

République Algérienne Démocratique et Populaire Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université des Sciences et de la Technologie Houari Boumediene Faculté d'Électronique et d'Informatique Département d'Informatique Mémoire de Master Spécialité : Informatique Visuelle

Détection et analyse automatique des sursaut radio de type 2 par des méthodes de deep learning

Supervisé par : Mr. khalil daifallah

Présenté par : ZEMAM ikram

Membre du Jurys : Amrous Imene Got Adel

Soutenu le : dd/mm/yyyy

Projet : MIV_13 / 2023

REMERCIEMENTS

Tout d'abord, je remercie Allah le Tout-Puissant de m'avoir donné le courage et la patience nécessaires pour mener à bien ce travail. Que les membres du jury trouvent ici l'ex-

pression de mes sincères remerciements pour l'honneur qu'ils me font en prenant le temps de lire et d'évaluer ce travail. J'aimerais remercier notre promoteur et chef de division de craag

, M. Kalil Daifallah, ainsi que l'équipe de CRAAG, pour leur aide compétente, leur patience et leur encouragement.

Je remercie également la commission de suivi, M. Adel Got, pour ses conseils et son soutien qui m'ont encouragé à ne pas abandonner.

Je tiens à remercier Ikram Djeghali, Mounir Dekali, Chakib Boudjema et le groupe Ameno pour leur immense aide et soutien tout au long de mon projet et de mon cursus.

Enfin, je remercie tous ceux qui ont contribué directement ou indirectement à la réalisation de ce travail.

Ikram

RÉSUMÉ

Une étude sur les sursauts radio de type II a été réalisée dans le but de développer un système de détection et d'analyse basé sur les méthodes de deep learning. Pour cela, nous avons effectué une collecte de données en rassemblant 4 426 échantillons à partir du site e-callisto. Les données ont ensuite été soumises à une étape de nettoyage et de prétraitement. Ensuite, nous avons construit des architectures de réseaux de neurones et les avons testées, obtenant une précision de 92% en classification. Enfin, nous avons procédé à une analyse des images .fits des sursauts radio de type II.

Mots-clés : sursaut radio Solaire, Neuron network , Apprentissage profond , Classification, Type II, traitement d'image , Segmentation, Deep larning.

ABSTRACT

A study on type II solar radio bursts was conducted with the aim of developing a detection and analysis system based on deep learning methods. To achieve this, we collected 4,426 samples from the e-callisto website. The collected data underwent a cleaning and preprocessing step. Subsequently, we constructed neural network architectures and tested them, achieving a classification accuracy of 92%. Finally, we performed an analysis of the .fits images of type II radio bursts.

Keywords : Solar radio burst, Neural network, Deep learning, Classification, Type II, Image processing, Segmentation, Deep learning.

_____TABLE DES MATIÈRES

In	trod	uction générale	1		
1 Généralités sur le Soleil					
	1.1	Soleil	3		
	1.2	L'activité solaire	4		
		1.2.1 Les taches solaires	4		
		1.2.2 Les éruptions solaires (Flare)	5		
		1.2.3 Les éjections de masse coronale (CMEs)	5		
	1.3	Sursaut radio solaire	6		
		1.3.1 Sursaut radio de typeII	$\overline{7}$		
	1.4	Réseau E-CALLISTO radio récepteurs solaires	8		
	1.5	Interaction Soleil - Terre (Space weather)	9		
	1.6	L'intérêt d'étudier les sursauts radio solaires de type II	10		
2	Éta	t de l'art	11		
	2.1	Introduction	11		
	2.2	Imagerie	11		
		2.2.1 Images et types de format d'images :	11		
		2.2.2 Prétraitement des images :	12		
	2.3	Apprentissage Profond	18		
		2.3.1 Réseaux de neurones	19		
		2.3.2 Fonctions d'activation	20		
		2.3.3 Fonctions de perte	20		
		2.3.4 Sur-apprentissage et Sous-apprentissage	21		
		2.3.5 Forward Propagation et Backpropagation	21		
	2.4	Convolutional Neural Networks	22		
		2.4.1 les couches du CNN	23		
		2.4.2 5 architectures bien connues de CNN	23		
		2.4.3 Étude comparatif	26		
		2.4.4 Travaux liés	27		

	2.5	Conclusion	28
3	Con	aception	29
	3.1	Introduction	29
	3.2	Workflow	30
	3.3	Partie 1 : data-set	30
	0.0	3 3 1 Extraction de données	30
		3.3.2 Les étapes suivit pour la collecte	34
		3.3.3 Analyse exploratoire des données	36
		3.3.4 Pré traitement et nettovage des données	37
	34	Partie 2 : entraînement des modèles	41
	0.1	3/1 Architecture proposé	/1
	35	Analyse des ondes radio type II	45
	0.0	3.5.1 Présentation de la structure des images fits	45
		2.5.2 Analyze des métadonnées	40
		2.5.2 Analyse des inetadonnees	40
		2.5.4 Máthada 1. A paluza automatique	41
	26	5.5.4 Methode I : Analyse automatique	48
	3.0	Conclusion	10
4	Réa	lisation et Evaluation	52
	4.1	Introduction	52
	4.2	Environnement de travail	52
		4.2.1 Environnement matériel	52
		4.2.2 Environnement logiciel	52
	4.3	Réalisation	54
	4.4	Évaluation	56
		4.4.1 Les métriques de performances	56
		4.4.2 Partie 1 : l'entraînement et validation	57
	4.5	Interface Utilisateur	61
	4.6	Conclusion	61
	1.0		01
Co	onclu	ision générale	62

TABLE DES FIGURES

1.1	Structure du soleil $[31]$	4
1.2	Exemple de taches solaires	4
1.3	Une image réelle d'une éruption solaire vue par trois longueur d'ondes le 14	
	Mars 2014 Crédit : NASA/SDO/Goddard Space Flight Center	5
1.4	Une image qui montre le déclenchement de deux CMEs dans différentes di-	
	rections le 08 Novembre 2000. Crédit : ESA/NASA/SOHO	6
1.5	Représentation d'un sursaut radio de type 2, où "H" désigne l'harmonique et	
	"F" la fondamentale.	8
1.6	Le réseau e-CALLISTO jusqu'à la date du 21 juin 2023	9
1.7	Représentation de l'interaction Soleil - Terre	9
2.1	Représentation de la structure d'une image de la format FITS	12
2.2	Exemple de dilatations sur des images binaires.[39]	13
2.3	Exemple d'érosion sur des images binaires.[39]	13
2.4	exemple d'application d'un filtre gaussien sur une image avec une matrice de	
	convolution de taille 5x5. [41] \ldots \ldots \ldots \ldots \ldots	14
2.5	Calcule du filtre médian sur une image. [39]	15
2.6	Représente la détection de contour et filtrage pour enlever le bruit [11]	16
2.7	Exemple d'images avant et après amélioration de contrast [42]	17
2.8	Représentation d'un resultat de la segmentation de k means avec un $\mathbf{k}=3$.	18
2.9	k means clustering	18
2.10	Architecture d'un réseau de neurones [44]	19
2.11	Calcul d'une fonction d'activation $[47]$	20
2.12	Représentation graphique de la propagation avant (forward propagation)	22
2.13	Représentation graphique de la propagation d'arrière (forward propagation) .	22
2.14	Représentation de l'opération de pooling $[53]$	23
2.15	Représentation de l'architecture de LeNet-5	24
2.16	Représentation de l'architecture de AlexNet	24
2.17	Représentation de l'architecture de $VGGnet(16)$	24
2.18	Représentation de l'architecture de Resnet	25

3.1	Diagramme de flux de travail avec 4 sections distinctes identifiées par des	20
0.0	$\begin{array}{c} \text{couleurs} \\ couleur$	30
3.2	Site callisto - (a) La section ou se trouvent les images a enregistrer - (b) la	
	section ou on trouve un classification realisee par des benevoles des images	01
	des sursauts radio, distinguant les sursauts radio de type II.	31
3.3	Dans la figure, à gauche, nous avons une partie de la première page affichée	
	de la section "Data Access". A droite, nous voyons le résultat obtenu après	
	avoir sélectionné une date spécifique.	31
3.4	Exemple d'une ligne de la page de "data access "	31
3.5	Différents pages de la section "Event lists 2010, 2011 and $\geq =2020$ "	32
3.6	La structure dans les fichiers de la section b	33
3.7	La différence entre les représentions des heures	33
3.8	Les étapes d'extraction d'information	35
3.9	Différents type d'image qu'on a collecté	36
3.10	Diagramme en barres - distribution des différents type d'image	37
3.11	Résumé du processus de nettoyage des données	38
3.12	Les résultats de chaque étape du pré-traitement	40
3.13	Une image avant et après le pré traitement	41
3.14	L'architecture du 1er modèle proposer	43
3.15	L'architecture du 2ème modèle proposer	44
3.16	L'image avant le traitement	47
3.17	L'image après le traitement	48
3.18	Une image qui montre la variation d'intensité des pixel	49
3.19	Un exemple de résultats de chaque étape	50
3.20	Un exemple des étape de la méthode semi-automatique	51
4.1	Le plot de la précision de l'entraînement et la validation du 1er modèle	57
4.2	Le plot de la précision de l'entraînement et la validation du 2eme modèle	57
4.3	Le plot de la perte de l'entraînement et la validation du 1er modèle	58
4.4	Le plot de la perte de l'entraînement et la validation du 2eme modèle	58
4.5	Les matrices de confusion des modèles 1 et 2	59
4.6	Le site web de la détection des sursauts radio	61

LISTE DES TABLEAUX

1.1	Tableau explicatif montrant les 5 classes de sursaut radio	7
2.1	Comparaison des modèles CNN	27
3.1	Informations sur le fichier	45
$4.1 \\ 4.2$	Paramètres du Modèle 1	$54 \\ 55$
4.3	Comparaison des modèles	59
4.4	Comparaison des modèles avec des architecture prédéfinie	60

INTRODUCTION GÉNÉRALE

Contexte

L'étude des sursauts radio solaires de type II revêt une grande importance pour notre compréhension de l'activité solaire et de ses impacts sur notre environnement spatial. Ces sursauts radio, qui sont des émissions intenses d'ondes radioélectriques générées lors de violentes éruptions solaires, ont des conséquences significatives sur notre environnement solaire.

Parmi ces conséquences, nous pouvons mentionner les perturbations des communications radio, qui se traduisent par des interférences et des perturbations des signaux. Ces perturbations peuvent affecter les réseaux de télécommunications, les radios et les systèmes de navigation tels que le GPS. De plus, les sursauts de type II peuvent perturber la structure de l'ionosphère, la couche supérieure de l'atmosphère terrestre qui joue un rôle crucial dans la réflexion des ondes radio.

Les sursauts radio de type II peuvent également engendrer des risques pour les astronautes en orbite et les satellites. Les particules énergétiques éjectées lors de ces sursauts peuvent endommager les équipements électroniques sensibles et accroître les doses de rayonnement auxquelles sont exposés les astronautes.

Enfin, ces sursauts peuvent induire des courants électriques dans les infrastructures terrestres, tels que les réseaux électriques, ce qui peut entraîner des perturbations et des dommages aux équipements.

Problématique

Ce phénomène est très complexe et variable, et la détection des sursauts radio de type II représente un défi majeur pour la communauté scientifique. La complexité de ces phénomènes, combinée à un manque de données appropriées, à cause le progrès dans ce domaine de recherche d'être lents

Objectif

L'objectif de notre recherche est de construire un système de détection et d'analyse des sursauts radio de type II en utilisant des méthodes avancées de deep learning. Notre travail consistera donc à répondre à ce besoin en :

- Identifier et collecter les données nécessaires pour entraîner les modèles de deep learning.
- Concevoir et mettre en place des architectures de deep learning adaptées à la détection des sursauts radio de type II.
- Utiliser des techniques d'imagerie pour analyser les caractéristiques temporelles et spatiales des sursauts radio de type II.

Présentation du mémoire

— Chapitre 01 : Généralités sur le Soleil Dans ce chapitre, nous allons aborder quelques concepts fondamentaux sur le Soleil, qui sont étroitement liés aux sursauts radio solaires de type II. Nous commencerons par introduire les caractéristiques générales du Soleil, puis nous expliquerons en détail les sursauts radio solaires de type II et on introduit d'autres types de sursauts radio.

— Chapitre 02 : État de l'art

Dans ce chapitre, nous passerons en revue l'état de l'art de notre problématique. Nous présenterons les différents concepts, approches et techniques algorithmiques qui ont été utilisés pour résoudre des problèmes similaires. Ensuite, nous examinerons les articles existants pertinents dans notre domaine de recherche, en discutant de leurs contributions et de leurs limites.

— Chapitre 03 :Étude conceptuelle

Dans ce chapitre, nous présenterons en détail les approches spécifiques développées dans le cadre de ce mémoire pour automatiser la détection et l'analyse des sursauts radio de type II. Nous expliquerons les modèles, les méthodes et les algorithmes choisis pour atteindre cet objectif

— Chapitre 04 : Réalisation

Ce chapitre est consacré à la mise en œuvre concrète des approches proposées et à leur évaluation. Nous décrirons la méthodologie utilisée pour mettre en place les approches, les paramètres choisis, ainsi que les résultats obtenus. Nous effectuerons également une comparaison entre les différentes méthodes, y compris des méthodes prédéfinies, afin d'évaluer leur performance.

CHAPITRE 1_

GÉNÉRALITÉS SUR LE SOLEIL

Introduction

Le Soleil, étoile la plus proche de la Terre, exerce une influence vitale sur notre planète. Sans lui, la vie ne serait pas possible. De plus, le Soleil joue un rôle crucial dans l'évolution technologique, conditionnant notre utilisation de l'énergie solaire et notre capacité à comprendre et atténuer les effets des éruptions solaires. En résumé, le Soleil est essentiel à notre existence et à notre progression en tant que société. Dans ce chapitre, nous aborderons des notions de physique solaire importantes pour comprendre le but du projet.

1.1 Soleil

Le Soleil est une étoile de taille moyenne âgée d'environ 4,5 milliards d'années, située à une distance d'environ 150 millions de kilomètres de la Terre. Il se présente sous la forme d'une sphère massive de gaz extrêmement chaud, appelé plasma, principalement composé d'hydrogène et d'hélium. Avec un rayon environ 109 fois plus grand que celui de la Terre, il constitue l'objet le plus imposant de notre système solaire.

Le Soleil produit son énergie par le biais de réactions nucléaires qui transforment l'hydrogène en hélium au sein de son noyau, où la température peut atteindre jusqu'à 15 millions de degrés Celsius. Cette énergie se propage ensuite des couches internes vers les couches externes, comme illustré dans la figure 1.1. [30]

L'activité solaire interagit directement avec le système solaire à travers des variations cycliques régulières, connues sous le nom de cycle solaire, d'une durée d'environ 11 ans. Cette activité se caractérise par des éruptions solaires et des éjections de masse coronale, qui libèrent un flux de particules hautement énergétiques ainsi que des orages magnétiques. Ces phénomènes peuvent avoir un impact sur l'environnement terrestre.



FIGURE 1.1 – Structure du soleil [31]

En raison de cette connexion étroite entre le Soleil et la Terre, la NASA et d'autres agences spatiales internationales surveillent en permanence le Soleil à l'aide d'une flotte de vaisseaux spatiaux. Ces missions permettent une observation continue, 24 heures sur 24 et 7 jours sur 7, afin de mieux comprendre et prévoir l'activité solaire.[32]

1.2 L'activité solaire

1.2.1 Les taches solaires

Les tâches solaires sont un élément très important du cycle d'activité solaire. Ce sont des régions relativement sombres qui apparaissent à la surface du Soleil, plus précisément dans la photosphère [33]. Ces tâches solaires sont générées par l'émergence d'une intense activité magnétique, formant une zone spécifique où le champ magnétique est environ 2500 fois plus puissant que celui de la Terre [34]. Cette structure magnétique entraîne une diminution de la température à l'intérieur de ces tâches par rapport à leur environnement, ce qui leur confère leur apparence distincte, comme le montre l'image 1.2.



FIGURE 1.2 – Exemple de taches solaires

Les tâches solaires sont généralement observées dans les zones actives du Soleil et peuvent se présenter sous la forme de tâches individuelles ou de groupes de tâches, avec des tailles et des formes très variables. Certaines tâches solaires peuvent même être plus grandes que la Terre. Leur nombre et leur fréquence varient au cours du cycle solaire [31]. La durée de vie de chaque tache solaire est différente, pouvant aller de quelques jours à plusieurs mois. En se déplaçant à la surface du Soleil, elles peuvent s'étendre ou se contracter, mais finissent par se décomposer[35].

