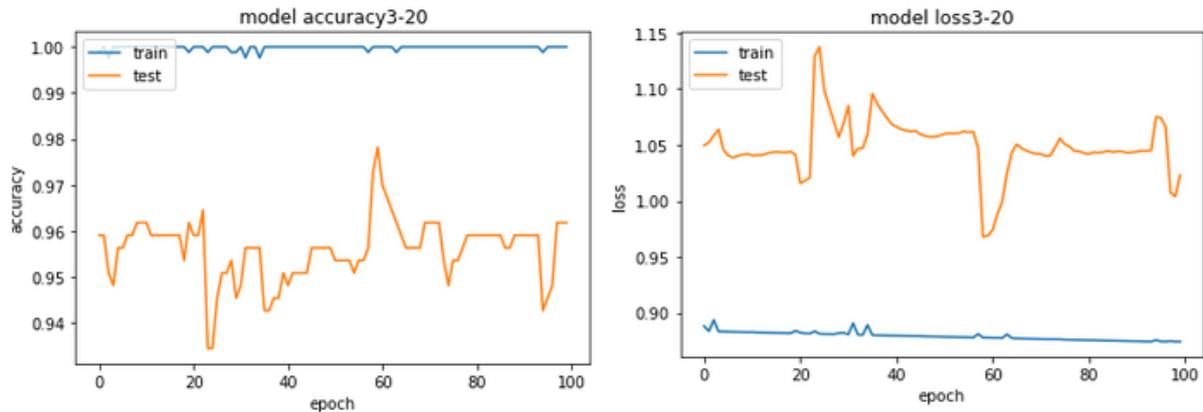


Figure 33: graphe de précision et de perte de la méthode 4.

4.6. Méthode 5

La cinquième méthode est très similaire au modèle de la méthode 4, la modification été au niveau du nombre de neurones de la première couche entièrement connecté. Cette dernière constitué de deux neurones, dont nous avons supprimé les trois couches entièrement connecté et les remplacées par deux autres couches formées de différents nombres de neurones (Figure 34) :

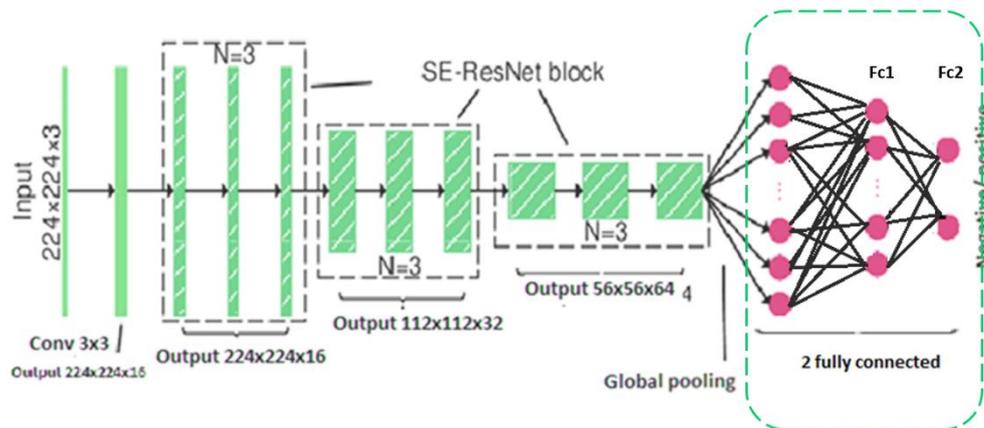


Figure 34: Architecture du transfert d'apprentissage avec la méthode 5.

- **Cas 01** : les deux couches entièrement connectéé : « fc1 » formée de 256 et « fc2 » formée de 2 neurones.

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

- **Cas 02** : les deux couches entièrement connectée : « fc1 » formée de 128 et « fc2 » formée de 2 neurones.
- **Cas 03** : les deux couches entièrement connectée : « fc1 » formée de 64 et « fc2 » formée de 2 neurones.

Nous avons entraîné ces modèles avec `batch_size =32` et `epoch=100`, car ces valeurs sont les valeurs qui nous a donné un bon taux de précision lors des méthodes précédentes.

Après un certain nombre d'itérations d'entraînement nous avons obtenu les taux de précision suivant : (Table 8)

Table 8 : Taux d'apprentissage en changeant le nombre de neurone des couches « fully connected » avec la méthode 5

Taux d'apprentissage (Méthode 5)			
	Cas 01	Cas 02	Cas 03
Accuracy	95.90%	97.27%	95.35%

D'après ce table nous pouvons voir que le meilleur taux de precision est obtenu dans le model du cas 02 qui a donnée un taux de précision de 97.27% avec 100 epochs et `batch_size 32`. La figure 25 montre le graphe de précision et de perte du modèle du cas 02. Nous pouvons remarquer que la précision était stable sur la valeur 59,29% jusqu'à l'époque 49 où la valeur augmente a 60,11% et elle a continué d'augmenter jusqu'au taux de 97.27%. Le graphe du taux de perte confirme le taux de précision. L'autre chose qui confirme la qualité de notre modèle est la similarité très apparaitre entre le graphe d'entraînement et le graphe de test que ce soit dans la précision ou la perte.

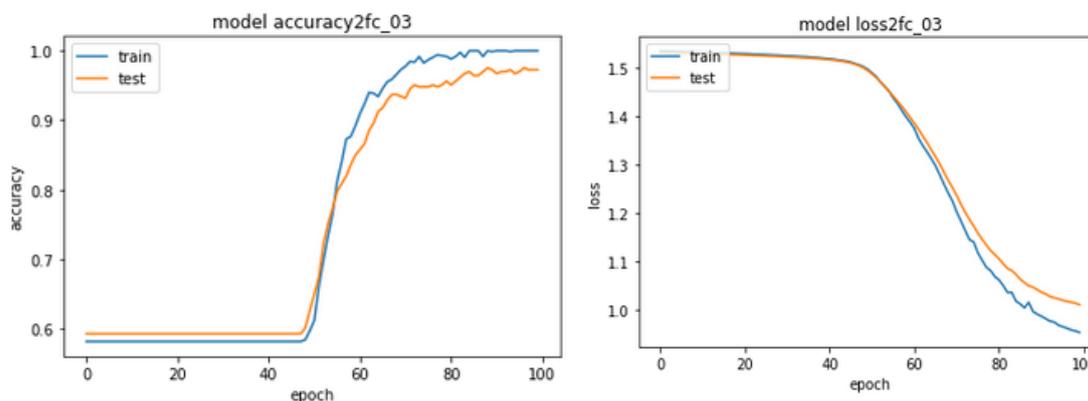


Figure 35: graphe de précision et de perte du modèle de cas 02 de la méthode 5.

5. Discussion

Dans notre travail, nous avons utilisé le modèle pré-entraîné BHCNet qui se caractérise par un taux de précision très élevé et nombre de paramètre très réduit malgré le nombre de couche utilisé. Notre travail consiste à geler, entraîner les couches du modèle BHCNet, aussi à supprimer est ajouter des couches dans le but de classifier notre base de données en E-cadhérine positive et E-cadhérine négative. Les paramètres de BHCNet ont été initialisés en poids pré-entraînés optimisés pour l'ensemble de données BreakHis avec un taux de précision de presque 98%. Dans un premier temps nous avons utilisé le modèle telle qu'il est, où nous avons obtenu un taux de précision de 88,52%. Pour mieux adapter le modèle à notre base de données nous avons utilisé l'apprentissage par transfert.

