

Titre: Vers l'utilisation de l'apprentissage profond aux soins intensifs pour assister la prise en charge de patients ayant subi un traumatisme crânio-cérébral sévère
Title:

Auteur: Sina Fartoumi
Author:

Date: 2024

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Fartoumi, S. (2024). Vers l'utilisation de l'apprentissage profond aux soins intensifs pour assister la prise en charge de patients ayant subi un traumatisme crânio-cérébral sévère [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/58318/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/58318/>
PolyPublie URL:

Directeurs de recherche: Benjamin De Leener, & Guillaume Emeriaud
Advisors:

Programme: Génie biomédical
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Vers l'utilisation de l'apprentissage profond aux soins intensifs pour assister la prise en charge de patients ayant subi un traumatisme crânio-cérébral sévère

SINA FARTOUMI

Institut de génie biomédical

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie biomédical

Avril 2024

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Vers l'utilisation de l'apprentissage profond aux soins intensifs pour assister la prise en charge de patients ayant subi un traumatisme crânio-cérébral sévère

présenté par **Sina FARTOUMI**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Philippe DOYON-POULIN, président

Benjamin DE LEENER, membre et directeur de recherche

Guillaume EMERIAUD, membre et codirecteur de recherche

Michaël SAUTHIER, membre

DÉDICACE

À mes parents.

REMERCIEMENTS

J'aimerais sincèrement remercier mes directeurs de recherches, Professeur Benjamin De Leener et Docteur Guillaume Emeriaud pour leur aide, leur confiance et leur soutien tout au long de ce projet. Sans leur expertise et leurs innombrables conseils, ce travail de recherche n'aurait pas été possible.

Je tiens aussi à remercier le regroupement en intelligence artificielle appliquée aux soins aigus de l'enfant du centre Hospitalier Universitaire de Sainte-Justine et plus particulièrement Professeur Rita Moumeir et Docteur Philippe Juvet pour l'aide financière qu'ils m'ont accordée.

RÉSUMÉ

Le traumatisme craniocérébral (TCC) sévère est une des principales causes de décès chez les enfants et jeunes adultes. Pour réduire les lésions cérébrales secondaires et les séquelles, le patient doit être pris en charge et stabilisé rapidement après la blessure initiale. Son état général est maintenu en surveillant ses fonctions hémodynamiques et respiratoires et l'état craniocérébral est monitoré par plusieurs signaux physiologiques. La complexité du TCC sévère, le contexte aigu, et la grande quantité d'information compliquent la prise en charge. Des outils informatiques d'aide à la décision clinique basés sur des techniques d'intelligence artificielle ont déjà été appliqués à différents aspects de la prise en charge de patients traumatisés aux soins intensifs, mais une aide diagnostique automatique et en temps réel de l'état craniocérébral précis du patient n'est pas disponible. Ce projet de recherche a pour objectif d'évaluer la capacité de réseaux de neurones artificiels profonds comme le Fully Convolutional Network et le Convolutional Neural Network-Long Short-Term Memory à produire un diagnostic en continu de l'état craniocérébral de patients après TCC sévère. Les signaux de pression intracrânienne, pression de perfusion cérébrale, oxygénation cérébrale et pression artérielle de 49 patients avec TCC sévère ont été enregistrés pour créer une base de données totalisant plus de 4000 heures d'enregistrement. Ces données ont été analysées rétrospectivement par des experts cliniciens et l'état craniocérébral a été classifié selon six états d'intérêt : situation contrôlée, hypertension intracrânienne (HTIC), HTIC avec risque d'ischémie, HTIC avec risque d'hyperhémie, ischémie ou hyperhémie sans HTIC. Cette base de données est utilisée pour entraîner et valider les algorithmes de classification supervisés à l'étude. Leurs performances sont mesurées en fonction de la précision du diagnostic, de la sensibilité et du score F1 pour chaque état. Les performances des deux réseaux de neurones sont comparées avec celles de classificateurs supervisés de type SVM et Random Forest. Les résultats obtenus montrent de faibles performances en ce qui concerne la classification des six états pathophysiologiques. Cependant, en simplifiant le problème de classification, les réseaux de neurones démontrent une performance satisfaisante pour la détection de l'hypertension intracrânienne. Il est toutefois recommandé de mettre au point une base de données regroupant un plus grand nombre de sujets avec TCC sévère et avec une classification diagnostique réalisée prospectivement par les experts afin de pouvoir mieux explorer l'utilité des techniques d'apprentissage automatique dans le diagnostic en temps réel de l'état de patient avec TCCS aux soins intensifs.

ABSTRACT

Severe traumatic brain injury (TBI) is one of the leading causes of death in children and young adults. To reduce secondary brain injury and sequelae, the patient must be treated and stabilized quickly after the initial injury. The general condition is maintained by monitoring the hemodynamic and respiratory functions, and the craniocerebral state is monitored by several physiological signals. The complexity of severe TBI, the acute context, and the large quantity of information complicate the management of these patients. Computer tools for clinical decision support based on artificial intelligence techniques have already been applied to different aspects of the management of traumatized patients in intensive care units, but automatic real-time diagnostic assistance of precise craniocerebral state of patients with severe TBI is not available. This research project aims to evaluate the capacity of deep learning artificial neural networks such as Fully Convolutional Network and Convolutional Neural Network-Long Short-Term Memory to produce a continuous diagnosis of the craniocerebral state of patients after severe TBI. Intracranial pressure, cerebral perfusion pressure, cerebral oxygenation and blood pressure signals from 49 patients with severe TBI were recorded to create a database totaling more than 4000 hours of recording. These data were retrospectively analyzed by clinical experts in craniocerebral trauma and the craniocerebral state was classified according to six states of interest: controlled situation, intracranial hypertension (IH), IH with risk of ischemia, IH with risk of hyperemia, ischemia or hyperemia without IH. This database was used to train and validate the supervised classification algorithms under study. Their performance is measured based on diagnostic accuracy, sensitivity and F1 score for each condition. The performances of the two neural networks are compared with those of supervised SVM and Random Forest classifiers. The results obtained show poor performance with regard to the classification of the six pathophysiological states. On the other hand, by simplifying the classification problem, neural networks demonstrate satisfactory performance for the detection of intracranial hypertension. However, it is recommended to develop a database bringing together a larger number of subjects with severe TBI, prospectively annotated by clinical experts to better explore the usefulness of machine learning techniques in the real-time diagnosis of the patient's condition with severe traumatic brain injury in the intensive care units.

TABLE DES MATIÈRES

DÉDICACE.....	III
REMERCIEMENTS	IV
RÉSUMÉ.....	V
ABSTRACT	VI
TABLE DES MATIÈRES	VII
LISTE DES TABLEAUX.....	XI
LISTE DES FIGURES.....	XII
LISTE DES SIGLES ET ABRÉVIATIONS	XIV
LISTE DES ANNEXES.....	XVI
CHAPITRE 1 INTRODUCTION.....	1
1.1 La problématique : le traumatisme crânio-cérébral.....	1
1.2 Projet OptiBrain	2
1.3 Objectif de recherche	3
1.4 Organisation du mémoire	3
CHAPITRE 2 REVUE DE LITTÉRATURE	4
2.1 Le traumatisme crânio-cérébrale.....	4
2.1.1 Causes.....	4
2.1.2 Type de blessure.....	5
2.1.3 Gravité de la blessure	5
2.1.4 Les signaux d'intérêts.....	6
2.1.5 Pathophysiologie	8
2.1.6 Prise en charge aux soins intensifs.....	9
2.2 Système d'aide à la décision clinique	9

2.2.1	Un bref historique.....	10
2.2.2	SADC pour le traumatisme crânio-cérébral	10
2.3	L'apprentissage automatique.....	11
2.3.1	Les séries temporelles	12
2.3.2	Classification des séries temporelles	12
2.3.3	Les réseaux de neurones artificiels.....	13
2.3.4	Convolutional Neural Network	14
2.3.5	Fully Convolutional Neural Network.....	15
2.3.6	Recurrent Neural Network	15
2.3.7	Méthodes classiques	15
2.4	L'apprentissage automatique pour le traitement du traumatisme crânio-cérébral	16
2.5	Résumé de la revue de littérature	19
CHAPITRE 3 MÉTHODOLOGIE.....		20
3.1	Objectifs	20
3.2	Données.....	20
3.2.1	Distribution par état pathophysiologique	21
3.2.2	Distribution par patient.....	22
3.3	Mise en forme des données	23
3.3.1	Séquences	23
3.3.2	Extraction des caractéristiques	24
3.3.3	Analyse en composantes principales.....	26
3.4	Sous échantillonnage de l'état contrôlé.....	26
3.5	Stratégie d'entraînement	27
3.6	Réseaux de neurones profonds	28

3.6.1	Le FCN	28
3.6.2	Le CNN-LSTM	29
3.6.3	Paramètres d'entraînement	30
3.6.4	Choix des hyperparamètres	30
3.7	Classificateurs classiques	31
3.7.1	Le SVM	31
3.7.2	Le Random Forest	31
3.8	Types d'expériences	31
3.8.1	Classification à six états	31
3.8.2	Classification binaire	31
3.9	Métriques de performance	32
3.10	Langage de programmation et librairies	34
3.11	Structure du programme	35
3.11.1	Traitement des données	36
3.11.2	Classification et performances	37
CHAPITRE 4	RÉSULTATS	38
4.1	Classification en six états	38
4.2	Détection de l'hypertension intrâ-cranienne	42
4.3	Détection de l'hyperhémie	46
4.4	Détection de l'ischémie	51
4.5	Résumé des résultats	58
CHAPITRE 5	DISCUSSION	61
5.1	Diagnostic à six états	61
5.2	Détection d'un état pathophysiologique	61

5.3	Comparaison entre classificateurs	62
5.4	Limitation de la base de données	63
5.5	Améliorations possibles concernant la méthode	64
5.5.1	Méthode création des séquences temporelles.....	64
5.5.2	Validation croisée.....	64
CHAPITRE 6	CONCLUSION	65
RÉFÉRENCES		67
ANNEXES		73

LISTE DES TABLEAUX

Tableau 2.1 Détail du score Glasgow.....	6
Tableau 4.1 Métriques du FCN sur le groupe de développement.....	39
Tableau 4.2 Métriques de la classification du réseau FCN pour six états.....	40
Tableau 4.3 Métriques de la classification du réseau CNN-LSTM pour six états	41
Tableau 4.4 Métriques pour la détection de l'HTIC du FCN.....	42
Tableau 4.5 Métriques pour la détection de l'HTIC du CNN-LSTM.....	43
Tableau 4.6 Métriques pour la détection de l'HTIC du SVM.....	44
Tableau 4.7 Métriques pour la détection de l'HTIC du RF	45
Tableau 4.8 Métriques du réseau FCN pour la détection de l'HYP.....	46
Tableau 4.9 Métriques du réseau FCN pour la détection de l'HYP v CTRL	47
Tableau 4.10 Métriques du réseau CNN-LSTM pour la détection de l'HYP	48
Tableau 4.11 Métriques du réseau CNN-LSTM pour la détection de l'HYP v CTRL.....	49
Tableau 4.12 Métriques du SVM pour la détection de l'HYP	50
Tableau 4.13 Métriques du RF pour la détection de l'HYP	51
Tableau 4.14 Métriques du FCN pour la détection de l'ISC.....	52
Tableau 4.15 Métriques du FCN pour la détection de l'ISC v CTRL	53
Tableau 4.16 Métriques du CNN-LSTM pour la détection de l'ISC	54
Tableau 4.17 Métriques du CNN-LSTM pour la détection de l'ISC v CTRL.....	55
Tableau 4.18 Métriques du SVM pour la détection de l'ISC.....	56
Tableau 4.19 Métriques du RF pour la détection de l'ISC.....	57
Tableau B.1 Métriques du CNN-LSTM sur le groupe de développement.....	76

LISTE DES FIGURES

Figure 3.1 Distribution des données associé à chaque état pathophysiologique.....	22
Figure 3.2 Nombre de patient associé à chaque état	23
Figure 3.3 Mise en forme des séquences temporelles.....	23
Figure 3.4 Étapes de développement des modèles de classificateurs	28
Figure 3.5 Architecture du réseau FCN retenue.....	29
Figure 3.6 Architecture du réseau CNN-LSTM retenue	30
Figure 3.7 Représentation d'une matrice de confusion.....	33
Figure 3.8 Processus de traitement des données	36
Figure 3.9 Processus de la classification et l'évaluation des performances.....	37
Figure 4.1 Matrice de confusion du FCN sur le groupe de développement.....	38
Figure 4.2 Matrice de confusion du réseau FCN pour la classification à six états	40
Figure 4.3 Matrice de confusion du CNN-LSTM pour six états.....	41
Figure 4.4 Matrice de confusion du FCN pour la détection de l'HTIC	42
Figure 4.5 Matrice de confusion du CNN-LSTM pour la détection de l'HTIC.....	43
Figure 4.6 Matrice de confusion du SVM pour la détection de l'HTIC	44
Figure 4.7 Matrice de confusion du RF pour la détection de l'HTIC	45
Figure 4.8 Matrice de confusion du réseau FCN pour la détection de l'HYP	46
Figure 4.9 Matrice de confusion du réseau FCN pour la détection de l'HYP v CTRL	47
Figure 4.10 Matrice de confusion du réseau CNN-LSTM pour la détection de l'HYP.....	48
Figure 4.11 Matrice de confusion du réseau CNN-LSTM pour la détection de l'HYP v CTRL ..	49
Figure 4.12 Matrice de confusion du SVM pour la détection de l'HYP.....	50
Figure 4.13 Matrice de confusion du RF pour la détection de l'HYP	51

Figure 4.14 Matrice de confusion du FCN pour la détection de l'ISC	52
Figure 4.15 Matrice de confusion du FCN pour la détection de l'ISC v CTRL	53
Figure 4.16 Matrice de confusion du CNN-LSTM pour la détection de l'ISC.....	54
Figure 4.17 Matrice de confusion du CNN-LSTM pour la détection de l'ISC v CTRL	55
Figure 4.18 Matrice de confusion du SVM pour la détection de l'ISC	56
Figure 4.19 Matrice de confusion du RF pour la détection de l'ISC	57
Figure 4.20 Comparaison des métriques de classification pour N-HTIC v HTIC	58
Figure 4.21 Comparaison des métriques de classification pour les états N-HYP v HYP	59
Figure 4.22 Comparaison des métriques de classification pour les états CTRL v HYP	59
Figure 4.23 Comparaison des métriques de classification pour les états N-ISC v ISC	60
Figure 4.24 Comparaison des métriques de classification pour les états CTRL v ISC.....	60
Figure A.1 Règles de décision pour l'HTIC	73
Figure A.2 Règles de décision pour l'ISC et l'HYP basé sur le Licox (PbTO ₂).....	73
Figure A.3 Règles de décision pour l'ISC et l'HYP basé sur la PPC	74
Figure A.4 Table de décision finale pour l'ISC et l'HYP basé sur le Licox et la PbTO ₂	74
Figure B.1 Apprentissage et de validation (fonction de coût et précision) : CNN-LSTM	75
Figure B.2 Apprentissage et de validation (fonction de coût et précision) : FCN	75
Figure B.3 Matrice de confusion du CNN-LSTM sur le groupe de développement	76

LISTE DES SIGLES ET ABRÉVIATIONS

AD	Arbre de Décision (Decision tree)
ATP	Adénosine Triphosphate
AUC	Area Under the Receiver Operatig Characteristic Curve
CNN	Convolutional Neural Network
CTRL	État Contrôlé
DC	Craniectomie Décompressive
DTW	Dynamic Time Warping
EEG	Électroencéphalometrie
FCN	Fully Convolutional Network
FN	Faux Négatif
FP	Faux Positif
GAP	Global Average Pooling
GPU	Processseur d'Unité Graphique (Graphical Processing Unit)
HTIC	Hypertension Intracrânienne
HTIC-H	Hypertension Intracrânienne avec Hyperhémie
HTIC-I	Hypertension Intracrânienne avec Ischémie
HYP	Hyperhémie
IA	Intelligence Artificielle
ISC	Ischémie
LCR	Liquide Céphalo-Rachidien
LSTM	Long Short-Term Memory
PAM	Pression Artérielle Moyenne
PCA	Analyse en Composante Principale (Principal Component Analysis)
PIC	Pression Intracrânienne
PPC	Pression de Perfusion Cérébrale
PPCO	Pression de Perfusion Cérébrale Optimale
PRx	Index d'Autorégulation Cérébrale
PbtO ₂	Oxygénation du tissu cérébral
RBF	Radial Basis Function (network)

RNA	Réseau de Neurones Artificiels
RF	Random Forest
SADC	Système d'Aide à la Décision Clinique
SI	Soins Intensifs
ST	Série Temporelle
SVM	Support Vector Machine
TCC	Traumatisme Crânio-Cérébral
TCCS	Traumatisme Crânio-Cérébral Sévère
VN	Vrai Négatif
VP	Vrai Positif

LISTE DES ANNEXES

ANNEXE A règles de décisions de l'état crânio-cérébral	73
ANNEXE B Résultats supplémentaires	75

CHAPITRE 1 INTRODUCTION

1.1 La problématique : le traumatisme crânio-cérébral

Le traumatisme crânio-cérébral (TCC) présente un fardeau énorme sur les systèmes de santé à travers le monde. Il est estimé qu'annuellement entre 27 et 69 millions de personnes souffrent d'une forme de TCC [1]. Sous sa forme la plus grave, le traumatisme crânio-cérébral sévère (TCCS), est une cause majeure de décès, tandis que les survivants voient leur qualité de vie grandement réduite dû aux séquelles causant des handicaps graves, limitant ainsi leur contribution socioéconomique. À titre d'exemple, aux États-Unis seulement, l'impact économique du TCC a été évalué à 76 milliards de dollars en coûts directs et indirects pour l'année 2010 [2].

En plus des dommages irréversibles infligés immédiatement après le traumatisme, des lésions secondaires se développent rapidement dans le reste du cerveau, aggravant son état tout en augmentant le risque de séquelles permanentes à long terme. Ces lésions secondaires sont aggravées par toutes les instabilités des fonctions hémodynamiques du patient et un mauvais contrôle de la pression et de perfusion sanguine du cerveau [3]. Une prise en charge optimale mais aussi très complexe devient donc primordiale durant les premiers jours après le traumatisme. Une surveillance neurologique multimodale aux soins intensifs (SI) spécialisés est essentielle pour guider les soins [4]. Cependant, l'analyse de la condition cérébrale peut être difficile pour un médecin non spécialisé puisque plusieurs variables doivent être prises en considération simultanément et que l'état du cerveau du patient peut changer très rapidement.

De plus, dans le contexte surchargé des SI où une très grande quantité d'informations doit être analysée rapidement, il existe un risque de mauvaise interprétation de l'état crânio-cérébral ce qui peut causer un délai dans l'identification d'un changement de l'état cérébral du patient. Bien que des lignes directrices pour la prise en charge de patients avec un TCCS existent, elles ne sont pas toujours respectées ce qui peut engendrer des soins de moins bonne qualité et donc un impact négatif sur le devenir du patient [5]. Les systèmes d'aide à la décision clinique (SADC) mettent à profits la capacité de calculs très rapide des systèmes informatiques dans l'analyse de grande quantité de données. Ceux-ci présentent donc un grand potentiel afin de venir en aide au personnel soignant dans des situations médicales complexes et délicates comme celles présentes aux SI.

1.2 Projet OptiBrain

De façon générale, les progrès technologiques permettent maintenant d'acquérir et analyser plus facilement une grande quantité d'informations relatives aux patients. La gestion de ces données et le développement d'outils informatiques capables de les interpréter afin de guider les professionnels de la santé lors de situations complexes sont des problèmes qui suscitent aujourd'hui beaucoup d'intérêts.

Le sujet de ce présent mémoire s'inscrit dans un projet d'envergure initié par le regroupement en IA pour les soins aigus de l'enfant au Centre Hospitalier Universitaire Sainte-Justine en collaboration avec Polytechnique Montréal visant au développement d'un SADC pour les TCCS chez l'enfant. Ce projet, qui a vu le jour en 2015, a pour objectif ultime d'optimiser la prise en charge des enfants avec TCCS et augmenter l'adhérence aux lignes directrices. Le système appelé OptiBrain mesure en temps réel la stabilité générale, l'état cérébral, et la qualité de la prise en charge du patient avec TCCS. Pour ce faire, le système analyse plusieurs signaux physiologiques provenant de divers instruments médicaux qui sont enregistrés dans une base de données cliniques dynamique.

L'état pathophysiologique crânio-cérébral du patient traumatisé étant un élément important de la prise en charge du TCCS, il doit être monitoré et stabilisé de façon continue. Un outil informatique capable de détecter automatiquement l'état crânio-cérébral d'un patient avec TCCS devient donc très intéressant pour guider le personnel soignant [6] d'autant que cette analyse est complexe. Ainsi, l'équipe d'OptiBrain travaille à la mise au point d'une fonctionnalité capable de fournir une recommandation sur l'évaluation de l'état cérébral du patient de façon précise, et en temps réel, avec suggestion de classification parmi six différents états : Contrôlé (CTRL), hypertension intracrânienne (HTIC), HTIC avec ischémie (HTIC-I), HTIC avec hyperhémie (HTIC-H), ischémie (ISC) et hyperhémie (HYP). Pour ce faire, des règles diagnostiques médicales, mises au point par des experts en TCC, ont été programmées et appliquées aux signaux physiologiques d'intérêts et une suggestion diagnostique est émise au 30 secondes. Les signaux utilisés sont : la pression intracrânienne, la pression de perfusion cérébrale, l'oxygénation cérébrale et la pression artérielle.

Avec l'émergence d'algorithmes d'intelligence artificielle (IA) de plus en plus performants comme les réseaux de neurones artificiels (RNA) à apprentissage profond (qui sont notamment

utilisés dans d'autres domaines de diagnostic médical), il est important d'évaluer si l'utilisation de tels algorithmes permettrait une approche par apprentissage profond pour assister le personnel soignant dans la prise de charge de patients avec un TCCS aux SI.

1.3 Objectif de recherche

Dans le cadre du projet de recherche présenté ici, nous nous intéressons à évaluer l'application de RNA à apprentissage profond sur des données temporelles physiologiques pour le diagnostic en continu de l'état crânio-cérébral. Plus particulièrement, nous nous intéressons à déterminer si l'apprentissage profond peut fournir une suggestion diagnostique de l'état crânio-cérébral similaire à celle d'un expert.