1.2.2 Les éruptions solaires (Flare)

Une des causes principales des tempêtes solaires sont les éruptions solaires (ou flares en anglais), des phénomènes explosifs et très énergétiques qui se produisent à la surface du Soleil au niveau des taches solaires. Ces éruptions sont provoquées par la libération soudaine d'énergie magnétique accumulée dans les lignes du champ magnétique d'une tache solaire. Cette énergie est libérée dans la chromosphère et la couronne solaire qui sont des couches externes de l'atmosphère du Soleil. Les flares émettent un large spectre de rayonnement électromagnétique, allant des ondes radio aux rayons X et gamma en passant par les UV, ainsi que des particules de haute énergie. A cause des particules ionisantes et du rayonnement électromagnétique très intense, les éruptions solaires peuvent avoir divers effets sur la Terre et l'espace environnant, tels que des perturbations électroniques au niveau des satellites, perturbations des communications sur Terre à cause de l'ionosphère, ainsi que des risques pour les astronautes en orbite [1].



FIGURE 1.3 – Une image réelle d'une éruption solaire vue par trois longueur d'ondes le 14 Mars 2014 Crédit : NASA/SDO/Goddard Space Flight Center.

1.2.3 Les éjections de masse coronale (CMEs)

Les éjections de masse coronale (CMEs) sont de grandes structures contenant du plasma et des champs magnétiques qui sont expulsées depuis la couronne solaire vers le milieu interplanétaire et au-delà de l'héliosphère. Ces dernières sont d'un grand intérêt car elles sont responsables des effets météorologiques spatiaux les plus extrêmes sur la Terre[2]. Elles peuvent provoquer des orages géomagnétiques, des aurores polaires, des perturbations des communications, des blackout radio, des dommages aux satellites et aux installations électriques, ainsi que des risques pour les astronautes en orbite. Ces phénomènes sont de grande ampleur, avec des tailles atteignant plusieurs dizaines de fois le rayon solaire, des vitesses de plusieurs millions de kilomètres par heure et une capacité à transporter des milliards de tonnes de matière. Les CMEs sont donc un élément clé de la météorologie spatiale, qui a pour but de prévoir et à atténuer les risques liés à l'activité solaire[3].



FIGURE 1.4 – Une image qui montre le déclenchement de deux CMEs dans différentes directions le 08 Novembre 2000. Crédit : ESA/NASA/SOHO.

1.3 Sursaut radio solaire

Un sursaut radio solaire est une émission intense et brève d'ondes radio provenant du Soleil[4]. Ces sursauts sont associés à divers phénomènes solaires, tels que les éruptions[4],es CMEs (chocs interplanétaires) [5]. Un sursaut radio intense peut avoir des effets importants sur la propagation des ondes radio dans l'ionosphère terrestre et ainsi perturber les communications, les radars et les systèmes de navigation sur Terre. Les sursauts radio solaires sont classés en plusieurs types selon leur forme spectrale, leur durée et leur polarisation. Ils sont observés par des radiotélescopes ou des réseaux d'antennes dédiés à l'étude du Soleil. 5 types de sursaut radio solaire sont classifiés dans le tableau ci-dessous [36] :

	caractéristiques	durée	Gamme de fréquence	phénomènes associés
Ι	courtes, sursaut de bande passante étroite, se produisent généralement en grand nombre avec un continuum sous-jacent.	sursaut unique : 1 seconde tempête : heures - jours	80 - 200 MHz	régions actives, éruptions, protubérances.
II	sursaut dérive de fréquence lente généralement accomplies par une (généralement plus forte intensité) seconde harmonique	3-30 minutes	fundamental 20-150 MHz	fusées éclairantes, émission de protons, ondes de choc magnétohydrodynamiques.
III	La dérive rapide de fréquence peut se produire singulièrement en groupes, ou des orages (souvent avec un continuum sous-jacent) peut être accompagné d'une seconde harmonique	Sursaut unique : 1-3 seconde groupe : 1-5 minutes tempête : minutes - heures	10 khz - 1 GHz	régions actives, fusées éclairantes.
IV	 type stationnaire 5 : continuum à large bande avec structure fine. type mobile 5 : large bande, dérive de fréquence lente, continuum lisse. flare continua : large bande, continuum fluid. 	 heures - jours 30min - heures 3.3- 45 mins 	1.20MHz - 2 GHz 2.20 - 400 MHz 3.25 - 200 MHz	 Fusées éclairantes, émission de protons. Proéminences éruptives, ondes de choc magnétohydrodynamiques Fusées éclairantes, émission de protons
V	Continuum lisse et de courte durée, suit certaines sursaut de type 3, ne se produit jamais isolément.	1-3 mins	10 - 200 MHz	Même que les phénomènes de sursaut type 3

TABLE 1.1 – Tableau explicatif montrant les 5 classes de sursaut radio

1.3.1 Sursaut radio de typeII

Les sursauts radio de type II ont été découverts par Payne-Scott, Yabsley et Bolton en 1947 [6]. Ils se produisent approximativement au même moment que le pic de rayons X générés lors d'une éruption solaire. Ces sursauts se caractérisent par une dérive lente vers des fréquences plus basses [7]. On observe deux bandes distinctes dans ce type de sursaut, la bande fondamentale et l'harmonique (comme illustré dans la figure 1.5). Généralement, la bande harmonique est souvent utilisée pour les calculs, car elle est plus visible et moins absorbée par l'ionosphère de la Terre que la bande fondamentale[8].



FIGURE 1.5 – Représentation d'un sursaut radio de type 2, où "H" désigne l'harmonique et "F" la fondamentale.

1.4 Réseau E-CALLISTO radio récepteurs solaires

CALLISTO est l'acronyme de : Compound Astronomical Low frequency Low cost Instrument for Spectroscopy and Transportable Observatory. Il a été concu en 2006 par Christian Monstein en tant que membre de l'ancien groupe de radioastronomie (RAG) de l'ETH Zurich en Suisse. CALLISTO est un récepteur hétérodyne programmable utilisé pour l'observation des sursauts radio solaires à des fins scientifiques, éducatives, de diffusion et de surveillance des radiofréquences. Il fonctionne dans la plage de fréquences de 45 à 870 MHz, offrant une résolution de fréquence de 62.5 kHz. Les données collectées par CALLISTO sont enregistrées sous forme de fichiers FIT contenant jusqu'à 400 fréquences par balayage. La résolution temporelle est de 0.25 seconde et la bande passante radiométrique est d'environ 300 kHz. La gamme dynamique totale est supérieure à 50 dB. Des routines de traitement des données ont été développées en IDL, PERL et Python pour faciliter leur analyse. CALLISTO est capable d'observer le spectre radio solaire 24 heures sur 24, toute l'année. L'ensemble des spectromètres CALLISTO forme le réseau e-CALLISTO, où les données collectées par chaque instrument sont automatiquement téléchargées sur un serveur central. [37]. La Figure ?? montre la distribution du réseau e-CALLISTO dans le monde jusqu'à la date du 21 Juin 2023.



FIGURE 1.6 – Le réseau e-CALLISTO jusqu'à la date du 21 juin 2023

1.5 Interaction Soleil - Terre (Space weather)

Le Soleil joue un rôle crucial en maintenant la Terre et les Le Soleil joue un rôle crucial en maintenant la Terre et les autres planètes en orbite, empêchant ainsi qu'elles ne se perdent dans l'espace. Outre cette interaction gravitationnelle bien connue, le Soleil exerce d'autres influences sur la Terre. En effet, il émet divers types de rayonnements électromagnétiques tels que les ondes radio, les ultraviolets (UV) et les rayons X, ainsi que des particules chargées et des éjections de masse coronales (CMEs) qui contiennent des particules à très haute énergie, accompagnées de champs magnétiques. Selon leur intensité, leur direction et leur durée, ces émissions peuvent avoir des conséquences diverses sur notre planète et son environnement.



FIGURE 1.7 – Représentation de l'interaction Soleil - Terre

Selon leur intensité, leur direction et leur durée, ces émissions peuvent avoir des conséquences diverses sur notre planète et son environnement.Parmi ces conséquences, on peut citer :

- Des perturbations des communications radio et des systèmes de navigation par satellite, en raison de l'ionisation de la haute atmosphère terrestre.
- Des aurores polaires, qui sont des manifestations visuelles de l'interaction entre les particules solaires et le champ magnétique terrestre.
- Des orages géomagnétiques, qui sont des fluctuations du champ magnétique terrestre causées par les variations du vent solaire. Ces orages peuvent endommager les réseaux électriques, les pipelines et les satellites artificiels.
- Des effets biologiques, tels que l'augmentation du risque de cancer de la peau ou des yeux pour les personnes exposées à un rayonnement ultraviolet plus intense, ou l'altération du rythme circadien pour les astronautes en orbite.

Il est donc essentiel de surveiller attentivement les émissions solaires, car elles peuvent avoir un impact significatif sur la vie et les activités humaines sur Terre et dans l'espace.

1.6 L'intérêt d'étudier les sursauts radio solaires de type II

L'interaction du vent solaire avec la couronne et le milieu interplanétaire provoque des ondes de chocs interplanétaires. Des sursauts radio de type II sont associés directement à ces chocs. Ces perturbations importantes arrivent en moyenne deux à trois jours après l'éruption et déclenchent des orages géomagnétiques sur Terre, alors que l'onde radio (sursaut radio de type II) arrive seulement 8 minutes après l'éruption, ce qui permet de prendre des précautions pour sauvegarder les instruments à bord des satellites au voisinage de la Terre, ainsi que les installations électriques et les pipelines sur Terre.

Conclusion

Ce chapitre commence par une présentation du Soleil et de ses diverses activités solaires. Ensuite, nous avons abordé les différents types d'ondes radio, ainsi que la relation entre la Terre et le Soleil.

CHAPITRE 2____

ÉTAT DE L'ART

2.1 Introduction

En 1958, le psychologue américain Frank Rosenblatt a conceptualisé une machine appelée le Perceptron dans le but de reproduire la réactivité de l'esprit humain. Aujourd'hui, les réseaux de neurones ont été développés pour résoudre des problèmes complexes. Dans ce chapitre, nous aborderons les notions du deep learning ainsi que d'autres concepts techniques et travaux connexes.

2.2 Imagerie

2.2.1 Images et types de format d'images :

L'image est une représentation visuelle d'un objet, d'un être vivant ou d'un concept. Elle peut prendre plusieurs formes telles que la photographie, la peinture, le dessin ou l'image numérique. L'image numérique se décline en plusieurs formats, parmi lesquels on trouve le format .FIT et .png :

- 1. Le format .png : (Portable Network Graphic) est un format d'image pixelisée. Ce qui le distingue des autres formats, c'est sa capacité à gérer jusqu'à 16 millions de couleurs.
- 2. Le format .fit :[38](Flexible Image Transport System) est largement utilisé en astronomie. Il comprend plusieurs dispositifs permettant de décrire les informations photométriques et spatiales de calibrage, ainsi que les métadonnées relatives à l'origine de l'image.



FIGURE 2.1 – Représentation de la structure d'une image de la format FITS

La structure de l'image FITS est représentée dans la figure 2.1, où l'en-tête principal est la première entête qui est suivie des données (DATA), qui peuvent être seulement une image ou vides.

Ensuite, on peut trouver des en-têtes d'extension (il peut y en avoir plusieurs) suivis des données qui peuvent être de différents types, tels que des images, des tableaux ASCII ou des tableaux binaires, peu importe le type.

2.2.2 Prétraitement des images :

Le prétraitement est un ensemble de techniques nécessaires pour améliorer la qualité de l'image, corriger les distorsions indésirables et les préparer efficacement pour une analyse et un traitement ultérieurs. Parmi les techniques les plus couramment utilisées, on retrouve :

a. Les filtres de morphologie :

Les filtres de morphologie utilisent la morphologie mathématique pour modifier la forme et la structure des objets présents dans une image. Cette technique repose sur deux opérations de base : la *dilatation* et l'*érosion*.

La dilatation consiste à ajouter des pixels à l'image en fonction de certains critères prédéfinis, ce qui permet d'agrandir l'objet dans l'image. Elle utilise un élément structurant de petite taille avec une forme prédéfinie, telle qu'un carré ou un cercle.

Chaque pixel de l'image est remplacé par la valeur maximale de son voisinage défini par l'élément structurant. Ce processus aide à combler les trous dans l'objet et à le rendre plus connecté [39].

On peut observer les résultats de l'opération de dilatation sur une image binaire avec trois éléments structurants différents dans la figure 2.2



FIGURE 2.2 – Exemple de dilatations sur des images binaires. [39]

Sur la colonne de gauche, de la figure 2.2 on voit l'image originale binaire : les pixels noirs sont considérés comme faisant partie de l'image. Sur la colonne du milieu, on voit les 3 éléments structurants

L'équation de la dilatation est définie comme suit :

$$\forall x, (f \oplus SE)(x) = \delta_{SE}(f)(x) = \max_{p \in SE} (x+p)$$

où f représente l'image binaire, SE est l'élément structurant, $\delta_{SE}(f)(x)$ représente la dilatation, et p désigne le pixel.

L'érosion, d'autre part consiste à supprimer des pixels de l'image, ce qui réduit la taille de l'objet dans l'image. Elle utilise également un élément structurant, mais remplace chaque pixel par la valeur minimale de son voisinage défini par l'élément structurant [39].

Un exemple de l'opération d'érosion peut être observé dans la figure 2.3



FIGURE 2.3 – Exemple d'érosion sur des images binaires.[39]

Sur la colonne de gauche, on voit l'image originale binaire : les pixels noirs sont considérés comme faisant partie de l'image. Sur la colonne du milieu, on voit les 3 éléments structurants

L'équation de l'érosion est définie comme suit :

$$\forall x, (f \ominus SE)(x) = \epsilon_{SE}(f)(x) = \min_{p \in SE}(x+p)$$

où f représente l'image binaire, SE est l'élément structurant, $\epsilon_{SE}(f)(x)$ représente l'érosion, et p désigne le pixel.

Les deux filtres de morphologie les plus utilisés sont :

- 1. Le filtre d'ouverture : Le filtre d'ouverture consiste à effectuer une opération d'érosion suivie d'une dilatation sur une image. Son objectif principal est de réduire le bruit, de lisser les contours et d'éliminer les structures trop fines. L'ouverture est utilisée pour améliorer la segmentation des objets en éliminant les détails indésirables.
- 2. Le filtre de fermeture : Le filtre de fermeture consiste à effectuer une dilatation suivie d'une érosion sur une image. Son objectif principal est de fermer les lacunes, de connecter les structures brisées, et de remplir les trous présents dans les objets. La fermeture est couramment utilisée pour restaurer des objets complets et réduire les discontinuités.

b. Les filtres de réduction de bruit :

Les filtres de réduction de bruit sont utilisés pour éliminer les artefacts indésirables présents dans une image, tels que le bruit électronique ou les pixels isolés. Deux des filtres les plus connus sont le *filtre médian* et le *filtre gaussien*.[9]

Le filtre gaussien est un filtre de lissage qui utilise la fonction gaussienne, nommée d'après le mathématicien Gauss. Ce filtre lisse l'image en remplaçant chaque pixel par la moyenne pondérée de ses voisins, en utilisant une distribution gaussienne. Cela est réalisé en appliquant une convolution entre un noyau gaussien et les pixels de l'image.

Le processus de convolution consiste à passer le noyau gaussien sur chaque pixel de l'image en calculant la moyenne pondérée des valeurs des pixels voisins selon la distribution gaussienne. Cela permet de réduire les variations brusques entre les pixels et de rendre l'image plus lisse[40].



FIGURE 2.4 – exemple d'application d'un filtre gaussien sur une image avec une matrice de convolution de taille 5x5. [41]

Le filtre gaussien produit un léger effet de flou tout en réduisant le bruit aléatoire présent dans l'image, comme illustré dans la figure 2.4

Le filtre médian est un filtre utilisant une fenêtre glissante qui remplace chaque pixel par la valeur médiane des pixels voisins situés dans cette fenêtre, comme indiqué dans la figure 2.5



 $23 \leq 42 \leq 111 \leq 123 \leq 134 \leq 145 \leq 154 \leq 155 \leq 176$

FIGURE 2.5 – Calcule du filtre médian sur une image. [39]

La fenêtre est représentée en bleu sur l'image d'origine (à gauche). Les valeurs dans la fenêtre sont triées (ligne du bas) et l'élément médian (134) est utilisé comme nouvelle valeur sur le résultat (à droite)

Ce filtre est particulièrement efficace pour éliminer les pixels isolés ainsi que le bruit de type "poivre et sel", tout en préservant les contours et les détails de l'image.