La première méthode consiste à geler les 3 blocs SE-ResNet afin que le « gradient descente » ne les modifie pas lors de l'apprentissage du nouveau modèle et d'entraîner uniquement la couche entièrement connecté. Nous avons obtenu une précision de 40,71%. Ce taux est très faible par rapport au taux initial du modèle.

Donc nous avons essayé une deuxième méthode « methode2.a » avec le même principe de geler les poids de les trois blocs, et remplacer les deux dernières couches du modèle (*dense_2*, *flatten_2*) par une autre couche de flatten et deux autres couches entièrement connecté entraînable. Le taux obtenu de cette méthode est égale à 86,61% que le considère comme étant une augmentation remarquable du taux.

Pour plus d'amélioration nous avons essayé la « methode 2.b » au principe d'entraîner plus des couches, plus précisément entraîner le 3^{ème} module du 3^{ème} bloc. Le taux obtenu est de (89.89%) après plus modification dans les options de *fitting*.

La méthode 3 consiste à ajouter une troisième couche et appliquer le principe de la methode2.b afin de déterminer si l'ajout de plusieurs couches entièrement connecté rend le modèle de transfert plus efficace ou non. Nous avons trouvé que le taux de précision à augmenter jusqu'à 93.99%.

Dans la méthode 4 nous avons dégelé toutes les couches du modèle de la méthode précédente. Ce dernier est utilisé dans la méthode 5, dans laquelle nous avons modifié leur nombre de neurones de sa première couche entièrement connecté et nous avons obtenu un taux de 97.27%.

Chapitre 3 : Conception et implémentation d'un modèle par le Transfer learning

Nous pouvons résumer les meilleurs résultats de chaque méthode dans le tableau suivant :

Table 9: les meilleurs taux de précision obtenus de chaque méthode

	Batch size	Epochs	Accuracy	Loss	Total de paramètres	Description
Méthode0 (BHCNet)	32	100	88.52%	1.48	14, 832,770	Toutes les couches sont entraînées
Méthode 1	64	100	40.71%	10.0234	14,832,770	Entraîner juste la couche FC
Méthode 2.a	64	100	86,61%	1.3511	14, 832,770	Entraîner les deux couches FC
Méthode 2.b	128	200	89.89%	1.1624	14, 832,776	Fixer les couches entre [-4,-10]
Méthode 3	32	100	93.99%	1.1142	14, 832,782	Ajout de 3eme FC et fixer les couches entre [-4,-10]
Méthode 4	32	100	95,90%	1.0229	14, 832,782	Entraîner toutes les couches du modèle précède
Méthode 5	32	100	97,27%	1.0101	14, 897,666	Changer le nombre de neurones de la couche fc1 (128)

Les résultats obtenus montrent que le choix du nombre d'époque et le batch_size ainsi que la profondeur du réseau ont une grande influence pour avoir des meilleurs résultats.

Nous remarquons que le meilleur taux de précision est celle de la méthode 5. C'est grâce à l'augmentation du nombre de neurones de la couche « fully connected ».

Nous pouvons remarquer aussi que le nombre de paramètres dans les trois premières méthodes est égal à 14832770 (le même nombre de paramètres du modèle pré-entraîné) paramètres car on n'a rien modifié au niveau de modèle pré entraîné. Au contraire aux autres méthodes, le nombre est augmenté à cause de l'ajout des nouvelles couches avec des différents nombres de neurones.

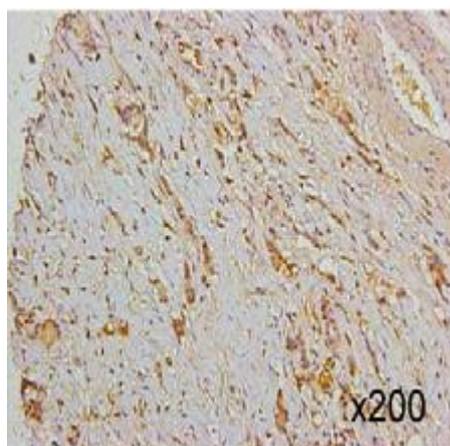
Alors nous pouvons dire que le taux le plus proche que le taux obtenu du modèle pré-entraîné (avec le base de données BreakHis) est celle de la méthode 5 avec un nombre de paramètres satisfaisant, n'est pas plus grand que le nombre de paramètre du modèle original.

6. Génération automatique de légende

Afin de valider notre modèle, nous avons opté pour faire une prédiction sur des nouvelles images à l'aide d'un expert du domaine médicale.

Donc nous avons chargé l'image et la redimensionner à la taille utiliser lors de l'entraînement et le test de 224×224 pixels. Ensuite, nous avons chargé le modèle et lancé la prédiction grâce à l'appel de la fonction « *Predict* ».

Ci-dessous, Les figures illustrent des exemples de légendes générées par notre modèle et des légendes proposées par des spécialistes.



Prédit: E-cadhérine positive (91.76%).

Original : E-cadhérine positive

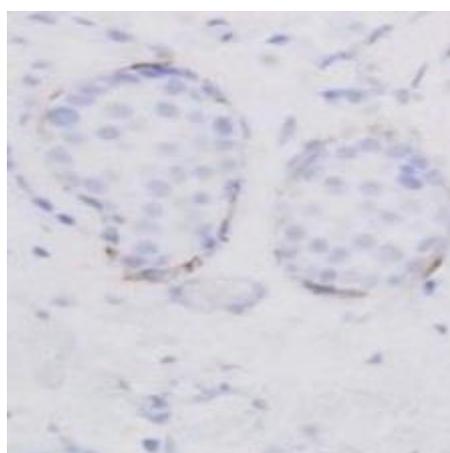


Figure 37: **Prédit:** E-cadhérine négative (88.68%). **Original :** E-cadhérine négative