1.4 Organisation du mémoire

Au chapitre suivant, une revue de littérature est proposée afin de mieux comprendre les particularités du TCCS. Des éléments théoriques sur l'apprentissage machine sont aussi abordés avant de discuter des travaux existants dans le domaine et de mieux situer le présent travail. Le chapitre 3 détaille la méthodologie adoptée dans le cadre de ce projet. Le chapitre 4 présente les résultats obtenus qui feront l'objet d'une discussion au chapitre 5. Enfin, une conclusion et des recommandations seront faites au chapitre 6.

CHAPITRE 2 REVUE DE LITTÉRATURE

Dans ce chapitre, le TCC est abordé plus en détail ; sa cause, le type de blessure et sa gravité, sa pathophysiologie, les signaux physiologiques d'intérêt et sa prise en charge clinique sont discutés. Puis, les outils informatiques existant pour le support clinique dans le traitement des TCC sont présentés. Ensuite, les notions théoriques nécessaires à la compréhension des algorithmes d'apprentissage machine utilisés dans ce travail sont présentées. Enfin, l'état de l'art en ce qui concerne les applications de l'intelligence artificielle à la prise en charge et au traitement du TCC sont discutés afin de mieux situer le présent projet de recherche.

2.1 Le traumatisme crânio-cérébrale

L'encéphale est la composante la plus importante du système nerveux central. En conjonction avec la moelle épinière et les nerfs, il contrôle l'activité de notre corps, nos mouvements, nos pensées, nos actions et interactions, et nos sentiments en analysant les informations sensorielles qui lui sont transmises. Il est constitué de 100 milliards de cellules nerveuses et bien que son poids ne représente que 2% de celle du corps, il requiert d'énormes ressources pour fonctionner normalement.

Le traumatisme crânio-cérébral survient lorsqu'une force externe appliquée à la tête cause un dommage au crâne ou au tissu cérébral altérant son fonctionnement normal. Au-delà de l'impact important sur le patient lui-même (risque de mortalité et de séquelles), son impact économique sur la société est considérable. Chaque cas de TCC induit des coûts annuels estimés à 400 000 \$US [7] et sachant qu'entre 100 et 300 admissions en SI par 100 000 personnes et par an sont liées à un TCC, il est facile de comprendre que le traumatisme crânio-cérébral représente un fardeau énorme pour les patients, les familles et le système de santé.

2.1.1 Causes

Parmi les causes du TCC, la simple chute est la plus fréquente, particulièrement chez les enfants et les personnes âgées de plus de 65 ans. Chez les enfants de moins de 14 ans, elle en est la cause dans 50% des cas. Les accidents de route, de sport, l'impact par différents types d'objets et des actes de violence (notamment par maltraitance chez les nourrissons) expliquent aussi les TCC [2].

2.1.2 Type de blessure

Le TCC est catégorisé par le mécanisme causal. Lorsque le TCC est causé par l'intrusion d'un objet ou d'un fracas d'os au-delà de la dure-mère des méninges, il se dit un TCC pénétrant ou ouvert. Ce type de TCC est souvent la cause d'un choc direct à la tête induisant une contusion localisée, c'est-à-dire que les dommages se limitent à la région bien précise de l'encéphale où l'impact a eu lieu. Lorsque le crâne et la dure-mère restent intacts, le TCC en question se dit fermé. Il est causé par un choc direct ou par un contrecoup du cerveau sur la paroi du crâne à la suite d'une accélération ou d'une décélération brusque. Ces types de blessures peuvent toucher une ou plusieurs régions. Dans le deuxième cas, la blessure est dite diffuse. Un type particulier de TCC fermé causé par accélération suivie d'une brusque décélération de la tête est la lésion axonale diffuse. Dans ce cas, les axones de la matière blanche du cerveau sont endommagés voire rompus du fait des distorsions et étirement du cerveau lors du traumatisme. Ce type de TCC est observé dans près de 50% des TCC sévères, touche particulièrement les enfants, et est associé à plus de séquelles permanentes et sévères [8].

2.1.3 Gravité de la blessure

Il est possible de classifier la sévérité du TCC en fonction de la présence de fracture du crâne et de la durée de la période d'inconscience à la suite de la blessure initiale. Le score de Glasgow est toutefois un indicateur plus précis qui permet d'évaluer l'état de conscience d'un patient. Il se base sur l'évaluation de trois critères : la capacité d'ouvrir les yeux, la réponse verbale ainsi que la réponse motrice. Selon ce système, une note globale comprise entre 3 et 15 est attribuée au patient. Elle correspond à la somme des scores des trois critères montrés au Tableau 2.1. Plus ce score est bas, plus l'état du patient est critique. Dans le cas d'un TCC, un score au-delà de 13 correspond à un traumatisme léger. Dans ce cas, la vie du patient n'est pas en danger et ne souffre généralement pas de séquelles graves à long terme. Un score compris entre 12 et 9 définit un traumatisme modéré qui peut être accompagné de séquelles à long terme. Un score de 8 ou moins indique la présence d'un TCC sévère qui met en danger la vie et le pronostic du patient. Il est d'ailleurs la première cause de mortalité chez les moins de 45 ans dans plusieurs pays [7,9].

La complexité de l'encéphale et les enjeux pronostiques font du TCCS une condition extrêmement difficile à gérer. Les dommages subis diffèrent en fonction du mécanisme ayant causé la blessure. Les régions touchées déterminent souvent les fonctions qui seront affectées. De

façon générale, les survivants d'un TCCS peuvent souffrir de handicaps physiques comme la paralysie partielle ou complète. Les séquelles incluent aussi surtout des troubles comportementaux, de mémoire, de concentration, et d'apprentissage qui limitent leur inclusion à part entière au sein de la société [6], d'où l'importance d'optimiser la prise en charge de tels patients au SI afin de maximiser leur chance de survie et de minimiser les séquelles à long terme.

Tableau 2.1 Détail du score Glasgow

Réponse oculaire		Réponse verbale		Réponse motrice	
Description	Score	Description	Score	Description	Score
Nulle	1	Nulle	1	Nulle	1
À la douleur	2	Incompréhensible	2	Extension stéréotypée	2
À la demande	3	Inappropriée	3	Flexion stéréotypée	3
Spontanée	4	Confuse	4	Évitement	4
		Normale	5	Orientée	5
				Aux ordres	6

2.1.4 Les signaux d'intérêts

À la suite de l'évaluation de la gravité de la blessure, les fonctions hémodynamiques et respiratoires du patient avec TCCS doivent être stabilisées et son état crânio-cérébral monitoré de façon continue en surveillant notamment les signaux physiologiques suivants : pression intracrânienne (PIC), pression artérielle moyenne (PAM), pression de perfusion cérébrale (PPC) et l'oxygénation cérébrale (PbtO₂). L'indice d'autorégulation cérébrale (PRx) et la pression de perfusion cérébrale optimale (PPCO) sont des variables qui sont calculées durant la prise en charge et permettent la personnalisation de la prise en charge qui est liée à un meilleur taux de survie chez les patients souffrant de TCCS [10].

2.1.4.1 Pression intracrânienne

Le crâne contient l'encéphale incompressible, le liquide céphalorachidien (LCR) circulant dans les cavités et autour de l'encéphale ainsi que le sang artériel et veineux. La somme du volume de ces éléments exerce une pression dans le crâne rigide et inextensible, qui s'appelle la pression intracrânienne et dont la valeur normale se situe en moyenne entre 10 et 15 millimètres de mercure. Les liquides contenus à l'intérieur du crâne ne peuvent être compressés mais leur volume peut changer. Toute augmentation de volume d'un de ces compartiments après un TCC (par exemple par œdème du tissu cérébral, par accumulation de LCR lors de blocage des mécanismes de réabsorption, ou par vaso-dilatation des vaisseaux cérébraux) provoque une augmentation de la PIC. Au-delà de 20 millimètres de mercure pour une période soutenue, il y a présence d'une HTIC. Si cette pression n'est pas contrôlée, elle peut avoir de graves conséquences sur le fonctionnement du cerveau jusqu'à mettre la vie du patient en danger. La PIC est mesurée habituellement de façon invasive à l'aide de cathéter intra parenchymateux ou intraventriculaire [11].

2.1.4.2 Pression artérielle moyenne

La pression artérielle moyenne est la pression du sang dans les artères de la circulation principale ou systémique. Elle est normalement maintenue autour de 90 millimètres de mercure.

2.1.4.3 Pression de perfusion cérébrale

La pression de perfusion cérébrale est la pression à l'origine de la perfusion sanguine cérébrale. Elle est définie par la différence entre la PAM et la PIC. La valeur normale de la PPC se trouve entre 60 et 70 millimètres de mercure chez l'adulte [10] et entre 40 et 70 chez l'enfant en fonction de l'âge.

2.1.4.4 Oxygénation du tissu cérébral

L'oxygénation locale du tissu cérébral (P_{btO_2}) est mesurée à l'aide de la sonde invasive du système Licox [12]. La plage normale se situe entre 20 et 35 millimètres de mercure. Cette mesure est prise en compte dans la détection de l'état ischémique ou d'hyperhémie cérébrale.

2.1.4.5 Index d'autorégulation cérébrale

L'index d'autorégulation cérébrale est défini comme le coefficient de corrélation mobile entre la PAM et la PIC. L'optimisation de son calcul est le sujet de plusieurs travaux de recherche [13] visant à déterminer la plage de temps et la fréquence d'échantillonnage optimale à considérer. De

façon générale, il est globalement accepté que la valeur moyenne sur une minute de la PAM et la PIC échantillonnée au moins à 60 Hz produit un PRx acceptable [14]. Sa valeur, comprise dans la plage $[-1,1]$, donne une indication des capacités du cerveau à maintenir une autorégulation cérébrale intacte après le TCCS. Une valeur inférieure à 0 indique une autorégulation cérébrale correcte alors que des valeurs supérieures à 0.3 sont associées à une augmentation de la morbidité [13].

2.1.4.6 Pression de perfusion cérébrale optimale

Le PRx varie selon la valeur de la PPC. Sur une période de temps donnée, la PPC correspondant à la valeur minimale du PRx est appelée la pression de perfusion optimale (PCCO). Son calcul nécessite souvent la considération de quelques heures d'enregistrement [10]. C'est donc une donnée qui n'est pas disponible dans les premières heures de la prise en charge du patient.

2.1.5 Pathophysiologie

Les lésions neuronales subies lors de l'impact sont appelées lésions primaires. Ces dommages encourus sont irréversibles et seules des mesures préventives peuvent permettre de les réduire.

La gestion des patients traumatisés se fait principalement en évitant les lésions secondaires. Celles-ci surviennent quelques heures à quelques jours après l'insulte primaire. Elles sont causées par la destruction des cellules nerveuses initialement saines mais en territoire fragilisé par le traumatisme. Elles sont favorisées par l'état instable du patient et par tout élément entraînant un manque d'apport sanguin dans les territoires fragilisés.

L'insulte primaire provoque un dérèglement du métabolisme cérébral ainsi que la perte de la régulation du débit sanguin cérébral. L'ischémie induite initialement provoque l'accumulation d'acide lactique dû à une glycolyse anaérobie et provoque la perte de l'adénosine triphosphate (ATP) des cellules nerveuses. Ces dernières perdent leur capacité à maintenir l'intégrité de leur barrière cytoplasmique induisant un influx d'eau et d'électrolytes à l'intérieur de la cellule. En second lieu, une dépolarisation des membranes terminales du neurone produit un excès de neurotransmetteurs tel que le glutamate et l'aspartate. S'en suit une augmentation de la concentration en ions calcium et sodium à l'intérieur des cellules ce qui déclenche un mécanisme de dégradation cellulaire.

Le cerveau est doté de la capacité d'autorégulation, un mécanisme lui permettant de réguler son débit sanguin en fonction du métabolisme cérébral et des fluctuations de la pression artérielle afin de maintenir un apport sanguin adéquat en tout temps. Le dérèglement voire la perte de l'autorégulation cérébrale est fréquente chez les traumatisés crâniens sévères et est associé à une augmentation de la pression intracrânienne ainsi qu'au développement d'ischémie neuronale, deux conditions liées à des pronostics négatifs. Le maintien de la PPC à l'intérieur de limites assurant un apport sanguin suffisant tout en évitant une augmentation de la pression intracrânienne est donc d'une grande importance dans la gestion du malade.

2.1.6 Prise en charge aux soins intensifs

La prise en charge du TCCS doit débuter le plus tôt possible sur le lieu même de l'accident ou du traumatisme. Les premiers soins ont pour but d'éviter principalement l'ischémie et l'hypotension tout en transférant le patient vers un centre hospitalier possédant idéalement une unité spécialisée en soins neuro intensifs. La prise en charge hospitalière initiale effectue un bilan lésionnel et une chirurgie d'urgence pour évacuation d'hématome intracrânien peut dans certains cas être effectuée [7]. Les fonctions hémodynamiques et respiratoires sont surveillées et l'état général du patient est stabilisé. La neuro-réanimation consiste à lutter contre le développement d'une HTIC tout en maintenant la PPC dans la zone optimale. Puis le monitoring des principaux signaux physiologiques présentés à la section 2.1.4 est fait en continu pour minimiser le développement de lésions secondaires et maintenir l'état crânio-cérébral du patient dans une zone contrôlée. De façon générale, des protocoles standardisés complexes ont été mis en place afin d'optimiser la prise en charge du TCCS [4]. En lien avec ces algorithmes de prise en charge, des règles de décisions sur l'état crânio-cérébral de patients avec TCCS ont été mises en place afin de classer l'état du patient parmi les suivants: l'état contrôlé, l'ischémie, l'hyperhémie, l'hypertension intracrânienne avec ischémie ou avec hyperhémie. Ces règles, présentées à l'annexe A, témoignent de la complexité du processus décisionnel requis pour optimiser la prise en charge du TCCS.

2.2 Système d'aide à la décision clinique

Le diagnostic de l'état crânio-cérébral de patient avec TCCS est donc un processus complexe impliquant l'analyse de multiples sources de données dans un environnement déjà surchargé. Le médecin traitant, outre sa connaissance approfondie et souvent très spécialisée, doit pouvoir

analyser les informations à sa disposition efficacement. Cette facette dépend souvent de son niveau d'expérience. Cependant, l'exposition à une énorme quantité de données peut mener à une diminution de sa capacité d'analyse et de synthèse. Cette diminution est liée aux limitations de la mémoire de travail et de la capacité attentionnelle dans la gestion d'information abondante [15]. En ce sens, les SADC sont des systèmes informatiques ayant pour but d'assister les médecins et autres professionnels de la santé durant différentes phases de la prise en charge de patients afin d'augmenter la qualité des soins.

2.2.1 Un bref historique

Vers la fin des années 50, Ledley et al. [16] analysent le processus décisionnel des professionnels de la santé afin de l'automatiser. En 1964, Warner et al. [17] proposent un premier prototype pour le diagnostic des anomalies congénitales du cœur basé sur le théorème de Bayes. Similairement, Lodwick et al. [18] proposent son prototype destiné au diagnostic des tumeurs osseuses. C'est toutefois dans les années 70 que trois systèmes considérés aujourd'hui comme les ancêtres des systèmes modernes ont vu le jour. Dombal et al. [19] ont mis au point le Leeds Abdominal Pain System qui permet de déceler la cause d'un malaise abdominal. Ce système, aussi basé sur le théorème de Bayes, considère une indépendance entre les différents facteurs ainsi qu'une exclusivité mutuelle de sept diagnostics possibles. Shortliffe et al. [20], proposent en 1973 leur système MYCIN, basé sur une base de connaissance composée de règles de décisions, pour l'identification du type de bactéries à la source d'une infection dans le but de prescrire des antibiotiques appropriés. Le système HELP, initialement conçu pour le diagnostic d'anomalies cardiaques et déployé en 1976, a permis de surpasser le taux de succès des cliniciens dans la prescription d'un plan de traitement avec 94% d'efficacité contre 77% [21]. Le système s'est depuis transformé en un outil de gestion de l'information complet dans les hôpitaux et permet l'intégration de différents modules pour l'aide à la décision dans plusieurs départements comme les laboratoires, la radiologie et les soins intensifs. Le système actuel permet de produire des alertes lors de suivi de patients et il émet aussi, à la demande du médecin, des suggestions sur le type de traitement à prescrire.

2.2.2 SADC pour le traumatisme crânio-cérébral

Dans cette section-ci, un résumé des différents SADC n'utilisant pas l'IA pour la prise en charge des TCC sévères est présenté. Le logiciel ICM+ développé par Smielewski et al. [22] est l'outil le

plus avancé pour l'enregistrement et la présentation des signaux relatifs au TCC. Bien qu'il permette d'enregistrer les signaux vitaux des patients en temps réel et certaines analyses et calculs (comme les paramètres d'autorégulation), il ne permet aucune forme d'analyse complexe de type classification diagnostique. Gomez et al. [23] proposent un outil similaire facilitant la déduction de variables d'intérêts telles que l'indice d'autorégulation cérébrale et la pression de perfusion cérébrale critique. Cependant, aucun de ces deux outils ne produit de diagnostic quant à l'état du patient.

Cet aspect est totalement laissé aux mains du personnel soignant. L'outil développé par Wilson et al [24], permet de recueillir les signaux en temps réel directement à partir des différents capteurs au chevet du patient. Il permet d'activer différentes alarmes en fonction de simples règles programmables. La complexité des règles programmables étant limitée, les événements détectables ne sont pas nombreux et ne couvrent que quelques situations cliniques simplistes telles qu'une augmentation de la pression intracrânienne chez un patient. Wu et al.[25] proposent de suivre l'état du patient en continu pour détecter toute déviation par rapport aux dernières recommandations médicales internationalement acceptées. En fonction des déviations détectées, une recommandation de traitement est émise par le système afin de stabiliser l'état du malade. Cependant, les variables prises en compte sont limitées et ne permettent donc pas d'analyser l'état de chaque patient de façon personnalisé.

2.3 L'apprentissage automatique

L'apprentissage automatique est un domaine combinant des techniques d'optimisation, de statistiques et d'informatique. Cette branche de l'intelligence artificielle s'intéresse au développement d'algorithmes capables d'acquérir des connaissances à partir d'observations et de données automatiquement sans instruction explicitement définie. Il existe trois principaux types d'apprentissage automatique : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. Le dernier type d'apprentissage est hors du contexte de ce travail de recherche et ne sera pas discuté ici.

Les algorithmes d'apprentissage supervisés utilisent des observations labellisées où les données en entrées sont associées à une valeur de sortie. Le processus consiste à apprendre les liens existants entre les variables d'entrées et le résultat associé. Concrètement, il s'agit d'approximer une fonction f qui puisse établir le lien entre l'espace d'entrées X et l'espace en sortie Y . Lorsque

sortie Y appartient au domaine des valeurs discrètes, c'est un problème de classification ou de reconnaissance de forme. S'il appartient au domaine des réels, c'est un problème de régression. Par ailleurs, les algorithmes non supervisés n'utilisent que des observations non labellisées. Ils s'intéressent plutôt à identifier les relations existantes entre ceux-ci afin de les séparer en groupes hétérogènes. Ces techniques sont aussi utiles pour la réduction de dimensionnalité de bases de données multivariées complexes.

Le diagnostic médical étant un problème similaire à la classification par méthode d'apprentissage automatique supervisé, des concepts de base permettant de mieux comprendre les méthodes à l'étude dans ce travail sont présentés dans les prochaines sections.

2.3.1 Les séries temporelles

Les données physiologiques enregistrées aux SI forment des séries temporelles (ST) ordonnées permettant de suivre l'évolution de l'état du patient. Par définition, une série temporelle X est une suite de valeurs ordonnée dans le temps : $X = [x_1, x_2, x_3, \dots, x_T]$ où X appartient au domaine des réels R^d , T est la longueur de la suite et d la dimension du vecteur. Si $d = 1$ la ST est dite univariée. Sinon, elle est multivariée.

2.3.2 Classification des séries temporelles

Soit une base de données d'entraînement $D = \{(X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)\}$ composée de n séries temporelles X_i ainsi que son étiquette y_i , le but d'un modèle de classification est d'apprendre la fonction f de sorte que $f(X_i) = y_i$ puisse prédire l'étiquette de nouvelles observations. La nature temporelle de ces séries implique habituellement une corrélation entre les points de données consécutifs. Cependant, toutes les méthodes d'apprentissage supervisées ne sont pas adaptées à cette particularité.

La classification de ST est un sujet qui suscite beaucoup d'intérêt et plusieurs techniques sont proposées dans la littérature. Ces techniques sont regroupables par leur mode de fonctionnement; estimation de métrique de distance, extractions de caractéristiques et les réseaux de neurones artificiels profonds.

Parmi les méthodes basées sur l'estimation de distance entre des séries temporelles, la métrique du Dynamic Time Warping (DTW) utilisée avec un simple classificateur de type plus proches voisins s'est établie comme une très bonne référence dans le domaine. Cependant, des recherches

ont démontré que des techniques utilisant un ensemble d'arbres de décision appelé Random Forest (RF) ou les Support Vector Machine (SVM) peuvent être de bonnes alternatives [26]. Ces techniques requièrent néanmoins une extraction des caractéristiques préalable. Avec l'émergence des réseaux de neurones profonds dans l'analyse du langage naturel qui présentent des similitudes avec les ST, les RNA suscitent aujourd'hui beaucoup d'intérêt dans les problèmes de classification de données temporelles.

2.3.3 Les réseaux de neurones artificiels

Les RNA sont des modèles d'apprentissage machine dont le fonctionnement est inspiré du fonctionnement du cerveau humain. De façon générale, ils peuvent être représentés comme l'interconnexion de plusieurs couches composées elles-mêmes d'éléments unitaires appelés des neurones. Le RNA comprend une couche d'entrée de données, une ou plusieurs couches de traitement cachée et une couche de sortie.