Synthèse comparatif :

Le filtre gaussien réduit le bruit de manière globale dans l'image, tandis que le filtre médian agit localement. Le filtre gaussien peut introduire un effet de flou et atténuer les contours et les détails fins de l'image, tandis que le filtre médian préserve ces éléments. En ce qui concerne la complexité de calcul, le filtre médian peut être plus exigeant en termes de ressources de calcul que le filtre gaussien. En résumé, le filtre gaussien est efficace pour réduire le bruit de manière globale, mais peut entraîner un certain flou et une perte de détails. Le filtre médian est préférable lorsque la préservation des contours et des détails est primordiale, même s'il peut nécessiter davantage de puissance de calcul.

c. Détection et filtrage de contours

La détection et le filtrage de contours sont des étapes essentielles pour trouver les objets présents dans une scène et différencier les zones d'une image. La détection de contours consiste à identifier les zones où il y a un changement d'intensité abrupte, ce qui se matérialise par un contour.Il existe différentes techniques pour détecter les contours, qui peuvent être regroupées en deux catégories : celles basées sur la recherche des extrema de la dérivée première et celles qui utilisent des méthodes plus avancées, comme l'algorithme de Canny.[10]



FIGURE 2.6 – Représente la détection de contour et filtrage pour enlever le bruit [11]

Une fois les contours détectés, il est nécessaire de supprimer les contours non significatifs ou indésirables afin de simplifier l'image. Cela facilite les étapes de traitement ultérieures. Une approche courante consiste à filtrer les contours en fonction de leur taille, en éliminant les parties qui sont trop petites pour être considérées comme significatives. La figure 2.6 est un résultat de la détection et du filtrage de contours.

d. Amélioration du contraste

Le contraste est la différence d'intensité et de luminosité entre différentes parties de l'image. L'amélioration du contraste vise à augmenter cette différence d'intensité et de luminosité afin de rendre l'image plus claire et plus informative. [12]

Il existe plusieurs méthodes pour améliorer le contraste, parmi lesquelles :

- 1. Filtre d'amélioration : Cette méthode améliore le contraste global de l'image sans altérer la structure générale de l'image en appliquant des filtres.
- 2. Amélioration globale : Comme son nom l'indique, cette méthode améliore le contraste global de l'image de manière uniforme.
- 3. Amélioration locale : Cette technique améliore le contraste dans des parties spécifiques de l'image. Elle tient compte des caractéristiques locales de l'image et ajuste de manière adaptative le contraste afin de préserver les détails.



FIGURE 2.7 – Exemple d'images avant et après amélioration de contrast [42]

À gauche, vous trouverez la photo non traitée, tandis qu'à droite, vous trouverez la photo après le traitement utilisant les techniques d'amélioration de contraste.

e. Segmentation d'image

La segmentation d'image est le processus de partitionnement d'une image numérique en plusieurs segments ou régions d'image, également appelés objets. L'objectif de la segmentation d'image est de simplifier l'image ou de changer sa représentation en quelque chose de plus significatif et plus facile à analyser. Voici quelques techniques couramment utilisées :

- Détection de contours : consiste à rechercher des discontinuités brusques dans les valeurs des pixels, qui indiquent généralement des contours. Ces contours peuvent définir des régions.[43]
- **Seuillage de couleur :** D'autres méthodes divisent l'image en régions en se basant sur les valeurs de couleur ou la texture.[43]
- Approches basées sur les régions : utilisent l'extraction de caractéristiques à l'aide d'un ensemble de réseaux de neurones convolutionnels (CNN).[13]
- Algorithmes d'apprentissage automatique : une collection d'algorithmes d'apprentissage automatique avec un ensemble de caractéristiques d'image sélectionnées pour produire des segmentations basées sur les pixels, on peut trouver plusieurs algorithme comme le k means[14]



FIGURE 2.8 – Représentation d'un resultat de la segmentation de k means avec un k = 3

K-means

Proposé par James McQueen en 1967, également connu sous le nom de regroupement des centroïdes, est un algorithme très réputé et largement utilisé en apprentissage non supervisé. Bien qu'il soit un peu sensible aux paramètres initiaux (k) fournis, il se distingue par sa rapidité et son efficacité.

Ce type d'algorithme regroupe les instances de données en fonction de plusieurs centroïdes présents dans les données. Chaque point de données est assigné à un cluster en se basant sur sa distance au carré par rapport au centroïde (la méthode euclidienne étant la plus couramment utilisée).[15]



FIGURE 2.9 – k means clustering

2.3 Apprentissage Profond

L'apprentissage automatique profond, également connu sous le nom de deep learning, représente un domaine essentiel de l'intelligence artificielle et un sous-domaine de l'apprentissage automatique (machine learning). Il offre une variété de capacités puissantes, telles que la reconnaissance d'images, la synthèse vocale, la traduction automatique et même la génération de texte. Son objectif principal est d'apprendre à partir de données complexes et non structurées en identifiant automatiquement les caractéristiques essentielles à la classification grâce à des réseaux neuronaux profonds. [16]

2.3.1 Réseaux de neurones

En réalité, le deep learning fait partie d'une nouvelle approche en intelligence artificielle : les réseaux de neurones ou "réseaux neuronaux artificiels" (RNA). Ces réseaux sont inspirés du fonctionnement du cerveau humain, où des neurones artificiels ou des nœuds sont interconnectés pour traiter des informations de manière hiérarchique et non linéaire, permettant de résoudre des problèmes complexes.

La structure d'un RNA est composée de plusieurs couches, notamment des couches d'entrée, des couches cachées et des couches de sortie. Les couches d'entrée reçoivent les données d'entrée, les couches cachées effectuent des calculs intermédiaires et les couches de sortie produisent les résultats finaux, comme représenté dans figure 2.11.



FIGURE 2.10 – Architecture d'un réseau de neurones [44]

L'apprentissage automatique profond repose sur des algorithmes d'optimisation avancés, des fonctions d'activation sophistiquées et des techniques de régularisation pour entraîner les réseaux de neurones. Ces méthodes permettent de prévenir les problèmes de sur-apprentissage et de sous-apprentissage. Un autre élément essentiel de l'entraînement est la fonction de perte, qui aide à évaluer le résultat final du modèle. Nous allons maintenant procéder à une brève explication de chacun des éléments

2.3.2 Fonctions d'activation

Une fonction d'activation est une fonction mathématique non linéaire qui compare la valeur d'entrée (input) avec un seuil (threshold) et décide si le nœud du réseau de neurones doit être activé ou non.

Cette fonction joue un rôle essentiel dans les réseaux de neurones en introduisant de la non-linéarité, ce qui leur permet de modéliser et d'apprendre des relations complexes dans les données.

Sans fonction d'activation, le réseau de neurones ne serait qu'une combinaison linéaire de ses entrées, ce qui limiterait sa capacité à représenter des modèles non linéaires, cette fonction est définie par l'équation :

$$a^{(i)} = \phi(W^{(i)}a^{(i-1)} + b^{(i)})$$

où :

 $a^{(i)}$ représente le vecteur de longueur n_i des activations de la couche i,

 $W^{(i)}$ est une matrice de taille $n_i \times n_{i-1}$ des paramètres de poids,

 $b^{(i)}$ est le paramètre de biais, et ϕ est la fonction d'activation.

Les poids sont des paramètres ajustables attribués à chaque connexion des neurones des couches adjacentes.

Le biais est également ajustable et il aide à influencer le produit final. [45]

Il existe différents types de fonctions d'activation utilisées dans les réseaux neuronaux. La fonction d'activation la plus couramment utilisée est la fonction ReLU (Rectified Linear Unit), qui renvoie l'entrée si elle est positive et zéro sinon. D'autres fonctions d'activation courantes comprennent les fonctions sigmoïdes et les fonctions tangentes.[46]



FIGURE 2.11 – Calcul d'une fonction d'activation [47]

2.3.3 Fonctions de perte

La fonction de perte évalue la performance du modèle en comparant la sortie prédite avec la valeur cible. Son objectif est de minimiser cette fonction, car une valeur élevée indique une mauvaise performance du modèle. La fonction de perte peut avoir des paramètres tels que le biais et le poids qui sont utilisés pour l'optimisation.[48]

Il existe différents types de fonctions de perte en fonction du problème à résoudre, tels que les fonctions de perte de régression pour les problèmes de régression, les fonctions de perte de classification binaire pour les problèmes de classification binaire, et les fonctions de perte de classification multiclasse pour les problèmes de classification multiclasse. La sélection de la fonction de perte appropriée dépend du contexte et de la tâche spécifique à accomplir. [49]

2.3.4 Sur-apprentissage et Sous-apprentissage

Le sur apprentissage (Overfitting en anglais) se produit lorsque le modèle mémorise la data set d'apprentissage par cœur, mais ne parvient pas à généraliser les motifs. Cela se traduit par de bons résultats lors de l'entraînement, mais de mauvais résultats lors des tests, car le modèle ne parvient pas à reconnaître les motifs dans de nouvelles données. [50]

Le sous-apprentissage (underfitting) se produit lorsque le modèle n'apprend pas suffisamment les motifs et ne parvient pas à trouver les relations entre les exemples d'entrée, ce qui se traduit par de mauvais résultats lors de l'entraînement et des tests.[51]

Après avoir abordé les différents éléments du réseau de neurones, tels que les neurones, les couches, les fonctions d'activation et de perte, et après avoir compris les défis de sousapprentissage et de sur-apprentissage, deux procédures essentielles et cruciales pour l'apprentissage du modèle : la propagation avant (forward propagation) et la rétropropagation (backward propagation). Ces deux mécanismes jouent un rôle clé dans l'optimisation du réseau de neurones.

2.3.5 Forward Propagation et Backpropagation

Forward propagation, également appelée la propagation avant, consiste à faire progresser les données d'entrée à travers un réseau neuronal dans un sens unidirectionnel. Son objectif est d'acheminer les données de la couche d'entrée vers la couche de sortie, en évitant les boucles qui ne génèrent pas de sortie.Pendant le processus de propagation avant, chaque neurone calcule une somme pondérée des entrées qu'il reçoit.Cette somme pondérée est obtenue en multipliant chaque entrée par son poids correspondant. Ensuite, la fonction d'activation est appliquée pour déterminer la sortie du neurone en fonction de la somme pondérée. Ce processus se répète jusqu'à atteindre la couche finale du réseau[52]. La figure 2.12 résume le processus de propagation avant dans un réseau neuronal.



FIGURE 2.12 – Représentation graphique de la propagation avant (forward propagation)

Après avoir effectué la propagation avant (forward propagation) et obtenu la sortie finale (output), l'erreur entre cette sortie et la valeur attendue est calculée à l'aide d'une fonction de perte.

Le processus de rétropropagation (back propagation) consiste ensuite à calculer le gradient de cette fonction d'erreur par rapport aux poids et aux biais du réseau. Ces gradients sont propagés en sens inverse à travers le réseau utilisés pour mettre à jour les poids et les biais du réseau de manière à minimiser l'erreur.



FIGURE 2.13 – Représentation graphique de la propagation d'arrière (forward propagation)

2.4 Convolutional Neural Networks

Les réseaux de neurones convolutifs (CNN) sont un type d'algorithme d'apprentissage profond qui a acquis une excellente réputation dans plusieurs domaines de la vision par ordinateur. Les CNN sont principalement utilisés pour la classification d'images et la reconnaissance d'objets, mais ils peuvent également être utilisés dans d'autres tâches telles que le traitement du langage naturel (NLP) et le traitement vidéo [17]. L'architecture du CNN est inspirée par le fonctionnement visuel du cerveau humain, qui comprend des cellules complexes et directes basées sur plusieurs couches alternatives.[18]

2.4.1 les couches du CNN

On trouve trois parties importantes dans . L'architecture du CNN :

- 1. La couche de convolution : le cœur du CNN. Elle supporte la majeure partie de la charge de calcul. Elle combine les données d'entrée (input) avec un filtre (kernel) de convolution et produit une représentation 2D de l'image appelée carte d'activation. Les filtres de cette couche sont ajustés en fonction des paramètres appris afin d'apprendre des informations pertinentes.[53]
- 2. Les couches de pooling : utilisées pour réduire les dimensions spatiales du volume de sortie. La méthode la plus populaire est le max pooling, qui prend le maximum de chaque région non chevauchante de l'entrée. [53]



FIGURE 2.14 – Représentation de l'opération de pooling [53]

3. Les couches entièrement connectées : utilisées pour connecter tous les neurones d'une couche à tous les neurones des autres couches. Le résultat est compressé en un vecteur unidimensionnel et transmis aux couches entièrement connectées suivantes, qui effectuent la classification finale.[53]

2.4.2 5 architectures bien connues de CNN

1. LeNet-5 : LeNet-5 est l'un des premiers CNN développés par Yann LeCun et son équipe, conçu pour la reconnaissance de caractères manuscrits[19]. Le modèle utilise des couches de convolution et de pooling, suivi de couches entièrement connectées. Bien que relativement simple, LeNet-5 a montré de bonnes performances dans diverses tâches comme la classification d'image [54] et la reconnaissance faciale.



FIGURE 2.15 – Représentation de l'architecture de LeNet-5

2. AlexNet : AlexNet a été développé par Alex Krizhevsky en collaboration avec Geoffrey Hinton dans le but de participer au concours ImageNet Large Scale Visual Recognition Challenge (ILSVRC) en 2012, où ils ont remporté la victoire[20]. AlexNet utilise 5 couches de convolution profondes avec des fonctions d'activation non linéaires, suivies par des couches de pooling, et enfin 3 couches entièrement connectées [55]. Il a grandement contribué à populariser l'utilisation des CNN au sein de la communauté de la vision par ordinateur.



FIGURE 2.16 – Représentation de l'architecture de AlexNet

3. VGGNet : VGGNet a été proposé par A. Zisserman et K. Simonyan du groupe Visual Geometry Group (VGG) de l'Université d'Oxford en 2014 [21] dans un article intitulé Very Deep Convolutional Networks for Large-Scale Image Recognition. Cette architecture se distingue par sa profondeur, avec des couches de convolution de petite taille (3x3) et un grand nombre de paramètres. VGGNet a obtenu d'excellents résultats sur le jeu de données ImageNet en 2014, ce qui a démontré l'importance d'augmenter la profondeur des réseaux . Parmi les variantes de VGGNet, on trouve VGG16, qui comprend un total de 16 couches, et VGG19, qui comprend 19 couches.



FIGURE 2.17 – Représentation de l'architecture de VGGnet(16)

- 4. GoogLeNet (Inception) :[22] GoogLeNet, également connu sous le nom d'Inception, a été développé par l'équipe Google Brain. Son architecture est composée de 22 couches et se caractérise par l'utilisation de modules d'inception. Ces modules combinent des convolutions de différentes tailles pour capturer des informations à différentes échelles spatiales[56]. GoogLeNet a introduit l'idée innovante de réduire le nombre de paramètres en utilisant des convolutions 1x1, ce qui a permis d'obtenir de bons résultats sur les jeux de données ImageNet.
- 5. **ResNet :** ResNet, ou Residual Network, a été proposé par Kaiming He, Xiangyu Zhang, Shaoqing Ren et Jian Sun en 2015 [23]. Cette architecture a introduit le concept de blocs résiduels, ce qui a permis de former des réseaux encore plus profonds tout en évitant le problème de la disparition du gradient qui avait entravé leurs performances dans les tâches de reconnaissance d'images à grande échelle. Les blocs résiduels permettent à l'information de "sauter" des couches pour atteindre plus facilement les couches profondes[57]. ResNet a remporté le concours ILSVRC en 2015 et a ouvert la voie à la conception de réseaux encore plus profonds.