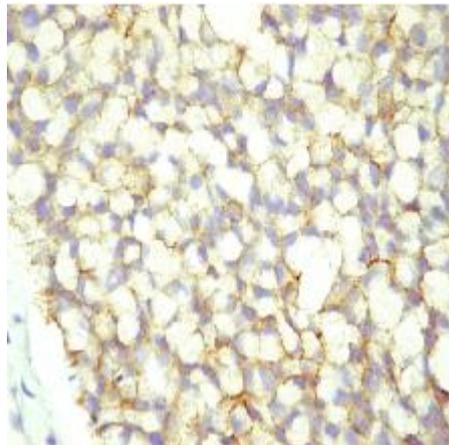
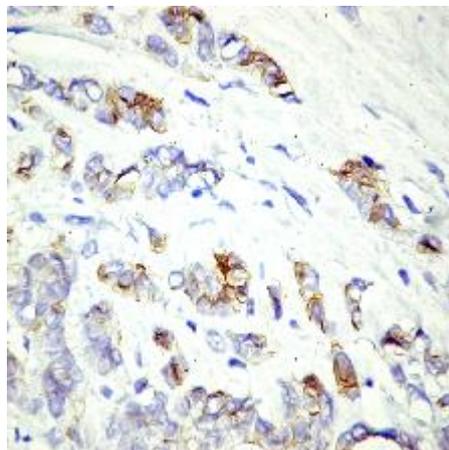
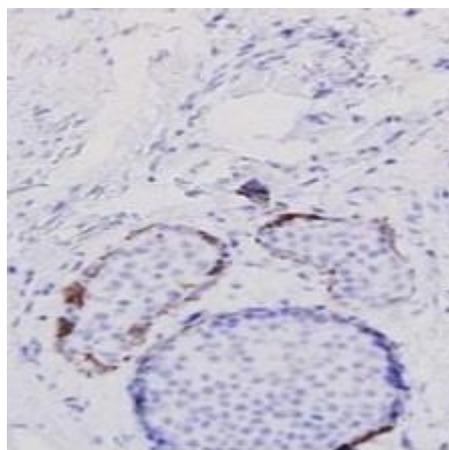


Figure 38: **Prédit:** E-cadhérine positive (91.88%). **Original :** E-cadhérine positive



Prédit : E-cadhérine négative (82.38%).

Original : Non décidable



Prédit : E-cadhérine négative (88.61%).

Original : E-cadhérine négative

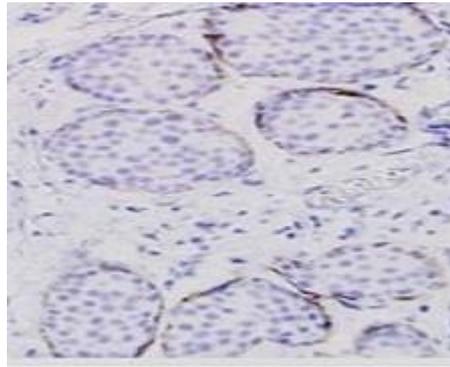


Figure 41: **Prédit**: E-cadhérine négative (84.74%). **Original** : E-cadhérine négative

Conclusion

Dans ce chapitre, nous avons décrit la démarche à suivre afin d'obtenir le classifieur le plus précis possible pour classifier notre base de données en E-cadhérine positive et E-cadhérine négative. Nous avons commencé par l'utilisation d'un modèle pré entraîné BHCNet telle qu'il est. Ce modèle se caractérise par un taux de précision très élevé et un nombre de paramètre très réduit malgré les nombres de couches utilisées. Afin d'améliorer le taux de précision obtenus, nous avons gelé et entraîné les couches du modèle BHCNet, ainsi qu nous avons supprimé est ajouté d'autres couches.

Après plusieurs tests, nous avons obtenu de bons résultats (un taux de précision de 97.27%) avec un nombre de paramètre très réduit (14897666 paramètres) et très proche du notre modèle de base SE-Small-ResNet .Nous espérons pouvoir les utiliser pour aider à choisir le traitement le plus approprié à partir des images histopathologiques colorées avec le E-cadhérine.

CONCLUSION GÉNÉRALE

L'immunohistochimie est une technique largement utilisée pour tester la sensibilité des cellules tumorales telles que le cancer du sein aux hormonothérapies. Cette étude nécessite l'utilisation des anticorps. L'interprétation de l'influence de ces anticorps sur le tissu est rendue par un médecin pathologiste. Cette interprétation nécessite du temps et elle souffre d'une mauvaise reproductibilité même entre les mains de pathologistes experts.

C'est dans ce cadre que s'inscrit notre travail, qui a pour objectif de classifier des images histopathologies colorées avec l'anticorps E-cadhérine pour aider les pathologistes dans leurs travaux à préciser la sensibilité des tissus aux hormonothérapies, afin de faciliter le travail aux oncologues pour l'utilisation du traitement le plus approprié.

Pour réaliser notre travail de classification on a utilisé le deep learning et nous avons choisi les réseaux de neurones convolutionnels et le transfer learning comme des techniques d'apprentissage qui met en évidence un grand succès dans la détection des cellules mitotiques à partir des images histopathologiques.

L'obtention des résultats de classification nécessite un grand nombre de données (images) la chose qui nous a mis en difficulté. Pour résoudre ce problème de manque de données d'apprentissage, nous avons opté à l'utilisation de la technique du « data augmentation » et l'apprentissage par transfert, pour créer un modèle de classification dédié aux lames anapaths des tissus du cancer du sein traité avec des anticorps E-cadhérine. Notre modèle est créé à partir d'un modèle atteint une meilleure performance de prédiction et de paramètres, pré-entraîné sur une large dataset d'images histopathologiques réelles (BreakHis). Le nouveau modèle est à leur tour ré-entraîné sur notre dataset des images histopathologiques colorées par l'anticorps E-cadhérine.

Les résultats obtenus sont très satisfaisants que ce soit sur le taux de précision ou on a atteint un taux de 97.27% avec un nombre de paramètres très réduit et très proche du notre modèle de base SE-Small-ResNet.

Les perspectives de notre travail sont d'arriver à interpréter la sensibilité des tissus cancéreux aux hormonothérapies au biais des lames anapath de base sans l'utilisation des anticorps, en appliquant d'autres méthodes d'augmentation des données. Les futurs objectifs

sont de gagner du temps qui va aider les patients à commencer leurs traitements le plutôt possible et aussi permet à l'état d'éviter le gaspillage d'argent destiné aux achats d'anticorps.