Le réseau de neurones profond se compose de L fonctions paramétriques. Ce sont les différentes couches qui sont considérées comme étant chacune une représentation du domaine d'entrée. Chaque couche l_i , de sorte que $i \in 1 \dots L$, est composé de neurones qui calculent un élément de sortie de chaque couche. Puis, chaque couche l_i utilise en entrée la sortie de la couche précédente l_{i-1} et y applique une non-linéarité (ex: la fonction sigmoïde ou la fonction redresseur) pour générer sa propre sortie. Ces transformations non linéaires sont contrôlées par un ensemble de paramètres θ_i appelés poids. Étant donnée une entrée x , la sortie f_L du réseau peut être représenté par l'équation 2.1 :

$$f_L(\theta_L, x) = f_{L-1}(\theta_{L-1}, f_{L-2}(\theta_{L-2}, \dots, f(\theta_1, x))) \quad 2.1$$

f_i est l'opération non-linéaire appliquée à la couche l_i . Cette équation représente la propagation avant (feedforward en anglais) dans les RNA. Le processus d'apprentissage du réseau consiste à présenter la base de données d'entraînement contenant une série d'observations labélisées afin de trouver les coefficients optimaux des paramètres. Pour ce faire, dans un premier temps les poids sont initialisés de façon aléatoire et une entrée de la base de données est présentée à la fonction représentée par l'équation 2.1. La sortie du réseau est obtenue en appliquant la propagation avant avec cette entrée. Elle correspond à un vecteur représentant une estimation des probabilités d'appartenance à chaque classe de l'entrée. Puis, la perte associée à la prédiction du réseau est

évaluée selon une fonction de coût. Pour les problèmes de classification à classes multiples, la fonction de coût d'entropie croisée (categorical cross-entropy en anglais) est un choix bien adapté. Finalement en utilisant l'algorithme de descente du gradient [27], les poids sont ajustés par la rétropropagation de l'erreur dans le réseau. En somme, en utilisant la propagation avant suivie d'une rétropropagation de façon itérative, il devient possible d'ajuster les paramètres du réseau afin de minimiser la perte sur la base de données d'entraînement. Durant la phase d'inférence ou de test, le modèle est testé sur de nouvelles observations. La propagation avant de ces nouvelles observations produit des probabilités d'appartenance à chaque classe. La prédiction du réseau correspond à la classe avec la plus grande probabilité.

2.3.4 Convolutional Neural Network

Les Convolutional Neural Network (CNN) sont des méthodes qui se sont montrées très performantes pour des tâches de reconnaissance d'images ou de traitement du langage naturel [28,29] puisqu'ils préservent l'aspect spatial des données d'entrées. Ils sont aussi utilisés dans l'analyse des séries temporelles. Le principe de fonctionnement de ce type de réseau repose sur la convolution de différents filtres sur des données. Dans le cas des séries temporelles, le filtre est d'une seule dimension (la dimension temporelle) et peut être considéré comme une transformation non-linéaire. En appliquant de multiples filtres sur une entrée, il devient possible d'y extraire différentes propriétés utiles à la classification. Les paramètres des filtres doivent être appris lors du processus d'apprentissage de sorte que la classification soit la plus optimale possible. Pour assurer que le CNN est insensible à de petites translations des données temporelles en entrée, une couche de regroupement (pooling en anglais) est ajoutée à la suite de la couche de convolution. Cette couche consolide les caractéristiques des couches de convolution et réduit la représentation spatiale. Ceci réduit aussi le nombre de paramètres du réseau déjà très complexe. Des couches de normalisation sont aussi souvent ajoutées pour faciliter la convergence lors de l'apprentissage. Une couche permettant la classification est nécessaire, elle prend en entrée les résultats de convolution pour donner une probabilité d'appartenance à une classe. Un réseau de neurones classique est typiquement utilisé à cette fin. Le processus d'entraînement des réseaux CNN est similaire à celui discuté dans la section précédente.

2.3.5 Fully Convolutional Neural Network

Le réseau Fully Convolutional Network (FCN) proposé par Wang et al. [30] pour la classification des séries temporelles se distingue principalement par le remplacement de la dernière couche dense du réseau CNN par une couche Global Average Pooling (GAP). Ceci réduit grandement le nombre de paramètres du réseau. Il réduit aussi l'aspect boîte noire typiquement associé aux RNA. En effet des techniques compatibles avec ces réseaux sont en mesure d'expliquer à un certain degré la prise de décision du FCN [31]. De plus, l'évaluation de ce réseau sur 44 différentes bases de données de séries temporelles démontre que cette architecture présente de meilleures performances de classification que certains réseaux beaucoup plus complexes.

2.3.6 Recurrent Neural Network

Les réseaux récurrents se différencient de l'architecture classique des réseaux de neurones comme discuté à la section 2.3.3. Le RNN possède des boucles de rétroactions entre les différents nœuds du réseau permettant de tenir compte d'informations antérieures dans une séquence donnée. Même s'il est adapté à l'analyse et la classification de séries temporelles, il est difficile à entraîner principalement à cause du phénomène appelé fuites de gradient (Vanishing gradient) [32] faisant qu'il ne peut pas apprendre les relations sur une longue période. Afin d'adresser ces limitations importantes, différentes variantes de RNN ont été proposées. Parmi eux, le réseau Long Short-Term Memory (LSTM) est une des alternatives les plus intéressantes [33].

Chaque unité d'un LSTM est composée de plusieurs mécanismes de déclenchement. À savoir un mécanisme de mise à jour de la mémoire, d'oubli de l'information et de gestion de la sortie qui permettent collectivement de retenir de l'information sur une plus grande durée.

2.3.7 Méthodes classiques

Contrairement aux algorithmes présentés plus haut, les principes de fonctionnement de plusieurs types de classificateurs ne sont pas adaptés aux problèmes des séries temporelles. Il n'est pas possible d'utiliser les séries temporelles brutes. Il est nécessaire d'extraire des caractéristiques statistiques de ces données temporelles afin d'utiliser des classificateurs comme le SVM ou le Random Forest.

2.3.7.1 SVM

De façon générale, un SVM est un classificateur supervisé binaire cherchant à établir un hyperplan capable de diviser les observations appartenant à deux classes de sorte que la distance entre ce plan et les instances les plus proches de chaque classe soit maximale.

En utilisant l'astuce du noyau [34], il devient possible d'utiliser un classificateur linéaire pour des problèmes non linéaires. Il suffit de transformer l'espace des données d'entrées dans un espace de plus grande dimension dans lequel un classificateur linéaire pourra faire une discrimination satisfaisante entre les observations des deux classes. Pour ce faire, plusieurs types de fonction de noyau ont été définis.

Ce type de classificateur a souvent démontré des performances égales ou même supérieures à certains réseaux de neurones dans plusieurs domaines expliquant sa popularité [26].

2.3.7.2 Random Forest

L'algorithme Random Forest est un classificateur qui fait partie des méthodes d'apprentissage par ensemble puisqu'il se compose de plusieurs classificateurs appelés arbre de décision. Chaque arbre de décision du RF est entraîné sur une sous-partie de la base de données d'entraînement. Chaque arbre de décision représente donc un classificateur spécialisé. Le RF considère le résultat de classification de chaque arbre afin de produire une prédiction plus robuste et bien souvent plus exacte.

De façon générale, l'entraînement d'un arbre de décision a pour but de construire une structure qui révèle les relations hiérarchiques entre les variables d'une base de données [35]. Elle a pour but d'expliquer une variable cible à partir de variables explicatives.

2.4 L'apprentissage automatique pour le traitement du traumatisme crânio-cérébral

Afin de faciliter la lecture, seules les applications utilisant des algorithmes similaires à ceux discutés à la section précédente sont présentés ici. Il n'en reste pas moins que le domaine de l'apprentissage automatique est très vaste et que plusieurs autres types d'algorithmes peuvent tout à fait convenir selon l'application.

Deux revues récentes et exhaustives de la littérature concernant l'application de l'intelligence artificielle dans la prise en charge du TCC aux SI faites par Rajaei et al. [36] ainsi que Kim et al.

[37] montrent que les techniques d'apprentissage automatique peuvent être appliquées à plusieurs étapes de la prise en charge des patients. La majorité des travaux suivants, utilise la métrique de performance Area Under the Receiver Operating Characteristic Curve (AUC). Plus cette valeur est proche de 1, plus les performances de classification du modèle sont bonnes.

Durant la phase préhospitalière et dès l'admission aux SI, il est important de rapidement pouvoir évaluer l'étendue et la gravité des blessures du patient. Hale et al. [38] ont entraîné un RNA capable, à partir d'images tomographiques, de déterminer avec une grande précision si un patient pédiatrique souffre d'un TCC assez important justifiant sa prise en charge aux SI. Guler et al. [39] utilisent un RNA pour déterminer la sévérité du TCC au SI en se basant sur le score Glasgow et les données d'électroencéphalogramme (EEG). Ce modèle est capable de reproduire la classification faite par des experts avec une précision de 87%. Abe et al. [40] utilisent les données recueillies par les premiers répondants sur le site de l'accident comme la pression artérielle, le rythme cardiaque, la température corporelle et le rythme respiratoire pour déterminer avec un arbre de décision (AD) si le patient souffre d'hémorragie intracrânienne. Le classificateur proposé a d'ailleurs montré de meilleures performances de diagnostic que plusieurs autres algorithmes d'apprentissage automatique (AUC = 0.80). Moyer et al. [41] proposent aussi une variante d'un AD (AUC = 0.81) afin de déterminer si les patients avec TCC modéré ou sévère nécessitent une neurochirurgie dans les premières 24h de la prise en charge. Pour se faire, l'algorithme se base sur quinze variables relevées durant la phase préhospitalière comme le score Glasgow, la pression artérielle et la saturation en oxygène du sang. Similairement, Habibzadeh et al. [42] proposent un modèle de type SVM (AUC = 0.93) pour prédire la nécessité de chirurgie chez les patients avec TCC modéré.

D'autres travaux s'intéressent à déterminer le meilleur moment pour réaliser une chirurgie telle que la craniectomie décompressive (DC). Hanko et al. [43] ont entraîné un modèle de type Random Forest dans le but de prédire la morbidité six mois après un TCC. Ces travaux montrent que la DC réalisée tôt dans la prise en charge permet de réduire l'hypertension intracrânienne ainsi que les hémorragies cérébrales et a une incidence positive sur le pronostic du patient.

Même avec un TCC léger ou modéré, l'état du patient peut rapidement se détériorer. Il est donc intéressant de pouvoir prédire le pronostic des patients. Plusieurs modèles ont été développés dans le but de prédire les probabilités de mortalité des patients ou même la gravité des séquelles à

long terme après traitement [44,45]. Ces outils d'aide à la décision clinique permettent de mieux orienter le traitement des patients et d'une certaine façon de personnaliser la prise en charge.

Dans un contexte déjà surchargé aux SI, le déclenchement de fausses alarmes durant les périodes les plus critiques de la prise en charge de patient peut augmenter la charge de travail du personnel et diminuer leur attention envers ceux-ci [46]. Elles sont souvent causées par des artéfacts de mouvement sur les sondes de mesure au chevet du patient. Les techniques d'éliminations d'artéfacts basées sur le filtrage de signaux neurophysiologiques se sont montrées insuffisantes indiquant la nécessité de modèles plus complexes pour cette tâche. Différents modèles de type Random Forest et SVM [46,47] ont démontré avec succès qu'il est possible de reconnaître des artéfacts sur les signaux de pression artérielle et d'électrocardiographie réduisant ainsi le taux de fausses alarmes au SI.

Comme discuté précédemment, il est primordial de monitorer les signaux physiologiques en continue afin de détecter des transitions vers des états pathophysiologiques le plus rapidement possible. En ce sens, plusieurs travaux s'intéressent à utiliser les signaux physiologiques en temps réel de patients avec TCC afin de prédire des événements dangereux nécessitant une réaction rapide du personnel soignant. Par exemple, Schweingruber et al. [48] ont entraîné un RNA profond de type LSTM pour la prédiction des périodes d'HTIC. En se basant sur des données démographiques ainsi que les enregistrements en continue de la PAM et de la PIC, ce modèle a pu prédire avec une très grande précision des périodes d'HTIC de durées différentes deux heures avant qu'elles ne se manifestent.

Les techniques d'apprentissage automatique sont aussi mises à profit pour la détection et la quantification d'hématomes et d'œdèmes cérébraux en utilisant des réseaux de type CNN dans l'analyse d'images tomodensitométriques [49,50].

2.5 Résumé de la revue de littérature

La revue de littérature présentée ci-dessus a mis en évidence la complexité de la prise en charge du TCCS ainsi que le manque d'outils automatiques permettant l'évaluation en temps réel de l'état physiopathologique d'un patient atteint de TCCS. En effet, la majorité des études existantes s'intéressent au pronostic du patient (taux de survie), à la prédiction de l'évolution du signal physiologique (ex: pression intracrânienne) et au développement d'outils pour l'établissement d'un diagnostic sur l'état initial du patient (ex: analyse d'images tomodensitométriques). Au regard de l'état de l'art, ce projet vise à explorer la capacité de systèmes d'apprentissage automatique pour l'identification des états physiopathologiques en temps réel. Des approches similaires ont déjà été explorées dans le contexte des soins intensifs [51], mais non spécifiquement au TCCS.

CHAPITRE 3 MÉTHODOLOGIE

À la lumière de la revue de littérature au chapitre précédent, nous présentons ici la méthodologie adoptée en vue du développement et l'évaluation de deux types de RNA à apprentissage profond pour la classification supervisée de l'état pathophysiologique crânio-cérébral de patients avec TCCS. Leurs performances diagnostiques respectives seront comparées à celles de classificateurs classiques.

3.1 Objectifs

L'objectif principal de ce projet est d'évaluer la capacité de modèles d'apprentissage automatique de diagnostiquer l'état physiopathologique crânio-cérébral, à partir de données physiologiques acquises en continu, de patients ayant subi un TCCS. Plus précisément, différents algorithmes d'apprentissage automatique, conventionnels et de type réseaux de neurones profonds, seront évalués sur une base de données adultes et leurs performances diagnostiques seront comparées.

3.2 Données

Étant donné qu'une base de données de patients pédiatriques n'est actuellement pas disponible au Centre Hospitalier Universitaire Sainte-Justine, des données physiologiques annotées *a posteriori* par des experts en TCC de 49 patients adultes obtenus de l'Hôpital du Sacré Cœur de Montréal sont utilisées ici. À noter que la durée du séjour aux SI de chaque patient est différente. Ces données servent à entraîner et valider des réseaux de type FCN et CNN-LSTM et leurs performances diagnostiques sont comparées avec celles des modèles de classification classique tels que le SVM et le RF.

La base de données à disposition totalise 4000 heures d'enregistrements continus de six signaux physiologiques d'intérêts enregistrés à un intervalle de 30 secondes. Ces signaux sont : la PIC, la PAM, la PPC, le PbtO₂, Le PRx et le PPCO. Un prétraitement est appliqué afin d'éliminer les valeurs aberrantes pour chaque signal physiologique. Les valeurs manquantes sur une période de quelques minutes ont été remplacées par la moyenne mobile. Seulement les enregistrements pour lesquels toutes les valeurs sont disponibles sont retenus lors de la classification.

Il est à noter que puisque le calcul du PPCO requiert un certain temps et pour maximiser les observations retenues, une valeur de PCCO de 60 mm de mercure a été ajoutée en accord avec les recommandations cliniques pour adultes qui suggèrent cette cible au début de la prise en charge.

En dehors du diagnostic de TCCS, les informations au sujet des patients tels que le sexe, l'âge, et le type et la sévérité des lésions n'étaient pas disponibles dans le cadre de ce travail.

Les signaux ont été analysés par deux experts du TCCS et les périodes correspondant aux dix étiquettes suivantes ont été identifiés :

1. État contrôlé
2. HTIC avec ischémie
3. HTIC avec hyperhémie
4. HTIC
5. Ischémie
6. Hyperhémie
7. Doute sur l'état pathophysiologique
8. Trop d'artéfact
9. Trop de données manquantes
10. Erreur de classification

Chaque point d'enregistrement est associé à une étiquette. Dans le présent travail, seules les étiquettes correspondant aux états pathophysiologiques sont conservées (étiquettes 1 à 6). Une comparaison entre les annotations des experts n'est pas disponible puisque chaque patient a été analysé par un seul des deux experts.

3.2.1 Distribution par état pathophysiologique

La Figure 3.1 présente la proportion des données associées à chaque état pathophysiologique. Les états présentant des anomalies critiques sont nettement sous-représentés (11%) alors que l'état contrôle compose 89% des données disponibles. C'est une sur-représentation de l'état contrôle qui doit être prise en considération pour le développement des modèles de classification supervisés.

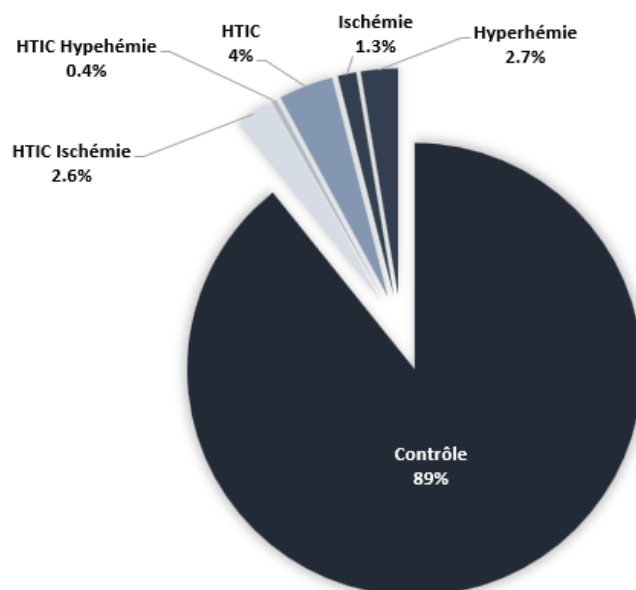


Figure 3.1 Distribution des données associée à chaque état pathophysiologique

3.2.2 Distribution par patient

La Figure 3.2 montre le nombre de patients associé à chaque état. La majorité des patients présentent des périodes de contrôle de l'état crânio-cérébrale tandis qu'un peu plus de la moitié ont subi des périodes d'HTIC. Les quatre autres états sont présents chez moins de la moitié des patients.

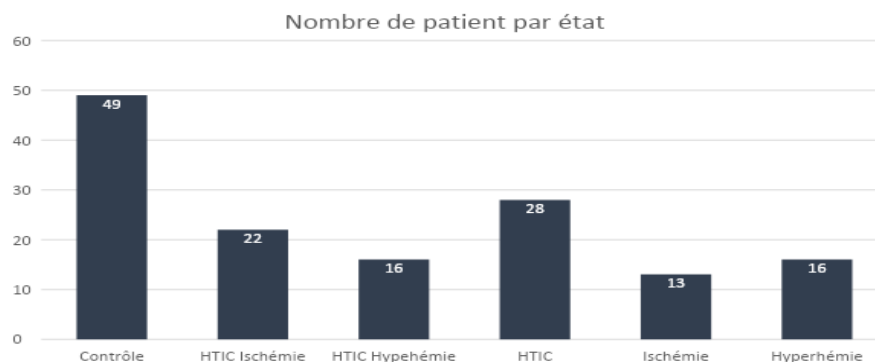


Figure 3.2 Nombre de patient associé à chaque état

3.3 Mise en forme des données

3.3.1 Séquences

Les signaux de la base de données sont mis en forme de façon qu'ils soient interprétables par les différents algorithmes de classification envisagés. Afin de tenir compte de l'aspect temporel des signaux physiologiques, ils sont séparés en sous-séquences. Comme le montre la figure 3.3, une fenêtre coulissante passe à travers les données physiologiques étiquetées et forme les séquences de longueur fixe. La fenêtre coulissante se déplace d'une itération à la fois. Ceci reflète bien la situation clinique où chaque nouvelle observation est disponible toutes les 30 secondes.

	T_1	T_2	T_3		T_L	T_{L+1}	T_{L+2}	T_{L+3}
PIC	PIC ₁	PIC ₂	PIC ₃		PIC _L	PIC _{L+1}	PIC _{L+2}	PIC _{L+3}
PPC	PPC ₁	PPC ₂	PPC ₃		PPC _L	PPC _{L+1}	PPC _{L+2}	PPC _{L+3}
PAM	PAM ₁	PAM ₁	PAM ₁		PAM ₁	PAM ₁	PAM ₁	PAM ₁
Licox	Licox ₁	Licox ₁	Licox ₁		Licox ₁	Licox ₁	Licox ₁	Licox ₁
PRX	PRX ₁	PRX ₁	PRX ₁		PRX ₁	PRX ₁	PRX ₁	PRX ₁
PPCO	PPCO ₁	PPCO ₁	PPCO ₁		PPCO ₁	PPCO ₁	PPCO ₁	PPCO ₁
	Séquence 1		Séquence 2		Séquence 3			
Code Diagnostique	1	1	1		1	2	2	2

Figure 3.3 Mise en forme des séquences temporelles

Des longueurs de séquence équivalentes à 5, 10, 15, 30 et 60 minutes d'enregistrement ont été considérées pour finalement retenir la durée de 30 minutes (60 observations par séquence) puisqu'elle présente globalement les meilleures performances de classification obtenues.

L'association de chaque séquence à un état pathophysiologique se fait en considérant le nombre d'étiquettes appartenant à chaque état dans celle-ci. Si toutes les observations de la séquence sont étiquetées à l'état contrôlé, alors la séquence est associée à cet état. Sinon, la séquence est associée à l'état pathophysiologique dont les observations sont majoritaires. Cette façon de labelliser les séquences permet de maximiser le nombre de séquences appartenant aux classes pathologiques, au risque de labelliser incorrectement certaines séquences.

Parmi les 49 patients disponibles, deux d'entre eux n'ont pas été retenus à cette étape car l'enregistrement de leur signal PbtO₂ était manquant durant tout le séjour aux SI.

Les séquences obtenues sont directement utilisées par les algorithmes d'apprentissage profond, mais sont transformées pour utilisation par les classificateurs de type SVM et RF, afin de faciliter leur apprentissage et la performance de ceux-ci.

3.3.2 Extraction des caractéristiques

Pour les classificateurs de type SVM et RF, des caractéristiques représentant chaque séquence temporelle sont extraites afin de faciliter l'apprentissage des modèles. Pour chacun des six signaux, différentes caractéristiques sont déduites : la moyenne, la médiane, l'écart-type, la variance, la valeur minimale, la valeur maximale, le coefficient d'acuité et le coefficient d'asymétrie. Le temps de déviation par rapport à la normale et la gravité de la déviation sont calculés pour les signaux PIC, PPC, PPCO, PbtO₂ et PRx seulement. Au total, pour chacune des séquences, 62 caractéristiques sont calculées.

- **Moyenne**

La moyenne est la somme de toutes les valeurs dans la séquence divisées par le nombre d'éléments (la longueur de la séquence).

- **Médiane**

La médiane M correspond à la valeur centrale de la séquence de façon que la moitié des valeurs de la suite sont inférieures ou égale à M et que 50% y sont supérieures ou égales.

- **Variance**

La variance est une mesure de la dispersion des valeurs dans la séquence. Elle correspond à la moyenne du carré des écarts à la moyenne.