FIGURE 2.18 – Représentation de l'architecture de Resnet

2.4.3 Étude comparatif

Nom	Avantage	Inconvénient	erreur
	-Simple et facile à	-Limité aux données petites	
	implémenter.	et simples.	
LeNet	- Adapté aux données	-Non adapté aux tâches de	0.7%
	petit et non compliquée.	reconnaissance complexes.	
	-Coût de calcul faible.	-Profondeur et capacité limitées.	
	-Architecture plus profonde	-Grand nombre de paramètres	
	que LeNet, ce qui lui	et coût computationnel élevé.	
	confere une mellieure capacité à extraire les		
AlexNet	caractéristiques.	-Susceptible de surapprentissage	15 907
Alexinet		de données.	13.370
	-La fonction d'activation		
	et aide à éviter le problème	-Demande beaucoup de mémoire	
	du gradient qui disparaît.	pour l'entrainement.	
	-A permis la construction	-Grand nombre de paramètres et	
	de réseaux de neurones de	cout computationnel eleve.	
	plus de 100 couches.	-Susceptible de surapprentissage	
VGGNet	-Les connexions résiduelles	avec de petits ensembles	7.0%
	améliorent l'apprentissage	de donnees.	
	du réseau en optimisant les	-Prend plus de temps dans	
	de gradients.	l'entraînement qu'autre	
	-Utilise des modules Incention	architecture.	
	qui réduisent le coût		
	computationnel de		
	l'entraînement d'un	-Coût computationnel élevé.	
GoogLeNet	grâce à la réduction		
(Inception	dimensionnelle.	-Susceptible de surapprentissage	6,7%
vi)		de données.	
	-Pratique pour les tâches de		
	avec des ressources		
	computationnelles limitées.		

	-A permis la construction des neuron networks de plus de	-Coût computationnel élevé.	
ResNet	100 couches.- Les connexions résiduelles améliorent l'apprentissage du	-Susceptible de surapprentissage avec de petits ensembles de données.	3.57%
	réseau en optimisant les représentations et le flux de gradients.	-Nécessite beaucoup de ressource mémoire	

TABLE 2.1 – Comparaison des modèles CNN

2.4.4 Travaux liés

En raison de la complexité des éruptions solaires de type II, seuls quelques articles se sont intéressés à leur détection. Cependant, ces dernières années, elles ont suscité davantage d'attention en raison de leur relation avec la prédiction de la météo spatiale et également le pic d'activité solaire qui est prévu entre 2023 et 2025. On peut résumer le contenu de ces articles comme suit : [24] [25] [26]

Smith et al. (2020) ont proposé une nouvelle approche adaptative de région d'intérêt (ROI) pour détecter les sursauts solaires de type II. Leur méthode a restreint la recherche aux régions qui suivent la courbure du sursaut à une fréquence donnée et comprenait une normalisation implicite des données pour réduire la variance de l'apparence du sursaut. Ils ont démontré l'efficacité de leur méthodologie en utilisant un détecteur Histogram of Oriented Gradients (HOG) et une régression logistique, ainsi que des techniques de segmentation basées sur le vote et la soustraction d'arrière-plan. Il convient de noter que leur étude n'a pas utilisé d'algorithmes avancés tels que l'apprentissage profond et l'ensemble de données utilisé était relativement restreint. [24]

Dans une étude réalisée par Mario F.R. et al, (2022) de l'université de Shandong, ces derniers se sont concentrés sur le développement d'un modèle permettant de classer rapidement et automatiquement les données sursauts et non-sursauts. Leur modèle a atteint une précision moyenne de classification de 98,73% et a montré une amélioration significative par rapport aux recherches antérieures, qui n'avaient atteint qu'un taux de vrais positifs de 67,4% pour la détection des sursauts. Cependant, cette étude n'a pas abordé la différenciation entre les différents types de sursauts.[25]

Le troisième article, publié en 2022 à l'université de North-West Potchefstroom, en Afrique du Sud, visait à créer un programme de réseau neuronal conventionnel pour détecter les sursauts radio de type II et III en utilisant des données brutes provenant de l'instrument e-callisto. Les résultats ont montré un rappel de 60,10% pour les sursauts de type et de 68,98% pour les sursauts de type III. Il convient de noter que cette étude ne s'est pas spécifiquement concentrée sur les sursauts de type et que l'ensemble de données utilisées était limité en raison de sa récente publication. De plus, l'étude a été menée pendant une période d'activité solaire élevée, ce qui a permis d'ajouter un plus grand nombre de sursauts solaires à l'ensemble de données [26]

En résumé, bien que des études antérieures aient été réalisées pour détecter les sursauts
solaires de type II, la disponibilité limitée des données et la complexité des sursauts ont posé des défis. Les recherches récentes ont montré des progrès dans l'amélioration de la précision de la détection.

2.5 Conclusion

Dans ce chapitre, nous avons posé les bases nécessaires pour entamer la conception de notre recherche. Nous avons également préparé le terrain pour le chapitre suivant, où nous détaillerons nos approches méthodologiques pour réaliser notre étude. CHAPITRE 3_____

CONCEPTION

3.1 Introduction

Dans ce chapitre, nous exposons en détail notre démarche méthodologique. Nous commençons par présenter la base de données que nous avons soigneusement constituée.Ensuite, nous nous concentrons sur la réalisation des modèles en mettant l'accent sur nos propres architectures.De plus, nous approfondissons l'analyse des images FITS des sursauts radio de type II en détaillant les méthodes et les algorithmes que nous avons développés. Nous explorons les caractéristiques uniques des images FITS et mettons en évidence leur pertinence dans l'analyse des sursauts radio.

3.2 Workflow



FIGURE 3.1 – Diagramme de flux de travail avec 4 sections distinctes identifiées par des couleurs

La figure 3.1 résume notre processus, qui s'est déroulé en quatre étapes. La première étape, qui était à la fois un défi et une priorité, consistait à collecter et préparer une base de données à cause de l'absence de base de données en ligne ou dans l'entreprise. La deuxième étape consistait à alimenter les données dans différentes architectures pour l'entraînement. L'objectif était d'obtenir de bons résultats afin de créer un modèle de prédiction capable de classifier de nouvelles images.

Enfin, nous avons procédé à une analyse approfondie des images avec l'extension .fits pour obtenir des informations importantes sur ce sursaut radio de type II.

3.3 Partie 1 : data-set

3.3.1 Extraction de données

Un aspect crucial de cette étude a été la collecte des données, étant donné qu'elles n'étaient pas disponibles auprès des sources existantes. Conscient de l'importance d'obtenir des données appropriées, j'ai pris l'initiative de réaliser moi-même le processus de collecte. Pour ce faire, j'ai mis en place une méthodologie d'extraction (data scraping) du site e-callisto network [https://www.e-callisto.org/Data/data.html] afin de récupérer les informations et les données nécessaires. Sur le site, on trouve une section appelée "DATA" où l'accès aux données et les informations associées sont disponibles.



FIGURE 3.2 – Site callisto - (a) La section où se trouvent les images à enregistrer - (b) la section ou on trouve un classification réalisée par des bénévoles des images des sursauts radio, distinguant les sursauts radio de type II.

Dans la section (a) "Data access", on trouve des images au format FITS et des images au format PNG pour chaque station intégrée dans le réseau e-CALLISTO à travers le monde, ordonnée par la date.



FIGURE 3.3 – Dans la figure, à gauche, nous avons une partie de la première page affichée de la section "Data Access". À droite, nous voyons le résultat obtenu après avoir sélectionné une date spécifique.

Chaque ligne de la page affiche , après avoir sélectionné une date, le nom de la station ou observatoire qui a capturé la photo, la date correspondantes au format AAAAMMJJ et l'heure au format HHMMSS . De plus, il y a deux hyperliens : un lien FITS qui permet de télécharger l'image FITS, et un lien image qui permet de visualiser l'image directement en cliquant dessus.



FIGURE 3.4 - Exemple d'une ligne de la page de "data access "

Dans la section (b) "Event lists 2010, 2011 and >=2020" on trouve des fichier que des bénévoles on crée , ou on voie la classification des sursaut radio entre eu les sursaut radio de type II avec la date et l'heure exact et les station qui correspond pour chaque mois