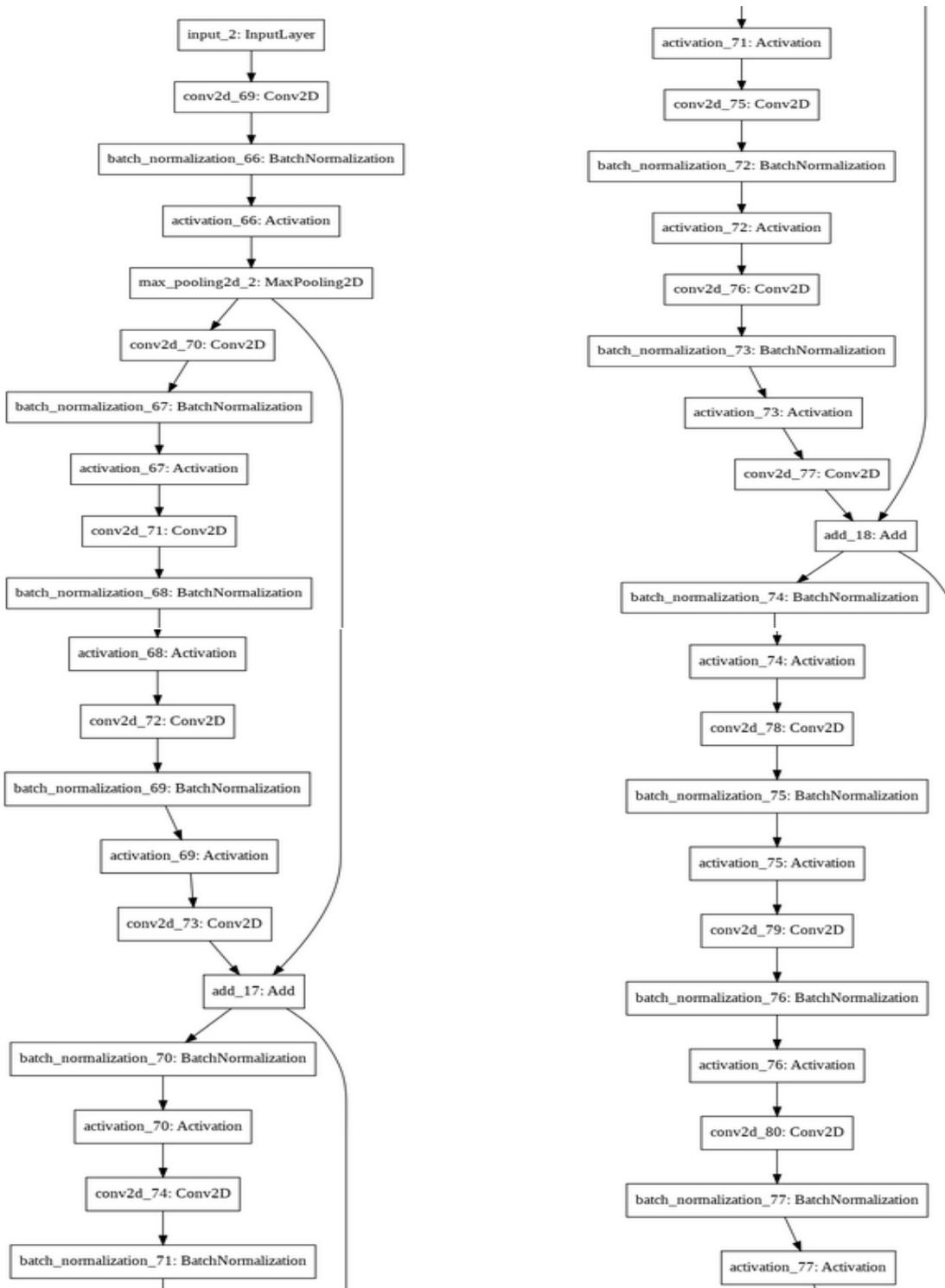
BIBLIOGRAPHIE

- [1] J.-F. Morère, F. Mornex, et D. Soulières, *Thérapeutique du cancer*, 2. éd. Paris: Springer, 2011.
- [2] C. R. Taylor et R. M. Levenson, « Quantification of immunohistochemistry? issues concerning methods, utility and semiquantitative assessment II », *Histopathology*, vol. 49, n° 4, p. 411-424, oct. 2006, doi: 10.1111/j.1365-2559.2006.02513.x.
- [3] B. Ben Cheikh, C. Bor-Angelier, et D. Racoceanu, « A model of tumor architecture and spatial interactions with tumor microenvironment in breast carcinoma », Orlando, Florida, United States, mars 2017, p. 101400C, doi: 10.1117/12.2254452.
- [4] « Cancer du sein - Causes, Symptômes, Traitement, Diagnostic - salutbonjour.ca ». <https://ressourcessante.salutbonjour.ca/condition/getcondition/cancer-du-sein> (consulté le juin 19, 2020).
- [5] C. M. Croce, « Oncogenes and Cancer », *New England Journal of Medicine*, vol. 358, n° 5, p. 502-511, janv. 2008, doi: 10.1056/NEJMra072367.
- [6] « A Gene-Expression Signature as a Predictor of Survival in Breast Cancer | NEJM ». <https://www.nejm.org/doi/full/10.1056/NEJMoa021967> (consulté le juin 21, 2020).
- [7] « glossaire Anatomie Pathologique UE Biopath Tiss 2014-2015.pdf ». .
- [8] M. Intra *et al.*, « Axillary sentinel lymph node biopsy in patients with pure ductal carcinoma in situ of the breast », *Arch Surg*, vol. 138, n° 3, p. 309-313, mars 2003, doi: 10.1001/archsurg.138.3.309.
- [9] « Cancer du sein : anatomie pathologique ». <http://www.dr-karazaitri-ma.com/pages/pour-les-professionnels/pathologies-du-sein/cancer-du-sein/cancer-du-sein-anatomie-pathologique-1.html> (consulté le juin 21, 2020).
- [10] « Manuel de techniques d'anatomo-cytopathologie - 1st Edition ». <https://www.elsevier.com/books/manuel-de-techniques-danatomo-cytopathologie/9782294708442> (consulté le juin 21, 2020).
- [11] « E-Cadherin - IHC Primary Antibodies ». https://shop.leicabiosystems.com/global_fr_FR/ihc-ish/ihc-primary-antibodies/pid-e-cadherin (consulté le juin 21, 2020).
- [12] B. Marr, « A Short History Of Deep Learning -- Everyone Should Read », *Forbes*. <https://www.forbes.com/sites/bernardmarr/2016/03/22/a-short-history-of-deep-learning-everyone-should-read/> (consulté le juin 23, 2020).
- [13] « Day-40 Deep Learning-5 (CNN-1) - PursuitData - Medium ». <https://medium.com/pursuitnotes/day-40-deep-learning-5-cnn-1-5f664998ea86> (consulté le juin 22, 2020).
- [14] Prabhu, « Understanding of Convolutional Neural Network (CNN) — Deep Learning », *Medium*, nov. 21, 2019. <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148> (consulté le juin 22, 2020).
- [15] Y.-W. Chen et L. C. Jain, Éd., *Deep Learning in Healthcare: Paradigms and Applications*, vol. 171. Cham: Springer International Publishing, 2020.