- **L'écart type**

L'écart type est une mesure de dispersion des valeurs dans la séquence. Elle représente l'ampleur de la variation des éléments par rapport à la moyenne. C'est aussi la racine carrée de la variance.

- **Valeur minimale**

C'est la valeur minimale dans une séquence.

- **Valeur maximale**

C'est la valeur maximale dans la séquence.

- **Coefficient d'acuité (kurtosis)**

Le coefficient d'acuité représente une mesure de l'aplatissement des valeurs dans la séquence.

- **Coefficient d'asymétrie**

Le coefficient d'asymétrie représente une mesure de l'asymétrie des valeurs dans la séquence.

- **Mesure du temps de déviation par rapport à la normale**

Cette mesure donne une indication du nombre d'observations qui sont à l'extérieur de la plage de normalité sur la séquence. Sa valeur est donc comprise entre zéro et la longueur de la séquence ([0,60]).

- **Mesure de la gravité de la déviation**

Cette mesure indique l'amplitude de la déviation des observations dans la séquence par rapport à la normalité. Pour chaque observation de la séquence, la déviation est calculée. Puis la somme de toutes les déviations sur la séquence est obtenue.

Les valeurs utilisées comme limites de la normalité pour la PIC, la PPC, le PbTO₂, la PRx et la PPCO proviennent des règles de décision établies par des experts en la matière (Annexe A). La valeur limite de la PIC est de 20 millimètres de mercure. La PbTO₂ a une limite inférieure de 25 et une limite supérieure de 35 millimètres de mercure. La limite pour le PRx est de 0. La valeur limite pour la PPC est de 50 millimètres si la PPCO n'est pas tenue en compte. Sinon elle a une limite inférieure et supérieure de ± 10 la valeur de la PPCO.

3.3.3 Analyse en composantes principales

L'analyse en composantes principales [52] est une technique qui consiste à projeter les observations d'un espace à n dimensions (variables) vers un espace à m dimensions de sorte que m soit plus petit que n tout en conservant le maximum d'information de l'ensemble de départ. Cette quantité d'information est quantifiée par la variance totale de l'ensemble de départ. Concrètement, cette méthode est utilisée afin de réduire la complexité du problème ainsi que le temps d'entraînement des modèles SVM et RF. Une réduction de dimension est effectuée avec la PCA sur l'ensemble de toutes les caractéristiques extraites. Une variabilité retenue de 99%, 95%, 80% et 70% ont été considérées. Finalement, la réduction de dimension choisie explique 99% de variabilité des données. Elle permet de réduire la dimension de la base d'entraînement de 45% tout en minimisant les pertes de performances des classificateurs SVM et RF.

3.4 Sous échantillonnage de l'état contrôlé

Entraîner un classificateur à l'aide d'un ensemble de données hautement déséquilibré comme celui disponible pour ce travail peut négativement influencer les performances. Les classificateurs auront tendance à être surentraînés sur les classes surreprésentées. Cependant, dans le cas qui nous intéresse, il est beaucoup plus critique de détecter les anomalies de l'état crânio-cérébrale que l'état contrôlé.

Une technique permettant de balancer la distribution des classes dans une base de données est le sous-échantillonnage de la classe majoritaire. L'algorithme NearMiss [53] est utilisé pour réduire le nombre de données associées à la classe de l'état contrôlé. Parmi les trois versions de cet algorithme, NearMiss-3 est considéré comme étant la plus robuste aux données bruitées et aux observations aberrantes.

Cet algorithme analyse les données en deux étapes. Dans un premier temps, pour chaque observation de la classe sous-représentée, M plus proches voisins de la classe sur-représentée sont retenues. Parmi ces M proches voisins, seuls ceux maximisant une distance moyenne avec N observations de la classe sous représentée sont conservés. Dans ce travail, à la suite des tests de différentes valeurs pour M et N , les valeurs par défaut de l'algorithme ont montré de meilleurs résultats ($M = N = 3$).

L'algorithme NearMiss-3 est donc utilisé pour sous-échantillonner les séquences appartenant à la classe majoritaire de la base de données; l'état contrôlé. À noter que l'implémentation de cet algorithme ne supporte pas directement l'analyse de séquences temporelles multivariées. C'est donc avec les caractéristiques statistiques associées à chaque séquence que cet algorithme a été utilisé. Les indices des observations retenues sont ensuite utilisés pour ne garder que les séquences pertinentes lors de l'entraînement des différents classificateurs étudiés.

3.5 Stratégie d'entraînement

Les 47 patients retenus sont séparés en trois groupes distincts : entraînement, validation et test. Chaque groupe de patients est formé de sorte qu'il contienne des observations relatives à chaque état pathophysiologique. Puisqu'il est primordial qu'à chaque étape les performances soient évaluées sur de nouvelles observations non connues du modèle, chaque patient ne se retrouve que dans un seul groupe.

Les groupes d'entraînement et de test sont formés de 37 et 10 patients respectivement. Parmi les 37 patients d'entraînement, 5 font partie du groupe de validation. De cette façon, environ 80% des patients sont utilisés dans la phase d'entraînement, 10% pour la validation des hyperparamètres et 20 % pour la phase de test.

Tel que montré à la Figure 3.4, des modèles avec différents hyperparamètres sont entraînés. Le modèle montrant les meilleures performances de classification sur le groupe de validation est ensuite sélectionné pour l'évaluation finale sur le groupe de test.

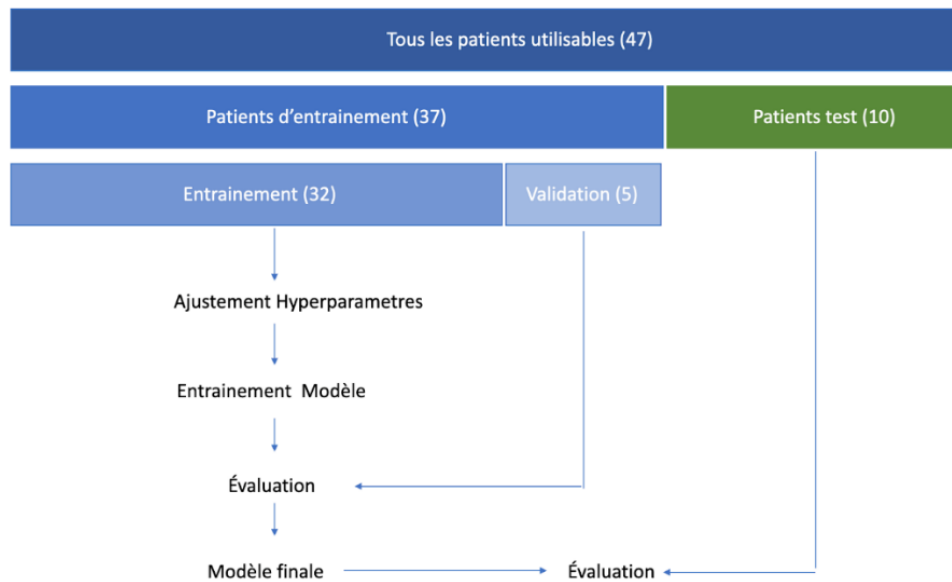


Figure 3.4 Étapes de développement des modèles de classificateurs

3.6 Réseaux de neurones profonds

Différentes architectures de complexités variables ont été considérées dans ce travail de recherche. Compte tenu des performances obtenues, les deux RNA profonds retenus sont le FCN et un réseau combinant un CNN avec un LSTM. Comme déjà discuté au chapitre 2, Ils sont tous deux adaptés au problème de classification de suites temporelles.

3.6.1 Le FCN

Le réseau FCN a été retenu dans cette étude puisqu'il a démontré que son architecture moins complexe comparé à plusieurs autres RNA profonds est en mesure de produire de bonnes performances de classification tel que validée sur différents types de base de données destinées à la classification de suites temporelles.

L'architecture du réseau FCN étudié ici comprend une couche d'entrée acceptant des séquences de dimension (6,60), soit une séquence de 60 observations des six signaux physiologiques. Trois couches de convolution à une dimension suivi d'une normalisation par batch et d'une fonction d'activation redresseur (ReLU) sont suivies d'une couche de type Global Average Pooling. Cette dernière couche est connectée à une couche dense qui permet la classification finale. La sorte

correspond à la probabilité d'appartenance à chaque classe. Cette architecture est similaire à l'architecture proposée par Wang et al. [30]. Ce réseau est montré à la Figure 3.5.

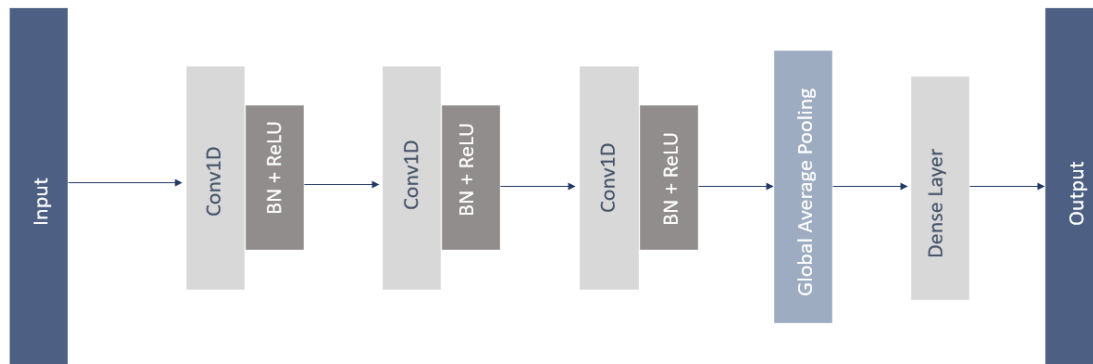


Figure 3.5 Architecture du réseau FCN retenue

3.6.2 Le CNN-LSTM

Le réseau CNN-LSTM est formé d'un premier réseau CNN permettant d'extraire les caractéristiques des suites temporelles présentées à son entrée automatiquement. Puis en y connectant un réseau LSTM, on effectue la classification tout en tenant compte des observations précédentes.

L'architecture du réseau CNN-LSTM étudié ici comprend une couche d'entrée acceptant des séquences de dimension (6,60), soit une séquence de 60 observations des six signaux physiologiques. Trois couches de convolution à une dimension avec une fonction d'activation redresseur (ReLU) et une couche pour la normalisation. Puis une couche dropout ayant pour but de diminuer le sur-apprentissage est devant la couche de pooling de type maximale. On trouve ensuite le réseau LSTM, une couche dropout et enfin une couche dense pour la classification finale. Ce réseau est montré à la Figure 3.6

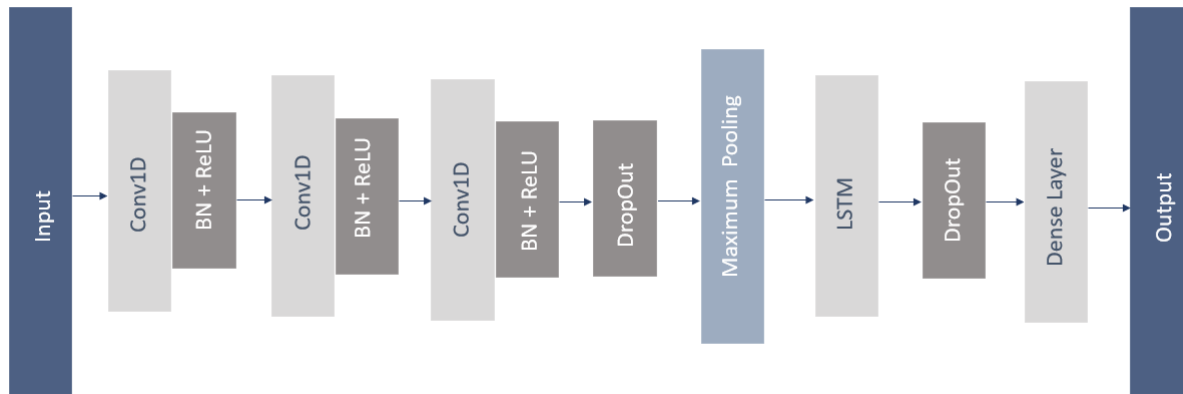


Figure 3.6 Architecture du réseau CNN-LSTM retenue

3.6.3 Paramètres d'entraînement

Dans les deux cas, la fonction de coût utilisée alors de l'entraînement est la Categorical Cross Entropy. L'optimisation Adam, une extension de la descente de gradient a été choisi. Ce sont deux choix globalement validés dans la littérature [54,55]. Afin de limiter le sur-apprentissage, un critère d'arrêt basé sur la fonction de coût des données de validation est spécifié avec une patience de 50 itérations.

3.6.4 Choix des hyperparamètres

Différent hyperparamètres ont été considérés lors de l'entraînement des deux RNA profonds: le nombre de filtre par couche de convolution et la taille des filtres. L'utilisation de couche de normalisation et de dropout ont aussi été évalué pour le réseau CNN-LSTM. Le nombre de lot d'entraînement (Batch size en anglais) a aussi été testé. Le nombre d'unités du LSTM a également été testé.

Finalement, le nombre de filtres par couche de convolution choisie est de 64 avec une taille de 3. La normalisation par batch est utilisée. Dans le cas plus spécifique du CNN-LSTM, le taux de dropout est fixé à 0.50 et le nombre d'unités LSTM est fixé à 200. Dans les deux cas le batch size choisi est de 128, qui est un bon compromis entre le temps d'entrainement et les performances des réseaux.

3.7 Classificateurs classiques

Deux types de classificateurs qui ne sont pas basés sur des réseaux de neurones sont aussi considérés à des fins de comparaison avec les RNA profond. La littérature démontre que ces deux types de classificateur ont démontré de bons résultats dans divers problèmes de classification de suite temporelle.

3.7.1 Le SVM

Le classificateur SVM utilisé dans ce travail a un noyau RBF. C'est un noyau polyvalent puisque qu'il a des similitudes avec la distribution gaussienne. Il a démontré de meilleures performances que le noyau linéaire aussi testé. Tous les autres paramètres utilisés ici sont gardés à leur valeur par défaut.

3.7.2 Le Random Forest

Pour le classificateur de type RF, des tests ont montrés que les performances de classification ne s'améliorent plus pour une valeur au-dessus de 1000 pour la profondeur maximale des arbres du classificateur. C'est cette valeur qui à été retenue. Tous les autres paramètres utilisés ici sont gardés à leur valeur par défaut.

3.8 Types d'expériences

Deux types de classification ont été étudiés : la classification en classe multiple et la classification binaire.

3.8.1 Classification à six états

Pour ce type de classification, chaque état pathophysiologique correspond à une classe. Le but est de distinguer chacun des cinq états critiques (HTIC, HTIC-I, HTIC-H, ISC et HYP) ainsi que l'état normal (CTL). Il s'agit d'un problème de classification à 6 classes qui ressemble au processus décisionnel du personnel soignant.

3.8.2 Classification binaire

Ce type de classification ne considère que deux classes dans la classification. Cette façon de faire permet de simplifier le problème de classification initiale en mettant au point des classificateurs spécialisés dans la détection d'un seul élément pathologique (ex: HTIC, ISC ou HYP) à la fois.

Deux stratégies ont été étudiées pour la classification binaire : le un versus le reste et le un versus un.

3.8.2.1 Un versus le reste

Cette façon de faire permet de détecter un élément pathologique (HTIC, ISC ou HYP) parmi tous les états. Par exemple pour la détection de l'ischémie, une classe comprend les états pathophysiologiques ISC et HTIC-I. La deuxième comprend les états CTRL, HTIC, HTIC-H et HYP. Pour la détection de l'hyperhémie, la première classe comprend les états HYP et HTIC-HYP. La deuxième comprend les états CTRL, HTIC, HTIC-I et ISC. Similairement, pour la détection de l'HTIC, la première classe comprend les états HTIC, HTIC-I et HTIC-H tandis que l'autre se compose des états CTRL, ISC et HYP.

3.8.2.1 Un versus un

Cette façon de faire permet de comparer simplement un état pathologique contre l'état contrôlé. Par exemple, la détection de l'ischémie se fait en considérant une classe comprenant les états ISC et HTIC-I contre une classe CTRL seulement. Pareillement, la détection de l'HYP devient une classe comprenant les états HYP et HTIC-H contre une classe CTRL. Similairement, la détection de l'HTIC se fait en comparant une classe comprenant HTIC, HTIC-I et HTIC-H contre la classe CTRL.

3.9 Métriques de performance

La matrice de confusion est une matrice qui permet d'évaluer visuellement les performances de classification d'un modèle de classificateur supervisée. Comme le montre la figure 3.7 les lignes de cette matrice correspondent aux instances de la classe réelle alors que chaque colonne correspond aux instances de la classe estimée par le modèle étudié.

		État Prédit	
		Positif	Négatif
Vrai État	Positif	Vrai Positif	Faux Négatif
	Négatif	Faux Positif	Vrai Négatif

Figure 3.7 Représentation d'une matrice de confusion

Il est possible de déduire des métriques de performances à partir de la matrice de confusion. Les métriques considérées dans ce travail sont : la précision, la sensibilité, le score F1 et l'exactitude. Ce sont des métriques globalement utilisées dans la littérature pour l'évaluation de modèles de classificateurs supervisés [56]. Dans ce travail, un score supérieur à 0.80 est considéré satisfaisant.

Il importe de connaître la signification de chaque case de la matrice de confusion :

- **Vrai Positif (VP)**

Une observation dans cette case signifie que l'état est correctement prédit comme positif.

- **Vrai Négatif (VN)**

Une observation dans cette case signifie que l'état est correctement prédit comme négatif.

- **Faux Positif (FP)**

Une observation dans cette case signifie que l'état est faussement prédit comme positif par le modèle. Elle appartient réellement à la classe négative.

- **Faux Négatif (FN)**

Une observation dans cette case signifie que l'état est faussement prédit comme négatif par le modèle. Elle appartient réellement à la classe positive.

Il est ensuite possible déduire les différentes métriques de la façon suivante:

- **Précision**

$$Précision = \frac{VP}{VP + FP} \quad 3.1$$

- **Exactitude (Accuracy)**

$$Exactitude = \frac{VP + VN}{VP + FP + VN + FN} \quad 3.2$$

- **Sensibilité**

$$Sensibilité = \frac{VP}{VP + FN} \quad 3.3$$

- **Score F1**

$$Score F1 = \frac{2VP}{2VP + FP + FN} \quad 3.4$$

3.10 Langage de programmation et librairies

Le programme développé dans le cadre de ce travail a été écrit sous le langage de programmation Python. En plus d'être relativement facile d'utilisation et portable sur différentes plateformes informatiques, il offre plusieurs librairies très bien adaptées à la manipulation et l'analyse de grande quantité de données, au traitement de signal, au développement d'applications d'apprentissage machine et à la science des données en général. Les principales librairies utilisées ici sont les suivantes :

Keras est une librairie destinée à l'apprentissage machine qui permet de facilement développer et tester des RNA de façon intuitive et modulaire. Il sert d'interface à la plateforme spécialisée en intelligence artificielle TensorFlow de Google. Lorsque Installée avec les pilotes appropriés, elle permet de mettre à profit la grande capacité de calcul des processeurs d'unité graphique (GPU) et ainsi d'accélérer les temps de calculs nécessaires au développement de modèles. Cette librairie est utilisée pour le développement des deux RNA à apprentissage profond FCN et CNN-LSTM.

Scikit-Learn est une autre librairie spécialisée pour l'apprentissage machine qui permet l'utilisation rapide de différents modèles classiques de classificateurs supervisés et non supervisés ainsi que des fonctionnalités permettant l'évaluation de leur performance. Cette librairie est utilisée ici pour les classificateurs SVM et RF.

Panda est une librairie facilitant l'analyse et la manipulation de données qui est très bien adaptée aux données de type séries temporelles. Ces données sont stockées sous forme de tableau à deux dimensions où les colonnes représentent chacune une variable d'intérêt alors que chaque ligne contient les valeurs enregistrées à un intervalle de temps donnée.

Imbalance-learn est une librairie offrant différentes méthodes afin de balancer la quantité de données associée aux différentes classes dans un ensemble de données destinées à la classification multi-classes. Cette librairie est utilisée ici afin de réduire la sur représentation de l'état contrôlé dans l'ensemble de données à l'étude.

SciPy est une librairie utilisée pour les manipulation mathématiques et statistiques. Elle est utilisée dans l'extraction des caractéristiques statistiques des séquences temporelles pour les classificateurs SVM et RF.

3.11 Structure du programme

Le programme développé peut être divisé en deux parties distinctes. Dans un premier temps, un module de traitement de données lit les données physiologiques des patients et les annotations médicales associées afin de les mettre en forme pour qu'elles soient interprétables par les différents algorithmes de classificateurs considérés. On appellera l'ensemble de ces données formatées la base de données de travail. Le deuxième module consiste simplement à la création des différents modèles de classificateurs, leur entraînement, validation et test en utilisant la base de données de travail ainsi que l'évaluation de leurs performances diagnostiques.

3.11.1 Traitement des données

Pour chacun des patients avec TCCS, deux fichiers sont disponibles. Le premier contient l'enregistrement de ses différents signaux physiologiques sous forme de séries temporelles échantillonnées au 30 secondes. Les signaux disponibles sont : la PIC, la PPC, la PAM, le PbtO₂, le PRX et la PPCO. La date et l'heure de chaque enregistrement est aussi disponible. Le deuxième fichier associé à chaque patient contient le diagnostic médical posé par deux experts en TCCS ayant analysé les données physiologiques de façon rétrospective. La date et l'heure du début et de la fin de chaque événement médical est identifiée et associée à un des états pathophysiologiques.

Dans un premier temps et pour chaque patient, les annotations médicales sont lues puis chaque ligne d'enregistrement des données physiologiques est associée à un code diagnostique. Ces données sont ensuite séparées en séquence de 30 minutes. Ensuite, chaque séquence est associée à un état pathophysiologique et les différentes caractéristiques statistiques sont extraites. Les séquences de tous les patients et leurs différentes caractéristiques associées forment la base de données de travail. C'est à partir de cette base que les différents classificateurs seront entraînés et testés. La Figure 3.8 résume le processus de traitement de données.