			Parent Directory	
			BurstCounts202011 png 2020_12_17 21:15 102K	
			DustCounts202011.phg 2020-12-17 21.15 102K	
			BurstCounts202012.png 2020-12-17 21:15 97K	
Parent Directo	ry		e-CALLISTO_2020_01.txt 2020-12-28 09:52 283	
<u>2010/</u>	2020-12-18 18	:07 -	e-CALLISTO 2020 02.txt 2020-12-28 09:52 283	
2011 /	2020-12-18-18	07 -		No burst list has been generated from 2012-2019 due to lack of man power
	2020-12-10 10		e-CALLISTO_2020_03.txt 2020-12-28 09:51 1.3K	Volunteers welcome to search for bursts in all Callisto data (very time consuming!)
2012-2019	2020-12-17 21	:42 334	e-CALLISTO_2020_04.txt 2020-12-28 09:51 282	Alternatively, you may have a look at SWPC/NOOA here:
<u>2014/</u>	2023-03-22 17	:42 -	e-CALLISTO 2020 05.txt 2020-12-28 09:51 1.3K	ftp://ftp.swpc.noaa.gov/pub/warehouse/
2020 /	2020-12-17 21	-16 -	CALLISTO 2020 06 tot 2020 12 28 00-51 1 7K	
	2020 12 17 21		e-CALLISTO 2020_00.1x1 2020-12-28 09:51 1.7K	1. Select year 2. Select YYYY events
2021/	2021-12-01 16	(41 -	e-CALLISTO_2020_07.txt 2020-12-28 09:51 1.1K	3. Select YYYYMMDDevents.txt and download
<u>2022/</u>	2022-12-01 09	:45 -	e-CALLISTO 2020 08.txt 2020-12-28 09:51 1.1K	
2023/	2023-06-01 17	19 -	CALLISTO 2020 09 tot 2020 12 28 09:51 022	
	2022 12 14 20	16 710	e-CALLISTO 2020 09.1X1 2020-12-28 09.51 923	un cas d'absence de benevole
readme.txt	2022-12-14 20	15 /18	e-CALLISTO_2020_10.txt 2020-12-28 09:51 2.7K	
			e-CALLISTO_2020_11.txt 2020-12-28 09:50 6.0K	
			CALLISTO 2020 12 tert 2021 01 01 10:17 5 2K	
			- CONTRACTO 2020 1234 2021-01-01 10.17 5.2K	
			*	
# Product: e-C	ALLISTO_2020_12.	txt	↓	
# Product: e-C # Prepared by	ALLISTO_2020_12. PI e-Callisto	.txt	¥	
<pre># Product: e-C # Prepared by # Please send #</pre>	ALLISTO_2020_12. PI e-Callisto comments and sug	.txt ggestions	¥ to monstein(at)irsol.ch	
<pre># Product: e-C # Prepared by # Please send # # Missing data</pre>	ALLISTO_2020_12. PI e-Callisto comments and sug : ####-####	.txt ggestions	¥ to monstein(at)irsol.ch	
<pre># Product: e-C # Prepared by # Please send # # Missing data # #Date</pre>	ALLISTO_2020_12. PI e-Callisto comments and sug : ####-#### Time	txt ggestions Type	↓ to monstein(at)irsol.ch Stations	
<pre># Product: e-C # Prepared by # Please send # # Missing data # Date #</pre>	ALLISTO_2020_12. PI e-Callisto comments and sug i: ####-#### Time	.txt ggestions Type	v to monstein(at)irsol.ch Stations	
<pre># Product: e-C # Prepared by # Please send # # Missing data # #Date #</pre>	ALLISTO_2020_12. PI e-Callisto comments and sug 1: ###-#### Time 04:53-05:53 05:39-05:39	txt ggestions Type I TIT	v to monstein(at)irsol.ch Stations AustraliarASSA	
<pre># Product: e-C # Prepared by # Please send # # Missing data # #Date #</pre>	ALLISTO 2020 12. PI e-Callisto comments and sug : ####-#### Time 04:53-05:53 05:39-05:39 07:05-07:05	txt ggestions Type I III III	v to monstein(at)irsol.ch Stations Australia-ASSA INDIA-GAURI, INDIA-DOTY, MRT1 INDIA-GAURI, INDIA-DOTY, MRT1	
# Product: e-C # Prepared by # Please send # Missing data #Date #	ALLISTO 2020 12. PI e-Callisto comments and sug : ####-#### Time 04:53-05:53 05:39-05:39 07:05-07:05 07:20-08:11	txt ggestions Type I III III I	v to monstein(at)irsol.ch Stations Australia-ASSA HUDIA-GAURT, HUDIA-COTY, MRT1 HUDIA-GAURT, HUDIA-COTY, MRT1 HUDIA-GAURT, HUDIA-COTY, MRT1 HUDIA-GAURT, HUDIA-COTY	
# Product: e-C # prepared by # Please send # #Date #	ALLISTO 2020 12. PI e-Callisto comments and sug : ####-#### Time 04:53-05:53 05:39-05:39 07:05-07:05 07:20-08:11 09:30-09:30	txt gestions Type I III III III III	to monstein(at)irsol.ch Stations Australia-ASSA INDIA-GONY, MRT1 INDIA-GONY, MRT1 INDIA-GONY, MRT1 INDIA-GONY, MRT1 INDIA-GONR, INDIA-OOYV	
# Product: e-C # Prepared by # Please send # #Date #Date #	ALLISTO 2020 12. PI e-Callisto comments and sug : ####-#### Time 04:53-05:53 05:39-05:39 07:05-07:05 07:20-08:11 09:30-09:30 11:00-11:38	txt gestions Type I III III I III I I	to monstein(at)irsol.ch Stations Australia-NSSA HUDIA-GAURT, NDIA-GOTY, MRTI HUDIA-GAURT, NDIA-GOTY, MRTI HUDIA-GAURT, NDIA-GOTY, MRTI HUDIA-GAURT, NDIA-GOTY HUDIA-GAURT, MRTI, SUGS-Landschlacht AUSTRIA-HTGERBACH	
# Product: e-C # prepared by # Please send # # Missing data # #Date #	ALLISTO 2020 12. PI e-Callisto comments and sug as ###-#### Time 04:53-05:53 05:39-05:39 07:85-07:05 07:20-08:11 09:30-09:30 11:00-11:38 22:15-23:59	txt ggestions Type I III III I I I I I I	v to monstein(at)irsol.ch Stations Australia-ASA TBDIA-GAURI, TBDIA-COTY, MRT1 TBDIA-GAURI, TBDIA-COTY, MRT1 TBDIA-GAURI, MBIA-COTY, MRT1 MDIA-GAURI, MBIA-STA MUSTALIA-NTOFELBACH AUSTRIA-NTOFELBACH AUSTRIA-NTOFELBACH AUSTRIA-NTOFELBACH	
# Product: e-C # Prepared by # Please send # Bate # Missing data # Date 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201	AllISTO 2020 12. PI e-Callisto comments and sup : ####=#### Time 04:53-05:53 05:39-05:39 07:05-07:05 07:05-07:05 07:05-07:05 10:00-11:38 22:15-23:59 23:58-23:58	txt ggestions Type I III III I I III I III	to monstein(at)irsol.ch Stations Australia-NSSA INDIA-GOURY, MRTI INDIA-GOURY, MRTI INDIA-GOURY, MRTI INDIA-GOURY, MRTI, SUSA-SLANdschlacht Australia-NSSA MKSART	
# Product: e-C # Prepared by # Please send # #Date #Date #Date #Date 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202	ALLISTO 2020_12. PT e-Callisto comments and sug :: ####=#### Time 04:53-05:53 07:30-05:39 07:20-08:11 09:30-09:30 11:00-11:38 22:15-23:59 23:56-23:58 00:00-06:15	txt ggestions Type I III III III I I III I I III I I I I	v to monstein(at)irsol.ch Stations Australia-ASA INDIA-GURI, INDIA-GOTY, MRT1 INDIA-GURI, INDIA-GURI, INDIA-GOTY, MRT1 INDIA-GURI, INDIA-GOTY, MRT1 INDIA-GURI, INDIA-GURI, INDIA-GOTY, MRT1 INDIA-GURI, INDIA-GURI, INDIA-G	
# Product: e-C # Prepared by # Please send # Missing data #Bate 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201	ALLISTO 2020 12. PT e-Callisto comments and sug : ####################################	txt ggestions Type I III III I I I I I I I I I I I I I I	to monstein(at)irsol.ch Stations Australia-ASSA HURLA-GAURT, NDIA-GOTY, MRT1 HURL-GAURT, NDIA-GOTY, MRT1 HURL-GAURT, NDIA-GOTY, MRT1 HURL-GAURT, NDIA-GOTY HURL-GAURT, MRT1, SUISS-Landschlacht Australia-ASSA MGXAT Australia-ASSA Augtralia-ASSA AugtraliaAugtralia-ASSA AugtraliaAugtraliaAug	
# Product: e-C # Prepared by # Please send # Missing data # 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202	ALLISTO 2020_12. PI e-Callisto comments and sug i: ####-#### Time 04:53-05:53 05:39-05:39 07:05-07:05 07:20-08:11 09:30-09:30 11:00-11:38 22:15-23:59 23:58-23:58 00:00-06:15 06:07-06:42 00:00-06:15 06:07-06:42	txt ggestions Type I III III I I I I I I I I I I I I I I	v to monstein(at)irsol.ch Stations Australia-ASA INDIA-GAURI, INDIA-COTY, MRT1 INDIA-GAURI, INDIA-COTY INDIA-GAURI, MRTI, SUSS-Landschlacht AUSTRIA-HOTERBACH AUSTRIA-TOTREBACH MCXART MCXART BTA BTA	
# Product: e-C # Prepared by Please send # # Missing data # #Date # 2001201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202 20201202	ALLISTO 2020 12. PI e-Callisto comments and sug : ####-#### 44:53-05:53 05:30-05:53 07:20-06:13 07:20-06:11 09:30-09:30 07:20-06:11 11:00-11:20 11:35-11:35	txt ggestions Type I III III I I I I I I I I I I I I I I	Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA BER GLASGOG, HMAIN, INDIA-GAURI BER BER GLASGOG, HMAIN, INDIA-GAURI, INDIA-ODTY, SMISS-Heiterswil, SMISS-Landsc	Jacht
# Product: e-C # Prepared by # Please send # Missing data #Date #Date 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202 20201202 20201202 20201202	ALLISTO 2020 12. PI e-Callisto comments and sug i: ####-#### Time 04:53-05:53 05:39-05:39 07:05-07:05 07:02-07:05 07:02-07:05 07:02-07:05 00:00-00:15 00:00-00:15 00:02-03:45	txt ggestions Type IIII III III III III III III III III	to monstein(at)irsol.ch Stations Australia-ASSA INDIA-GAURI, INDIA-GOTY, MRT1 INDIA-GAURI, INDIA-GOTY, MRT1 INDIA-GAURI, MRT1, SUISS-Landschlacht Australia-ASSA AufsrLak-Totschader, AuATY, INDIA-GAURI BR BR, GLASGOM, HMMAIN, INDIA-GAURI IBR BR, GLASGOM, HMMAIN, INDIA-GAURI IBR BR, GLASGOM, HMMAIN, INDIA-GAURI, INDIA-GOTY, SMISS-Heiterswil, SMISS-Landsc Australia-ASSA	lacht
# Product: e-C # Prepared by Please send # Missing data # 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202 20201202 20201202 20201202	ALLISTO 2020 12. PT e-Callisto comments and sug time 04:53-05:53 04:53-05:53 05:39-05:39 05:39-05:39 05:39-05:39 05:39-05:39 05:39-05:30	txt ggestions I III III I I I I I I I I I I I I I I	Australia-ASA Mustralia-ASA Mustralia-ASA Mustralia-ASA MUDIA-GAURI, MDIA-COTY, MRT1 MUDIA-GAURI, MDIA-COTY, MRT3 MUDIA-GAURI, MRT1, SAUSS-Landschlacht Australia-ASSA MEXART MESART	slæcht
# Product: e-C # Prepared by Please send # dissing data # 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202 20201204	ALLISTO 2020 12. PT e-Callisto comments and succession time 04:153-05:133 05:139-05:33 07:02-08:11 09:330-09:330 11:00-11:38 22:152-23:59 00:09-00:15 23:158-23:58 00:09-00:15 00:125-01:56 00:125-01:56 00:125-01:56 00:125-01:56 00:125-01:56 00:125-01:57 00:125-01:	txt ggestions Type IIII III III I I III I I I I I I I I	V Stations Australia-ASSA HDDIA-GAURI, HDDIA-GOUY, MRT1 HDDIA-GAURI, MDIA-GOUY, MRT1 HDDIA-GAURI, MRT1, SUISS-Landschlacht AUSTRIA-HICHERACH Australia-ASSA HOZA-GAURI, MRT1, SUISS-Landschlacht Australia-ASSA MCART MASTRIA-HICHERACH Australia-ASSA MCART MASTRA-HICHERACH Australia-ASSA MASTRI MARTANG, AUNATY, HDIA-GAURI BIR BIRSA Australia-ASSA MASTRIA-HICHACHARI, INDIA-GAURI, INDIA-COUTY, SMISS-Heiterswil, SMISS-Landsch MASTRIA-IN-SSA MIT, SMISS-Landschlacht	slacht
# Product: e-C # Prepared by # Please send # #Bases send # #Date # 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201205 20200 20200 20201205 20200 2000000	ALLISTO 2020 12. PT e-Callisto comments and sup time 04:53-05:53 05:33-05:53 07:35-07:55	txt Type I III III III III III III III	Australia-ASA HDDI-GUME, HDDI-GOTY, NRTI HDDI-GUME, HDDI-GOTY, NRTI HDDI-GUME, HDDI-GOTY, NRTI HDDI-GUME, HDDI-GOTY, NRTI HDDI-GUME, HDDI-GUME Australia-ARGA MUSTRIA-ARGAENERA Australia-ARGAENERA MEXART Australia-CRAGG, AUMIY, HDIA-GUMEI BR BR, GLAGGOM, HMPATH, HDIA-GUMEI, HDIA-GOTY, SMISS-Heiterswil, SMISS-Landsc MUTI, SMISS-Landschlacht BR, GLASGOM, HMPATH, HDIA-GUMEI, SMISS-Landschlacht, TRIEST	slacht
# Product: e-C # Prepared by Please send # dissing data # 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202 20201204 20201204	ALLISTO 2020 12. PT e-Callisto comments and sug : ####-#### Hime 44:53-05:53 05:59-05:59 07:20-06:11 09:30-09:30 07:20-06:11 09:30-09:30 07:20-06:11 09:30-09:30 00:00-06:15	txt Type Type I III III III III III III III	V Stations JOURT-AGURT, NOTA-COTY, MRT1 INDIA-GURT, MSTA-COTY, MRT1 INDIA-GURT, MSTA-STATUS INDIA-GURT, INDIA-COTY, MRT1 INDIA-GURT, INDIA-GURT INDIA-GURT, INDIA-GURT BR BR BR BR, GLASGON, HMATN, INDIA-GURT, INDIA-COTY, SWISS-Heiterswil, SWISS-Landschacht BR, GLASGON, HMATN, INDIA-COTY, MRT1, SWISS-Landschlacht, TRIEST BR, GLASGON, HMATN, MRT1, SWISS-Landschlacht, TRIEST	ılacht
# Product: e-C # Prepared by Please send # Missing data #Date #Date #Date 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202 20201202 20201202 20201202 20201203 20201204 20201204 20201204 20201204 20201204 20201204 20201204	ALLISTO 2020 12. PT e-Callisto commerts and sup i: ####-#### Time 04:53.05:53 05:39.05:39 07:05-07:05 07:20-06:11 09:30-07:20 07:20-06:11 09:30-07:20 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:62-06:22 11:20-11:20	txt Type Type I III III III III III III III	V Stations Australia-ASSA INDIA-GAURI, INDIA-GOTY, NRTI Statustralia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Rastralia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA BIR, GLASGON, HWAIN, INDIA-GAURI, INDIA-OOTY, SMISS-Heiterswil, SMISS-Landsch RIM, JORA-ONT, NOTA-ONT, NOTA-SOTY, SMISS-Landschlacht, TRIEST BIR, GLASGON, NRTI, SMISS-Landschlacht, TRIEST BIR, GLASGON, NRTI, SMISS-Landschlacht	llacht
# Product: e-C # Preparad by Please send # Missing data # Missing data # Boate # 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201203 20201203 20201204 20201204 20201204 20201204 20201204	ALLISTO 2020 12. PI e-Callisto Comments and sup comments and sup rime 04:53 09:53 09:39-09:53 09:39-09:53 09:39-09:53 09:30-09:13 09:30-09:13 09:30-09:13 09:30-09:13 09:30-09:13 01:09-11:35 00:42-03:45 01:25-09:36 09:37-09:37 10:52-10:52 11:45-11:55 00:42-03:45 01:25-01:45 11:45-11:55 01:25-01:35 11:45-11:55 11:45-11:45 11:45-11:55 11:45-11:55 11:45-11:55 11:45-11:55 11:45-11:45 11:45-11	txt Type Type I III III III III III III III	Australia-ASA Australia-ASA TDUI-AGMUR, IDUIA-OOTY, MRTI IDUIA-GMUR, IDUIA-OOTY, MRTI IDUIA-GMUR, INDIA-OOTY, MRTI IDUIA-GMUR, INDIA-OOTY, MRTI AUSTRIIA-ATGOREBACH AUSTRIIA-ATGOREBACH AUStralia-ASSA AUStralia-ASSA AUStralia-ASSA AUSTRIIA-CAMAG, AUARTY, INDIA-GAURI IBR, GLASGOM, MMRJN, INDIA-GAURI AUStralia-ASSA AUStralia-ASSA AUStralia-ASSA AUSTRIIA-SSA AUSTRIIA-SSA MRTI, SUISS-Landschlacht BR, GLASGOM, MRTI, SUISS-Landschlacht, TRIEST BR, GLASGOM, MRTI, SUISS-Landschlacht	alacht
# Product: e-C # Prepared by Please send # Missing data #Date #Date 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201202 20201202 20201202 20201202 20201204 2020000000000	ALLISTO 2020 12. PT e-Callisto commerts and sup i: ####-#### Time 04:53-05:53 05:39-05:39 07:20-06:11 09:30-09:30 07:20-06:11 09:30-09:30 00:00-06:11 00:00-07:05	txt ggestions Type I III III III III III III III	Australia-ASSA Australia-ASSA Mustralia-ASSA MUDIA-GAURI, MDIA-GOTY, MRT1 INDIA-GAURI, MDIA-GOTY, MRT1 INDIA-GAURI, MRT1, SUISS-Landschlacht Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA Australia-ASSA	lacht
# Product: e-C # Prepared by Please send # Hissing data # #Date # 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201201 20201204 20201204 20201204 20201204 20201204 20201204	ALLISTO 2020 12. PT e-Callisto Commerts and sup time 44:53-05:53 05:39-05:39 07:49-07:49 07:49-07:49 07:49-07:49 07:59-07:49 07:59-07:49 11:09-11:38 07:59-07:49 23:58-23:58 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:60-06:15 00:62-03:65 00:7-06:32 10:52-10:52 11:24-11:24 00:60-06:00 01:12:39 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12:12 11:24-11:24 01:12:12 0	txt Type Type Tur TIT TIT TIT TIT TIT TIT TIT TI	Australia-ASA Australia-ASA TBDIA-GAURI, TBDIA-GOTY, NRTI TBDIA-GAURI, TBDIA-GOTY, NRTI TBDIA-GAURI, TBDIA-GOTY, NRTI TBDIA-GAURI, TBDIA-GOTY, NRTI TBDIA-GAURI, TBDIA-GOTY, NRTI Australia-ASA Australia-ASA Australia-ASA Australia-ASA Australia-ASA Australia-ASA Australia-ASA MUTA, SUISS-Landschacht BTB, GLASGOM, NRTI, SUISS-LANdschlacht, TRIEST BTB, GLASGOM, NRTI, SUISS-Landschlacht RTI BTB, GLASGOM, SUISS-Landschlacht BTB, GLASGOM, SUISS-Landschlacht RTI BTB, GLASGOM, SUISS-Landschlacht BTB, GLASGOM, SUISS-Landschlacht RTI BTB, GLASGOM, SUISS-Landschlacht BTB, GLASGOM, SUISS-Landschlacht RTI BTB, GLASGOM, SUISS-Landschlacht BTB, GLASGOM, SUISS-Landschlacht RTI BTB, GLASGOM, SUISS-Landschlacht	slacht

FIGURE 3.5 – Différents pages de la section "Event lists 2010, 2011 and >=2020"

La structure des fichier de cette section ce compose d'une en-tête ou en trouve le nom du fichier, qui a préparé les données et une adresse mail pour réclamer des erreur le corp ou on trouve des ligne de data qui nous porte les informations suivante :

- Date : La date de l'observation au format AAAAMMJJ.
- Time : La plage horaire de l'observation au format HH :MM-HH :MM.
- Type : La classification ou le type de l'observation.
- Stations : Les stations ou observatoires impliqués dans l'observation. Les stations sont séparées par des virgules.
- Pied de page pour expliquer les annotation

On peut voir cette structure clairement dans la figure 3.6



FIGURE 3.6 – La structure dans les fichiers de la section b

La relation entre les deux page :

La date et les stations sont écrites de la même façon dans les deux sections. Cependant, on trouve un changement dans la représentation de l'heure entre les fichiers "Event lists 2010, 2011 and ≥ 2020 " et les noms des images dans "Data access".

Dans le premier fichier, on trouve l'heure exacte du début et de la fin d'un sursaut radio solaire au format HH :MM-HH :MM. En revanche, dans le deuxième fichier, on trouve l'heure pour chaque intervalle de 15 minutes (ou autre, selon la station), au format HHMMSS. Par conséquent, seule l'heure a été utilisée sans tenir compte des minutes et des secondes



FIGURE 3.7 – La différence entre les représentions des heures

3.3.2 Les étapes suivit pour la collecte

L'extraction des champs nécessaire

Pour pouvoir enregistrer les images de type II de la section "Data access", nous avons besoin de la date, de l'heure et des stations qui les ont capturées. Ces informations se trouvent dans des fichiers séparés. Pour cela, nous avons suivi les étapes suivantes :

Algorithm 1 Téléchargement et Extraction des Données
1: Télécharger les fichiers de 2020 à 2023 de la section "Event lists 2010, 2011 and $>=2020$ "
dans un dossier.
2: Ouvrir le dossier et, pour chaque fichier du dossier :
3: for all fichier dans le dossier do
4: Obtenir le chemin du fichier.
5: Ouvrir le fichier courant.
6: for all ligne du fichier courant do
7: Séparer chaque ligne en champs.
8: if le troisième champ (Type) est égal à 'II' then
9: Extraire la date, l'heure et la station.
10: Enregistrer les champs extraits dans un autre fichier .txt.
11: else
12: Passer à la ligne suivante.
13: end if
14: end for
15: end for



Le fichier . txt

FIGURE 3.8 – Les étapes d'extraction d'information

L'enregistrement des images

Après avoir obtenu toutes les dates, heures et stations où un onde radio de type II s'est manifesté, nous enregistrons ces informations dans un fichier texte. Nous allons utiliser ce fichier pour enregistrer les images nécessaires de la section "Data access". Voici les étapes :

Algorithm 2 Nettoyage et Téléchargement des Images

- 1: Nettoyer le fichier de toutes les stations qui sont entre crochets [], parenthèses () ou doubles parenthèses (()).
- 2: Créer un dossier pour le téléchargement des images.
- 3: Parcourir le fichier .txt et, pour chaque ligne :
- 4: for all lignes dans le fichier .txt do
- 5: Récupérer la date et construire l'URL correspondante. Par exemple : http://so-leil.i4ds.ch/solarradio/callistoQuicklooks/?date=20230617
- 6: Récupérer la page HTML correspondante et l'enregistrer.
- 7: for all lignes dans le contenu de la page HTML do
- 8: Extraire la première colonne et la diviser en 3 champs : station, date, heure.
- 9: **if** la station ou l'heure correspondent à ceux de la date précise **then**
- 10: Télécharger l'image.
- 11: else
- 12: Passer à la ligne suivante.
- 13: end if
- 14: **end for**
- 15: **end for**

3.3.3 Analyse exploratoire des données



FIGURE 3.9 – Différents type d'image qu'on a collecté

On peut voir que nos données sont très riches en termes de variété d'images, comprenant différents types de sursauts radio solaires ainsi que différentes classes d'images. Cependant, il est important que nos données présentent également des intrusions et des artefacts.



FIGURE 3.10 – Diagramme en barres - distribution des différents type d'image

Dans le diagramme en barres, on peut observer une répartition équitable entre les différents types d'images. Il est intéressant de noter que parmi ces types, nous trouvons des images étiquetées comme étant vides et représentant une grande proportion de notre ensemble de données, ce qui correspond à des intrusions et des artefacts. Il est crucial que notre modèle les reconnaisse, car ils représentent une partie significative de ce que le système e-callisto capture.

3.3.4 Pré traitement et nettoyage des données

Nettoyage

Après avoir collecté les données, on a été confronté à un ensemble de 5 000 images qui comportaient des éruptions solaires de type 2, ainsi que des types 3 et 5. Cependant, cet ensemble de données comprenait également de nombreuses images vides, ainsi que des images présentant des artefacts, du bruit et des perturbations extrêmes liées aux tempêtes. Pour analyser les différentes facettes de ces données, on a fait une étape de nettoyage manuel approfondi. On a eu l'occasion de bénéficier de l'expertise des physiciens du centre de recherche CRAAG pour examiner chaque image individuellement et séparer spécifiquement les images de type 2 dans un dossier distinct, tout en éliminant les images présentant un bruit d'intrusion excessif ou une perturbation trop importante.