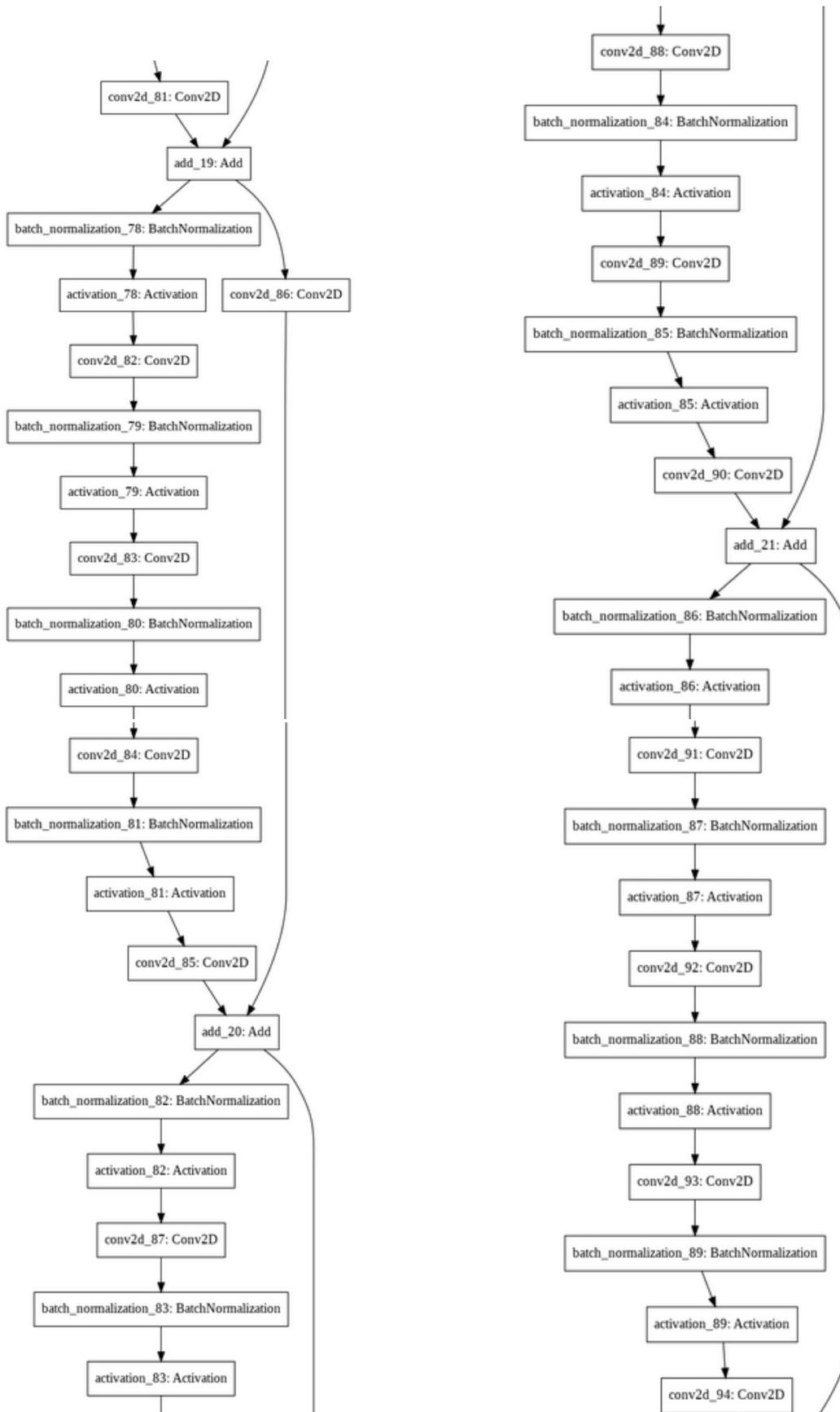
- [16] J. Brownlee, « Transfer Learning in Keras with Computer Vision Models », *Machine Learning Mastery*, mai 14, 2019. <https://machinelearningmastery.com/how-to-use-transfer-learning-when-developing-convolutional-neural-network-models/> (consulté le juin 18, 2020).
- [17] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, et Li Fei-Fei, « ImageNet: A large-scale hierarchical image database », in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, juin 2009, p. 248-255, doi: 10.1109/CVPR.2009.5206848.
- [18] « Krizhevsky et al. - 2017 - ImageNet classification with deep convolutional ne.pdf ». .
- [19] C. Szegedy *et al.*, « Going Deeper with Convolutions », *arXiv:1409.4842 [cs]*, sept. 2014, Consulté le: juin 18, 2020. [En ligne]. Disponible sur: <http://arxiv.org/abs/1409.4842>.
- [20] K. Simonyan et A. Zisserman, « Very Deep Convolutional Networks for Large-Scale Image Recognition », *arXiv:1409.1556 [cs]*, avr. 2015, Consulté le: juin 18, 2020. [En ligne]. Disponible sur: <http://arxiv.org/abs/1409.1556>.
- [21] K. He, X. Zhang, S. Ren, et J. Sun, « Deep Residual Learning for Image Recognition », *arXiv:1512.03385 [cs]*, déc. 2015, Consulté le: juin 18, 2020. [En ligne]. Disponible sur: <http://arxiv.org/abs/1512.03385>.
- [22] J. Hu, L. Shen, S. Albanie, G. Sun, et E. Wu, « Squeeze-and-Excitation Networks », *arXiv:1709.01507 [cs]*, mai 2019, Consulté le: juin 18, 2020. [En ligne]. Disponible sur: <http://arxiv.org/abs/1709.01507>.
- [23] G. Cheng, L. Guo, T. Zhao, J. Han, H. Li, et J. Fang, « Automatic landslide detection from remote-sensing imagery using a scene classification method based on BoVW and pLSA », *International Journal of Remote Sensing*, vol. 34, n° 1, p. 45-59, janv. 2013, doi: 10.1080/01431161.2012.705443.
- [24] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, et C. Bray, « Visual Categorization with Bags of Keypoints », p. 16.
- [25] C.-C. Chang et C.-J. Lin, « LIBSVM: A library for support vector machines », *ACM Trans. Intell. Syst. Technol.*, vol. 2, n° 3, p. 1-27, avr. 2011, doi: 10.1145/1961189.1961199.
- [26] « randomforest2001.pdf ». .
- [27] G. Litjens *et al.*, « A Survey on Deep Learning in Medical Image Analysis », *Medical Image Analysis*, vol. 42, p. 60-88, déc. 2017, doi: 10.1016/j.media.2017.07.005.
- [28] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, et R. Monczak, « Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images », *Computers in Biology and Medicine*, vol. 43, n° 10, p. 1563-1572, oct. 2013, doi: 10.1016/j.combiomed.2013.08.003.
- [29] F. A. Spanhol, L. S. Oliveira, C. Petitjean, et L. Heutte, « A Dataset for Breast Cancer Histopathological Image Classification », *IEEE Trans. Biomed. Eng.*, vol. 63, n° 7, p. 1455-1462, juill. 2016, doi: 10.1109/TBME.2015.2496264.
- [30] P. Wang, X. Hu, Y. Li, Q. Liu, et X. Zhu, « Automatic cell nuclei segmentation and classification of breast cancer histopathology images », *Signal Processing*, vol. 122, p. 1-13, mai 2016, doi: 10.1016/j.sigpro.2015.11.011.
- [31] M. Felsberg, A. Heyden, et N. Krüger, Éd., *Computer Analysis of Images and Patterns*, vol. 10425. Cham: Springer International Publishing, 2017.
- [32] F. A. Spanhol, L. S. Oliveira, C. Petitjean, et L. Heutte, « Breast cancer histopathological image classification using Convolutional Neural Networks », in *2016 International Joint Conference on*