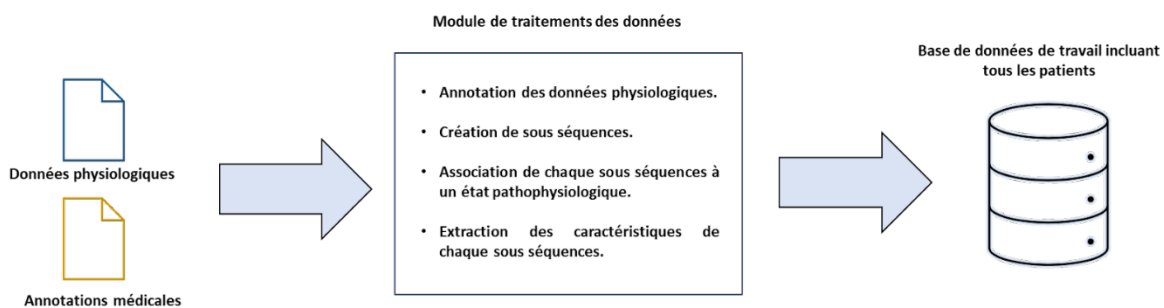


Figure 3.8 Processus de traitement des données

3.11.2 Classification et performances

Le module d'apprentissage machine partitionne la base de travail en trois parties distinctes de façon que les données de chaque patient ne se retrouvent que dans l'une de celles-ci pour l'entraînement, le développement et l'évaluation des classificateurs créés. Les performances d'entraînement et de diagnostic de chaque classificateur sont générées permettant une comparaison entre les différentes méthodes de classification. La Figure 3.9 résume le processus du module.

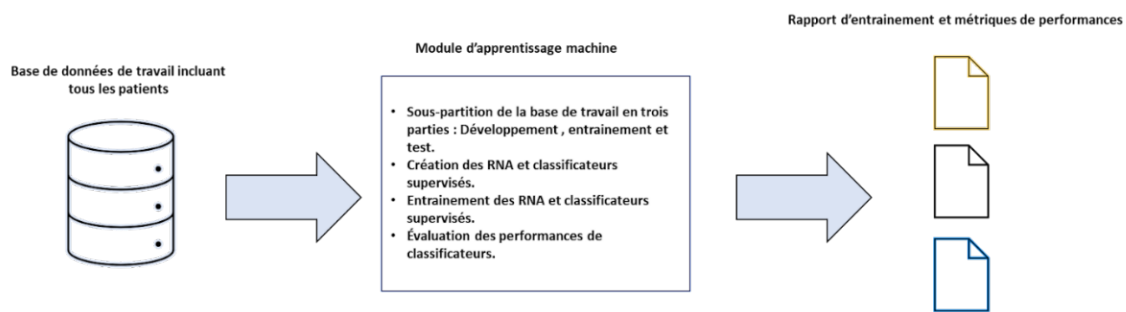


Figure 3.9 Processus de la classification et l'évaluation des performances

CHAPITRE 4 RÉSULTATS

Ce chapitre présente les résultats des différentes expériences décrites dans la méthodologie à la section 3. Les résultats sont assemblés par type d'expérience, en commençant par la classification de 6 états distincts : contrôle (CTRL), hypertension intracrânienne avec ischémie (HTIC-I), hypertension intracrânienne avec hyperhémie (HTIC-H), hypertension intracrânienne (HTIC), ischémie (ISC), hyperhémie (HYP), suivi par la classification des états en un nombre plus restreints de catégories.

4.1 Classification en six états

Deux types de réseaux de neurones ont été évalués pour classifier les six états physiopathologiques des patients : un réseau FCN et un réseau CNN-LSTM. Ces deux réseaux ont été choisis par leur capacité à analyser des séries temporelles.

La Figure 4.1 et le Tableau 4.1 montrent les résultats de classification du réseau FCN retenu sur les patients du groupe de validation. Des résultats similaires ont été obtenus pour le réseau de type CNN-LSTM et sont présentés en annexe B. Les courbes d'apprentissage et de validation (fonction de coût et précision) des deux réseaux sont également présentées en annexe B et démontrent bien que dans les deux cas, les réseaux sont raisonnablement bien entraînés. Les performances de classification indiquent que les états avec HTIC sont les mieux détectés parmi tous les états physiopathologiques.

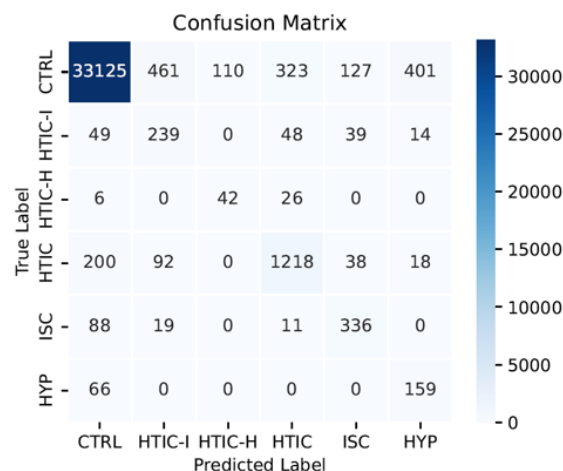


Figure 4.1 Matrice de confusion du FCN sur le groupe de développement

Tableau 4.1 Métriques du FCN sur le groupe de développement

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.99	0.96	0.97	34547
HTIC-I	0.29	0.61	0.40	389
HTIC-H	0.28	0.57	0.37	74
HTIC	0.75	0.78	0.76	1566
ISC	0.62	0.74	0.68	454
HYP	0.27	0.71	0.39	225
accuracy			0.94	37255
macro average	0.53	0.73	0.6	
weighted average	0.96	0.94	0.95	

La Figure 4.2 présente les résultats du réseau FCN. Tel qu'appuyé par le Tableau 4.2, les résultats montrent que le réseau est en mesure de classifier l'état contrôle avec une grande précision. Le réseau est également en mesure de bien classifier les états d'hypertension intracrânienne (HTIC) mais n'est pas capable de classifier correctement les autres états.

La Figure 4.3 et le Tableau 4.3 présentent les résultats de la classification des états par le réseau CNN-LSTM. À l'instar du réseau FCN, les résultats présentés ici montrent que ce réseau est aussi en mesure de classifier l'état contrôle avec une grande précision. Il classifie également les états d'hypertension intracrânienne (HTIC) de façon acceptable, mais n'est pas capable de reconnaître correctement les quatre autres états.

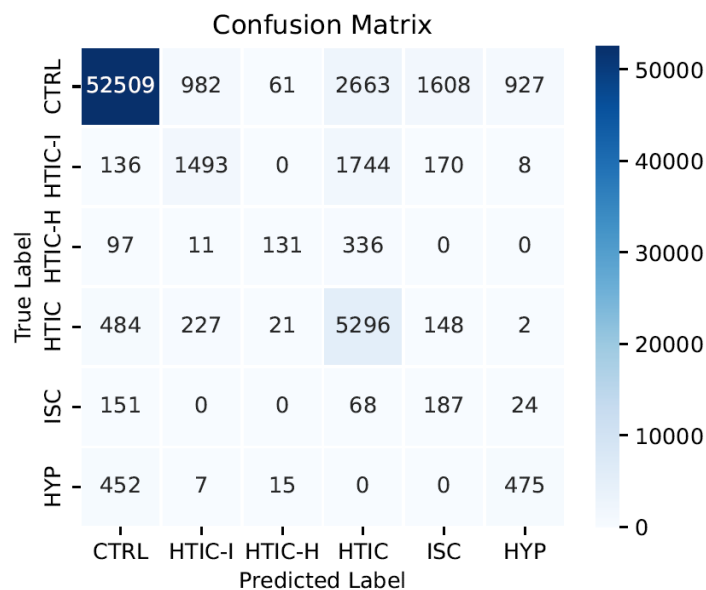


Figure 4.2 Matrice de confusion du réseau FCN pour la classification à six états

Tableau 4.2 Métriques de la classification du réseau FCN pour six états

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.98	0.89	0.93	58750
HTIC-I	0.55	0.42	0.48	3551
HTIC-H	0.57	0.23	0.33	575
HTIC	0.52	0.86	0.65	6178
ISC	0.09	0.43	0.15	430
HYP	0.33	0.50	0.40	949
accuracy			0.85	70433
macro average	0.51	0.56	0.49	
weighted average	0.9	0.85	0.87	

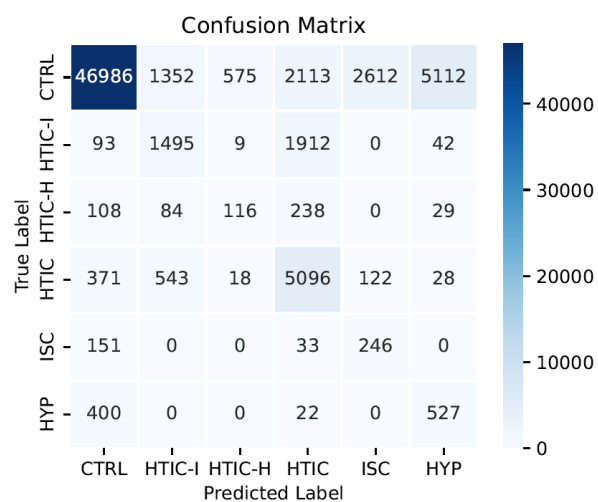


Figure 4.3 Matrice de confusion du CNN-LSTM pour six états

Tableau 4.3 Métriques de la classification du réseau CNN-LSTM pour six états

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.98	0.80	0.88	58750
HTIC-I	0.43	0.42	0.43	3551
HTIC-H	0.16	0.20	0.18	575
HTIC	0.54	0.82	0.65	6178
ISC	0.08	0.57	0.14	430
HYP	0.09	0.56	0.14	949
accuracy			0.77	70433
macro average	0.38	0.56	0.41	
weighted average	0.89	0.77	0.82	

4.2 Détection de l'hypertension intrâ-cranienne

Au vu des performances limitées de réseaux complexes à classifier un grand nombre d'états pathophysiologiques, des stratégies d'entraînement plus simples ont été explorées, en classifiant des regroupements d'états. Les Figures 4.4 et 4.5 et les Tableaux 4.4 et 4.5 présentent les résultats des réseaux FCN et CNN-LSTM pour la classification de l'état HTIC versus les autres états, identifiés ici comme non-HTIC (N-HTIC). Ces résultats montrent que les deux réseaux sont capables de distinguer avec une bonne performance les deux groupes d'états.

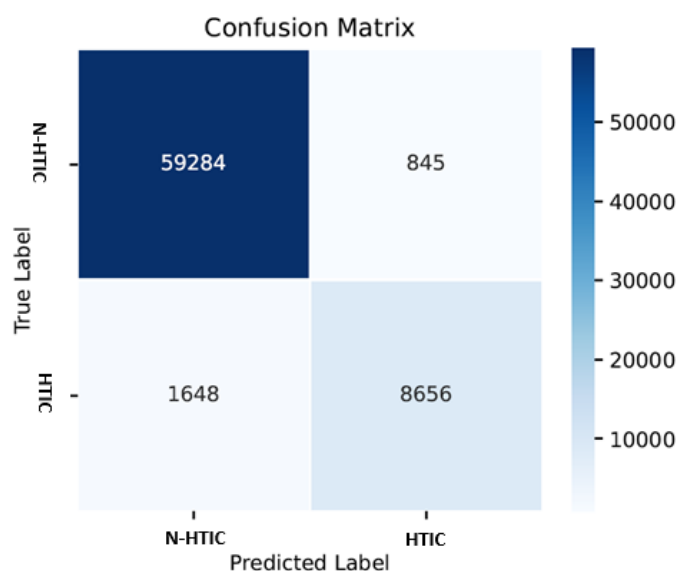


Figure 4.4 Matrice de confusion du FCN pour la détection de l'HTIC

Tableau 4.4 Métriques pour la détection de l'HTIC du FCN

États	Précision	Sensibilité	Score F1	Nombre
N-HTIC	0.97	0.99	0.98	60129
HTIC	0.91	0.84	0.87	10304
Accuracy			0.96	70433
Macro Average	0.94	0.91	0.93	
Weighted Average	0.96	0.96	0.96	

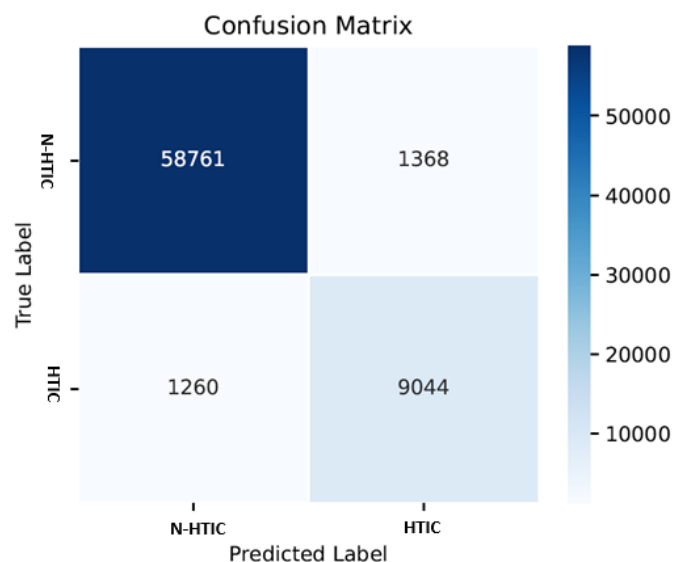


Figure 4.5 Matrice de confusion du CNN-LSTM pour la détection de l'HTIC

Tableau 4.5 Métriques pour la détection de l'HTIC du CNN-LSTM

États	Précision	Sensibilité	Score F1	Nombre
N-HTIC	0.98	0.98	0.98	60129
HTIC	0.87	0.88	0.87	10304
Accuracy			0.96	70433
Macro Average	0.92	0.93	0.93	
Weighted Average	0.96	0.96	0.96	

La Figure 4.6 et le Tableau 4.6 présentent les résultats d'un modèle plus simple pour cette classification de l'HTIC : un modèle de type SVM. Les résultats montrent que ce classificateur est capable d'atteindre des performances similaires, voire supérieures, à celles de classifieurs plus complexes. La Figure 4.7 et le Tableau 4.7 présentent les résultats d'un modèle RF pour cette même classification, montrant des performances sensiblement différentes. La précision de la détection de l'HTIC est moindre par rapport aux autres résultats, à cause d'un plus grand nombre

de faux positifs. Nous pouvons cependant remarquer que l'entraînement de ce réseau résulte en très peu de faux négatifs, laissant penser qu'il s'agit d'un enjeu de régularisation.

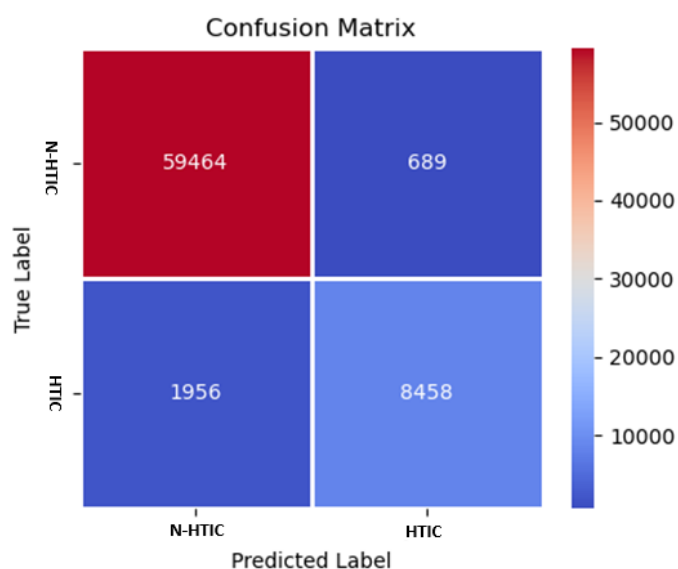


Figure 4.6 Matrice de confusion du SVM pour la détection de l'HTIC

Tableau 4.6 Métriques pour la détection de l'HTIC du SVM

États	Précision	Sensibilité	Score F1	Nombre
N-HTIC	0.97	0.99	0.98	60153
HTIC	0.92	0.81	0.86	10414
Accuracy			0.96	70567
Macro Average	0.95	0.90	0.92	
Weighted Average	0.96	0.96	0.96	

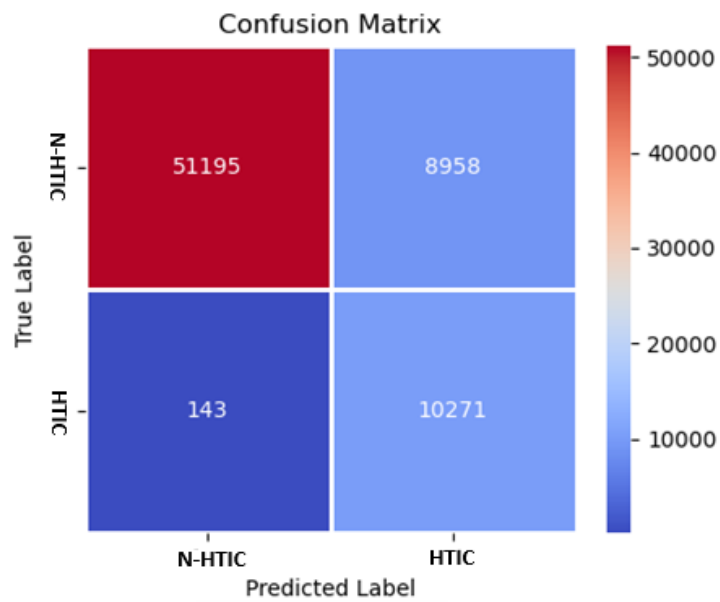


Figure 4.7 Matrice de confusion du RF pour la détection de l'HTIC

Tableau 4.7 Métriques pour la détection de l'HTIC du RF

États	Précision	Sensibilité	Score F1	Nombre
N-HTIC	1.00	0.85	0.92	60153
HTIC	0.53	0.99	0.69	10414
Accuracy			0.87	70567
Macro Average	0.77	0.92	0.81	
Weighted Average	0.93	0.97	0.89	

4.3 Détection de l'hyperhémie

Des stratégies d'entraînement simplifiées ont aussi été évaluées pour la détection de l'état HYP. La Figure 4.8 ainsi que le Tableau 4.8 présentent les résultats du réseau FCN pour la classification de l'état HYP versus tous les autres états, appelés ici N-HYP. Ces résultats montrent que cette stratégie ne permet pas d'améliorer les performances de détection de l'HYP sur le réseau FCN.

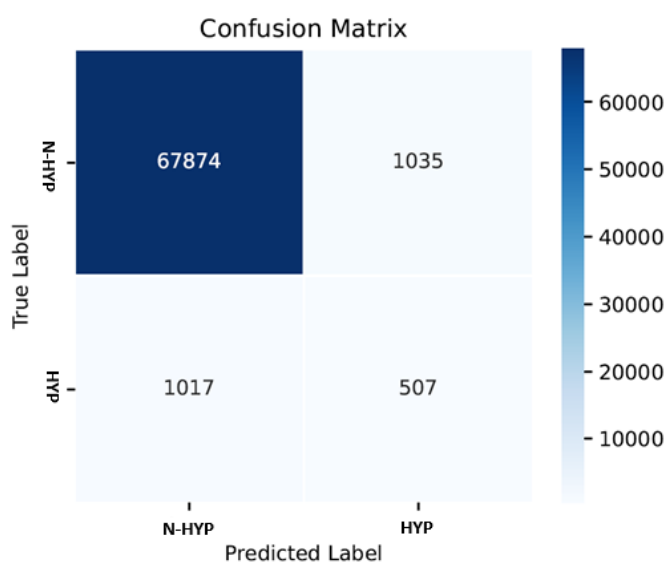


Figure 4.8 Matrice de confusion du réseau FCN pour la détection de l'HYP

Tableau 4.8 Métriques du réseau FCN pour la détection de l'HYP

États	Précision	Sensibilité	Score F1	Nombre
N-HYP	0.99	0.98	0.99	68909
HYP	0.33	0.33	0.33	1524
Accuracy			0.97	70433
Macro Average	0.66	0.66	0.66	
Weighted Average	0.97	0.97	0.97	

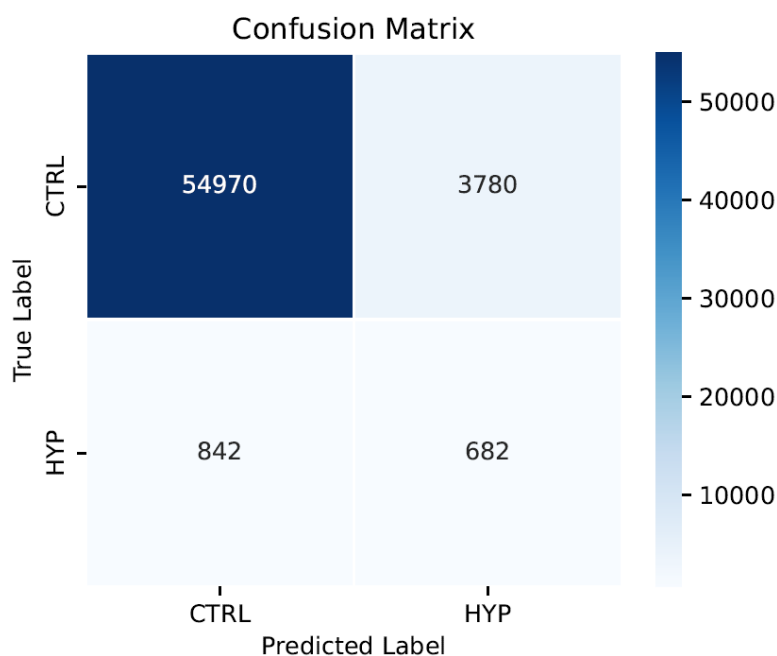


Figure 4.9 Matrice de confusion du réseau FCN pour la détection de l'HYP v CTRL

Tableau 4.9 Métriques du réseau FCN pour la détection de l'HYP v CTRL

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.98	0.94	0.96	58750
HYP	0.15	0.45	0.23	1524
Accuracy			0.92	60274
Macro Average	0.57	0.69	0.59	
Weighted Average	0.96	0.92	0.94	

La Figure 4.9 et le Tableau 4.9 présentent les résultats du même réseau FCN pour une classification entre l'état CTRL et l'HYP sans considérer l'ischémie ni l'hypertension intracrânienne. Bien qu'on obtienne une meilleure sensibilité de détection de l'HYP, la précision est grandement réduite par rapport à la stratégie précédente. La Figure 4.10 et le Tableau 4.10

présentent les résultats de classification du réseau CNN-LSTM entraîné selon la stratégie simplifiée de l'HYP versus tous les autres états. Ces résultats montrent que cette stratégie permet d'atteindre une meilleure précision que le réseau CNN-LSTM plus complexe sans toutefois améliorer la sensibilité.