Cette approche a été essentielle pour garantir la fiabilité et la validité de nos résultats de recherche. En travaillant avec des données soigneusement sélectionnées et nettoyées, on a pu

étudier les divers aspects de ces éruptions solaires de type II avec une grande précision. En conséquence, on a obtenu deux dossiers distincts comme résultat. Le premier dossier contient exclusivement les sursauts radio de type II, tandis que le second dossier contient différentes images que les stations capturent régulièrement et que l'on a considérées comme des images négatives.



FIGURE 3.11 – Résumé du processus de nettoyage des données

Cette procédure nous a permis d'obtenir des données uniques et bien sélectionnées, introuvables ailleurs sur Internet.

Pré-traitement

Notre capteur se situe dans différente partie de la terre il est donc indispensable que différent intrusion sont possible on peut citer :

- Bruit électronique : Le bruit électronique se produit lorsque des signaux indésirables sont captés par le capteur, ce qui peut entraîner des distorsions ou des artefacts dans les données.
- Perturbations atmosphériques : Les perturbations atmosphériques, telles que les interférences causées par les conditions météorologiques ou les phénomènes ionosphériques, peuvent affecter la qualité des mesures et introduire du bruit supplémentaire dans les données.
- Artefacts de transmission : Des problèmes de transmission des données peuvent entraîner des artefacts, des pertes de données ou des déformations des images capturées par le capteur Callisto.

- Interférences radiofréquences : Les interférences radiofréquences provenant d'autres appareils électroniques ou de sources de signal extérieures peuvent perturber les mesures du capteur et introduire des erreurs dans les données.
- Instabilités du capteur : Des instabilités ou des dysfonctionnements du capteur lui-même peuvent entraîner des erreurs de mesure ou des anomalies dans les données enregistrées

Il est donc indispensable de faire un pré-traitement pour diminuer ces intrusion

Algorithm 3 Prétraitement de l'image

- 1: Convertir l'image en niveaux de gris
- 2: Appliquer un filtre morphologique d'ouverture avec un noyau de taille 5×5 de 1
- 3: Appliquer un filtre morphologique de fermeture avec un noyau de taille 5×5 de 1
- 4: Supprimer le bruit en utilisant un filtre médian avec un noyau de taille 5×5
- 5: Binariser l'image en appliquant un seuillage
- 6: Rechercher les contours dans l'image binarisée et les dessiner sur l'image originale
- 7: Éliminer les contours ayant une surface inférieure à un certain seuil
- 8: Réaliser une opération de ET logique entre l'image de fermeture et l'image binarisée
- 9: Redimensionner l'image et la recadrer pour ne conserver que la partie importante
- 10: Améliorer le contraste de l'image en utilisant l'égalisation d'histogramme adaptative limitée au contraste (CLAHE)





Dans le résultat final, on peut constater que certaines intrusions n'ont pas été éliminées, car notre prétraitement ne vise pas spécifiquement le bruit de RFI (Radio Frequency



FIGURE 3.13 – Une image avant et après le pré traitement

Interference) et les signaux de calibration qui se présentent sous la forme de lignes horizontales et verticales. Nous nous reposons sur la capacité du modèle à apprendre leur forme caractéristique. Ce phénomène est plus fréquent dans la figure 3.20.

3.4 Partie 2 : entraînement des modèles

La taille limitée des données a posé un défi lors de la conception du modèle d'apprentissage profond. Les réseaux neuronaux profonds nécessitent généralement une quantité plus importante de données d'entraînement pour généraliser efficacement et éviter le surapprentissage. De plus, la nature unique des données, sans équivalent disponible , a rendu difficile l'utilisation du transfert d'apprentissage avec des architectures CNNs pré entraînées. Dans ce contexte, nous avons proposé deux architectures inspirées de VGG16, AlexNet et LeNet, qui sont adaptées à nos besoins spécifiques.

3.4.1 Architecture proposé

Modèle 1

Une version simplifiée inspirée de l'architecture de VGG16 présente quelques modifications :

- Profondeur : VGG16 a une architecture plus profonde avec 16 couches, tandis que ce modèle n'a que trois blocs de convolution. Cette modification a été faite pour simplifier le modèle afin d'éviter le surapprentissage en raison de la quantité limitée de données.
- Tailles de filtres : VGG16 utilise des filtres de taille 3x3 dans l'ensemble du réseau, tandis que ce modèle utilise une combinaison de filtres(3x3 et 2x2). Cette combinaison permet de capturer différents motifs de l'image, en tenant compte de leur échelle et de leur complexité.
- Nombre de filtres : VGG16 utilise généralement plus de filtres dans chaque couche de convolution par rapport à ce modèle. Le choix des filtres dans ce modèle est adapté

aux spécificités des données utilisées, ce qui permet d'obtenir des caractéristiques plus pertinentes pour la tâche de classification.

— Couches entièrement connectées : VGG16 possède des couches entièrement connectées à la fin du réseau avec un grand nombre d'unités, tandis que ce modèle n'a que deux couches denses pour la classification. Cette simplification est réalisée dans le but de réduire le risque de surapprentissage, en évitant une complexité excessive du modèle et en favorisant une généralisation plus efficace.

On peut voir l'architecture de ce modèle dans la figure 3.14



FIGURE 3.14 – L'architecture du 1er modèle proposer

Modèle 2

inspirer de le l'architecture de le net et Alex Net aussi utiliser dans un article qui entame une problématique similaire cette architecture a été conçu on considerons le suivant : l'arrangement des couches :Le modèle commence par trois ensembles de couches Conv2D suivis de couches MaxPooling2D. Ce schéma d'alternance de couches est fait pour extraire et réduire les caractéristiques. Nombre de filtres : un nombre de filtres moins élevé pour réduire le coût computationnel. En utilisant moins de filtres, le modèle nécessite moins de calculs pour extraire les caractéristiques des données d'entrée. Simplicité : Le modèle utilise une architecture plus simple avec des blocs de convolution et de pooling répétés, ce qui le rend plus facile à comprendre et à mettre en œuvre.

on peut voir l'architecture de ce modèle dans la figure 3.17



FIGURE 3.15 – L'architecture du 2ème modèle proposer

Comme on peut le voir dans les figures 3.17 3.14, l'agencement des couches des deux modèles est différent, mais ils ont tous les deux le même type de couches :

- 1. Couches de convolution : Les modèles comprennent plusieurs couches Conv2D. Ces couches effectuent des opérations de convolution sur les images d'entrée, extrayant des caractéristiques en appliquant des filtres sur l'image.
- 2. Couches de MaxPooling : Après chaque couche Conv2D, une couche MaxPooling2D est ajoutée. Le Max Pooling réduit les dimensions spatiales des cartes de caractéristiques en effectuant un échantillonnage sous-échantillonné de la représentation, en utilisant une fenêtre de 2x2.
- 3. Couches de Dropout : Des couches de Dropout sont ajoutées après certaines couches Conv2D. Le Dropout est une technique de régularisation qui met aléatoirement à zéro une fraction des unités d'entrée pendant l'entraînement. Cela aide à prévenir le surapprentissage en réduisant les interdépendances entre les neurones.
- 4. Couche de mise à plat (Flatten) : Après les couches de convolution, une couche de mise à plat est ajoutée pour convertir les cartes de caractéristiques en 3D en un vecteur 1D. Cela prépare les données pour les couches entièrement connectées.
- 5. **Couches denses :** Le modèle comprend deux couches Dense. Ces couches effectuent la classification en fonction des caractéristiques extraites par les couches de convolution.

Pour introduire la notion de non-linéarité, des fonctions d'activation sont utilisées :

- ReLU (Rectified Linear Unit) : Elle est utilisée dans les couches de convolution pour capturer les caractéristiques complexes de l'image et améliorer la capacité du modèle à apprendre et généraliser.
- Sigmoid : Elle est utilisée dans la dernière couche Dense pour la classification binaire, car elle produit une sortie entre 0 et 1 représentant la probabilité d'appartenir à une classe.

3.5 Analyse des ondes radio type II

3.5.1 Présentation de la structure des images fits

Pour cette partie on utilise les images .fit car il contient des information qu'on a besoin pour l'analyse.

l'image fit télechargée du sit e-callisto contient les informations suivantes :

No.	Nom	Version	Type	Métadonnées	Dimension	Format
0	PRIMARY	1	PrimaryHDU	40	(3600, 200)	uint8
1	/	1	BinTableHDU	16	1 R x 2 C	[3600D8.3, 200D8.3]

TABLE 3.1 – Informations sur le fichier

Notre analyse se fait sur la partie "PRIMARY". Les données du HDU PRIMAIRE sont stockées dans un format uint8, qui représente des entiers non signés sur 8 bits (valeurs allant de 0 à 255). Il contient 40 cartes d'en-tête (métadonnées) et a des dimensions de (3600, 200).

3.5.2 Analyse des métadonnées

Les métadonnées fournissent des informations sur l'observation, l'instrument, l'objet, l'emplacement et d'autres détails pertinents , on peut mentionner quelqu'une :

- **EXTEND** : l'ensemble de données FITS peut contenir des extensions
- **DATE** : Date et heure de l'observation.
- **CONTENT** : Date, heure et nom de la station.
- **ORIGIN** : Nom de l'organisation.
- **TELESCOP** : Type d'instrument.
- **INSTRUME :** Nom du spectromètre.
- OBJECT : Description de l'objet capturé, dans notre cas, c'est le Soleil.
- **DATE-OBS** : Date de début de l'observation.
- **TIME-OBS** : Heure de début de l'observation.
- **DATE-END** : Date de fin de l'observation.
- **TIME-END** : Heure de fin de l'observation.

De plus, nous trouvons des informations plus détaillées sur l'image et la corrélation entre les pixels, la fréquence (axe des y) et le temps (axe des x) que nous allons utiliser pour notre analyse :

- CRVAL1 : La valeur sur l'axe 1 (axe des x) au pixel de référence [en secondes du jour], indique la quantité physique associée à chaque pixel le long de l'axe des x.
- CRPIX1 : Le pixel de référence de l'axe 1 (axe des x), indique la position du pixel sur l'axe des x qui correspond à la valeur de CRVAL1.
- CDELT1 : L'écart entre le premier et le deuxième élément de l'axe des x. Il indique le changement incrémentiel dans la quantité physique représentée par chaque pixel le long de l'axe des x. Par exemple, si la valeur de cette métadonnée est de 0,25, cela signifie qu'il y a une incrémentation de 0,25 secondes du jour entre les pixels adjacents le long de l'axe des x.
- **CRVAL2** : La même chose que CRVAL1, mais sur l'axe 2 (axe des y).
- **CRPIX2** : La même chose que CRPIX1, mais sur l'axe 2 (axe des y).
- **CDELT2** : La même chose que CDELT1, mais sur l'axe 2 (axe des y).

Nous pouvons également voir ces deux informations :

- **DATAMIN** : Élément minimum dans l'image= 0
- **DATAMAX** : Élément maximum dans l'image= 254.

Cela indique que l'image est en niveaux de gris (grayscale).

3.5.3 Le plot de l'image :

Pour extraire les informations nécessaires de l'image, il faut redimensionner l'image de manière à ce que chaque position d'un pixel représente une valeur temporelle sur l'axe des x et une valeur de fréquence sur l'axe des y. Cela peut être réalisé en utilisant une séquence d'opérations mathématiques basées sur les métadonnées CRVAL, CRPIX et CDELT. Voici l'algorithme correspondant :

Algorithm 4 Calculer les valeurs de fréquence et de temps

Étape 1 : Obtenir les dimensions de l'image
Définir la longueur de l'axe des x de l'image.
Définir la longueur de l'axe des y de l'image.
Étape 2 : Calculer les valeurs de fréquence
Initialiser un tableau axisFreq pour stocker les valeurs de fréquence.
Pour chaque indice de pixel i le long de l'axe des y :
Calculer la valeur de fréquence pour le pixel à l'indice i en utilisant la formule suivante :
$freq = CRVAL2 + (CRPIX2 - 1 + i) \times CDELT2$
Ajouter la valeur de freq calculée au tableau axisFreq.
Étape 3 : Calculer les valeurs de temps
Initialiser un tableau axisTime pour stocker les valeurs de temps.
Pour chaque indice de pixel j le long de l'axe des x :
Calculer la valeur de temps pour le pixel à l'indice j en utilisant la formule suivante :
$time = (CRPIX1 - 1 + j) \times CDELT1$
Ajouter la valeur de time calculée au tableau axisTime.

ou:

- i, j : c'est l'index du pixel courant

- $(\operatorname{CRPIX}(1/2) 1 + i)$: calcule la position relative du pixel actuel par rapport au pixel de référence. On soustrait 1 à $\operatorname{CRPIX}(1/2)$ pour tenir compte de l'indexation à partir de zéro, et on ajoute *i* pour déterminer la position du pixel actuel par rapport au pixel de référence.
- $(\operatorname{CRPIX}(1/2)-1+i) \times \operatorname{CDELT}(1/2)$: calcule le changement incrémentiel de la quantité physique représentée par le pixel actuel le long de l'axe y. On multiplie la position relative par $\operatorname{CDELT}(1/2)$, qui est la taille de pas entre les pixels le long de l'axe y, pour déterminer le changement incrémentiel de la quantité physique par rapport au pixel de référence.



FIGURE 3.16 – L'image avant le traitement



FIGURE 3.17 – L'image après le traitement

3.5.4 Méthode 1 : Analyse automatique

Cette méthode repose sur des pré définition importante des sursaut radio solaire :

- L'analyse des sursaut radio type II se fait généralement sur la courbe fondamentale qui a une gamme de fréquence de 20-150 MHz.
- Dans les images des sursaut radio on trouve que l'intensité du pixel qui fait partie du sursaut radio est plus grande que les autres comme on peut voir dans la figure3.18.



FIGURE 3.18 - Une image qui montre la variation d'intensité des pixel

Les 3 algorithmes suivants, résument les étapes accomplies de l'analyse automatique

Algorithm 5 Étape 1 : Prétraitement

- 1: Amélioration du contraste de l'image à l'aide de l'égalisation d'histogramme adaptative limitée au contraste (CLAHE).
- 2: L'application d'un filtre morphologique d'ouverture.
- 3: L'application d'un filtre morphologique de fermeture.
- 4: Binarisation de l'image (application d'un seuillage).
- 5: Recherche des contours dans l'image binarisée et les dessiner sur l'image originale.
- 6: Élimination des contours ayant une surface inférieure à un certain seuil.
- 7: Réalisation d'une opération de ET logique entre l'image de fermeture et l'image seuillée.
- 8: Appliquer l'algorithme du redimensionnement.
- 9: Supprimer les pixels qui représentent une fréquence plus grande que 170 MHz et moins que 20 MHz.

Algorithm 6 Étape 2 : Segmentation et Extraction

- 1: Appliquer l'algorithme de k-means sur l'ensemble des pixels de l'image en utilisant un nombre de clusters (k) égal à 5.
- 2: Calculer l'intensité maximale parmi tous les pixels de l'image.
- 3: Pour chaque cluster, calculer le pourcentage d'intensité maximale en comparant l'intensité de chaque pixel avec l'intensité maximale calculée précédemment.
- 4: Extraire le cluster qui présente le plus grand pourcentage d'intensité maximale. Ce cluster contient les pixels les plus représentatifs.

Algorithm 7 Étape 3 : Extraction des Fréquences par le Temps

- 1: Définition de la fonction non linéaire (expo) qui sera utilisée pour l'ajustement.
- 2: Conversion de l'image en niveaux de gris en une image binaire en utilisant un seuil de valeur pour déterminer les pixels noirs.
- 3: Obtention des indices des pixels non noirs de l'image.
- 4: Extraction des coordonnées x et y des pixels sélectionnés à partir des indices obtenus.
- 5: Création de tableaux myDates et myFreqs contenant respectivement les coordonnées x et y des pixels sélectionnés.
- 6: Réalisation de la régression non linéaire en utilisant les coordonnées x (myDates), les coordonnées y (myFreqs) et une estimation initiale.
- 7: Obtention des paramètres optimisés à partir du résultat de la régression.
- 8: Génération de la courbe ajustée en utilisant les paramètres optimisés et les coordonnées x (myDates) à l'aide de la fonction.
- 9: Finalement, extraire les coordonnées y et x de la courbe comme résultat final.