- Neural Networks (IJCNN)*, Vancouver, BC, Canada, juill. 2016, p. 2560-2567, doi: 10.1109/IJCNN.2016.7727519.
- [33] F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, et L. Heutte, « Deep features for breast cancer histopathological image classification », in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Banff, AB, oct. 2017, p. 1868-1873, doi: 10.1109/SMC.2017.8122889.
- [34] Benzhenq Wei, Zhongyi Han, Xueying He, et Yilong Yin, « Deep learning model based breast cancer histopathological image classification », in *2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, Chengdu, China, avr. 2017, p. 348-353, doi: 10.1109/ICCCBDA.2017.7951937.
- [35] H. Garud *et al.*, « High-Magnification Multi-views Based Classification of Breast Fine Needle Aspiration Cytology Cell Samples Using Fusion of Decisions from Deep Convolutional Networks », in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, juill. 2017, p. 828-833, doi: 10.1109/CVPRW.2017.115.
- [36] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, et S. Li, « Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model », *Scientific Reports*, vol. 7, n° 1, Art. n° 1, juin 2017, doi: 10.1038/s41598-017-04075-z.
- [37] S. Akbar, M. Peikari, S. Salama, S. Nofech-Mozes, et A. Martel, « Transitioning between Convolutional and Fully Connected Layers in Neural Networks », *arXiv:1707.05743 [cs]*, juill. 2017, Consulté le: juin 18, 2020. [En ligne]. Disponible sur: <http://arxiv.org/abs/1707.05743>.
- [38] « Zhi et al. - 2017 - Using Transfer Learning with Convolutional Neural .pdf ». .
- [39] X. Zhang *et al.*, « Classifying Breast Cancer Histopathological Images Using a Robust Artificial Neural Network Architecture », in *Bioinformatics and Biomedical Engineering*, vol. 11465, I. Rojas, O. Valenzuela, F. Rojas, et F. Ortuño, Éd. Cham: Springer International Publishing, 2019, p. 204-215.
- [40] D. M. Vo, N.-Q. Nguyen, et S.-W. Lee, « Classification of breast cancer histology images using incremental boosting convolution networks », *Information Sciences*, vol. 482, p. 123-138, mai 2019, doi: 10.1016/j.ins.2018.12.089.
- [41] Y. Jiang, L. Chen, H. Zhang, et X. Xiao, « Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module », *PLOS ONE*, vol. 14, n° 3, p. e0214587, mars 2019, doi: 10.1371/journal.pone.0214587.