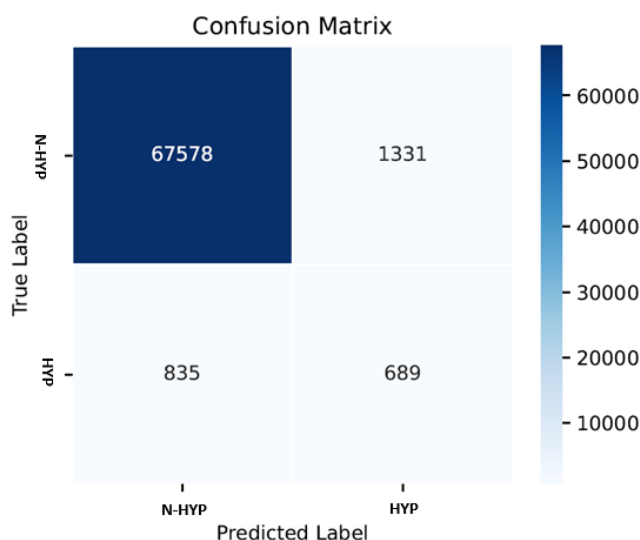


Figure 4.10 Matrice de confusion du réseau CNN-LSTM pour la détection de l'HYP

Tableau 4.10 Métriques du réseau CNN-LSTM pour la détection de l'HYP

États	Précision	Sensibilité	Score F1	Nombre
N-HYP	0.99	0.98	0.98	68909
HYP	0.34	0.45	0.39	1524
Accuracy			0.97	70433
Macro Average	0.66	0.72	0.69	
Weighted Average	0.97	0.97	0.97	

La Figure 4.11 présente cette fois les résultats du même réseau CNN-LSTM pour une classification entre l'état contrôle et la présence d'hyperémie sans considérer l'ischémie ni

l'hypertension intracrânienne. Tel qu'appuyé par le Tableau 4.10, les résultats montrent que cette stratégie augmente encore une fois la sensibilité de détection mais avec une réduction considérable de la précision.

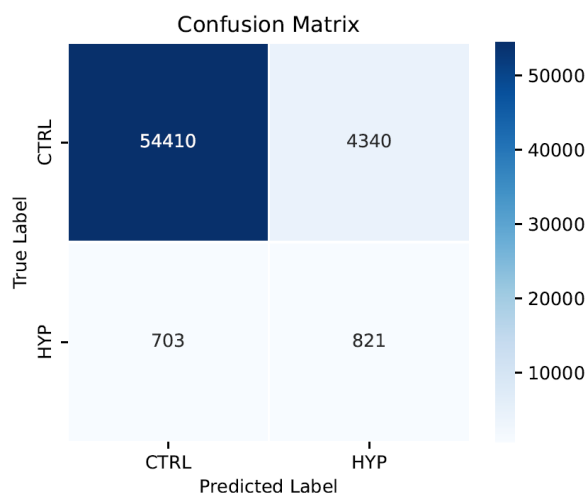


Figure 4.11 Matrice de confusion du réseau CNN-LSTM pour la détection de l'HYP v CTRL

Tableau 4.11 Métriques du réseau CNN-LSTM pour la détection de l'HYP v CTRL

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.99	0.93	0.96	58750
HYP	0.16	0.54	0.25	1524
Accuracy			0.92	60274
Macro Average	0.57	0.73	0.6	
Weighted Average	0.97	0.92	0.94	

La Figure 4.12 et le Tableau 4.12 présentent les résultats d'un modèle de type SVM entraîné selon la stratégie HYP versus tous les autres états. Les résultats montrent que ce type de classificateur simplifié présente des performances de classification beaucoup plus faibles que les réseaux de neurones.

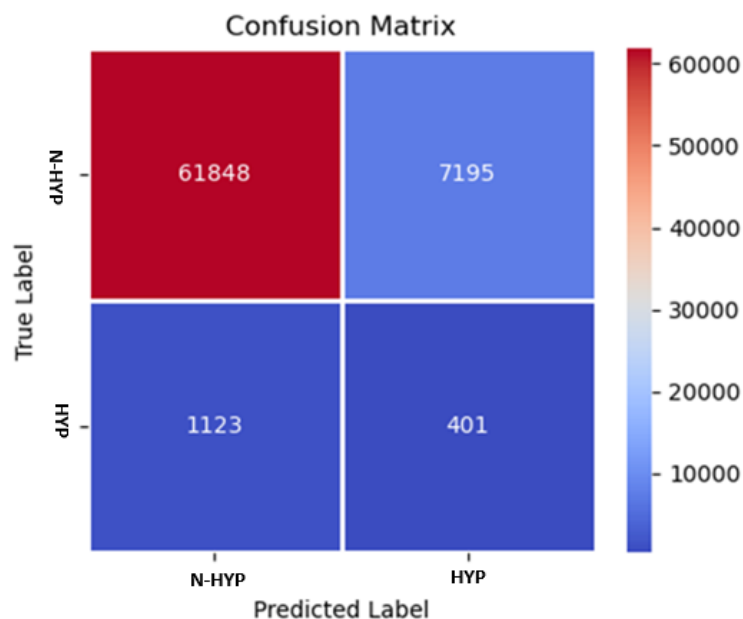


Figure 4.12 Matrice de confusion du SVM pour la détection de l'HYP

Tableau 4.12 Métriques du SVM pour la détection de l'HYP

États	Précision	Sensibilité	Score F1	Nombre
N-HYP	0.98	0.90	0.94	69043
HYP	0.05	0.26	0.09	1524
Accuracy			0.88	70567
Macro Average	0.52	0.58	0.51	
Weighted Average	0.96	0.88	0.92	

La Figure 4.13 et le Tableau 4.13 présentent les résultats du modèle RF entraîné pour la détection de l'HYP versus tous les autres états. Les résultats montrent que ce modèle présente une plus grande sensibilité de détection que les autres classificateurs, mais sans avoir une bonne précision.

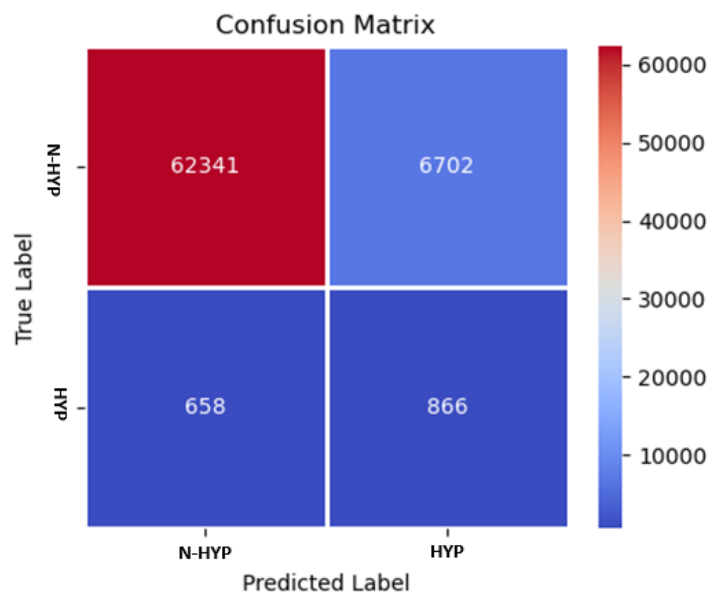


Figure 4.13 Matrice de confusion du RF pour la détection de l'HYP

Tableau 4.13 Métriques du RF pour la détection de l'HYP

États	Précision	Sensibilité	Score F1	Nombre
N-HYP	0.99	0.90	0.94	69043
HYP	0.11	0.57	0.19	1524
Accuracy			0.90	70567
Macro Average	0.55	0.74	0.57	
Weighted Average	0.97	0.90	0.93	

4.4 Détection de l'ischémie

Tout comme les deux états pathophysiologiques précédents, les mêmes stratégies d'entraînements ont été appliquées pour la détection de l'état ischémique.

La Figure 4.14 et le Tableau 4.14 présentent les résultats du réseau FCN entraîné pour la classification de l'état ISC versus tous les autres états possibles. Il est intéressant de constater que

lors de cette expérience, aucune séquence ischémique n'a été détectée correctement, suggérant que le réseau n'est pas en mesure de discriminer les deux groupes de données.

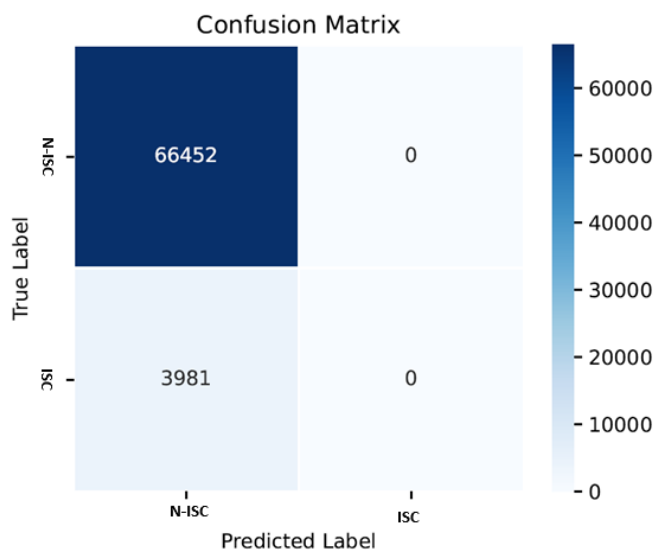


Figure 4.14 Matrice de confusion du FCN pour la détection de l'ISC

Tableau 4.14 Métriques du FCN pour la détection de l'ISC

États	Précision	Sensibilité	Score F1	Nombre
N-ISC	0.94	1.0	0.97	66452
ISC	0.0	0.0	0.0	3981
Accuracy			0.94	70433
Macro Average	0.47	0.50	0.49	
Weighted Average	0.89	0.94	0.92	

La Figure 4.15 et le Tableau 4.15 présentent les résultats du même réseau FCN pour une classification entre l'état contrôle et la présence d'ischémie sans considérer l'hyperémie ni l'hypertension intracrânienne. Ils montrent une légère amélioration de la détection d'ischémie par rapport à la stratégie précédente mais sans être significative. La détection d'ischémie reste très difficile.

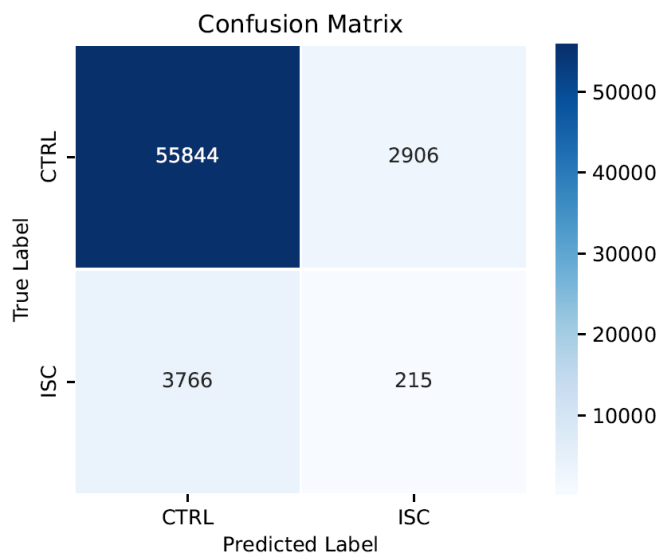


Figure 4.15 Matrice de confusion du FCN pour la détection de l'ISC v CTRL

Tableau 4.15 Métriques du FCN pour la détection de l'ISC v CTRL

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.94	0.95	0.94	58750
ISC	0.07	0.05	0.06	3981
Accuracy			0.89	62731
Macro Average	0.5	0.5	0.5	
Weighted Average	0.88	0.89	0.89	

La Figure 4.16 et le tableau 4.16 présentent les résultats du réseau CNN-LSTM avec la stratégie d'entraînement ISC versus N-ISC. Ces résultats montrent que le réseau CNN-LSTM, bien qu'il soit capable de détecter l'ISC, présente une sensibilité très faible et loin d'être acceptable.

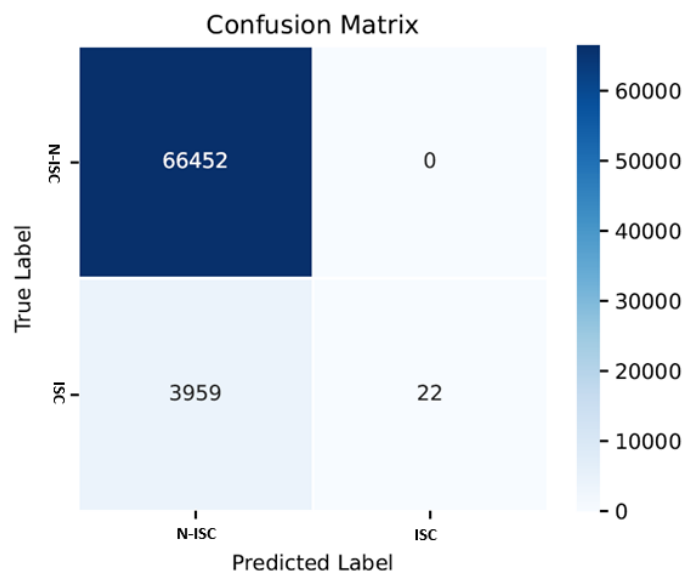


Figure 4.16 Matrice de confusion du CNN-LSTM pour la détection de l'ISC

Tableau 4.16 Métriques du CNN-LSTM pour la détection de l'ISC

États	Précision	Sensibilité	Score F1	Nombre
N-ISC	0.94	1.0	0.97	66452
ISC	1.0	0.01	0.01	3981
Accuracy			0.94	70433
Macro Average	0.97	0.5	0.49	
Weighted Average	0.95	0.94	0.92	

La Figure 4.17 et le Tableau 4.17 présentent les résultats du même réseau pour une classification entre l'état contrôle et la présence d'ischémie sans considérer l'hyperémie ni l'hypertension intracrânienne. Ces résultats montrent que le réseau présente une amélioration de la sensibilité de détection ainsi que de sa précision. La reconnaissance d'ischémie reste toutefois assez difficile.

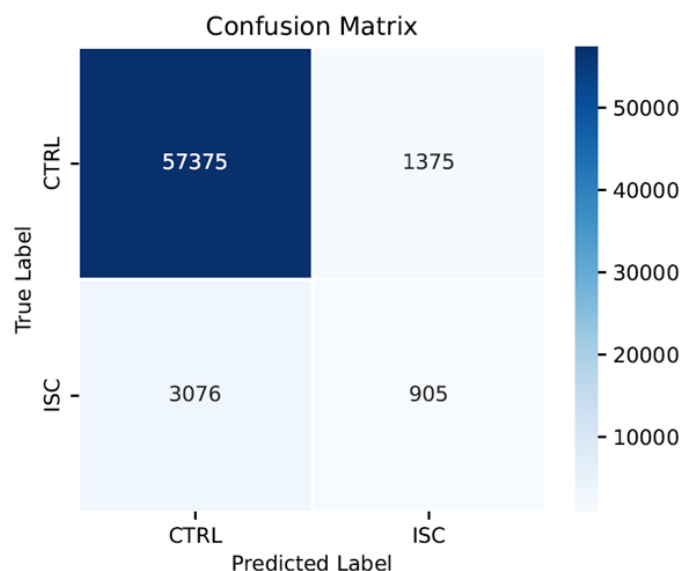


Figure 4.17 Matrice de confusion du CNN-LSTM pour la détection de l'ISC v CTRL

Tableau 4.17 Métriques du CNN-LSTM pour la détection de l'ISC v CTRL

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.95	0.98	0.96	58750
ISC	0.40	0.23	0.29	3981
Accuracy			0.93	62731
Macro Average	0.67	0.6	0.63	
Weighted Average	0.91	0.93	0.92	

La Figure 4.17 et le Tableau 4.17 présentent les résultats du modèle de type SVM avec une stratégie de classification ISC versus tous les autres états. On constate que les performances de classification pour cet état surpassent celles des deux réseaux de neurones avec toutefois une précision qui reste assez basse.

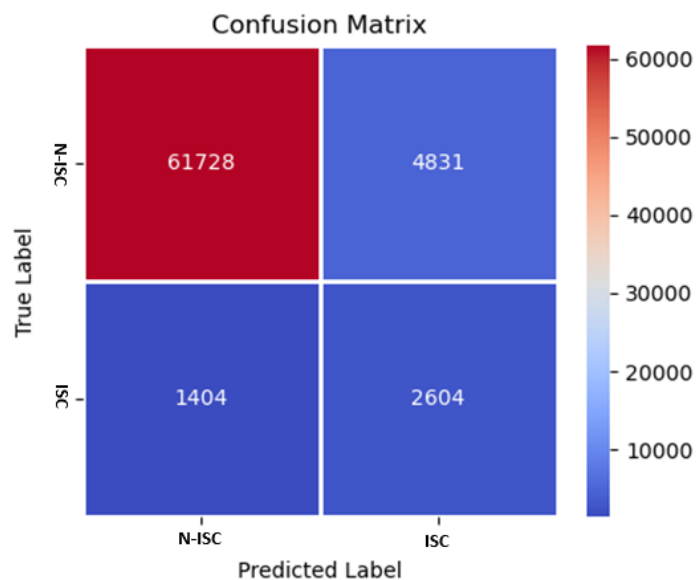


Figure 4.18 Matrice de confusion du SVM pour la détection de l'ISC

Tableau 4.18 Métriques du SVM pour la détection de l'ISC

États	Précision	Sensibilité	Score F1	Nombre
N-ISC	0.98	0.93	0.95	66559
ISC	0.35	0.65	0.46	4008
Accuracy			0.91	70567
Macro Average	0.66	0.79	0.70	
Weighted Average	0.94	0.91	0.92	

La Figure 4.19 et le Tableau 4.19 présentent les résultats du modèle RF de la même stratégie d'entraînement. Les résultats montrent que ce type de classificateur présente une bonne sensibilité à la détection d'ischémie mais la précision demeure basse et moins bonne que le SVM.

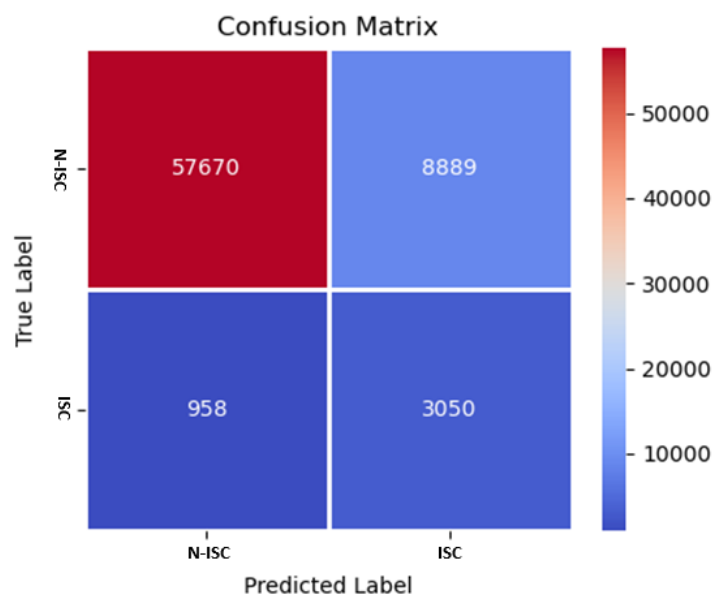


Figure 4.19 Matrice de confusion du RF pour la détection de l'ISC

Tableau 4.19 Métriques du RF pour la détection de l'ISC

États	Précision	Sensibilité	Score F1	Nombre
N-ISC	0.98	0.87	0.92	66559
ISC	0.26	0.76	0.38	4008
Accuracy			0.86	70567
Macro Average	0.62	0.81	0.65	
Weighted Average	0.94	0.86	0.89	

4.5 Résumé des résultats

Les Figures 4.20 à 4.24 résument la comparaison des performances de classifications pour les différentes expériences. De façon générale, nous avons pu mettre en évidence que différents types de classificateurs sont en mesure de détecter avec une grande performance la présence d'une HTIC (Figure 4.20). Ces résultats ne sont cependant pas observés pour d'autres états pathophysiologiques, laissant penser que la discrimination de ces états par rapport aux autres états est un enjeu plus complexe compte tenu des données acquises.