(a) Un résultats du pré-traitement





(b) Un résultats du k means segmentation d'image avec k=5

(c) Un résultats de la régression non linéaire

FIGURE 3.19 – Un exemple de résultats de chaque étape

Méthode 2 : Analyse Semi-Automatique

A cause de la difficulté de généraliser la première méthode on a créé une méthode semi automatique ou l'utilisateur clique sur les point qu'il veuille étudier pour minimiser l'erreur lors de l'étude des sursaut radio type II.

Algorithm 8 Analyse Semi-Automatique

- 1: Définition de la fonction non linéaire (expo) qui sera utilisée pour l'ajustement.
- 2: Amélioration du contraste de l'image à l'aide de l'égalisation d'histogramme adaptative limitée au contraste (CLAHE).
- 3: Appliquer l'algorithme de redimensionnement.
- 4: Créer une interface où l'utilisateur peut cliquer sur les points qu'il souhaite (limité à 30 points).
- 5: Récupérer les points et créer un nuage de points.
- 6: Réalisation de la régression non linéaire en utilisant les coordonnées x, les coordonnées y et une estimation initiale.
- 7: Génération de la courbe.
- 8: Extraire les coordonnées y et x de la courbe comme résultat final.





la régression non linéaire

FIGURE 3.20 – Un exemple des étape de la méthode semi-automatique

3.6 Conclusion

Dans ce chapitre de conception, nous avons présenté les différentes approches que nous avons essayées, et nos choix de modèles et de méthodes. Pour le chapitre suivant, nous allons passer à la réalisation et aux tests de ces approches, afin de les comparer entre elles

CHAPITRE 4_

RÉALISATION ET EVALUATION

4.1 Introduction

Ce chapitre abordera notre environnement de travail, les métriques utilisées, les résultats obtenus, une étude comparative et une interface utilisateur, offrant ainsi une vision globale de notre projet et de ses différentes composantes.

4.2 Environnement de travail

Dans cette partie nous allons présenter les bibliothèques et outils que nous avons utiliser tout au long de notre projet, ainsi que le matériel avec lequel nous avons travailler.

4.2.1 Environnement matériel



Lenovo ThinkPad, 4GB of RAM, 64-bit operating system, x64 processor, Intel(R) Core(TM) i5-5300U CPU @ 2.30GHz 2.29 GHz processor.

4.2.2 Environmement logiciel



Python 3 : est un langage de programmation interprété de haut niveau qui a été publié en 2008. populaire et prend en charge plusieurs paradigmes tels que l'orienté objet, impératif, fonctionnel et procédural. Python 3 dispose de nombreuses fonctionnalités qui facilitent l'écriture de code clair et concis, telles que l'indentation, le typage dynamique, les structures de données intégrées et les modules. Python 3 est largement utilisé pour le développement web, l'analyse de données, l'apprentissage automatique et l'automatisation [27].







NumPy



Tenserflow : est un framework open-source populaire pour l'apprentissage automatique et l'apprentissage profond. Il permet aux utilisateurs de créer, d'entraîner et de déployer des réseaux neuronaux et d'autres modèles à l'aide de divers outils et langages. TensorFlow prend en charge plusieurs plates-formes telles que Windows, Linux, macOS, Android et iOS. TensorFlow fournit également une API de haut niveau appelée Keras, qui simplifie le processus de construction et d'exécution des modèles[28].

Opencv: est une bibliothèque de fonctions de programmation qui peut être utilisée pour la vision par ordinateur et le traitement d'images. Elle prend en charge plusieurs langages tels que C++, Python, Java et MAT-LAB. OpenCV peut effectuer des tâches telles que la détection de visage, la reconnaissance d'objets, l'extraction de caractéristiques et la manipulation d'images[29].

Matplotlib est une bibliothèque Python permettant de créer et d'afficher différents types de graphiques et de diagrammes. Elle peut être utilisée pour la visualisation de données, le calcul scientifique et les graphiques interactifs. Matplotlib prend en charge de nombreux moteurs graphiques différents, tels que Tkinter, Qt et les navigateurs Web. Matplotlib offre également une interface de haut niveau appelée pyplot, qui simplifie les tâches courantes de création de graphiques.[58]

NumPy : est une bibliothèque Python qui offre des outils puissants pour le calcul scientifique. Elle permet de créer et de manipuler des tableaux multidimensionnels, d'effectuer des opérations d'algèbre linéaire, de générer des nombres aléatoires, d'appliquer des transformations de Fourier, et bien plus encore. NumPy est open source, bien optimisée et compatible avec d'autres bibliothèques[59].

Keras est un framework d'apprentissage profond qui permet de construire et d'entraîner des réseaux neuronaux facilement. Il fournit des API de haut niveau pour des tâches courantes telles que le prétraitement des données, la création de modèles, l'optimisation et l'évaluation. Keras prend également en charge plusieurs moteurs d'exécution, tels que TensorFlow, MXNet et Theano, vous permettant ainsi de choisir celui qui convient le mieux à vos besoins [60].



React est une bibliothèque JavaScript pour la construction d'interfaces utilisateur. Elle permet de créer des composants réutilisables qui peuvent modifier leur apparence et leur comportement en fonction des données et des interactions utilisateur. React est déclaratif, ce qui signifie que vous décrivez ce que vous souhaitez voir à l'écran et React mettra à jour et rendra les éléments de manière efficace[61].



JavaScript est un langage de programmation populaire qui peut être utilisé pour créer des pages web dynamiques et interactives. Il peut également s'exécuter sur d'autres plates-formes, telles que des serveurs, des bases de données et des applications de bureau. Il est facile à apprendre et possède de nombreuses fonctionnalités, telles que les fonctions, les objets, les tableaux et les événements[62].

Google Colab est un service en ligne qui permet d'écrire et d'exécuter du code Python dans votre navigateur. Il est basé sur Jupyter Notebook, un outil open source pour le calcul interactif. Avec Google Colab, vous pouvez accéder à des ressources informatiques gratuites, y compris des GPU et des TPU, et partager vos notebooks avec d'autres utilisateurs. Google Colab est idéal pour les projets de machine learning, de science des données et d'éducation [63].

4.3 Réalisation

Après l'implémentation des deux architectures en utilisant les bibliothèques de python , on peut voir le résulté de ces architecture en utilisant model.summary () le résultat est le suivant pour les 2 architecture :

Layer (type)	$\begin{array}{c} \text{Output Shape} \\ \text{Param } \# \end{array}$		
conv2d 12 (Conv2D)	(None, 300, 300, 32)		
conv2d 13 (Conv2D)	(None, 300, 300, 32)		
$\max_{\text{pooling2d}_6} (MaxPooling2D)$	(None, 150, 150, 32)		
dropout_8 (Dropout)	(None, 150, 150, 32)		
conv2d 14 (Conv2D)	(None, 150, 150, 64)		
$conv2d_{15} (Conv2D)$	(None, 150, 150, 64)		
max_pooling2d_7 (MaxPooling2D)	(None, 75, 75, 64)		
dropout_9 (Dropout)	(None, 75, 75, 64)		
$conv2d_{16} (Conv2D)$	(None, 75, 75, 128)		
$conv2d_17$ (Conv2D)	(None, 75, 75, 128)		
max_pooling2d_8 (MaxPooling2D)	(None, 25, 25, 128)		
dropout_10 (Dropout)	(None, 25, 25, 128)		
$flatten_2$ (Flatten)	(None, 80000)		
$dense_4$ (Dense)	(None, 512)		
dropout_11 (Dropout)	(None, 512)		
$dense_5$ (Dense)	(None, 1)		
Total params : 41,247,457			
Trainable params : $41,247,457$			
Non-trainable params : 0			

 TABLE 4.1 – Paramètres du Modèle 1

Paramètres :

Le nombre total de paramètres entraı̂nables dans le modèle est déterminé par le nombre de poids et de biais :

- Dans cette architecture, le nombre total de paramètres est de 1 721 409.
- La majorité des paramètres proviennent des deux couches Dense, la première couche Dense ayant 262 656 paramètres et la dernière couche Dense ayant 513 paramètres.
- Les couches de convolution contribuent à un nombre significatif de paramètres, avec un total de 1 458 240 paramètres. L'allocation spécifique des paramètres parmi les couches Conv2D dépend du nombre de filtres et de la taille des filtres utilisés.

Les caractéristiques extraites :

On peut calculer les caractéristiques extraites en additionnant le nombre des filtres :

- La première couche Conv2D possède 32 filtres.
- La deuxième couche Conv2D possède 32 filtres.
- La troisième couche Conv2D possède 32 filtres.
- La quatrième couche Conv2D possède 64 filtres.
- La cinquième couche Conv2D possède 64 filtres.
- La quatrième couche Conv2D possède 128 filtres.
- La cinquième couche Conv2D possède 128 filtres.

Donc le nombre des caractéristique extrait est 32+32+64+64+128+128 = 448

Lavar (tura)	Output Shape		
Layer (type)	Param $\#$		
conv2d (Conv2D)	(None, 298, 298, 32)		
max_pooling2d (MaxPooling2D)	(None, 149, 149, 32)		
$conv2d_1 (Conv2D)$	(None, 147, 147, 32)		
max_pooling2d_1 (MaxPooling2D)	(None, 73, 73, 32)		
$conv2d_2$ (Conv2D)	(None, 71, 71, 32)		
max_pooling2d_2 (MaxPooling2D)	(None, 35, 35, 32)		
dropout (Dropout)	(None, 35, 35, 32)		
$conv2d_3$ (Conv2D)	(None, 33, 33, 64)		
max_pooling2d_3 (MaxPooling2D)	(None, 16, 16, 64)		
dropout_1 (Dropout)	(None, 16, 16, 64)		
$conv2d_4$ (Conv2D)	(None, 14, 14, 64)		
max_pooling2d_4 (MaxPooling2D)	(None, 7, 7, 64)		
dropout_2 (Dropout)	(None, 7, 7, 64)		
flatten (Flatten)	(None, 3136)		
dense (Dense)	(None, 512)		
dense_1 (Dense)	(None, 1)		
Total params : 1,680,897			
Trainable params : 1,680,897			
Non-trainable params : 0			

TABLE 4.2 – Paramètres du Modèle 2

Paramètres :

Le nombre total de paramètres entraînables dans le modèle est déterminé par le nombre de poids et de biais :

- Dans cette architecture, le nombre total de paramètres est de 1,680,897.
- La majorité des paramètres proviennent des deux couches Dense, la première couche Dense ayant 1 606 144 paramètres et la dernière couche Dense ayant 513 paramètres.
- Les couches de convolution contribuent à un nombre significatif de paramètres, avec un total de 36 928 paramètres.
- Le modèle possède un nombre relativement faible de paramètres (1 680 897), ce qui suggère une architecture plus efficiente en termes de paramètres. Cela peut présenter des avantages en termes de réduction des besoins en mémoire, des ressources de calcul et du temps d'entraînement.

Les caractéristiques extraites :

- La première couche Conv2D possède 32 filtres.
- La deuxième couche Conv2D possède 32 filtres.
- -La troisième couche Conv2D possède 32 filtres.
- La quatrième couche Conv2D possède 64 filtres.
- La cinquième couche Conv2D possède 64 filtres.
- -La quatrième couche Conv2D possède 64 filtres.

Donc le nombre des caractéristique extrait est 32+32+32+64+64=224

4.4 Évaluation

4.4.1 Les métriques de performances

- Vrai positif (**TP**) : lorsque le modèle prédit que c'est positif et que c'est réellement positif.
- Faux positif (**FP**) : lorsque le modèle prédit que c'est positif mais que c'est réellement négatif.
- Vrai négatif (**TN**) : lorsque le modèle prédit que c'est négatif et que c'est réellement négatif.
- Faux négatif (**FN**) : lorsque le modèle prédit que c'est négatif mais que c'est réellement positif.

La précision (**accuracy**) : décrit la performance globale du modèle sur toutes les classes. Elle est calculée en divisant le nombre de prédictions correctes par le nombre total de prédictions :

$$\operatorname{accuracy} = rac{TP + TN}{TP + TN + FP + FN}$$

L'exactitude (**précision**) : la précision se concentre sur l'exactitude ou la cohérence des prédictions. Elle est calculée en divisant le nombre de vrais positifs par la somme des vrais positifs et des faux positifs :

$$ext{précision} = rac{TP}{TP + FP}$$

Le rappel (**recall**) : également appelé la sensibilité, il décrit la capacité du modèle à prédire les vrais positifs :

$$ext{recall} = rac{TP}{TP + FN}$$

La matrice de confusion : est une représentation tabulaire qui montre les prédictions du modèle par rapport aux valeurs réelles. Elle présente le nombre de vrais positifs, de vrais négatifs, de faux positifs et de faux négatifs. Elle permet d'avoir une vision globale des performances du modèle en termes de classifications correctes et d'erreurs.

4.4.2 Partie 1 : l'entraînement et validation

Comparaison des deux modèles proposés :



FIGURE 4.1 – Le plot de la précision de l'entraînement et la validation du 1er modèle



FIGURE 4.2 – Le plot de la précision de l'entraînement et la validation du 2eme modèle

Les deux graphes illustrent la précision de l'entraînement et la validation des deux modèles proposer le y axis représente la précision et le x axis le nombre d'époque.

On remarque que, à mesure que le nombre d'époques d'entraînement augmente, à la fois la précision de l'entraînement et celle de la validation s'améliorent régulièrement. Cela indique que le modèle apprend efficacement à partir des données d'entraînement et généralise bien aux données invisibles.

La précision de l'entraînement commence à une valeur relativement basse, mais augmente rapidement, démontrant ainsi la capacité du modèle à capturer les motifs sous-jacents dans l'ensemble d'entraînement.

La précision de la validation suit une tendance similaire, ce qui montre que le modèle ne se sûr-apprend pas aux données d'entraînement, mais obtient de bonnes performances sur de nouvelles données invisibles.



FIGURE 4.3 – Le plot de la perte de l'entraînement et la validation du 1er modèle



FIGURE 4.4 – Le plot de la perte de l'entraînement et la validation du 2eme modèle

Les deux graphes illustrent la perte de l'entraînement et la validation des deux modèles proposés le y axis représente la perte et le x axis le nombre d'époque.

Tout au long du processus d'entraînement pour les deux graphes, on observe quelques fluctuations dans les courbes de perte. Ces fluctuations peuvent être causées par divers facteurs, tels que la complexité de l'ensemble de données ou les caractéristiques spécifiques de l'architecture du modèle. Cependant, dans l'ensemble, les valeurs de perte diminuent régulièrement, ce qui indique que le modèle parvient à réduire les erreurs. Il est à noter que la perte de validation suit de près la perte d'entraînement tout au long du processus d'entraînement. Cela suggère que le modèle généralise bien aux données non utilisées pour l'entraînement, car la perte de validation, qui mesure les performances sur des données non utilisées pour l'entraînement, est alignée avec la perte d'entraînement dans le modèle proposé 2 et très proche dans le modèle VGG 16 simplifiée.

Modèle	Précision	L'exactitude	Le rappel
Modèle 1	92%	81%	90%
Modèle 2	92%	84%	86%

TABLE 4.3 – Comparaison des modèles

Du tableau on peut voir que les deux modèles affichent un taux de précision similaire ce qui indique des performances globales similaires en termes de classification. Cependant, le 2ème Modèle Proposé surpasse le Modèle 1 en termes d'exactitude, avec une valeur de 84% par rapport à 81%. D'un autre côté, le Modèle 1 présente un taux de rappel légèrement plus élevé de 90% par rapport aux 86% du 2ème Modèle proposé. on peut voir plus de détaille dans les matrices de confusion suivant



(a) La matrice de confusion de 2ème modèle

(b) La matrice de confusion de 1er modèle

FIGURE 4.5 – Les matrices de confusion des modèles 1 et 2

De ces de matrice en remarque le suivant :

- Vrais Positifs (VP) :

Le Modèle 1:109

Le Modèle 2:104

- Le Modèle 1 présente un nombre légèrement plus élevé de vrais positifs, ce qui signifie qu'il prédit correctement un peu plus d'instances positives que le Modèle 2.

— Vrais Négatifs (VN) :

Le Modèle 1:295

Le Modèle 2:300

- Le Modèle 2 : a un nombre légèrement plus élevé de vrais négatifs, ce qui indique qu'il prédit correctement un peu plus d'instances négatives que le Modèle 1.

— Faux Positifs (FP) :

Le Modèle 1 : 24

Le Modèle 2:19

- Le Modèle 2 a un nombre légèrement plus faible de faux positifs, ce qui signifie qu'il fait moins d'erreurs en prédisant des instances positives lorsque la classe réelle est négative.