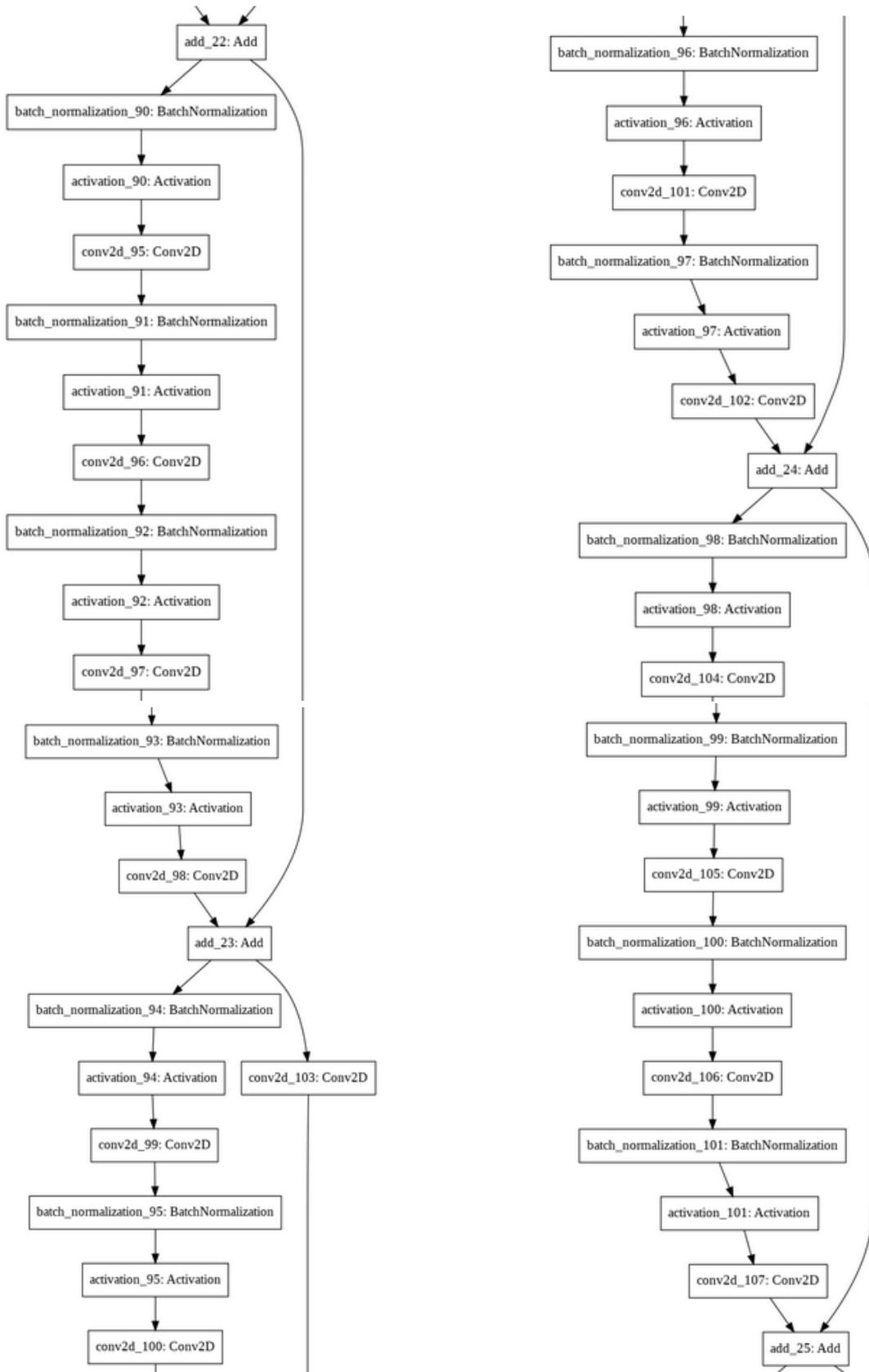
Annexe



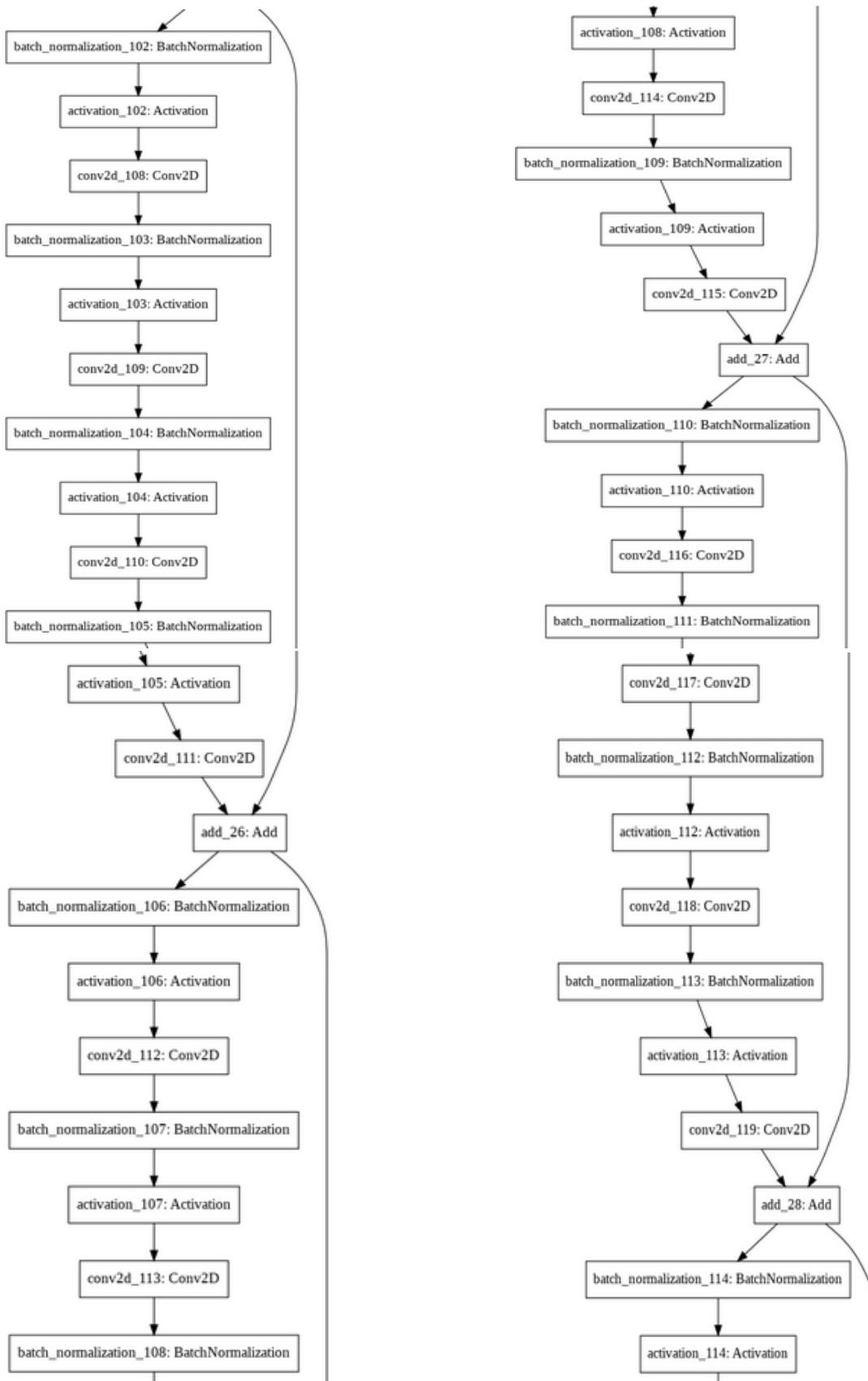
Annexe 1: Architecture et détail du modèle pré entraîné BHCNet.



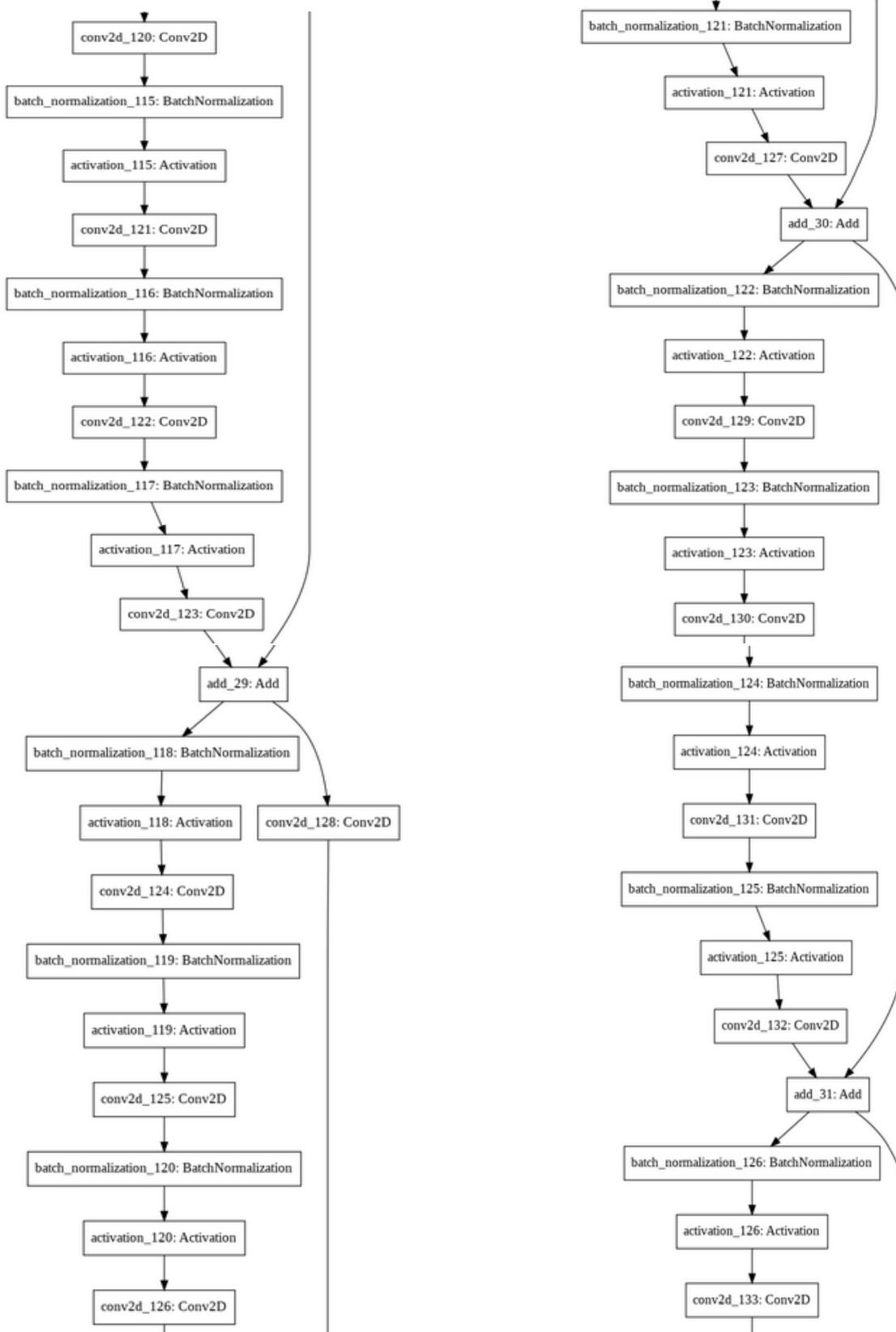
Annexe 1: Architecture et détail du modèle pré entraîné (suite).