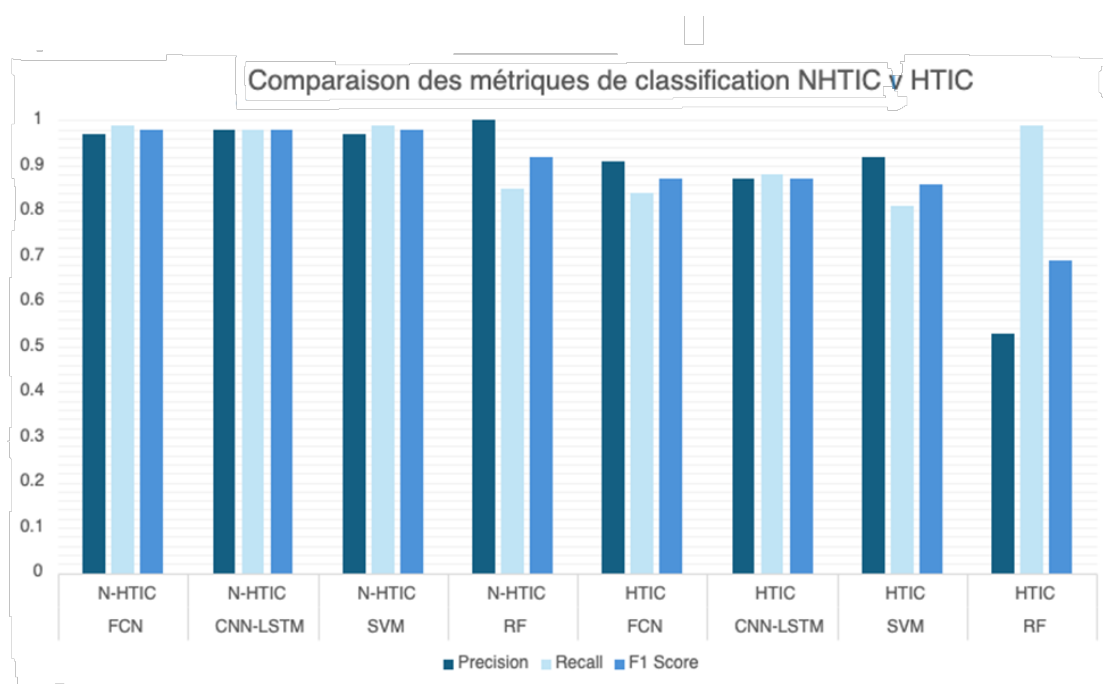


Figure 4.20 Comparaison des métriques de classification pour N-HTIC v HTIC

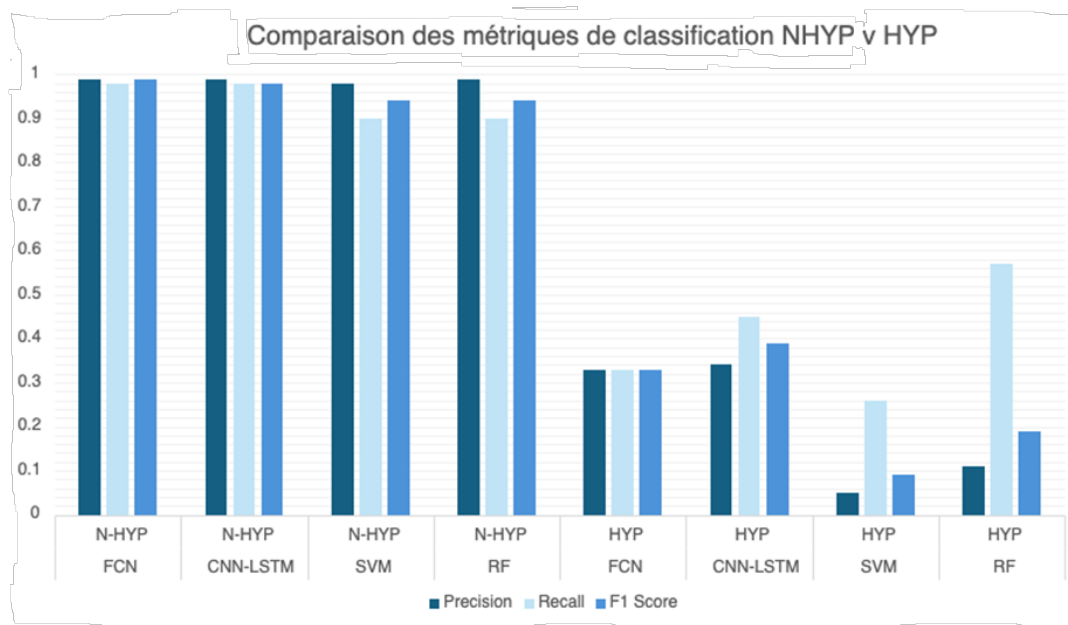


Figure 4.21 Comparaison des métriques de classification pour les états N-HYP v HYP

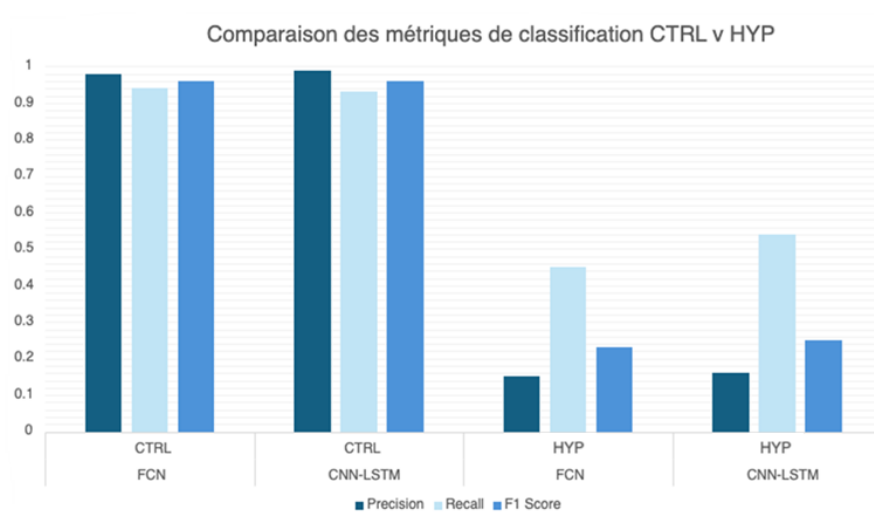


Figure 4.22 Comparaison des métriques de classification pour les états CTRL v HYP

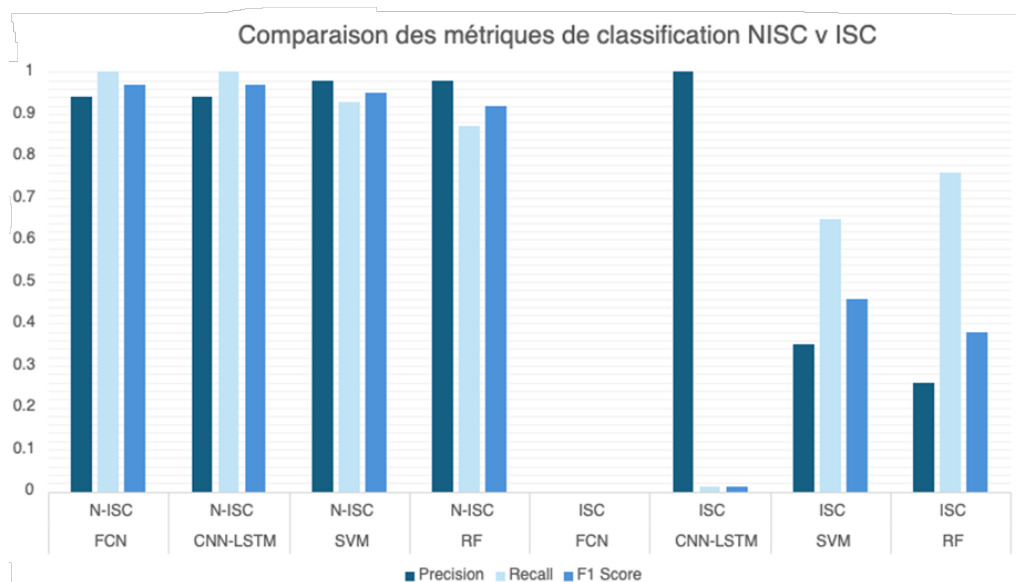


Figure 4.23 Comparaison des métriques de classification pour les états N-ISC v ISC

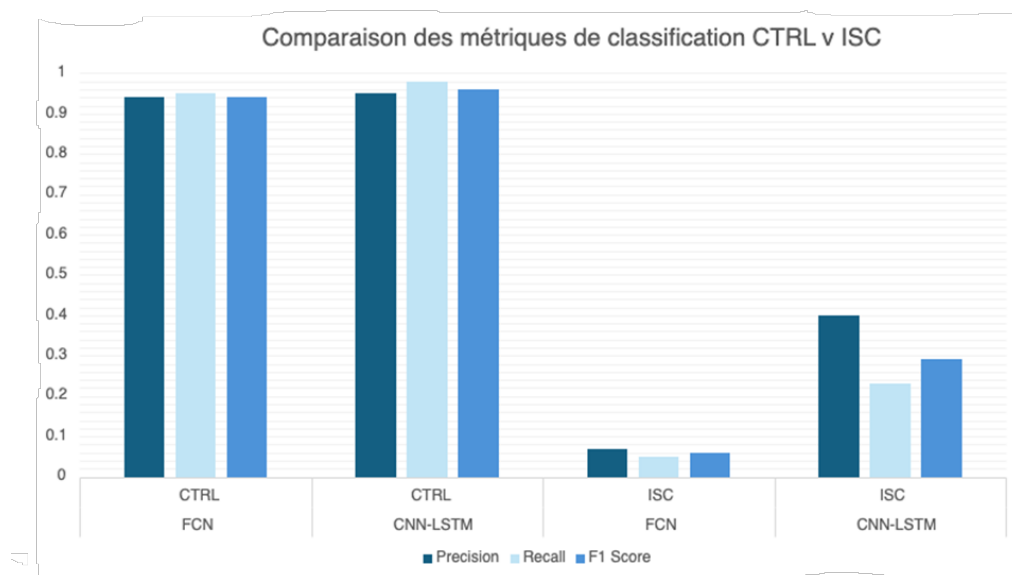


Figure 4.24 Comparaison des métriques de classification pour les états CTRL v ISC

CHAPITRE 5 DISCUSSION

Dans ce chapitre, une analyse des résultats présentés à la section précédente est faite afin de déterminer si les algorithmes d'apprentissage machine étudiés dans le cadre de ce travail de recherche permettent de produire une détection automatique de l'état crânio-cérébral chez les patients ayant subi un TCCS. Les limitations de ce travail de recherche sont aussi abordées et des pistes d'amélioration sont proposées.

5.1 Diagnostic à six états

Quel que soit le type de RNA à apprentissage profond, les résultats obtenus lors de cette expérience démontrent que la précision de détection pour la majorité des états pathophysiologiques est faible ce qui laisse croire qu'avec la base de données à disposition, il n'est pas envisageable de déployer de tels type de classificateurs complexes. Les performances de détection des états ISC et HYP sont notamment beaucoup plus basses que celles des états présentant une composante d'hypertension intracrânienne. Ceci peut s'expliquer par la nature plus complexe et ambiguë des états ISC et HYP alors que l'HTIC dépend uniquement de la pression intracrânienne. Il est à noter que parmi les états présentant de l'HTIC, HTIC sans ISC ni HYP est l'état le mieux détecté. Même s'il semble que les RNA étudiés ne puissent pas reproduire le raisonnement d'expert en la matière dans le cas de la classification à six états, il est toutefois possible de mettre à profit ces méthodes afin d'assister les professionnels de la santé en considérant des scénarios moins complexes et donc dans des situations cliniques plus spécifiques.

5.2 Détection d'un état pathophysiologique

En simplifiant le problème de classification en la détection de trois composantes indépendantes : la présence d'hypertension intracrânienne, d'ischémie et d'hyperémie, il est nécessaire d'entraîner trois classificateurs indépendamment au lieu d'un seul lors de la classification à six états. Cependant, les performances de classification se voient améliorées par rapport aux réseaux plus généralistes. Bien que la sensibilité de détection des états ISC et HYP n'est toujours pas satisfaisante, une amélioration de leur précision de détection est observée. Pour la composante HTIC, les performances de classification obtenue sont très bonnes et une distinction assez claire entre l'état contrôlé et la présence d'HTIC devient possible à détecter avec une grande précision. Bien que la classification à six états présente aussi de bonnes performances sur la détection de

l'état HTIC, il existe beaucoup de faux négatifs (HTIC-I et HTIC-H). En simplifiant la stratégie de classification à la présence ou non d'hypertension intracrânienne, la tâche du médecin peut se voir simplifiée, alors qu'il peut se concentrer à la détection d'ischémie ou d'hyperhémie. C'est en soit, une légère aide à la charge de travail, mais compte tenu du contexte surchargé des SI, cela est intéressant et est sans doute la meilleure application des algorithmes d'apprentissage machine au diagnostic du TCCS. Par exemple, dans une situation où un patient souffrirait de HTIC-I ou HTIC-H, la détection de la composante HTIC avec un haut niveau de certitude permet de réduire la charge de travail du personnel soignant en leur permettant de concentrer leur attention sur les composantes ischémique ou d'hyperhémie qui sont plus complexe à diagnostiquer. Cette aide est d'autant plus intéressante dans les centres hospitaliers ne possédant pas de surspécialisation en TCC.

5.3 Comparaison entre classificateurs

Il est aussi intéressant de noter que le réseau FCN qui est moins complexe que le CNN-LSTM présente des performances de classification similaire à celui-ci. Ceci est en ligne avec la littérature qui démontre que le FCN atteint souvent de bonnes performances comparées à des réseaux plus complexes sur différentes bases de données. Ce résultat montre que compte tenu des données disponibles l'ajout d'un réseau récurrent pour la prise en considération de l'aspect temporel des signaux physiologiques n'est pas primordial et que la simple architecture proposée par le réseau FCN est suffisante. Le modèle FCN est d'autant plus intéressant puisqu'il n'est pas totalement une boîte noire comparé au CNN-LSTM. Sa compatibilité avec des méthodes explicatives permet une plus grande transparence sur la classification. Cela peut contribuer à améliorer la compréhension du fonctionnement du modèle et ainsi augmenter la confiance des utilisateurs.

Pour la stratégie de classification binaire, il faut souligner que les classificateurs classiques à l'étude dans ce travail, soit le SVM à noyau RBF et le RF présentent parfois de meilleures performances de détection que les réseaux à apprentissage profond et ce notamment pour la détection de l'ischémie. Cela laisse croire qu'une combinaison de plusieurs types de classificateurs devrait être envisagée afin de couvrir plus qu'une détection de l'HTIC.

5.4 Limitation de la base de données

Comme tout problème de classification par apprentissage machine, la nature et la qualité des données disponibles a un impact important sur les résultats obtenus. Dans le cadre de ce projet, non seulement la quantité de patients disponibles est limitée, mais de façon plus importante, les observations concernant les états pathologiques sont grandement sous-représentées. Bien qu'une technique de balancement des classes ait été utilisée, il est évident qu'un plus grand nombre d'observations des cas pathologiques avec une plus grande variabilité améliorerait l'apprentissage de ces cas et augmenterait la capacité de détection. Aussi, comme cela a été mentionné dans la méthodologie, l'analyse expert pour chaque patient reflète l'avis d'un seul spécialiste. Idéalement, chaque patient devrait être annoté par tous les experts. En comparant ensuite les annotations, il devient plus facile d'éliminer les erreurs de diagnostic potentielles et d'uniformiser les annotations. De plus, l'annotation par expert des données disponibles a été faite de façon rétrospective. C'est -à-dire qu'aucune connaissance du contexte clinique n'était disponible au moment de leur analyse. Ceci peut bien entendu avoir un impact sur la classification expert. Idéalement les données recueillies devraient être annotées à même le contexte clinique. Pour se faire, un outil d'annotation automatique pourrait être mis en place basé sur les règles cliniques pour faire l'annotation sans augmenter la surcharge de travail du personnel soignant.

Idéalement, une base de données spécialisée pour le TCC devrait être mise en place contenant des séquences de signaux physiologiques ainsi que tout autre type d'informations utiles au diagnostic et à l'exploration de données comme le type de blessure subie, historique des traitements effectués en clinique ainsi que les médicaments prescrits par exemple ou les notes médicales au dossier. Ceci permettrait d'utiliser des méthodes non supervisées afin de découvrir des liens entre la condition des patients et les traitements recommandés.

En ce qui concerne les séquences temporelles de signaux physiologiques, il est préférable d'inclure des séquences qui représentent bien un état physiologique donné et pas nécessairement toutes les séquences enregistrées lors de la prise en charge de patients. Bien que la quantité de données soit importante, il faut aussi s'assurer que les informations disponibles représentent bien les pathophysiologies visées. Pour ce faire, un travail de sélection de patients pertinents mais surtout de séquences de données de longueur fixe devrait être fait pour construire une base de patients avec TCC qui puisse devenir une référence dans le domaine de la classification de l'état

crânio-cérébral de patients avec TCCS. Évidemment un tel travail nécessite beaucoup de ressources matérielles et humaines, mais est primordial afin de créer une référence pour ce type de diagnostic.

5.5 Améliorations possibles concernant la méthode

5.5.1 Méthode création des séquences temporelles

La méthode utilisée dans ce travail de recherche pour la création et l'étiquetage des séquences est assez simpliste. Associer une séquence d'une durée de 30 minutes à un état pathologique dès la présence d'un enregistrement pathologique augmente le nombre de séquences associées aux groupes d'état critique et favorise la détection dès les premiers signes de déviation. Cependant, elle crée aussi une ambiguïté entre les états contrôle et pathologiques du fait que les séquences d'un état contrôle deviennent similaires à un grand nombre de séquences critiques et augmente le nombre de faux négatifs. Des méthodes plus complexes devraient être envisagées. La création de séquences de séries temporelles et leur étiquetage est un domaine complexe et devrait être exploré davantage.

5.5.2 Validation croisée

Dans ce travail, la séparation des données patients a été faite en trois groupes distincts. Vu la faible quantité de données des états pathologiques disponibles ce type de séparation influence les résultats obtenus. Un apprentissage sur des patients présentant une trop grande variabilité par rapport aux patients tests résultent en de mauvais scores de classification, alors que dans le cas contraire les métriques laissent croire à tort que le modèle performe très bien.

Une meilleure façon de procéder serait l'utilisation de la validation croisée permettant de mieux évaluer la capacité de généraliser du modèle. Cette méthode se fait en plusieurs étapes appelées "K-fold". À chaque étape, une partie des patients est utilisée pour l'apprentissage alors que le reste est utilisé pour le test. De cette façon, toutes les données disponibles sont utilisées lors de l'entraînement et du test. Cependant, c'est un processus assez long à performer puisque le modèle à l'étude doit être entraîné k fois. Lorsque utilisée avec une grande base de données et des modèles complexes, cette méthode devient très coûteuse en temps d'entraînement.

CHAPITRE 6 CONCLUSION

Dans ce mémoire, l'objectif était d'évaluer la capacité d'algorithmes d'apprentissage automatique, notamment les réseaux de neurones à apprentissage profond, pour assister le personnel soignant dans le diagnostic en temps réel de l'état crânio-cérébral de patients ayant subi un TCC sévère aux soins intensifs. Les réseaux FCN et CNN-LSTM ont été choisis sur base de leur capacité à analyser des séries temporelles de données physiologiques. Ces deux réseaux ont été entraînés et testés avec les données d'enregistrement de signaux physiologiques de 49 patients adultes avec TCCS durant leur prise en charge à l'hôpital Sacré Cœur de Montréal. Ces données avaient été annotés *a posteriori* par deux experts intensivistes qui ont classifié l'état crânio-cérébral des patients en considérant cinq états pathophysiologiques d'intérêt (HTIC, HTIC-I, HTIC-H, ISC et HYP) ainsi que l'état contrôlé. Dans un premier temps, les faibles performances de classification de ces deux réseaux pour catégoriser les six états laissent croire que cette tâche est trop complexe pour les réseaux à l'étude compte tenu des données d'entraînements disponibles. Cependant, la simplification du problème en plusieurs classifications binaires entre l'état contrôle et la présence d'hypertension intracrânienne, d'ischémie ou d'hyperémie montre qu'il est possible de détecter la présence d'hypertension intracrânienne de façon satisfaisante avec les réseaux de neurones à apprentissage profond. La comparaison des performances des réseaux FCN et CNN-LSTM avec ceux de ces classificateurs moins complexes tels que le SVM et le RF montre toutefois que le classificateur de type SVM performe mieux dans la détection de l'ischémie. Bien que des méthodes d'apprentissage automatiques aient déjà été mises à profit dans le contexte des SI afin d'aider la prise en charge de patients traumatisés crâniens de diverses façons, aucune application de détection automatique en temps réel de l'état précis d'un patient basé sur l'évaluation de plusieurs signaux physiologiques n'a été relevé dans la littérature à ce jour. Les résultats obtenus ici suggèrent qu'il n'est pas envisageable de déployer en clinique des réseaux de neurones pour une tâche aussi complexe que le diagnostic de l'état crânio-cérébral de patients avec TCCS à ce stade. Pour le moment, les efforts devraient s'orienter vers l'acquisition à grande échelle de données relatives aux patients avec TCCS aux SI et la mise en place d'une base de données maximisant les exemples de cas pathophysiologiques la plus diversifiée possible. Par la suite, il serait intéressant d'optimiser le processus d'annotation des données recueillies et prendre soin d'inclure différents types d'informations relatives aux patients dans cette base de données. Cette façon de faire

facilitera l'exploration des données et permettra d'y appliquer plus facilement différents types d'algorithmes d'apprentissage automatique de complexités variées pour la découverte de nouvelles connaissances ou la classification médicale. La mise en place d'une telle base de données nécessite évidemment de grandes ressources financières, matérielles et humaines mais présente une étape primordiale à l'avancement du projet OptiBrain si l'intelligence artificielle veut y être intégrée. Additionnellement, une approche d'apprentissage en continu pourrait être envisagée, permettant d'ajuster et de développer une base de données tout en améliorant la performance de modèles d'apprentissage machine à l'identification d'états physiopathologiques.