— Faux Négatifs (FN) :

Le Modèle 1 : 11

Le Modèle 2:16

- Le le Modèle 1 a un nombre légèrement plus faible de faux négatifs, ce qui suggère qu'il fait moins d'erreurs en prédisant des instances négatives lorsque la classe réelle est positive.

En conclusion, le choix entre les deux modèles dépend de nos besoins spécifiques. Les performances des deux modèles sont assez similaires. Étant donné que notre objectif est de détecter les sursauts radio de type II de manière précise et d'éviter un grand nombre d'autres types de sursauts ou image étiquetés faussement comme étant de type II.

Il est important de noter que le 2ème modèle proposé présente un nombre relativement faible de paramètres, ce qui le rend moins lourd en termes de complexité. Cela peut être avantageux en termes d'efficacité de calcul et de temps d'exécution.

Finalement dans notre cas, le choix du 2ème modèle proposé semble plus approprié. Il offre des performances similaires tout en étant plus léger et en répondant mieux à notre objectif spécifique de détection précise des sursauts radio de type II.

Comparaison avec des architecture prédéfinie :

Les résultats suivant sont une comparaison entre les résultats de l'évaluation des architecture prédéfinie avec nos architecture utilisant les métriques précision exactitude et rappel :

Modèle	précision	L'exactitude	Le rappel	
Le Net	24%	24%	100%	
Alex Net	22%	22%	100%	Base
VGG16	72%	0	0	
Modèle 1	92%	81%	90%	Trained from seratch
Modèle 2	92%	84%	86%	

TABLE 4.4 – Comparaison	des	$mod \`eles$	avec	des	architecture	prédéfinie
-------------------------	-----	--------------	------	-----	--------------	------------

En fonction des résultats présentés dans le tableau, on peut voir une amélioration des performances du modèle 1 et du 2ème modèle proposé par rapport aux architectures prédéfinies de CNN. La performance améliorée de ces modèles suggère que les modifications et les améliorations apportées à l'architecture de base ont été bénéfiques.

4.5 Interface Utilisateur

Cette interface web est composée de deux parties. La première partie affiche à l'utilisateur les sursauts de type II détectés pour la journée, s'il y en a.

La deuxième partie affiche toutes les données de la journée capturées par e-callisto du CRAAG en Algérie.



FIGURE 4.6 – Le site web de la détection des sursauts radio

4.6 Conclusion

Dans ce dernier chapitre, nous avons évalué nos approches de détection des sursauts radio de type II. Nous avons effectué une évaluation approfondie de chaque approche, en analysant les performances en termes de précision, d'exactitude, de rappel et en utilisant la matrice de confusion. Nous avons également discuté des résultats obtenus et identifié les forces et les limites de chaque modèle.et on a fait une comparaison entre les architecture proposer et entre autre architecture prédéfinie.

CONCLUSION GÉNÉRALE

Dans le cadre de ce projet, nous nous sommes intéressés à la création d'un système de prédiction et d'analyse des sursauts radio solaires de type II en utilisant des méthodes d'apprentissage profond. Avant tout, nous avons dû constituer l'ensemble de données nécessaire pour poursuivre notre projet. Nous avons créé un ensemble de données unique et comprenant 4 426 échantillons. Nous avons développé deux architectures de réseaux de neurones profonds adaptées à notre ensemble de données et les avons entraînées dessus. Nous avons testé les performances et ajusté les hyperparamètres jusqu'à obtenir un résultat satisfaisant pour les deux modèles, on a fait une comparaison et on a choisi le model qui a donner une précision de 92%, une exactitude de 84% et un taux de classification de 86% pour les images de type II.

Nous avons pris les images FITS des sursauts radio de type II et les avons analysées en utilisant une segmentation par la méthode de K-means et une fonction FITS. Deux méthodes ont été proposées : une méthode entièrement automatisée dont nous n'avons pas pu évaluer la généralisation en raison du manque d'échantillons de sursauts radio de type II analysables, et une méthode semi-automatisée que peut être utilisée pour minimiser les erreurs. Enfin, nous avons mis en place un site web qui récupère les données fournies par Callisto CRAAG Algérie et prédit les sursaut type quotidiennement.

Les résultats obtenus sont des taux de précision et d'exactitude élevés pour la classification des images de sursauts de type II, démontrant l'efficacité du système proposé. Cependant, des améliorations supplémentaires peuvent être envisagées, telles que l'augmentation progressive de la taille de l'ensemble de données, la collecte de plus de sursauts radio de type II analysables et l'exploration approfondie des hyperparamètres afin d'éliminer les fausses alertes

Dans l'ensemble, cette recherche représente une contribution significative dans le domaine, offrant des perspectives sur la prédiction et l'analyse des sursauts radio solaires de type II. Le système développé présente un potentiel élevé pour des applications pratiques et des avancées supplémentaires dans la recherche en physique solaire.

BIBLIOGRAPHIE

- [1] Zdenek SVESTKA. Solar flares. T. 8. Springer Science & Business Media, 2012.
- [2] David F WEBB et Timothy A HOWARD. « Coronal mass ejections : Observations ». In : Living Reviews in Solar Physics 9.1 (2012), p. 1-83.
- [3] Propagation du Soleil à la TERRE. « Ejections de masse coronales : propagation et interaction avec la magnétosphère terrestre ». In : ().
- [4] Xinan YUE et al. « The effect of solar radio bursts on GNSS signals ». In : *Extreme Events in Geospace*. Elsevier, 2018, p. 541-554.
- [5] JV WIJESEKERA, KPSC JAYARATNE et J ADASSURIYA. « Analysis of type II and type III solar radio bursts ». In : Journal of Physics : Conference Series. T. 1005. 1. IOP Publishing. 2018, p. 012046.
- [6] Silja POHJOLAINEN et Nasrin TALEBPOUR SHESHVAN. « Formation of Isolated Radio Type II Bursts at Low Frequencies ». In : *Solar Physics* 296.5 (2021), p. 81.
- [7] Stephen M WHITE. « Solar radio bursts and space weather ». In : Asian Journal of Physics 16 (2007), p. 189-207.
- [8] NAM NORSHAM et ZS HAMIDI. « An Analysis on the Formation of Solar Radio Burst Type II, III and IV by Using e-CALLISTO, IUGONET and Space Weather Data ». In : Journal of Physics : Conference Series. T. 1298. 1. IOP Publishing. 2019, p. 012018.
- [9] Abdalla Mohamed HAMBAL, Zhijun PEI et Faustini Libent ISHABAILU. « Image noise reduction and filtering techniques ». In : International Journal of Science and Research (IJSR) 6.3 (2017), p. 2033-2038.
- [10] Jean-Hugh THOMAS. « Détection de contours par opérateur de Roberts. » In : (2009).
- [11] Gilles AUBERT et Pierre KORNPROBST. « Traitement des images numériques ». In : ().
- [12] Max MIGNOTTE. « TRAITEMENT D'IMAGES ». In : ().
- [13] Satoru MASUBUCHI et al. « Deep-learning-based image segmentation integrated with optical microscopy for automatically searching for two-dimensional materials ». In : npj 2D Materials and Applications 4.1 (2020), p. 3.
- [14] Kaiwen CHANG. « Machine learning for image segmentation ». Thèse de doct. Université Paris sciences et lettres, 2019.
- [15] J MACQUEEN. « Some methods for classification and analysis of multivariate observations ». In : Proc. 5th Berkeley Symposium on Math., Stat., and Prob. 1965, p. 281.
- [16] Ian GOODFELLOW, Yoshua BENGIO et Aaron COURVILLE. Deep learning. MIT press, 2016.
- [17] Muhammad Asif SALEEM et al. « Comparative analysis of recent architecture of Convolutional Neural Network ». In : *Mathematical Problems in Engineering* 2022 (2022).
- [18] Ahmed Al MAASHRI et al. « Accelerating neuromorphic vision algorithms for recognition ». In : Proceedings of the 49th annual design automation conference. 2012, p. 579-584.
- [19] Yann LECUN et al. « Gradient-based learning applied to document recognition ». In : Proceedings of the IEEE 86.11 (1998), p. 2278-2324.
- [20] Jayant BOKEFODE, MV Panduranga RAO et G KOMARASAMY. « Ensemble Deep Learning Models for Lung Cancer Diagnosis in Histopathological Images ». In : Procedia Computer Science 215 (2022), p. 471-482.
- [21] Karen SIMONYAN et Andrew ZISSERMAN. « Very deep convolutional networks for large-scale image recognition ». In : arXiv preprint arXiv :1409.1556 (2014).
- [22] Christian SZEGEDY et al. « Going deeper with convolutions ». In : Proceedings of the IEEE conference on computer vision and pattern recognition. 2015, p. 1-9.
- [23] Kaiming HE et al. « Deep residual learning for image recognition ». In : *Proceedings* of the IEEE conference on computer vision and pattern recognition. 2016, p. 770-778.
- [24] Joseph JENKINS et al. « Physics-informed detection and segmentation of type II solar radio bursts ». In : British Machine Vision Virtual Conference. 2020.
- [25] JC GUO et al. « A deep learning method for the recognition of solar radio burst spectrum ». In : *PeerJ Comput Sci* 8 (jan. 2022), e855. DOI : 10.7717/peerj-cs.855.
- [26] Herman le ROUX et al. « DEEP LEARNING APPROACH TO CLASSIFYING SO-LAR RADIO BURSTS ». In : ().
- [27] Alberto ARTASANCHEZ et Prateek JOSHI. « Artificial Intelligence with Python ». In : 2017.
- [28] Martín ABADI et al. « TensorFlow : A System for Large-Scale Machine Learning ». In : CoRR abs/1605.08695 (2016). arXiv : 1605.08695. URL : http://arxiv.org/ abs/1605.08695.
- [29] Alexander ZELINSKY. «Learning OpenCV—Computer Vision with the OpenCV Library (Bradski, G.R. et al.; 2008)[On the Shelf] ». In : *IEEE Robotics & Automation Magazine* 16 (2009), p. 100-100.

WEBOGRAPHIE

- [30] Fleur OLAGNIER. Le Soleil, notre bonne étoile. LeSoleil, notrebonneÃltoile. consulté le 10-02-2023.
- [31] Philippe GARCELON. Le Soleil. https://pg-astro.fr/astronomie/le-systemesolaire/le-soleil.html. consulté le 10-02-2023.
- [32] Solar system EXPLORATION. Sun. https://solarsystem.nasa.gov/solar-system/ sun/overview/#:~:text=Our%20Sun%20is%20a%204.5,here%20on%20our%20home% 20planet.. consulté le 10-02-2023.
- [33] SPACEWEATHERLIVE. Que sont les taches solaires ? https://www.spaceweatherlive. com/fr/aide/que-sont-les-taches-solaires.html. consulté le 12-02-2023.
- [34] National weather SERVICE. The Sun and Sunspots. https://www.weather.gov/fsd/ sunspots. consulté le 12-02-2023.
- [35] Libretexts GLOBAL. Le cycle solaire. https://query.libretexts.org/Francais/ Livre_:_Astronomy_(OpenStax)/15:_Le_soleil,_une_star_de_la_variÃl'tÃl'_ des_jardins/15.02:_Le_cycle_solaire. consulté le 12-02-2023.
- [36] World Data CENTER. Spectrograph : Solar Radio Burst Classifications. https://www. sws.bom.gov.au/World_Data_Centre/1/9/3. consulté le 16-02-2023.
- [37] E-CALLISTO. International Network of Solar Radio Spectrometers, a Space Weather Instrument Array, Goal : Understanding Transient Phenomenon in the Solar Corona. https://www.e-callisto.org/index.html. consulté le 20-02-2023.
- [38] HST Data Handbook for WFPC2. FITS File Format. https://www.stsci.edu/ instruments/wfpc2/Wfpc2_dhb/intro_ch23.html#:~:text=AdatafileinFITS, componentimmediatelyfollowstheheader. consulté le 25-02-2023.
- [39] Benjamin PERRET. Morphologie Mathématique. https://perso.esiee.fr/~perretb/ I5FM/TAI/morpho/index.html. consulté le 25-02-2023.
- [40] F-LEGRAND. *Filtres gaussiens*. https://www.f-legrand.fr/scidoc/docmml/image/ filtrage/gaussien/gaussien.html. consulté le 27-02-2023.

- [41] open CLASSROOM. Initiez-vous aux traitements de base des images numériques. https: //openclassrooms.com/fr/courses/5060661-initiez-vous-aux-traitementsde-base-des-images-numeriques/5217251-analysez-le-filtrage-spatial-etla-convolution-par-masque.consulté le 01-03-2023.
- [42] Sharp NEC Display SOLUTIONS. Contrast Enhancement. https://www.sharp-necdisplays.com/global/technology/tec_sv.html. consulté le 02-03-2023.
- [43] Math WORKS. Image Segmentation. https://www.mathworks.com/help/images/ image-segmentation.html. consulté le 03-03-2023.
- [44] TIBCO. Qu'est-ce qu'un réseau neuronal? https://www.tibco.com/fr/referencecenter/what-is-a-neural-network. consulté le 10-03-2023.
- [45] Kuruva Satya GANESH. What's The Role Of Weights And Bias In a Neural Network. https://towardsdatascience.com/whats-the-role-of-weights-and-bias-ina-neural-network-4cf7e9888a0f. consulté le 10-03-2023.
- [46] Vineet JOSHI. Activation Functions. https://www.google.com/url?q=https: //www.geeksforgeeks.org/activation-functions/&sa=D&source=docs&ust= 1687363083992706&usg=AOvVaw2M7DxsxTGg-riBeXVpqMZG.consulté le 11-03-2023.
- [47] MONCOACHDATA. Comprendre les réseaux de neurones. https://moncoachdata.com/ blog/comprendre-les-reseaux-de-neurones/G. consulté le 13-03-2023.
- [48] Vishal YATHISH. Loss Functions and Their Use In Neural Networks. https://towardsdatascience. com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9#:~: text=Alossfunctionisa, thepredictedandtargetoutputs.. consulté le 13-03-2023.
- [49] Jason BROWNLEE. How to Choose Loss Functions When Training Deep Learning Neural Networks. https://machinelearningmastery.com/how-to-choose-lossfunctions-when-training-deep-learning-neural-networks/. consulté le 14-03-2023.
- [50] ALEXANDRA TWIN. Understanding Overfitting and How to Prevent It. https: //www.investopedia.com/terms/o/overfitting.asp. consulté le 14-03-2023.
- [51] Amazon AWS. Model Fit : Underfitting vs. Overfitting. https://docs.aws.amazon. com/machine-learning/latest/dg/model-fit-underfitting-vs-overfitting. html. consulté le 14-03-2023.
- [52] H2O WIKI. What is a Neural Network. https://h2o.ai/wiki/forward-propagation/. consulté le 15-03-2023.
- [53] Mayank MISHRA. Convolutional Neural Networks, Explained. https://towardsdatascience. com/convolutional-neural-networks-explained-9cc5188c4939. consulté le 17-03-2023.
- [54] Mohammad Badhruddouza KHAN. Summary : LeNet-1, LeNet-4, LeNet-5, Boosted LeNet-4 (Image Classification). https://bootcamp.uxdesign.cc/summary-lenet-1-lenet-4-lenet-5-boosted-lenet-4-image-classification-df5673a61eb2. consulté le 17-03-2023.

- [55] Khush PATEL. Architecture comparison of AlexNet, VGGNet, ResNet, Inception, DenseNet. https://towardsdatascience.com/architecture-comparison-of-alexnetvggnet-resnet-inception-densenet-beb8b116866d. consulté le 23-03-2023.
- [56] DOLLY_Vaishnav. ML / Inception Network V1. https://www.geeksforgeeks.org/mlinception-network-v1/. consulté le 20-03-2023.
- [57] Gaudenz BOESCH. Deep Residual Networks (ResNet, ResNet50) 2023 Guide. https: //viso.ai/deep-learning/resnet-residual-neural-network/. consulté le 23-03-2023.
- [58] Matplotlib Development TEAM. *Matplotlib Documentation*. https://matplotlib.org/. Year Accessed.
- [59] NumPy CONTRIBUTORS. NumPy Documentation. https://numpy.org/doc/stable/. Year Accessed.
- [60] François CHOLLET et al. Keras Documentation. https://keras.io/. Year Accessed.
- [61] React CONTRIBUTORS. React Documentation. https://react.dev/. Year Accessed.
- [62] JavaScript.info CONTRIBUTORS. JavaScript.info Documentation. https://javascript. info/. Year Accessed.
- [63] GEEKFLARE. A Beginner's Guide to Google Colab. https://geekflare.com/googlecolab/. Year Accessed.