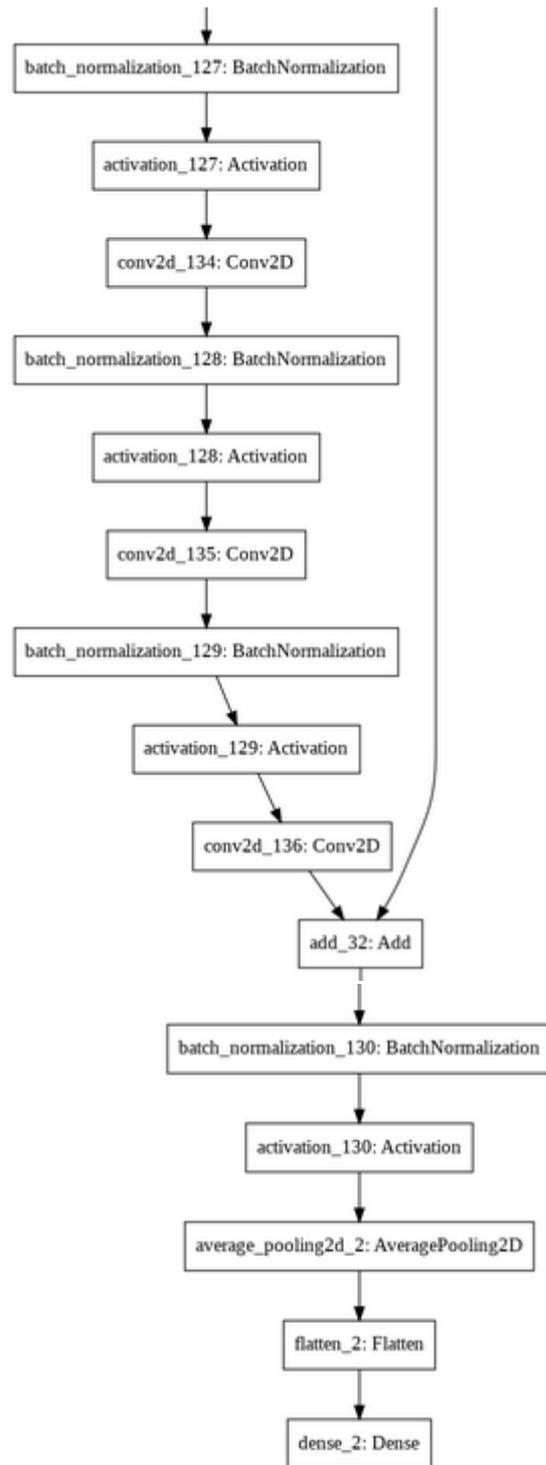
Annexe 1: Architecture et détail du modèle pré entraîné (suite).



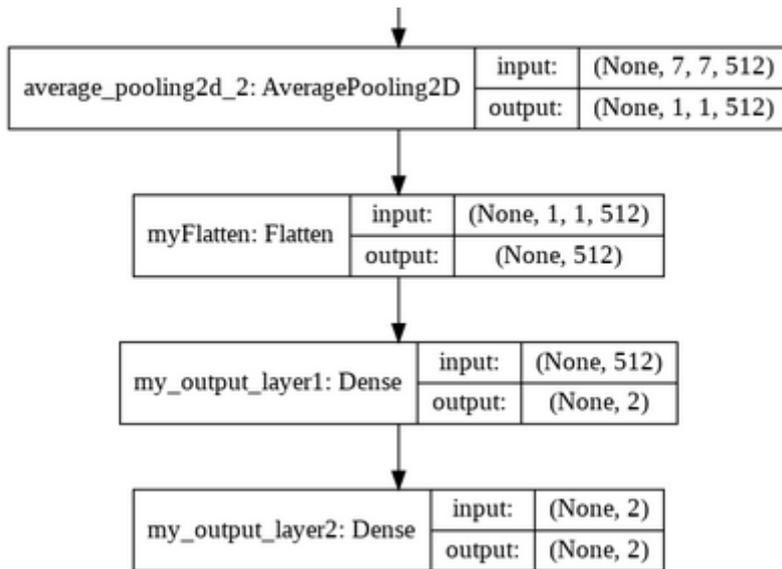
Annexe 1: Architecture et détail du modèle pré entraîné (suite).



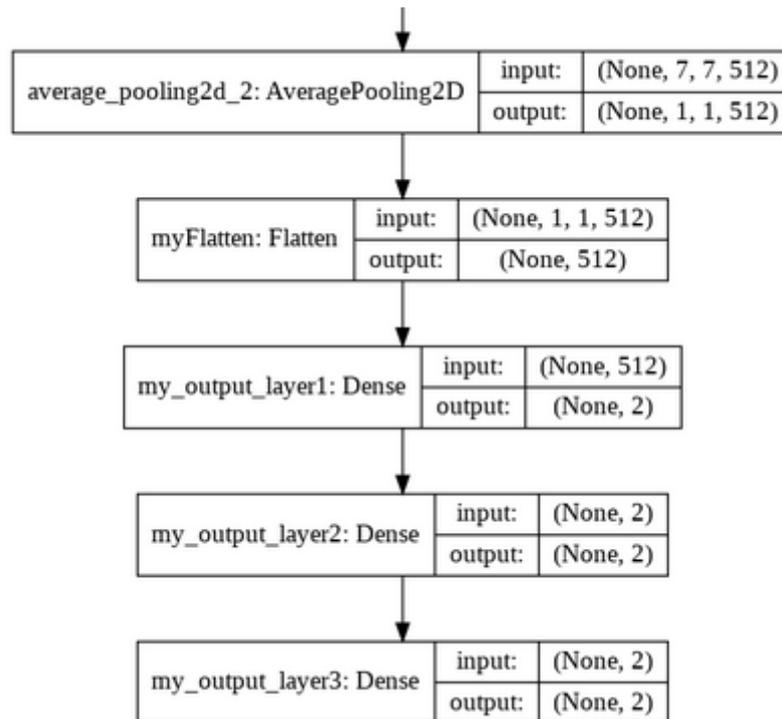
Annexe 1: Architecture et détail du modèle pré entraîné (suite).



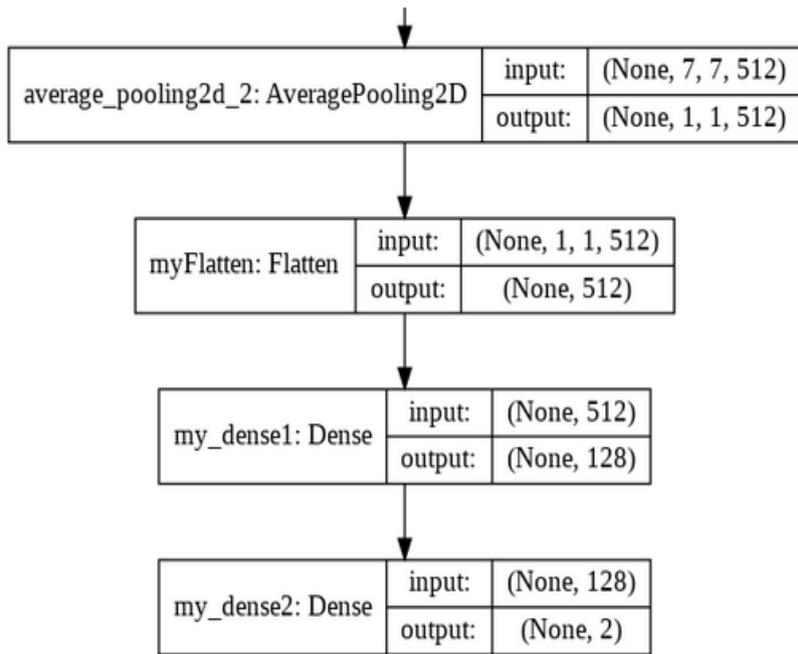
Annexe 1: Architecture et détail du modèle pré entraîné (suite)



Annexe 2: Architecture de dernière partie de modèle de la méthode 2.



Annexe 3: Architecture de dernière partie de modèle de la méthode 3.



Annexe 4: Architecture de dernière partie de modèle de la méthode 5.