RÉFÉRENCES

- [1] G.B.D.T.B. Injury et C. Spinal Cord Injury, "Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016," *Lancet Neurol*, vol. 18, no 1, p. 56-87, Jan 2019. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/30497965>
- [2] Centers for Disease Control and Prevention, "Moderate and severe TBI", [En ligne]. Disponible: <https://www.cdc.gov/traumaticbraininjury/moderate-severe/index.html> (Consulté le 4 avril 2024)
- [3] G. Emeriaud, G. Pettersen et B. Ozanne, "Pediatric traumatic brain injury: an update," *Curr Opin Anaesthesiol*, vol. 24, no 3, p. 307-13, Jun 2011. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/21494129>
- [4] P. Le Roux et al., "Consensus summary statement of the International Multidisciplinary Consensus Conference on Multimodality Monitoring in Neurocritical Care: a statement for healthcare professionals from the Neurocritical Care Society and the European Society of Intensive Care Medicine," *Intensive Care Med*, vol. 40, no 9, p. 1189-209, Sep 2014. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/25138226>
- [5] D. Gupta et al., "Guideline Adherence and Outcomes in Severe Adult Traumatic Brain Injury for the CHIRAG (Collaborative Head Injury and Guidelines) Study," *World Neurosurg*, vol.89, p. 169-79, May 2016. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/26806065>
- [6] S. Fartoumi et al., "Computerized Decision Support System for Traumatic Brain Injury Management," *J Pediatr Intensive Care*, vol. 5, no 3, p. 101-107, Sep 2016. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/31110893>
- [7] J. V. Rosenfeld et al., "Early management of severe traumatic brain injury," *Lancet*, vol. 380, no 9847, p. 1088-98, Sep 22, 2012. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/22998718>
- [8] D. H. Smith, D. F. Meaney et W. H. Shull, "Diffuse axonal injury in head trauma," *J Head Trauma Rehabil*, vol. 18, no 4, p. 307-16, Jul-Aug 2003. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/16222127>

- [9] R. L. Gruen et M. C. Fitzgerald, "Towards a maturation of trauma research," *The Lancet*, vol. 380, no 9847, p. 1033-1034, Sept 2012. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0140673612614184>
- [10] M. J. Aries et al., "Continuous determination of optimal cerebral perfusion pressure in traumatic brain injury," *Crit Care Med*, vol. 40, no 8, p. 2456-63, Aug 2012. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/22622398>
- [11] K. B. Evensen et P. K. Eide, "Measuring intracranial pressure by invasive, less invasive or non-invasive means: limitations and avenues for improvement," *Fluids and Barriers of the CNS*, vol. 17, no 1, p. 34, May 2020. [En ligne]. Disponible: <https://doi.org/10.1186/s12987-020-00195-3>
- [12] V. Stambolija et al., "PbtO(2) monitoring in normobaric hyperoxia targeted therapy in acute subarachnoidal hemorrhage," *Surg Neurol Int*, vol. 9, p. 46, 2018. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/29541487>
- [13] Z. A. Tsigaras et al., "The pressure reactivity index as a measure of cerebral autoregulation and its application in traumatic brain injury management," *Critical Care and Resuscitation*, vol. 25, no 4, p. 229-236, Dec 2023. [En ligne]. Disponible: <https://www.sciencedirect.com/science/article/pii/S1441277223022202>
- [14] B. Depreitere et al., "Pressure autoregulation monitoring and cerebral perfusion pressure target recommendation in patients with severe traumatic brain injury based on minute-by-minute monitoring data," *J Neurosurg*, vol. 120, no 6, p. 1451-7, Jun 2014. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/24745709>
- [15] C. D. Wickens et al., *Engineering Psychology and Human Performance*. London, États-Unis: Taylor & Francis Group, 2012.
- [16] R. S. Ledley et L. B. Lusted, "Reasoning foundations of medical diagnosis; symbolic logic, probability, and value theory aid our understanding of how physicians reason," *Science*, vol. 130, no 3366, p. 9-21, Jul 1959.
- [17] H. R. Warner et al., "A Mathematical Approach to Medical Diagnosis: Application to Congenital Heart Disease," *JAMA*, vol. 177, no 3, p. 177-183, 1961. [En ligne]. Disponible: <https://doi.org/10.1001/jama.1961.03040290005002>

- [18] G. S. Lodwick, "Radiographic diagnosis and grading of bone tumors, with comments on computer evaluation," *Proc Natl Cancer Conf*, vol. 5, p. 369-80, 1964. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/5856167>
- [19] F. T. De Dombal, "Computer-aided decision support in acute abdominal pain, with special reference to the EC concerted action," *Int J Biomed Comput*, vol. 26, no 3, p. 183-8, Sep 1990. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/2242936>
- [20] E. H. Shortliffe et al., "An artificial intelligence program to advise physicians regarding antimicrobial therapy," *Comput Biomed Res*, vol. 6, no 6, p. 544-60, Dec 1973.
- [21] R. M. Gardner, T. A. Pryor et H. R. Warner, "The HELP hospital information system: update 1998," *International journal of medical informatics*, vol. 54 3, p. 169-82, 1999.
- [22] P. Smielewski et al., "ICM+, a flexible platform for investigations of cerebrospinal dynamics in clinical practice," *Acta Neurochir Suppl*, vol. 102, p. 145-51, 2008. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/19388307>
- [23] H. Gomez et al., "Development of a multimodal monitoring platform for medical research," *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2010, p. 2358-61, 2010. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/21097226>
- [24] J. A. Wilson, L. A. Shutter et J. A. Hartings, "COSBID-M3: a platform for multimodal monitoring, data collection, and research in neurocritical care," *Acta Neurochir Suppl*, vol.115,p.67-74,2013.[Enligne].Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/22890647>
- [25] F. Wu et al., "A modular Clinical Decision Support System Clinical prototype extensible into multiple clinical settings," communication présentée à *2009 3rd International Conference on Pervasive Computing Technologies for Healthcare*, 1-3 April 2009, p. 1-4.
- [26] A. Bagnall et al., "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances," *Data Mining and Knowledge Discovery*, vol. 31, no 3, p. 606-660, May 2017. [En ligne]. Disponible: <https://doi.org/10.1007/s10618-016-0483-9>
- [27] Y. A. LeCun et al., "Efficient BackProp," dans *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr et K.-R. Müller, édit. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, p. 9-48.

- [28] C. Szegedy et al., "Going deeper with convolutions," communication présentée à *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7-12 June 2015, 2015, p. 1-9.
- [29] Y. LeCun, Y. Bengio et G. Hinton, "Deep learning," *Nature*, vol. 521, no 7553, p. 436-444, Mai 2015. [En ligne]. Disponible: <https://doi.org/10.1038/nature14539>
- [30] Z. Wang, W. Yan et T. Oates, "Time series classification from scratch with deep neural networks: A strong baseline," *2017 International Joint Conference on Neural Networks (IJCNN)*, p. 1578-1585, 2016.
- [31] B. Zhou et al., "Learning deep features for discriminative localization," communication présentée à *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, p. 2921-2929.
- [32] R. Pascanu, T. Mikolov et Y. Bengio, "On the difficulty of training recurrent neural networks," communication présentée à *International conference on machine learning*, 2013, p. 1310-1318.
- [33] S. Hochreiter et J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no 8, p. 1735-1780, 1997.
- [34] A. Ben-Hur et al., "Support vector machines and kernels for computational biology," *PLoS Comput Biol*, vol. 4, no 10, p. e1000173, Oct 2008.
- [35] Y. Y. Song et Y. Lu, "Decision tree methods: applications for classification and prediction," *Shanghai Arch Psychiatry*, vol. 27, no 2, p. 130-5, Apr 25, 2015.
- [36] F. Rajaei et al., "AI-Based Decision Support System for Traumatic Brain Injury: A Survey," *Diagnostics (Basel)*, vol. 13, no 9, May 2023. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/37175031>
- [37] K. A. Kim et al., "Artificial Intelligence-Enhanced Neurocritical Care for Traumatic Brain Injury: Past, Present and Future," *J Korean Neurosurg Soc*, Jan 2024. [En ligne]. Disponible: <https://www.ncbi.nlm.nih.gov/pubmed/38186369>
- [38] A. T. Hale et al., "Using an artificial neural network to predict traumatic brain injury," *J Neurosurg Pediatr*, vol. 23, no 2, p. 219-226, Nov 2018. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/30485240>

- [39] İ. Güler, Z. Gökçil et E. Gülbandilar, "Evaluating of traumatic brain injuries using artificial neural networks," *Expert Systems with Applications*, vol. 36, no 7, p. 10424-10427, Sep 2009. [Enligne]. Disponible: <https://www.sciencedirect.com/science/article/pii/S0957417409000773>
- [40] D. Abe et al., "A Prehospital Triage System to Detect Traumatic Intracranial Hemorrhage Using Machine Learning Algorithms," *JAMA Netw Open*, vol. 5, no 6, p. e2216393, Jun 2022. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/35687335>
- [41] J. D. Moyer et al., "Machine learning-based prediction of emergency neurosurgery within 24 h after moderate to severe traumatic brain injury," *World J Emerg Surg*, vol. 17, no 1, p. 42, Aug 2022.
- [42] A. Habibzadeh et al., "Machine learning-based models to predict the need for neurosurgical intervention after moderate traumatic brain injury," *Health Sci Rep*, vol. 6, no 11, p. e1666, Nov 2023.
- [43] M. Hanko et al., "Random Forest-Based Prediction of Outcome and Mortality in Patients with Traumatic Brain Injury Undergoing Primary Decompressive Craniectomy," *World Neurosurg*, vol. 148, p. e450-e458, Apr 2021.
- [44] K. C. Tu et al., "A Computer-Assisted System for Early Mortality Risk Prediction in Patients with Traumatic Brain Injury Using Artificial Intelligence Algorithms in Emergency Room Triage," *Brain Sci*, vol. 12, no 5, May 2022.
- [45] T. Tunthanathip et T. Oearsakul, "Application of machine learning to predict the outcome of pediatric traumatic brain injury," *Chin J Traumatol*, vol. 24, no 6, p. 350-355, Nov 2021.
- [46] M. Cvach, "Monitor alarm fatigue: an integrative review," *Biomed Instrum Technol*, vol. 46, no 4, p. 268-77, Jul-Aug 2012.
- [47] S. B. Lee et al., "Artifact removal from neurophysiological signals: impact on intracranial and arterial pressure monitoring in traumatic brain injury," *J Neurosurg*, vol. 132, no 6, p. 1952-1960, May 10, 2019.
- [48] N. Schweingruber et al., "A recurrent machine learning model predicts intracranial hypertension in neurointensive care patients," *Brain*, vol. 145, no 8, p. 2910-2919, Aug 2022.

- [49] M. Grewal et al., "RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, p. 281-284, 2017.
- [50] P. D. Chang et al., "Hybrid 3D/2D Convolutional Neural Network for Hemorrhage Evaluation on Head CT," *AJNR Am J Neuroradiol*, vol. 39, no 9, p. 1609-1616, Sep 2018.
- [51] D. Apiletti et al., "Real-time analysis of physiological data to support medical applications," *IEEE Trans Inf Technol Biomed*, vol. 13, no 3, p. 313-21, May 2009.
- [52] I. T. Jolliffe et J. Cadima, "Principal component analysis: a review and recent developments," *Philos Trans A Math Phys Eng Sci*, vol. 374, no 2065, p. 20150202, Apr 13 2016.
- [53] Imbalanced-learn, "NearMiss".[EnLigne].Disponible:https://imbalanced-learn.org/stable/references/generated/imblearn.under_sampling.NearMiss.html (Consulté le 4 avril 2024)
- [54] A. Demirkaya, J. Chen et S. Oymak, "Exploring the Role of Loss Functions in Multiclass Classification," communication présentée à *2020 54th Annual Conference on Information Sciences and Systems (CISS)*, 18-20 March 2020 2020, p. 1-5.
- [55] E. M. Dogo et al., "A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks," communication présentée à *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 21-22 Dec. 2018, p. 92-99.
- [56] M. Sokolova et G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no 4, p. 427-437, July 2009.

ANNEXE A RÈGLES DE DÉCISION DE L'ÉTAT CRÂNIO-CÉRÉBRAL

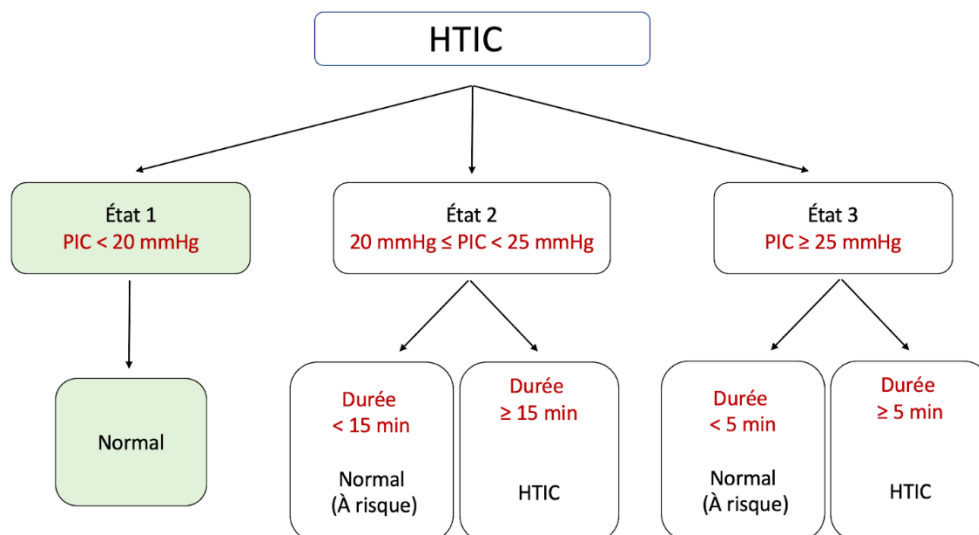


Figure A.1 Règles de décision pour l'HTIC

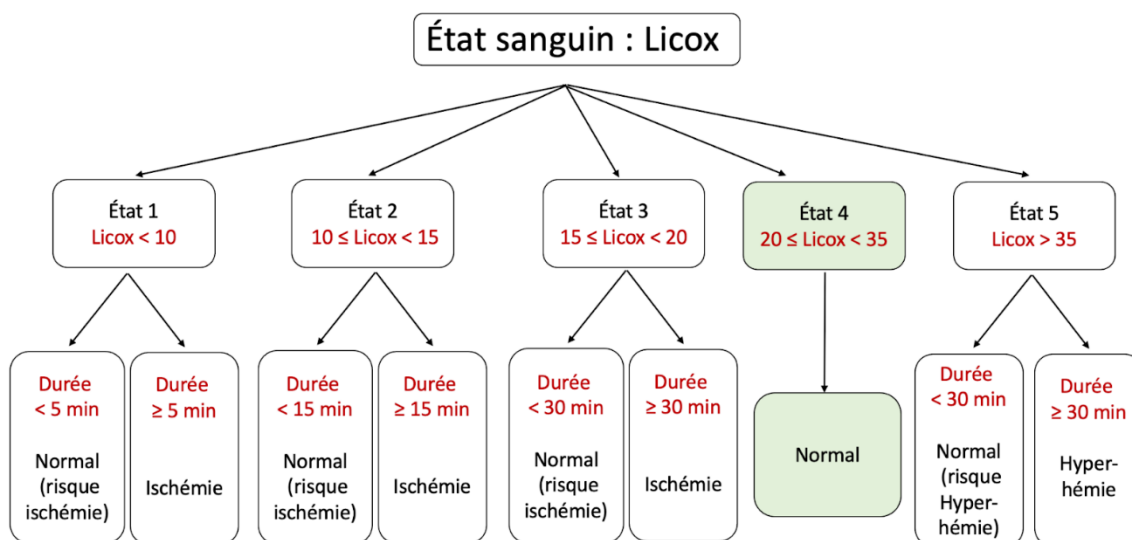


Figure A.2 Règles de décision pour l'ISC et l'HYP basé sur le Licox (PbtO₂)

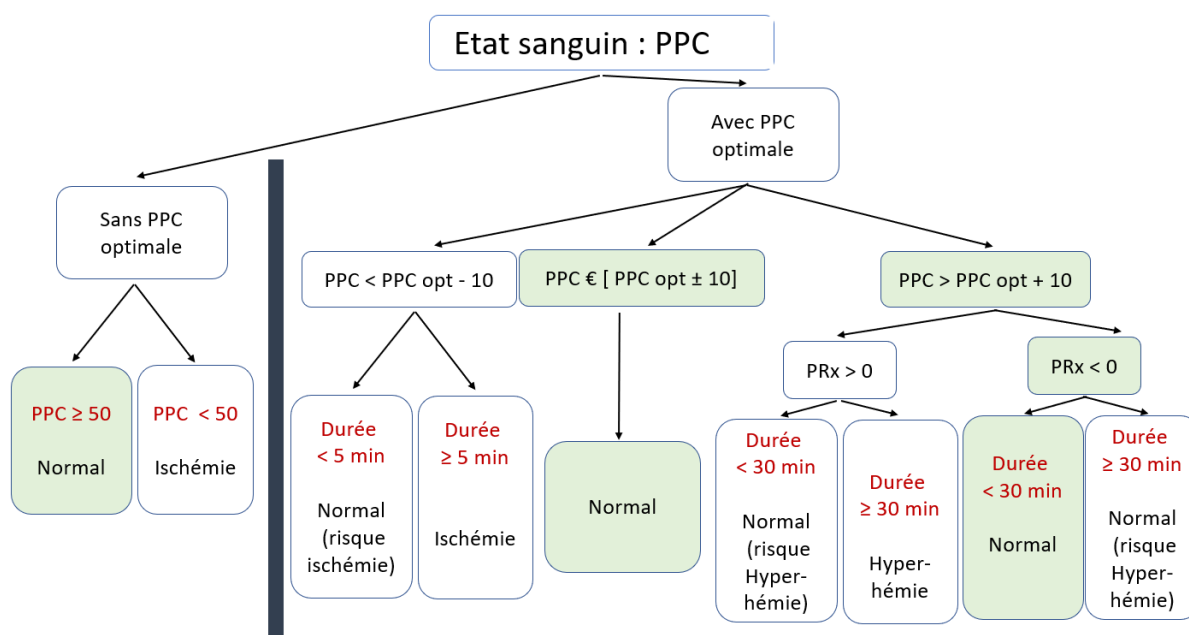


Figure A.3 Règles de décision pour l'ISC et l'HYP basé sur la PPC

État sanguin Final					
Ischémie		Hyperhémie		Ischémie + Hyperhémie	
États (Licox+ PPC)	Décision	États(Licox+ PPC)	Décision	États(Licox+ PPC)	Décision
N + N	Normal	N + N	Normal	Ri + Rh	Normal
N + Ri	Normal	N + Rh	Normal	Ri + H	Normal
N + I	Normal	N + H	Normal	I + Rh	Ischémie
Ri + N	Normal	Rh + N	Normal	I + H	Ischémie
Ri + Ri	Normal	Rh + Rh	Normal		
Ri + I	Ischémie	Rh + h	Hyperhémie	H + Ri	Normal
I + N	Ischémie	H + N	Normal	H + I	Normal
I + Ri	Ischémie	H + Rh	Hyperhémie	Rh + I	Normal
I + I	Ischémie	H + H	Hyperhémie	Rh + Ri	Normal

N : Normal I : Ischémie Ri : Risque Ischémie H: Hyperhémie Rh : Risque Hyperhémie

Figure A.4 Table de décision finale pour l'ISC et l'HYP basé sur le Licox et la PbTO₂

ANNEXE B RÉSULTATS SUPPLÉMENTAIRES

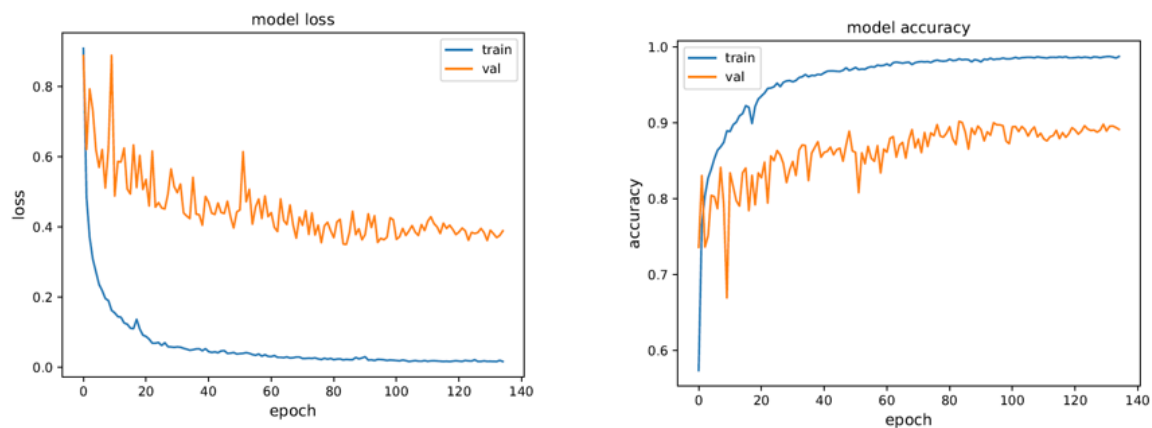


Figure B.1 Apprentissage et de validation (fonction de coût et précision) : CNN-LSTM

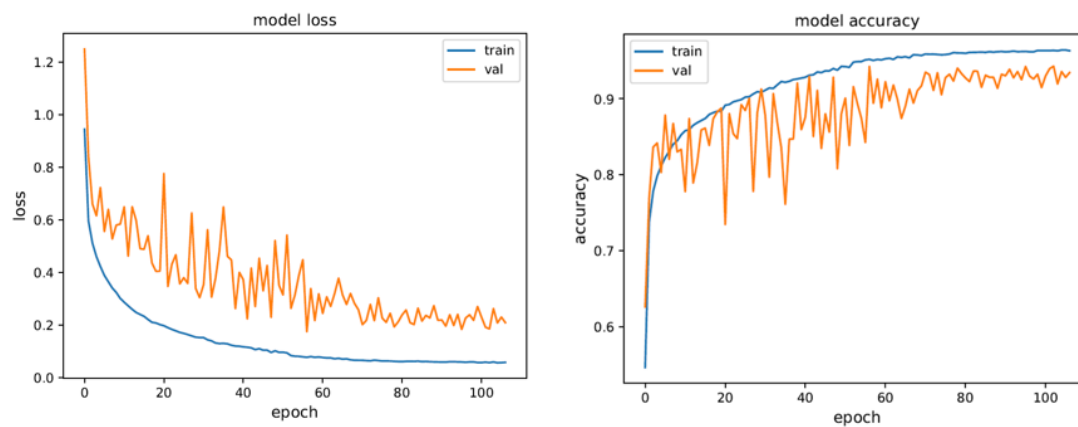


Figure B.2 Apprentissage et de validation (fonction de coût et précision) : FCN

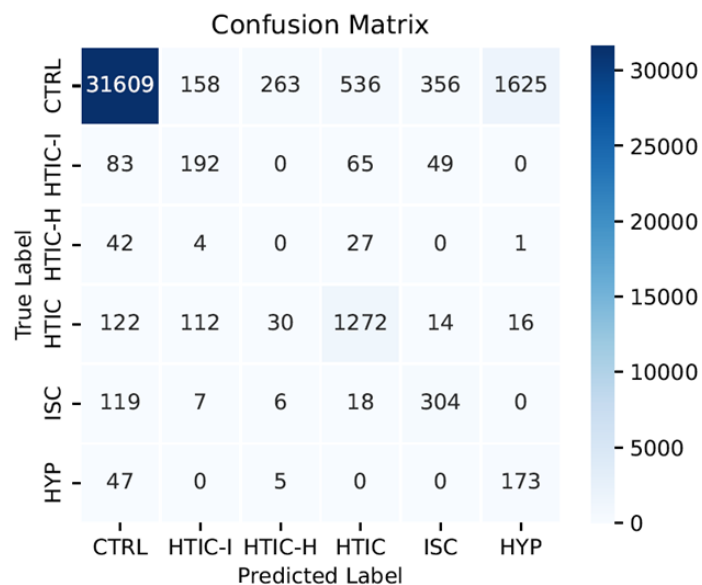


Figure B.3 Matrice de confusion du CNN-LSTM sur le groupe de développement

Tableau B.1 Métriques du CNN-LSTM sur le groupe de développement

États	Précision	Sensibilité	Score F1	Nombre
CTRL	0.99	0.91	0.95	34547
HTIC-I	0.41	0.49	0.45	389
HTIC-H	0.0	0.0	0.0	74
HTIC	0.66	0.81	0.73	1566
ISC	0.42	0.67	0.52	454
HYP	0.10	0.77	0.17	225
Accuracy			0.90	37255
macro average	0.43	0.61	0.47	
weighted average	0.95	0.90	0.92	